

UNIVERSIDAD CATOLICA DE SALTA



PROYECTO FINAL

"BUSINESS INTELLIGENCE PARA SUBPROGRAMA DE BECAS PROVINCIALES"

Carrera: Ingeniería en Informática

Alumno: Arroyo, Pablo Gastón



Año: 2019

**"BUSINESS INTELLIGENCE PARA SUBPROGRAMA DE BECAS
PROVINCIALES"**

Profesor Guía:

Nombre: Ing. Fernando Rivera Bernsdorff

Firma: _____

Integrantes del Tribunal Evaluador:

Nombre: _____

Firma: _____

Nombre: _____

Firma: _____

Nombre: _____

Firma: _____

Fecha de Exposición del trabajo:

____/____/____

DEDICATORIAS Y AGRADECIMIENTOS

Poder al fin, dedicar esta tesis, es un sueño hecho realidad. Lo hago con inmensa alegría y consciente de que ha sido posible gracias al cariño y apoyo incondicional de las personas a las que va dedicada.

A mis padres, por tantos años de amor infinito que me ha llevado hasta aquí. Por su apoyo y paciencia incondicional, pero sobre todo por ser el gran motor en todos estos años de estudio y hacer de mí una persona de bien. Este logro es para ustedes. Nunca terminare de agradecerles, los amo profundamente.

A mis queridos abuelos Enrique, Guillermo, María y América, por ser raíces de amor y afecto para mis dos familias, que me han enseñado el lado dulce y no amargo de la vida.

A mi hermana, por su ayuda y cariño durante toda mi vida, por hacerme reír, por ser mi ejemplo a seguir como persona y como futura profesional.

Al resto de la familia, en especial a mi Tía Lita por siempre brindarme un espacio en su hogar, a mi primo Martín por ayudarme siempre que lo necesite.

A mi director de tesis Fernando, por su compromiso, entrega y total guía para obtener este logro, siempre le estaré agradecido por todo el tiempo y las atenciones invertidas en mi proyecto de grado.

A los docentes de la Universidad Católica de Salta, que me acompañaron y orientaron durante los años de la carrera, que hicieron que cada día ame más el camino que elegí.

A mis amigos y compañeros, por acompañarme en este proceso y escuchar de ellos siempre palabras que me llenan de confianza e impulso para culminar esta etapa.

GRACIAS a todos los que aportaron comentarios, consejos, tiempo para el logro de este trabajo.

A todos ellos, muchas gracias.



Índice

ÍNDICE	4
ÍNDICE DE GRÁFICOS	7
ÍNDICE DE TABLAS	10
DEFINICIONES, ACRÓNIMOS O ABREVIACIONES.....	11
ABSTRACT	12
CAPÍTULO 1: INTRODUCCIÓN.....	13
BREVE DESCRIPCIÓN DEL PROBLEMA E IMPORTANCIA	13
MOTIVACIÓN	13
PROCEDIMIENTO A REALIZAR.....	14
CRITERIOS DE ÉXITO.....	14
CAPÍTULO 2: EL ESTADO DE LA CUESTIÓN	15
BUSINESS INTELLIGENCE	15
<i>Orígenes de Business Intelligence.....</i>	<i>16</i>
<i>Componentes de Business Intelligence</i>	<i>18</i>
<i>¿Sector Público o Privado?</i>	<i>22</i>
<i>Beneficios de BI.....</i>	<i>23</i>
<i>Requerimientos en una Herramienta de BI.....</i>	<i>24</i>
BUSINESS INTELLIGENCE EN EL MUNDO	25
BUSINESS INTELLIGENCE Y SU APLICACIÓN	26
PROCESO DE BI	27
CAPÍTULO 3: EL PROBLEMA	29
DEFINICIÓN DEL PROBLEMA	29
OBJETIVOS	29
<i>Objetivos General</i>	<i>29</i>
<i>Objetivos Específicos</i>	<i>29</i>
ORGANISMO SOBRE EL QUE SE REFERENCIA EL TRABAJO	29
<i>Misión del Organismo.....</i>	<i>29</i>
<i>Visión del Organismo.....</i>	<i>30</i>
ANÁLISIS FODA.....	30
<i>Fortalezas.....</i>	<i>30</i>
<i>Debilidades.....</i>	<i>30</i>
<i>Oportunidades.....</i>	<i>31</i>
<i>Amenazas.....</i>	<i>31</i>
<i>Matriz FODA.....</i>	<i>31</i>
SISTEMA DE INFORMACIÓN EXISTENTE.....	32
SISTEMA DE EVALUACIÓN ACTUAL	32
INFRAESTRUCTURA INFORMÁTICA	33
RESTRICCIONES O LIMITACIONES DE LA INVESTIGACIÓN	33
ALCANCE	34
ALTERNATIVAS TECNOLÓGICAS	34
<i>Alternativas de Solución.....</i>	<i>34</i>
<i>Alternativas del motor de Base de Datos</i>	<i>35</i>
<i>Alternativas de Herramientas BI.....</i>	<i>37</i>
<i>Alternativas de Hardware</i>	<i>38</i>



Alternativas de Software Base.....	39
Elección de las Alternativas.....	40
CAPÍTULO 4: SOLUCIÓN PROPUESTA	42
JUSTIFICACIÓN	42
GESTIÓN DE LOS RIESGOS DEL PROYECTO	42
<i>Identificación de los Riesgos</i>	42
<i>Prevención de Riesgos</i>	43
<i>Contrarrestar Riesgos</i>	43
ESTRUCTURA DE DESGLOSE DE TRABAJO (EDT).....	45
EQUIPO DE TRABAJO	46
MATRIZ DE RESPONSABILIDADES	46
<i>Planificación</i>	46
<i>Recolección de requerimientos</i>	47
<i>Diseño solución BI</i>	47
<i>Desarrollo</i>	48
<i>Instalación y pruebas</i>	48
<i>Finalización</i>	48
DIAGRAMA DE GANTT	48
PRESUPUESTO	52
FLUJO DE EGRESOS	54
<i>VAN y TIR</i>	54
BASE DE DATOS	55
<i>Modelo Relacional</i>	55
<i>Modelo lógico y Generación de Datos</i>	60
<i>Migración</i>	66
METODOLOGÍAS PARA EL DESARROLLO DE BI	73
<i>Metodología de Kimball</i>	73
<i>Metodología Inmon</i>	74
<i>Metodología CRISP-DM</i>	75
<i>Metodología DWEP (Data Warehouse Engineering Process)</i>	77
<i>Comparativa de Metodologías</i>	78
<i>Selección de Metodología</i>	78
METODOLOGÍA HEFESTO	79
<i>Descripción</i>	79
<i>Características</i>	80
ANÁLISIS DE REQUERIMIENTOS	80
<i>Preguntas de negocio</i>	80
<i>Indicadores y Perspectivas</i>	81
<i>Modelo Conceptual</i>	81
ANÁLISIS DE DATA SOURCES	82
<i>Hechos e Indicadores</i>	82
<i>Mapeo</i>	83
<i>Granularidad</i>	84
<i>Modelo Conceptual Ampliado</i>	85
MODELO LÓGICO DEL DW.....	86
<i>Tipología</i>	86
<i>Tabla de Dimensiones</i>	86
<i>Tabla de Hechos</i>	87
<i>Uniones</i>	88



INTEGRACIÓN DE DATOS.....	89
<i>Carga Inicial</i>	89
<i>Actualización</i>	93
CUBO MULTIDIMENSIONAL.....	96
<i>Creación Cubo</i>	96
<i>Presentación Cubo</i>	109
CAPÍTULO 5: RESULTADOS.....	114
GENERACIÓN DE DATOS	114
RESULTADOS DE LAS PRUEBAS.....	115
CAPÍTULO 6: CONCLUSIONES.....	117
CONCLUSIONES	117
FUTURAS LÍNEAS DE INVESTIGACIÓN	119
RECOMENDACIONES	120
REFERENCIAS BIBLIOGRÁFICAS	121
ANEXOS.....	124
ANEXO I: REGLAMENTO PROGRAMA DE BECAS	124
ANEXO II: FORMULARIO BECAS.....	128
ANEXO III: CAPTURAS SISTEMA BECAS.....	133
ANEXO IV: ENTREVISTA.....	141
ANEXO V: DATA WAREHOUSE Y CUBO	144
<i>Diseño Físico DW</i>	144
<i>ETL</i>	146
<i>Programación de tareas con SQL Server Agent</i>	152



Índice de Gráficos

Ilustración 1: Cuanto más datos, más conocimientos. Tomado de (Hey, J., The Data, Information, Knowledge, Wisdom Chain: The metaphorical link, 2004)	15
Ilustración 2: Evolución del BI	18
Ilustración 3: Proceso ETL – Extracción. Elaboración propia	19
Ilustración 4: Proceso ETL – Transformación. Elaboración propia	20
Ilustración 5: Proceso ETL – Carga. Elaboración propia	20
Ilustración 6: Business Intelligence	22
Ilustración 7: Proceso BI. Elaboración propia	27
Ilustración 8: EDT. Elaboración propia	45
Ilustración 9: Diagrama de Gantt - Parte 1. Elaboración propia	50
Ilustración 10: Diagrama de Gantt - Parte 2. Elaboración propia	51
Ilustración 11: DER Parte izquierda. Elaboración propia	57
Ilustración 12: DER Parte central. Elaboración propia	58
Ilustración 13: DER Parte derecha. Elaboración propia	59
Ilustración 14: Ejemplo make:Migration. Elaboración propia	61
Ilustración 15: Ejemplo Migration Nacionalidades. Elaboración propia	62
Ilustración 16: Ejemplo make:seeder. Elaboración propia	63
Ilustración 17: Ejemplo Seeder Nacionalidades. Elaboración propia	64
Ilustración 18: Ejemplo Seeder Faker. Elaboración propia	65
Ilustración 19: Backup de base de datos. Elaboración propia	67
Ilustración 20: Estableciendo conexión MySQL. Elaboración propia	67
Ilustración 21: Data import/Restore. Elaboración propia	68
Ilustración 22: Datos de Restauración. Elaboración propia	68
Ilustración 23: Base de datos restaurada. Elaboración propia	69
Ilustración 24: Proyecto Migración. Elaboración propia	69
Ilustración 25: Migración - Conexión MySQL. Elaboración propia	70
Ilustración 26: Migración - Conexión SQL Server. Elaboración propia	70
Ilustración 27: Migración - Convertir Schema. Elaboración propia	71
Ilustración 28: Migración - Sincronizar base de datos. Elaboración propia	71
Ilustración 29: Migración - Sincronización. Elaboración propia	72
Ilustración 30: Migración - Migrando Datos. Elaboración propia	72
Ilustración 31: Migración - Reporte. Elaboración propia	72
Ilustración 32: Migración - SQL Server. Elaboración propia	73
Ilustración 33: Modelo Dimensional Kimball	74
Ilustración 34: Modelo Relacional Inmon	75
Ilustración 35: Modelo CRISP-DM	76
Ilustración 36: Proceso Unificado y DWEP	77
Ilustración 37: Metodología HEFESTO	79
Ilustración 38: Modelo Conceptual. Elaboración propia	82
Ilustración 39: Mapeo entre modelos. Elaboración propia	83
Ilustración 40: Modelo Conceptual Ampliado. Elaboración propia	86
Ilustración 41: Dimensión InstitucionTipos. Elaboración propia	86
Ilustración 42: Dimensión Municipios. Elaboración propia	87
Ilustración 43: Dimensión CondicionTipos. Elaboración propia	87
Ilustración 44: Dimensión Fechas. Elaboración propia	87
Ilustración 45: Tabla de Hecho Evaluaciones. Elaboración propia	88
Ilustración 46: Modelo Dimensional. Elaboración propia	88
Ilustración 47: ETL - Carga Inicial. Elaboración propia	89
Ilustración 48: CUBO. Elaboración propia	96



<i>Ilustración 49: Nuevo Proyecto para CUBO. Elaboración propia.....</i>	<i>97</i>
<i>Ilustración 50: Nuevo Proyecto Analysis Services. Elaboración propia</i>	<i>97</i>
<i>Ilustración 51: CUBO - Orígenes de datos. Elaboración propia</i>	<i>98</i>
<i>Ilustración 52: CUBO - Asistente para orígenes de datos. Elaboración propia</i>	<i>98</i>
<i>Ilustración 53: CUBO - Administrador de conexiones. Elaboración propia.....</i>	<i>99</i>
<i>Ilustración 54: CUBO - Información de suplantación. Elaboración propia</i>	<i>99</i>
<i>Ilustración 55: CUBO - Selección de conexión. Elaboración propia</i>	<i>100</i>
<i>Ilustración 56: CUBO - Resumen de conexión. Elaboración propia.....</i>	<i>100</i>
<i>Ilustración 57: CUBO - Vistas del origen de datos. Elaboración propia.....</i>	<i>101</i>
<i>Ilustración 58: CUBO - Nueva vista de origen de datos. Elaboración propia</i>	<i>101</i>
<i>Ilustración 59: CUBO - Selección tablas y vistas. Elaboración propia</i>	<i>102</i>
<i>Ilustración 60: CUBO - Resumen de vista de origen de datos. Elaboración propia</i>	<i>102</i>
<i>Ilustración 61: CUBO - Vista de origen de datos. Elaboración propia.....</i>	<i>103</i>
<i>Ilustración 62: CUBO - Nuevo cubo. Elaboración propia</i>	<i>103</i>
<i>Ilustración 63: CUBO - Método de creación. Elaboración propia</i>	<i>104</i>
<i>Ilustración 64: CUBO - Selección tablas de grupo de medida. Elaboración propia</i>	<i>104</i>
<i>Ilustración 65: CUBO - Selección medidas. Elaboración propia.....</i>	<i>105</i>
<i>Ilustración 66: CUBO - Selección nuevas dimensiones. Elaboración propia</i>	<i>105</i>
<i>Ilustración 67: CUBO - Resumen del nuevo cubo. Elaboración propia.....</i>	<i>106</i>
<i>Ilustración 68: CUBO - Diseño de Cubo. Elaboración propia.....</i>	<i>106</i>
<i>Ilustración 69: CUBO - Proceso. Elaboración propia</i>	<i>107</i>
<i>Ilustración 70: CUBO - Procesar cubo BECAS. Elaboración propia.....</i>	<i>108</i>
<i>Ilustración 71: CUBO - Resultados. Elaboración propia.....</i>	<i>108</i>
<i>Ilustración 72: CUBO - SSMS. Elaboración propia</i>	<i>109</i>
<i>Ilustración 73: Reportes Generados. Elaboración propia.....</i>	<i>109</i>
<i>Ilustración 74: Servicios. Elaboración propia</i>	<i>110</i>
<i>Ilustración 75: Estado de servicios de Analysis Services. Elaboración propia.....</i>	<i>110</i>
<i>Ilustración 76: Conexión Base de datos SSAS. Elaboración propia.....</i>	<i>111</i>
<i>Ilustración 77: Elección CUBO. Elaboración propia.....</i>	<i>111</i>
<i>Ilustración 78: Evaluaciones Aceptadas. Elaboración propia</i>	<i>112</i>
<i>Ilustración 79: Evaluaciones Rechazadas. Elaboración propia.....</i>	<i>112</i>
<i>Ilustración 80: Dashboard. Elaboración propia.....</i>	<i>113</i>
<i>Ilustración 81: Dashboard con datos filtrados. Elaboración propia.....</i>	<i>113</i>
<i>Ilustración 82: Generación de datos - problema. Elaboración propia.....</i>	<i>114</i>
<i>Ilustración 83: Seeder problema. Elaboración propia</i>	<i>115</i>
<i>Ilustración 84: Seeder solución. Elaboración propia</i>	<i>115</i>
<i>Ilustración 85: CUBO con Dimensión CARRERA.....</i>	<i>119</i>
<i>Ilustración 86: CUBO con Dimensión PERSONA.....</i>	<i>120</i>
<i>Ilustración 87: Gestión de Inscripciones - Pagina Inicio. Elaboración propia.....</i>	<i>133</i>
<i>Ilustración 88: Nueva Inscripción - Datos Personales. Elaboración propia.....</i>	<i>133</i>
<i>Ilustración 89: Nueva Inscripción - Grupo Familiar 1. Elaboración propia</i>	<i>134</i>
<i>Ilustración 90: Nueva Inscripción - Grupo Familiar 2. Elaboración propia</i>	<i>134</i>
<i>Ilustración 91: Nueva Inscripción - Datos Académicos 1. Elaboración propia</i>	<i>135</i>
<i>Ilustración 92: Nueva Inscripción - Datos Académicos 2. Elaboración propia</i>	<i>135</i>
<i>Ilustración 93: Nueva Inscripción - Resumen Evaluación. Elaboración propia</i>	<i>136</i>
<i>Ilustración 94: Nueva Inscripción - Servicio Social. Elaboración propia.....</i>	<i>136</i>
<i>Ilustración 95: Renovaciones - Inicio. Elaboración propia.....</i>	<i>137</i>
<i>Ilustración 96: Renovación - Beneficiario. Elaboración propia.....</i>	<i>137</i>
<i>Ilustración 97: Renovación - Evaluación. Elaboración propia</i>	<i>138</i>
<i>Ilustración 98: Planes de Estudio - Inicio. Elaboración propia</i>	<i>138</i>



<i>Ilustración 99: Nuevo Plan de Estudio. Elaboración propia</i>	<i>139</i>
<i>Ilustración 100: Nueva Institución. Elaboración propia</i>	<i>139</i>
<i>Ilustración 101: Tipo de Institución. Elaboración propia</i>	<i>140</i>
<i>Ilustración 102: Selección de Provincia. Elaboración propia</i>	<i>140</i>
<i>Ilustración 103: Nueva Carrera. Elaboración propia</i>	<i>140</i>
<i>Ilustración 104: ETL - Nuevo proyecto Integration Services. Elaboración propia</i>	<i>146</i>
<i>Ilustración 105: ETL - Administrador de conexiones. Elaboración propia</i>	<i>146</i>
<i>Ilustración 106: ETL - Paquetes SSIS. Elaboración propia</i>	<i>147</i>
<i>Ilustración 107: ETL - Carga Inicial. Elaboración propia</i>	<i>148</i>
<i>Ilustración 108: ETL - Flujo de datos. Elaboración propia</i>	<i>148</i>
<i>Ilustración 109: ETL - Editor de origen de datos. Elaboración propia</i>	<i>149</i>
<i>Ilustración 110: ETL - Editor de destino de datos. Elaboración propia</i>	<i>149</i>
<i>Ilustración 111: ETL - Actualización Diaria. Elaboración propia</i>	<i>150</i>
<i>Ilustración 112: ETL - Editor de tarea Ejecutar SQL. Elaboración propia</i>	<i>150</i>
<i>Ilustración 113: ETL - Actualización anual. Elaboración propia</i>	<i>151</i>
<i>Ilustración 114: Servicio SQL Server Agent. Elaboración propia</i>	<i>152</i>
<i>Ilustración 115: Integration Services Catalogs. Elaboración propia</i>	<i>152</i>
<i>Ilustración 116: Creación catálogo SSISDB. Elaboración propia</i>	<i>153</i>
<i>Ilustración 117: Catálogo SSISDB. Elaboración propia</i>	<i>153</i>
<i>Ilustración 118: Implementar proyecto Integration Services. Elaboración propia</i>	<i>154</i>
<i>Ilustración 119: Destino implementación proyecto IS. Elaboración propia</i>	<i>154</i>
<i>Ilustración 120: Resultados implementación proyecto IS. Elaboración propia</i>	<i>155</i>
<i>Ilustración 121: Paquetes SSISDB. Elaboración propia</i>	<i>155</i>
<i>Ilustración 122: SSA crear nuevo trabajo. Elaboración propia</i>	<i>156</i>
<i>Ilustración 123: SSA Nuevo trabajo. Elaboración propia</i>	<i>156</i>
<i>Ilustración 124: SSA Nuevo paso actualización anual DW. Elaboración propia</i>	<i>157</i>
<i>Ilustración 125: SSA nueva programación actualización anual DW. Elaboración propia</i>	<i>157</i>
<i>Ilustración 126: SSA paquetes SSIS. Elaboración propia</i>	<i>158</i>
<i>Ilustración 127: SSA nueva programación actualización diaria DW. Elaboración propia</i>	<i>158</i>
<i>Ilustración 128: SSMS Credenciales. Elaboración propia</i>	<i>159</i>
<i>Ilustración 129: SSA Proxy. Elaboración propia</i>	<i>159</i>
<i>Ilustración 130: SSAS Proceso Proyecto Analysis Services. Elaboración propia</i>	<i>160</i>
<i>Ilustración 131: SSAS Script proceso proyecto Analysis Services. Elaboración propia</i>	<i>160</i>
<i>Ilustración 132: SSA nuevo paso actualización CUBO. Elaboración propia</i>	<i>161</i>
<i>Ilustración 133: SSA programación actualización CUBO. Elaboración propia</i>	<i>161</i>



Índice de Tablas

Tabla 1: Matriz FODA	31
Tabla 2: Identificación de riesgos	43
Tabla 3: Prevención de riesgos	43
Tabla 4: Contrarrestar riesgos.....	44
Tabla 5: Matriz de responsabilidades - Planificación.....	47
Tabla 6: Matriz de responsabilidades - Recolección de requerimientos	47
Tabla 7: Matriz de responsabilidades - Diseño solución BI.....	48
Tabla 8: Matriz de responsabilidades - Desarrollo	48
Tabla 9: Matriz de responsabilidades - Instalación y pruebas.....	48
Tabla 10: Matriz de responsabilidades - Finalización	48
Tabla 11: Honorarios.....	52
Tabla 12: Costos Hardware	52
Tabla 13: Costos Software	52
Tabla 14: Presupuesto Desglosado.....	53
Tabla 15: Flujo de Egresos	54
Tabla 16: VAN y TIR.....	55
Tabla 17: Comandos Artisan.....	60
Tabla 18: Migrations y Seeders.....	64
Tabla 19: Comparativa metodologías BI	78
Tabla 20: Descripción de Dimensión dimInstitucionTipos.....	144
Tabla 21: Descripción de Dimensión dimMunicipios.....	144
Tabla 22: Descripción de Dimensión dimCondicionTipos	144
Tabla 23: Descripción de Dimensión dimFechas	144
Tabla 24: Descripción de Tabla de Hechos factEvaluaciones.....	145



Definiciones, Acrónimos o Abreviaciones

DW: Data Warehouse

DM: Data Mart

OLTP: On Line Transaction Processing

ETL: Extraer, Transformar, Cargar

OLAP: On Line Analytical Processing

BI: Business Intelligence

TIC: Tecnologías de la información y la comunicación

IT: Tecnología de la información

DWH: Data Warehousing

DER: Diagrama Entidad Relación

ERP: Enterprise Resource Planning

SQL: Structured Query Language

SGDB: Sistema gestor de base de datos

BBDD: Bases de datos

SSIS: SQL Server Integration Services

SSDT: SQL Server Data Tools

SSAS: SQL Server Analysis Services

SSISDB: SQL Server Integration Services Data Bases

SSA: SQL Server Agent

SSMS: SQL Server Management Studio



Abstract

Hoy en día las empresas y las organizaciones dirigen gran parte de sus esfuerzos en revelar áreas de mejora que les permitan mejorar su desempeño, con la finalidad de mantenerse en el nivel competitivo deseado. Uno de los aspectos que mayor importancia refleja, hace referencia al uso de la información dentro de la empresa, de tal manera que a través de su eficiente administración sea posible la toma de decisiones certera y oportuna, que la conduzcan a alcanzar las metas y objetivos planeados.

El uso de nuevas tecnologías como medio para comunicarse con el ciudadano aumenta día a día. El reto al que se enfrentan los organismos públicos es transformar el gran volumen de datos de que disponen en información útil y fácil de usar para los ciudadanos. Otro gran reto es la necesidad de transparencia que no se mide tanto por la cantidad de la información sino por su calidad, lo que obliga a disponer de herramientas que permitan profundizar y flexibilizar la información.

Entre los organismos públicos que poseen un gran volumen de información sin analizar, encontramos al Subprograma de Becas Universitarias y Terciarias de la Provincia de Salta.

El proyecto en cuestión plantea la necesidad de implantar un sistema de Business Intelligence para la gestión integrada del Subprograma de Becas, basado en el estudio de los datos de los ciudadanos que solicitan este beneficio que puedan llegar o no a obtenerlo.

Como resultado permitirá generar un entorno apto en el ámbito de la toma de decisiones en el Subprograma de Becas Universitarias y Terciarias de la Provincia de Salta.



Capítulo 1: Introducción

El presente trabajado corresponde a la elaboración del Proyecto de Grado a presentar en la Facultad de Ingeniería de la Universidad Católica de Salta.

El organismo público sobre la que se referencia el mencionado trabajo es el Subprograma de Becas Universitarias y Terciarias de la Provincia de Salta. El proyecto consiste en el planteo de una propuesta para dar solución a una serie de problemas que se presentan al momento de llevar a cabo la toma de decisiones sobre los procesos generales del mismo. Todo ello pensando en la mejor manera de obtener resultados, apuntando tanto al uso presente como el futuro.

El objetivo, en forma general, es diseñar e implementar en el organismo una herramienta de *Business Intelligence* que permita automatizar la generación de reportes y consultas sobre información relevante para los usuarios del sistema.

Breve Descripción del Problema e Importancia

Para éste proyecto de grado se desarrolla una solución informática para el Subprograma de Becas Universitarias y Terciarias de la Provincia de Salta, como medida para resolver inflexiones que se presentan en la preselección, evaluación y selección de personas que soliciten becas, en lo que respecta a la administración de la información para tomar decisiones.

Esta solución informática planea ser una herramienta útil para éste organismo, a través del cual posibilitara desarrollar avances en la toma de decisiones justificadas en datos reales, actualizados y métodos confiables para la recuperación de dicha información, sin perder el foco y respetando que los datos de los cuales se obtengan los reportes sean de calidad con un alto nivel de respuesta.

La importancia del desarrollo de éste proyecto en la que se aplicarán técnicas y herramientas de ingeniería en un organismo público apunta a que el primer beneficiado sea el propio ciudadano de la provincia de Salta a través de las mejoras en los procesos para la obtención de una beca que serán aprobadas por los ejecutivos gracias a las mejoras que se obtendrán con Business Intelligence. De esta manera, una persona podría llegar a obtener más posibilidades de conseguir éste beneficio.

Además, la implementación de esta solución puede reducir los tiempos asignados a la toma de decisiones, dejando a los usuarios enfocarse en otras tareas.

Motivación

El interés y la motivación a realizar éste proyecto fue la de poder aplicar los conocimientos adquiridos durante la carrera ayudándome a consolidarlos y por lo que el éxito del mismo puede dar lugar a la inserción de éste modelo de sistemas de Business Intelligence en otros tipos de Organismos Públicos y Privados del Norte Argentino. Así mismo el poder ayudar en el progreso de las áreas educativas, motivar a miles de personas a estudiar una carrera terciaria o universitaria con el objetivo de cumplir la meta ser un profesional, dadas por una buena organización y un adecuado orden.



Procedimiento a realizar

El desarrollo de las actividades que respectan al cumplimiento de los objetivos del proyecto debe respetarse al pie de la letra, acatándose a los tiempos fijados y realizando una administración eficaz de los recursos.

Se debe realizar un análisis de requisitos como trabajo inicial para fijar cuáles serán las necesidades del organismo en el funcionamiento interno. Con el fin de conseguir un panorama de estas necesidades se procederá a realizar entrevistas con las personas que trabajan a diario en el organismo, de la misma manera se realizarán observaciones sobre el manejo del sistema para ir pensando cual será la mejor solución que se adapte a la situación actual del mismo.

La selección de las metodologías a emplear es una tarea que se debe realizar con anticipación al inicio al proyecto, de esta manera se pueden orientar los objetivos y metas en una misma dirección.

Es imprescindible reconocer los límites y restricciones sobre los recursos con los cuales se cuenta, estudiar toda la información obtenida y poder alcanzar un entendimiento beneficioso para así fijar una estrategia que pueda ser ejecutada correctamente.

Se desarrollará a través de la metodología Hefesto la construcción e implementación del Data warehouse y cubo de datos debido a que los objetivos y resultados finales de cada fase se diferencian fácilmente y son sencillos de comprender.

Se mostrarán a los usuarios los resultados de las estructuras diseñadas y los datos que devuelvan las diferentes herramientas, que se desarrollaran para verificar y controlar que es lo que realmente necesitaban para respaldar la toma de decisiones dentro de la organización, y si hace falta corregirlas.

Criterios de Éxito

El proyecto estará encaminado hacia el éxito en el momento que se den por cumplidos los siguientes factores:

- Tener una visualización de la información de lo que ha pasado en periodos de tiempos determinados de varios años lectivos.
- Ayudar a comprender la razón o las causas de lo que ha pasado en ese periodo de porque hay menos o más becados que años anteriores.
- Obtener una predicción de lo que sucederá el próximo año en la evaluación de solicitudes de becas.
- Obtener reportes de solicitudes rechazadas, aprobadas y reconsideraciones que luego serán útiles para realizar informes.
- Presentar la información en forma oportuna y consistente, para brindar soporte para la toma de decisiones.

Capítulo 2: El Estado de la Cuestión

Business Intelligence

Business Intelligence o también conocida como Inteligencia de Negocios es una herramienta TIC (Tecnologías de la Información y la Comunicación), que permite a directivos y ejecutivos la toma de decisiones, a través de la aplicación de varios de sus componentes, metodologías, aplicaciones y tecnologías que actúan sobre información que genera la propia organización. Esta herramienta permite procesar, resumir, reportar y distribuir la información relevante a tiempo, gestionando y compartiendo el conocimiento del negocio con la organización.¹

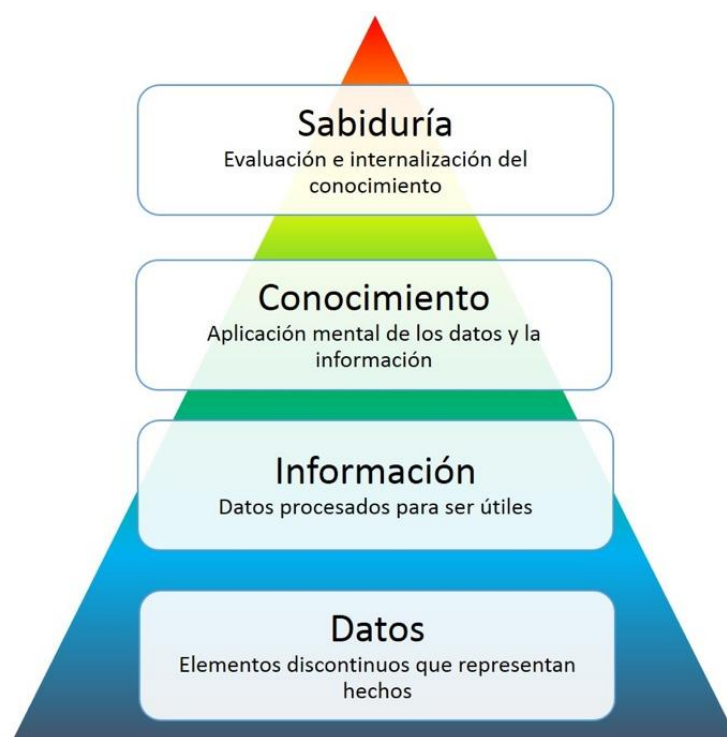


Ilustración 1: Cuanto más datos, más conocimientos. Tomado de (Hey, J., The Data, Information, Knowledge, Wisdom Chain: The metaphorical link, 2004)

El retorno positivo que ofrecen los procesos de Business Intelligence son²:

- Datos homogéneos. - Es decir, la unificación de datos de distintas soluciones y/o aplicaciones.
- Información en tiempo real. - Obtener información oportuna para la toma de decisiones.
- Compartir información.
- Identificar problemas u oportunidades.
- Simplificar el trabajo del personal informático.

¹ Josep Lluís Cano 2007. Business Intelligence: Competir con Información.

² Alberto Duchi 2017. Generalidades de la Inteligencia de negocios (Business Intelligence – BI).



Orígenes de Business Intelligence

Todas las empresas cuentan con grandes cantidades de representaciones simbólicas (numérica, alfabética, algorítmica), atributos o características de una entidad, denominado datos. Estos datos pueden ser de diferentes tipos, como productos, personas, lugares, entre otros. [Stair & Reynolds, 1999]

Los miembros de las diferentes organizaciones van a conseguir ciertas respuestas para sus negocios a partir de esos datos. Sin embargo, los datos de manera aislada pueden no contener información relevante. [Laudon & Laudon, 1997]

El uso esporádico del término inteligencia de negocios se remonta a la década de 1960. Sin embargo, El consulto Howard Dresner se acredita con la primera propuesta en 1989 como una categoría para la aplicación de técnicas de análisis de datos para apoyar los procesos de toma de decisiones empresariales³. Pero para llegar a ello, en un principio se tuvo que comenzar a realizar análisis en los pequeños y rudimentarios almacenes de datos.

Antes de que aparezcan los primeros ordenadores, los datos eran almacenados en estructuras conocidas como archivadores.

Durante la década de 1950 y comienzos de los años '60, se produjo un gran cambio con la aparición de las primeras computadoras, lugar donde los datos comenzaron a ser almacenados. Este almacenamiento resultó ser algo riesgoso y difícil de administrar. Como resultado de las actividades de almacenamiento y administración, aparece un concepto sumamente importante, el de las Base de Datos [Codd, 1970]. Terminología utilizada frecuentemente en la actualidad.

Las bases de datos proporcionaban una nueva forma para que los datos de los negocios puedan ser almacenados. Sin embargo, esta tarea no resultaba tan sencilla, porque para el almacenamiento de los datos se esperaba de cierta experiencia y conocimiento sobre el tema.

Los negocios requerían de una vía para poder llevar a cabo el almacenamiento, es por eso que en los años '70 aparecen las primeras Aplicaciones de Negocios [Turban & Aronson, 2001]. Estas aplicaciones permitían realizar la recolección de datos de una mejor manera de cómo se la trabajaba, al mismo tiempo que posibilitaban acceso a las bases de datos. Igualmente, este acceso no era fácil, debido a que los datos provenían de múltiples lugares. La obtención de la información resultaba ser un proceso muy lento, donde se creaban reportes que contenían información unidimensional, dando como resultado que ésta información se encuentre fragmentada.

Como se pretendía conseguir un acceso mucho más eficaz hacia los datos originados de múltiples localizaciones, a comienzos de los años '80 apareció un nuevo concepto, el de Almacén de Datos (Data Warehouse). El Data Warehouse (DW) permitió extraer, transformar, consolidar, integrar y centralizar todos los datos que una organización generaba en las diferentes tareas diarias que ésta realizaba [Inmon, 2005].

A partir de ahora, la información podía ser administrada y accedida de manera más simple, a través de datos útiles provenientes de diferentes lugares:

Acceso a Datos + Administración = CONSUMO → Business Intelligence 1.0

³ Extraído de <https://analyticboard.com/blog/evolucion-del-uso-de-la-inteligencia-de-negocios/>



De esta manera, se empezó a dar a conocer el concepto de Business Intelligence 1.0⁴, el cual permitía a través del acceso a los datos, y una mejor administración, generar demanda para su utilización.

Es así, como en el año 1989, se utilizó por primera vez el término Business Intelligence, como consecuencia de la aparición en el mercado de las primeras herramientas que brindaban la posibilidad de generar reportes y analizar información. Durante estos años, los vendedores de software empezaron a prometer mayor acceso de datos, a partir de múltiples lugares, lo cual permitían realizar un mejor análisis de los datos, con el único objetivo de concebir una mayor demanda.

Las organizaciones comenzaron a necesitar que los datos se generen de manera más rápida, es por eso que, en un primer momento, el BI se basó en el Rendimiento.

Con la aparición de Internet, el número de datos generados empezó a crecer de manera abrupta por tal motivo los negocios necesitaban mayores respuestas de manera más rápida.

Fue así como en los años '90 aparecieron diferentes herramientas de BI que permitieron realizar ciertos tipos de actividades, para resolver las distintas problemáticas que habían surgido:

- Herramientas para generar reportes.
- Herramientas para análisis de información.
- Herramientas para gestionar rendimientos.

Esas herramientas de BI resultaban ser difíciles de mantener, por el cual los que ofrecían este tipo de productos garantizaban mayor funcionalidad a un menor costo.

A raíz de este cambio, se puede decir que aparece la primera plataforma de BI. Las diferentes empresas existentes en el mercado, empezaron a fusionarse y comenzaron a adquirir diferentes firmas, dando lugar a nuevas compañías.

A partir de aquí se empieza a hablar de Business Intelligence 2.0⁵, caracterizándose por ser un BI más funcional, y lograr una mejor consolidación en el mercado de la Inteligencia de Negocio.

- Business Intelligence más funcionales
 - Consolidación del Mercado BI
- 

A pesar de la aparición del BI 2.0, las organizaciones no obtenían las respuestas que muchas veces deseaban o necesitaban. En esa época, los encargados de acceder a la información eran los departamentos de Tecnología de la Información (IT) de las empresas. El 70% de los miembros de las organizaciones realizaban solicitudes de datos, produciéndose ciertas frustraciones, ya que las herramientas de BI no estaban diseñadas para ser intuitivas, y el personal de IT no disponía del tiempo necesario para conseguir y analizar la información provista por las herramientas.

⁴ Extraído de <http://www.businessintelligence.info/definiciones/historia-business-intelligence.html>

⁵ Extraído de <http://www.businessintelligence.info/definiciones/historia-business-intelligence.html>

Para hacer frente a esa problemática, se comenzó a trabajar en la idea de que el BI esté más relacionado con la Utilidad, fusionándose con la característica de Rendimiento, lograda en el BI 1.0.

Las organizaciones necesitaban información útil, en el momento preciso. Se buscaba convertir datos en formatos correctos, para que a las personas les resulte más fácil el consumo de la información, como también que la administración sea más amigable para el departamento de IT.

Otro de tantos inconvenientes surgidos, estuvo relacionado con los denominados datos no estructurados, es decir, todos aquellos datos que no se encontraban en las bases de datos. Estos provenían de diferentes orígenes como: documentos, presentaciones, videos, entre otros.

Comenzó a analizarse la posibilidad de que BI necesitaba estar más centrado en las personas, y que el departamento de IT tenga la posibilidad de brindar información a un número mayor de usuarios, ya que son ellos quienes toman las mejores decisiones, y no así los sistemas.

Es así como el BI se lo fue perfeccionando a través de las necesidades que iban naciendo en las organizaciones, tratando de superar los diferentes inconvenientes surgidos desde su creación.

En la actualidad nos encontramos con numerosas e importantes herramientas de BI, las cuales brindan diferentes tipos de posibilidades a miembros de las organizaciones [Cano, 2007].



Ilustración 2: Evolución del BI⁶

Componentes de Business Intelligence

Fuente de Datos

Se podría decir que es la materia prima a partir de la cual proviene toda la información tanto de sistemas operacionales (ERP, CRM, SCM, Legacy, etc.), sistemas departamentales

⁶ Extraído de: <https://www.beeva.com/beeva-view/estrategia-negocio/un-viaje-en-el-tiempo-por-la-historia-del-business-intelligence/>

(Provisiones, Presupuestos, Hojas de cálculo, otros) y fuentes externas (Estudios de mercado, Índices de la Industria, Estadísticas y censos, Internet) ya que la herramienta debe poder reunir información completa por más que se encuentre separada para resumirla y utilizarla para análisis profundos, con bases y datos actualizados.

ETL⁷

ETL consisten en la Extracción, Transformación y Carga de la información que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, Data Mart o Data Warehouse, o en otro sistema para actuar como soporte de otro proceso de negocio.

El proceso ETL se encuentra condicionado por los requerimientos de los usuarios, quienes expresan que preguntas desea responder y quienes nos pueden indicar de donde viene esta información.

- Extraer: La primera parte del proceso ETL consisten en extraer los datos desde los sistemas de origen. La mayoría de los procesos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Los datos extraídos son analizados, de lo que resulta un chequeo para verificar si los datos cumplen la estructura que se espera conseguir. De no ser así los datos son rechazados.

Si los datos a extraer son muchos, el sistema de origen se podría ralentizar hasta inclusive llegar a colapsar, logrando que el sistema no pueda utilizarse con normalidad para su uso cotidiano. Por esta razón, en sistemas grandes las operaciones suelen programarse en tiempos determinados donde el impacto sea nulo o mínimo.

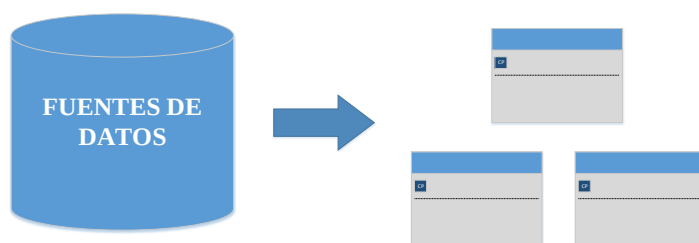


Ilustración 3: Proceso ETL – Extracción. Elaboración propia

- Transformar: Es la fase del proceso que implica un mayor esfuerzo ya que se procede con la limpieza de los datos para que cumplan con los requisitos especificados antes de iniciar el proceso.

⁷ Bernabeu R. Dario, 2010. "HEFESTO" DATA WAREHOUSING: Investigación y Sistematización de Conceptos; HEFESTO: Metodología para la construcción de un Data Warehouse.

Entre las principales transformaciones que se pueden hacer, son cambios en los tipos de datos, obtener nuevas columnas de las existentes, manejo de los nulos, entre muchas cosas.

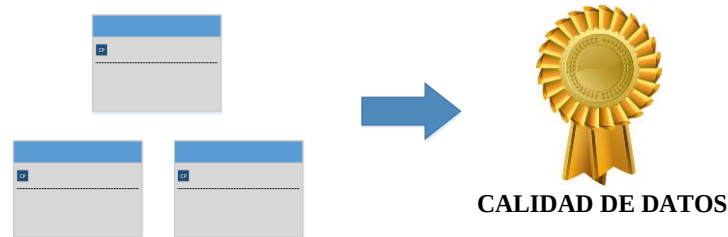


Ilustración 4: Proceso ETL – Transformación. Elaboración propia

- **Cargar:** La etapa de carga es el momento en el cual los datos de la fase anterior son cargados en el sistema de destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. Los Data Warehouse mantienen un historial de los registros de manera que se pueda hacer una auditoría de los mismos y disponer de un rastro de toda la historia de un valor a lo largo del tiempo.

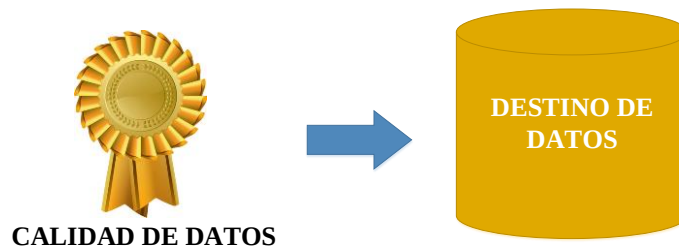


Ilustración 5: Proceso ETL – Carga. Elaboración propia

Data Warehouse

“Un Data Warehouse es una colección de información creada para soportar las aplicaciones de toma de decisiones.” [Hugh J. Watson, 2006].

“Es un conjunto no volátil de información que soporta la toma de decisiones, orientado sobre un área, integrado al tiempo.” [Bill Inmon, 2005].

“El Data Warehouse da acceso a la información de la corporación o área funcional. El alcance del Data Warehouse puede ser bien un departamento o bien corporativo.” [Ralph Kimball, 1996].

“Data Warehousing (DWH) es el proceso completo de extraer información, transformarla y cargarla en un Data Warehouse y el acceso a esta información por los usuarios finales y las aplicaciones.” [Hugh J. Watson, 2002].

“Descubres el valor real de un Data Warehouse cuando alguien puede encontrar los detalles importantes en la información, y te dice algo que puede generar la diferencia.” [Sharon Sibighroth, 2001].



Es el proceso de extraer datos de distintas fuentes para que una vez filtradas y especialmente ordenadas sean dispuestas en un depósito de datos consolidado para el análisis del negocio. Requiere una combinación de metodologías, técnicas, hardware y los componentes de software que proporcionan en conjunto la infraestructura para soportar el proceso de información. Esta estructura debe ser capaz de responder a las necesidades y características del negocio, con el tiempo se irá convirtiendo en la memoria de la empresa que permitirá describir tanto el pasado como el presente de la empresa. Un Data Warehouse no descubre ni predice.

Para la construcción del mismo se pueden establecer tres etapas; la primera está dedicada a observar y analizar el Diagrama de Entidad Relación, generando los esquemas multidimensionales que podrían ser utilizados posteriormente.

La segunda etapa, consiste en recoger los requisitos de usuario por medio de entrevistas, observaciones, encuestas y registro de informes para obtener información acerca de las necesidades de análisis de estos.

Y por último en la tercera etapa se constata la información que se obtuvo en la etapa anterior, con los esquemas multidimensionales formados en la primera etapa.

Herramientas de acceso a datos

Estas herramientas son utilizadas y aplicadas una vez que se ha desarrollado e implementado un Data Warehouse en base al mismo, con el fin de poder generar conocimiento para aquel que las utilice a partir de la información que brinde el Data Warehouse. Algunas de las herramientas que podemos mencionar son:

- **Generadores de informes:** Es el componente esencial para las personas que toman decisiones dentro de la organización, ya que les permite observar el comportamiento de la misma.
- **Herramientas OLAP:** A diferencia de las herramientas OLTP encargadas de procesar una cantidad considerable de pequeñas transacciones de captura de información que se dan día a día en la organización, las herramientas OLAP se dedican al análisis de enormes cantidades de información estática que permiten tratar la información de manera multidimensional para explorarla desde distintas perspectivas y periodos de tiempo.
El historial de datos es a largo plazo, normalmente de dos a cinco años. Las bases de datos OLAP se suelen alimentar de información procedente de los sistemas operacionales existentes, mediante un proceso ETL.
- **Herramientas Dashboards:** Está destinado a medir el desempeño de los procesos de negocio para asegurar que cumplan los objetivos establecidos, el dashboard contiene métricas que permiten medir el cumplimiento de los objetivos proporcionando información para la toma de decisiones sobre acciones que encaminen al cumplimiento de los objetivos organizacionales.

- Herramientas Scorecard: Son usados para alinear las operaciones con la estrategia del negocio. Este tipo de presentación de la información contiene principalmente los KPI (Key Performance Indicator) y está dirigido a los altos mandos.
- Herramientas de planificación, modelización y consolidación: Permiten crear planes de negocio, simulaciones, planificación, presupuestos, previsiones, etc. Proveen a los dashboards y scorecard de objetivos y umbrales de las métricas.
- Herramientas Datamining: Es la extracción de patrones de conocimiento interesantes de cantidades enormes de datos. Permite la gestión en tiempo real de manera eficaz, es una herramienta aplicable a cualquier tipo de empresa. Datamining surge para intentar ayudar a comprender el contenido de un repositorio de datos. Con este fin, hace uso de prácticas estadísticas y, en algunos casos, de algoritmos de búsqueda próximos a la Inteligencia Artificial y a las redes neuronales.

Quizás la manera más fácil de comprender este término y lo que significa el Business Intelligence sea con la siguiente figura:

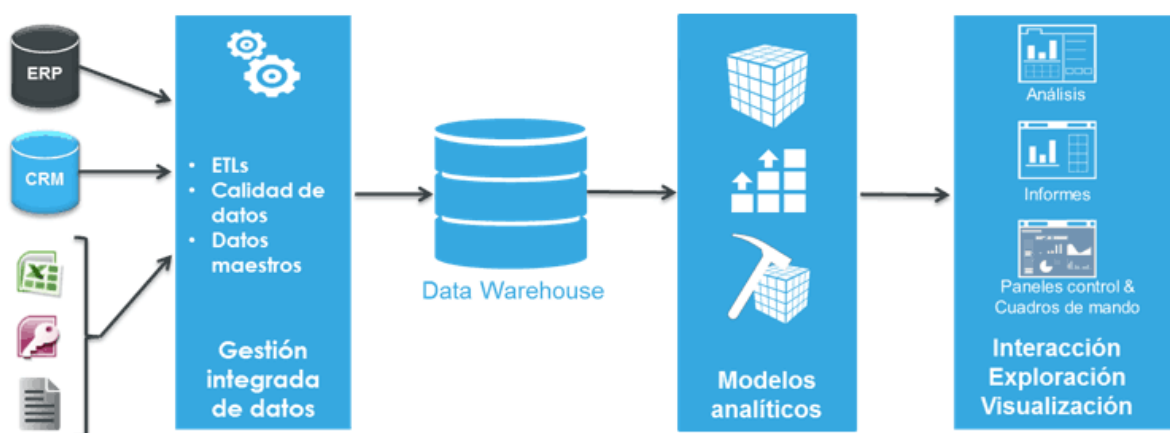


Ilustración 6: Business Intelligence⁸

¿Sector Público o Privado?

Con la consolidación del Business Intelligence en el mercado, iban surgiendo interrogantes, uno de los cuales estaba relacionado con obtener una respuesta sobre el sector en el cual se podía aplicar este tipo de solución.

¿Es el Business Intelligence propio de las empresas que se desempeñan en el sector privado?

La respuesta para ésta pregunta es NO. Walter Jones, director de WePlan, partner exclusivo del líder Microstrategy opina que “el mercado público se desarrollará más que el

⁸ Extraído de <http://aaronza.blogspot.com/2015/10/que-es-el-business-intelligence.html>



mercado privado, debido a la necesidad de gestionar de forma colaborativa volúmenes de información cada vez más grandes, más complejos y provenientes de diferentes frentes de la realidad de la gestión”⁹. En la actualidad, se puede observar con mayor frecuencia cómo diferentes gobiernos hacen uso de esta solución, la cual les permite dirigir sus estrategias en un contexto cada vez más confuso. Con éste tipo de análisis, se buscan tomar mejores decisiones y así poder establecer el camino hacia su misión, al igual que lo hacen las empresas privadas para llevar a cabo sus desiguales objetivos corporativos.

Sirkka Jarvenpaa y Thomas Davenport¹⁰ mantienen que “existen áreas importantes en las cuales las aplicaciones de análisis en organismos gubernamentales se pueden llevar a cabo”. Por ejemplo, la misión y gestión de las áreas de atención de salud, logística, gestión de riesgos, y la inteligencia a la que el análisis se ha aplicado. Si bien las posibilidades de análisis para mejorar la eficiencia y la eficacia en el gobierno parecen ilimitadas, hay mucho menos claridad acerca de la disposición del gobierno a adoptar estas herramientas y plataformas de análisis. Davenport y Jarvenpaa son cuidadosos al señalar que “el uso del análisis en inteligencia de negocios requiere innovación de gestión”.

Beneficios de BI¹¹

Las herramientas de BI pueden ofrecer una gran variedad de beneficios a las empresas incluyendo la reducción de costos, mayores ingresos, realización de pronósticos y la detección de problemas. Al usar estas herramientas, es posible para empleados comunes crear reportes y responder preguntas propias sin la necesidad de programadores o analistas. El análisis de estos reportes puede revelar costos u oportunidades ocultas.

Otros beneficios que pueden aportar la implementación de sistemas BI son:

- Reducir el tiempo mínimo que se necesita para recoger todos los datos relevantes de un tema específico.
- Automatizar la actualización de datos en la fuente integrada, a través de procesos predefinidos.
- Proporcionar herramientas de análisis que servirán para establecer comparaciones y explorar los datos, con lo cual se mejorará notablemente la toma de decisiones.
- Dar el sustento necesario para pasar de la decisión a la acción.
- Independizar a los usuarios de los tradicionales informes pre-programados a partir de análisis y reportes dinámicos y/o definidos por el mismo en dicho momento.
- Permite dar respuesta a preguntas de forma inmediata. También ayuda a la formulación de nuevas preguntas cuyas respuestas son clave para mejorar el desempeño de la organización.
- Permite acceder, analizar y monitorear directamente los indicadores críticos de la organización.

⁹ Extraído de <http://www.mercado.com.ar/notas/8018217>

¹⁰ “Strategic Use of Analytics in Government” - Introducción. Managing for Performance and Results Series, Thomas Davenport, Sirkka L. Jarvenpaa, IBM Center for The Business Of Government, 2008.

¹¹ Bernabeu R. Dario Y García M. Mariano, 2017. “HEFESTO - DATA WAREHOUSING”, Guía completa de aplicación teórico-práctica; metodología Data Warehouse.



- Identificar cuáles son los factores que inciden en el buen o mal funcionamiento de la organización.
- Se pueden detectar situaciones fuera de lo normal o potencialmente fuera de su rumbo.
- Permite predecir el comportamiento futuro con un alto porcentaje de certeza, basado en el entendimiento del pasado.
- Los usuarios podrán consultar y analizar los datos de manera sencilla e intuitiva.

Requerimientos en una Herramienta de BI

Las empresas tienen una gran presión para mantenerse competitivas y requieren de acceso de datos en tiempo real a medida que buscan oportunidades para mejorar la eficiencia operacional. Es por esto que la necesidad de métricas de desempeño visualizadas en tiempo real y dashboards con alertas, son muy necesarias.

La escalabilidad y el rendimiento son algunos de los principales requerimientos cuando se consideran herramientas que transfieren datos en tiempo real y que deben ser rápidamente presentados y analizados.

Basado en la encuesta Macig Quadrant de Gartner (2017) hay varias características que deben de poseer una plataforma de BI¹²:

- En general:
 - La plataforma debe ofrecer todas sus capacidades de manera integrada y no a través de varios productos con poca integración entre sí.
 - Brindar facilidad en el manejo de la plataforma para administrar y desplegar los datos, crear contenido, e interactuar con el contenido. Además, es atractiva visualmente (lo cual facilita el trabajo y el análisis).
 -
- Infraestructura:
 - La plataforma debe ser segura y ofrecer recursos para la administración de usuarios, auditorías, gestión de accesos, optimización del desempeño, y recuperación ante cualquier desastre que puede ocurrir.
 - Consultar, desarrollar y gestionar analytics en la nube, con base de datos que están tanto en la nube como en la empresa.
 - Facilitar a que los usuarios se conecten con datos estructurados y no estructurados, desde diferentes plataformas de almacenamiento, tanto en la empresa como en la nube.
- Gestión de Datos:
 - Ofrecer herramientas robustas y centralizadas que permitan buscar, capturar, guardar, reusar y publicar metadatos, por ejemplo: dimensiones, jerarquías, medidas, métricas de desempeños y otros.

¹² Extraído de <https://www.gartner.com/doc/reprints?id=1-3TYE0CD&ct=170221&st=sb>



- Permitir integrar, transformar y cargar datos en un motor de rendimiento autónomo, siendo capaz de indexar y gestionar cargas de datos y actualizar la programación.
 - Poder seleccionar y “arrastrar” (drag and drop) de diferentes fuentes para crear tus propios modelos analíticos. Permitir aplicar machine learning para el descubrimiento de frases, perfiles, jerarquías y combinaciones de datos.
- Análisis y creación de contenido:
 - Poder crear cuadros de mando y contenido altamente interactivos con exploración visual, análisis avanzados y geoespaciales, con la posibilidad de compartir con personas externas a la organización.
 - Permitir explorar datos de manera visual, más allá de los gráficos conocidos de pastel, barra y línea para incluir mapas de calor y diagramas de árbol, mapas geográficos, diagramas de dispersión y otros efectos especiales. Estas herramientas permiten analizar y manipular los datos visualmente, de manera más intuitiva y fácil.
 - Automáticamente encontrar, visualizar y narrar hallazgos importantes tales como correlaciones, excepciones, clusters, enlaces y predicciones en datos que son relevantes, sin la necesidad de construir modelos o escribir algoritmos. Poder explorar los datos a través de visualizaciones, narraciones generadas por lenguaje natural, búsqueda y tecnologías NLQ (Consulta de Lenguaje natural).
 - Desarrollar y entregar contenido a través de dispositivos móviles, aprovechando las capacidades nativas de los dispositivos móviles, como la pantalla táctil, la cámara y la ubicación.
- Opciones para compartir hallazgos:
 - Facilitar un kit para el desarrollador de software con API (Interfaz de Programación de Aplicaciones) y soporte para estándares abiertos, de manera de poder crear y modificar contenido analítico, visualizaciones y aplicaciones, que incluso permita incrustar en otros procesos del negocio, aplicaciones o en un sitio web.
 - Ofrecer la capacidad de publicar, implementar y distribuir resultados de análisis, al mismo tiempo que ofrece soporte para búsqueda de contenido, programación y alertas. Permitir compartir, discutir y rastrear información, contenido analítico y decisiones a través de foros de discusión, chats y notas.

Business Intelligence en el Mundo¹³

En la actualidad la gran mayoría de las organizaciones cuenta con un sistema de información que soporta gran parte de las actividades diarias propias del sector de negocios en

¹³ Scientia et Technica Año XVI, No 44, Abril de 2010. Universidad Tecnológica de Pereira. ISSN 0122-1701



donde se esté desempeñando, este sistema puede ser sencillo o robusto todo depende de las exigencias del negocio, con el transcurso del tiempo estas aplicaciones llegan a tener la historia de la organización, los datos almacenados en las bases de datos, pueden ser utilizados para argumentar la decisión que se quiera tomar.

Un estudio realizado en Europa por Information Builders Ibéric mostró el costo que tiene la falta de sistemas de toma de decisiones en las organizaciones, según estos datos, el empleado europeo medio pierde una media de 67 minutos diariamente buscando información de la compañía, lo que equivale a un 15,9% de su jornada laboral. Para una organización de 1.000 empleados que gane unos 50.000 euros al día esto equivale a 7,95 millones de euros al año de salario perdido, todo ello por la búsqueda de información para tomar una decisión. [Zúmel 2008]

El poder competitivo que puede tener una empresa se basa en la calidad y cantidad de la información que sea capaz de usar en la toma de decisiones; mediante la implementación de Inteligencia de Negocios se proporcionan las herramientas necesarias para aprovechar los datos almacenados en las bases de datos de los sistemas transaccionales para utilizar la información como respaldo a las decisiones, reduciendo el efecto negativo que puede traer consigo una mala determinación.

Dentro del mercado de las herramientas de BI se están produciendo distintos movimientos:

- Los desarrolladores de soluciones ERP están comercializando sus propios productos.
- Los fabricantes de BI se están especializando en soluciones concretas.
- Los fabricantes de motores de BBDD tienen sus propias soluciones de Business Intelligence.
- Disminución de los precios de algunas soluciones.
- Aparición de las soluciones Open Source.
- Futuras compras entre fabricantes.

Hay que estar atento a cualquier cambio que se pueda producir dentro de lo que sería el mercado de las Tecnologías de la Información y de la Comunicación.

Existen consultoras independientes que evalúan continuamente la evolución de los productos: Gartner, Forrester, etc., a las que podemos dirigirnos cuando tengamos que decidir en qué herramienta invertir.

Business Intelligence y su aplicación

Al igual que grandes y/o medianas organizaciones, hay empresas que sin tener en cuenta su envergadura y complejidad, estudian toda la información que poseen a través de la aplicación de Business Intelligence con el fin de convertir todo en conocimiento. Aquellas personas que están involucradas en niveles críticos de la organización hacen uso de este conocimiento empleándolas en la toma de decisiones y produciendo un impacto positivo en las tareas diarias de un sector o de todo el organismo.

BI es aplicado en la mayoría de las ramas de las ciencias (Astronomía, Bioinformática, Descubrimiento de medicamentos, etc.), al igual que en temas relacionados a negocios



(Marketing dirigido, E-commerce, deportes/entretenimiento, etc.). Otras aplicaciones en las que se ve involucrado el BI es asuntos de seguridad hacia la sociedad (Vigilancia, detección de crímenes, perfilar estafadores de impuestos, etc.). Por lo que se puede observar BI puede estar destinado a todo tipo asignaturas del cual se puede sacar provecho en la información que aportan y ayudar al crecimiento de una organización e incluso a mejorar el estilo de vida en la sociedad de diferentes maneras.

En este proyecto se aplicará BI a un programa gubernamental que surge con el propósito de dar una ayuda monetaria a través de la adjudicación de Becas a personas que estudian una carrera de nivel superior. Se investigó y no se encontró información sobre algún tipo de caso en el que se haya aplicado BI a un programa de Becas o similar.

Proceso de BI¹⁴

El proceso mediante el cual una organización obtiene rédito de sus datos (los utiliza de forma inteligente) se denomina Proceso BI.

A continuación, se enumeran sus etapas:

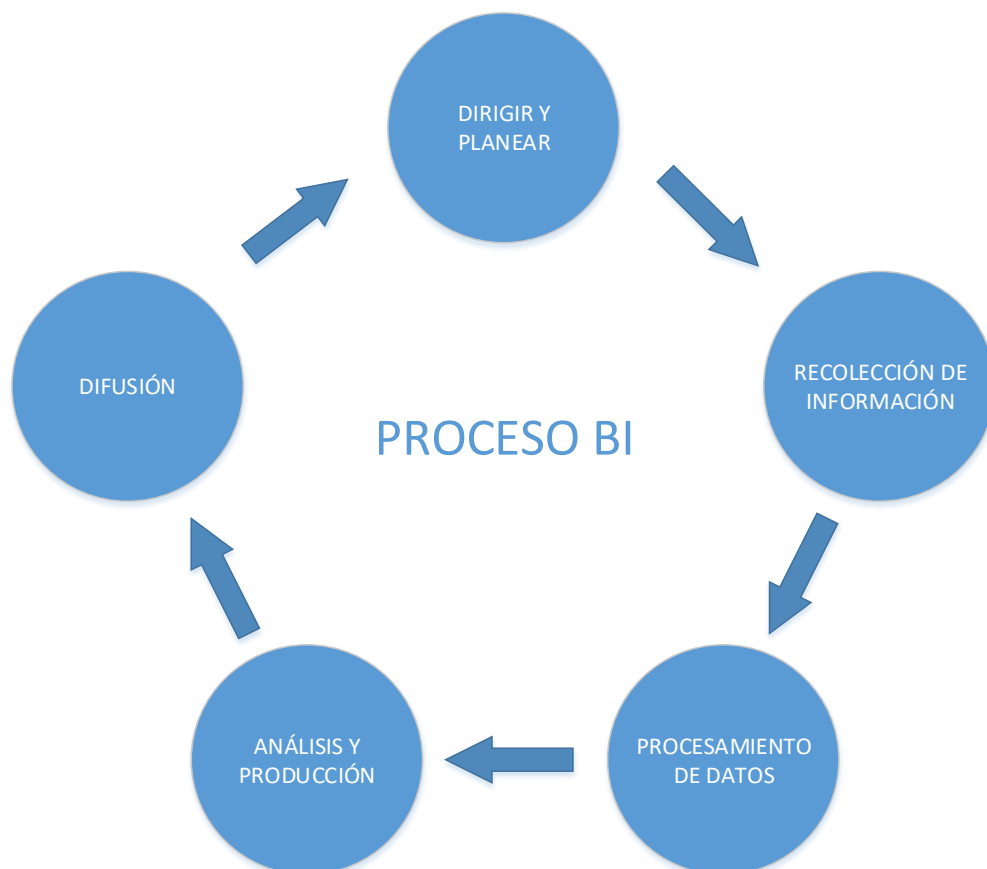


Ilustración 7: Proceso BI. Elaboración propia

¹⁴ Extraído de <http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/1-business-intelligence/13-proceso-bi>



1. **Dirigir y Planear:** Esta etapa es el inicio y el fin del proceso. Es el inicio porque involucra redactar los requerimientos específicos. Y es el final porque contesta preguntas que guían a otras nuevas. El proceso de BI empieza con los usuarios y aquí se generan las preguntas que les va ayudar a ellos a alcanzar sus objetivos. Las necesidades son presentadas por diversos analistas de negocios que formulan los requerimientos de los usuarios y dirigen un plan para la recolección de la información y para solución de las respuestas.
2. **Recolección de Información:** Hay diversas fuentes de información dentro de una compañía. La automatización de los procesos ha creado una fuente de recursos como son: Puntos de ventas, ERP, CRM, SFA, Aplicaciones de Servicios al cliente, etc. Los diferentes sistemas crean, procesan y almacenan diferentes tipos de información. Este es un proceso continuo y es importante entender que los datos de esas fuentes son simplemente información y no inteligencia. Los datos sin haberlos procesado por un algún tipo de análisis, frecuentemente son incompletos y confusos. La información se convierte en Inteligente a través de procesarla y de analizarla. El proceso de recolección es cuando las diferentes fuentes son analizadas para determinar los datos necesarios para encontrar las respuestas a las preguntas.
3. **Procesamiento de Datos:** Esta etapa es la integración de datos sin procesar a un formato utilizable para el análisis. Esto puede ser posible, creando una nueva base de datos, agregar datos o base de datos existente o consolidando información. Esta etapa generalmente es vista como ETL (Extracción, Transformación y Carga) que ocurre en los ambientes de BI.
4. **Análisis y Producción:** El grupo de análisis de negocios utiliza herramientas y técnicas para ordenar sobre los datos y crear inteligencia. El resultado final es la producción de respuestas “inteligentes”, en un contexto propio. En algunos casos es un proceso simple como la creación de un reporte. En otros casos, son la creación de indicadores. Tal vez en esta etapa, sean generados requerimientos adicionales pues los analistas puede que encuentren nuevas preguntas que necesiten ser contestadas.
5. **Difusión:** Esta etapa de difusión, es entregar productos inteligentes a los diversos clientes que lo requieren. Esto básicamente implica el uso de herramientas BI para la publicación de gráficos, reportes o la posibilidad de tener herramientas de fácil uso para que los mismos usuarios tengan la capacidad de revisar los datos de manera rápida y sencilla.



Capítulo 3: El Problema

Definición del Problema

El Subprograma de Becas Provinciales está dedicado a brindar becas universitarias y terciarias a personas que pertenezcan al Territorio de la provincia de Salta y cumplan con una cierta cantidad de requisitos personales, académicos, socioeconómicos y del grupo familiar. Al ser una beca que proporciona el estado, puede correr el riesgo de ser disuelta por el mismo gobierno de turno como consecuencia de que cada vez haya menos aspirantes y/o beneficiarios de la misma.

La problemática surge por el hecho de no existir un sistema de toma de decisiones que brinde sondeos sobre lo que ocurre en cada año, por lo que resulta difícil saber en qué puntos y por qué razones las personas no obtienen o deciden abandonar dicha beca.

Objetivos

Objetivo General

A través del desarrollo de las herramientas y tecnologías que proporciona el *Business Intelligence* se busca generar conocimiento sobre los hechos que competen al organismo de Subprograma de Becas Provinciales implementando un sistema de apoyo de toma de decisiones que proporcione información sobre los datos que ya poseen.

Para ellos se busca diseñar e implementar en dicho organismo una herramienta de *Business Intelligence* que permita automatizar la generación de reportes y consultas sobre información relevante para los usuarios del sistema.

Objetivos Específicos

- Analizar la relación que existe entre la información del programa de becas universitarias y terciarias y toma de decisiones.
- Determinar los principales elementos de los requerimientos que la organización busca en una herramienta de *Business Intelligence*.
- Identificar los procesos y/o métodos para definición, manipulación y explotación de información del programa de becas universitarias y terciarias.
- Desarrollar un dashboard basado en la información del programa de becas universitarias y terciarias.
- Diseñar e implementar un prototipo funcional que permita satisfacer la principal necesidad del organismo.

Organismo sobre el que se referencia el trabajo

Misión del Organismo



El Subprograma de Becas busca:

- a) promover la igualdad de oportunidades en el ámbito de la educación superior, a través de la implementación de un sistema de becas que favorezca el acceso, permanencia y egreso de estudiantes con particulares condiciones de vulnerabilidad socioeconómica familiar y que presenten un rendimiento académico y regularidad en sus estudios en los términos exigidos por el presente;
- b) fomentar la promoción efectiva del nivel medio o secundario, y
- c) contribuir al desarrollo de la provincia, promoviendo la formación de recursos humanos. -

Visión del Organismo

Un organismo al servicio del estudiante, con tecnología de avanzada para responder ante las necesidades de la población estudiantil de manera ágil, oportuna, de calidad, mediante procesos de diagnóstico, análisis, reflexión y toma de decisiones, con personal actualizado y en formación permanente.

Análisis FODA

Fortalezas

- **F1.** Ventajas en costos.
- **F2.** Genera cambios reales en la toma de decisiones al implementar el sistema de Business Intelligence.
- **F3.** El proyecto podrá ser montado sobre el equipo existente del programa de Becas.
- **F4.** El proyecto podrá utilizar la licencia de SQL Server que posee la organización.
- **F5.** Mejor comprensión de la información que revela la aplicación de BI a los datos que posee la organización.
- **F6.** Experiencia en el rubro por parte de los administrativos.
- **F7.** Posibilidad de cruzar datos de diferentes fuentes para profundizar el análisis.

Debilidades

- **D1.** Falta de capacidades sobre manejo de herramientas informáticas.
- **D2.** Inconsistencia de datos a la hora de analizar la base de datos original.
- **D3.** Poca motivación por parte del directivo para llevar adelante la implementación del proyecto.
- **D4.** Falta de dirección en la toma de decisiones.
- **D5.** Poca integración con herramientas basadas en redes sociales.
- **D6.** Página web obsoleta en caso de integración del programa con la aplicación de Business Intelligence.



Oportunidades

- **O1.** Expandir el programa a nivel nacional.
- **O2.** Incentiva una mejora al nivel educativo.
- **O3.** Entrar en otros segmentos del sector público.
- **O4.** Alta disponibilidad de datos para tomar decisiones en todo momento y en todo lugar.

Amenazas

- **A1.** Incremento de las solicitudes para una beca y no poseer el equipo necesario para afrontarlo.
- **A2.** Cambio de coordinador y/o administrativos.
- **A3.** Suspensión del programa por cambio de gobierno.
- **A4.** Distintos modelos conceptuales entre el prototipo y la versión con la base de datos original.

Matriz FODA

MATRIZ FODA	Fortalezas	Debilidades
	F1 F2 F3 F4 F5 F6 F7	D1 D2 D3 D4 D5 D6
Oportunidades	Estrategias FO	Estrategias DO
O1 O2 O3 O4 O5	Afianzando F2 y F5 podremos demostrar que este programa tiene todos los puntos necesarios para expandirse a un nivel nacional y que el número de beneficiarios sea aún mayor.	Expandir el programa a nivel nacional podría realizar un cambio de actitudes de parte de los directivos e involucrarse más dentro del programa y transformar D3 y D4 en fortalezas.
Amenazas	Estrategias FA	Estrategias DA
A1 A2 A3 A4	Debido a que en este momento no se afrontan grandes costos para la solución BI no será un gran desafío hacer frente al aumento de solicitudes para una beca ya que el aumentar equipamiento no implicara un gasto que afecte al programa. Demostrar que el programa produce un cambio positivo (F2 y F5) se podrá mantener la solución (A2 y A3) y lograr que A4 no sea una molestia.	Superando D1, D3 y D4, y mostrando a la sociedad que el programa junto a la solución BI son realmente útil podre minimizar A2 y A3.

Tabla 1: Matriz FODA. Elaboración Propia



Sistema de Información Existente

Las oficinas del Programa de Becas cuentan con un servidor principal que se encuentra dentro de una red LAN con conexión a Internet.

Cada persona que trabaja dentro del Programa de Becas cuenta con una computadora personal, también conectadas a la misma red LAN en la que se encuentra el servidor antes mencionado, que es utilizada para realizar sus tareas diarias, además hay dos computadoras que se utilizan para realizar la carga de datos de las personas que se inscriben o desean renovar la beca.

Poseen un sistema desarrollado en Visual Basic 2005 con Microsoft SQL Server 2005 como base de datos, que permite cargar los datos de las personas que se postulan a dicha beca con un módulo de evaluación, para ver quiénes serán los pre-seleccionados y posteriormente becados del corriente año.

Una vez que una persona es evaluada, los resultados son mostrados automáticamente en la página web del Programa de Becas, pudiendo ser visualizadas únicamente accediendo con un usuario y contraseña de cada persona inscripta en el Programa.

No existe un sistema o forma que les permita obtener reportes o información relevante sobre la base de datos que maneja el sistema de carga y evaluación. Los reportes que a veces deben ser realizados son trabajados a simple vista de los datos cargados, por lo que resulta tedioso para las personas que trabajan ahí. La persona principal encargada de realizar reportes son los administrativos que son destinadas al Coordinador para que el las analice y las presente al Ministerio de Educación, Ciencia y Tecnología.

Sistema de evaluación actual

Las evaluaciones que realiza el sistema están basadas en los datos que las personas deben proporcionar al completar el formulario Inscripción o Renovación según la situación particular de la persona, que deben descargar de la página web del Programa de Becas.

La postulación consiste en llenar un formulario en la que se piden datos personales, académicos, socioeconómicos y familiares, además se deberán presentar documentos adicionales expedidos por otros organismos.

Según serán preseleccionadas aquellas personas con nacionalidad argentina menores de 26 años y tengan por lo menos tres años de residencia en la provincia, también deben poseer un promedio general igual a 7 o mayor, sin adeudar ninguna asignatura previa del colegio. Otro ítem importante en la selección es que el ingreso del grupo familiar no debe superar ciertos límites dependiendo la cantidad de personas que conformen dicho grupo.

Tampoco debe poseer otro beneficio de carácter similar.

Al ser seleccionado se acreditan 4 cuotas mensuales y se realiza un control académico en el mes de agosto que debe ser completado por el beneficiario y si llegase a cumplir se otorgan las otras 4 cuotas restantes mensuales.

Para consultar el formulario de Inscripción o Renovación ver en **Anexo II**.



Infraestructura Informática

El servidor principal que se encuentra en las mismas oficinas que el resto de las computadoras que usan el sistema. Dicho servidor tiene las siguientes características:

- Software base “Windows Server 2003 Enterprise Edition Service Pack 2”.
- Conexión a red local e Internet.
- Procesadores Intel Core 2.
- 4GB de memoria RAM.
- 2TB de Disco Rígido.
- Mouse y teclado.
- La base de datos está desarrollada en Microsoft SQL Server 2000.

Las computadoras de los administrativos, técnicos y servicio social que son los que manejan la mayor parte del tiempo el sistema, tienen las siguientes características:

- Software base “Windows XP SP2” o “Windows 7”.
- Conexión a red local e Internet.
- Procesadores Intel Pentium Dual Core.
- 2GB de memoria RAM.
- 250GB de Disco Rígido.
- Mouse, teclado e Impresora.

Restricciones o Limitaciones de la Investigación

Una de las principales restricciones es impuesta por la Ley de Protección de Datos Personales 25.326 que impide poder trabajar con los datos reales de las personas que se encuentran en la base de datos principal de la organización para poder desarrollar la solución informática.

Otra restricción importante es la falta de apoyo del Coordinador del Programa de Becas para obtener la estructura de la base de datos ya sean, DER (Diagrama de entidad-relación), diccionario de datos, etc., sin hacer intromisión en los datos reales que se encuentran reflejados en la misma, por lo que se deberá crear y completar una base de datos ficticia basada en el estudio de cómo funciona la aplicación de carga y evaluación que utilizan actualmente.

A su vez la tecnología de la infraestructura del programa puede ser muy obsoleta, pero al trabajar con tecnologías nuevas para poder desarrollar Business Intelligence en este organismo no habrá un impacto muy grande que afecte al proyecto.

Se pretende incursionar en una capacitación del personal, siendo que los usuarios poseen un conocimiento limitado en cuanto al uso de sistemas de Business Intelligence, es por ello que el plan de implementación referente a la capacitación de los usuarios se extiende en el tiempo.

Así mismo no se podrá afectar las tareas diarias del personal para desarrollar el proyecto y tampoco ocupar horas laborales para capacitarlos.



A alcance

Con el fin de desplegar una solución al problema en cuestión, y cumplir con los requisitos en el Programa de Becas, se implementó una solución basada en el concepto de Business Intelligence haciendo uso de tecnologías de Data Warehouse, análisis OLTP y reportes.

En primera instancia, se implementó una solución que permite obtener información concisa e histórica mediante la implementación de un Data Warehouse formado por un Data Mart de Inscripción de Ingresantes y Reinscriptos que se dan año a año. Este almacén de datos garantiza responder a las interrogantes de negocio:

- Número de evaluaciones aceptadas de cada tipo de institución de cada municipio en un tiempo determinado.
- Número de evaluaciones rechazadas de cada municipio que no cumplen algún tipo de condición y/o requisito en un tiempo determinado.

Los Data Marts fueron diseñados y desarrollados a través de la metodología de HEFESTO. Las fases de desarrollo de esta metodología fueron uno de los puntos por lo que fue elegida y permitió desarrollar los conocimientos adquiridos en la materia Análisis Estratégicos de Datos.

El desarrollo de la base de datos ficticia fue realizado en el framework Laravel 5.4 junto con MySQL lo que permitió poder generar datos aleatorios. La migración de la base de datos de MySQL a SQL Server se obtuvo con Microsoft SQL Server Migration Assistant for MySQL. El DW fue implementado con la herramienta SQL Server Data Tools para Visual Studio 2013. El análisis de los datos y creación de cubos multidimensionales fue realizado con la misma herramienta. La explotación de este análisis fue desarrollada con la aplicación Power BI, por lo que se diseñó un proceso de capacitación al usuario para la creación y manejo de tablas dinámicas, que permitirá la personalización de sus propios reportes y cálculos.

Este software permitirá ser la entrada principal del Data Warehouse propuesto, el cual se integrará completamente con la base de datos real, en la que se podrá observar un circuito completo. Las operaciones se almacenarán en la base de datos y luego, mediante algoritmos de extracción, transformación y carga se integrará con el almacén de datos que permite mantener información histórica.

Alternativas Tecnológicas

Al tener en cuenta los objetivos, el alcance y las restricciones del proyecto mencionadas en secciones anteriores se pueden considerar un conjunto de alternativas tecnológicas, las cuales deberán ser adecuadamente elegidas que también que se acoten a las mismas necesidades de la organización.

Alternativas de Solución



Como parte del proceso de resolución de problemas llegamos al momento en que tenemos que generar alternativas de solución, las cuales luego de un proceso de evaluación nos llevarán a tomar una decisión sobre la solución a implementar.

Para que las alternativas de solución sean válidas deben ser operativas y concretas, es decir que puedan ejecutarse de manera efectiva, de las cuales podemos desatacar las siguientes alternativas:

- Data Mining: exploración y análisis de grandes cantidades de datos con el fin de descubrir patrones con significado
- Business Intelligence – Cubo: estudio de datos en función de múltiples dimensiones.
- Business Intelligence – Dashboard: grafica basada en la información de la organización que permite visualizar de manera general la situación de la misma.
- Business Intelligence – Scoreboard: cuadro de mando que condensan grandes volúmenes de información para determinar si una compañía consigue sus objetivos.

Alternativas del motor de Base de Datos

Seleccionar el motor de Base Datos es importante sobre todo al momento de pensar en el tipo de servicio que se ofrecerá a los usuarios y el rendimiento que necesitamos, sin ella la solución no podría funcionar, por eso es una de las piezas fundamentales en cualquier proyecto. Además, no debemos pensar solo en el momento de que la instalamos sino también en el futuro, al crecer el volumen crecerán los problemas y puede traducirse en procesos con un consumo muy elevado de recursos.

1. SQL Server 2014: Es un sistema de administración y análisis de base de datos relacionales de Microsoft, de las más seguras del mundo y su sistema de almacenamiento permite un rendimiento de consultas muy superior a lo habitual. Todos los procesos de análisis, consulta, limpieza, formateo de datos y acceso se realizan a una gran velocidad.



2. MySQL 5.0: Herramienta de gestión libre, que presenta gran dinamismo en el desarrollo web. Presenta una gran cantidad de tipos de datos que brindan mayor flexibilidad en el diseño de tablas.





3. Oracle PL/SQL: Las unidades de programa PL/SQL son compiladas por el servidor de Oracle Database y almacenadas dentro de la base de datos. Y en tiempo de ejecución, tanto PL/SQL como SQL se ejecutan dentro del mismo proceso de servidor, brindando una eficiencia óptima. PL/SQL hereda automáticamente la solidez, seguridad y portabilidad de la base de datos Oracle.



4. Access 2013: Access es un gestor de datos que recopila información relativa a un asunto o propósito particular. Está pensado en recopilar datos de otras utilidades (Excel, SharePoint, etc.) y manejarlos por medio de las consultas e informes.



5. PostgreSQL 11: Es un potente sistema de base de datos relacional de objetos de código abierto que utiliza y amplía el lenguaje SQL combinado con muchas características que almacenan y escalan de forma segura las cargas de trabajo de datos más complicadas.



6. IBM Informix: Ofrece un sistema de base de datos híbrido capaz de soportar datos relacionales y no relacionales, que le da la capacidad de almacenar JSON y tablas relacionales en el mismo motor de almacenamiento.



7. SQLite: Es un motor de base de datos SQL incorporado. A diferencia de la mayoría de las otras bases de datos SQL, SQLite no tiene un proceso de servidor por separado.



8. MariaDB: MariaDB es una rama compatible hacia atrás de MySQL Database Server. Incluye soporte para la mayoría de los motores de almacenamiento Open Source, y además para el propio motor de almacenamiento Maria.



Alternativas de Herramientas BI

Estas son las herramientas que se encuentran actualmente en el mercado que serán utilizadas para cumplir con la solución BI elegida.

Si bien la metodología propuesta es absolutamente independiente de herramientas de software específicas, se presenta un conjunto de herramientas que satisfagan las necesidades específicas.

1. Pentaho: Es una suite de Business Intelligence pionera que ofrece servicios críticos como calendarización, seguridad, integración, navegación por contenido y provee otras funcionalidades (Data Integration, Reporting, Analysis, Dashboard, Data Mining).



2. MicroStrategy: Provee soluciones a clientes de cualquier industria y/o área funcional con el fin de ayudarlos en la obtención de un mayor conocimiento sobre la información manejada en su empresa.



3. COGNOS: Es un software que ofrece la funcionalidad de análisis y toma de decisiones. Cuenta con una herramienta especial para modelación, pronóstico del negocio.



4. Microsoft SQL Server Data Tools Business Intelligence: Es una herramienta de desarrollo para crear base de datos SQL Server, bases de datos SQL de Azure, modelos de datos de Analysis Services, paquetes de Integration Services e informes de Reporting Services. Permite diseñar e implementar cualquier tipo de contenido de SQL Server con la misma facilidad con la que se desarrollaría una aplicación en Visual Studio.



5. Microsoft Power BI Desktop: está diseñado para el analista. Combina visualizaciones interactivas de vanguardia, con consultas y modelado de datos líderes en la industria incorporados. Permite crear y publicar informes en Power BI. Power BI Desktop ayuda a capacitar a otros con información crítica oportuna, en cualquier momento y en cualquier lugar.



6. Jaspersoft: Es una plataforma de Business Intelligence que está destinada para generar soluciones en pequeñas y medianas empresas. Esta plataforma cuenta con herramientas que permiten hacer el desarrollo completo de una solución de BI, partiendo desde la extracción de la información de las fuentes de datos para ser almacenadas en un repositorio de datos, para que posteriormente estos sean explotados con herramientas de análisis para luego ser visualizados por los usuarios.



Alternativas de Hardware

Se destinará una PC para la gestión que se encargara de manejar todo lo relacionado a la solución.



- Placa Madre
 1. Asus H11m-d Box
 2. Gigabyte Intel
 3. Asrock 1151
- Procesador
 1. Intel Core i3
 2. Intel Core i5
 3. Intel Core i7
- Memoria RAM
 1. Kingston 4GB DDR4
 2. Kingston 8GB DDR4
 3. Kingston 16GB DDR4
- Disco Rígido
 1. Western Digital Blue 1TB
 2. Toshiba 1TB
 3. Seagate 1TB

Alternativas de Software Base

1. Windows 7: Sistema operativo de Microsoft con licencia paga, que se encuentra dentro de los operativos más estables de la familia de SOs Windows, que se caracteriza por ser más rápido que sus antecesores (Windows XP y Vista), consume menos recursos y es fácil de usar.



2. Windows 8.1: Sistema operativo planteado para cambiar totalmente la interfaz de las versiones anteriores, puede resultar un poco costoso para los usuarios acostumbrados a las interfaces clásicas de Windows y con un mayor consumo de recursos.



3. Windows 10: Es el último sistema operativo desarrollado por Microsoft, con optimización de recursos gráficos, ofrece una gama de funciones potentes para las necesidades de la organización, todo en un paquete seguro.





4. Linux Debían: Es un sistema operativo libre que puede instalarse en una computadora de medios/bajos recursos. Las personas no están muy habituados a este tipo de sistema.



5. Linux Ubuntu: Es un sistema operativo libre enfocado a la facilidad de uso e instalación, pensado para el usuario promedio.



6. Linux Mint: Es una distribución de Linux de escritorio más populares y utilizada por millones de personas, ya que es extremadamente fácil de usar. Es a la vez gratuito y de código abierto



7. MacOS: Sistema operativo creado por Apple para computadoras Apple, tienen un mejor funcionamiento y pueden procesar información más rápido. Poseen un alto costo.



Elección de las Alternativas

La capacidad para desarrollar alternativas es a menudo tan importante como la habilidad para elegir correctamente entre ellas. A la hora de elegir entre las diferentes alternativas vamos a seguir como referencia los factores limitantes que se interpone en el camino del cumplimiento de un objetivo deseado.

Luego del análisis detallado se detectó la necesidad del poder analizar los datos y no crear nuevo conocimiento en esta etapa, por lo cual no se elige Data Mining y se identifica como óptimas soluciones, desarrollar Business Intelligence – Cubo y Dashboard, que estará



dirigido a un procesamiento global de los datos de la organización a través de indicadores que serán elegidos estratégicamente.

El motor de base de datos que dispondrá esta solución es el SQL Server 2005 debido a que es el motor que se está utilizando actualmente en la organización, el administrador en cambio será SQL 2014, ya que permitirá contralar grandes almacenes de datos, por su velocidad, eficiencia y por la posibilidad de que la organización en un futuro desee migrar sus datos a un motor con una versión más actualizada y no se produzcan incompatibilidades de versiones diferentes.

Para el desarrollo de la solución BI se opta por las herramientas de Microsoft, SQL Server Data Tools (SSDT) y Power BI Desktop, que permitirán implementar los diseños y mostrar los resultados del estudio de la organización y satisfacer las necesidades. Ambos se seleccionaron principalmente por la compatibilidad con el sistema gestor de base de datos (SGDB) que actualmente posee la organización. Además, Power BI resultara muy fácil de comprender para los usuarios.

En lo que respecta a Hardware para el desarrollo de la solución se adquirirá una PC con los siguientes componentes:

- Placa madre Asus H11m-d Box
- Procesador Core I5
- Memoria RAM Kingston DDR4 8GB
- Disco Rígido Western Digital Blue 1TB

El sistema operativo apto para la solución será Windows 10 ya que es la versión más actualizada, cuenta con soporte y consume menos recursos que versiones anteriores de Windows, a pesar de que Windows XP y 7 sean los más usados por la organización, se opta por Windows 10 para agilizar la potencia de procesamiento de los programas que serán utilizados para la solución que cumplirán los objetivos propuestos.

Windows Server 2003 se mantiene y seguirá siendo utilizado como software base para el servidor ya que hasta el día de la fecha no fueron reportados inconvenientes, además cuentan con las licencias necesarias.



Capítulo 4: Solución Propuesta

Justificación

Al Subprograma de Becas Provinciales le resulta importante implementar una herramienta capaz de dar solución permanente a la problemática, para evitar que aquellos que si poseen el beneficio o deseen obtenerlo en años posteriores no puedan hacerlo y no puedan afrontar los gastos de desarrollar un estudio de nivel superior, por el solo hecho de la reducción de presupuesto a este programa o por la disolución del mismo.

Es por lo antes expuesto que el organismo requiere un sistema de soporte para la toma de decisiones basado en Business Intelligence.

El desarrollo de este proyecto permitirá a este organismo:

- Mayor entendimiento de la información para identificar oportunidades y problemas.
- Los usuarios serán capaces de acceder y apalancar una vasta cantidad de información y analizar sus relaciones y entender las tendencias que últimamente están apoyando las decisiones de negocio.
- Incrementar la eficiencia organizacional y la efectividad.

La finalidad de este proyecto es abastecer al Subprograma de Becas de una herramienta para acceder a los datos de este organismo, que residirán en almacenes de datos operacionales para proporcionar reportes, análisis y visualizaciones a los usuarios.

Gestión de los Riesgos del Proyecto

Consideramos fundamental desarrollar un plan de riesgo que pudiera minimizar la incertidumbre de ejecución del proyecto.

Identificación de los Riesgos

Se identificarán los riesgos y se calificarán para determinar su gravedad; dicha calificación se realiza con base en dos variables: probabilidad de ocurrencia y la importancia de que ocurra, obteniendo el peso que tendría de que ocurriese. La probabilidad de ocurrencia varía en un rango de [0; 1] y la importancia en un rango de [0; 10].

Nº	Descripción	Probabilidad de ocurrencia	Importancia	Peso
R1	Posibilidad de que se destruyan, dañen, borren, deterioren, alteren o supriman datos.	0.3	10	3
R2	No conseguir el DER de la base de datos original.	0.5	10	5
R3	No poseer el diccionario de datos.	0.6	6	3.6
R4	No contar con la base de datos.	0.6	9	5.4
R5	Rotación de personal en la dirección de la organización.	0.2	6	1.2
R6	Cambios en el mandato del gobierno.	0.8	5	4



R7	No poder ejecutar tareas programadas ETL en SQL Server Express.	0.4	7	2.8
R8	La aplicación de BI no se usa.	0.3	9	2.7
R9	Escaso involucramiento de los usuarios finales que les lleva a sentir cierta frustración con los resultados obtenidos.	0.3	9	2.7
R10	Desconocimiento de las herramientas.	0.3	9	2.7
R11	Los datos de origen no están limpios.	0.7	7	4.9
R12	Fallas con instalaciones y configuraciones de Software.	0.7	9	6.3
R13	Equipamiento defectuoso.	0.2	7	1.4
R14	Metodología de BI no se adapta a la organización.	0.4	6	2.4

Tabla 2: Identificación de riesgos. Elaboración Propia

Prevención de Riesgos

Conjunto de actividades, medidas adaptadas o previstas con el fin de evitar o disminuir las posibilidades de que un riesgo se manifieste y afecte al proyecto.

Nº	Método de prevención
R1 R2 R3 R4	Realizar un acuerdo, posiblemente un contrato para asegurar la obtención del DER, diccionario de datos o la base de datos, indicando que no se va a hacer mal uso de las mismas y asegurando la integridad de los mismos.
R5	Dejar por escrito con firma de un directivo actual, de que se realizara el proyecto sin intervención de la próxima persona que tome ese cargo.
R6	Dejar implementado el proyecto antes de un cambio de gobierno.
R7	Instalar y configurar un SQL Server que contenga la característica de SQL Agent.
R8 R9	Analizar cuáles serían las causas de porque no se usa y realizar un plan de contingencia, ya sean capacitaciones, pruebas, cambios en el aplicativo.
R10	Usar herramientas que manejen la mayoría de los usuarios finales.
R11	Entrevistar al DBA y conseguir un informe detallado de cuáles son los campos que pudieran a llegar a tener los mayores defectos en la base de datos.
R12	Capacitarme y realizar pruebas de las instalaciones y configuraciones en otro equipo anteriormente.
R13	Soporte técnico para dichos equipos.
R14	Investigación y estudio de todas las metodologías posibles que se adecuen a la organización.

Tabla 3: Prevención de riesgos. Elaboración propia

Contrarrestar Riesgos

Se presentan los métodos que pueden ser implementados en caso de que unos de los posibles riesgos lleguen a presentarse, nos permitirán saber cómo actuar de manera astuta y que permita proseguir normalmente el desarrollo y funcionamiento de la solución.

Nº	Método para contrarrestar
R1	Se realizara un backup diario de la base de datos original para evitar que cualquier de los datos puedan ser alterados de cualquier forma.



R2 R3 R4	Al no contar con ninguna de las tres, se estudiara el sistema de información que se encuentra conectado a la base de datos del programa, se analizara el reglamento y estatutos, se entrevistara a los usuarios sobre cómo actúa el sistema a posibles comportamientos y se procederá a construir una base de datos desde cero, comenzando con un DER, eligiendo los tipos de datos de cada campo que mejor se adapten al mismo y cargar con datos de pruebas.
R5	Presentar el aplicativo BI al nuevo directivo y mostrar los impactos positivos que tuvo al introducirlos dentro del programa.
R6	Presentar el proyecto en otro programa de becas u otra parte del sector público que se encuentre interesado en esta solución.
R7	Utilizar tareas programadas que se ejecuten a través de la terminal de Windows.
R8	Aplicar la solución BI en otras plataformas, por ejemplo: en la página web para que pueda ser consultada en otros momentos fuera del horario de trabajo.
R9	Realizar más capacitaciones sobre el funcionamiento de la solución BI y avocarnos más sobre cuáles serían los impactos positivos sobre su utilización.
R10	Capacitación sobre el funcionamiento de las herramientas utilizadas.
R11	A través el proceso ETL aseguramos la consistencia de los datos que llegaran al almacén de datos.
R12	Instalar y configurar las herramientas una por una para no tener todos los posibles errores al mismo tiempo.
R13	Comprar nuevo equipamiento.
R14	Investigar casos que hayan utilizado la misma metodología y aplicar tareas que puedan facilitar el desarrollo de la aplicación BI.

Tabla 4: Contrarrestar riesgos. Elaboración propia



Estructura de Desglose de Trabajo (EDT)

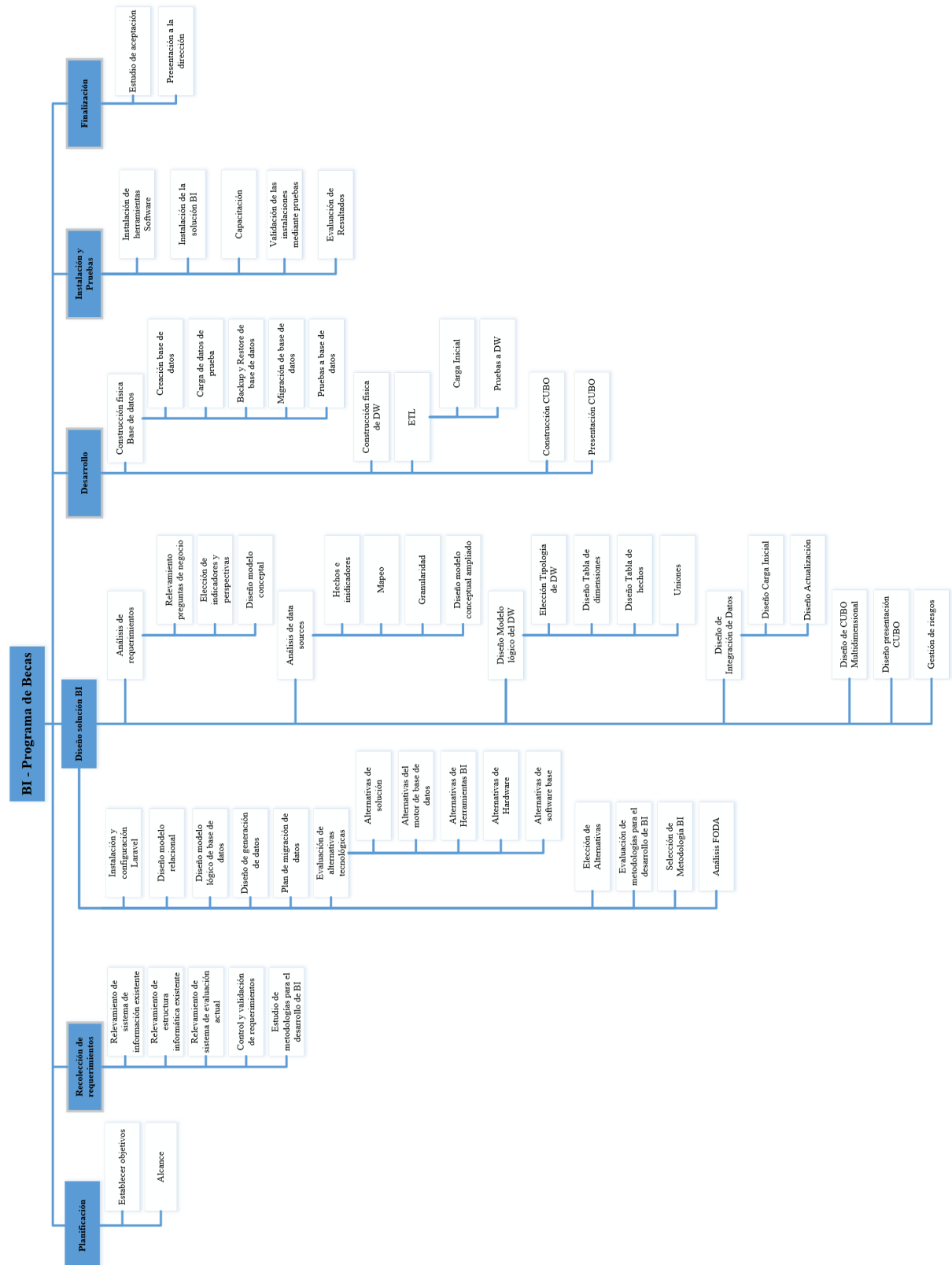


Ilustración 8: EDT. Elaboración propia



Equipo de trabajo

- **Director de proyecto:** Se encargará de supervisar todos los aspectos generales del proyecto, sin profundizar demasiado en cada tema específico, pero siempre manteniendo el orden, la organización y la coordinación entre los subordinados.
- **Analista Funcional:** Estará a cargo de sus subordinados (Desarrollador de aplicación y técnico) supervisando, organizando y coordinando su desempeño. Tendrá una alta participación en el desarrollo de la solución: es el principal actor de diseño y análisis, como así también, el gestor de ideas y quien se encargará de la recolección de información útil y fundamental para los programadores.
- **Desarrollador de Aplicación:** Esta pura y exclusivamente relacionado con la creación de la solución. Será el encargado de desarrollar los algoritmos necesarios para el correcto funcionamiento de la solución, y realizar una integración entre los programas existentes relacionados con la solución a implementar.
- **Técnico:** Es el encargado de realizar todos los estudios correspondientes sobre el área en donde se implementará la solución. También participará de la implementación del proyecto instalando y realizando las pruebas del sistema relacionadas a la instalación de hardware y software.

Matriz de responsabilidades

Tipo de responsabilidad:

- **E (Encargado):** este rol realiza el trabajo y es responsable por su realización. Debe existir sólo un Responsable, si existe más de uno, entonces el trabajo debería ser subdividido a un nivel más bajo.
- **R (Responsable):** este rol se encarga de aprobar el trabajo finalizado y a partir de ese momento, se vuelve responsable por él.
- **C (Consultado):** este rol posee la información y/o capacidad necesaria para terminar el trabajo.
- **I (Informado):** este rol debe ser informado sobre el progreso y los resultados del trabajo.

Referencias:

- PM: Director de proyecto
- AF: Analista funcional
- DA: Desarrollador de aplicación
- T: Técnico

Planificación

	PM	AF	DA	T	Días
Establecer objetivos	R, E				1
Alcance	R, E				2



Tabla 5: Matriz de responsabilidades – Planificación. Elaboración propia

Recolección de requerimientos

	PM	AF	DA	T	Días
Relevamiento de sistema de información existente	E, I	E			1
Relevamiento de estructura informática existente	E, I	E			1
Relevamiento de sistema de evaluación actual	E, I	E			2
Control y validación de requerimientos	E, I	E			1
Estudio de metodologías para el desarrollo de BI	E, I	E			2

Tabla 6: Matriz de responsabilidades - Recolección de requerimientos. Elaboración propia

Diseño solución BI

	PM	AF	DA	T	Días
Instalación y configuración proyecto Laravel	C, I		E		6
Diseño modelo relacional base de datos	C, I	E			2
Diseño modelo lógico de base de datos	C, I	E			2
Diseño de generación de datos	C, I	E			3
Plan de migración de datos	C, I	E			1
Evaluación de alternativas tecnológicas	C, I	E			5
Alternativas de solución	C, I	E			2
Alternativas del motor de base de datos	C, I	E			1
Alternativas de Herramientas BI	C, I	E			2
Alternativas de Hardware	C, I	E		C	1
Alternativas de software base	C, I	E			1
Elección de alternativas	C, I	E			2
Evaluación de metodologías para el desarrollo de BI	C, I	E			4
Selección de Metodología BI	C, I	E			1
Análisis FODA	C, I	E			1
Análisis de requerimientos	C, I	E			3
Relevamiento preguntas de negocio	C, I	E			2
Elección de indicadores y perspectivas	C, I	E			1
Diseño modelo conceptual	C, I	E			1
Análisis de data sources	C, I	E			4
Hechos e indicadores	C, I	E			1
Mapeo	C, I	E			1
Granularidad	C, I	E			1
Diseño modelo conceptual ampliado	C, I	E			1
Diseño modelo lógico del DW	C, I	E			3
Elección Tipología de DW	C, I	E			1
Diseño tabla de dimensiones	C, I	E			1
Diseño tabla de hechos	C, I	E			1
Uniones	C, I	E			1
Diseño de Integración de datos	C, I	E			10
Diseño carga inicial	C, I	E			4
Diseño actualización	C, I	E			6



Diseño de CUBO multidimensional	C, I	E			2
Diseño de presentación CUBO multidimensional	C, I	E			2
Gestión de Riesgos	C, I	E			1

Tabla 7: Matriz de responsabilidades - Diseño solución BI. Elaboración propia

Desarrollo

	PM	AF	DA	T	Días
Construcción física base de datos	C, I		E		5
Creación base de datos	C, I		E		1
Carga de datos de prueba	C, I		E		2
Backup y Restore de base de datos	C, I		E		1
Migración de base de datos	C, I		E		1
Pruebas a base de datos	C, I	C	E		1
Construcción física de DW	C, I		E		1
Creación DW	C, I		E		1
ETL	C, I		E		3
Carga inicial	C, I		E		1
Pruebas a DW	C, I	C	E		2
Construcción CUBO	C, I		E		4
Creación CUBO	C, I		E		1
Carga Inicial CUBO	C, I		E		1
Pruebas CUBO	C, I	C	E		2
Presentación CUBO	C, I		E		2

Tabla 8: Matriz de responsabilidades – Desarrollo. Elaboración propia

Instalación y pruebas

	PM	AF	DA	T	Días
Instalación de herramientas software	I		C	E	2
Instalación de la solución BI	I		E		2
Capacitación	C	E			2
Validación de las instalaciones mediante pruebas	C	E			2
Evaluación de resultados	C	E			1

Tabla 9: Matriz de responsabilidades - Instalación y pruebas. Elaboración propia

Finalización

	PM	AF	DA	T	Días
Estudio de aceptación	R, I	E			2
Presentación a la dirección	R, E	C	C	C	2

Tabla 10: Matriz de responsabilidades – Finalización. Elaboración propia

Diagrama de Gantt

En el desarrollo del esquema el calendario del trabajo será el siguiente:

- 1 día de trabajo tendrá 8 horas.



- Los días laborables serán de lunes a viernes.
- Se tomarán en cuenta los días feriados como no laborables.

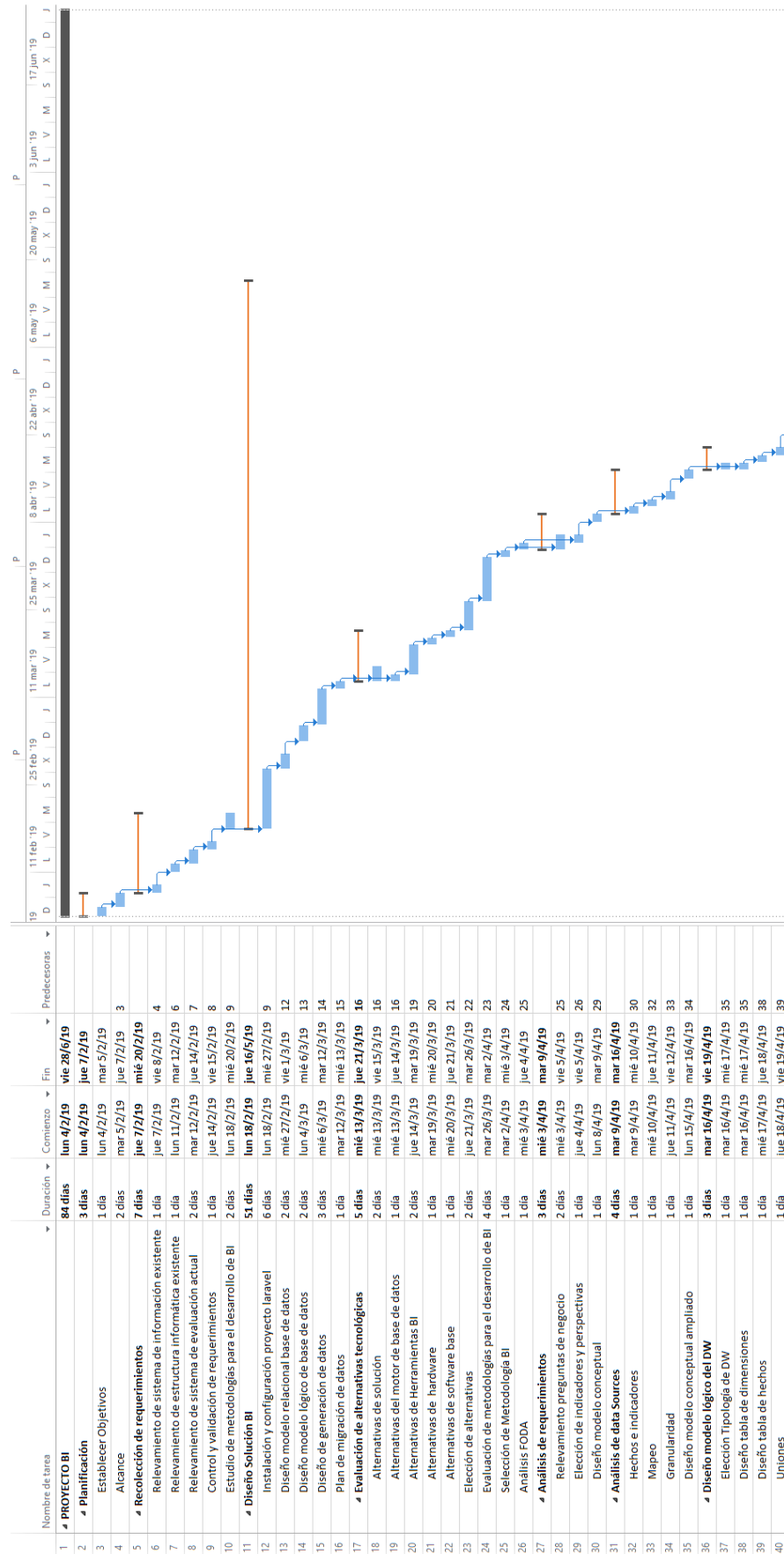


Ilustración 9: Diagrama de Gantt - Parte 1. Elaboración propia

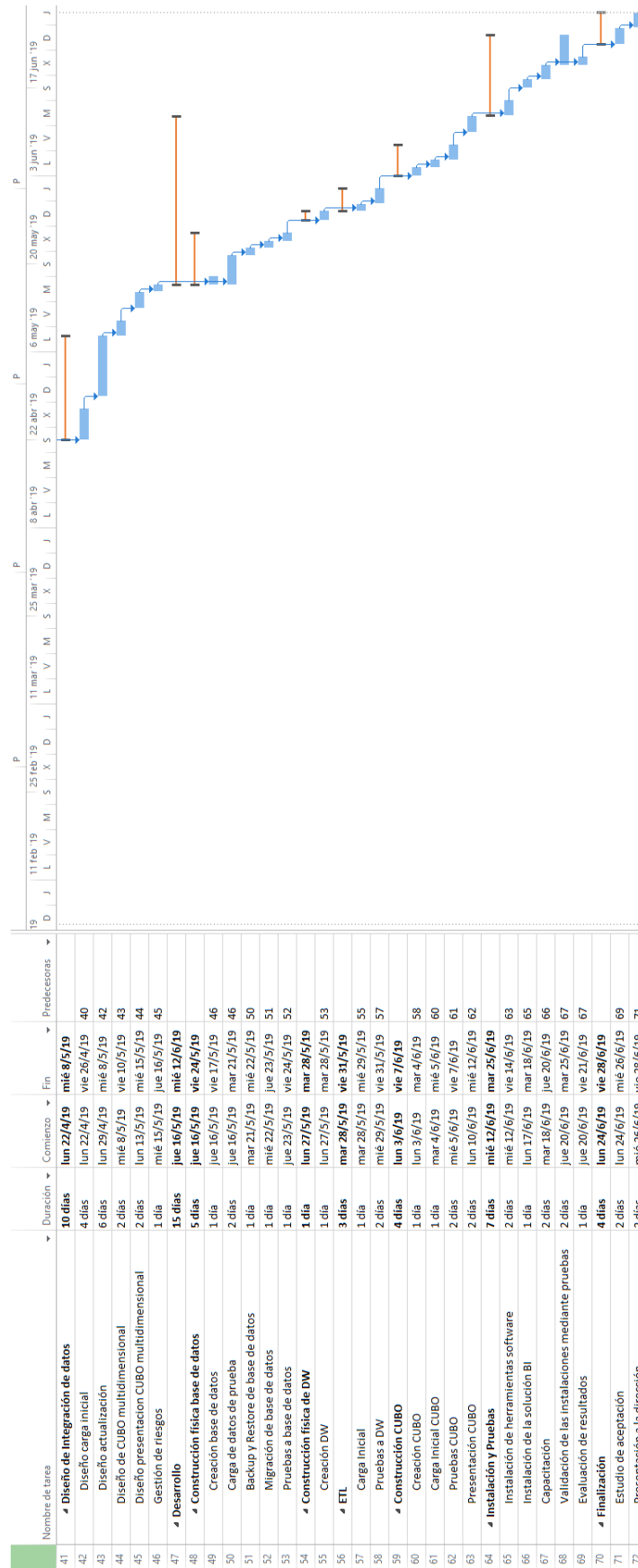


Ilustración 10: Diagrama de Gantt - Parte 2. Elaboración propia



Presupuesto

En el siguiente presupuesto, se tomarán como referencia el tipo de responsabilidad que asumieron en cada trabajo.

Cuando un participante tenga la responsabilidad del tipo Encargado o Responsable se tomará como trabajado el día completo, cuando la persona tenga la responsabilidad del tipo Consultado se tomará como trabajado cuatro horas por la cantidad de días.

Costos equipo de trabajo¹⁵:

Cargo	Tiempo en Hs.	Honorario por hora en \$	Cantidad	Subtotal \$
Director de proyecto	420	450	1	189.000
Analista Funcional	540	300	1	168.000
Desarrollador de aplicaciones	208	200	1	41.600
Técnico	28	150	1	4.200
			TOTAL	396.800

Tabla 11: Honorarios. Elaboración propia

Costos Hardware:

Hardware	Costo en \$	Cantidad	Subtotal \$
PC Asus Core I5 8GB RAM Disco 1TB	21.615	1	21.615
Cable UTP Cat 5e 3mts.	66	1	66
		TOTAL	21.681

Tabla 12: Costos Hardware. Elaboración propia

Costos Software:

Software	Costo en \$	Cantidad	Subtotal \$
SQL Server 2014 Standard	(U\$S 875) 31.937,5	1	31.937,5
Windows 10 Pro	900	1	900
		TOTAL	32.837,5

Tabla 13: Costos Software. Elaboración propia

Etapas	Valor parcial	Valor Total de la etapa
1. Planificación		10.800
Objetivos y Alcance	10.800	
2. Recolección de requerimientos		42.000
Relevamiento de sistema de información existente	6.000	
Relevamiento de estructura informática existente	6.000	

¹⁵ Honorarios extraídos de <http://www.copaipa.org.ar/informatica/>



Relevamiento de sistema de evaluación actual	12.000	
Control y Validación de requerimientos	6.000	
Estudio de metodologías BI	12.000	
3. Diseño Solución BI		231.000
Instalación y configuración proyecto Laravel	20.400	
Diseño modelo relacional base de datos	8.400	
Diseño modelo lógico de base de datos	8.400	
Generación de datos	12.600	
Plan de Migración	4.200	
Evaluación de alternativas tecnológicas	30.000	
Elección de alternativas	8.400	
Evaluación y selección de metodología BI	21.000	
Análisis FODA	4.200	
Análisis de requerimientos	16.800	
Análisis de data Sources	16.800	
Diseño modelo lógico del DW	16.800	
Diseño de integración de datos	42.000	
Diseño de CUBO	8.400	
Diseño de presentación de datos	8.400	
Gestión de Riesgos	4.200	
4. Desarrollo		60.400
Construcción física base de datos	21.600	
Construcción física de DW	3.400	
ETL	12.600	
Construcción CUBO	16.000	
Presentación CUBO	6.800	
5. Instalación y Pruebas		28.200
Instalación de herramientas	4.000	
Instalación de BI	3.200	
Capacitación	8.400	
Validación de las instalaciones	8.400	
Evaluación de resultados	8.400	
6. Finalización		24.400
TOTAL COSTOS		396.800
SQL Server 2014 Standard	(U\$S 875) 31.937,5	
PC Asus Core I5 8GB RAM Disco 1TB	21.615	
Windows 10 Pro	900	
Cable UTP Cat 5e 3mts.	66	
Gastos Generales		54.518,5
TOTAL		451.318,5

Tabla 14: Presupuesto Desglosado. Elaboración propia



Flujo de Egresos

Actividad		Mes				
		1	2	3	4	5
1 - Planificación	Tiempo					
	Costo	10800				
2 - Recolección de Requerimientos	Tiempo					
	Costo	42000				
3 - Diseño Solución BI	Tiempo					
	Costo	37200	89300	67200	37300	
4 - Desarrollo	Tiempo					
	Costo				53600	6800
5 - Instalación y Pruebas	Tiempo					
	Costo					28200
6 - Finalización	Tiempo					
	Costo					24400
Gastos Generales	Tiempo					
	Costo	10.903,7	10.903,7	10.903,7	10.903,7	10.903,7
TOTAL DESEMBOLSO		100.903,7	100.203,7	78.103,7	101.803,7	70.303,7

Tabla 15: Flujo de Egresos. Elaboración propia

VAN y TIR

A continuación, con el fin de analizar la factibilidad económica de la inversión de este proyecto, vamos a hacer uso de dos herramientas numéricas esenciales:

- Valor Actual Neto (VAN)
- Tasa Interna de Retorno (TIR)

Con una tasa de descuento mensual del 1% y un horizonte de inversión de dos años, procedemos a obtener los siguientes resultados.

Mes	Ingreso	Egreso	Total	Saldo		Flujo Descontado	Saldo Descontado
0	0	0	0	0		0	0
1	0	-100903.7	-100903.7	-100903.7		-100401.692	-100401.692
2	0	-100203.7	-100203.7	-201107.4		-99209.1285	-199610.82
3	0	-78103.7	-78103.7	-279211.1		-76943.7632	-276554.583
4	0	-101803.7	-101803.7	-381014.8		-99792.8246	-376347.408
5	0	-70303.7	-70303.7	-451318.5		-68572.1669	-444919.575
6	26000		26000	-425318.5		25233.47	-419686.105
7	26000		26000	-399318.5		25107.9304	-394578.174
8	26000		26000	-373318.5		24983.0153	-369595.159
9	26000		26000	-347318.5		24858.7217	-344736.437
10	26000		26000	-321318.5		24735.0465	-320001.391
11	26000		26000	-295318.5		24611.9865	-295389.404
12	26000		26000	-269318.5		24489.5388	-270899.865



13	26000		26000	-243318.5		24367.7003	-246532.165
14	26000		26000	-217318.5		24246.468	-222285.697
15	26000		26000	-191318.5		24125.8388	-198159.858
16	26000		26000	-165318.5		24005.8097	-174154.049
17	26000		26000	-139318.5		23886.3779	-150267.671
18	26000		26000	-113318.5		23767.5402	-126500.131
19	26000		26000	-87318.5		23649.2937	-102850.837
20	26000		26000	-61318.5		23531.6355	-79319.2014
21	26000		26000	-35318.5		23414.5627	-55904.6387
22	26000		26000	-9318.5		23298.0723	-32606.5663
23	26000		26000	16681.5		23182.1615	-9424.40482
24	26000		26000	42681.5		23066.8274	13642.4226
VAN		TIR					
13642.42		1%					

Tabla 16: VAN y TIR. Elaboración propia

A partir del último mes del proyecto se puede ahorrar \$26.000 mensual del trabajo de una persona que analiza la información y que ahora tendrá toda la información disponible. Se realiza esto para tener un ahorro que permita justificar este proyecto y que en menos de 2 años se pueda recuperar la inversión justificado por un VAN positivo y TIR positiva.

Al obtener un VAN mayor a 0 (cero), en este caso el VALOR supera al COSTO por lo que se puede decir que el proyecto agrega valor económico al programa de becas.

En cuanto al obtener una TIR del 1%, estamos en una situación, en la que la inversión podrá llevarse a cabo si efectivamente se produce un cambio positivo en la organización.

Base de Datos

Modelo Relacional

Al no contar con el DER (Diagrama Entidad Relación) del sistema que maneja la organización y sobre el cual se aplicara Business Intelligence, es necesario realizar el diseño de la Estructura de la base de datos en el cual se almacenan los datos de cientos de personas que se registran todos los años.

En base a lo relevado se buscó armar los datos lo más representativo de la realidad, pero igualmente por ser datos no reales, los mismos no serán suficientes para un buen análisis con el prototipo.

Para proceder a la construcción de la base de datos y obtener el mejor resultado, se tomará en cuenta las siguientes condiciones:

- El modo en cómo interactúa la base de datos con la aplicación que hace uso de la misma.
- La correcta relación entre los diferentes tipos de entidades.
- Se respetarán los años de vigencia de la Organización.



- Determinados tipos de datos y/o tablas tendrán sus valores por defecto, por ejemplo: Estado Civil, Sexo, etc.
- El enfoque de la base de datos estará dado sobre el objetivo de la organización, por lo tanto, no se tomarán en cuenta los usuarios que utilizan y realizan la carga de datos de las personas que presentan una solicitud.
- Los términos que se establecen en el REGLAMENTO GENERAL DE BECA DE ESTUDIOS.

Para consultar el REGLAMENTO GENERAL DE BECA DE ESTUDIOS ver en **Anexo I.**

Para ver capturas del sistema de becas ver en **Anexo III.**

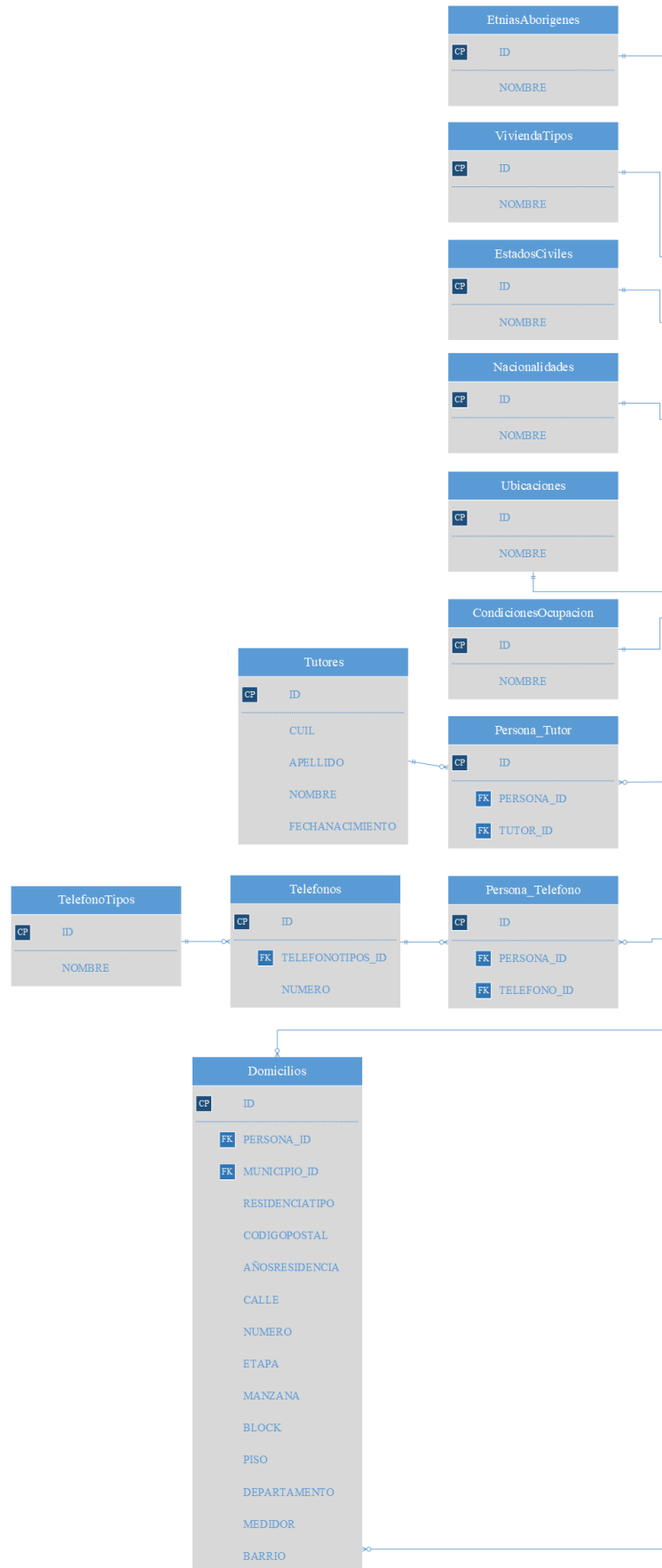


Ilustración 11: DER Parte izquierda. Elaboración propia

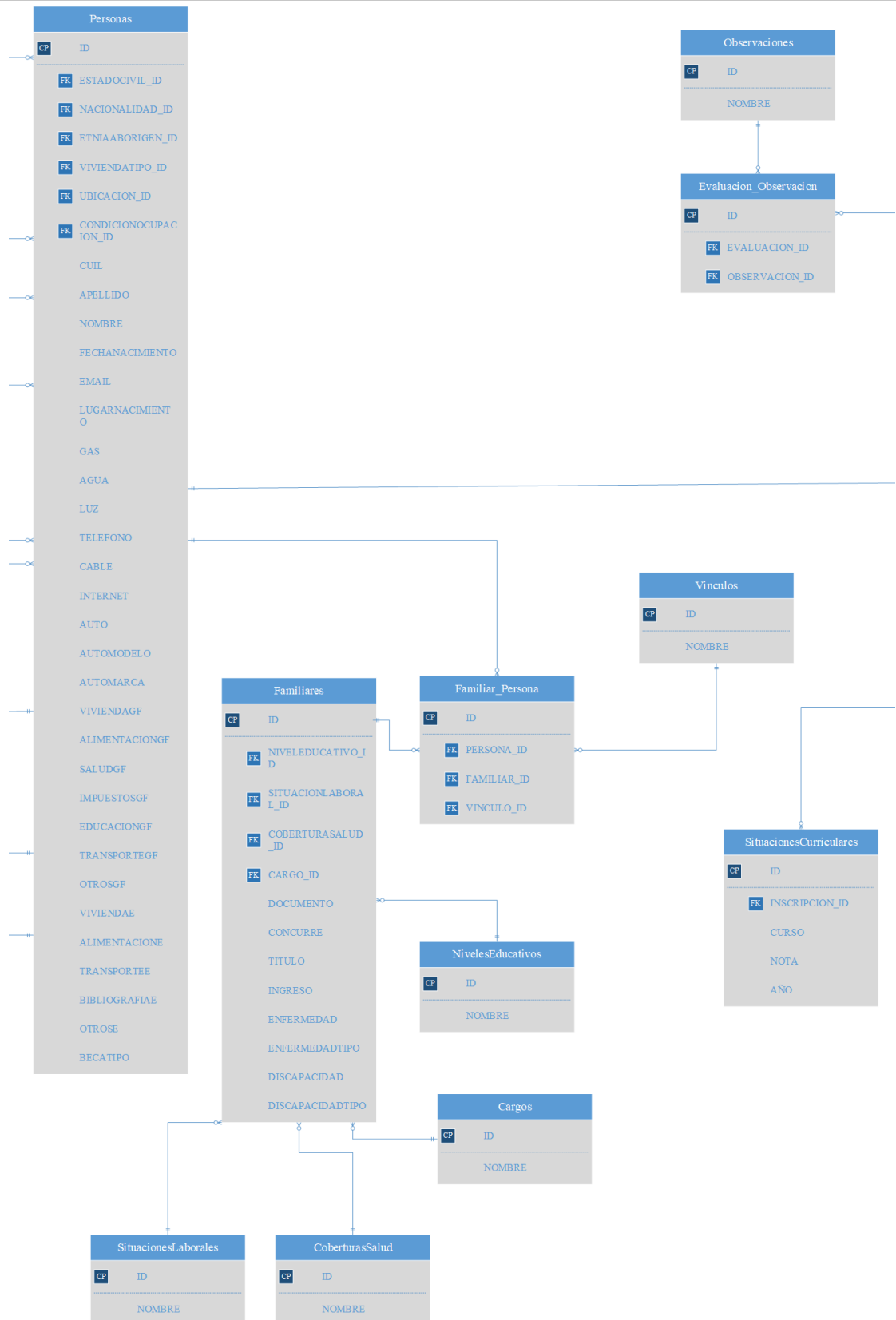


Ilustración 12: DER Parte central. Elaboración propia

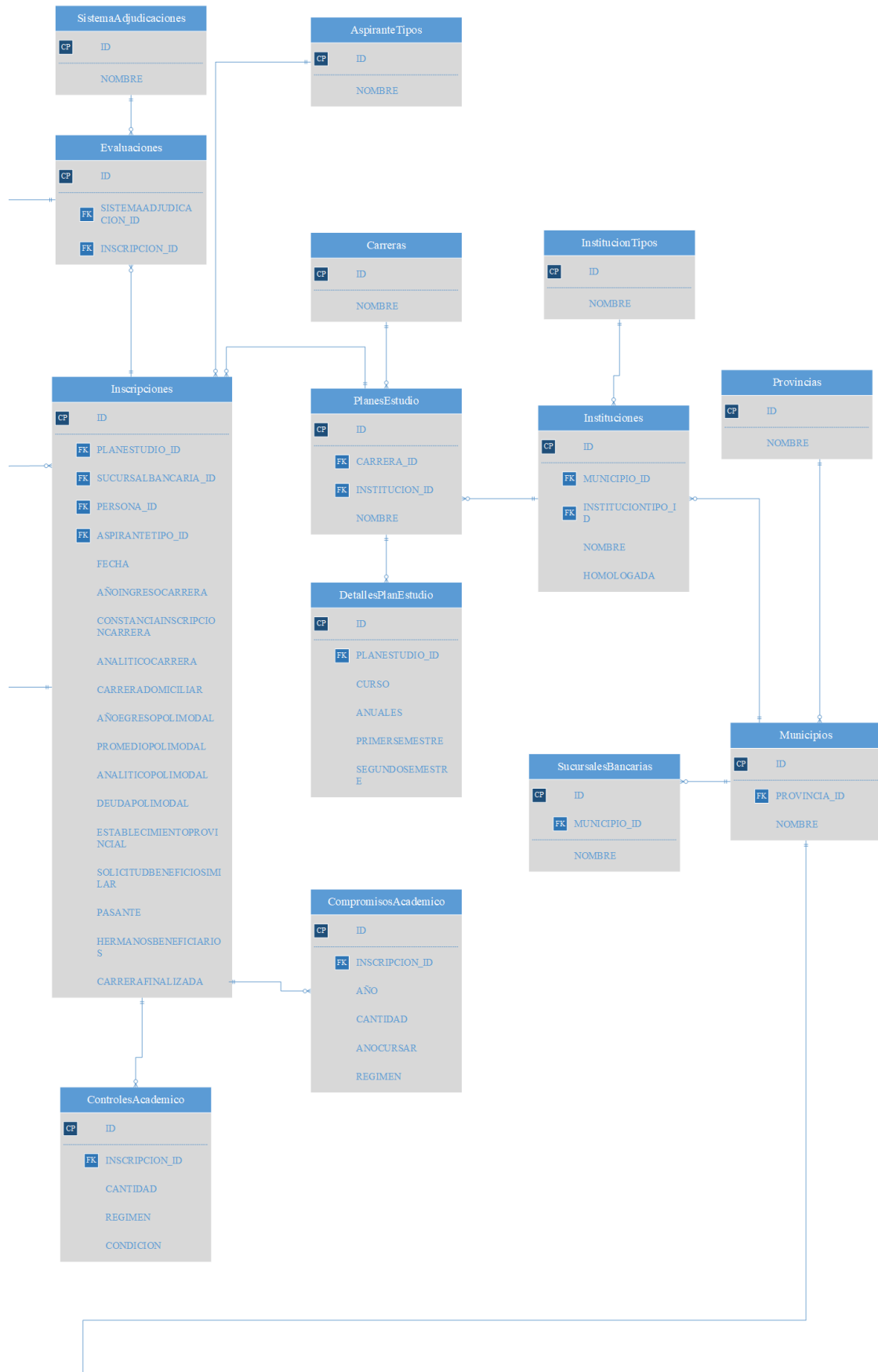


Ilustración 13: DER Parte derecha. Elaboración propia



Modelo lógico y Generación de Datos

Para traducir al modelo relacional en un modelo lógico se lo realizó a través de Laravel, un framework de código abierto utilizado para desarrollar aplicaciones y servicios web con PHP 5 y PHP 7, que trabajará en conjunto con MySQL como SGBD (Sistema Gestor de Base de Datos) realizado en Linux Mint 19 “Tara”.

Se utilizó Laravel con el propósito de dar solución a la problemática planteada en el capítulo anterior, el no poseer los datos reales de la base de datos para luego ser trabajada en los procesos de Business Intelligence. Se utilizarán datos de prueba que servirán como referencia para aquellos resultados que se obtengan de la Solución Informática que se fijaron como objetivos.

Con una base de datos vacía que fue diseñada previamente, se procederá a llenarla con instancias de datos de prueba en determinadas tablas que se ven afectadas por no poseer los datos reales que posee la organización. Para lograr esto se utilizó ciclos y evaluaciones condicionales asociados a una librería de Laravel llamada **Faker**, que permite generar distintos tipos de datos de prueba de manera aleatoria.

Para llegar al punto de generar estos datos, primero debemos definir ciertos conceptos:

- Consola Artisan
- Migrations (Migraciones)
- Seeders (Semilleros)

Consola Artisan

Artisan es la interfaz de línea de comandos incluida con Laravel. Proporciona una serie de útiles comandos que pueden ayudar mientras se construye una base de datos o una aplicación.

Lista de comandos que usaremos y para qué sirven:

	Comando	Utilidad
1	php artisan make:migration	Crear un Migration
2	php artisan make:seeder	Crear un Seeder respecto a los tipos de datos de un Migration
3	php artisan migrate	Ejecutar los migrations que se hayan creado
4	php artisan migrate:refresh	Refrescar la estructura de los migrations creados
5	php artisan db:seed	Inicializar los datos de la base de datos
6	php artisan migrate:refresh --seed	Refrescar la estructura de los migrations y los seeder con los datos de prueba

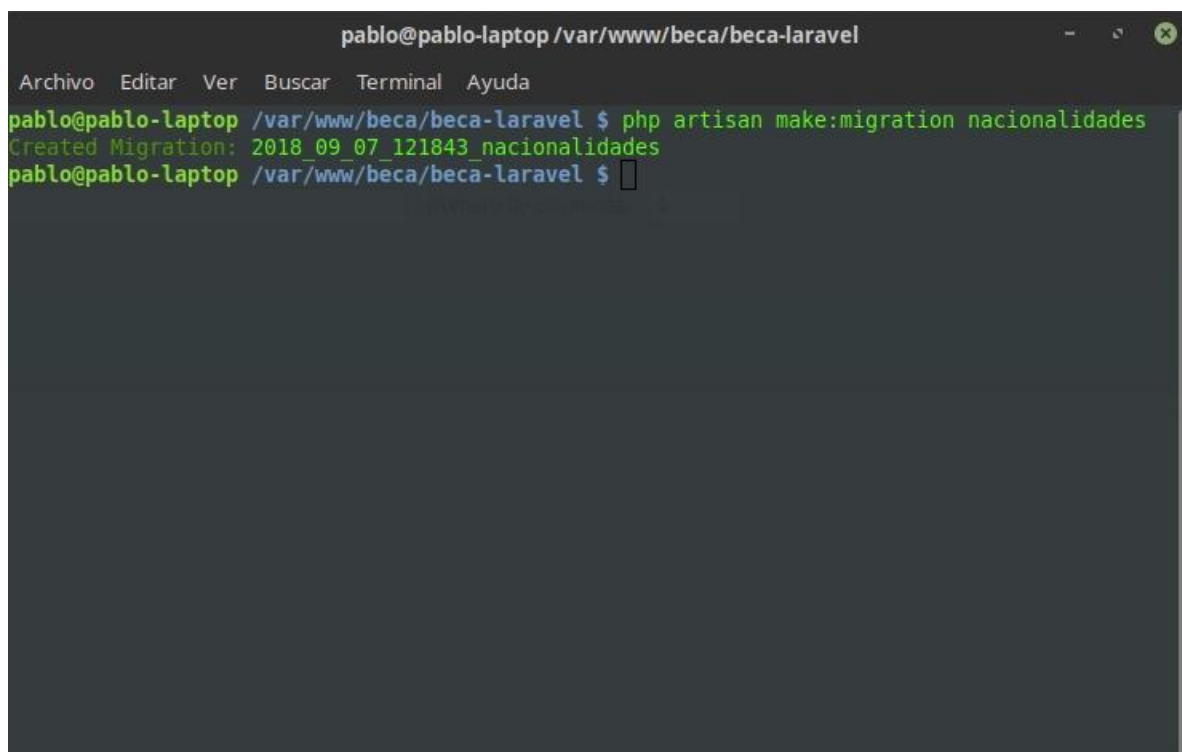
Tabla 17: Comandos Artisan. Elaboración propia



Tablas y tipo de datos - Migraciones

Para generar las tablas se utilizan Migrations (Migraciones) en Laravel, que son como el control de versiones para una base de datos, lo que permite modificar el esquema de la base de datos de la aplicación. Las migraciones suelen combinarse con el generador de esquemas de Laravel para crear fácilmente el esquema de la base de datos. Permite agregar una columna a una tabla de una base de datos sin la necesidad de hacerlo manualmente ejecutando un simple comando Artisan que refresca la base de datos con las nuevas modificaciones.

Para crear una migración se usa el comando Artisan **make:migration**. Por ejemplo, para la tabla de nacionalidades se ejecuta el siguiente comando:



```
pablo@pablo-laptop /var/www/beca/beca-laravel
Archivo  Editar  Ver  Buscar  Terminal  Ayuda
pablo@pablo-laptop /var/www/beca/beca-laravel $ php artisan make:migration nacionalidades
Created Migration: 2018_09_07_121843_nacionalidades
pablo@pablo-laptop /var/www/beca/beca-laravel $
```

Ilustración 14: Ejemplo make:Migration. Elaboración propia

Un Migration contiene dos métodos: un método **up** y otro **down**. El método **up** es usado para agregar una nueva tabla en la base de datos, mientras que el método **down** invierte las operaciones realizadas por el método **up**.

Dentro de estos dos métodos, utilizando librerías propias de Laravel, se construyen o modifican tablas de manera expresiva. Por ejemplo, este Migration crea una tabla **nacionalidades**:



```
<?php

use Illuminate\Support\Facades\Schema;
use Illuminate\Database\Schema\Blueprint;
use Illuminate\Database\Migrations\Migration;

class CreateNacionalidadTable extends Migration
{
    public function up()
    {
        Schema::create('nacionalidades', function (Blueprint $table) {
            $table->increments('id');
            $table->string('nombre');
            $table->timestamps();
            $table->softDeletes();
        });
    }

    public function down()
    {
        Schema::drop('nacionalidades');
    }
}
```

Ilustración 15: Ejemplo Migration Nacionalidades. Elaboración propia

Para crear una nueva tabla se usa el método **create**. El método acepta dos argumentos. El primero es el nombre de la tabla y el segundo es la estructura de la tabla, en el cual se definen los atributos y el tipo de dato del atributo.

Para borrar la tabla a través de este Migration se usa el método **drop**, pasando como único argumento el nombre de la tabla a eliminar.

Datos de prueba – Seeder

Laravel incluye un método simple para llenar con datos de prueba una base de datos utilizando Seeders (semilleros). Los Seeders se generan ejecutando el comando de Artisan **make:seeder**.



```
pablo@pablo-laptop /var/www/beca/beca-laravel
Archivo Editar Ver Buscar Terminal Ayuda
pablo@pablo-laptop /var/www/beca/beca-laravel $ php artisan make:seeder nacionalidades
Seeder created successfully.
pablo@pablo-laptop /var/www/beca/beca-laravel $
```

Ilustración 16: Ejemplo make:seeder. Elaboración propia

Un Seeder contiene un método por defecto: **run**. Este método es llamado cuando se ejecuta el comando Artisan **db:seed**. Dentro de este método se insertan los datos que deseamos a las tablas de la base de datos. Por ejemplo, en el Seeder nacionalidades, se define un registro de datos que luego son insertados en la tabla nacionalidades al usar el método **insert** ingresando como parámetro el nombre del registro definido anteriormente.



```
<?php

use App\Modules\Nacionalidad\Models\Nacionalidad;
use Illuminate\Database\Seeder;

class NacionalidadesTableSeeder extends Seeder
{
    public function run()
    {
        $nacionalidades = [
            [
                'nombre' => 'Argentino/a',
            ],
            [
                'nombre' => 'Extranjero/a'
            ],
        ];

        Nacionalidad::insert($nacionalidades);
    }
}
```

Ilustración 17: Ejemplo Seeder Nacionalidades. Elaboración propia

A partir del modelo relacional se pudieron crear los siguientes Migrations y Seeders:

Tablas con datos ¹⁶	
AspiranteTipos	Nacionalidades
Cargos	Observaciones
Carreras	PersonaTelefono
CoberturaSalud	PersonaTutor
CompromisoAcademico	Personas
ControlAcademico	PlanesEstudio
DetallesPlanesEstudio	Provincias
Domicilios	SistemaAdjudicaciones
EstadosCiviles	SituacionesCurriculares
EtniasAborigenes	SituacionesLaborales
EvaluacionObservacion	SucursalesBancarias
Evaluaciones	Telefonos
FamiliarPersona	TelefonoTipos
Familiares	Tutores
Inscripciones	Ubicaciones
Instituciones	Users
InstitucionTipos	Vinculos
Municipios	ViviendaTipos

Tabla 18: Migrations y Seeders. Elaboración propia

¹⁶ Adjunto Proyecto de Laravel que contiene los Migrations y Seeders para visualizar su código dirigirse a **beca-laravel > database**



Una vez definidos estos conceptos podemos pasar a ver cómo actúa el paquete **Faker** en los Seeders.

Faker es una librería de PHP que genera datos de prueba, basta con usar la sentencia **Faker\Factory::create()** para crear e inicializar un generador de datos de prueba, el cual puede generar datos indicando el tipo de dato que se necesite.

Por ejemplo, el seeder Evaluaciones consta de tres atributos con tipo de dato de número entero, en los cuales se asigna que los números enteros se generen entre ciertos números usando el formateador `numberBetween()` y en el caso de `inscripcion_id` se agrega un modificador `unique()` para que los valores ingresados no se repitan.

```
<?php
use Illuminate\Database\Seeder;

class EvaluacionesTableSeeder extends Seeder
{
    public function run()
    {
        $dummy = Faker\Factory::create();

        for ($i=0; $i < 200; $i++) {
            App\Modules\Evaluacion\Models\Evaluacion::create([
                'sistemaAdjudicacion_id' => $dummy->numberBetween(1, 4),
                'inscripcion_id' => $dummy->unique()->numberBetween(1, 200),
                'persona_id' => $dummy->numberBetween(1, 200),
            ]);
        }
    }
}
```

Ilustración 18: Ejemplo Seeder Faker. Elaboración propia

Esta es una lista de los formateadores que se usaron para la construcción de la base de datos:

Formateador	Tipo de dato que genera
<code>numberBetween()</code>	Números enteros entre ciertos valores
<code>streetName</code>	Nombres de calles
<code>buildingNumber</code>	Número de un domicilio
<code>boolean</code>	Falso o Verdadero
<code>randomFloat</code>	Numero decimales entre ciertos valores
<code>firstName</code>	Nombres de personas
<code>lastName</code>	Apellidos de personas
<code>dateTimeBetween()</code>	Fechas entre ciertos valores
<code>freeEmail</code>	Emails
<code>regexify()</code>	Conjunto de caracteres
<code>tollFreePhoneNumber</code>	Números de teléfonos
<code>isbn10</code>	Números con una longitud de diez dígitos
<code>date</code>	Fechas

Además, se hace uso del modificador `unique()` para evitar datos repetidos de un atributo.



Migración

Al haber creado esta base de datos en MySQL dentro Linux se deberá migrarla a SQL Server en Windows para luego ser trabajada por las herramientas BI seleccionadas. Se decidió realizar este proceso debido a que la base de datos del sistema que maneja la organización está en SQL Server.

Para realizar con éxito esta migración fueron necesario de la instalación en Windows de ciertas herramientas:

- MySQL Workbench 6.3 CE: Proporciona modelado de datos, desarrollo de SQL y herramientas integrales de administración para la configuración del servidor, administración de usuarios, respaldo y mucho más.
- Microsoft SSMS (SQL Server Management Studio): Es un entorno gratuito e integrado para obtener acceso, configurar, administrar y desarrollar todos los componentes de SQL Server. Permite administrar instancias del motor de base de datos de SQL Server.
- Microsoft SSMA (SQL Server Migration Assistant) para MySQL: Herramienta libre de Microsoft que simplifica el proceso de migración de bases de datos desde MySQL a SQL Server. SSMA automatiza todos los aspectos de la migración, incluido el análisis de evaluación de la migración, el esquema y la conversión de las declaraciones SQL, la migración de datos y las pruebas de migración.

Backup y Restore de base de datos¹⁷

En principio fue necesario realizar un backup completo de la base de datos alojada en MySQL en Linux para luego restaurarla en el servidor de MySQL en Windows. Para ello debemos situarnos en cualquier carpeta dentro del directorio de Linux por terminal y luego ejecutar el siguiente comando:

```
mysqldump -h [hostname] -u [username] -p[password] [dbname] > [backupfile.sql]
```

- **mysqldump**: comando que acompaña al servidor MySQL.
- **[hostname]**: indica el nombre del host en el que se encuentra la base de datos.
- **[username]**: indica el nombre de usuario que administra la base de datos.
- **[password]**: indica la contraseña de la base de datos.
- **[dbname]**: el nombre de la base de datos.
- **[backupfile.sql]**: nombre del archivo que almacena el backup de la base de datos.

¹⁷ Adjunto backup de MySQL en carpeta de **Backup Base de datos MySQL**



```
pablo@pablo-laptop /var/www/beca/beca-laravel
Archivo Editar Ver Buscar Terminal Ayuda
pablo@pablo-laptop /var/www/beca/beca-laravel $ mysqldump -h localhost -u root -p becas > backupfile.sql
```

Ilustración 19: Backup de base de datos. Elaboración propia

Una vez completado el backup de la base de datos se procede a restaurar la base de datos en el servidor de MySQL en Windows.

Antes de proceder a restaurar la base de datos, debemos tener en cuenta que es necesarios establecer la conexión con el servidor de MySQL que hayamos creado. En este proyecto se crea una conexión en MySQL simplemente instalando e iniciando XAMPP, siendo esta una distribución de Apache completamente gratuita que contiene MariaDB, PHP y Perl. Se puede descargar ingresando a <https://www.apachefriends.org/>.

Ingresamos a MySQL Workbench y seleccionamos al host al que nos queremos conectar y en el cual vamos a proceder con la restauración.

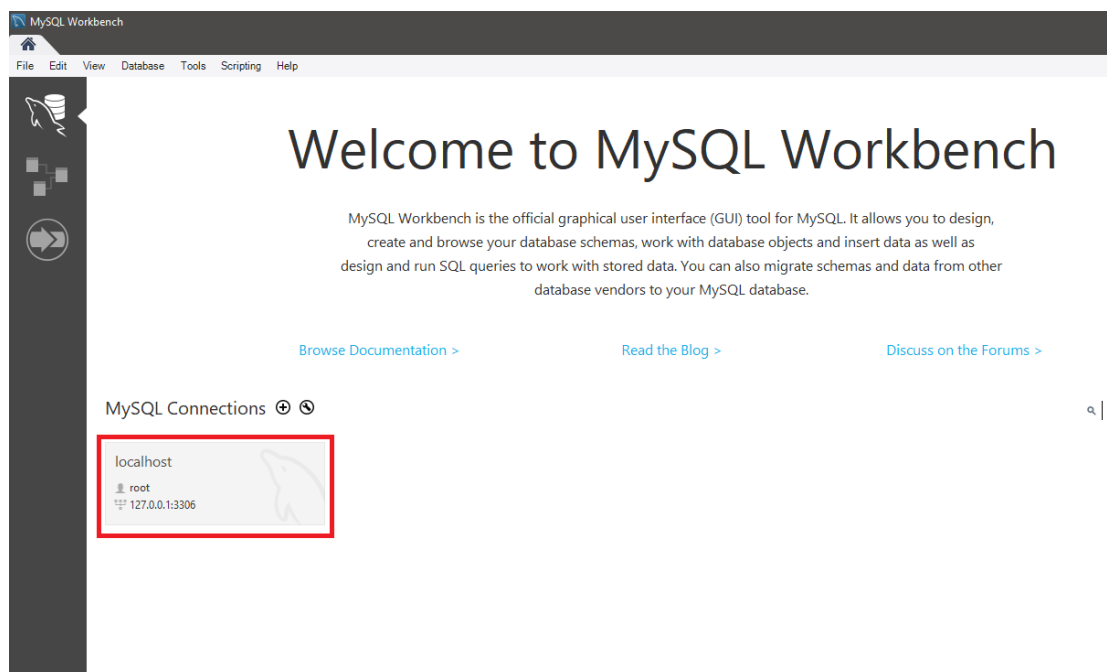


Ilustración 20: Estableciendo conexión MySQL. Elaboración propia

Para restaurar la base de datos debemos seleccionar en la opción **Navigator>Management>Data Import/Restore**:

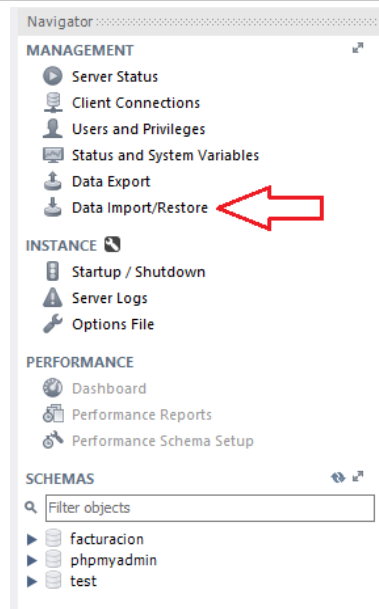


Ilustración 21: Data import/Restore. Elaboración propia

Como el archivo que hemos generado de backup es un archivo SQL, seleccionaremos la opción de “**Import from Self-Contained File**”. Como la base de datos no existe, debemos crearla con el botón de “**New...**”. Le daremos el nombre de “**becas**” y pulsaremos “**Start Import**”:

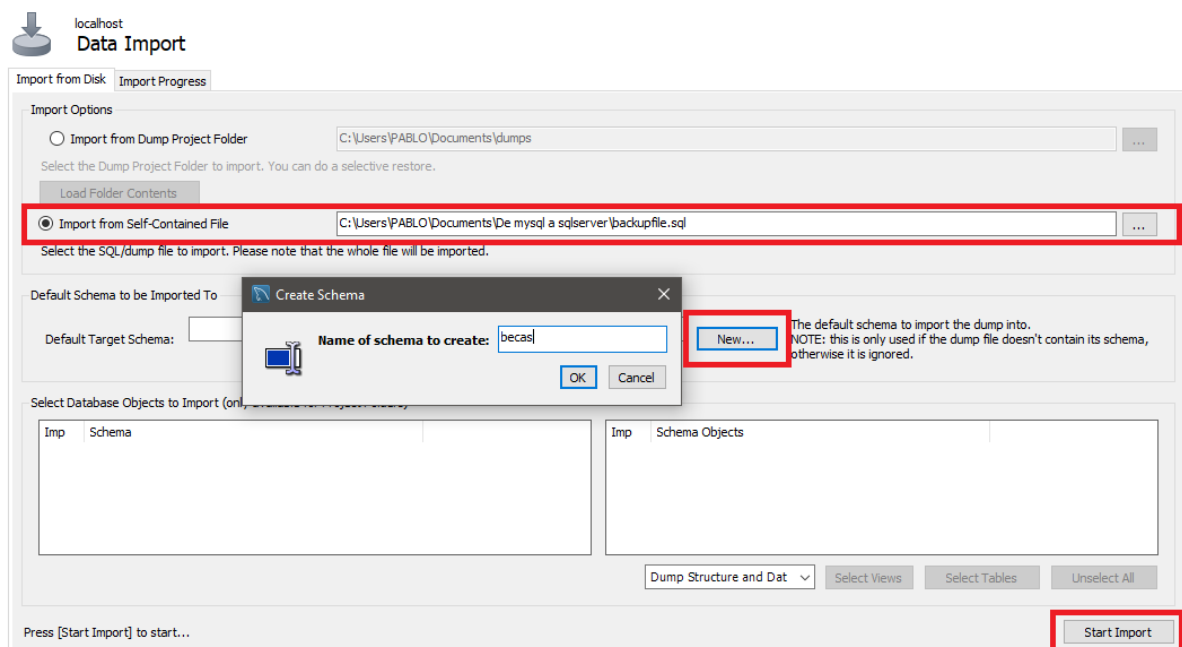


Ilustración 22: Datos de Restauración. Elaboración propia

En la parte izquierda de MySQL Workbench, podemos ver todas las bases de datos. Con el clic derecho daremos en **Refresh** y veremos que la base de datos se ha restaurado satisfactoriamente.

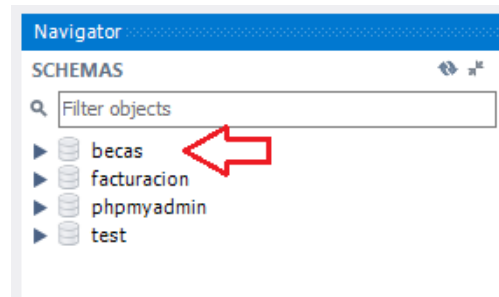


Ilustración 23: Base de datos restaurada. Elaboración propia

Migrar datos a SQL Server

Antes que nada, debemos tener creado también un servidor en Microsoft SQL Server y debemos conocer las credenciales de acceso de usuario administrador. También debemos crear una base de datos vacía, con la herramienta de Microsoft SQL Server Management Studio.

Ingresamos a la herramienta de migración Microsoft SQL Server Migration y debemos crear un nuevo proyecto indicando a la versión de SQL Server a la que vamos a migrar la base de datos de MySQL.

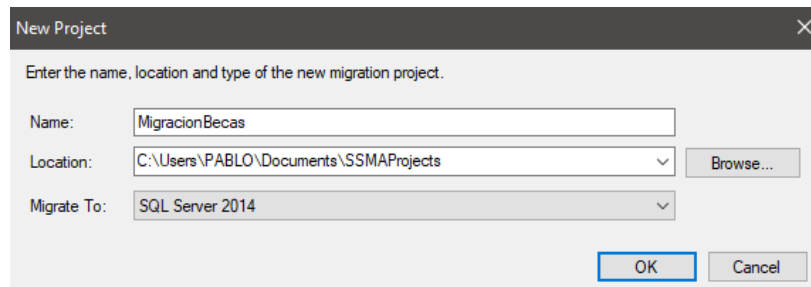


Ilustración 24: Proyecto Migración. Elaboración propia

Una vez creado el proyecto, como primera instancia debemos conectarnos al servidor MySQL. Cuando trabajamos sobre la misma maquina donde reside el del servidor MySQL, podemos utilizar la dirección interna 127.0.0.1 o localhost, para referirnos a nuestra máquina, de lo contrario tendría que ingresar la dirección IP externa del servidor. El puerto predeterminado para MySQL es 3306.

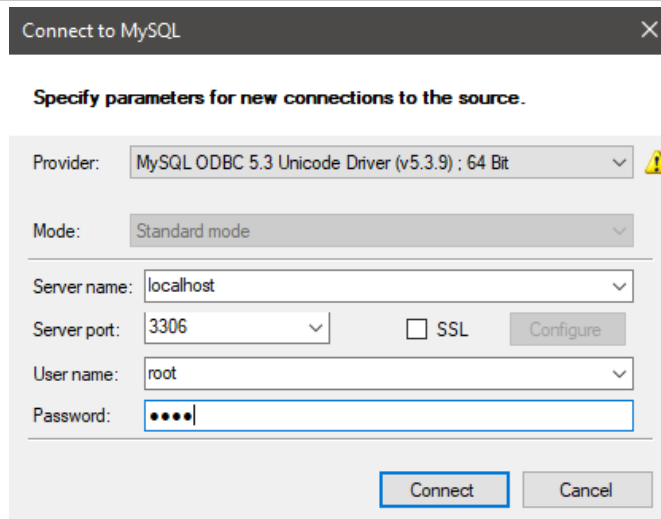


Ilustración 25: Migración - Conexión MySQL. Elaboración propia

Una vez conectados a MySQL, nos conectamos a SQL Server seleccionando el nombre del servidor y eligiendo la base de datos en los cuales serán migrados los datos.

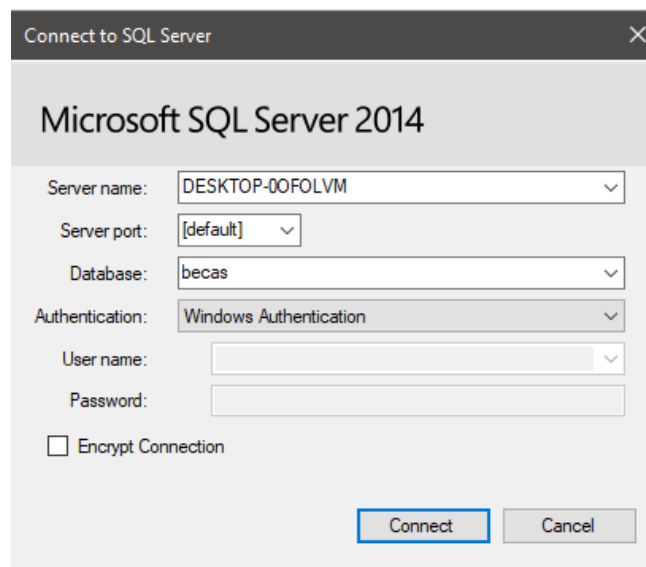


Ilustración 26: Migración - Conexión SQL Server. Elaboración propia

Ahora procedemos a crear el **Schema**, que será un script con la estructura de la base de datos que se va a exportar, esto lo haremos seleccionado tanto en el explorador de MySQL como de SQL Server la base de datos **becas** y le damos clic en el botón **Convert Schema**.

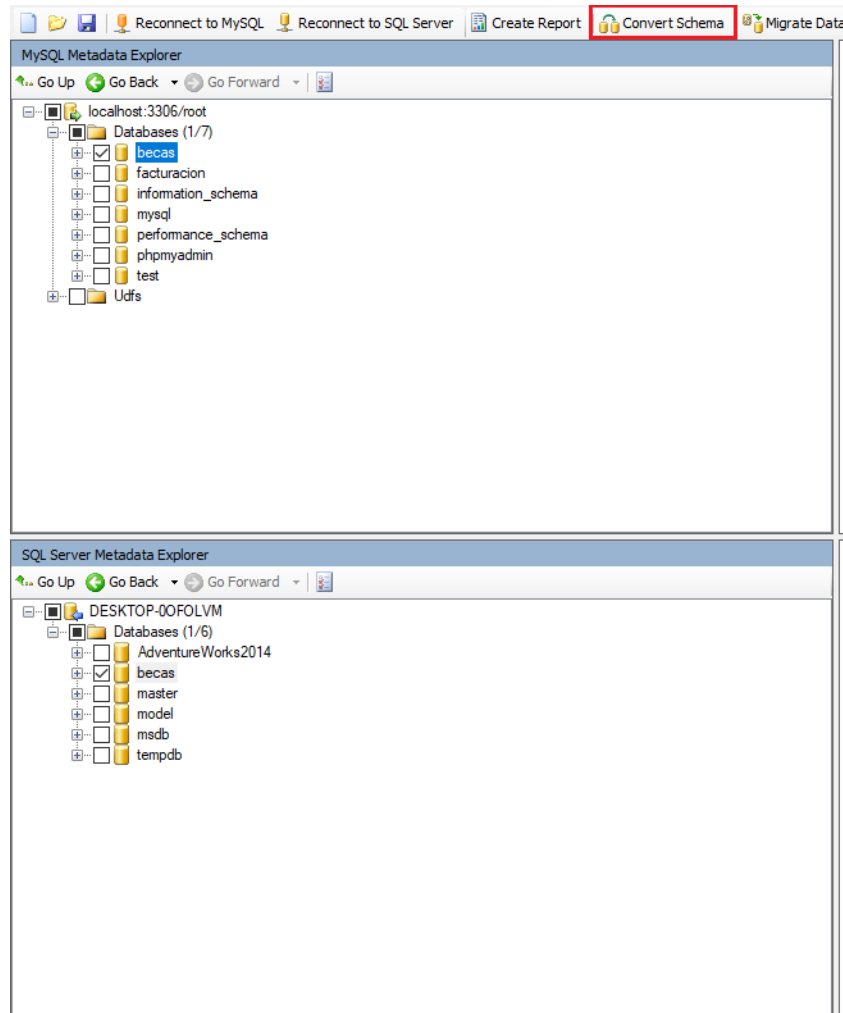


Ilustración 27: Migración - Convertir Schema. Elaboración propia

Ya creado el schema, vamos a sincronizar las bases de datos, dando clic derecho sobre la base de datos a la cual queremos sincronizar en el explorador de SQL Server y seleccionado la opción **Synchronize with Database**.

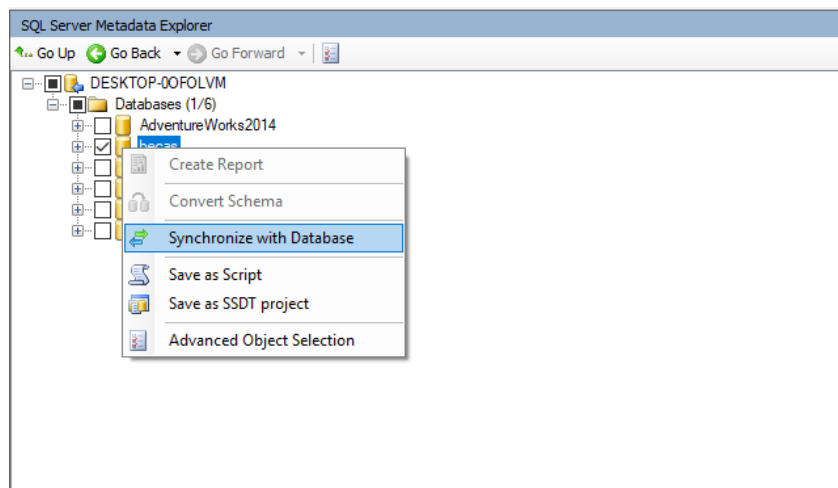


Ilustración 28: Migración - Sincronizar base de datos. Elaboración propia



Nos mostrará una ventana en la que apareciera el schema que se va a utilizar y sobre cual base de datos se va sincronizar. A esto le damos **OK** y se procedera a la sincronización.

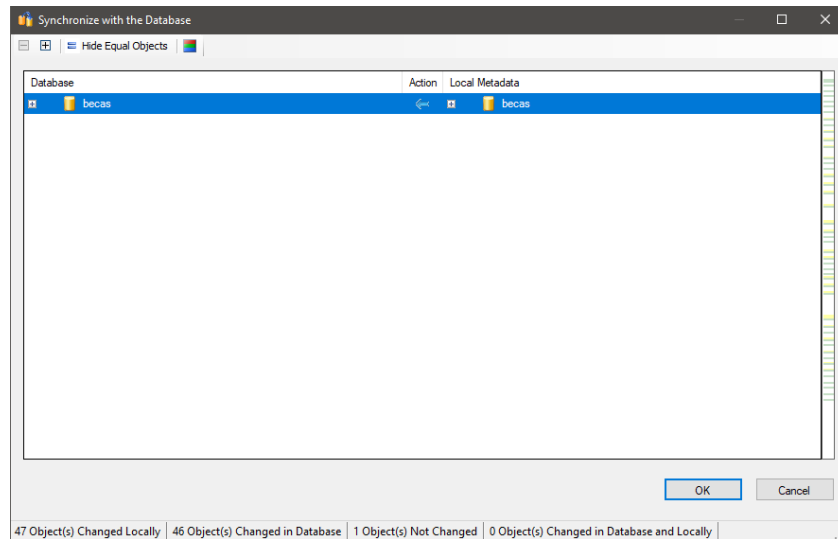


Ilustración 29: Migración - Sincronización. Elaboración propia

Para finalizar, procederemos a exportar los datos de la base de datos MySQL a SQL Server dando clic en **Migrate Data**. Sera necesario volver a identificarse en ambos servidores.

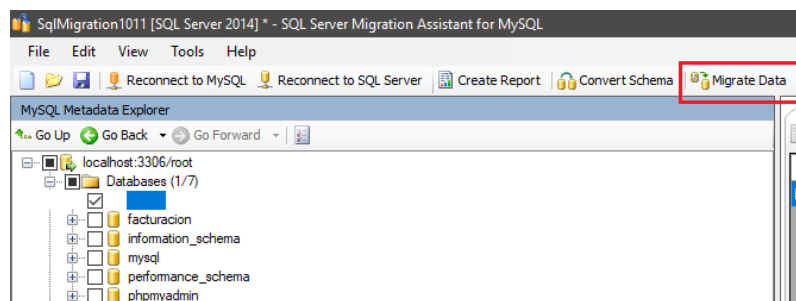


Ilustración 30: Migración - Migrando Datos. Elaboración propia

Al terminar la migración nos mostrará un reporte de los datos migrado con el nombre de la tabla y los registros.

Status	From	To	Total Rows	Migrated Rows	Success Rate
	'becas'. 'aspirantestpos'	[becas] [becas] [aspirantestpos]	2	2	100.00 %
	'becas'. 'becatpos'	[becas] [becas] [becatpos]	2	2	100.00 %
	'becas'. 'cargos'	[becas] [becas] [cargos]	8	8	100.00 %
	'becas'. 'cameras'	[becas] [becas] [cameras]	138	138	100.00 %
	'becas'. 'coberturassalud'	[becas] [becas] [coberturassalud]	4	4	100.00 %
	'becas'. 'compromisoacademico'	[becas] [becas] [compromisoacademico]	1200	1200	100.00 %
	'becas'. 'condicionesocupacion'	[becas] [becas] [condicionesocupacion]	8	8	100.00 %
	'becas'. 'controlesacademico'	[becas] [becas] [controlesacademico]	1200	1200	100.00 %

Ilustración 31: Migración - Reporte. Elaboración propia

Debemos asegurarnos y probar que realmente se migraron los datos, para eso, vamos a SQL Server Management Studio y lo comprobamos.¹⁸

¹⁸ Adjunto backup de base de datos migrada en carpeta **Backup Base de datos SQL Server**

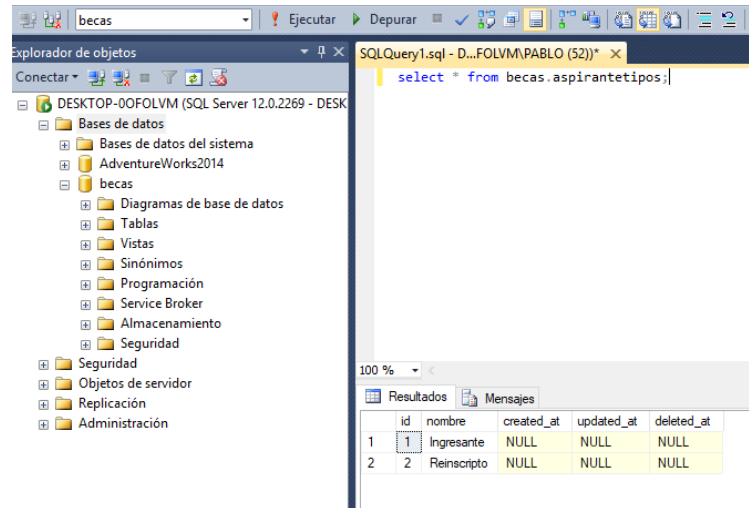


Ilustración 32: Migración - SQL Server. Elaboración propia

Metodologías para el desarrollo de BI

Existen varias tecnologías, utilizadas para el desarrollo e implementación de BI, entre las que encontramos:

Metodología de Kimball

En la presente metodología los Data Warehouse se forman de trabajar el grupo de todos los Data Marts generados existentes en una organización. La información siempre se guarda en un modelo dimensional. “Bottom-up” es denominado a este tipo de metodologías.

En esta metodología un Data Warehouse es considerado como un almacén en el cual se realizan las consultas y análisis de los datos transaccionales que fueron copiados en el mismo. La metodología de Kimball, también conocida como Modelo Dimensional, está basado en el Ciclo de Vida Dimensional del Negocio.

Para almacenar los datos transaccionales, las estructuras de los Data Warehouses precisan una presentación dimensional. En el modelo Dimensional se construyen modelos de tablas y relaciones con el objetivo de optimizar la toma de decisiones, basándose en las sentencias SQL realizadas en una base de datos relacional que están enlazadas con la medición o conjunto de mediciones de los resultados de los procesos de negocio.

Cada Modelo Dimensional está basado en un modelo estrella, donde los valores numéricos asociados a los procesos de negocios están concentrados en una tabla de valores llamada tabla de hechos como núcleo del modelo, y con un conjunto de tablas más pequeñas llamadas tablas de dimensiones referidas a la información contextual que se encuentran conectadas a la tabla de hechos. Los elementos de estas tablas se pueden definir de la siguiente manera:

- Hechos: es una colección de piezas de datos y datos de contexto. Cada hecho representa una parte del negocio, una transacción o un evento.
- Dimensiones: es una colección de miembros, unidades o individuos del mismo tipo.

- Medidas: son atributos numéricos de un hecho que representan el comportamiento del negocio relativo a una dimensión.

Cada punto de entrada a la tabla de hechos está conectado a una dimensión, lo que permite determinar el entorno de los hechos.

Una base de datos dimensional se puede comprender como un cubo de tres o cuatro dimensiones (OLAP), en el que los usuarios pueden acceder a una porción de la base de datos a lo largo de cualquiera de sus dimensiones.

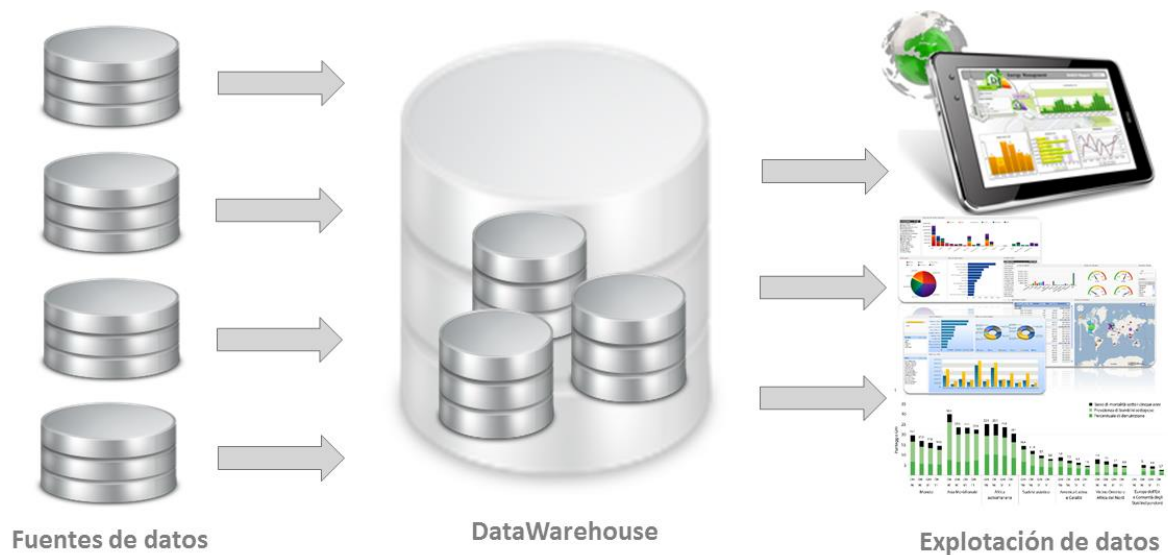


Ilustración 33: Modelo Dimensional Kimball¹⁹

Metodología Inmon

Bill Inmon ve la necesidad de transferir la información de los diferentes sistemas transaccionales de las organizaciones a un lugar de los datos puedan ser utilizados para el análisis.

Para esta metodología un Data Warehouse es considerado como un conjunto de datos con las siguientes características:

- Orientado a temas: Los datos en la base de datos están organizados de manera que todos los elementos de datos relativos al mismo evento u objeto del mundo real queden unidos entre sí.
- Integrado: La base de datos contiene los datos de todos los sistemas operacionales de la organización, y dichos datos deben ser consistentes.

¹⁹ Extraído de: <https://blog.bi-geek.com/arquitectura-el-enfoque-de-ralph-kimball/>

- No volátil: La información no se modifica ni se elimina, una vez almacenado un dato, este se convierte en información de solo lectura, y se mantiene para futuras consultas.
- Variante en el tiempo: Los cambios producidos en los datos a lo largo del tiempo quedan registrados para que los informes que se puedan generar reflejen esas variaciones.

El objetivo de este almacén es ser de base importante para la toma de decisiones estratégicas. El Data Warehouse es una parte de un sistema de BI y los Data Marts obtienen su información a partir de este Data Warehouse y antes de que los datos sean almacenados deben ser normalizados.

Debido a las enormes cargas de datos que deben producirse, Inmon considera que el ambiente de origen de los datos y el ambiente de acceso de datos deben estar físicamente separados en diferentes bases de datos y en equipos separados, y los datos deben estar cuidadosamente gestionados y condensados. En esta metodología los Data Marts son considerados como capas del Data Warehouse y los Data Marts son dependientes del depósito central de datos o Data Warehouse corporativo y por lo tanto se construyen después de él.

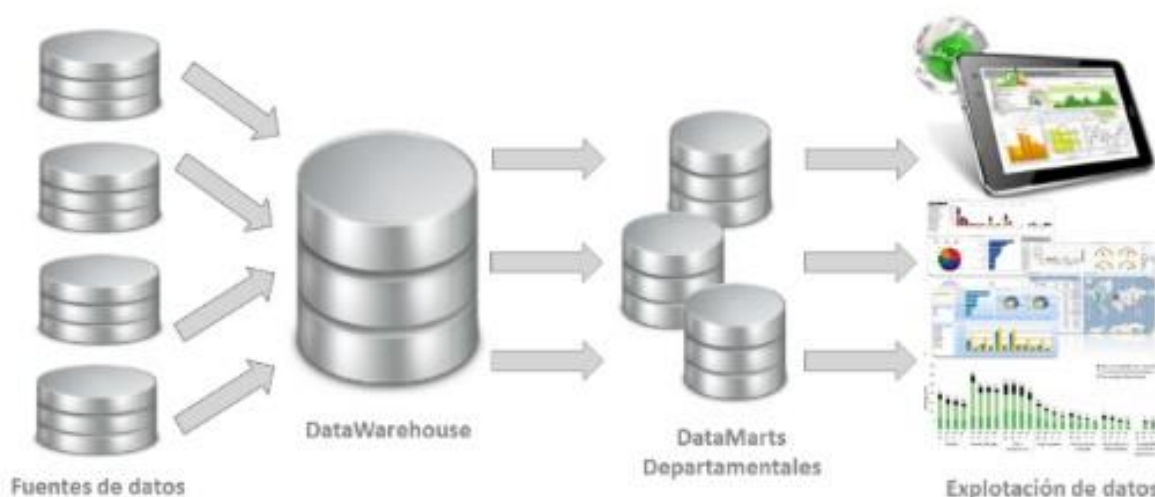


Ilustración 34: Modelo Relacional Inmon²⁰

Metodología CRISP-DM

CRISP-DM es una de las principales metodologías que siguen los analistas en la inteligencia de negocios, donde se puede rescatar primordialmente Data Warehouse y Data Mining.

Esta metodología se encuentra sustentada en modelos internacionales que reflejan la robustez de sus procesos y que facilitan la unificación de sus fases en una estructura confiable y amigable para el usuario. Además de ello, esta tecnología interrelaciona las diferentes fases del proceso entre sí, de tal manera que se consolida un proceso iterativo y recíproco. Otro

²⁰ Extraído de: <https://blog.bi-geek.com/arquitectura-enfoque-de-william-h-inmon/>

aspecto fundamental de esta tecnología es que es planteada como una metodología imparcial o “neutra respecto a la herramienta que utilice para el desarrollo del proyecto de Data Warehouse o Data Mining siendo su distribución libre y gratuita”²¹.

El ciclo de vida del proyecto según la metodología CRISP está basado en seis fases variables entre si no concluyentes, lo cual se lo considera como un ciclo en constante movimiento y mejora. Las fases son las siguientes:

1. Comprensión del negocio: Se trata de entender claramente los requerimientos y objetivos del proyecto siempre desde una visión de negocio.
2. Comprensión de los datos: Es conseguir y habituarse con los datos, reconocer las dificultades en la calidad de los datos y reconocer también las fortalezas de estos mismos que pueden servir en el proceso de análisis.
3. Preparación de los datos: Es analizar los datos realmente importantes en el proceso de selección, depuración y transformación.
4. Modelado: Es la aplicación de técnicas de modelado o de Data Warehouse
5. Evaluación: Esta fase es muy importante y decisiva, pues corresponde a la evaluación de la escogencia de los modelos anteriores y la toma de decisión respecto a si realmente son útiles en el proceso.
6. Despliegue o divulgación: Es la fase de implementación o de divulgación de los modelos anteriormente escogidos y evaluados.

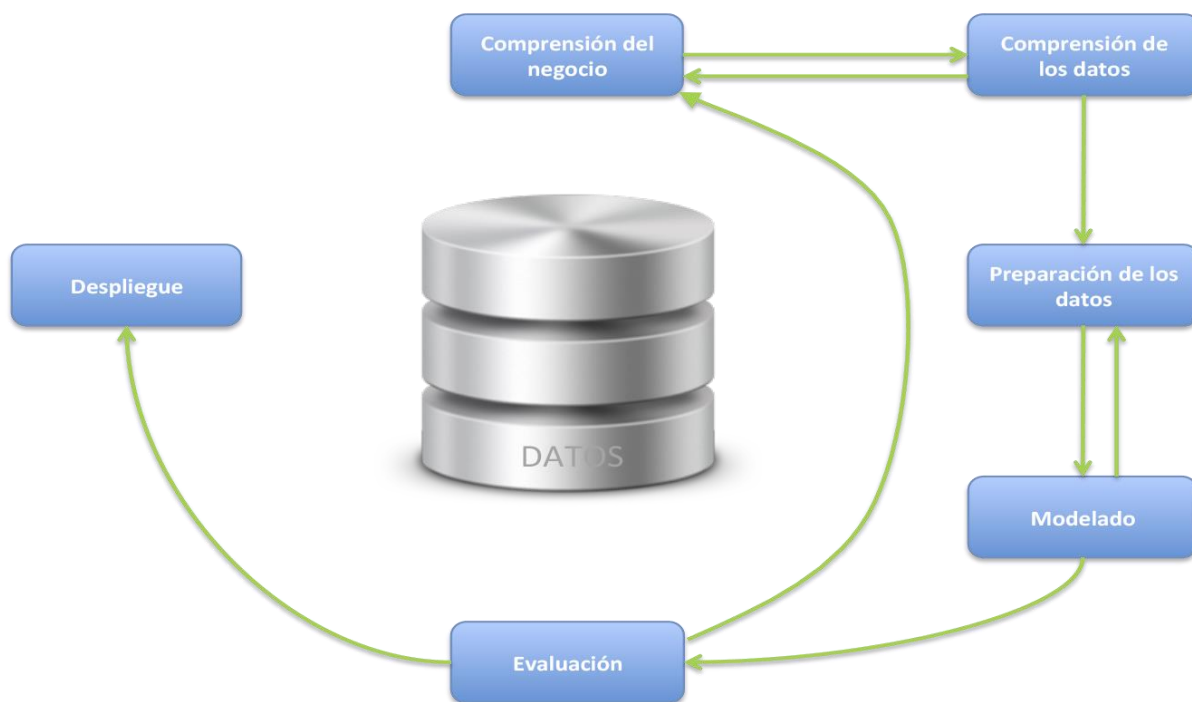


Ilustración 35: Modelo CRISP-DM²²

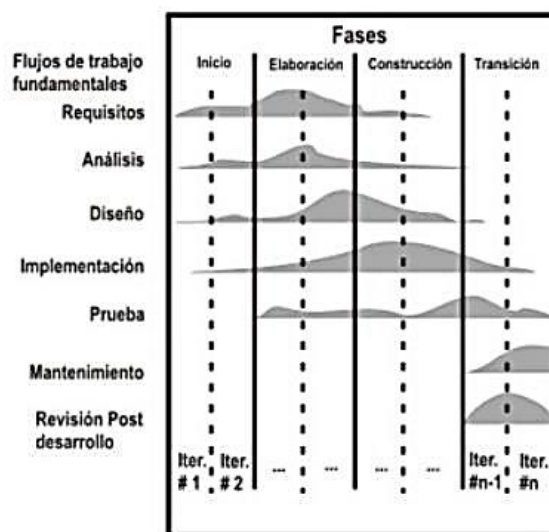
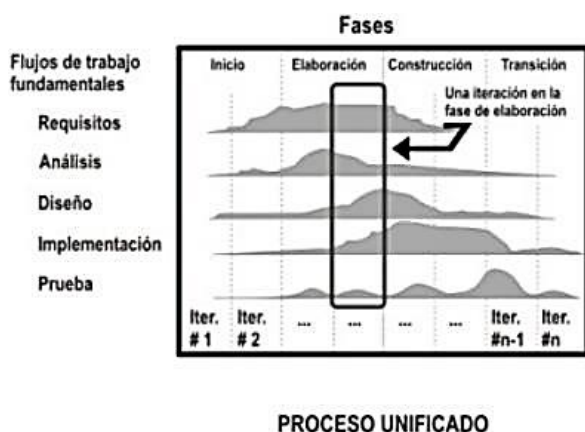
²¹ J. Hernández, “Minería de Datos”. [En línea]. Disponible: <http://www.estadistico.com/arts.html?20040426>

²² Extraído de: <http://www.diegocalvo.es/modelo-crisp-drm-data-mining/>

Metodología DWEP (Data Warehouse Engineering Process)

Esta metodología, según (Luján Mora & Trujillo, 2004), está basada en la metodología de Proceso Unificado (PU) y hace uso de los artefactos propuestos por UML (Lenguaje unificado de modelado), propuesto por Sergio Lujan-Mora y Juan Trujillo. Esta metodología hace uso del ciclo de vida propuesto por PU (Incepción, elaboración, construcción y transición) además de sus flujos de trabajo (Requerimientos, análisis, diseño, implementación y pruebas) y agrega dos flujos de trabajo Mantenimiento y Revisión post-desarrollo.

1. **Requerimientos:** Durante esta fase se hace uso de los diagramas de caso de uso para modelar los requerimientos de usuario, además con su especificación respectiva.
2. **Análisis:** En esta fase se busca depurar y estructurar los requerimientos de usuario, definir las fuentes de datos y de ser necesario documentarlas con la finalidad de tener claras las fuentes de datos.
3. **Diseño:** Durante esta fase se realiza el diseño del data warehouse conceptualmente y el mapeo que será usado para generar la data que se guardará en el data warehouse desde los sistemas fuente.
4. **Implementación:** Durante este proceso se realiza la construcción del data warehouse, se realiza la carga inicial de datos y se optimiza el performance del data warehouse.
5. **Pruebas:** Se realizan las pruebas del data warehouse creado para verificar que el diseño sea el correcto y se realizan las modificaciones necesarias.
6. **Mantenimiento:** Esta fase busca mantener el data warehouse actualizado con los nuevos datos generados en las fuentes de data. Además, se captan nuevos requerimientos y realizan nuevos desarrollos.
7. **Revisión post-desarrollo:** Esta fase busca la obtención de buenas prácticas y lecciones aprendidas, así como la adecuación del proceso de desarrollo para mejorar el mismo.



PROCESO DE INGENIERÍA DE ALMACÉN DE DATOS

Ilustración 36: Proceso Unificado y DWEP²³

²³ Extraído de: https://www.researchgate.net/figure/Figura-3-a-Proceso-Unificado-y-b-DWEP_fig2_316634417



Comparativa de Metodologías

Metodología	Kimball	Inmon	CRISP-DM	DWEP	HEFESTO
Autor	Ralph Kimball	Bill Inmon	SPSS, Teradata, Daimler AG, NCR y Ohra	Luján Mora & Trujillo	Bernabeu Ricardo Darío
Énfasis	Data Mart	Data Warehouse	Data Warehouse y Data Mining	Data Warehouse	Data Warehouse
Resumen	1. Planeación del proyecto 2. Requerimientos del negocio 3. Diseño Arquitectur a Técnica 4. Selección e instalación de herramienta s 5. Modelo dimensional 6. Diseño físico 7. Especificaciones herramienta analítica 8. Desarrollo herramienta analítica 9. Área de staging 10. Desarrollo 11. Mantenimiento	1. Implementación Data Warehouse 2. Integración de datos 3. Pruebas 4. Programación sobre datos 5. Diseño DSS 6. Análisis resultados 7. Entendimiento	1. Comprensión del negocio 2. Comprensión de los datos 3. Preparación de los datos 4. Modelado 5. Evaluación 6. Despliegue	1. Requerimientos 2. Análisis 3. Diseño 4. Implementación 5. Pruebas 6. Mantenimiento 7. Revisión post-desarrollo	1. Análisis de requerimientos 2. Análisis OLTP 3. Modelo lógico Data Warehouse 4. Integración de datos
Arquitectura	Bottom-up	Top-Down	Ciclo en Espiral	UML y PU	Híbrida
	Está basada en experimentos y prototipos, es más flexible y menos costosa	Los problemas que se van a solventar ya son conocidos de antemano, tiene un mayor costo ya que cada fase alcanza un alto nivel de detalle.	Esta más orientada hacia los proyectos de data Mining	Desarrollado principalmente para la creación de software desde cero, requiere demasiados recursos.	Se enfoca en el análisis de los requerimientos para identificar indicadores, perspectivas y realizar el análisis de los datos.

Tabla 19: Comparativa metodologías BI. Elaboración propia

Selección de Metodología

Se ha seleccionado la Metodología HEFESTO, debido a que las fases que esta comprende, se acopla al desarrollo y está enfocada al análisis de requerimientos, identificación de identificadores y perspectivas, que permitirán realizar el análisis de datos respectivos y obtener resultados que apoyen a la toma de definiciones de la organización.

Metodología HEFESTO

La metodología HEFESTO, es un proyecto comunitario liderado por el Ing. Bernabeu Ricardo Dario. La idea principal de esta metodología, según lo expresa su autor, es “comprender cada paso que se realizará, para no caer en el tedio de tener que seguir un método al pie de la letra sin saber exactamente qué se está haciendo ni por qué” [Bernabeu, 2010]. El autor explica en la documentación de la metodología HEFESTO, que puede ser embebida en cualquier ciclo de vida de desarrollo de software, siempre y cuando se adapte para trabajar con la metodología seleccionada. Además, resalta la importancia de tener fases de desarrollos cortos que pueden demostrar las ventajas de usar un data warehouse y de esta forma motivar a los usuarios. La metodología HEFESTO puede resumirse a través de la siguiente figura:

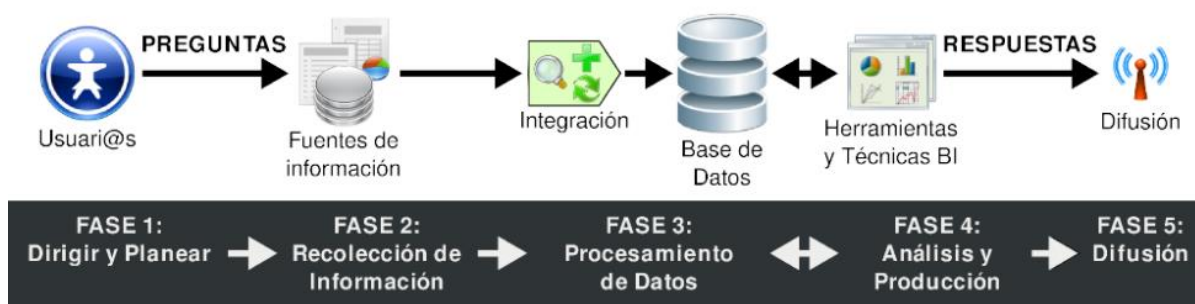


Ilustración 37: Metodología HEFESTO²⁴

Descripción

La metodología HEFESTO inicia con la recolección de las necesidades de información de los usuarios obteniendo así las preguntas claves del negocio. Luego, se deben identificar los indicadores resultantes y sus perspectivas de análisis, mediante las cuales se construirá el modelo conceptual de datos del AD, se analizarán las bases de datos de los OLTP o fuentes externas para señalar las correspondencias con los datos fuentes y seleccionar los campos de estudio de cada perspectiva. Una vez hecho eso, se pasará a la construcción del modelo lógico, tomando en cuenta las jerarquías que intervendrán. Por último, se definirán los procesos de carga, transformación, extracción y limpieza de los datos fuente.

A continuación, podemos ver cada uno de los pasos correspondientes a la metodología HEFESTO y que resumen la descripción antes mencionada, sobre los procesos que se siguen.

1. Análisis de requerimientos
 - a. Preguntas del Negocio
 - b. Indicadores y Perspectivas
 - c. Modelo Conceptual
2. Análisis de Data Sources
 - a. Hechos e indicadores
 - b. Mapeo
 - c. Granularidad
 - d. Modelo Conceptual Ampliado

²⁴ Bernabeu R. Dario Y García M. Mariano, 2017. “HEFESTO - DATA WAREHOUSING”, Guía completa de aplicación teórico-práctica; metodología Data Warehouse.



3. Modelo lógico del DW
 - a. Tipología
 - b. Tablas de dimensiones
 - c. Tablas de hechos
 - d. Uniones
4. Integración de datos
 - a. Carga inicial
 - b. Actualización

Características

La metodología HEFESTO posee las siguientes características:

- Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- La piedra fundamental la constituyen los requerimientos de los usuarios, por lo cual se adapta con facilidad y rapidez a los cambios del negocio.
- Reduce drásticamente la resistencia al cambio, ya que involucra a los usuarios finales en cada etapa para que tomen decisiones respecto al comportamiento y funciones del DW, y además expone resultados inmediatos.
- Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
- Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.
- Es independiente del software/hardware que se utilicen para implementación.
- Es independiente de las estructuras físicas que contengan el DW y de su respectiva distribución.
- Cuando se culmina con una fase, los resultados obtenidos se constituyen en la entrada de la fase siguiente.
- Se aplica en Data Warehouse y en Data Mart.

Análisis de Requerimientos

Se realizaron reuniones con el personal que trabaja en el programa de Becas, para determinar las necesidades y requerimientos con el fin de poder definir qué información necesitan obtener del resultado de aplicar BI a la organización una vez realizado el desarrollo.

Preguntas de negocio

Se realizó una entrevista con el coordinador del programa de Becas haciendo preguntas acerca de la información pertinente que se necesita para la formulación del Data Warehouse,



para determinar qué información requieren revisar constantemente, de tal manera que facilite sus actividades diarias y no lo realicen de manera manual u operativa.

Las preguntas de negocio obtenidas fueron las siguientes:

- Deseamos saber cuál es la cantidad de personas por tipo de institución de los municipios de Salta que reciben una beca en cierto periodo. O, en otras palabras: Número de evaluaciones aceptadas de cada tipo de institución de cada municipio en un tiempo determinado.
- Por último, también quisiéramos saber por municipio cuales condiciones son las que impiden a una persona obtener una beca en cierto periodo. O, en otras palabras: Número de evaluaciones rechazadas de cada municipio que no cumplen algún tipo de condición y/o requisito en un tiempo determinado.

Indicadores y Perspectivas

A continuación, se analizarán las preguntas obtenidas en el paso anterior y se detallarán cuáles son sus indicadores y perspectivas.

Número de evaluaciones aceptadas de cada tipo de institución de cada municipio en un tiempo determinado.

Número de evaluaciones rechazadas de cada municipio que no cumplan algún tipo de condición en un tiempo determinado.

Los indicadores son:

- Número de evaluaciones aceptadas
- Número de evaluaciones rechazadas

Perspectivas

- Tipo de institución
- Municipio
- Tipo de condición
- Tiempo

Modelo Conceptual

En esta etapa, se construirá un Modelo conceptual a partir de los Indicadores y Perspectivas obtenidas en el paso anterior.



Ilustración 38: Modelo Conceptual. Elaboración propia

La relación mediante la cual se unen las diferentes perspectivas, para obtener como resultado los Indicadores requeridos por los usuarios, es precisamente Evaluación.

Análisis de Data Sources

Se analizará las fuentes OLTP, las cuales ayudarán a determinar cómo se debe calcular los indicadores y establecer las respectivas correspondencias que existen entre el modelo conceptual creado y las fuentes de datos que se tiene.

Hechos e Indicadores

Los indicadores se calcularán de la siguiente manera:

- Indicador: Número de evaluaciones aceptadas
 - Hechos: Número de evaluaciones aceptadas
 - Función de agregación: COUNT
 - Aclaración: el Indicador Número de evaluaciones aceptadas representa la sumatoria de evaluaciones de personas que cumplieron todas las condiciones y consiguieron la beca.

- Indicador: Número de evaluaciones rechazadas
 - Hechos: Número de evaluaciones rechazadas
 - Función de agregación: COUNT
 - Aclaración: el Indicador Número de evaluaciones rechazadas representa la sumatoria de evaluaciones de personas que no lograron conseguir la beca por algún tipo de razón que no lograron cumplir.

Maapeo

En el Data Source del programa de Becas, el proceso de evaluación está representado por el DER expuesto anteriormente. A continuación, se expondrá el Maapeo entre el modelo relacional y el modelo conceptual:

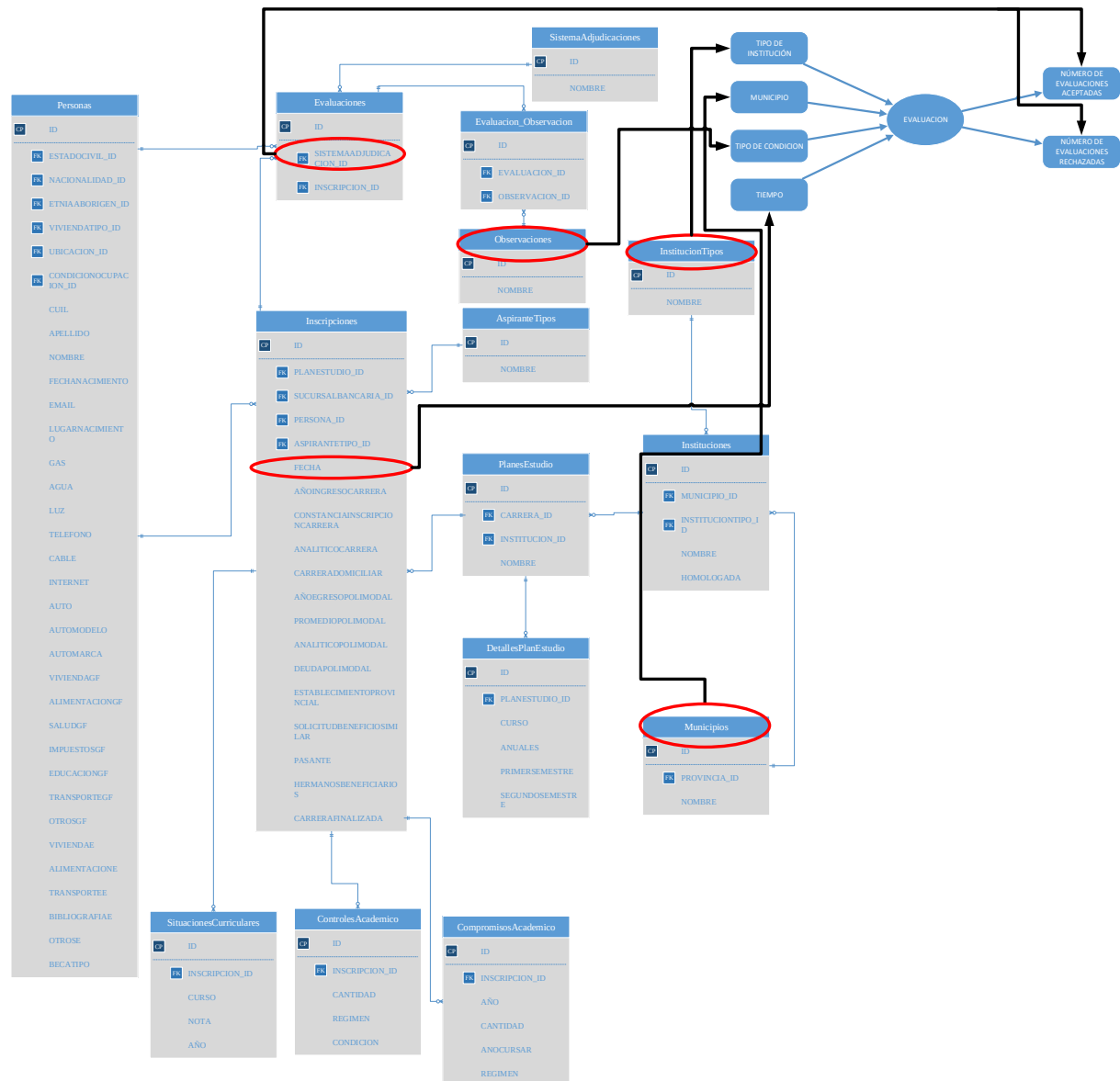


Ilustración 39: Maapeo entre modelos²⁵. Elaboración propia

El Maapeo realizado es el siguiente:

- La perspectiva Tipo de institución se relaciona con la tabla InstitucionTipos.
- La perspectiva Municipio se relaciona con la tabla Municipios.
- La perspectiva Tipo de condición se relaciona con la tabla Observaciones.

²⁵ El modelo relacional en esta imagen posee una menor cantidad de tablas para que se pueda apreciar mejor el mapeo.



- La perspectiva Tiempo se relaciona con el atributo FECHA de la tabla Inscripciones debido a que el proceso de evaluación esta por la fecha en la que se carga la inscripción.
- El indicador Número de Evaluaciones Aceptadas se relaciona con el atributo SISTEMAADJUDICACION_ID de la tabla Evaluaciones, quedando como fórmula de cálculo la siguiente:
 - COUNT(SISTEMAADJUDICACION_ID)
- El indicador Numero de Evaluaciones Rechazadas se relaciona SISTEMAADJUDICACION_ID de la tabla Evaluaciones, quedando la fórmula de cálculo como sigue:
 - COUNT(SISTEMAADJUDICACION_ID)

Granularidad

De acuerdo al mapeo realizado, se analizaron los campos que constituyen cada tabla a la que se hace referencia a través de dos métodos diferentes. Como ya se mencionó en apartados anteriores, una de las limitaciones que se presentaron en este proyecto, es la de no contar con la base de datos original del sistema. Por lo que la fuente de datos de la cual se plantea desarrollar esta solución es una base de datos creada a partir de cómo funciona el sistema, reglamento de requisitos y condiciones, etc.

Observando cómo funciona el sistema y cuáles son los datos que se piden a la hora de ingresar nuevas inscripciones se pudo constatar cuáles serían los datos y su tipo.

Con respecto a la perspectiva Tipo de Institución, los datos disponibles son los siguientes:

- ID: es la clave primaria de la tabla InstitucionTipos, y representa unívocamente a un tipo de institución.
- NOMBRE: representa la denominación del tipo de institución a la que se hace referencia.

En la perspectiva Municipio, los datos disponibles son los siguientes:

- ID: es la clave primaria de la tabla Municipios, y representa unívocamente a un Municipio
- PROVINCIA_ID: representa a través de una clave foránea la provincia a la que pertenece el municipio.
- NOMBRE: nombre del Municipio.

Con respecto a la perspectiva Tipo de Condición, los datos que se pueden utilizar son los siguientes:

- ID: es la clave primaria de la tabla Observaciones, y representa unívocamente a una observación.
- NOMBRE: expresión que representa al tipo de observación.

Con respecto a la perspectiva Tiempo, que es la que determinará la granularidad del DW, los datos más típicos que pueden emplearse son los siguientes:



- Año
- Semestre
- Número de mes
- Nombre del mes
- Quincena
- Semana
- Número de día
- Nombre del día

Una vez finalizada la recolección de la información pertinente, y consultados los usuarios sobre los datos que consideraban de interés para analizar los indicadores, los resultados obtenidos fueron los siguientes:

Perspectiva Tipo de Institución:

- NOMBRE de la tabla InstitucionTipos. Ya que este hace referencia a la denominación del tipo de institución.

Perspectiva Municipio:

- NOMBRE de la tabla Municipios. Ya que este hace referencia al nombre del municipio.

Perspectiva Tipo de Condición:

- OBSERVACIONES de la tabla Observaciones. Ya que esta hace referencia al atributo nombre de la tabla Observaciones.

Perspectiva Tiempo:

- Año

Modelo Conceptual Ampliado

Se amplía el diseño modelo conceptual con los campos seleccionados en el paso anterior:

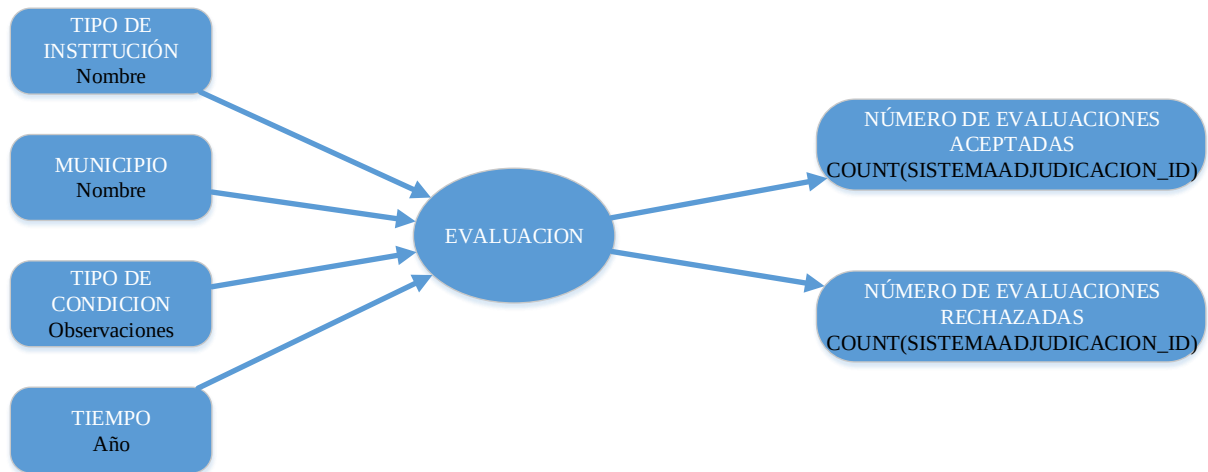


Ilustración 40: Modelo Conceptual Ampliado. Elaboración propia

Modelo Lógico del DW

A continuación, se confeccionará el Modelo Lógico de la estructura del DW, teniendo como base el Modelo Conceptual que ya ha sido creado.

Tipología

Se ha seleccionado el Esquema en Estrella ya que cumple con los requerimientos planteados y es simple de implementar y comprender.

El diseño de esquemas en estrella permite implementar la funcionalidad de una base de datos multidimensional utilizando una clásica base de datos relacional. Otra razón para utilizar los esquemas en estrella es su simplicidad desde el punto de vista del usuario final. Las consultas no son complicadas, ya que las condiciones y las uniones necesarias solo involucran a la tabla de hechos y a las de dimensiones.

Tabla de Dimensiones

Perspectiva Tipo de Institución:

- La nueva tabla de dimensión tendrá el nombre de dimInstitucionTipos
- Se le agregará una clave principal con el nombre de idInstitucionTipo
- Se modificará el nombre del campo NOMBRE por institucionTipo

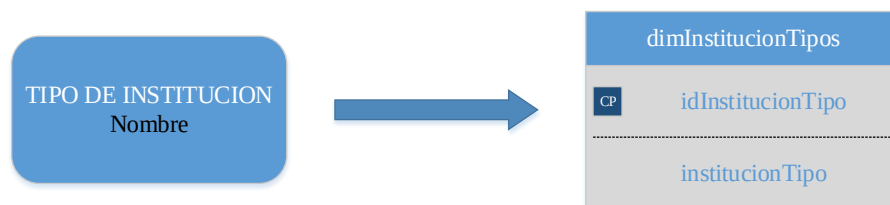


Ilustración 41: Dimensión InstitucionTipos. Elaboración propia

Perspectiva Municipio:

- La nueva tabla de dimensión tendrá el nombre de dimMunicipios
- Se le agregará una clave principal con el nombre de idMunicipio
- Se modificará el nombre del campo NOMBRE por municipio

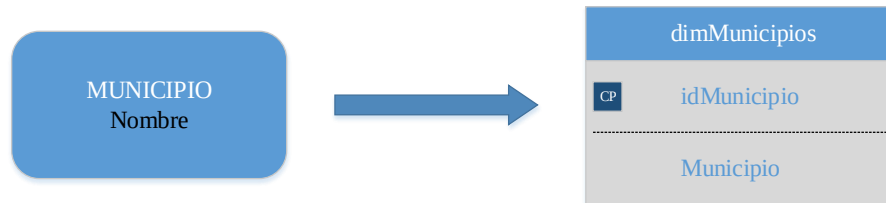


Ilustración 42: Dimensión Municipios. Elaboración propia

Perspectiva Tipo de Condición:

- La nueva tabla de dimensión tendrá el nombre de dimCondiciónTipos
- Se le agregará una clave principal con el nombre de idCondicionTipo
- Se modificará el nombre del campo OBSERVACIONES por condicionTipo

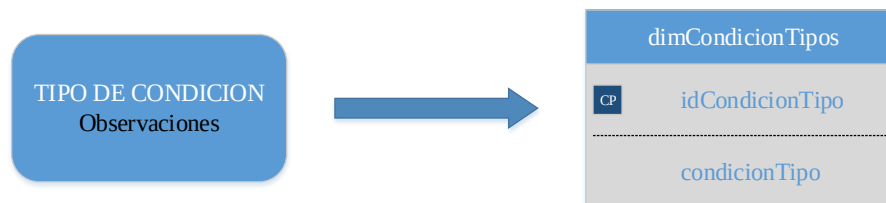


Ilustración 43: Dimensión CondicionTipos. Elaboración propia

Perspectiva Tiempo:

- La nueva tabla de dimensión tendrá el nombre de dimFechas
- Se le agregará una clave principal con el nombre de idFechas
- Se le modificará el nombre del campo AÑO por anio

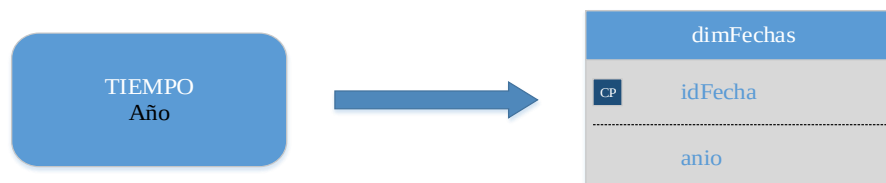


Ilustración 44: Dimensión Fechas. Elaboración propia

Tabla de Hechos

A continuación, se confeccionará la tabla de Hechos:

- La tabla de Hechos tendrá el nombre de factEvaluaciones.
- Su clave principal será la clave primaria de la tabla evaluaciones.

- Tendrá de claves foráneas idInstitucionTipo heredada de la dimensión dimInstitucionTipos, idMunicipio heredada de la dimensión dimMunicipios, idCondicionTipo heredada de la dimensión dimCondicionTipos e idFecha heredada de la dimensión dimFechas.
- Se crearán dos Hechos, que se corresponden con los dos indicadores:
 - Número de Evaluaciones Aceptadas será renombrado como aceptada
 - Número de Evaluaciones Rechazadas será renombrado como rechazada

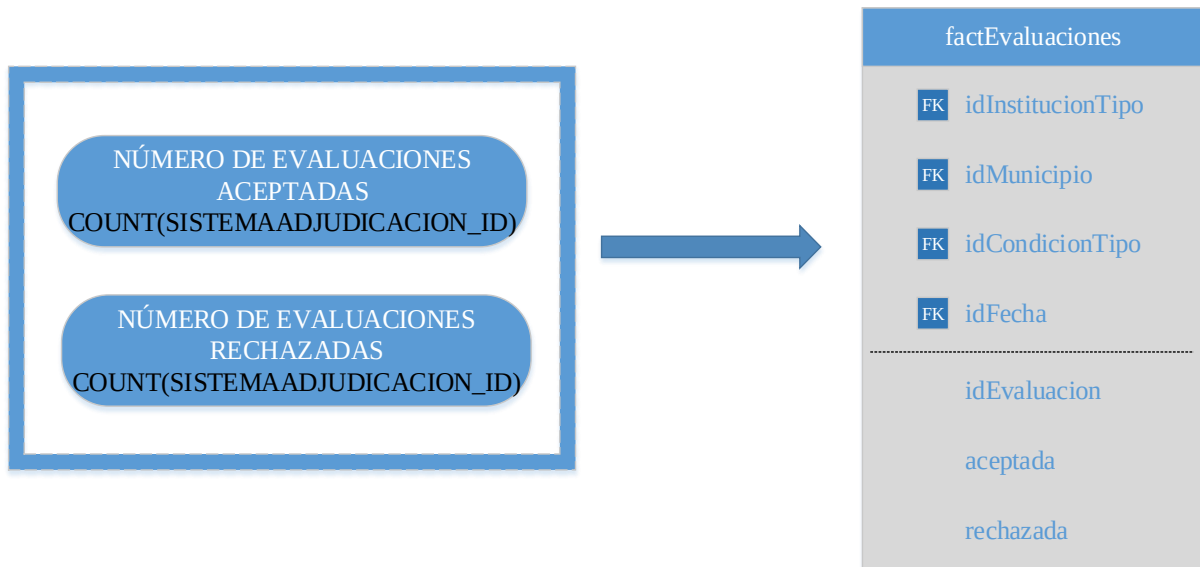


Ilustración 45: Tabla de Hecho Evaluaciones. Elaboración propia

Uniones

Se realizarán las uniones necesarias según corresponda:

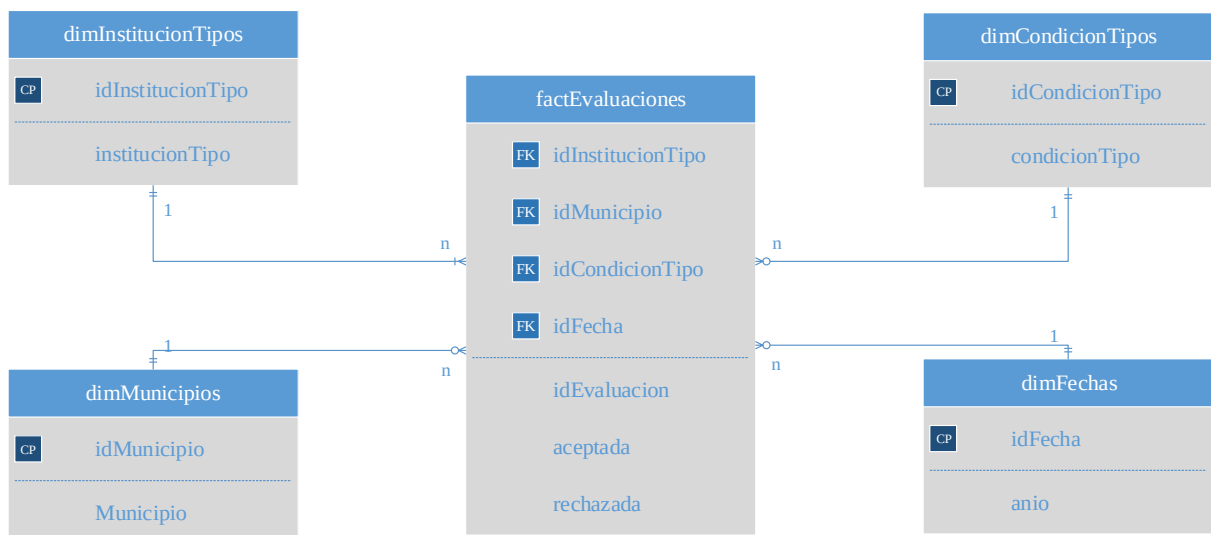


Ilustración 46: Modelo Dimensional. Elaboración propia

Integración de Datos

En esta parte del proceso se procede a cargar el Modelo Lógico con datos, utilizando procesos ETL.

Se realizan los procesos que se llevarán a cabo para lograr cumplir las reglas y políticas de actualización necesarias.

Todos estos procesos serán realizados con la herramienta SQL Server Data Tools para Visual Studio.

Carga Inicial

Proceso ETL Principal

El proceso ETL principal planteado para la Carga Inicial es el siguiente:

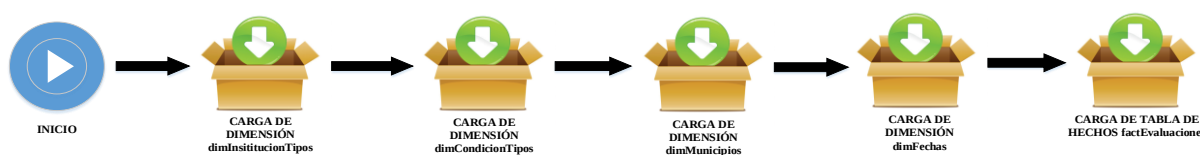


Ilustración 47: ETL - Carga Inicial. Elaboración propia

Las tareas que lleva a cabo este proceso son:

- Inicio: inicia la ejecución de los pasos en el momento en que se le indique.
- Carga de Dimensión dimInstitucionTipos: ejecuta el contenedor de Pasos que cargará la tabla de Dimensión dimInstitucionTipos (detallado más adelante).
- Carga de Dimensión dimMunicipios: ejecuta el contenedor de Pasos que cargará la tabla de Dimensión dimMunicipios (detallado más adelante).
- Carga de Dimensión dimCondicionTipos: ejecuta el contenedor de Pasos que cargará la tabla de Dimensión dimCondicionTipos (detallado más adelante).
- Carga de Dimensión dimFechas: ejecuta el contenedor de Pasos que cargará la tabla de Dimensión dimFechas (detallado más adelante).
- Carga de Tabla factEvaluaciones: ejecuta el contenedor de Pasos que cargará la tabla de Hechos factEvaluaciones (detallado más adelante).

Carga de Dimensión dimInstitucionTipos

A continuación, se especificarán las tareas llevadas a cabo por Carga de Dimensión dimInstitucionTipos. Este Paso es un Contenedor de Pasos, así que incluye las siguientes tareas:

- Obtener datos de Data Source (OLTP): obtiene a través de una consulta SQL los datos del Data Source necesarios para cargar la tabla de Dimensión dimInstitucionTipos.

Se toma como fuente de datos la tabla **institucionTipos** del Data Source como habíamos mencionado anteriormente.



Se determinó con el coordinador del programa y se confirmó con un usuario administrativo, que los Cursos de capacitación no serán incluidas dentro de los tipos de instituciones, debido a que existen muy pocas instituciones de este tipo y se está planteando para futuros años descartarlas del programa.

Los cursos de capacitación están referenciados en la tabla como **Curso de Capacitacion** con ID igual a 1.

A continuación, se expondrá la sentencia SQL configurada en este Paso:

```
SELECT
    id AS idInstitucionTipo,
    nombre AS institucionTipo
FROM
    becas.instituciontipos
WHERE
    id <> 1
ORDER BY
    id,
    nombre
```

- Cargar dimInstitucionTipos: almacena en la tabla de Dimensión dimInstitucionTipos los datos obtenidos en el Paso anterior.

Carga Dimensión dimCondicionTipos

Se especificarán las tareas llevadas a cabo por Carga de Dimensión dimCondicionTipos. Este Paso es un Contenedor de Pasos así que incluye las siguientes tareas:

- Obtener datos de Data Source (OLTP): obtiene a través de una consulta SQL los datos del Data Source para cargar la Dimensión dimCondicionTipos.

Se empleará como fuente de la dimensión, la tabla de **observaciones**.

A continuación, se expondrá la sentencia SQL configurada en este Paso:

```
SELECT
    id AS idCondicionTipo,
    nombre AS condicionTipo
FROM
    becas.observaciones
ORDER BY
    id,
    nombre
```

- Cargar dimCondicionTipos: almacena en la tabla de Dimensión dimCondicionTipos los datos obtenidos en el Paso anterior.

Carga de Dimensión dimMunicipios



Se especificarán las tareas llevadas a cabo por Carga de Dimensión dimMunicipios. Este Paso es un Contenedor de Pasos así que incluye las siguientes tareas:

- Obtener datos de Data Source (OLTP): obtiene a través de una consulta SQL los datos del Data Source necesarios para cargar la Dimensión dimMunicipios.

La fuente que será empleada para esta dimensión, es la tabla de **municipios**.

Es necesario aclarar que la base de datos el sistema solo contiene los municipios de la provincia de Salta, aunque en la tabla tenga una clave foránea para indicar la provincia a la que pertenece el municipio. Pero pensando en un futuro en caso de que este programa provincial se convierta en un programa de escala nacional, se decide agregar la condición de que se tomen solos los municipios que pertenecen a la provincia de Salta.

En la tabla **provincias** aparece referenciado en el campo **nombre** como **Salta**.

La sentencia SQL configurada en este paso es la siguiente:

```
SELECT
    id AS idMunicipio,
    nombre as municipio
FROM
    becas.municipios
WHERE
    provincia_id = (
        SELECT
            id
        FROM
            becas.provincias
        WHERE
            nombre = 'Salta'
    )
ORDER BY
    id,
    nombre
```

- Cargar dimMunicipios: almacena en la tabla de Dimensión dimMunicipios los datos obtenidos en el Paso anterior.

Carga Dimensión dimFechas

Se especificarán las tareas llevadas a cabo por Carga de Dimensión dimFechas. Este Paso es un Contenedor de Pasos, así que incluye las siguientes tareas:

- Obtener datos de Data Source (OLTP): obtiene a través de una consulta SQL el campo **año** de la Dimensión dimFechas.

La fuente empleada para esta dimensión es el campo **fecha** de la tabla **inscripciones**.



El clave primaria **idFecha** será formado por los 3 dígitos de la izquierda del año de la fecha correspondiente.

El campo **anio** será resuelto tomando solo el año de la fecha.

La sentencia SQL configurada en este paso es la siguiente:

```
SELECT DISTINCT
    RIGHT(YEAR(fecha),3) AS idFecha,
    YEAR(fecha) AS anio
FROM
    becas.inscripciones
ORDER BY
    YEAR(fecha)
```

- Cargar dimFechas: almacena en la tabla de Dimensión dimFechas los datos obtenidos en el Paso anterior.

Carga de Tabla de Hechos factEvaluaciones

Se especificarán las tareas llevadas a cabo por Carga de Tabla de Hechos factEvaluaciones. Este Paso es un Contenedor de Pasos, así que incluye las siguientes tareas:

- Obtener datos de Data Source: obtiene a través de una consulta SQL los datos del Data Source necesarios para cargar la tabla de Hechos factEvaluaciones.

La tabla de Hechos fue conformada con las tablas evaluaciones, inscripciones, personas, domicilios, planesEstudio, instituciones, institucionTipos e evaluacion_observacion. Se analizaron las condiciones que deben cumplir los datos para considerarse de interés. En este caso se trabajaran con aquellas evaluaciones que hayan sido aprobadas y rechazadas.

El campo que brinda esta información es **sistemaAdjudicacion_id** que hace referencia a un **id** de la tabla **sistemaAdjudicaciones**:

- Cuando **sistemaAdjudicacion_id** es igual a 1 hace relación a una evaluación aprobada
- Cuando **sistemaAdjudicacion_id** es igual a 2 hace relación a una evaluación rechazada

El campo **idEvaluacion** será el campo **id** obtenida de la tabla evaluaciones.

Se formará de la misma manera el campo **idFecha** como se la construyo en la dimensión **dimFechas**.

Se decide aplicar una pre-agregación a los Hechos que formarán parte de la tabla de Hechos; es por esta razón que se utilizará la cláusula **GROUP BY** para agrupar todos los registros a través de las claves primarias y secundarias de esta tabla.

La sentencia SQL configurada en este Paso es la siguiente:

SELECT



```
evaluacion.id AS idEvaluacion,  
domicilio.municipio_id AS idMunicipio,  
institucion.institucionTipo_id AS idInstitucionTipo,  
evaluacionObservacion.observacion_id AS idCondicionTipo,  
right(year(inscripcion.fecha),3) AS idFecha,  
count(case evaluacion.sistemaAdjudicacion_id when 1 then 1 else 0 end) AS  
aceptada,  
count(case evaluacion.sistemaAdjudicacion_id when 2 then 1 else 0 end) AS  
rechazada  
FROM  
becas.evaluaciones AS evaluacion  
inner join becas.inscripciones AS inscripcion  
ON evaluacion.inscripcion_id = inscripcion.id  
inner join becas.personas AS persona  
ON inscripcion.persona_id = persona.id  
inner join becas.domicilios AS domicilio  
ON domicilio.persona_id = persona.id  
inner join becas.planesEstudio AS planestudio  
ON planestudio.id = inscripcion.planEstudio_id  
inner join becas.instituciones AS institucion  
ON institucion.id = planestudio.institucion_id  
inner join becas.institucionTipos AS instituciontipo  
ON instituciontipo.id = institucion.institucionTipo_id  
left join becas.evaluacion_observacion AS evaluacionObservacion  
ON evaluacionObservacion.evaluacion_id = evaluacion.id  
  
GROUP BY  
evaluacion.id,  
domicilio.municipio_id,  
institucion.institucionTipo_id,  
evaluacionObservacion.observacion_id,  
inscripcion.fecha  
ORDER BY  
idEvaluacion,  
idFecha,  
idMunicipio,  
idInstitucionTipo,  
idCondicionTipo
```

- Cargar factEvaluaciones: almacena en la tabla de Hechos factEvaluaciones los datos obtenidos en el Paso anterior

Actualización

Este proceso se ejecutará una vez realizada la carga inicial del DW, estableciendo políticas y estrategias de actualización periódica.

Las políticas de actualización que se establecieron son las siguientes:

- La información se actualizará: todos los días durante el periodo de evaluación a las 18hs.
- Los datos de las tablas de dimInstitucionTipos y dimMunicipios no tendrán modificaciones.
- Los datos de la tabla dimFechas se cargarán al comenzar un nuevo año lectivo.



- Los datos de la tabla dimCondicionTipos serán cargados en su totalidad al iniciar un nuevo año lectivo, modificando o agregando nuevos requisitos que puedan llegar a surgir si es que el reglamento es modificado.
- Los datos de la tabla de Hechos factEvaluaciones que corresponden al último año a partir de la fecha actual, serán renovados cada vez.

Habrán dos procesos ETL para la actualización del DW es similar al de Carga Inicial, con las siguientes diferencias:

- Anual:
 - Inicio: comenzará la ejecución de los pasos todos los años los primeros de enero.
 - No se cargará de nuevo las dimensiones dimInstitucionTipos y dimMunicipios.
 - Agregar nuevo año a Dimensión dimFechas: se ejecutará una consulta SQL en la que se agregará un nuevo registro con el año lectivo del corriente año a la Dimensión dimFechas. Consulta SQL:

```
IF (SELECT count(CASE anio WHEN YEAR(GETDATE()) THEN 1 ELSE 0 END)
FROM dimFechas) = 0
BEGIN
    INSERT INTO dimFechas VALUES(right(YEAR(GETDATE()),
3),YEAR(GETDATE()))
END
```

- Actualizar Dimensión dimCondicionTipos: Si se produjo algún cambio en algún registro o se agregaron nuevas condiciones impactaran en la dimensión del DW. Consulta SQL:

```
CREATE TABLE ControlTable(idControl int, nombreControl
nvarchar(255));

INSERT INTO ControlTable(idControl,nombreControl)
SELECT id AS idCondicionTipo, nombre AS condicionTipo FROM
becas.becas.observaciones EXCEPT SELECT idCondicionTipo,
condicionTipo FROM dimCondicionTipos

DECLARE @TableID INT

While EXISTS (SELECT * FROM ControlTable)
BEGIN

select top 1 @TableID = idControl
from ControlTable
order by idControl asc

IF EXISTS(SELECT * FROM dimCondicionTipos WHERE idCondicionTipo =
@TableID)
BEGIN
    UPDATE dimCondicionTipos
    SET condicionTipo = (SELECT nombreControl FROM ControlTable
WHERE idControl = @TableID)
    WHERE idCondicionTipo = @TableID
END
ELSE
```



```
BEGIN
    INSERT INTO dimCondicionTipos (idCondicionTipo,
condicionTipo)
    VALUES ((SELECT idControl FROM ControlTable WHERE idControl
= @TableID), (SELECT nombreControl FROM ControlTable WHERE idControl
= @TableID))
END

DELETE FROM ControlTable
WHERE idControl = @TableID

END

DROP TABLE ControlTable
```

- Diario
 - Inicio: comenzará la ejecución de los pasos todos los días a las 6 de la tarde durante el periodo que dura la evaluación en el programa.
 - Eliminar datos de factEvaluaciones:
 - Carga Incremental de tabla de Hechos factEvaluaciones: a la serie de pasos que realiza esta tarea, se distinguirá en la obtención de los datos una condición que traerá solo los datos del corriente año. Consulta SQL de Obtener Datos de Data Source:

```
SELECT
    evaluacion.id AS idEvaluacion,
    domicilio.municipio_id AS idMunicipio,
    institucion.institucionTipo_id AS idInstitucionTipo,
    evaluacionObservacion.observacion_id AS idCondicionTipo,
    right(year(inscripcion.fecha),3) AS idFecha,
    count(case evaluacion.sistemaAdjudicacion_id when 1 then 1 else 0
end) AS aceptada,
    count(case evaluacion.sistemaAdjudicacion_id when 2 then 1 else 0
end) AS rechazada
FROM
    becas.evaluaciones AS evaluacion
    inner join becas.inscripciones AS inscripcion
        ON evaluacion.inscripcion_id = inscripcion.id
    inner join becas.personas AS persona
        ON inscripcion.persona_id = persona.id
    inner join becas.domicilios AS domicilio
        ON domicilio.persona_id = persona.id
    inner join becas.planesEstudio AS planestudio
        ON planestudio.id = inscripcion.planEstudio_id
    inner join becas.instituciones AS institucion
        ON institucion.id = planestudio.institucion_id
    inner join becas.institucionTipos AS instituciontipo
        ON instituciontipo.id = institucion.institucionTipo_id
    left join becas.evaluacion_observacion AS evaluacionObservacion
        ON evaluacionObservacion.evaluacion_id = evaluacion.id
WHERE
    year(inscripcion.fecha) = year(GETDATE())
GROUP BY
    evaluacion.id,
    domicilio.municipio_id,
    institucion.institucionTipo_id,
    evaluacionObservacion.observacion_id,
    inscripcion.fecha
```



ORDER BY

```
idEvaluacion,  
idFecha,  
idMunicipio,  
idInstitucionTipo,  
idCondicionTipo
```

Cubo Multidimensional

Un cubo de datos permite a los datos ser modelados y vistos en diferentes dimensiones para lo cual es necesario que trabaje bajo un modelo multidimensional, basado en dimensiones y hechos. En otras palabras, un cubo procesa la información de acuerdo a un diseño específico que empata con los requisitos de información determinado de tal manera que nos permite ver de acuerdo a las necesidades establecidas, por ejemplo, podemos ver las evaluaciones aceptadas de Institutos de educación terciaria durante el año 2017 en Salta Capital, de esta manera combinamos 3 dimensiones y un hecho.

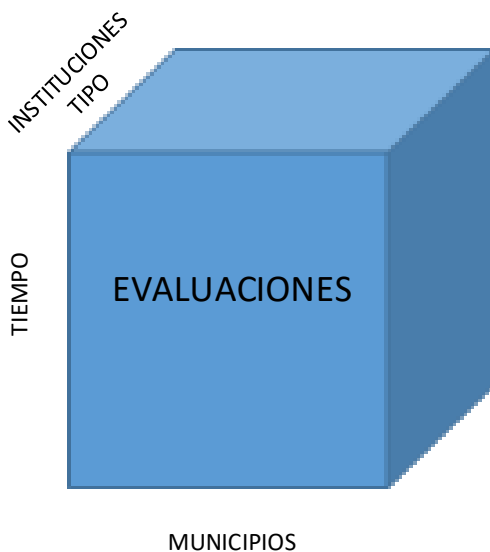


Ilustración 48: CUBO. Elaboración propia

Esta organización de los datos nos permitirá conocer la cantidad de becas que están siendo otorgadas dentro de un territorio específico para ver cuáles son los principales obstáculos que están impidiendo que una persona reciba el beneficio, identificar cuáles son los municipios que tiene la mayor cantidad de becarios.

El cubo multidimensional está conformado en base al modelo lógico del DW creado anteriormente.

Creación Cubo²⁶

²⁶ Adjunto Proyecto de Analysis Services en la carpeta **Proyecto Analysis Services** y adjunto un backup del cubo en la carpeta **Backup CUBO**.



Para la creación del Cubo se hace uso de la herramienta Microsoft SQL Server Data Tools Business Intelligence trabajando con la extensión de Analysis Services.

Una vez iniciada la herramienta se inicia un nuevo proyecto.



Ilustración 49: Nuevo Proyecto para CUBO. Elaboración propia

Se seleccionará un Proyecto multidimensional y de minería de datos de Analysis Services dentro de las plantillas de Analysis Services incluidas en Business Intelligence, se especifica el nombre del proyecto “Proyecto Analysis Services” y la ubicación donde se guardará dicho proyecto.

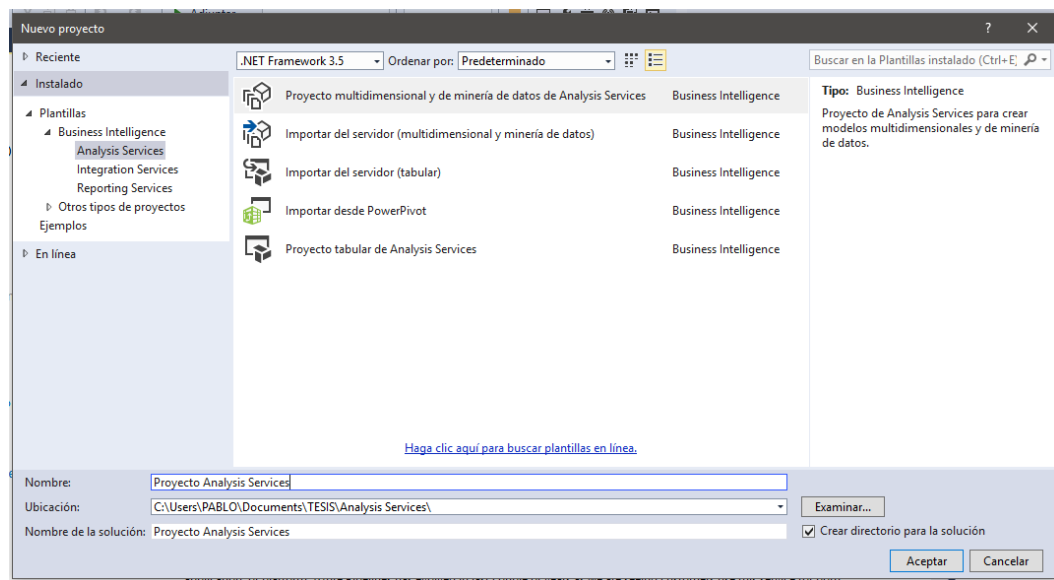


Ilustración 50: Nuevo Proyecto Analysis Services. Elaboración propia

Ya una vez creado el proyecto, en el Explorador de soluciones agregamos la fuente de datos de la cual se obtendrán los mismos para el Cubo que se va a diseñar, realizando clic izquierdo en Orígenes de datos seleccionamos un Nuevo origen de datos.

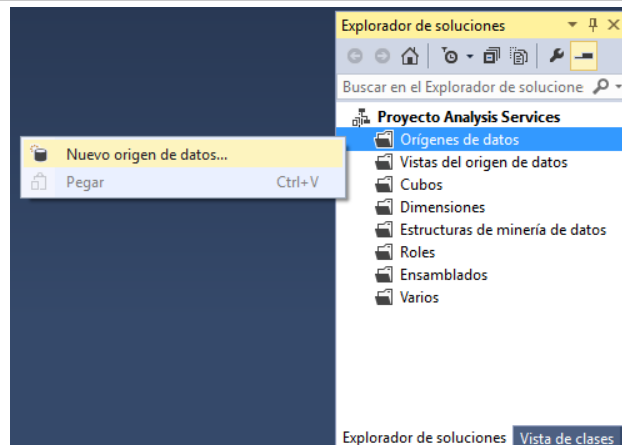


Ilustración 51: CUBO - Orígenes de datos. Elaboración propia

En el asistente para orígenes de datos seleccionamos en **Nueva** para agregar la conexión con el DW.

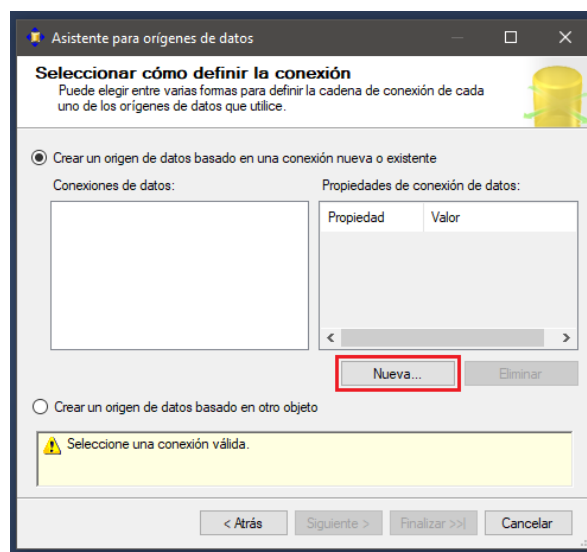


Ilustración 52: CUBO - Asistente para orígenes de datos. Elaboración propia

Debemos seleccionar el nombre del servidor en este caso DESKTOP-00FOLVM²⁷ y la base de datos becasDW.

²⁷ DESKTOP-00FOLVM es el nombre de la PC en el cual se desarrolló la tesis.

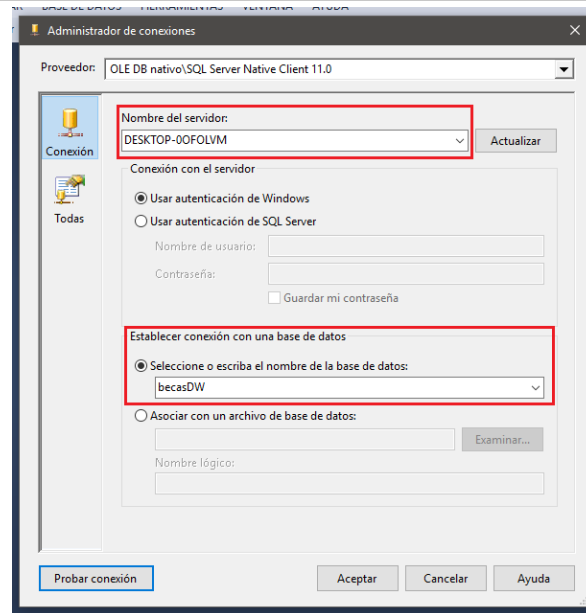


Ilustración 53: CUBO - Administrador de conexiones. Elaboración propia

Al Aceptar esta conexión se deben ingresar las credenciales que tiene esta base de datos, siendo las que se usar para iniciar sesión en el servidor.

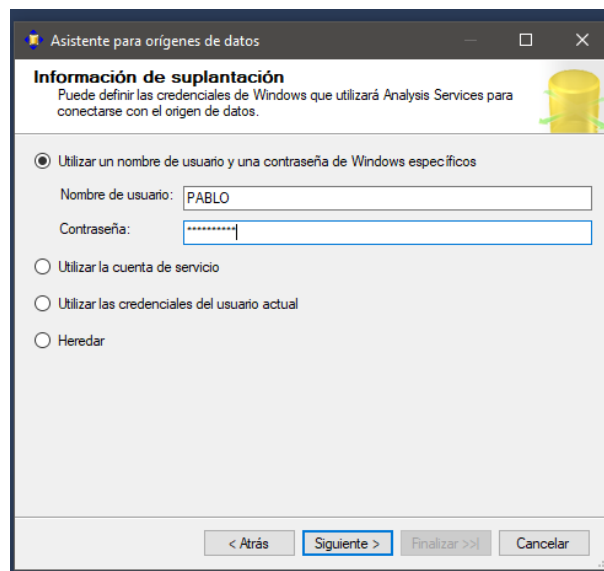


Ilustración 54: CUBO - Información de suplantación. Elaboración propia

Damos en siguiente y seleccionamos el origen de datos con la nueva conexión creada a becasDW.

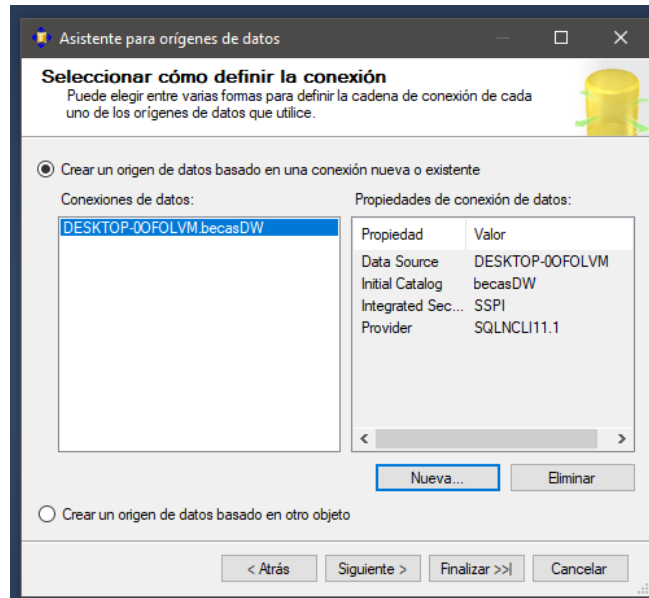


Ilustración 55: CUBO - Selección de conexión. Elaboración propia

Y para finalizar le asignamos un nombre al origen de datos Becas DW.

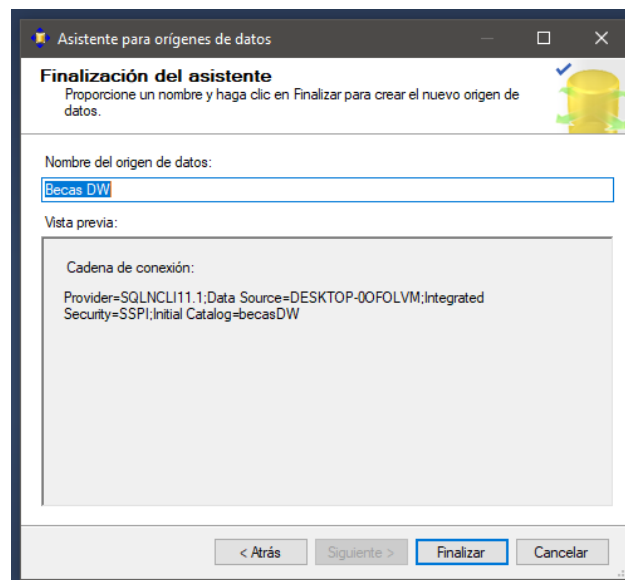


Ilustración 56: CUBO - Resumen de conexión. Elaboración propia

Se creará una vista del origen de datos de Becas DW, para ello damos clic derecho en Vistas del origen de datos e iniciamos una Nueva vista del origen de datos.

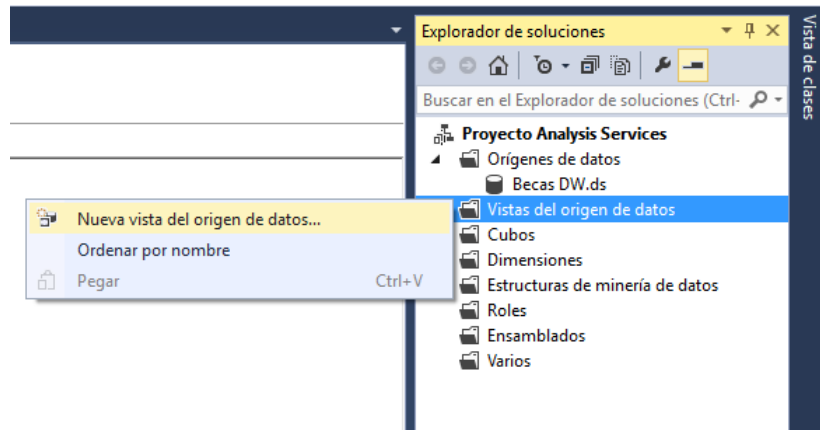


Ilustración 57: CUBO - Vistas del origen de datos. Elaboración propia

En el asistente para vistas del origen de datos seleccionamos Becas DW y damos en Siguiente.

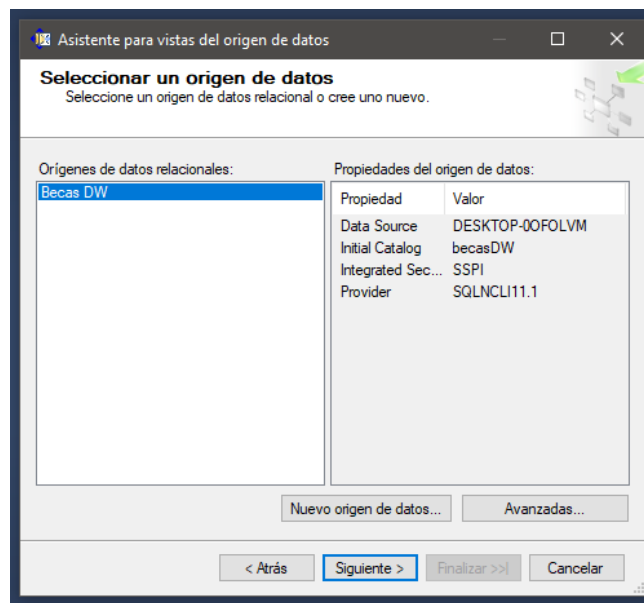


Ilustración 58: CUBO - Nueva vista de origen de datos. Elaboración propia

Incluimos todas las tablas (dimensiones y tabla de hechos) de lista de Objetos disponibles, con los cuales se va a trabajar el Cubo.

Una vez agregadas todas las tablas damos en siguiente.

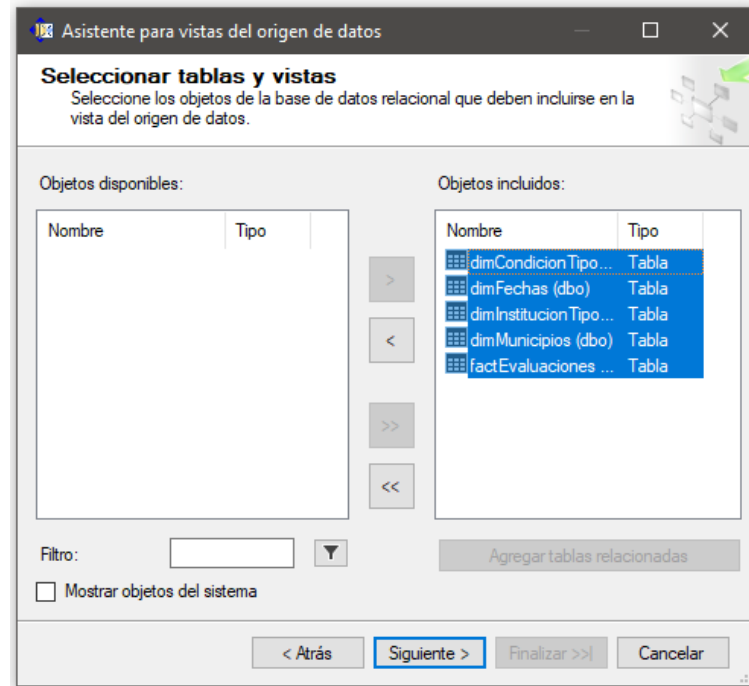


Ilustración 59: CUBO - Selección tablas y vistas. Elaboración propia

Y asignamos el mismo nombre que el origen de datos: Becas DW.

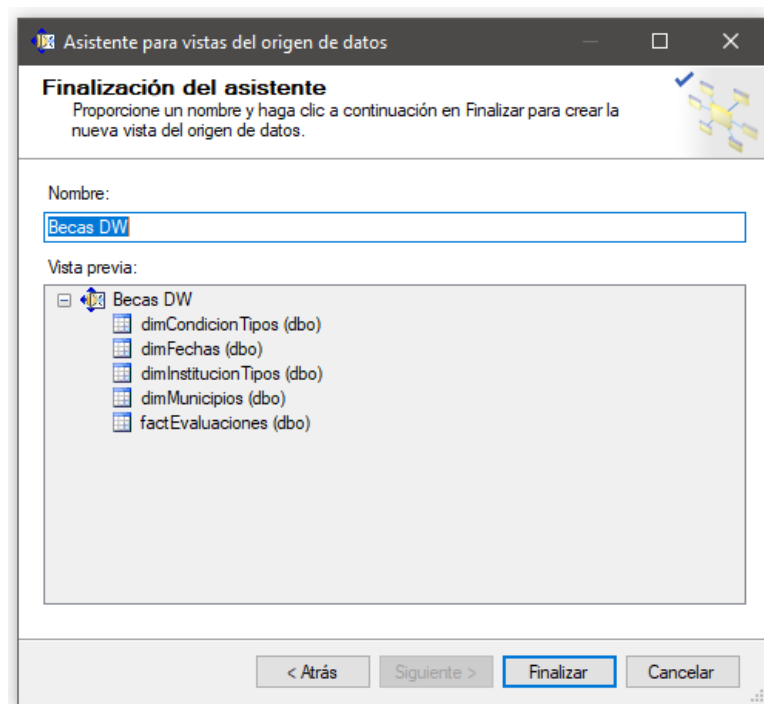


Ilustración 60: CUBO - Resumen de vista de origen de datos. Elaboración propia

Podremos observar el Diseño de la vista del origen de datos.

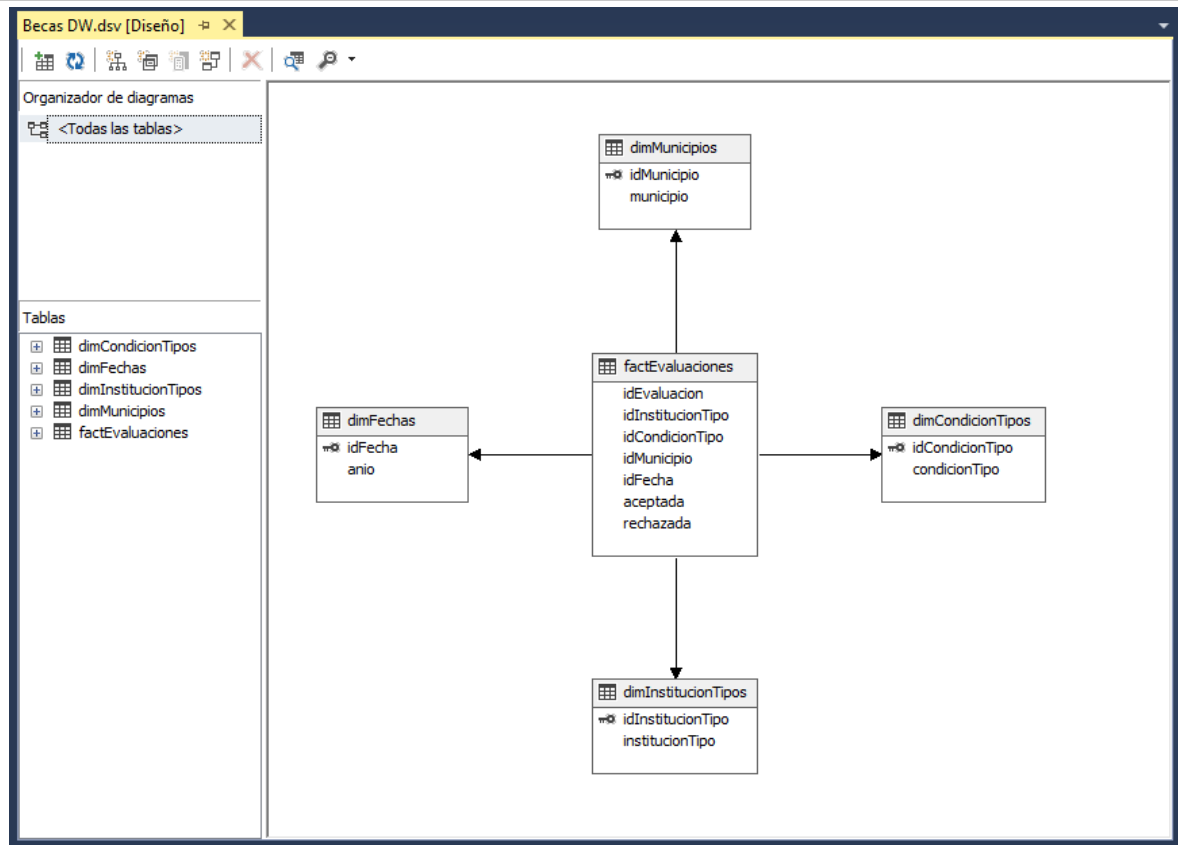


Ilustración 61: CUBO - Vista de origen de datos. Elaboración propia

Ahora procedemos a crear el Cubo, dando clic derecho en Cubos y seleccionamos en Nuevo cubo para instanciarlo.

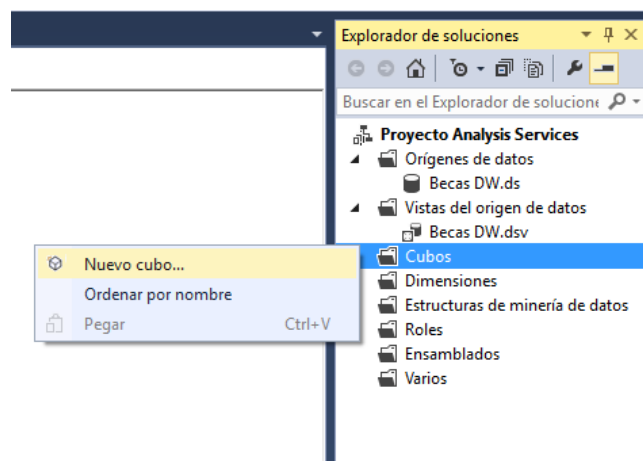


Ilustración 62: CUBO - Nuevo cubo. Elaboración propia

Nos llevará al asistente para cubos, en el cual debemos elegir “Usar tablas existentes” como método de creación para el cubo.

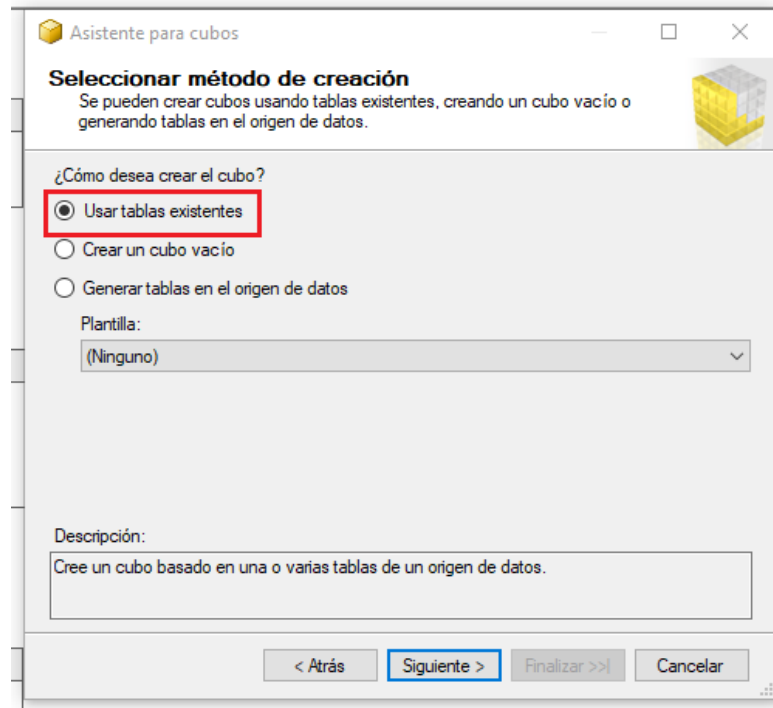


Ilustración 63: CUBO - Método de creación. Elaboración propia

Prosiguiendo con la creación del cubo seleccionamos la tabla factEvaluaciones como tabla de grupo de medida

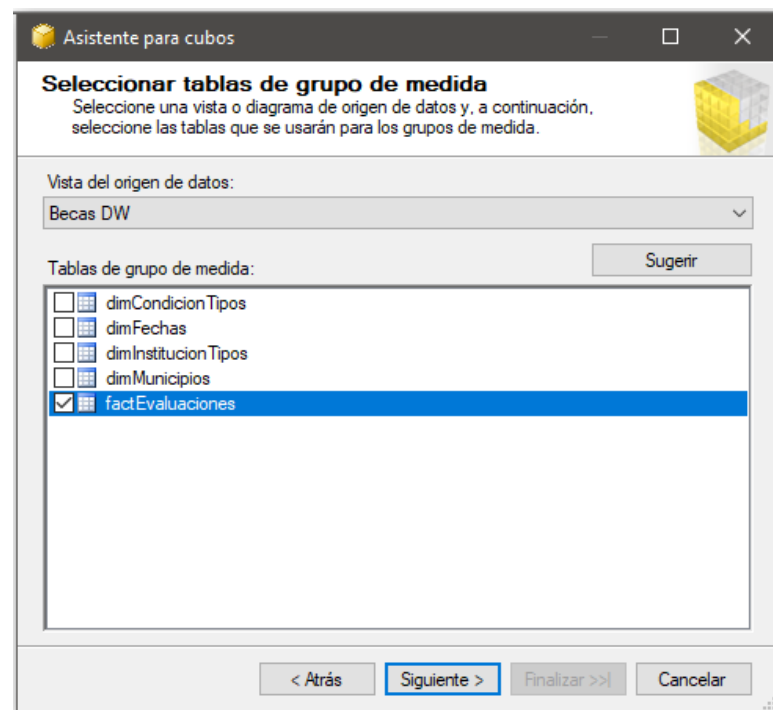


Ilustración 64: CUBO - Selección tablas de grupo de medida. Elaboración propia

De las cuales nos darán para elegir las medidas que se generan de la misma tabla factEvaluaciones: Aceptada, Rechazada y Recuento FactEvaluaciones.

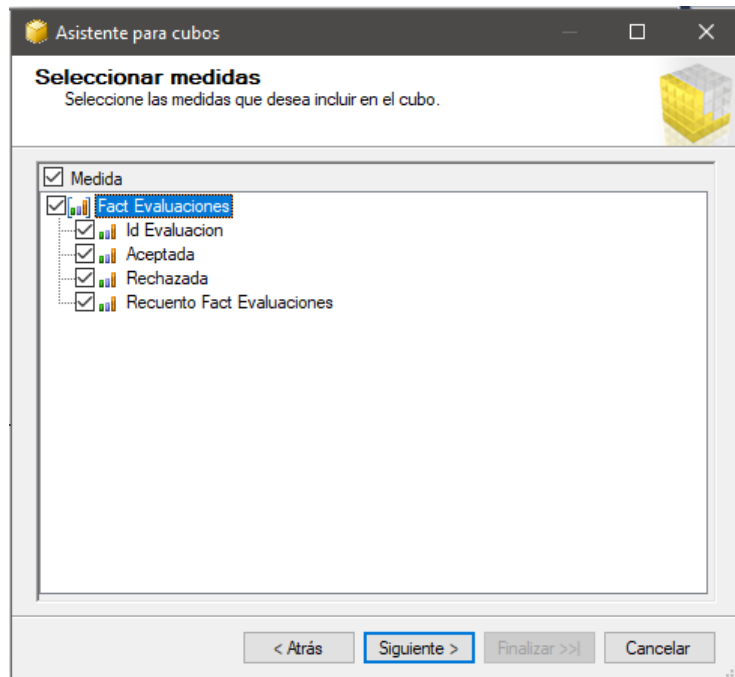


Ilustración 65: CUBO - Selección medidas. Elaboración propia

También seleccionamos las dimensiones del cubo que hacen referencia a las tablas de dimensiones del DW: dimMunicipios, dimInstitucionTipos, dimCondicionTipos y dimFechas.

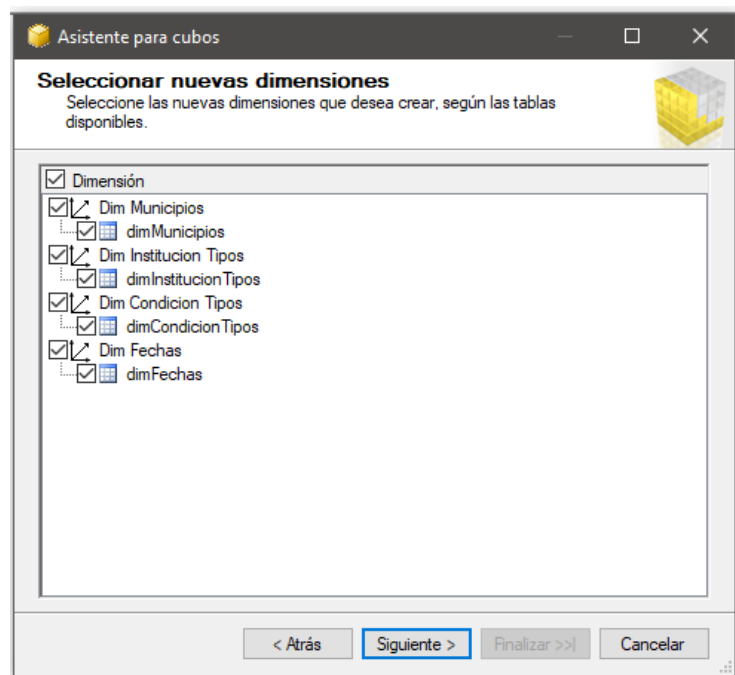


Ilustración 66: CUBO - Selección nuevas dimensiones. Elaboración propia

Se finalizará el armado de la estructura del cubo asignándole el nombre de CUBO BECAS.

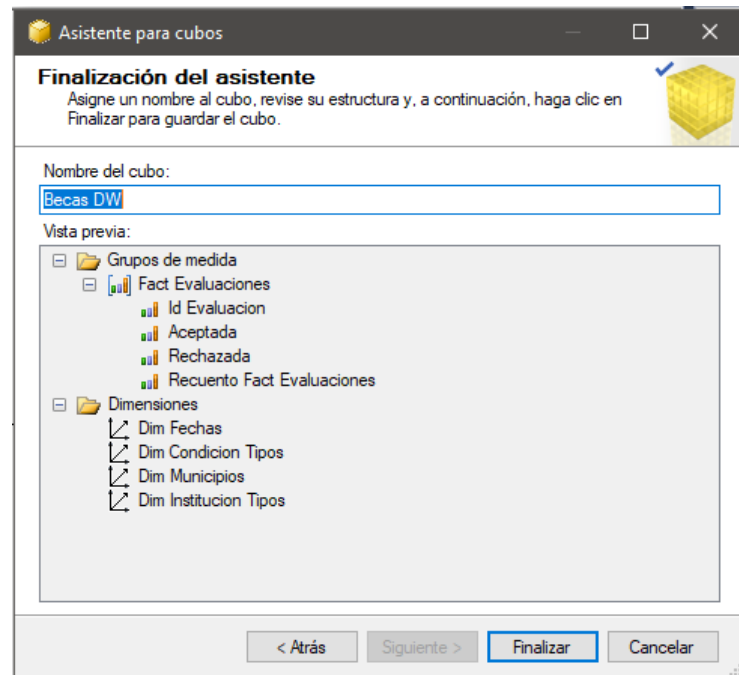


Ilustración 67: CUBO - Resumen del nuevo cubo. Elaboración propia

En la vista del Diseñador se podrá observar la estructura que tendrá el CUBO BECAS

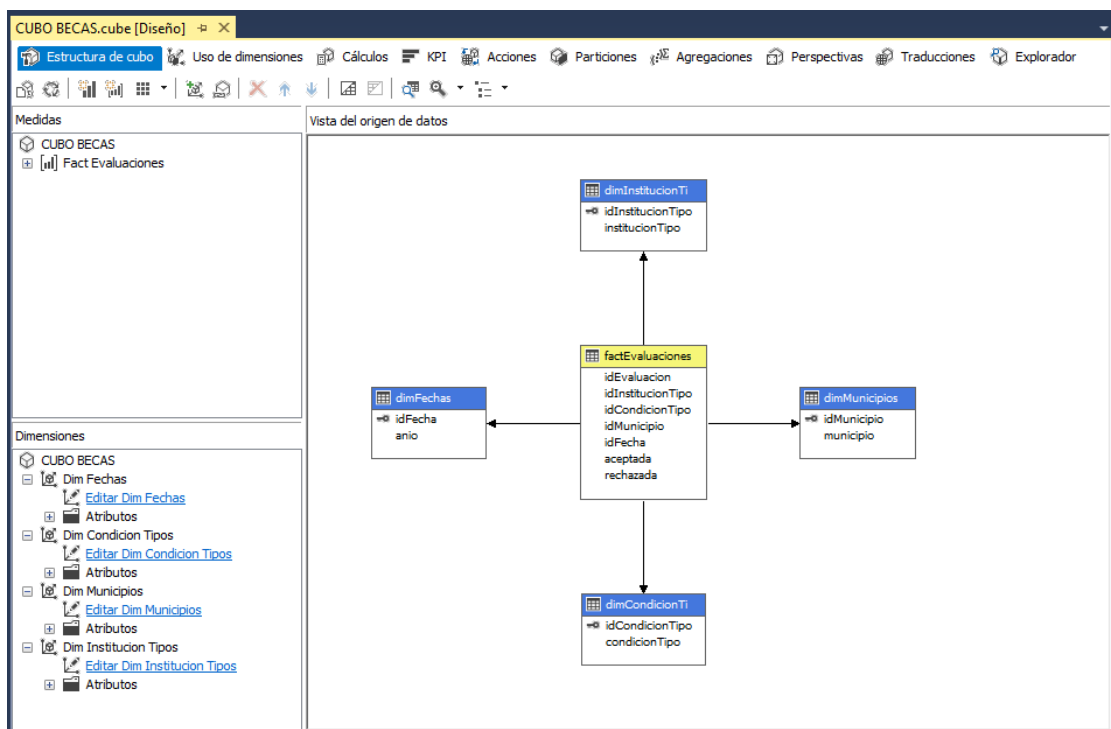


Ilustración 68: CUBO - Diseño de Cubo. Elaboración propia



Ahora seguimos a procesar los datos del DW para llenar las dimensiones y las medidas del Cubo, realizando clic derecho en el CUBO BECAS y seleccionamos en Proceso.

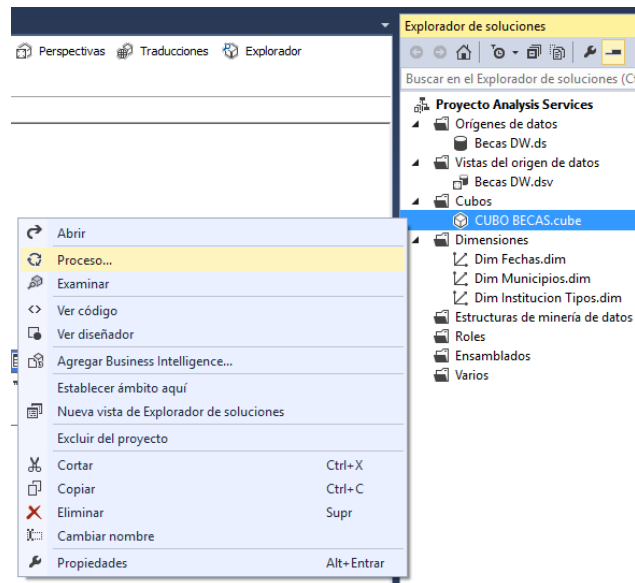


Ilustración 69: CUBO - Proceso. Elaboración propia

Aparecerá una ventana emergente indicándonos que el contenido del servidor ha caducado y le damos clic en Sí para generar e implementar el proyecto, esto sucede cada vez que se inicia un nuevo proyecto de Analysis Services.

Se muestra en otra ventana el CUBO BECAS a procesar en la lista de objetos que pertenecen al proyecto y procedemos a Ejecutar la carga del cubo.

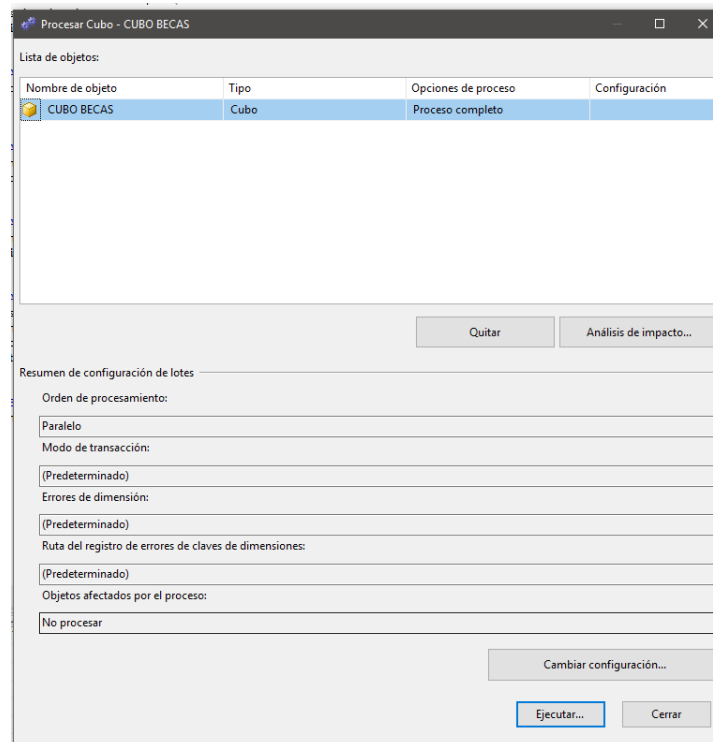


Ilustración 70: CUBO - Procesar cubo BECAS. Elaboración propia

Y como resultado nos mostrará el progreso del Cubo y que se ha finalizado correctamente.

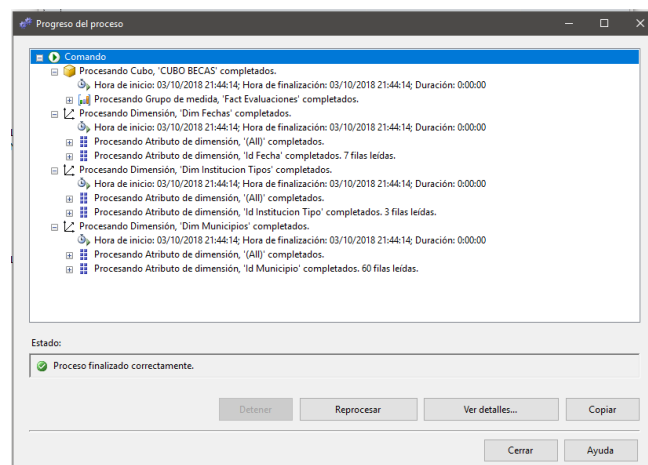


Ilustración 71: CUBO - Resultados. Elaboración propia

Al realizar el procesamiento del CUBO BECAS podremos observar en SQL Server Management Studio dentro del servidor de Analysis Services el proyecto creado anteriormente.

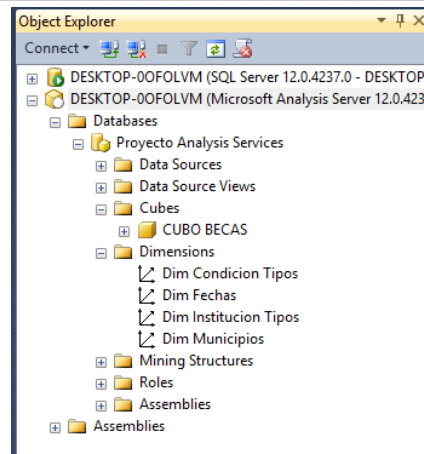


Ilustración 72: CUBO - SSMS. Elaboración propia

Con esto estará completo la creación del Cubo que posteriormente será utilizado para ser presentada a través de tablas y gráficos, que servirán como fuente de información para los usuarios e incursionar en la toma de decisiones.

Presentación Cubo²⁸

Una vez generado el cubo multidimensional es posible generar los reportes necesarios mediante la utilización de sus elementos. En este caso se generaron dos reportes y un dashboard a través de la utilización de la herramienta Power BI Desktop.



Ilustración 73: Reportes Generados. Elaboración propia

Para comenzar el desarrollo de los reportes y el dashboard se debe verificar que los servicios de Analysis Services estén activos que es donde se encuentra alojado el cubo multidimensional del cual se van a obtener los datos. Para ello debemos ingresar a los servicios de Windows simplemente apretando la tecla de Windows, escribir “Servicios” y acceder a la herramienta.

²⁸ Adjunto proyecto de Power BI en la carpeta **Presentación de CUBO**

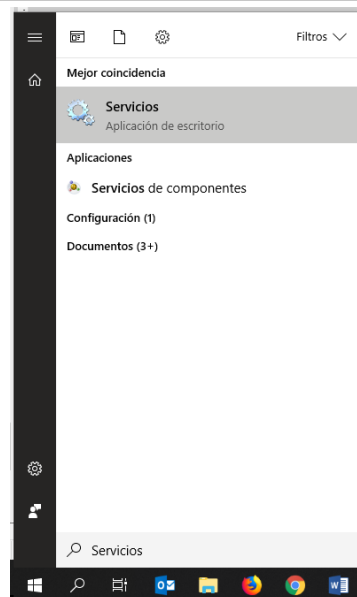


Ilustración 74: Servicios. Elaboración propia

Una vez en la herramienta Servicios debemos buscar la instancia con el nombre de SQL Server Analysis Services y si hay más de una instancia del mismo servicio elegir en el cual se encuentra el cubo, verificar su estado y si no se encuentra en ejecución, iniciar el mismo.

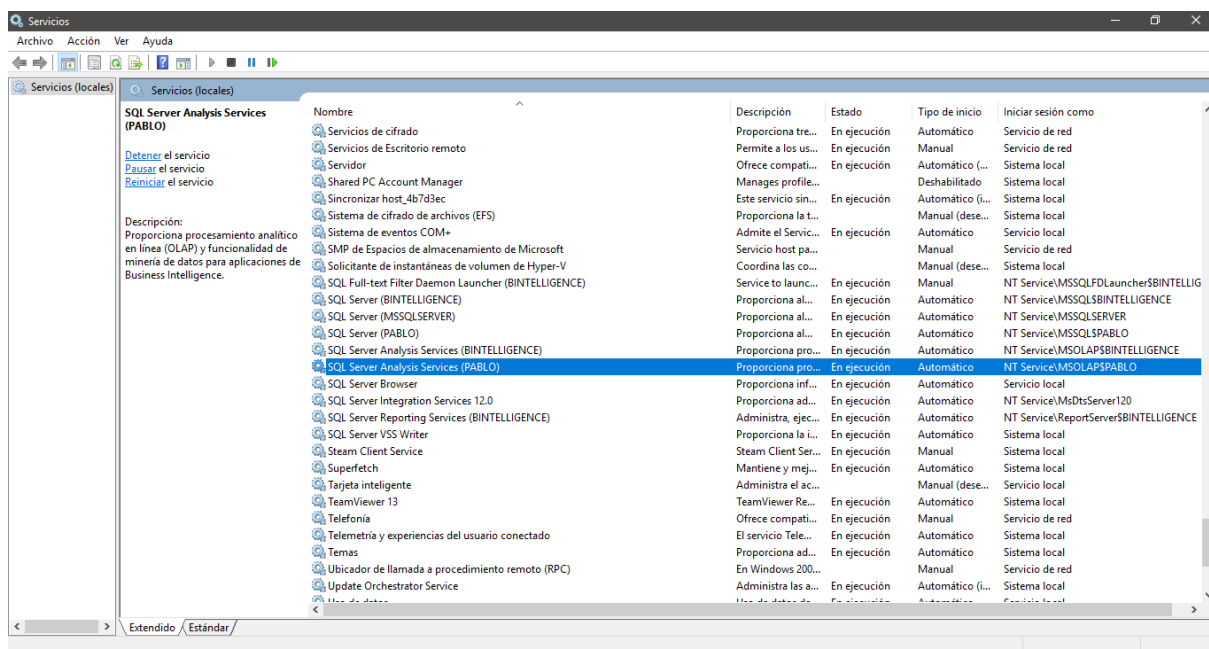


Ilustración 75: Estado de servicios de Analysis Services. Elaboración propia

Al iniciar el proyecto en Power BI Desktop debemos indicar cuál es el servidor con el que vamos a trabajar, con el fin de tener una conexión dinámica con el cubo.

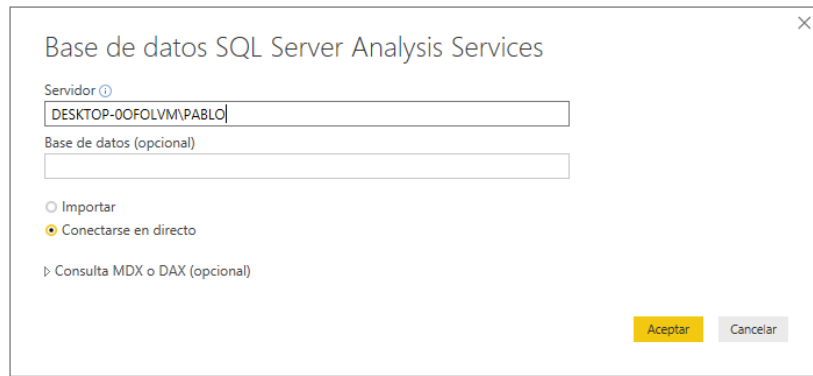


Ilustración 76: Conexión Base de datos SSAS. Elaboración propia

Al aceptar la conexión con SQL Server Analysis Services nos desplegará un navegador con los objetos que posee el mismo, en el cual debemos elegir el cubo con el que vamos a trabajar.

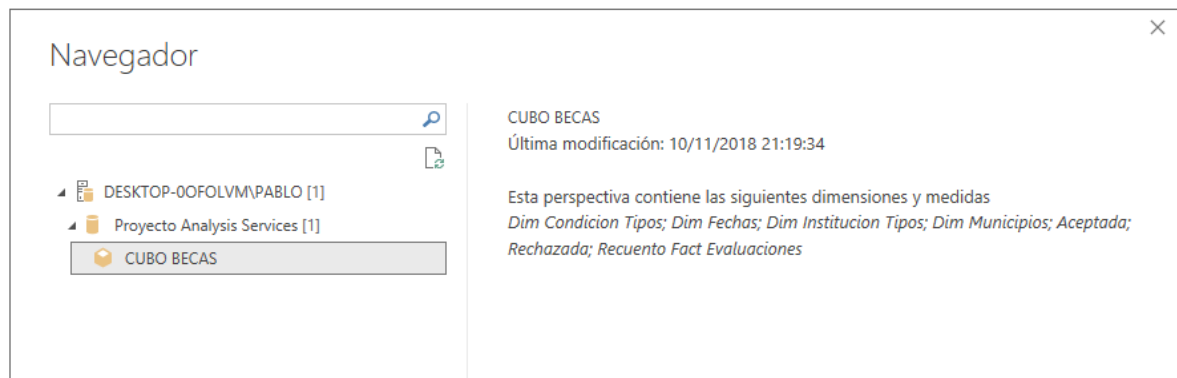


Ilustración 77: Elección CUBO. Elaboración propia

Uno de ellos muestra principalmente en que años, que municipios y cuáles son los tipos de instituciones que tienen la mayor cantidad de evaluaciones aceptadas, esto sirve de indicador para poder conocer porque ciertos municipios tienen un menor número de aceptación en comparación a otros y poder recabar información cuales son las razones de esto.



EVALUACIONES ACEPTADAS POR TIPO DE INSTITUCION DE CADA MUNICIPIO EN UN PERIODO DETERMINADO

Año	Nivel Superior Terciario	Nivel Superior Universitario	Total
2013	2	4	6
El Bordo	1		1
El Jardín	0		0
La Caldera	1		1
La Merced		0	0
La Poma		1	1
Rosario de La Frontera		1	1
San José de Metán		0	0
Santa Victoria Oeste		1	1
Secantás		0	0
Tolar Grande		1	1
2014	3	5	8
Cafayate		0	0
Colonia Santa Rosa		0	0
El Bordo	0		0
El Jardín	0		0
General Güemes		0	0
La Caldera	1		1
La Merced		1	1
La Poma		0	0
Los Toldos		0	0
Rivadavia Banda Norte	1		1
Rivadavia Banda Sur		0	0
Total	21	81	102

Visualización: Tabla de evaluaciones aceptadas por municipio y año. El panel de filtros a la derecha muestra que se han seleccionado 'Fact Evaluaciones' (Aceptada, Rechazada, Recuento Fact...), 'Dim Fechas' (Año, Id Fecha), 'Dim Institucion Tipos' (Institucion Tipo) y 'Dim Municipios' (Municipio).

Ilustración 78: Evaluaciones Aceptadas. Elaboración propia

Así mismo se generó un reporte para reconocer cuales son los requisitos reglamentados en las evaluaciones rechazadas que no llegan a ser cumplidos por la mayoría de personas solicitantes para obtener una beca. Esto ayudaría a reformular tal vez las medidas o las condiciones que se deben cumplir para conseguir el beneficio.

EVALUACIONES RECHAZADAS DE CADA MUNICIPIO DE POR TIPO DE CONDICIÓN QUE NO CUMPLEN POR PERIODO

Año	Adeuda materias	Carrera universitaria o terciaria finalizada	Es mayor a la edad permitida	Es pasante pago	Los ingresos del grupo familiar es mayor que el permitido	Materias aprobadas menor del sesenta y seis por ciento
2013	1	1	1	1		
El Bordo						
El Jardín		1				
La Caldera						
La Merced	1					
La Poma						
Rosario de La Frontera						
San José de Metán			1			
Santa Victoria Oeste						
Secantás				1		
Tolar Grande						
2014	2	1		1		1
Cafayate						
Colonia Santa Rosa						
El Bordo						
El Jardín						
General Güemes	1					
La Caldera						
La Merced						
La Poma			1			
Los Toldos						1
Rivadavia Banda Norte						
Rivadavia Banda Sur						
Rosario de La Frontera						
Santa Capital						
San José de Metán						
San Ramón de la Nueva Orán						
Santa Victoria Este						
Santa Victoria Oeste	1			1		
Secantás						
Tartagal						
Tolar Grande						
2015	1	2	1			2
Total	8	10	9	10		10

Visualización: Tabla de evaluaciones rechazadas por municipio y año, categorizadas por condición. El panel de filtros a la derecha muestra que se han seleccionado 'Fact Evaluaciones' (Rechazada, Recuento Fact...), 'Dim Fechas' (Año, Id Fecha), 'Dim Institucion Tipos' (Institucion Tipo) y 'Dim Municipios' (Municipio).

Ilustración 79: Evaluaciones Rechazadas. Elaboración propia

Contando con estos reportes se pudo generar un Dashboard en el cual se muestra los datos que se obtienen del cubo a través de distintos gráficos, aplicando ciertos filtros que permiten las dimensiones del mismo almacén multidimensional.

Los dashboards se pueden entender como una colección de reportes, consultas y análisis interactivos que hacen referencia a un tema en particular y que están relacionados entre sí.



UCS – Facultad de Ingeniería
Proyecto de Grado: Business Intelligence para Subprograma de Becas Provinciales
Capítulo 4: Solución Propuesta

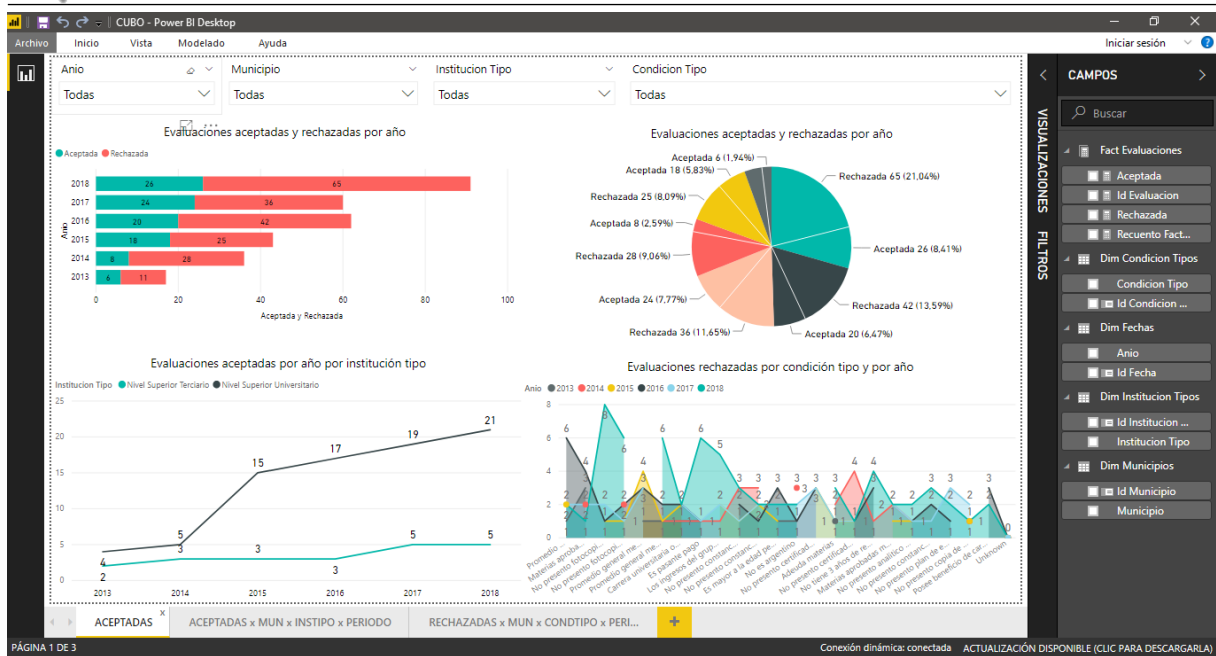


Ilustración 80: Dashboard. Elaboración propia

En este dashboard podemos filtrar los resultados a través de los campos de Año, Municipio, Institución Tipo y Condición Tipo.

Podemos observar en el siguiente ejemplo la aplicación de los siguientes filtros:

- Año: 2017
- Municipio: Salta Capital
- Institución Tipo: Nivel Superior Universitario
- Condición Tipo: Todas

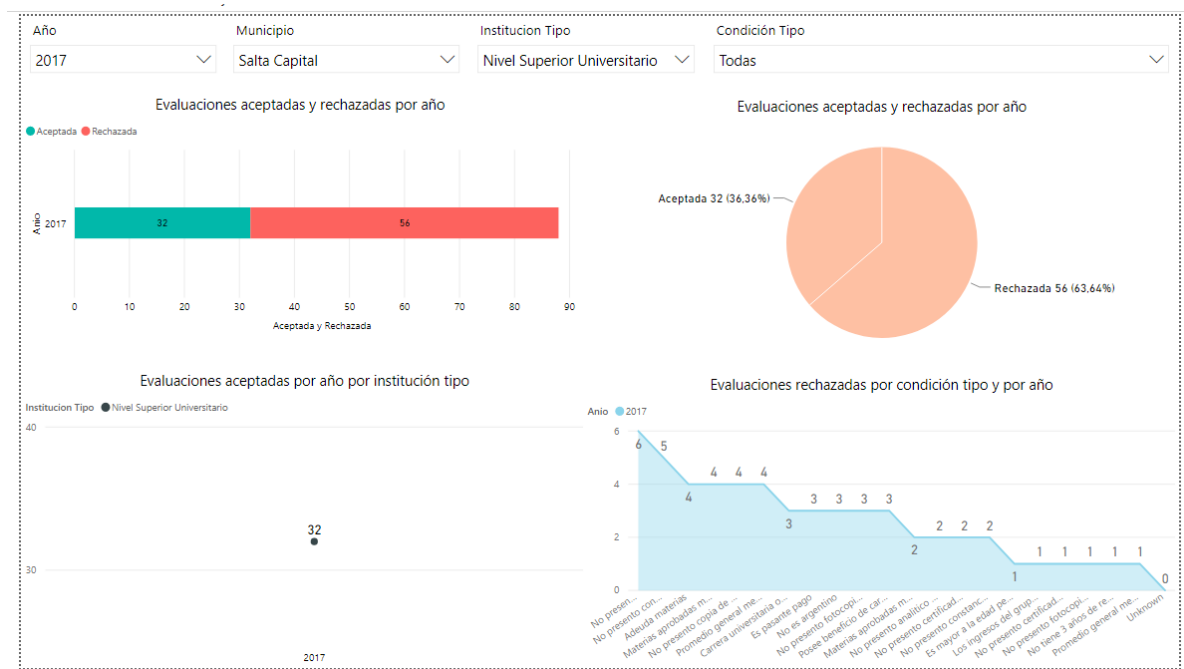


Ilustración 81: Dashboard con datos filtrados. Elaboración propia



Capítulo 5: Resultados

En la siguiente sección se comentarán los resultados obtenidos de este proyecto.

Generación de datos

A fin de poder analizar la solución, se decidió realizar pruebas para el prototipo planteado. Se utilizó datos de hasta 20000 inscripciones, de personas que pertenecen a uno de los 59 Municipios de Salta que tuvieron su respectiva evaluación en un rango de seis años 2013, 2014, 2015, 2016, 2017 y 2018. Los datos de estas inscripciones son datos de prueba y fueron generados de manera aleatoria para los propósitos de este trabajo y procesados para poder ser cargados en el Data Warehouse. Finalmente se obtuvo un Data Warehouse con información de seis años y se diseñó un CUBO a partir de este Data Warehouse, proporcionando de esta manera una plataforma de consultas para satisfacer las necesidades de los usuarios a partir de la generación de reportes y un Dashboard que servirá para la toma de decisiones.

En un principio se planteó utilizar un total de 6000 inscripciones por año dando un total de 36000 inscripciones, pero el paquete encargado de generar estos datos de prueba presenta un límite de 10000 iteraciones en la búsqueda de un número aleatorio distinto a los elegidos anteriormente, el problema se da cuando no quedan números aleatorios para elegir después de que se seleccionaron los primeros 10000 registros terminando el proceso de la generación de datos.

```
pablo@pablo-laptop /var/www/beca/beca-laravel
Archivo Editar Ver Buscar Terminal Ayuda

[OverflowException]
Maximum retries of 10000 reached without finding a unique value

Exception trace:
() at /var/www/beca/beca-laravel/vendor/fzaninotto/faker/src/Faker/UniqueGenerator.php:51
Faker\UniqueGenerator->call() at /var/www/beca/beca-laravel/database/seeds/DomiciliosTableSeeder.php:32
DomiciliosTableSeeder->run() at /var/www/beca/beca-laravel/vendor/laravel/framework/src/Illuminate/Database/Seeder.php:43
Illuminate\Database\Seeder->call() at /var/www/beca/beca-laravel/database/seeds/DatabaseSeeder.php:35
DatabaseSeeder->run() at /var/www/beca/beca-laravel/vendor/laravel/framework/src/Illuminate/Database/Console/Seeds/SeedCommand.php:63
Illuminate\Database\Console\Seeds\SeedCommand->Illuminate\Database\Console\Seed{s}{closure}() at /var/www/beca/beca-laravel/vendor/laravel/framework/src/Illuminate/Database/Eloquent/Model.php:2292
Illuminate\Database\Eloquent\Model::unguarded() at /var/www/beca/beca-laravel/vendor/laravel/framework/src/Illuminate/Database/Console/Seeds/SeedCommand.php:64
Illuminate\Database\Console\Seeds\SeedCommand->fire() at n/a:n/a
call user func array() at /var/www/beca/beca-laravel/vendor/laravel/framework/s
```

Ilustración 82: Generación de datos - problema. Elaboración propia

Se pasó de generar de la siguiente manera:

```
public function run()
{
    $dummy = Faker\Factory::create();
    for ($i=0; $i < 20000; $i++) {
        $k = rand(0, 1);
        if ($k == 0) {
            App\Modules\Inscripcion\Models\Inscripcion::create([
                'fecha' => '2018-03-15',
                'añoIngresoCarrera' => '2018',
                'constanciaInscripcionCarrera' => $dummy->boolean,
                'carreraDomiciliar' => $dummy->boolean,
                'añoEgresoPolimodal' => $dummy->numberBetween(2014, 2017),
                'promedioPolimodal' => $dummy->randomFloat(2, 6, 10),
                'analiticoPolimodal' => $dummy->boolean,
                'deudaPolimodal' => $dummy->boolean,
                'establecimientoProvincial' => $dummy->boolean,
                'solicitudBeneficioSimilar' => $dummy->boolean,
                'pasante' => $dummy->boolean,
                'hermanosBeneficiarios' => $dummy->boolean,
                'persona id' => $dummy->unique()->numberBetween(1, 20000),
                'aspiranteTipo id' => '1',
                'sucursalBancaria id' => '16',
                'planEstudio id' => $dummy->numberBetween(1, 1000),
            ]);
        }
    }
}
```

Ilustración 83: Seeder problema. Elaboración propia

A cambiar la aleatoriedad de algunos generadores por la dependencia hacia otras variables se pudo lograr el cometido sin tener más problemas:

```
public function run()
{
    $dummy = Faker\Factory::create();
    for ($i=0; $i < 20000; $i++) {
        $k = rand(0, 1);
        if ($k == 0) {
            App\Modules\Inscripcion\Models\Inscripcion::create([
                'fecha' => '2018-03-15',
                'añoIngresoCarrera' => '2018',
                'constanciaInscripcionCarrera' => $dummy->boolean,
                'carreraDomiciliar' => $dummy->boolean,
                'añoEgresoPolimodal' => $dummy->numberBetween(2014, 2017),
                'promedioPolimodal' => $dummy->randomFloat(2, 6, 10),
                'analiticoPolimodal' => $dummy->boolean,
                'deudaPolimodal' => $dummy->boolean,
                'establecimientoProvincial' => $dummy->boolean,
                'solicitudBeneficioSimilar' => $dummy->boolean,
                'pasante' => $dummy->boolean,
                'hermanosBeneficiarios' => $dummy->boolean,
                'persona id' => $i+1,
                'aspiranteTipo id' => '1',
                'sucursalBancaria id' => '16',
                'planEstudio id' => $dummy->numberBetween(1, 1000),
            ]);
        }
    }
}
```

Ilustración 84: Seeder solución. Elaboración propia

Resultados de las pruebas

Para la validación del prototipo con usuarios se realizó una reunión en el cual se explicó la finalidad del proyecto y las acciones que éstos debían ejecutar en el mismo.

Los usuarios participantes en las pruebas del sistema evaluaron de forma positiva lo siguiente:

- Facilidad de leer y de entender tablas y gráficas.
- Rapidez para obtener resultados.



- Para la carga del Data Warehouse, la no alteración de la base de datos que actúa como fuente.

Junto con esto indicaron las siguientes falencias o dificultades:

- Selección de datos históricos incómoda, debido al ingreso de valores por año para todos los gráficos, siendo más intuitivo poder seleccionar por rangos de fecha.
- La espera de un día para poder visualizar nuevas modificaciones en los gráficos planteados en la presentación del cubo.
- Falta de conocimiento sobre el uso del programa Power BI Desktop, tener que aprender a usarlo para generar nuevos gráficos que hagan referencias a nuevos indicadores.

También propusieron los siguientes cambios al sistema:

- Presentar estos gráficos en su página web, que se encuentre disponible para los usuarios que trabajan en ella.
- Agregar nuevos indicadores orientados a obtener más datos relevantes de las personas que se postulan en la misma.

Las personas que participaron de la prueba del sistema presentan diversas características, dos personas que están en la parte de Servicio Social, otros dos administrativos y el coordinador del programa. El usuario menos avanzado, fue uno de los administrativos, el cual posee un nivel básico de Windows, Excel y Word. Otro de los participantes, posee un nivel avanzado en estadísticas, por el cual le resultó de entender perfectamente el Dashboard, al instante de visualizar los gráficos. El coordinador trabajó en otras ocasiones con sistemas que manejaban grandes cantidades de reportes y gráficos, por ejemplo: Tango Gestión. Los otros dos usuarios que pertenecen a servicio social, también con nivel básico de herramientas de oficina, expresaron su contento al obtener este tipo de reportes, encargados de analizar aquellas evaluaciones que no son aprobadas.

El sistema demostró permitir a sus usuarios realizar menos tareas de recopilación de información lo que les significó una disminución de tiempos muertos, además gracias al registro de los datos en el Data Warehouse, se pudo automatizar este proceso y brindar datos históricos.



Capítulo 6: Conclusiones

El presente capítulo resume los objetivos que se han logrado y que han sido descritos en el presente documento. A continuación de éstos, se exponen ciertas líneas futuras que extenderán los resultados obtenidos aquí presentados.

Conclusiones

Durante la elaboración de este proyecto se conoció que las organizaciones en la actualidad, sin importar su tamaño, conservan grandes cantidades de datos generados por los sistemas de información que permiten llevar adelante sus tareas de rutina.

Esta cantidad de información que resulta invaluable para los encargados de tomar decisiones, llevó a plantear una de las problemáticas que suele presentarse a los directivos de las empresas: la de pasar más tiempo recolectando información que gestionándola.

A partir de esa problemática, se demostró como la aplicación de una solución BI permite mejorar sus operaciones.

El valor de la información como soporte de la gestión de la calidad no reside en la información de la misma, si no en la capacidad de análisis que tengamos sobre ésta. La gestión de la calidad surge de la capacidad de tomar las decisiones adecuadas basadas en información útil, oportuna y correcta.

Las fuentes de información que soportan las tareas operacionales de una empresa no soportan los procesos de toma de decisiones, pues no están diseñadas con ese objetivo. Las organizaciones deben introducir nuevas tecnologías de información que les permitan integrar sus fuentes de información operacionales en sistemas de información orientados a la gestión.

Las organizaciones públicas no son distintas al resto de las organizaciones, generan un volumen de información operacional menor en el caso de este programa, pero su necesidad de gestión de calidad es igual; sin embargo, en general, las organizaciones públicas desconocen o incluso a veces menosprecian el valor agregado que genera el manejo correcto de la información sobre sus procesos de toma de decisiones, y por lo tanto no realizan los esfuerzos necesarios para mejorar sus procesos de toma de decisiones mediante la introducción de nuevas tecnologías de información.

Durante este trabajo se observó que los principales factores que impiden a las organizaciones públicas mejorar la calidad de su gestión mediante la introducción de nuevas tecnologías de información son:

- **Confidencialidad de la información:** Las organizaciones públicas no pueden revelar su información que se encuentran alojadas en su base de datos por la ley de protección de datos personales de las personas que llenaron un formulario y presentaron su solicitud. Para muchas, éste es su principal activo y no están dispuestas a compartirlo con asesores externos.
- **Visión estrecha:** En general los mandos medios y superiores en las organizaciones públicas no visualizan el impacto que produce la mejora en la calidad de la gestión producto de la introducción de nuevas tecnologías de información. Las



organizaciones públicas no consideran que la introducción de nuevas tecnologías sea crítica para el crecimiento de su empresa y son reacios a abordar proyectos de mediano a largo plazo. Tal vez por desconocimiento de las mismas.

- **Alto costo de la tecnología:** Sumando a lo anterior los productos de software asociados a la gestión tienen un alto costo de licenciamiento, básicamente porque están enfocados a otro segmento de empresas.

La metodología propuesta abarca todos los puntos conocidos de la inteligencia de negocios para ser aplicados, engloba aspectos de negocio, un poco de administración de proyectos y los tres puntos modulares de la inteligencia de negocios que son la construcción de un Data Warehouse, la construcción de los procesos ETL que permiten poblar el Data Warehouse, los procesos de análisis como son cubos de información y la explotación de dicha información por medio de un Dashboard, los cuales permitirán a los usuarios finales tomar las decisiones necesarias para alcanzar los objetivos establecidos durante la familiarización con la organización.

La utilización de la Metodología HEFESTO, permitió identificar fácilmente los resultados alcanzados, los cuales son sencillos de comprender, con lo que se determinó las necesidades de la organización, basándose en los requerimientos del usuario, involucrándolo así en cada etapa para la toma de decisiones oportunas ante los cambios de la organización, gracias a su estructura adaptable.

El desarrollo de soluciones de inteligencia de negocios con herramientas open Source puede implicar un proceso complicado con una curva de aprendizaje bastante amplia, además de no existir toda la documentación que uno quisiera y la que existe no siempre suele ser tan clara. Es por eso que a mi consideración es importante que se aproveche que las empresas que ofrecen herramientas propietarias estén viendo a las organizaciones como un mercado potencial, abriéndoles las puertas al mundo de la toma de decisiones ofreciendo versiones de sus herramientas de manera gratuita. De esta manera se puede aprovechar la estabilidad que ofrece una herramienta propietaria en comparación de una herramienta libre.

No es necesario estar atado a una sola herramienta, en el desarrollo de este trabajo se utilizaron en conjunto una herramienta libre como LARAVEL para la generación de datos que utiliza como base de datos a MySQL, SQL Server como motor de base de datos para utilizarlo como almacén del DW y del cubo, por último, con una versión gratuita de Power BI de Windows.

Con la utilización de estas herramientas se consiguió que gran parte del volumen de datos que podría llegar a poseer la organización, se presenten a través de reportes dinámicos que reflejan solamente la información indispensable de una manera ágil y versátil, por ejemplo:

- Cual fue el municipio con la mayor tasa de evaluaciones aceptadas durante el último año.
- Cuál o cuáles son las razones por las que se rechazan la mayoría de las evaluaciones.
- Durante qué año se produjo un mayor grado de inscripciones ya sea con evaluaciones aceptadas o rechazadas.

Mediante el acceso fácil y directo a la información, los miembros de las organizaciones pueden encontrar respuestas a todas las preguntas que surjan en el ejercicio de sus actividades.

Estos resultados demuestran que es factible la aplicación de una solución de BI orientadas a este tipo de organizaciones. La idea principal es que las organizaciones públicas



no inviertan en soluciones que pudieran estar muy alejadas a las necesidades de la misma, si no en buscar la solución que mejor se adapte tanto a sus requerimientos, como a sus necesidades y recursos.

Para la instalación de las herramientas necesarias para desarrollar la solución BI, se investigó el funcionamiento de la herramienta SSDT BI, SSIS, SSAS, SSMS, Power BI, en base a lectura de papers, búsquedas en línea y manuales, para configurar de una manera adecuada en el entorno de trabajo para el desarrollo de la aplicación.

Estas herramientas son intuitivas, de fácil manejo, multiplataforma, sencilla al momento de integrar los datos.

El proceso de extracción, transformación y carga de datos son los apropiados siempre y cuando estén basados en la información requerida por los involucrados en la toma de decisiones.

El uso de la interfaz de Power BI, permite un manejo intuitivo y sencillo a los usuarios finales en la generación de sus reportes y análisis acorde a las necesidades de la organización.

Futuras líneas de Investigación

El trabajo descrito en este documento abre todo un conjunto de ideas sobre las cuales se puede continuar la investigación. Algunas de ellas no han sido abordadas por estar fuera de los objetivos inicialmente planteados, y otras han aparecido como consecuencia de las aportaciones aquí realizadas. A continuación, se presentan algunas de las líneas que considere más importante de cara a futuras aportaciones.

El proyecto no realiza un trabajo de minería de datos, lo que podría ser importante para la búsqueda de patrones en las bases de datos de estudiantes, teniendo un enfoque más amplio, como utilizar información de estudiantes de diferentes instituciones educativas, para así llevar la búsqueda de buenas prácticas en educación a un universo con más opciones.

El proyecto puede ser ampliado tomando en cuenta otros factores que influyen en la obtención o el poder mantener el beneficio más de un año de los estudiantes, como información de las familias de los estudiantes.

Profundizar en el proyecto añadiendo otras dimensiones que permitiese diseñar y crear nuevos cubos para así aumentar la capacidad de toma de decisiones en la organización. Se podría agregar como dimensión las carreras que ofrecen las instituciones terciarias y universitarias para tener conocimiento sobre cuales son aquellas con el mayor número de solicitantes.

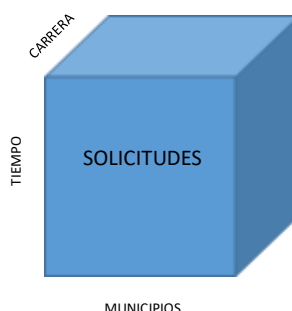


Ilustración 85: CUBO con Dimensión CARRERA

Otra dimensión podría ser las personas en sí mismas para analizar cuál es el promedio de edades de aquellas que buscan obtener una beca.

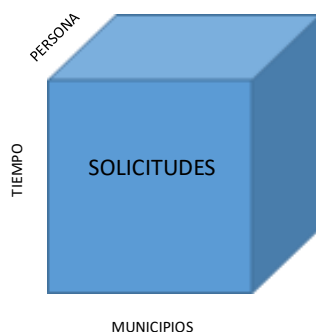


Ilustración 86: CUBO con Dimensión PERSONA

Recomendaciones

Para la obtención y comprensión de la lógica de la organización se debe mantener una relación estrecha con el usuario en cada una de las fases que comprende la metodología aplicada, para obtener la información necesaria e identificar los puntos clave para el desarrollo de la solución.

Se debería mantener reuniones constantes con los custodios de las bases de datos, para definir claramente los significados de los campos necesarios para obtener los indicadores y así precisar que tablas contienen los datos requeridos para el análisis, puesto que no se pudo contar con la base de datos ni con la estructura de la misma, se procedió a construir una base de datos con la información reunida que se obtuvo de otras fuentes, por lo que es necesario verificar que los datos entregados sean los requeridos para realizar el análisis antes de ejecutar la carga inicial.

Al inicio del proyecto es necesario definir los recursos que se dispondrán para preparar el ambiente sobre el cual se construirá la solución, en el cual se instalaran todos los utilitarios necesarios y se hará uso de la base de datos existente.



Referencias Bibliográficas

- Bernabeu Dario R. Y García Mattío Mariano. 2017. HEFESTO DATA WAREHOUSING. Guía completa de aplicación teórico-práctica; metodología Data Warehouse. 182 Páginas.
- Parra Salcedo Octavio J.; Galeano R. M. Y Rodríguez B. Luis. G. 2010. Metodología crisp para la implementación Data Warehouse. Revista Tecnura. Volumen 14. Número 26. Páginas 35 – 48.
- Prat Iglesias Lorenzo R. 2015. El Business Intelligence aplicado al Ministerio de Defensa. Revista Atlántica de Economía. Volumen 2. Páginas 1 – 47.
- Rivadera Gustavo R. 2010. La metodología de Kimball para el diseño de almacenes de datos (Data warehouse). U Católica de Salta. Páginas 16.
- Cano Josep L. 2007. Business Intelligence: Competir con información. 393 Páginas. Banesto, Fundación Cultural [i.e. Cultural].
- Stair Ralph M. Y Reynolds George W. 1999. Principios de sistemas de información: Un enfoque administrativo. 704 páginas.
- Laudon Kenneth C. Y Laudon Jane P. 1997. Sistemas de información Gerencial. 684 Páginas.
- Kimball Ralph Y Inmon William H. 2013. The Data Warehouse Toolkit. The Definitive Guide to Dimensional Modeling. 600 Páginas.
- Webb Chris. 2017. Expert Cube Development with SSAS Multidimensional Models. 402 Páginas.
- Ferrari Alberto. 2015. The Definitive Guide to DAX. Business Intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI. 592 Páginas.
- Russo Marco. 2017. Tabular Modeling in Microsoft SQL Server Analysis Services. 512 Páginas.
- Codd Edgar F. 1970. A relational Model of Data for Large Shared Data Banks. Revista Communications of the ACM. Volumen 13. Páginas 377 – 387.
- Turban Efraim Y Aronson Jay E. 2001. Decision Support Systems And Intelligent Systems. 867 Páginas.
- Inmon William H. 2005. Building The Data Warehouse. 576 Páginas.
- Alberto Duchi. 2017. Generalidades de la inteligencia de negocios. 14 Páginas.
- Gomez R. Alveiro A. Y Rico B. Dewar Willmer. 2010. Inteligencia de Negocios: Estado del Arte. Revista Scientia et Technica Año XVI. Número 44. Páginas 321 – 326. ISSN 0122-1701.
- Davenport Thomas Y Jarvenpaa Sirkka L. 2008. Strategic Use of Analytics in Government. Introducción. Managing for Performance and Results Series. Páginas 42.
- Zorrilla Marta. 2010/11. Introducción al Business Intelligence [Diapositivas]. España: Universidad de Cantabria.



- ITMadrid. 2008. Qué es inteligencia de negocios (Business Intelligence). <http://www.itmadrid.com/que-es-inteligencia-de-negocios-business-intelligence/>. Página vigente al 11/08/2018.
- Hernández J. 2018. Minería de Datos. <http://www.estadistico.com/arts.html?20040426>. Página vigente al 01/09/2018.
- Castro B. Yudi; Leonard B. Eric I. Y Soler P. Yolanda. 2014. Almacén de Datos para la Gestión Energética en el Ministerio de la Agricultura. Revista de Arquitectura e Ingeniería. Volumen 8. Número 3. Páginas 1 – 13.
- Gómez Rosado; Alonso Alveiro; Bautista Rico Y Willmer Dewar. 2010. Inteligencia de Negocios. Revista Scientia et Technica Año XVI. Número 44. Páginas 321 – 326.
- Watson Hugh J. 2006. Data Warehousing Stages of Growth. Revista Information Systems Management. Volumen 18. Páginas 42 – 50.
- Watson Hugh J. 2002. Recent Developments in Data Warehousing. Revista Communications of the Association for Information Systems. Volumen 8. Artículo 1.
- Sitios web:
 - [https://www.wikipedia.org/] Wikipedia.
 - [http://inteligenciadenegociosval.blogspot.com/] Inteligencia de Negocios.
 - [http://dbi.io/es/blog/] Data Business Intelligence.
 - [https://analyticboard.com/] Evolución Inteligencia de Negocios.
 - [https://blogs.imf-formacion.com/blog/tecnologia/] Blog de Tecnología de IMF Business School.
 - [https://www.beeva.com] Estrategia – Negocio.
 - [http://www.mercado.com.ar/] Revista Mercado.
 - [https://saimasolutions.com/]
 - [http://www.dataprix.com/]
 - [https://www.microstrategy.com/]
 - [https://www.ibm.com/]
 - [http://community.jaspersoft.com/]
 - [http://www.pentaho.com/]
 - [https://powerbi.microsoft.com/]
 - [https://blog.bi-geek.com/]
 - [https://www.businessintelligence.info/] Business Intelligence Fácil – Urquizu Pau.
 - [http://albertoduchi.info/category/bi/] Business Intelligence – Ing. Duchi Alberto
 - [https://www.gartner.com/]
 - [http://www.diegocalvo.es/]
 - [https://www.researchgate.net/]
 - [https://www.esan.edu.pe/apuntes-empresariales/] Business School
 - [https://es.workmeter.com/blog] Como hacer crecer tú negocio.



- [<https://www.gestiopolis.com/inteligencia-de-negocios/>] Inteligencia de Negocios – Ing. Carmen Zarate Gallardo.
- [<http://www.oracle.com/>]
- [<https://www.postgresql.org/>]
- [<https://www.debian.org/>]
- [<https://www.microsoft.com/>]
- [<https://www.mysql.com/>]
- [<https://blog.powerdata.es>] El valor de la gestión de datos.
- [<https://github.com/fzaninotto/Faker>] Faker.
- [<https://larevel.com>]
- [<https://geeks.ms/>]
- [<http://www.copaipa.org.ar/informatica/>] Copaipa Honorarios Informática.
- [<https://www.sqlshack.com/>]
- [<http://php.net/>]



Anexos

Anexo I: Reglamento programa de Becas

REGLAMENTO GENERAL DE BECAS DE ESTUDIOS

NIVEL SUPERIOR UNIVERSITARIO Y NO UNIVERSITARIO

CAPITULO I

DISPOSICIONES GENERALES

Artículo 1º.- Objetivos del Subprograma de Becas: El Subprograma de Becas tiene por objetivo: a) promover la igualdad de oportunidades en el ámbito de la educación superior, a través de la implementación de un sistema de becas que favorezca el acceso, permanencia y egreso de estudiantes con particulares condiciones de vulnerabilidad socioeconómica familiar y que presenten un rendimiento académico y regularidad en sus estudios en los términos exigidos por el presente ; b) fomentar la promoción efectiva del nivel medio o secundario, y c) contribuir al desarrollo de la provincia, promoviendo la formación de recursos humanos.-

Artículo 2º.- Becas. Denominación. Alcance: Denomínese BECA al apoyo económico anual que brinda el Gobierno de la Provincia de Salta, a estudiantes que cursan regularmente carreras de grado de Nivel Superior Universitario, dictada por Universidades Publicas y Privadas del Territorio de la República Argentina y de Nivel Superior No Universitario, dictadas por Institutos dependientes de las Direcciones Generales de Educación Superior y de Educación Privada del Ministerio de Educación, Ciencia y Tecnología del Gobierno de la Provincia de Salta.-

Artículo 3º.- Becas, caracteres: La Beca tiene los siguientes caracteres:

- a) Personal.-
- b) Intransferible.-
- c) Incompatible con cualquier otro beneficio de carácter similar.-

Artículo 4º- Reglamentación: Los montos de las becas, números de cuotas en que se paguen, los periodos de convocatoria, la documentación a presentar por los postulantes, el sistema de ponderación socio-económico, los convenios de becas individuales y toda otra cuestión no



prevista expresamente por la presente Resolución, será determinada mediante Instrumento Legal emitido por la Autoridad inmediata Superior del Ministerio de Educación, Ciencia y Tecnología de la Provincia de Salta.-

CAPITULO II

CONDICIONES Y REQUISITOS

Artículo 5º.- Condiciones personales: Todos los solicitantes deberán acreditar que cumple con las siguientes condiciones:

- a) Ser argentino.-
- b) No ser mayor de veinticinco (25) años de edad en caso de ser ingresante, ni mayor de treinta (30) en caso de ser reinscripto a carreras de Nivel Superior. En todos los casos la edad establecida deberá acreditarse al cierre de la convocatoria anual del Subprograma.-

Por alumnos reinscriptos se entiende aquellos que al momento de la inscripción ya han cursado un año de carrera mínimamente.

- c) Tres (3) años de residencia inmediatos anteriores a la postulación en la Provincia de Salta del solicitante y su grupo familiar.
- d) No poseer otro beneficio de carácter similar, ni ser pasante pago en dependencias de la Administración Pública Nacional, Provincial y/o Municipal, o en entidades de carácter privado.-

Artículo 6º.- Requisitos académicos: Para solicitar una Beca, el estudiante que curse una carrera de Nivel Superior en alguna de las Instituciones establecidas en el Art. 2º, deberá cumplimentar los requisitos académicos correspondientes a las siguientes situaciones:

1) Ingresantes a carreras de Nivel Superior:

- a) Haber cursado y concluido el Nivel Medio, Polimodal o Secundario, en cualquiera de sus modalidades, en establecimientos de la Provincia de Salta, sin adeudar ninguna materia.-
- b) Haber obtenido un promedio general de calificaciones mínimo de siete (07) puntos.
- c) Estar inscripto como alumno regular en una carrera de Nivel Superior y haber superado satisfactoriamente las condiciones de admisión a la Institución a la cual pretenda ingresar.-

2) Reinscriptos en carreras de Nivel Superior Universitario:



a) Tener un promedio general de calificaciones equivalente a cinco (5) computándose aplazos.-

b) Haber aprobado un mínimo de materias equivalentes al sesenta y seis (66 %) del total de materias que debería haber aprobado de acuerdo al Plan de Estudio, computadas desde el año de ingreso a la carrera hasta el inmediato anterior a la postulación.-

c) No haber finalizado una carrera de grado de Nivel Superior Universitario o No Universitario. Podrán postularse cuando el título obtenido sea intermedio y la nueva carrera implique una continuidad y unicidad en la formación académica.

d) Cursar regularmente la carrera para la que postuló, sea en modalidad presencial o semi-presencial.-

3) Reinscriptos en carreras de Nivel Superior No Universitario:

a) Tener un promedio general de calificaciones mínimo de seis (6) puntos incluyendo aplazos.

b) Haber aprobado un mínimo de materias equivalentes al setenta y cinco por ciento (75 %) de materias, de acuerdo al Plan de Estudio y computadas desde el año de ingreso a la carrera hasta el inmediato anterior a la postulación.-

c) No haber finalizado una carrera de grado de Nivel Superior Universitario o No Universitario. Podrán postularse cuando el título obtenido sea intermedio y la nueva carrera implique una continuidad y unicidad en la formación académica.

d) Cursar regularmente la carrera para la que postuló, sea en modalidad presencial o semi-presencial.-

Artículo 7º.- Condiciones socioeconómicas: El Solicitante deberá acreditar que su grupo familiar conviviente presenta situación socio-económica insuficiente para afrontar los gastos que demandan sus estudios. Las variables e indicadores de tal situación serán previamente determinadas/os y ponderadas/os por el Equipo Profesional Interdisciplinario del Subprograma de Becas a los efectos de evaluar el cumplimiento de dicha condición, y será objeto de Instrumento Legal emitido por la Autoridad inmediata Superior del Ministerio de Educación, Ciencia y Tecnología de la Provincia de Salta.-

Artículo 8º.- Facultad del Subprograma de Becas: Existiendo postulantes y/o renovantes que no cumplan alguno o los requerimientos académicos establecidos en el Artículo 5º y presenten alguna situación de inequidad y/o vulnerabilidad socioeconómica, familiar, de salud o cultural, según criterio fundado del Equipo Profesional Interdisciplinario del Subprograma de Becas, podrá resolverse su inclusión en el sistema adjudicatario o reconsiderar favorablemente su otorgamiento o renovación, según la oportunidad de que se trate, cuando el rendimiento académico sea acorde a dicha situación.-

Sin perjuicio de lo anterior, se consideraran como especiales situaciones de inequidad y/o vulnerabilidad, las siguientes:



a).- Estudiantes con discapacidad: Alumnos que padezcan una alteración funcional permanente o prolongada que, en relación a su edad y medio social impliquen desventajas considerables para su integración familiar, social, educacional o laboral. La condición de discapacidad deberá ser acreditada por las dependencias oficiales que oportunamente se indique, conforme el sistema de protección integral establecido por la Ley N° 22.431.

b).- Estudiantes pertenecientes a Pueblos Originarios: Alumnos que provengan de una comunidad originaria, es decir, de un conjunto de familias que se reconozcan como tales por el hecho de descender de poblaciones que habitaban el territorio nacional en la época de la conquista o colonización, conforme lo estipulado por la Ley N° 23.302. La pertenencia a una comunidad originaria deberá ser avalada formalmente la autoridad de la comunidad, reconocida como perteneciente a Etnia Originaria por el Instituto Nacional de Asuntos Indígenas.



Anexo II: Formulario Becas



GOBIERNO DE LA PROVINCIA DE SALTA.
Ministerio de Educación, Ciencia y Tecnología.

Domicilio: Mitre N° 274 1 Piso, Ciudad de Salta.
Tel./Fax: 0387-4370606 - Internos 14-15

SUBPROGRAMA DE BECAS UNIVERSITARIAS Y TERCARIAS SOLICITUD DE BECA DE ESTUDIO 2018

Artículo 10°.- Valor de declaración jurada. Conocimiento: La información suministrada en los medios y/o instrumentos de inscripción o renovación de Beca que refiere el artículo anterior, tendrá valor de declaración jurada.

La inscripción o tramite de renovación efectuados por los postulantes o renovantes, importará el conocimiento y aceptación del presente Reglamento de Becas y de la/s Resolución/es que lo complementa.-

Artículo 20°.- Falsedad de datos y/o adulteración de documentación: La falsedad u ocultamiento de los datos requeridos en los medios o instrumentos de inscripción o por el Subprograma de Becas, y la adulteración o falseamiento de la documentación presentada, hará pasible al firmante de las sanciones administrativas, académicas, civiles y penales correspondientes.-

Datos Personales del Estudiante

Apellido/s		Nombre/s	
------------	--	----------	--

CUIL		-								-		Nacionalidad	
------	--	---	--	--	--	--	--	--	--	---	--	--------------	--

Fecha de Nacimiento	/	/	Edad		Estado Civil	
---------------------	---	---	------	--	--------------	--

Cel.		Teléfono		E-mail	
------	--	----------	--	--------	--

Pertenece a un Pueblo Originario?	No	Si	Etnia:	
¿Tiene alguna Incapacidad o Discapacidad?	No	Si	Cuál?:	

Residencia del Grupo Familiar:

Municipio		C. Postal		Años de Residencia del Grupo Familiar en la Provincia de Salta	
Calle					
Barrio					



UCS – Facultad de Ingeniería
Proyecto de Grado: Business Intelligence para Subprograma de Becas Provinciales
Anexos

GRUPO FAMILIAR: Se entiende por tal a los miembros de la familia del estudiante que residen bajo el mismo techo y comparten ingresos y gastos.												
GRUPO FAMILIAR			EDUCACIÓN			TRABAJO E INGRESOS			SALUD			
Nº	Vínculo con el Solicitante *	D.N.I. Nº	Nivel Educativo Alcanzado*	Concurre		Título obtenido	Situación Laboral*	Cargo o Actividad que desempeña*	Ingresos Mensuales en pesos*	Cobertura de Salud	¿Padece enfermedad?	¿Padece Discapacidad?
				SI	NO							
1	Solicitante											
2												
3												
4												
5												
6												
7												
8												
9												
10												
	*Vínculo Madre/Padrastro o • Hermana/o • Hija/o • Tía/o • Sobrina/o • Abuela/o • Esposa/o • Concubina/o • Cuñada/o • Suegra/o		Nivel Educativo • Inicial • Prim Incompleto • Prim Completo • Sec Incompleto • Sec Completo • Terc Incompleto • Terc Completo • Univ Incompleto • Univ Completo				*Situación Laboral Ocupado (Ingresos Fijos). Sub-Ocupado (Ingresos esporádicos) Desocupado (sin ingresos) Jubilado/Pensionado	*Cargo o actividad que desempeña • Negocio y/o empresa • propia • Gerente • Empleado • Independiente c/ CUIT • Trabajador esporádico • Independiente s/ CUIT • Estudiante • Ama de casa	*Total	*cobertura • Obra Social • Prepaga • Hosp. Público	Marcar con una X lo que Corresponda. Si posee, presentar certificado Médico correspondiente que acredite tal situación.	

IMPORTANTE: El estudiante deberá llenar la información solicitada atendiendo *únicamente* a las alternativas presentadas.



GRUPO FAMILIAR (MARQUE CON X LO QUE CORRESPONDA)

VIVIENDA		EGRESOS MENSUALES	
Ubicación Zonal		GRUPO FAMILIAR Egresos mensuales del grupo familiar en pesos.	
Rural		Vivienda	\$
Urbano		Alimentación	\$
Servicios de la Vivienda		Salud	\$
Gas		Educación	\$
Agua		Impuestos	\$
Luz Eléctrica		Transportes	\$
Teléfono		Otros	\$
Cable		Subtotal \$	
Internet			
Condición de ocupación de la Vivienda		ESTUDIANTE (Completar si el estudiante deja el domicilio familiar para cursar sus estudios.) Egresos mensuales del estudiante en pesos.	
Propietario		Vivienda	\$
Más de una Propiedad?		Alimentación	\$
Inquilino		Transporte	\$
Ocupante de hecho		Bibliografía	\$
		Otros	\$
		Subtotal \$	
		Total Egresos \$	

ACTUACIONES ADMINISTRATIVAS

Firma y sello del profesional



GOBIERNO DE LA PROVINCIA DE SALTA.
Ministerio de Educación, Ciencia y Tecnología.

Domicilio: Mitre N° 274 1 Piso, Ciudad de Salta.
Tel./Fax: 0387-4370606 - Internos 15-21

DATOS ACADEMICOS

Institución Académica Donde Estudia					
Carrera					
Provincia		Localidad		Año de ingreso a la Carrera	

Compromiso Académico 2018

Complete el siguiente cuadro, detallando las materias que se compromete a cursar en el presente período lectivo

N°	Nombre de la Materia	Año de la Carrera a la que corresponde la materia según Plan de Estudio	Marcar con X lo que Corresponda		
			Anual	1° Cuatri.	2° Cuatri.
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					

PLAZO DE PRESENTACIÓN DEL FORMULARIO

La presente solicitud, junto con la documentación repaldatoria, deberá presentarse desde el 01 febrero al 02 de marzo del corriente año, en el domicilio del Programa Becas (Salta- Capital Mitre N° 274, 1° piso) o en las casas Municipales, en caso de alumnos del interior provincial que no puedan o se les dificulte hacerlo en nuestras oficinas.-

Firma del Estudiante o tutor: _____

Aclaración: _____

DNI N°: _____



 GOBIERNO DE LA PROVINCIA DE SALTA. Ministerio de Educación, Ciencia y Tecnología.	Domicilio: Mitre N° 274 1 Piso, Ciudad de Salta. Tel./Fax: 0387-4370606 - Internos 15-21
---	---

DOCUMENTACIÓN A PRESENTAR

PERSONAL

- Fotocopia del D. N. I. del Solicitante (1ª y 2ª Hoja).
- Constancia de CUIL: Si ya lo tiene generado, lo obtiene en el sitio web del ANSES: www.anses.gov.ar o en el Correo Argentino; de lo contrario debe concurrir a la oficina del ANSES mas cercana a su domicilio. (Su falta de presentación es **excluyente**).-
- Certificado Original de Residencia y Convivencia del grupo familiar expedido por autoridad Policial.

SOCIOECONÓMICA FAMILIAR

Constancias de ingresos económicos de los integrantes del grupo familiar

- Constancia de CUIL de ambos padres.
- **Trabajadores en relación de Dependencia, Jubilados y/o Pensionados** Recibo de sueldo.
- **Trabajadores Independientes:** constancia de monotributo, declaración jurada de haberes adjuntando **Certificación NEGATIVA del ANSES**
- **Desocupados:** Los miembros de la Familia mayores de 18 años que se declaren desocupados, deben presentar **CONSTANCIA NEGATIVA de ANSES incluyendo al solicitante.**-

En caso de corresponder:

- **Por enfermedad** del solicitante y/o miembro del grupo familiar: Certificado Medico expedido por Hospital Público, original, firmado y actualizado.
- **Por discapacidad:** Certificado oficial de Discapacidad.
- **Padres Separados:** Copia de Sentencia de Divorcio o Exposición policial de Separación de Hecho.
- **Por Defunción de algunos de los padres:** fotocopia del Acta.
- **Estudiantes pertenecientes a ETNIAS de Pueblos Originarios** deberán presentar constancia o declaración Jurada

DOCUMENTACION ACADÉMICA

Alumnos Reinscriptos (estudiantes que cursaran desde el segundo año en adelante su carrera)

- Estado Curricular original o Certificado Analítico actualizado, firmado y sellado por la autoridad competente.
- Plan de Estudio (discriminado por años de carrera y cantidad de materias en cada año



Anexo III: Capturas sistema becas

Administración de Inscripciones

Año Lectivo: 2018 Documento: Nombre: Buscar

Carrera: Institución: Municipio:

☒ Evaluación WEB ☒ Seleccionado p/Beca
☒ Evaluación por Sistema ☒ Excluidos

Documento	Apellido y Nombre	T/U	I/R	Resultado	Municipio
-----------	-------------------	-----	-----	-----------	-----------

Sin Evaluación Seleccionado Excluidos

Ilustración 87: Gestión de Inscripciones - Pagina Inicio. Elaboración propia

Agregar Datos de la Inscripción

Datos Personales Grupo Familiar I Grupo Familiar II Datos Académ. I Datos Académ. II Resumen Servicio Social

Datos Personales del Solicitante

N° de CUIL: Apellido: Nombre: Nacionalidad: Argentino/a Fecha Nac.: Edad: Estado Civil: N° Celular: N° Teléfono: E-Mail: Lugar de Nac.: Salta Etnia Aborigen: (Ninguna) ☐ Dejó domicilio del Grupo Familiar para Estudiar ☐ Discapacitado

Residencia Grupo Familiar

Municipio: SALTA Capital Código Postal: Años de Residencia: Calle: N°: Etapa: Mza: Block: Piso: Dpto: Med: Barrio:

Tutor (Menores de 18 Años)

CUIL: Apellido: Nombre: F.Nac.:

Ilustración 88: Nueva Inscripción - Datos Personales. Elaboración propia



Agregar Datos de la Inscripción

Tab: **Grupo Familiar I**

Composición del Grupo Familiar

Vínculo	Documento	Sit. Laboral	Cargo/Ocup	Ingr. Mensual	Nivel Educ.	Cobertura Salud
---------	-----------	--------------	------------	---------------	-------------	-----------------

Integrantes Grupo Familiar: 0 Total de Ingresos Grupo Familiar: 0,00

Datos del Familiar

Vínculo: [] Documento: [0][0][0]

Nivel Educ.: [] ☐ Concurre? Título Obtenido: [] ☒

Sit. Laboral: [] Cargo/Ocup.: [] Ingr. Mens.: 0,00

Cobertura Salud: [] ☐ Enfermedad? [] ☐ Discapacidad? []

Ilustración 89: Nueva Inscripción - Grupo Familiar 1. Elaboración propia

Agregar Datos de la Inscripción

Tab: **Grupo Familiar II**

Vivienda

Ubicación: []

Tipo: []

Condición de Ocupación: []

Servicios

☐ Gas ☐ Agua

☐ Luz Eléctrica ☐ Teléfono

☐ Cable ☐ Internet

Automóvil

☐ Posee Automóvil Modelo: [] Marca: []

Egresos del Grupo Familiar

Concepto	Importe
Vivienda	0,00
Alimentación	0,00
Salud	0,00
Educación	0,00
Impuestos	0,00
Transporte	0,00
Otros	0,00

Total Egresos Grupo Familiar: 0,00

Egresos del Estudiante

Concepto	Importe
Vivienda	0,00
Alimento	0,00
Transporte	0,00
Bibliografía	0,00
Otros	0,00

Total de Egresos del Estudiante: 0,00

Ilustración 90: Nueva Inscripción - Grupo Familiar 2. Elaboración propia

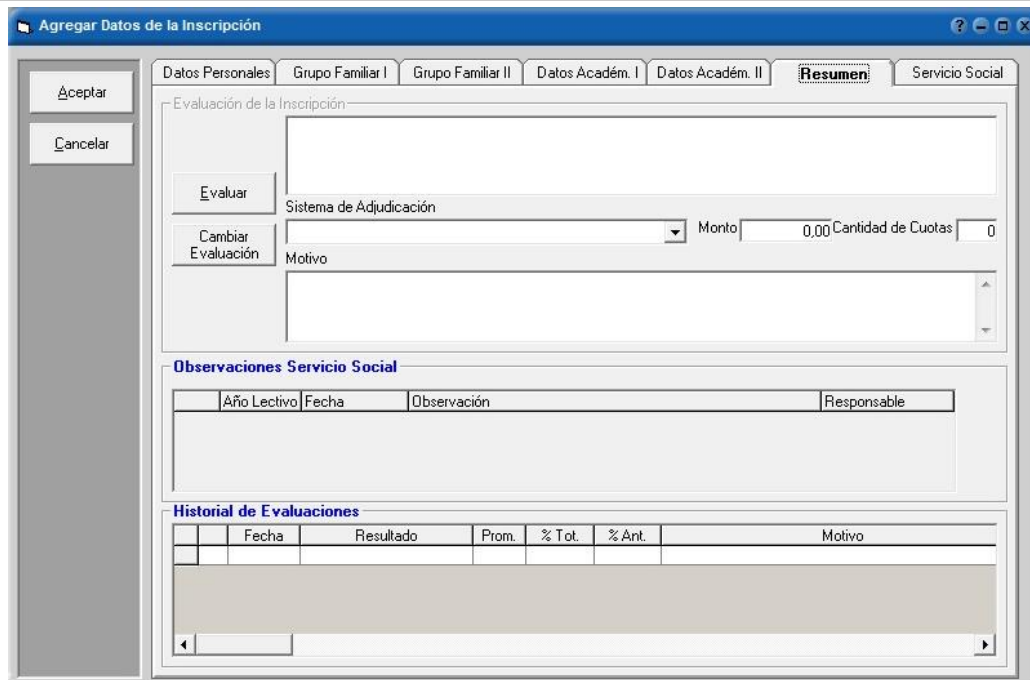


The screenshot shows the 'Agregar Datos de la Inscripción' window with the 'Datos Académ. I' tab selected. The form includes fields for 'Año Lectivo' (2016), 'Fecha de Inscripción' (11/10/16), and 'Clase de Aspirante' (Ingresante). The 'Datos de la Carrera' section contains fields for 'Plan de Estudios', 'Carrera', 'Institución', 'Provincia', 'Año de Ingreso a la Carrera' (2016), 'Tipo de Institución', and 'Suc. Bancaria'. There are checkboxes for 'No Presentó Constancia de Inscripción en la Carrera', 'No Presentó Certificado Analítico', and 'Carrera e Institución del Mismo caracter existente en el Domicilio del Grupo Familiar'. The 'Información de Polimodal' section has fields for 'Año Egreso', 'Promedio', and checkboxes for 'No Presentó Analítico o Constancia con Promedio', 'Adeuda Materias del Polimodal', and 'Establecimiento Polimodal o Ex-Medio no es de la Provincia de Salta'. The 'Cuestionario' section contains several checkboxes regarding benefits, bursaries, and previous studies.

Ilustración 91: Nueva Inscripción - Datos Académicos 1. Elaboración propia

The screenshot shows the 'Agregar Datos de la Inscripción' window with the 'Datos Académ. II' tab selected. The form is divided into three main sections: 'Plan de Estudios Discriminado', 'Compromiso Académico', and 'Situación Curricular'. The 'Plan de Estudios Discriminado' section contains a table with columns: Curso, Anuales, 1º Cuatrim., 2º Cuatrim., and Total. The 'Compromiso Académico' section contains a table with columns: Año, Cant., Año a Cursar, and Regimen. The 'Situación Curricular' section contains a table with columns: Curso, Nota, and Año. Below these tables are fields for 'Total Materias', 'Mat. Aprob. Año Acad. Anterior' (0), 'Promedio' (0.00), and 'Mat. Apr.' (0). The 'Residencia Mientras Estudia' section contains fields for 'Calle', 'Nº', 'Etapas', 'Mza.', 'Block', 'Piso', 'Dpto', 'Med.', 'Barrio', 'Entre Calle', 'Provincia' (Salta), 'Municipio' (Salta Capital), and 'Código Postal'.

Ilustración 92: Nueva Inscripción - Datos Académicos 2. Elaboración propia



Agregar Datos de la Inscripción

Aceptar
Cancelar

Datos Personales Grupo Familiar I Grupo Familiar II Datos Académ. I Datos Académ. II **Resumen** Servicio Social

Evaluación de la Inscripción

Evaluar

Sistema de Adjudicación

Cambiar Evaluación

Monto 0,00 Cantidad de Cuotas 0

Motivo

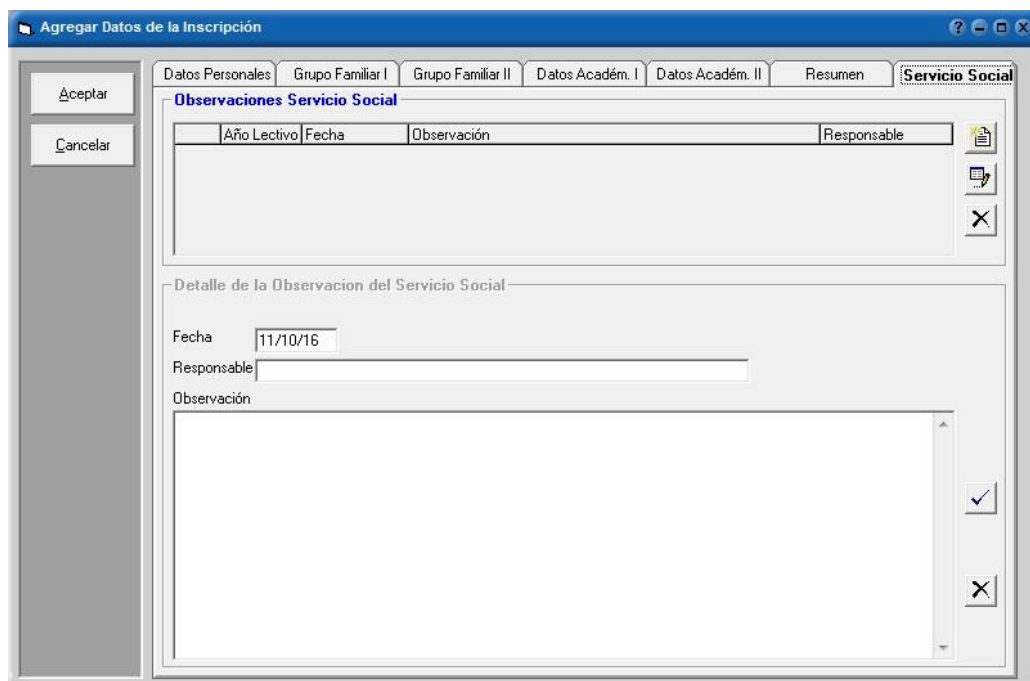
Observaciones Servicio Social

Año Lectivo	Fecha	Observación	Responsable
-------------	-------	-------------	-------------

Historial de Evaluaciones

	Fecha	Resultado	Prom.	% Tot.	% Ant.	Motivo
--	-------	-----------	-------	--------	--------	--------

Ilustración 93: Nueva Inscripción - Resumen Evaluación. Elaboración propia



Agregar Datos de la Inscripción

Aceptar
Cancelar

Datos Personales Grupo Familiar I Grupo Familiar II Datos Académ. I Datos Académ. II Resumen **Servicio Social**

Observaciones Servicio Social

Año Lectivo	Fecha	Observación	Responsable
-------------	-------	-------------	-------------

Detalle de la Observación del Servicio Social

Fecha 11/10/16

Responsable

Observación

Ilustración 94: Nueva Inscripción - Servicio Social. Elaboración propia



Administración de Renovaciones

Año Lectivo: 2018
Documento:
Nombre:
Carrera:
Institución:
☒ Renuevan el Beneficio ☒ No Renuevan el Beneficio ☒ Sin Evaluar

Documento	Apellido y Nombre	T/U	I/R	Resultado	Municipio
-----------	-------------------	-----	-----	-----------	-----------

Sin Evaluación Renueva el Beneficio No Renueva el Beneficio

Ilustración 95: Renovaciones - Inicio. Elaboración propia

Modificar Datos de Beneficiario

Datos Personales

Documento: 37636520 CUIL: 37636520 5 Suc. Bco. Salta - 100 - España N° 550 - SALTA - Capital
Apellido: ARROYO Nombre: PABLO GASTÓN
N° de Cuenta: 410009451866413 ☒ Firmó Convenio 18/04/16
Método: Renueva el Beneficio Monto: 800,00
N° Celular: 155859178 N° Teléfono: 4231514 E-Mail: arroyo.pablogaston@gmail.com
Etnia Aborigen: (Ninguna)

Residencia Mientras Estudia

Calle: Dionisio Ramos 1859 - B° Policial
Provincia: SALTA Municipio: SALTA Capital
Código Postal: 4400

Historial de Evaluaciones

Fecha	Resultado	Prom.	% Tot.	% Ant.	Motivo
12/02/14	No Renueva el Beneficio	7,82	78,00 %	0,00 %	No se puede realizar evaluación socio económica: del
21/03/13	Renueva el Beneficio	7,33	78,00 %	0,00 %	Debe presentar original del estado curricular.
01/03/12	Renueva el Beneficio	0,00	0,00 %	0,00 %	
31/01/12	Renueva el Beneficio	6,33	70,00 %	0,00 %	
01/03/11	Seleccionado para una Beca	8,26	0,00 %	0,00 %	Seleccionado para una Beca de Estudio.
04/02/11	Preselec. para una Beca de	8,26	0,00 %	0,00 %	Hasta el 25 de Febrero de 2011 deberá presentar este

Ilustración 96: Renovación - Beneficiario. Elaboración propia



Modificar Datos de Beneficiario

Beneficiario | Inscript/Renov | Grupo Familiar | **Resoluciones** | Acreditaciones | Datos Académicos | Servicio Social

Bajas Provisorias

Fecha	Motivo de Baja Provisoria
-------	---------------------------

Historial de Resoluciones

Fecha	Nº	Tipo	Artículo/Anexo	Observaciones
19/04/16	001/16	Alta	II - RENOVANTES	Renueva el Beneficio
11/04/15	001/15	Alta	II - RENOVANTES CAPITAL	Renovacion Capital
10/07/14	009/14	Alta	ANEXO I - Seleccionados 2014	Seleccionado para una Beca de Estudio.
23/05/13	008/13	Alta	Anexo II - Renov. por 8 Ctas.	
04/04/12	010-12	Alta	II - Renovantes Provincia	Renueva el Beneficio
18/04/11	017/11	Alta	I - Seleccionados Provincia	Seleccionado para una Beca de Estudio.

Ilustración 97: Renovación - Evaluación. Elaboración propia

Administración de Planes de Estudio

Seleccionar | Modificar | Agregar | Eliminar | Consultar | Salir

Universidad: Carrera: Plan de Estudios: Buscar

☒ Solo Instituciones Homologadas ☒ Solo con Suc. Bancaria Registrada ☒ Solo con Municipio Registrado

ID	Institución	Carrera	Plan de Estudios	Años	Mat.	Hom.
						☒

Ilustración 98: Planes de Estudio - Inicio. Elaboración propia



Agregar Datos del Plan de Estudios

Institución ...

Institución Homologada ☐

Carrera ...

Plan de Estudios

Duración (Años) Total de Materias

Plan de Estudios Discriminado

	Año	Anuales	1º Cuatrim.	2º Cuatrim.	Total
*					

Ilustración 99: Nuevo Plan de Estudio. Elaboración propia

Agregar Datos de Institución Superior

Nombre

Tipo de Institución ... ☐ Institución Homologada

Provincia ...

Municipio

Suc. Bancaria mas cercana

Ilustración 100: Nueva Institución. Elaboración propia



ID	Tipo de Institución
0003	CURSO DE CAPACITACION
T	Terciario
U	Universidad

Ilustración 101: Tipo de Institución. Elaboración propia

Ilustración 102: Selección de Provincia. Elaboración propia

Ilustración 103: Nueva Carrera. Elaboración propia



Anexo IV: Entrevista

Entrevista a Coordinador Programa de Becas Lic. Guillermo Saavedra

P: ¿Cuál o cuáles son las actividades que realiza su organización?

E: Recibir, evaluar y designar beneficiarios de Becas universitarias y terciarias.

P: ¿Qué procesos necesitan para ello?

E: Primero la persona se preinscribe en nuestra página web, luego descarga un formulario y lo debe entregar completo en nuestras oficinas para su comprobación y evaluación. Luego se publican los resultados.

P: ¿Quiénes los realizan?

E: Tenemos dos personas en el área administrativa que reciben la documentación del postulante y un área técnica y servicio social que realiza la evaluación, allí tenemos cuatro empleados, dos en técnica y dos en servicio social.

P: ¿Son suficientes esas personas?

E: El Área administrativa está bien en cuanto al personal, pero al área técnica le vendría bien contar con dos empleados más.

P: ¿Cuáles son los papeles que verifican en el área administrativa?

E: Se controla que se presenten todas las hojas correctamente completadas del formulario descargado más la documentación adicional que se especifica en el mismo, como ser recibos de sueldo, certificado de residencia y convivencia, etc.

P: ¿Qué pasa si los formularios están incompletos?

E: Se lo devuelve a quien pretende entregar para que complete o agregue lo que falta, según se le indica en el momento del rechazo.

P: ¿Qué clases de becas otorgan?

E: En ambas becas tenemos tanto solicitudes como renovaciones.

P: ¿Estas becas llegan a todas las personas de la provincia de Salta?

E: Si, el programa abarca a todas las personas que pertenezcan a un municipio dentro de la provincia de Salta.

P: ¿Puede una persona que estudia en otra provincia postularse por una beca?

E: Si, siempre y cuando cumplan el proceso de evaluación.

P: ¿Cuál o cuáles serían los indicadores más representativos del proceso de evaluación?

E: Deseamos saber cuál es la cantidad de personas por tipo de institución de los municipios de Salta que reciben una beca en cierto periodo. Así mismo también queremos conocer cuál es la cantidad de personas por municipio que dejan de percibir el beneficio por no cumplir el control académico en un periodo determinado. Por último, también quisiéramos saber por municipio cuales condiciones son las que impiden a una persona obtener una beca.



P: ¿Qué pasos realizan en la evaluación? ¿Qué condiciones se necesitan para aprobarla?

E: Luego de recibir los documentos del postulante por el área administrativa, se evalúa:

Primero, condiciones personales y/o familiares después, se revisa su condición, y para finalizar se evalúan los aspectos socioeconómicos del postulante y su grupo familiar.

P: ¿Qué evalúan en las condiciones personales y/o familiares?

E: Comprobar que sea menor a 26 años, también la nacionalidad argentina, su residencia en la provincia de por lo menos tres años y que no posea otro beneficio similar.

P: ¿Cómo comprueban estas condiciones?

E: Deberá presentar una fotocopia del documento.

P: ¿Y que evalúan en las condiciones académicas?

E: Se comprueba si tiene una inscripción en alguna institución de nivel superior, si tiene el secundario finalizado con promedio general superior a 7 (siete) y si adeudó materias durante el cursado

P: ¿Cómo comprueban estas condiciones?

E: El postulante debe entregar junto al formulario una copia del título autenticado por el colegio con el promedio general del secundario.

P: ¿Qué aspectos socioeconómicos evalúan?

E: Dependiendo de cuantas personas formen el grupo familiar los ingresos posibles de ese grupo tendrán cierto límite que no deben superar, no cumpliendo esto quedaran fuera del beneficio.

P: ¿Cómo controlan estas condiciones?

E: El postulante deberá presentar una constancia de convivencia y residencia expedido por la policía de salta, también deberá presentar un recibo de sueldo o una declaración jurada si es independiente.

P: ¿Cuánto es el tiempo promedio de formularios evaluados por hora que realiza una persona?

E: En promedio se evalúan 5 formularios por hora.

P: ¿Cuánto tiempo lleva la evaluación de todos los postulantes y renovaciones?

E: Aproximadamente lleva entre 2 o 2 meses y medio

P: ¿Qué se necesita para renovar la beca?

E: Para renovar la beca el beneficiario deberá presentar un certificado analítico de materias expedido por la misma universidad en la que estudia que deberá mostrar si aprobó el 66% de las materias en la que se inscribió y además deberá tener un promedio general al menos de 6 (seis).

P: ¿Se aplica de la misma manera para los que estudian en un terciario?

E: En el caso de la cantidad de materias sí, pero deberán tener un promedio de al menos de 7 (siete).

P: ¿Se realiza algún tipo de control durante el beneficio? ¿Cuál sería?



E: Si, antes del pago del quinto mes, se deberá completar un formulario con las materias anuales que cursa, las regularizadas y las aprobadas durante el primer semestre.

P: ¿Cómo notifican a los nuevos beneficiarios? ¿Y a los de renovación?

E: Se los notifica a ambos por una carta enviada a su domicilio para firmar un convenio en una fecha elegida por el programa.



Anexo V: Data Warehouse y CUBO

Diseño Físico DW²⁹

Dimensiones y tablas de hechos

Tabla	dimInstitucionTipos								
Tipo de Tabla	Dimensional								
Descripción	Tabla de registros del tipo de instituciones								
Destino							Origen		
Columna	Descripción	Tipo de Dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo de Dato
idInstitucionTipo	Código del tipo de institución	integer	(2^31) a 2^31-1	pk	factEvaluaciones	No	institucionTipos	id	bigint
institucionTipo	Clasificación del tipo de institución	nvarchar	255	-	-	No	institucionTipos	nombre	nvarchar

Tabla 20: Descripción de Dimensión dimInstitucionTipos. Elaboración propia

Tabla	dimMunicipios								
Tipo de Tabla	Dimensional								
Descripción	Tabla de registros de municipios								
Destino							Origen		
Columna	Descripción	Tipo de Dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo de Dato
idMunicipio	Código que identifica al municipio	integer	(2^31) a 2^31-1	pk	factEvaluaciones	No	municipios	id	bigint
municipio	Nombre del municipio	nvarchar	255	-	-	No	municipios	nombre	nvarchar

Tabla 21: Descripción de Dimensión dimMunicipios. Elaboración propia

Tabla	dimCondicionTipos								
Tipo de Tabla	Dimensional								
Descripción	Tabla de registros del tipo de condiciones								
Destino							Origen		
Columna	Descripción	Tipo de Dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo de Dato
idCondicionTipo	Código que identifica el tipo de condición	integer	(2^31) a 2^31-1	pk	factEvaluaciones	No	observaciones	id	bigint
condicionTipo	Requisito o condición a cumplir para una evaluación	nvarchar	255	-	-	No	observaciones	nombre	nvarchar

Tabla 22: Descripción de Dimensión dimCondicionTipos. Elaboración propia

Tabla	dimFechas								
Tipo de Tabla	Dimensional								
Descripción	Tabla de registros de periodos								
Destino							Origen		
Columna	Descripción	Tipo de Dato	Tamaño	Clave	FK para	Null	Tabla	Campo	Tipo de Dato
idFecha	Código que identifica el periodo elegido	integer	(2^31) a 2^31-1	pk	factEvaluaciones	No	inscripciones	fecha	date
anio	Año Lectivo correspondiente	integer	(2^31) a 2^31-1	-	-	No	inscripciones	fecha	date

Tabla 23: Descripción de Dimensión dimFechas. Elaboración propia

²⁹ Adjunto backup de DW en carpeta de **Backup DW**



Tabla	factEvaluaciones					
Tipo de Tabla	Dimensional					
Descripción	Tabla de hechos, conformada por las claves primarias de las dimensiones y las métricas					
Destino						
Columna	Descripción	Tipo de Dato	Tamaño	Clave	FK para	Null
idEvaluacion	Código que identifica un registro de la tabla de hechos	Integer	(2^31) a 2^31-1	-	-	No
idInstitucionTipo	FK – tabla dimInstitucionTipos	Integer	(2^31) a 2^31-1	fk	-	No
idCondicionTipo	FK – tabla dimCondicionTipos	Integer	(2^31) a 2^31-1	fk	-	Si
idMunicipio	FK – tabla dimMunicipios	Integer	(2^31) a 2^31-1	fk	-	No
idFecha	FK – tabla dimFechas	Integer	(2^31) a 2^31-1	fk	-	No
aceptada	Campo contador de evaluaciones aceptadas	Integer	(2^31) a 2^31-1	-	-	Si
rechazada	Campo contador de evaluaciones rechazadas	Integer	(2^31) a 2^31-1	-	-	Si

Tabla 24: Descripción de Tabla de Hechos factEvaluaciones. Elaboración propia

Consultas SQL para la creación del DW respetando las descripciones anteriores:

```
CREATE DATABASE becasDW;
```

```
USE becasDW;
```

```
CREATE TABLE dimInstitucionTipos (
    idInstitucionTipo INT PRIMARY KEY NOT NULL,
    institucionTipo NVARCHAR(255) NOT NULL,
);
```

```
CREATE TABLE dimFechas (
    idFecha INT PRIMARY KEY NOT NULL,
    anio INT NOT NULL,
);
```

```
CREATE TABLE dimMunicipios (
    idMunicipio INT PRIMARY KEY NOT NULL,
    municipio NVARCHAR(255) NOT NULL,
);
```

```
CREATE TABLE dimCondicionTipos (
    idCondicionTipo INT PRIMARY KEY NOT NULL,
    condicionTipo NVARCHAR(255) NOT NULL,
);
```

```
CREATE TABLE factEvaluaciones (
    idEvaluacion INT NOT NULL,
    idInstitucionTipo INT FOREIGN KEY REFERENCES
dimInstitucionTipos(idInstitucionTipo) NOT NULL,
    idCondicionTipo INT FOREIGN KEY REFERENCES dimCondicionTipos(idCondicionTipo)
NULL,
    idMunicipio INT FOREIGN KEY REFERENCES dimMunicipios(idMunicipio) NOT NULL,
    idFecha INT FOREIGN KEY REFERENCES dimFechas(idFecha) NOT NULL,
    aceptada INT NULL,
    rechazada INT NULL
);
```

Para los procesos ETL se realizaron usando la herramienta Microsoft SQL Server Data Tools Business Intelligence trabajando con la extensión de Integration Services.

Una vez iniciada la herramienta se inicia un nuevo proyecto.

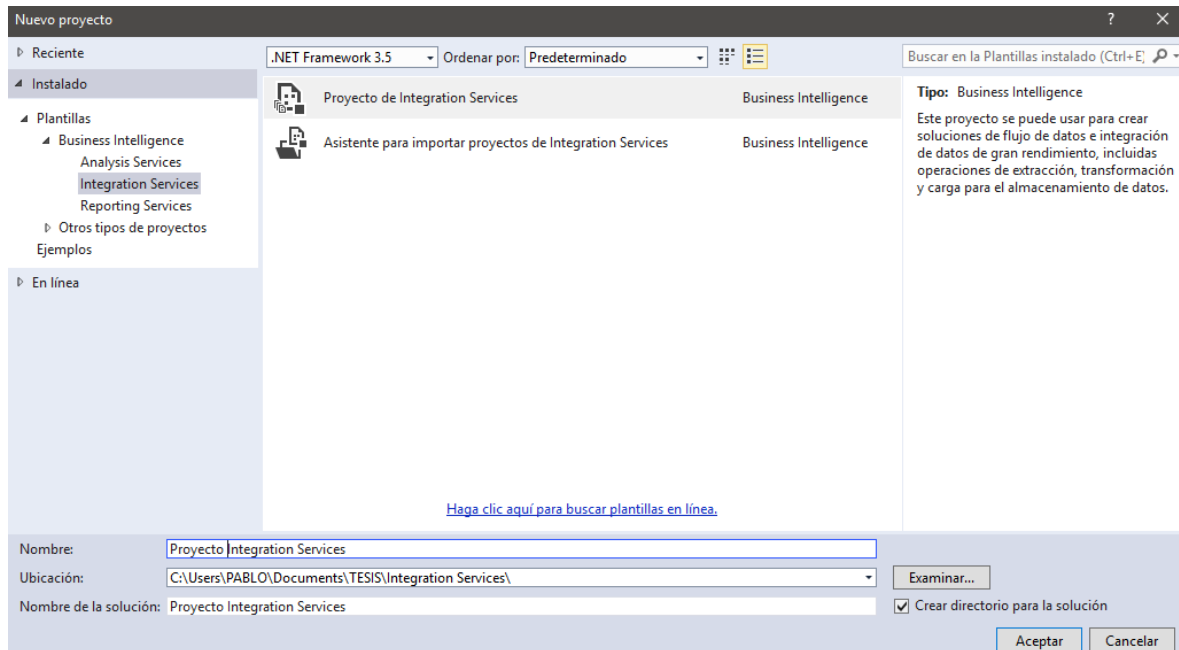


Ilustración 104: ETL - Nuevo proyecto Integration Services. Elaboración propia

Se seleccionará un Proyecto de Integration Services dentro de las plantillas de Analysis Services incluidas en Business Intelligence, se especifica el nombre del proyecto “Proyecto Integration Services” y la ubicación donde se guardará dicho proyecto.

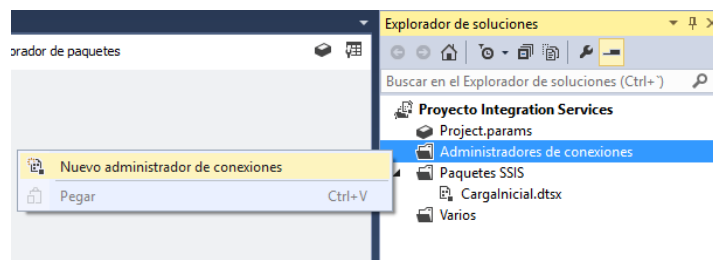


Ilustración 105: ETL - Administrador de conexiones. Elaboración propia

Ya una vez creado el proyecto, en el Explorador de soluciones agregamos la fuente y el destino de datos con los cuales se realizarán los procesos ETL, realizando clic izquierdo en Administradores de conexiones seleccionamos un Nuevo administrador de conexiones. A este

³⁰ Adjunto proyecto en la carpeta **Proyecto Integration Services**



paso lo realizaremos dos veces, una para conectar la base de datos que actuará como fuente “becas” y la otra para conectar la base de datos “becasDW” que tendrá el papel de destino.

Para las conexiones se deberán ingresar las credenciales que tienen estas bases de datos, siendo las que se usaran para iniciar sesión en el servidor.

Crearemos tres paquetes especificando en su nombre a cuál proceso ETL hacen referencia por el cual serán:

- Carga Inicial
- Actualización Diaria
- Actualización Anual

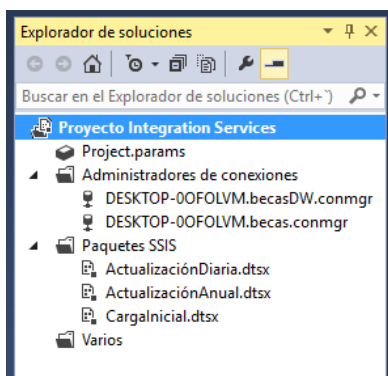


Ilustración 106: ETL - Paquetes SSIS. Elaboración propia

En estos paquetes se agregarán los flujos de datos especificados en el proceso ETL de la solución propuesta.

Carga Inicial

En el caso de la carga inicial del DW los flujos de datos se acomodarán de manera secuencial, colocando al final la Carga de la Tabla de Hechos factEvaluaciones ya que esta depende de las dimensiones del DW.

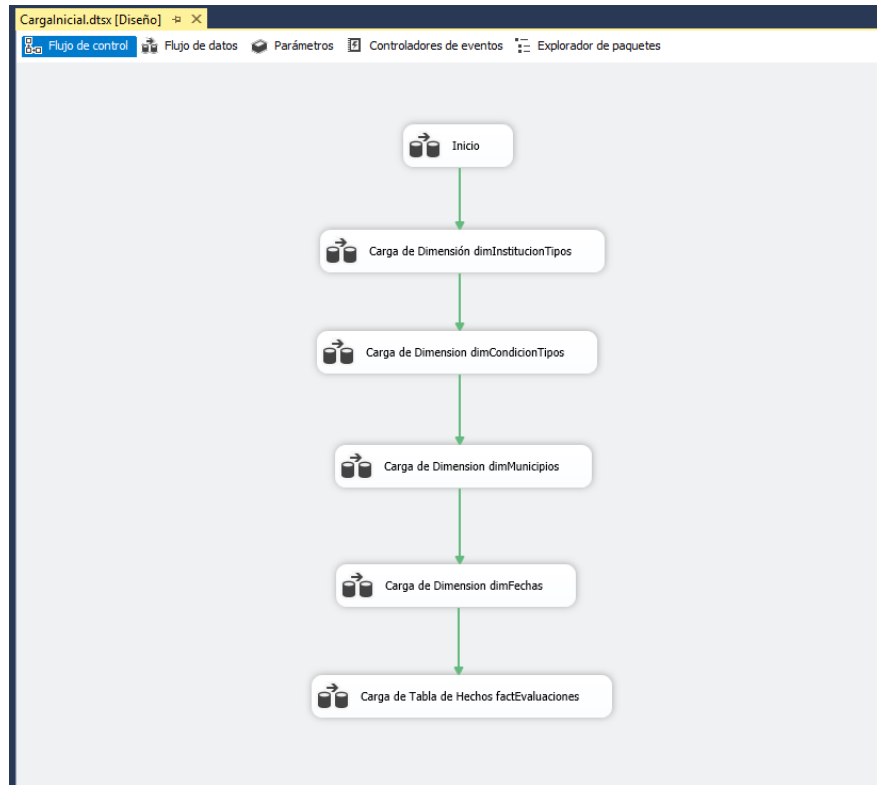


Ilustración 107: ETL - Carga Inicial. Elaboración propia

Dentro de cada flujo de datos agregar las sub-tareas de cada una de las Cargas de las dimensiones y tabla de hechos.

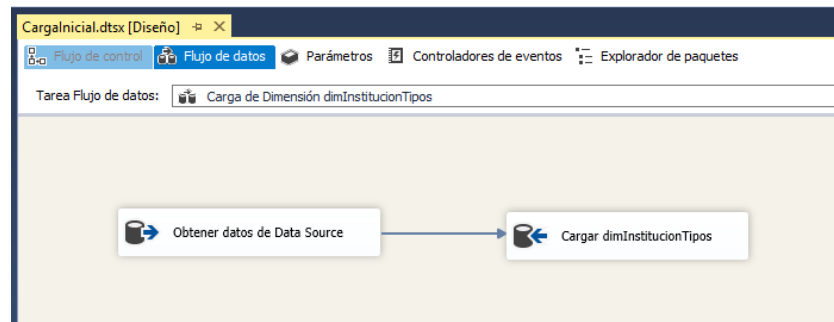


Ilustración 108: ETL - Flujo de datos. Elaboración propia

En la sub-tarea de Obtener datos de Data Source se especifica a “becas” como la fuente de datos en el Administrador de conexiones. El modo de acceso a los datos será a través de un Comando SQL y se deberá ingresar la consulta especificada en la Solución propuesta en el Texto de comando SQL.

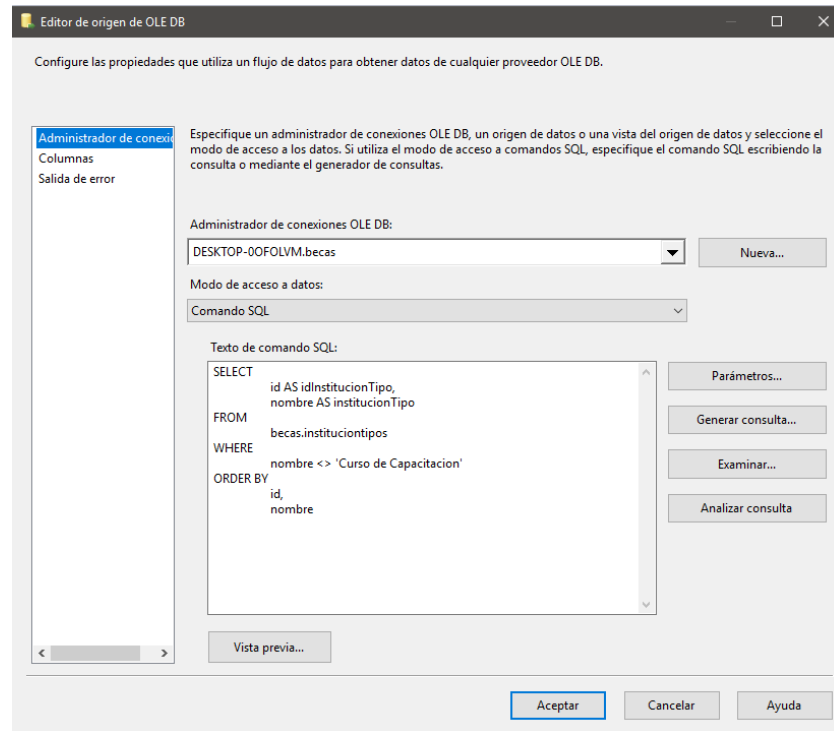


Ilustración 109: ETL - Editor de origen de datos. Elaboración propia

De la misma manera en la sub-tarea de Cargar la dimensión se tomarán los datos obtenidos a partir de la consulta de la anterior sub-tarea. En el administrador de conexiones se seleccionará “becasDW” donde serán depositados los datos de la consulta, el modo de acceso a los datos será por medio de Tabla o Vista, especificando el nombre de la dimensión a la que corresponde.

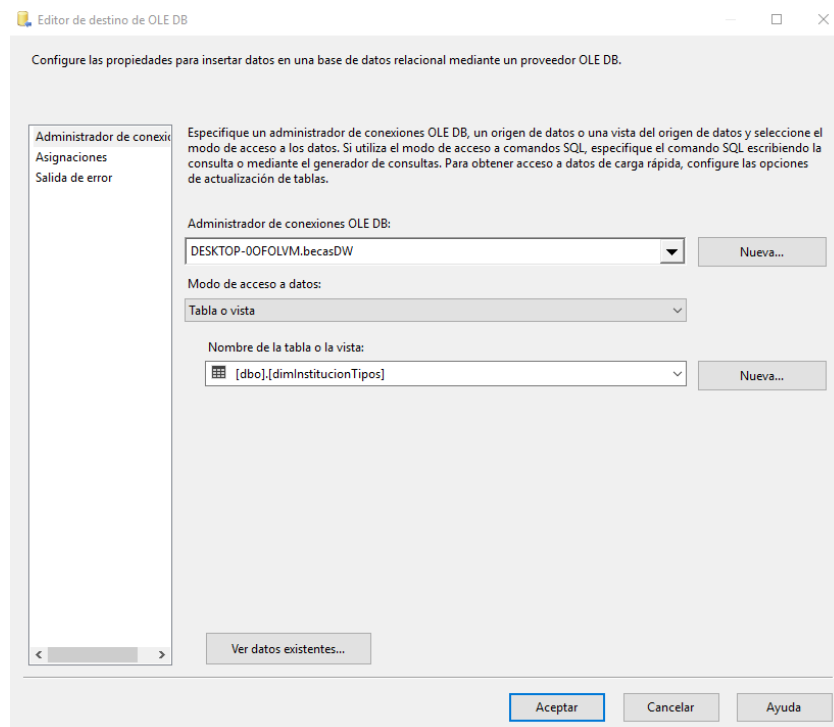


Ilustración 110: ETL - Editor de destino de datos. Elaboración propia

Actualización Diaria

En las actualizaciones diarias se agregará una Tarea para ejecutar una consulta SQL, en el cual se borrarán ciertos registros de la tabla de hechos factEvaluaciones y luego una tarea de flujo de datos para cargarla.

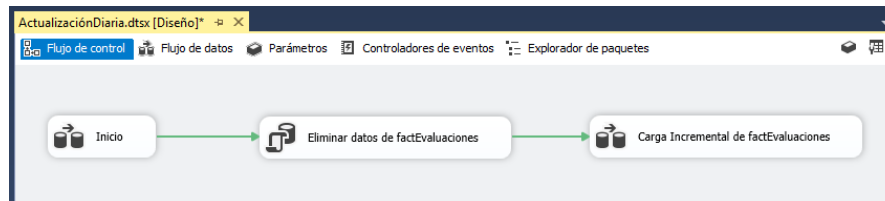


Ilustración 111: ETL - Actualización Diaria. Elaboración propia

En la tarea de eliminar datos de factEvaluaciones se especificarán los siguientes parámetros:

- **Connection:** se especifican el servidor y a la base de datos en el cual impactara la consulta SQL.
- **SQLStatement:** se detalla la consulta que se ejecutara al llegar a esta tarea. La consulta se encuentra en la solución propuesta.
- **BypassPrepare:** Se indica si la tarea debe preparar la consulta antes de ejecutarla. Como esto no es necesario, la dejamos configurada en falso.

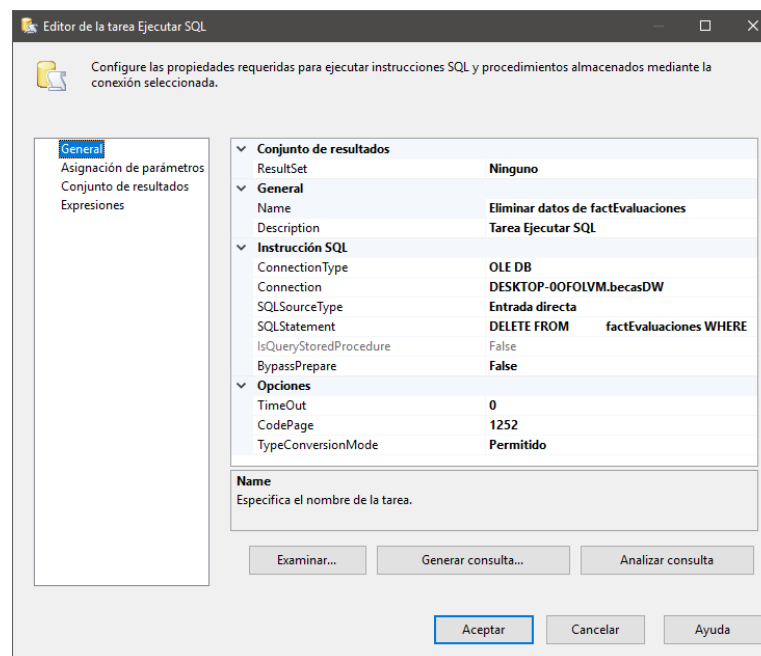


Ilustración 112: ETL - Editor de tarea Ejecutar SQL. Elaboración propia

Y luego de esta tarea se realiza el mismo flujo de datos realizado en la carga inicial con la diferencia en la condición en la consulta para obtener los datos de la fuente de datos.



Actualización Anual

Las actualizaciones anuales afectaran a las dimensiones dimFechas y dimCondicionTipos.



Ilustración 113: ETL - Actualización anual. Elaboración propia

Al igual que en la tarea de Eliminar datos de factEvaluaciones de las actualizaciones diarias se ejecutarán dos tareas para ejecutar SQL:

- Agregar nuevo año a Dimensión dimFechas: solo se agregará un nuevo año si es que no se encuentra ya agregado.
- Actualizar dimCondicionTipos: Se comparan las diferencias entre las tablas de la base de datos de origen y la base de datos del DW.



Programación de tareas con SQL Server Agent

Con el servicio de SQL Server Agent, podemos programar tareas administrativas de Integration Services y Analysis Services, para que se ejecuten en el orden y a las horas que necesitemos. Con este servicio vamos a automatizar la Actualización diaria y anual del proceso ETL, como así también el procesamiento del CUBO BECAS.

Antes de empezar a desarrollar los trabajos que ejecutaran estos paquetes, se deberá corroborar que el Servicio de SQL Server Agent se encuentre activo, con lo cual deberemos ir a la herramienta de Windows donde se encuentran listados los servicios del mismo y buscar el servicio “SQL Server Agent (Nombre de la instancia del motor de la base de datos)” y cambiar su estado para que se encuentre “En ejecución”.

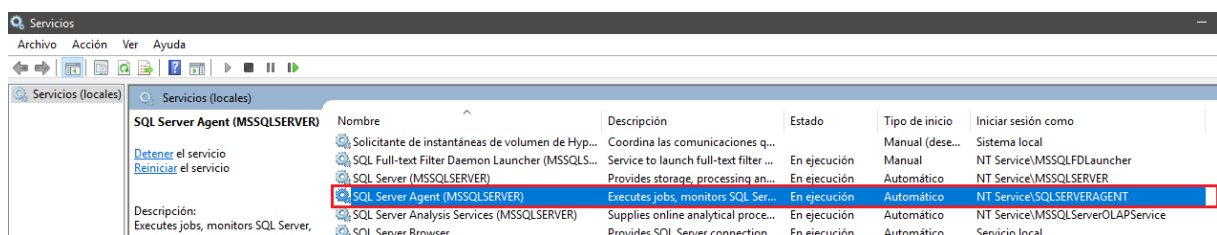


Ilustración 114: Servicio SQL Server Agent. Elaboración propia

También deberemos implementar un contenedor de base de datos para todos los paquetes implementados en el proyecto de Integration Services a través de la creación de un catálogo de SSIS. Para esto deberemos estar logueados con un usuario que tenga permisos de administrador para crear el catalogo.

En SQL Server Management Studio (SSMS) creamos el catalogo realizando clic derecho en la carpeta “Integration Services Catalogs” y seleccionando “Create Catalog”.

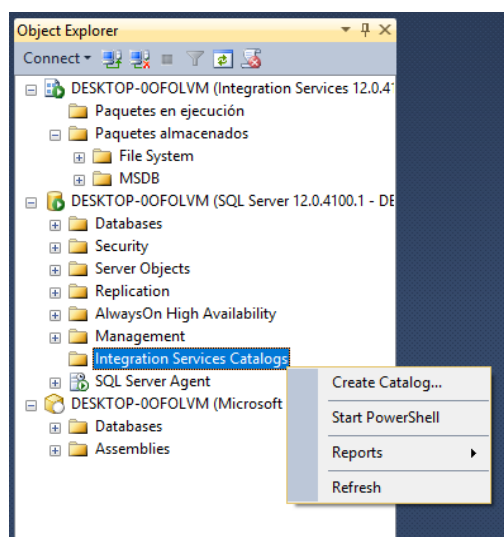


Ilustración 115: Integration Services Catalogs. Elaboración propia

La opción seleccionada abrirá la ventana para crear el catalogo. Algunas configuraciones son necesarias para realizar su creación. La Integración de CLR debe



habilitarse en esta instancia. En el inicio de SQL Server, la ejecución automática de procedimientos almacenados de SSIS no necesariamente requiere que se encuentre habilitada, pero para evitar errores la dejaremos habilitada. Lo más importante de este proceso, es proveer una contraseña segura para la gestión de este nuevo contenedor de base de datos.

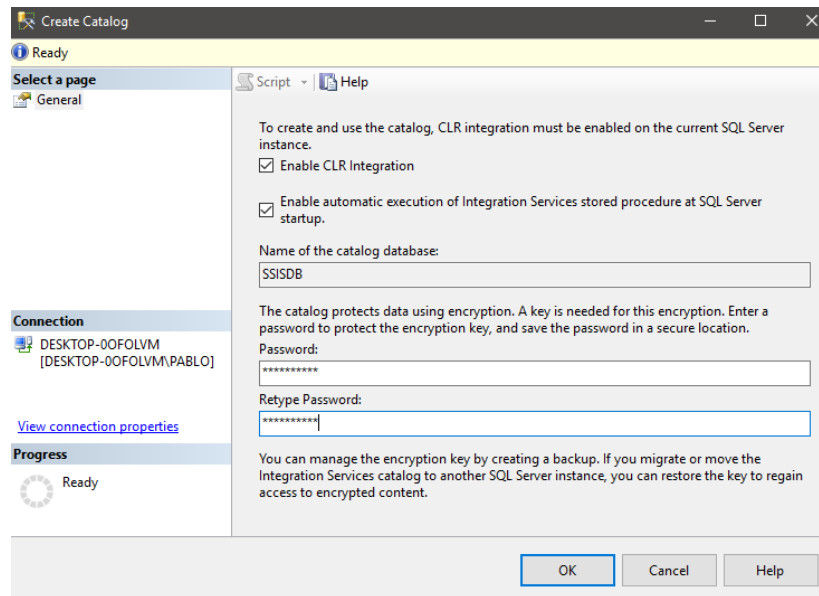


Ilustración 116: Creación catálogo SSISDB. Elaboración propia

Una vez realizada la configuración, una nueva base de datos se creará en esta instancia denominada SSISDB.

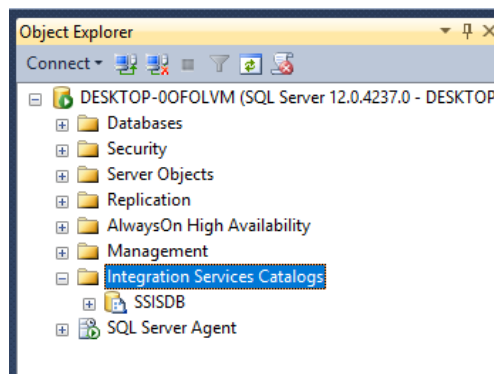


Ilustración 117: Catálogo SSISDB. Elaboración propia

Un usuario no necesita ser un miembro de SysAdmin para poder implementar proyectos (y paquetes) en el Catálogo. Existe una función de base de datos denominada `ssis_admin` en la base de datos SSISDB. Esta función ya contiene los permisos para implementar proyectos de SSDT.

La implementación del proyecto SSIS, junto con los paquetes del proyecto, se agregan a SQL Server, así como en el Catálogo SSIS. Esta organización de paquetes en un proyecto SSIS permite compartir objetos, propiedades y valores entre los paquetes de un proyecto.

El proyecto ETL desarrollado en Integration Services contiene varios paquetes. Para implementar el proyecto de manera completa con todos sus paquetes se realiza clic derecho en el proyecto en el explorado de soluciones de SSDTS.

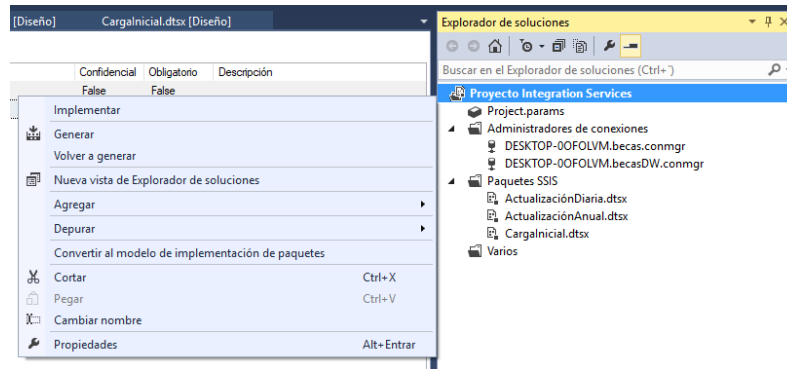


Ilustración 118: Implementar proyecto Integration Services. Elaboración propia

El asistente para la implementación de Integration Services mostrará una pantalla de bienvenida. El paso de “Seleccionar origen” predeterminará el proyecto seleccionado mientras estemos dentro del proyecto en SSDT. El asistente nos lleva directo a la página “Seleccionar destino”, pero podemos volver al paso anterior si es necesario y elegir otra fuente. En el “Seleccionar destino” se selecciona el servidor y la ruta del catálogo de SSIS.

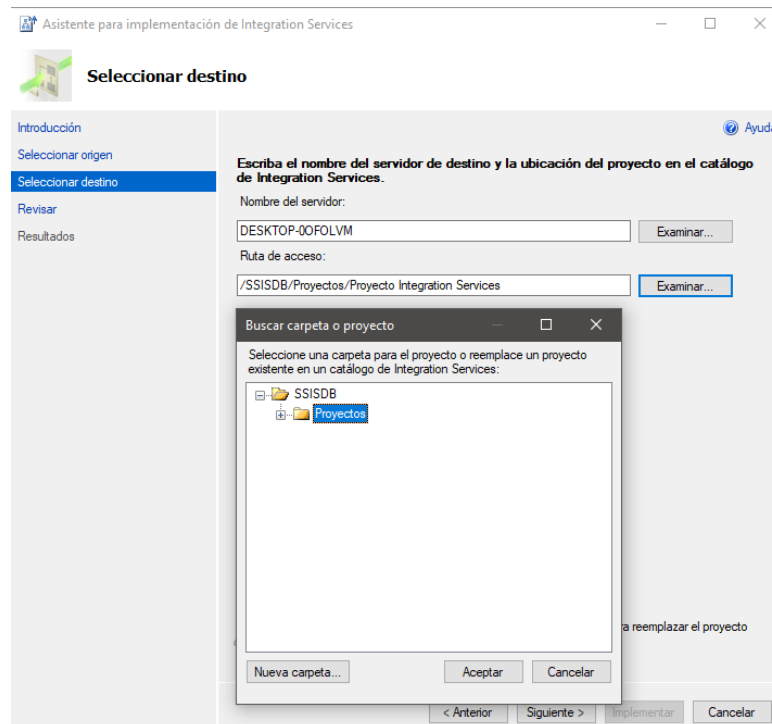


Ilustración 119: Destino implementación proyecto IS. Elaboración propia

El nombre del servidor es la instancia en la que se creó el Catálogo. La ruta puede ser una existente o se puede crear un Nuevo Camino. El mismo camino puede ser utilizado para diferentes proyectos. Todos los paquetes de proyectos estarán juntos bajo el proyecto de la ruta especificada.

La pantalla “Revisar” proporciona las opciones para ver el origen y el destino seleccionados. Una vez seleccionado el botón Implementar en la pantalla Revisar, la pantalla Resultados muestra la lista de éxitos o una indicación de los fallos durante la implementación.

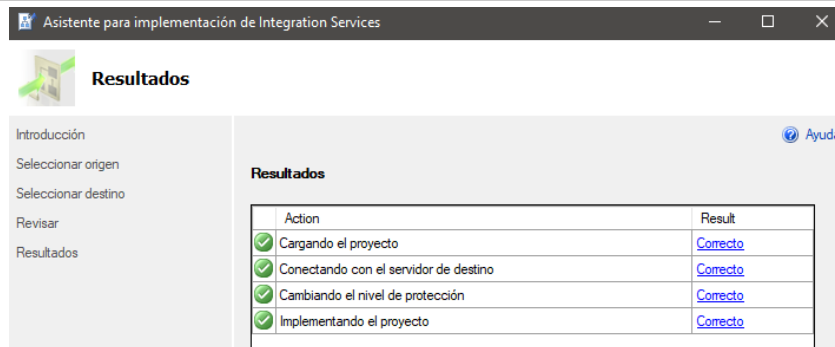


Ilustración 120: Resultados implementación proyecto IS. Elaboración propia

Un paquete puede ser lanzado de múltiples formas. En SQL Server Management Studio, la carpeta Integration Services Catalogs se puede expandir para ver las rutas creadas en el Catálogo, en nuestro proyecto podremos observar los paquetes de Actualización Anual, Actualización Diaria y la Carga Inicial.

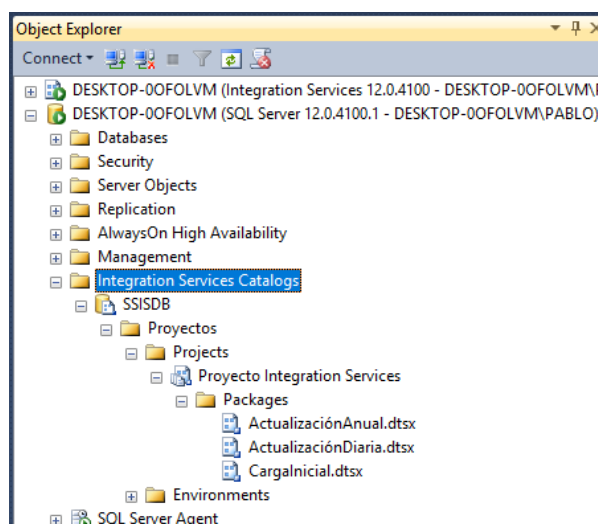


Ilustración 121: Paquetes SSISDB. Elaboración propia

Ahora procediendo con el servicio de SQL Server Agent vamos a crear un trabajo por cada tarea automatizada que necesitemos, de los cuales serán:

- Actualización anual del DW
- Actualización diaria del DW
- Procesamiento diario del cubo

Continuando con las actualizaciones del DW crearemos un trabajo denominado Actualización Anual DW realizando clic derecho sobre la carpeta Jobs de SQL Server Agent.

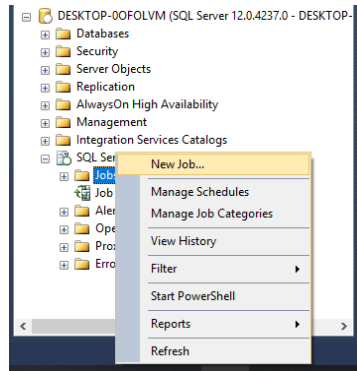


Ilustración 122: SSA crear nuevo trabajo. Elaboración propia

Lo que nos llevara a un asistente para crear el trabajo en el cual nos configura por defecto como propietario al usuario que se encuentra logueado en ese momento.

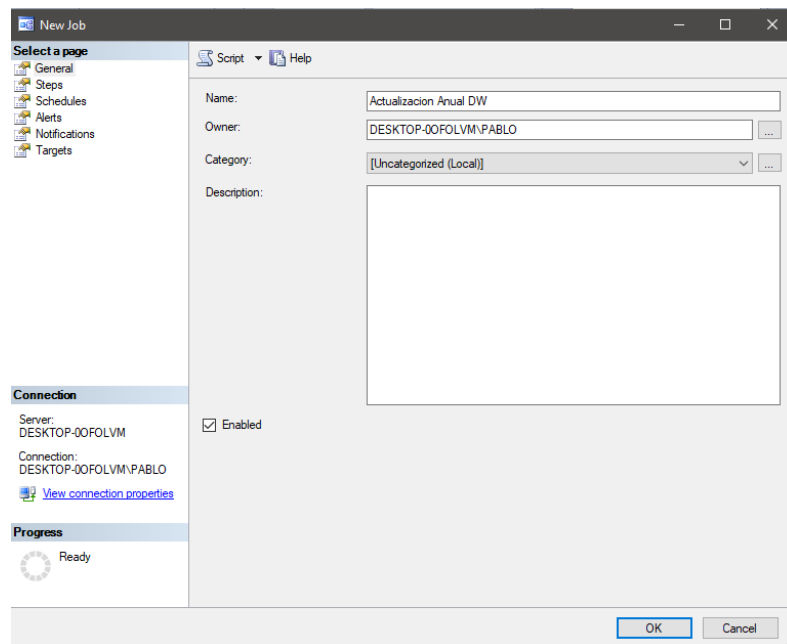


Ilustración 123: SSA Nuevo trabajo. Elaboración propia

En “Select a page” nos vamos a “Steps” en el cual agregaremos un nuevo paso al cual denominaremos Procesar actualización anual DW. Configuramos para que el paso tenga un tipo SQL Server Integration Package y la ejecutaremos como un usuario de SQL Server Agent Service Account. El Package Source será un catálogo de SSIS que creamos anteriormente, el servidor será donde se encuentra el mismo y deberemos elegir el paquete que vamos a ejecutar en este proceso.

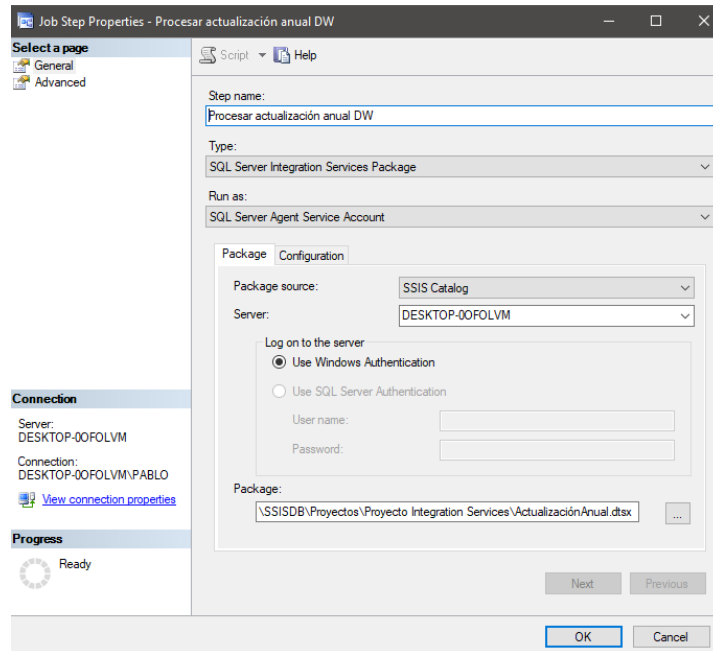


Ilustración 124: SSA Nuevo paso actualización anual DW. Elaboración propia

Una vez agregado este paso nos iremos a “Schedules” en “Select a page” para agregar una nueva programación de este trabajo en el cual se configura en que momentos precisos se ejecutara.

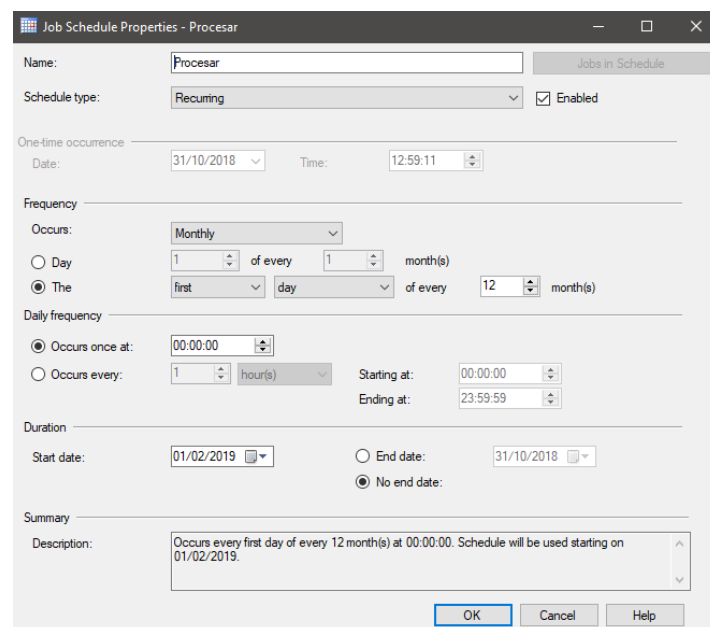


Ilustración 125: SSA nueva programación actualización anual DW. Elaboración propia

Configuración de la programación del trabajo:

- Nombre: necesitara un nombre.
- Schedule type: Recurring (Periódico).
- Frequency: Con qué frecuencia se va a ejecutar.
- Daily Frequency: En que horarios del día se va a ejecutar.



- Duration: Durante cuánto tiempo se va a ejecutar el trabajo, un día de inicio y un día de finalización.
- Summary: Resumen de las opciones elegidas.
 - En el caso de este trabajo de Actualización Anual del DW se ejecutará, Cada primer día de cada 12 meses a 0:00:00 Hs., iniciando el primero de febrero del 2019.

Con este finalizaremos la creación de este trabajo. De la misma manera realizaremos la creación del trabajo Actualización Diaria DW, con la diferencia que seleccionamos el paquete ActualizaciónDiaria del Catálogo del SSISDB.

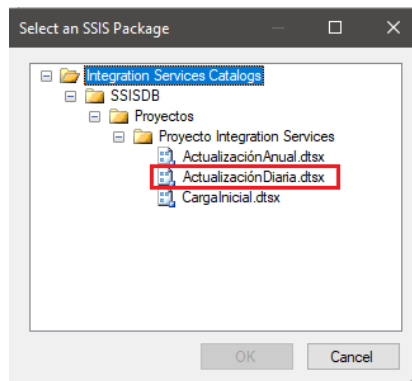


Ilustración 126: SSA paquetes SSIS. Elaboración propia

Y también la programación del trabajo se ejecutará en otros momentos específicos.

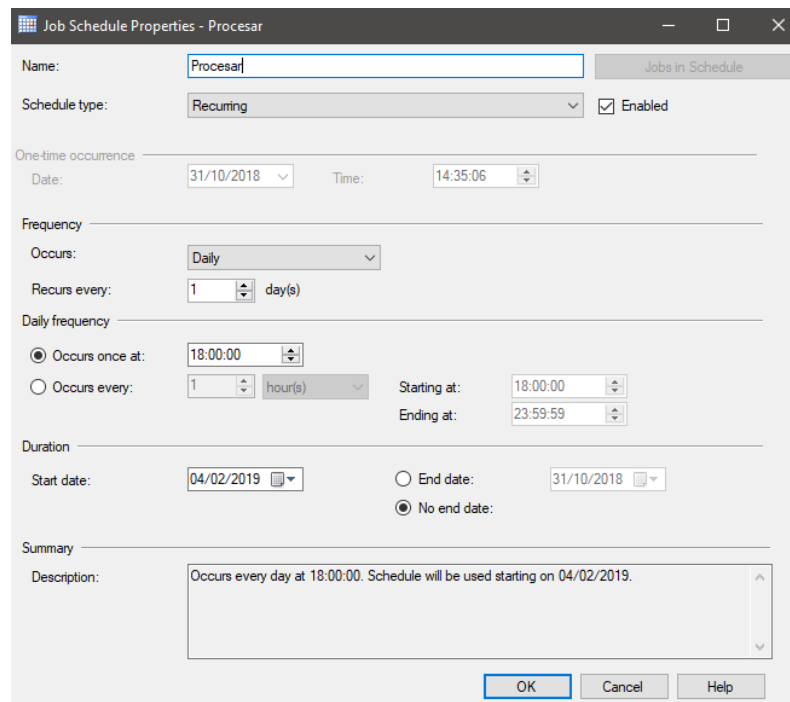


Ilustración 127: SSA nueva programación actualización diaria DW. Elaboración propia

Summary: Se ejecutará cada día a 18:00:00 Hs. Iniciando el lunes 04 de febrero de 2019 siendo el primer día de actividades del programa de Becas.

Con estos trabajos creados tendremos completa las actualizaciones del DW que garantizarán la consistencia de los datos del mismo, y permitirán proceder con la automatización del procesamiento del cubo.

Antes de crear el trabajo para la automatización del procesamiento del cubo deberemos asegurarnos que existen credenciales que tengan los permisos suficientes para poder ejecutar un trabajo de Analysis Services.

Para este proyecto creamos la credencial CREDENCIAL_DB_SSAS con los permisos del usuario administrador del equipo en el cual se corre el proyecto.

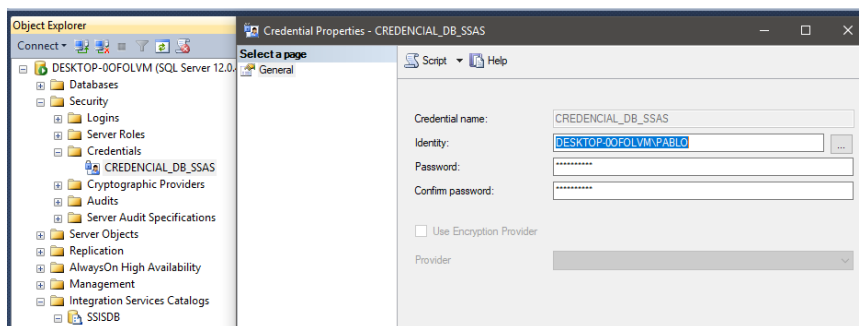


Ilustración 128: SSMS Credenciales. Elaboración propia

Con esto realizado, crearemos un nuevo Proxy dentro de la carpeta Proxies de SQL Server Agent que tenga las credenciales necesarias para este trabajo y que se encuentre activo a subsistema de SQL Server Analysis Services Command.

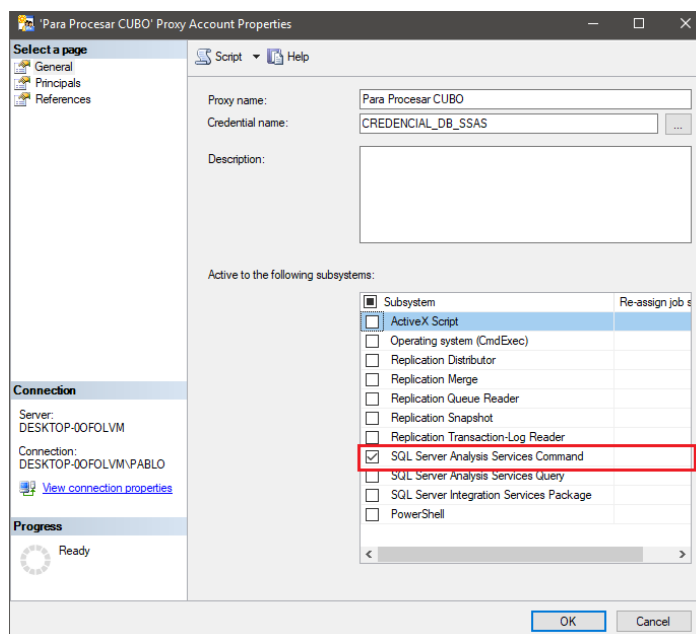


Ilustración 129: SSA Proxy. Elaboración propia

Nos conectaremos al servidor de Analysis Services, deberemos hacer clic derecho y elegir la opción Process en el Proyecto Analysis Services que se creó al realizar el primer procesamiento del CUBO BECAS cuando se construyó el mismo. Lo que nos llevara a un cuadro de dialogo y deberemos seleccionar la opción Script.

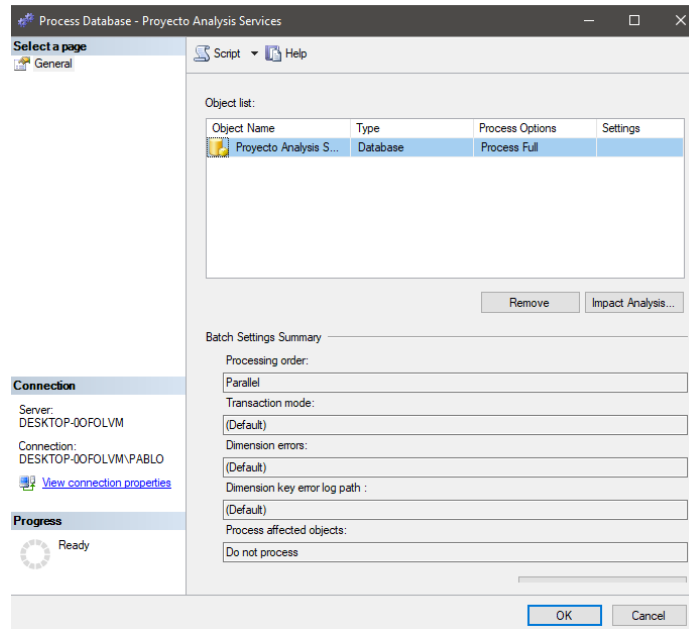


Ilustración 130: SSAS Proceso Proyecto Analysis Services. Elaboración propia

Este paso abre una ventana de Consulta XML, el cual contiene el script XMLA que procesa el proyecto. Copiaremos este Script y nos dirigiremos a SQL Server Agent para proceder a crear el trabajo que ejecutara este Script.

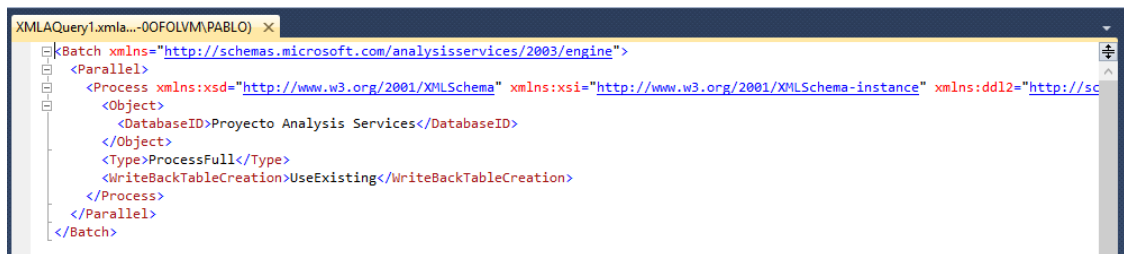


Ilustración 131: SSAS Script proceso proyecto Analysis Services. Elaboración propia

De la misma manera que realizamos la creación de los otros trabajos anteriores deberemos elegir un propietario del trabajo, darle un nombre, crear un paso que ejecute el script y agregar una programación el cual determinara cuando se ejecutara.

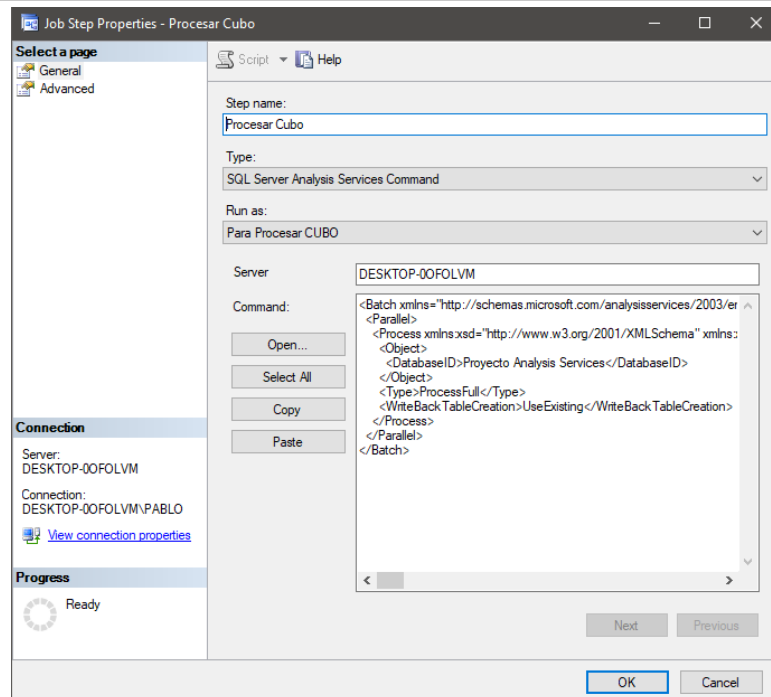


Ilustración 132: SSA nuevo paso actualización CUBO. Elaboración propia

En las propiedades del paso elegiremos tipo de trabajo que se ejecutará siendo de tipo SQL Server Analysis Services Command, ejecutándose como el Proxy creado que tendrá los permisos para llevar a cabo el proceso del cubo. El servidor como siempre será en el cual se encuentra el Proyecto de Analysis Services y el comando será el Script generado anteriormente.

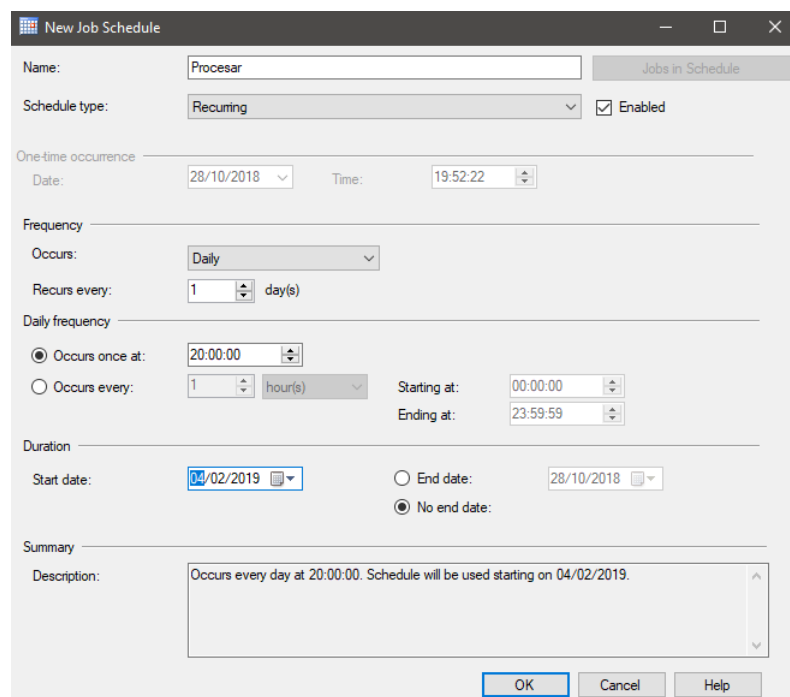


Ilustración 133: SSA programación actualización CUBO. Elaboración propia

En la programación del trabajo tendrá la siguiente configuración:

- Summary: Se ejecutará cada día a 20:00:00 Hs. Iniciando el lunes 04 de febrero de 2019.