



**UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO**

**ESCUELA DE POSGRADO**



**MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN GERENCIA  
DE TECNOLOGÍAS DE LA INFORMACIÓN Y GESTIÓN DEL SOFTWARE**

**IMPLEMENTACIÓN DEL MODELO DE ANÁLISIS DE SENSIBILIDAD PARA  
DETERMINAR LOS PERFILES FINANCIEROS DE LOS ESTUDIANTES EN LOS  
DIFERENTES PROGRAMAS DE MAESTRÍA Y DOCTORADO DE LA ESCUELA  
DE POSGRADO UTILIZANDO BUSINESS INTELLIGENCE**

**TESIS**

**TESIS PARA OPTAR EL GRADO DE MAESTRO EN INGENIERÍA DE  
SISTEMAS CON MENCIÓN EN GERENCIA DE TECNOLOGÍAS DE LA  
INFORMACIÓN Y GESTIÓN DEL SOFTWARE**

**PRESENTADO POR:**

**ING. LUIS ALBERTO MIJA CAMARGO**

**Lambayeque, Setiembre 2018**

Implementación del Modelo de Análisis de sensibilidad para determinar los perfiles financieros de los estudiantes en los diferentes programas de Maestría y Doctorado de la Escuela de Posgrado utilizando Business Intelligence.

---

Ing. Luis Alberto Mija Camargo  
Autor

---

Ing. Mg. Ernesto Karlo Celi Arévalo  
Asesor

Presentada a la Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo para optar el Grado de MAESTRO EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN GERENCIA DE TECNOLOGÍAS DE LA INFORMACIÓN Y GESTIÓN DEL SOFTWARE.

APROBADO POR:

---

M.Sc. Edward Ronald Haro Maldonado  
Presidente del Jurado

---

M.Sc. Gilberto Carrión Barco  
Secretario del Jurado

---

M.Sc. Percy Celis Bravo  
Vocal del Jurado

**Lambayeque, Setiembre 2018**

## DEDICATORIA

A Dios, Que durante toda mi vida me acompaña,  
y que no me deja solo, siempre estas presente, gracias  
Señor por todo lo que me has dado y me das.

A mis padres Edith y Francisco mis mejores amigos,  
por su apoyo y comprensión constante, ustedes son los  
pilares de mi vida, siempre los llevo en mi corazón y los  
tengo en mi mente, los amo.

A mis maestros y amigos por las veces que pedía  
ayuda y comprensión durante mis estudios de  
Pregrado, Postgrado y compañeros de trabajo que  
de alguna forma me ayudaron con sus ideas gracias.

A mi hijo, Sebastián que es la razón de mi vida el  
impulso en todas las cosas que hago para poder  
esforzarme todos los días, siempre estás en mi  
mente y en mi corazón.

## **AGRADECIMIENTOS**

A Dios por darme la vida y poner en mi camino a personas tan maravillosas.

A mi familia por su apoyo desinteresado, en cada parte de mi formación como persona.

A mi Asesor de tesis el Ing. Ernesto Karlo Celi Arévalo, por su paciencia y comprensión, por ser el guía que me encamino a conseguir mi objetivo.

A cada uno de mis amigos y compañeros de estudio, por su cordial amistad, por su estima, y por los momentos vividos; a cada uno de ellos me llevo un grato recuerdo.

## ÍNDICE

<b>CAPÍTULO I: ANÁLISIS DEL OBJETO DE ESTUDIO .....</b>	<b>15</b>
<b>1.1. Ubicación.....</b>	<b>16</b>
<b>1.2. Planteamiento del Problema .....</b>	<b>17</b>
<b>1.3. Formulación del Problema .....</b>	<b>17</b>
<b>1.4. Objetivos .....</b>	<b>18</b>
1.4.1. Objetivo General .....	18
1.4.2. Objetivos Específicos.....	18
<b>CAPÍTULO II: MARCO TEÓRICO .....</b>	<b>19</b>
<b>2.1. Antecedentes .....</b>	<b>20</b>
<b>2.2 Fundamentos Teóricos.....</b>	<b>22</b>
2.2.1. Perfil financiero.....	22
2.2.2. Inteligencia de Negocios (Business Intelligence).....	22
2.2.3. Metadatos .....	24
2.2.4. Data Warehouse .....	25
2.2.5. DataMart.....	26
2.2.6. Datamart OLAP.....	27
2.2.7. Datamart OLTP .....	27
2.2.8. Metodología de Bill Inmon .....	28
2.2.9. Metodología de Ralph Kimball .....	30
2.2.9.1. Ciclo de Vida Ralph Kimball .....	31
2.2.9.2. Comparación y selección de la Metodología.....	35
2.2.10. Gestor de base de datos MySQL .....	36
2.2.10.1. Herramienta de Administración phpMyAdmin.....	36
2.2.11. Ventajas y Desventajas de MySQL.....	37
2.2.12. Pentaho BI Suite.....	38
2.2.13. Servidor OLAP.....	41
2.2.14. Data Mining.....	42
2.2.15. ETL .....	43
2.2.16. Pentaho Report Designer.....	44
<b>CAPÍTULO III: MARCO METODOLÓGICO.....</b>	<b>47</b>

3.1 Tipo y Diseño de la Investigación.....	48
3.1.1 Tipo de Investigación .....	48
3.2 Variables .....	48
3.2.1 Variable Independiente .....	48
3.2.2 Variable Dependiente .....	48
<b>CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS.....</b>	<b>49</b>
<b>4.1 Metodología .....</b>	<b>50</b>
4.1.1. Diseño de la Arquitectura.....	50
4.1.1.1. Requerimientos del negocio .....	50
4.1.1.2. Diseño de la Arquitectura del Modelo .....	52
4.1.1.3. Instalación y Configuración del Producto .....	53
<b>4.2. Diseño de modelamiento dimensional.....</b>	<b>54</b>
4.2.1 Modelamiento Dimensional .....	54
4.2.2 Diseño Físico.....	58
4.2.3 Creación de Cubos .....	60
4.2.4 Diseño de ETL .....	69
<b>4.3. Definición de los procedimientos de mantenimiento y crecimiento del Data Mart .....</b>	<b>80</b>
4.3.1. Despliegue.....	80
4.3.2. Mantenimiento .....	94
<b>4.4. Interpretación de los resultados obtenidos del Data Mart para encontrar el perfil financiero adecuado.....</b>	<b>95</b>
<b>4.5. Discusión de Resultados .....</b>	<b>115</b>
<b>CAPITULO V: CONCLUSIONES Y RECOMENDACIONES .....</b>	<b>117</b>
5.1. Conclusiones .....	118
5.2. Recomendaciones.....	119
<b>CAPÍTULO VI: REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>120</b>
<b>ANEXOS .....</b>	<b>123</b>
Anexo 1: Diseño de Extracción.....	124
Anexo 2: Instalación de la Plataforma Pentaho 7.1.0.0-12 .....	126
Anexo 3: Instalación de Schema Workbench Pentaho.....	130
Anexo 4: Instalación de Pentaho Designer .....	131

## ÍNDICE DE FIGURAS

<b>Figura NRO. 1: Organigrama de la Escuela de Postgrado de la UNPRG .....</b>	<b>16</b>
<b>Figura NRO. 2: Business Intelligence.....</b>	<b>23</b>
<b>Figura NRO. 3: Metadatos .....</b>	<b>24</b>
<b>Figura NRO. 4: Representación del DataWarehouse .....</b>	<b>25</b>
<b>Figura NRO. 5: Representación del DataMart.....</b>	<b>26</b>
<b>Figura NRO. 6: DataMart .....</b>	<b>27</b>
<b>Figura NRO. 7: Metodología Bill Inmon.....</b>	<b>29</b>
<b>Figura NRO. 8: Marco Metodológico - Metodología de Ralph Kimball.....</b>	<b>31</b>
<b>Figura NRO. 9: Metodología de Ralph Kimball .....</b>	<b>32</b>
<b>Figura NRO. 10: Herramienta de Administración phpMyAdmin .....</b>	<b>37</b>
<b>Figura NRO. 11: Arquitectura de una solución Business Intelligence .....</b>	<b>38</b>
<b>Figura NRO. 12: Pentaho Reporting.....</b>	<b>39</b>
<b>Figura NRO. 13: Pentaho Dashboards.....</b>	<b>40</b>
<b>Figura NRO. 14: Pentaho Open BI Suite.....</b>	<b>41</b>
<b>Figura NRO. 15: Análisis OLAP .....</b>	<b>42</b>
<b>Figura NRO. 16: Data Mining .....</b>	<b>43</b>
<b>Figura NRO. 17: ETL Kettle.....</b>	<b>44</b>
<b>Figura NRO. 18: Pentaho Report Designer .....</b>	<b>45</b>
<b>Figura NRO. 19: Ejemplo Pentaho Report Designer.....</b>	<b>46</b>
<b>Figura NRO. 20: Diseño de la Arquitectura .....</b>	<b>52</b>
<b>Figura NRO. 21: Modelo estrella usado en el registro de transacciones.....</b>	<b>55</b>
<b>Figura NRO. 22: Modelo estrella usando como cuadro estadístico semanal .....</b>	<b>55</b>
<b>Figura NRO. 23: Diseño físico de Base de Datos transaccional .....</b>	<b>59</b>
<b>Figura NRO. 24: Diseño físico de Base de Datos estadístico .....</b>	<b>59</b>
<b>Figura NRO. 25: Ejecutando la herramienta schema Workbench portable a traves del archivo workbench.bat, descargado de la pagina de pentaho.....</b>	<b>60</b>
<b>Figura NRO. 26: Menú opción para la conexión con la base de datos MySQL .....</b>	<b>61</b>
<b>Figura NRO. 27: Configurando conexión con MySQL .....</b>	<b>61</b>
<b>Figura NRO. 28: Crear nuevo esquema a través del menú File .....</b>	<b>62</b>
<b>Figura NRO. 29: Creación del nuevo cubo .....</b>	<b>62</b>
<b>Figura NRO. 30: Aquí se realiza la asignando de nombre a nuestro cubo .....</b>	<b>63</b>
<b>Figura NRO. 31: Agregar nueva tabla de hechos fact_operacion a nuestro cubo .....</b>	<b>63</b>
<b>Figura NRO. 32: Seleccionar el nombre de la tabla de hechos, de una tabla de hechos creada .....</b>	<b>64</b>
<b>Figura NRO. 33: Agregando la dimensión a nuestro cubo perfil .....</b>	<b>64</b>
<b>Figura NRO. 34: Asignar nombre a la dimensión.....</b>	<b>65</b>
<b>Figura NRO. 35: Agregando una tabla a la dimensión correspondiente .....</b>	<b>65</b>
<b>Figura NRO. 36: Seleccionando nombre a la tabla asignada a la dimensión Alumno .....</b>	<b>66</b>
<b>Figura NRO. 37: Asignar PRIMARY KEY en la jerarquía .....</b>	<b>66</b>
<b>Figura NRO. 38: Asignamos la ForeignKey en la dimensión Alumno.....</b>	<b>67</b>
<b>Figura NRO. 39: Agregamos los niveles a la Jerarquía.....</b>	<b>67</b>

<b>Figura NRO. 40: Asignar el tipo de dato .....</b>	<b>68</b>
<b>Figura NRO. 41: Agregar las medidas que se utilizaran en los cálculos del cubo.....</b>	<b>68</b>
<b>Figura NRO. 42: Asignar un nombre a la medida .....</b>	<b>69</b>
<b>Figura NRO. 43: Creación de una nueva transformación.....</b>	<b>69</b>
<b>Figura NRO. 44: Creación de una nueva conexión a la base .....</b>	<b>70</b>
<b>Figura NRO. 45: Prueba de la conexión a la base de datos MySQL .....</b>	<b>70</b>
<b>Figura NRO. 46: Prueba de la conexión a la base de datos MSSQL.....</b>	<b>71</b>
<b>Figura NRO. 47: CSV file input.....</b>	<b>71</b>
<b>Figura NRO. 48: Selección del archivo CSV .....</b>	<b>72</b>
<b>Figura NRO. 49: En este paso se configura la tabla de destino, hacia donde irán los datos que se importaran.....</b>	<b>73</b>
<b>Figura NRO. 50: A continuación se muestra como se importa la información hacia su tabla de destino.....</b>	<b>73</b>
<b>Figura NRO. 51: Diseño de la forma como se importara la información deseada.....</b>	<b>74</b>
<b>Figura NRO. 52: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración.....</b>	<b>74</b>
<b>Figura NRO. 53: Tabla de destino lista y preparada para recibir la data deseada .....</b>	<b>75</b>
<b>Figura NRO. 54: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim_partida .....</b>	<b>75</b>
<b>Figura NRO. 55: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración.....</b>	<b>76</b>
<b>Figura NRO. 56: Tabla de destino lista y preparada para recibir la data deseada .....</b>	<b>76</b>
<b>Figura NRO. 57: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim_programa.....</b>	<b>77</b>
<b>Figura NRO. 58: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración.....</b>	<b>77</b>
<b>Figura NRO. 59: Tabla de destino lista y preparada para recibir la data deseada .....</b>	<b>78</b>
<b>Figura NRO. 60: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim_semestre.....</b>	<b>78</b>
<b>Figura NRO. 61: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración.....</b>	<b>79</b>
<b>Figura NRO. 62: Tabla de destino lista y preparada para recibir la data deseada .....</b>	<b>79</b>
<b>Figura NRO. 63: Diseño del Data Mart para la migracion de Base de Datos operativa de la informacion deseada .....</b>	<b>80</b>
<b>Figura NRO. 64: Aquí podemos ver la consola de administración y la url de la herramienta Pentaho Server desde donde se ingresaran los datos necesarios para iniciar sesión <a href="http://localhost:8080/Login">http://localhost:8080/Login</a>.....</b>	<b>85</b>
<b>Figura NRO. 65: Lista de usuarios que administraran el aplicativo Web.....</b>	<b>86</b>
<b>Figura NRO. 66: Roles que serán asignados según los privilegios a cada usuario.....</b>	<b>86</b>
<b>Figura NRO. 67: Conexión a la Base de datos del Data Mart desde donde se crear el cubo el los reportes.....</b>	<b>87</b>
<b>Figura NRO. 68: Consulta de datos en la dimensión tiempo .....</b>	<b>89</b>
<b>Figura NRO. 69: Consulta de datos de la dimensión programa .....</b>	<b>89</b>
<b>Figura NRO. 70: Consulta de datos de la dimensión alumno .....</b>	<b>89</b>



<b>Figura NRO. 71: Consulta de datos de la dimensión semestre .....</b>	<b>90</b>
<b>Figura NRO. 72: Consulta de datos de la dimensión forma de pago.....</b>	<b>90</b>
<b>Figura NRO. 73: Consulta de datos de la dimensión partida.....</b>	<b>91</b>
<b>Figura NRO. 74: Consulta de datos de la tabla de hechos operación.....</b>	<b>91</b>
<b>Figura NRO. 75: Consulta de datos de la tabla de hechos cuadro estadístico .....</b>	<b>92</b>
<b>Figura NRO. 76: Establecer parámetros de conexión .....</b>	<b>92</b>
<b>Figura NRO. 77: Establecer el origen de datos de la conexión .....</b>	<b>92</b>
<b>Figura NRO. 78: Establecer la consulta SQL para obtener el reporte deseado (listado de importes por mes de los últimos 4 años).....</b>	<b>93</b>
<b>Figura NRO. 79: Manipulación de los campos para su respectiva visualización en el JPivot .</b>	<b>93</b>
<b>Figura NRO. 80: Visualizar nuestro reporte .....</b>	<b>93</b>
<b>Figura NRO. 81: Visualizar los datos elegidos en la creación del reporte (listar de importes por meses de los últimos 4 años) .....</b>	<b>94</b>
<b>Figura NRO. 82: Postulantes por procesos de admisión .....</b>	<b>95</b>
<b>Figura NRO. 83: Cantidad de alumnos por sexo de los proceso de admisión 2015-I, 2017-I y 2017-II .....</b>	<b>96</b>
<b>Figura NRO. 84: Reporte de ingresos por medio de pago del año 2017 y 2018.....</b>	<b>97</b>
<b>Figura NRO. 85: Reporte de ingresos por mes de los últimos 4 años.....</b>	<b>97</b>
<b>Figura NRO. 86: Reporte de ingresos de los últimos 4 años .....</b>	<b>98</b>
<b>Figura NRO. 87: Reporte de ingresos por partidas de los últimos 4 años .....</b>	<b>98</b>
<b>Figura NRO. 88: Reporte de ingresos por partidas en los últimos 4 años .....</b>	<b>99</b>
<b>Figura NRO. 89: Reporte de ingresos por programas del admisión 2017-I .....</b>	<b>100</b>
<b>Figura NRO. 90: Cantidad de alumnos por sexo de los programas de admisión 2017-I.....</b>	<b>101</b>
<b>Figura NRO. 91: Reporte de ingresos por forma de pago de los programas de admisión 2017-I .....</b>	<b>101</b>
<b>Figura NRO. 92: Reporte de ingresos por forma de pago de los programas de admisión 2017-II.....</b>	<b>103</b>
<b>Figura NRO. 93: Reporte de ingresos por programas de admisión 2017-II de los conceptos matrícula y pensión de enseñanza.....</b>	<b>103</b>
<b>Figura NRO. 94: Reporte de ingresos por mes de los programas de admisión 2017-II por conceptos matrícula y pensión de enseñanza .....</b>	<b>104</b>
<b>Figura NRO. 95: Reporte del rango de edades de los alumnos del proceso de admisión 2017-I Y 2017-II .....</b>	<b>105</b>
<b>Figura NRO. 96: Cantidad de alumnos por universidad de procedencia del proceso de admisión 2017-I y 2017-II.....</b>	<b>105</b>

## ÍNDICE DE TABLAS

<b>Tabla 1: Análisis comparativo de las Metodologías .....</b>	<b>35</b>
<b>Tabla 2: Requerimientos del Negocio .....</b>	<b>51</b>
<b>Tabla 3: Descripción comparativa entre diferentes Tecnologías .....</b>	<b>53</b>
<b>Tabla 4: Requerimientos de Hardware .....</b>	<b>81</b>
<b>Tabla 5: Requerimientos de Software .....</b>	<b>82</b>
<b>Tabla 6: Plantilla del Caso de Prueba .....</b>	<b>83</b>
<b>Tabla 7: Caso de uso de prueba manipulación de la información .....</b>	<b>84</b>
<b>Tabla 8: Pruebas de seguridad.....</b>	<b>88</b>
<b>Tabla 9: Cantidad de postulantes por proceso de admisión.....</b>	<b>95</b>
<b>Tabla 10: Resumen de proporciones por sexo de los procesos admisión 2015-I, 2017-I y 2017-II.....</b>	<b>96</b>
<b>Tabla 11: Importes por medio de pago del año 2017 y 2018 .....</b>	<b>96</b>
<b>Tabla 12: Lista donde se muestran los ingresos por año .....</b>	<b>98</b>
<b>Tabla 13: Aquí se muestran los importes por partidas en los últimos 4 años.....</b>	<b>99</b>
<b>Tabla 14: Listado de importes por programas de admision 2017-I.....</b>	<b>100</b>
<b>Tabla 15: Aquí se muestran los ingresos por forma de pago de los programas de admisión 2017-II .....</b>	<b>102</b>
<b>Tabla 16: Rango de las edades de los alumnos del proceso de admisión 2017-I y 2017-II .....</b>	<b>104</b>
<b>Tabla 17: Cantidad de alumnos por universidad de procedencia del proceso de admisión 2017-I y 2017-II .....</b>	<b>105</b>
<b>Tabla 18: Cantidad de profesiones por tipo de universidad del proceso de admisión 2017-I</b>	<b>106</b>
<b>Tabla 19: Cantidad de profesiones por tipo de universidad del proceso de admisión 2017-II .....</b>	<b>107</b>
<b>Tabla 21: Cantidad de alumnos aptos admisión 2017-I.....</b>	<b>108</b>
<b>Tabla 22: Cantidad de alumnos aptos 2017-II.....</b>	<b>108</b>

## **RESUMEN**

La información almacenada en múltiples orígenes de datos es muy rica y abundante en conocimientos e ignorada por muchas entidades públicas y privadas, con la evolución de las herramientas tecnológicas y el Internet of Things (IOT) se están rompiendo los límites del conocimiento que crecen a pasos enormes y que nos proporcionan mucha información relacionada con finanzas, estadísticas, económicas, etc.

Actualmente las entidades utilizan la inteligencia de negocios para tratar de captar más clientes, aumentar sus ventas, reducir sus costos e invertir en maquinaria y equipos, modernizando de esta manera su infraestructura y aumentar su capital.

La Escuela de Posgrado como entidad del más alto nivel en la UNPRG, no tenía un modelo o usaba alguna metodología para administrar su información, la misma que era obtenida a través de sus procesos de admisión que durante muchos años brindaron ingresos muy importantes para el crecimiento y mejora de sus actividades académicas y procesos administrativos, de tal forma no se conocía quienes eran sus clientes más puntuales, y a quienes estar enfocados laboralmente o académicamente para brindarles el servicio.

Las entidades académicas públicas o privadas no están exentas o libres de este tema, por lo que necesitan utilizar herramientas de Inteligencia de Negocio(BI), para obtener datos de tipo académico como, cantidad de alumnos, egresados, ubicación geográfica, sexo, número de alumnos hábiles, alumnos retirados, etc. Y en el campo financiero conocer cuáles son los alumnos deudores quienes son los mejor clientes, etc.

Este presente trabajo de investigación propone un modelo y metodología, haciendo uso de la suite de Pentaho(ETL, Data Integración, Analisis Services, Pentaho Reporting), que le permiten conocer en tiempo real datos estadísticos, académicos y financieros que son de mucha ayuda en la toma de decisiones(gestión y análisis de la información), para poder invertir en varias de sus áreas como la Unidad de Tele-Educación, Unidad de Admisión, Unidad de Logística, Unidad de Contabilidad, Unidad de Asuntos Académicos, etc.

Como resultado del modelo propuesto se obtuvo los perfiles financieros deseados de acuerdo a las características recolectadas a través del registro web y de esta forma trabajar en tiempo real con el modelo propuesto y Metodología de Ralph Kimball.

**Palabras claves:** herramientas Business Intelligence, Pentaho BI, minería de datos.

## **ABSTRACT**

The information stored in multiple data sources is very rich and abundant in knowledge and ignored by many public and private entities, with the evolution of technological tools and the Internet of Things (IOT), the limits of knowledge that grow in steps are being broken. huge and that provide us with a lot of information related to finance, statistics, economics, etc.

Currently, entities use business intelligence to try to attract more customers, increase their sales, reduce their costs and invest in machinery and equipment, thus modernizing their infrastructure and increasing their capital.

The Graduate School as an entity of the highest level in UNPRG, did not have a model or used any methodology to manage their information, which was obtained through their admission processes that for many years provided very important income for growth and improvement of their academic activities and administrative processes, in such a way it was not known who their most punctual clients were, and to whom they were focused labor or academically to provide them with the service.

Public or private academic entities are not exempt or free of this topic, so they need to use Business Intelligence (BI) tools to obtain academic data such as number of students, graduates, geographic location, gender, number of Skilled students, retired students, etc. And in the financial field, know which debtor students are the best clients, etc.

This research work proposes a model and methodology, making use of the Pentaho suite (ETL, Data Integration, Analytics Services, Pentaho Reporting), which allow you to know in real time statistical, academic and financial data that are very helpful in making decisions (management and analysis of information), to be able to invest in several of its areas such as the Tele-Education Unit, Admission Unit, Logistics Unit, Accounting Unit, Academic Affairs Unit, etc.

As a result of the proposed model, the desired financial profiles were obtained according to the characteristics collected through the web registry and in this way to work in real time with the proposed model and Ralph Kimball's Methodology.

Keywords: Business Intelligence tools, Pentaho BI, data mining.

## INTRODUCCIÓN

Desde sus inicios la Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo ha utilizado procesos de tipo manual, para el registro de inscripción de sus procesos de admisión, lo que no permite obtener de forma oportuna, reportes presupuestales y tomar decisiones rápidamente. Es ahí donde nace la idea de utilizar inteligencia de negocios para agilizar los procesos manuales y obtener perfiles idóneos que ayuden a seleccionar, crear y direccionar, programas de maestría y doctorado hacia personas idóneas que contribuyan con el fortalecimiento de sus habilidades, desempeño laboral y académico.

Actualmente la EPG no utiliza estas herramientas BI que le ayuden a satisfacer las necesidades de información (clasificar, seleccionar, agrupar) en tiempo real.

Actualmente y durante el desarrollo del proyecto se utilizó como herramientas la suite de Pentaho Data Integration, Pentaho Dashboard Designer, Pentaho Report Designer, Pentaho Schema Workbench y como Metodología la de Ralph Kimball. Que por la experiencia y demanda de estas herramientas en el mercado internacional y su metodología, hacen de este proyecto una buena alternativa a utilizar.

Una de las muchas limitaciones que se obtuvo durante el desarrollo de la propuesta fue los ítems o campos necesarios para obtener los reportes deseados es decir que no se consideraban algunos datos y quedaban archivados en papel, datos que podrían proporcionar información valiosa, por este motivo se vio en la obligación de crear más campos que permitieran obtener los perfiles deseados y otro reportes relacionados a situaciones académicas y financieras

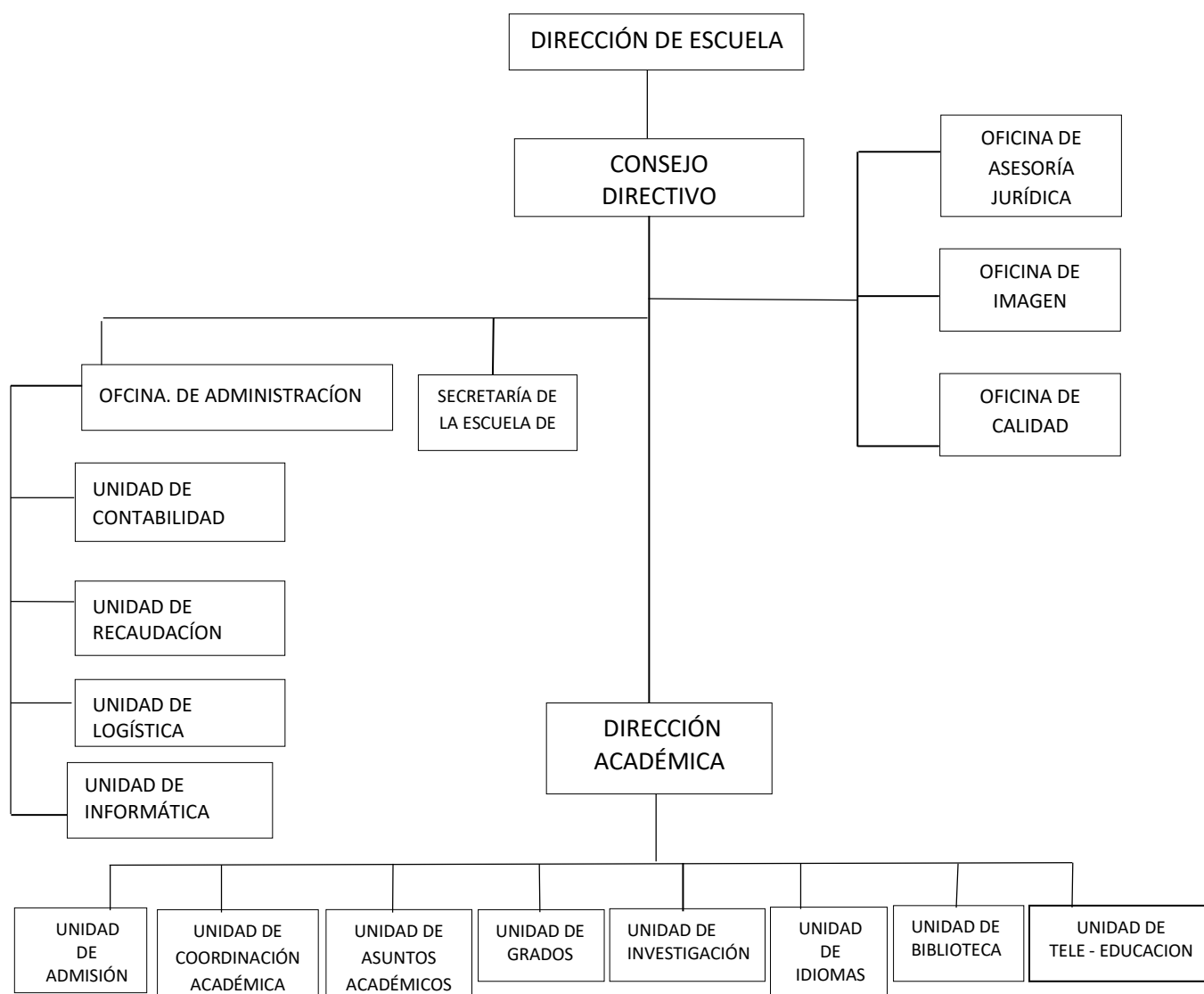
En este presente proyecto se abordan temas del capítulo 1 como el análisis del objeto de estudio donde se realiza la ejecución del proyecto, en el capítulo 2 como marco teórico se describen todos los conceptos metodologías y herramientas a utilizar, en el capítulo 3 el marco metodológico el desarrollo de la propuesta con su metodología.

# **CAPÍTULO I: ANÁLISIS DEL OBJETO DE ESTUDIO**

## 1.1. Ubicación

La Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo fue creada mediante resolución N° 1184-91-R de fecha 27 de noviembre de 1991, es la Unidad de más alto nivel, lleva acabo dos procesos de admisión al año. Donde permite que diferentes profesionales de todos los campos del conocimiento estudien las maestrías y doctorados para elevar su nivel académico y profesional acorde con la nueva ley universitaria 30220 y con los mejores docentes del país.

La Escuela de Posgrado (EPG), se encuentra ubicado en la av. Huamachuco #1130 en la ciudad de Lambayeque, departamento de Lambayeque.



**Figura NRO. 1: Organigrama de la Escuela de Postgrado de la UNPRG**

*Fuente: [Elaborado por Escuela de Postgrado]*



## **1.2. Planteamiento del Problema**

La Universidad Nacional Pedro Ruiz Gallo a través de su Escuela de Posgrado realiza 2 procesos de admisión al año, donde obtiene información constante de sus postulantes.

En esta etapa no se tiene en consideración el perfil del postulante o demanda de los programas, nivel económico, zonas geográficas, o a quien está dirigido porque no cuenta con una metodología o herramienta tecnológica para el análisis de los datos. Lo importante para la EPG solo es conocer los inscritos a los diversos programas que brinda para el desarrollo profesional de los postulantes y no el perfil de cada potencial alumno.

Es muy importante resaltar también, la tomar decisiones sobre el comportamiento de sus ingresos que se obtienen producto de las pensiones de enseñanza, etc.

Desde su creación la EPG con cada año ha incrementado su población estudiantil, lo que le ha permitido crecer rápidamente gracias al prestigio de sus docentes, alumnos, egresados y graduados que con su trayectoria enriquecen y fortalecen la imagen de la Institución.

Podemos decir también que con el crecimiento de la EPG, ha crecido la deuda por parte de los alumnos, Actualmente existe una deuda de más de 15 millones por conceptos de pensiones de enseñanza y matrícula.

Como respuesta al problema se podría solucionar utilizando la inteligencia de negocios para encontrar patrones de comportamiento que ayuden a crear perfiles financieros adecuados través del modelo propuesto, que cumplan el objetivo deseado por parte de la institución y así reducir la deuda actual.

## **1.3. Formulación del Problema**

¿Qué perfiles de comportamiento financiero se obtendrán mediante el modelo de análisis de sensibilidad basado en inteligencia de negocios, que permitieran determinar los perfiles financieros de los estudiantes de los diferentes programas de Maestría y Doctorado de la Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo?

## **1.4. Objetivos**

### **1.4.1. Objetivo General**

Implementar un modelo de análisis de sensibilidad basado en Inteligencia de Negocios que permita determinar los perfiles financieros basados en patrones de comportamiento de los estudiantes en los diferentes programas de Maestría y Doctorado de la Escuela de Postgrado de la Universidad Nacional Pedro Ruiz Gallo – Lambayeque.

### **1.4.2. Objetivos Específicos**

- a) Realizar un análisis del funcionamiento de los procedimientos académicos y financieros actuales en el proceso de admisión a la EPG, con la finalidad de determinar los perfiles financieros.
- b) Analizar y diseñar un modelo de minería de datos que permita determinar los perfiles financieros a partir de los procesos de admisión y seguimiento de pagos desde el año 2015 al 2018, bajo el enfoque de inteligencia de negocios.
- c) Implementar un modelo de minería de datos diseñado para encontrar patrones de comportamiento financiero a partir de la información obtenida de las transacciones históricas de la EPG, utilizando una herramienta de software libre.
- d) Interpretar y obtener perfiles financieros deseados de los alumnos, a través de reportes resultantes del análisis financiero obtenido de las operaciones realizadas en línea.

## **CAPÍTULO II: MARCO TEÓRICO**

## **2.1. Antecedentes**

- a) En la Tesis de Maestría (Olavarria Paz, 2014); propone un Modelo de explotación de datos históricos de los registros de Inscripción a los procesos de admisión, basado en Business Intelligence (BI) y Customer Relationship Management (CRM), para el incremento de postulantes en el proceso de admisión de la Universidad Nacional Pedro Ruiz Gallo – Lambayeque – Perú. El presente trabajo de investigación se ha enfocado en el desarrollo de un Modelo que permita el incremento de postulantes en los procesos de admisión de la Universidad Nacional Pedro Ruiz Gallo, el autor del trabajo de investigación ha dividido esta tesis en varios capítulos para una mejor descripción, comprensión y entendimiento de las necesidades actuales que surgieron de la Oficina Central de Admisión. Donde describe la teoría de inteligencia de negocio, CRM, comparación y selección de la metodología y herramientas a utilizar para el desarrollo del modelo de explotación de datos. Además utiliza la metodología Kimball y realiza un análisis y discusión de resultado apoyándose en la construcción del modelo con la Plataforma Open Suite de Pentaho dando solución a las necesidades de consultas y generación de reportes de los datos históricos de los procesos de admisión de la UNPRG, y contrastar la hipótesis del proyecto.
  
- b) En la Tesis de Maestría (Falcón, 2012); el autor propone la investigación sobre el Desarrollo de una Solución de Inteligencia de Negocios en el Manejo de Estadísticas de Control en la venta de repuestos realizado en la empresa Talleres Ambamazda S.A., ubicada en la ciudad de Ambato sector Mall de los Andes dedicada a la comercialización de repuestos nuevos de la marca Mazda. La empresa de repuestos automotrices se dirige al segmento de usuarios finales exclusivamente aquellos usuarios que poseen un vehículo de la marca Mazda. Debido a este tipo de negocio y sobre todo a su alto movimiento de stock de repuestos se hace necesario el desarrollo e implementación de una tecnología poco aprovechada por las empresas del sector automotriz como es la Inteligencia de Negocios, de eso trata este trabajo de investigación a mejorar significativamente la toma de decisiones de la gerencia comercial para optimizar recurso de personal y económico.

- c) En la Tesis Doctoral (Lopez, 2008); El impacto de la Data Warehouse e Inteligencia de Negocios en el desempeño de las empresas: investigación empírica en PERÚ, como país en vías de desarrollo. Se realizó una investigación con la finalidad de estimar el impacto que tienen la Data Warehouse (DW) y la Inteligencia de Negocios (BI) en el desempeño de las empresas en un país en vías de desarrollo.

Se establecieron las preguntas de investigación y se utilizaron dos modelos para resolver las mismas. El primero un modelo Cualitativo Exploratorio, mediante entrevistas semi-estructuradas, y el segundo un modelo Cuantitativo, mediante cuestionarios. El modelo Cualitativo Exploratorio utilizó 23 entrevistas en 16 empresas de diversos segmentos de negocios, que utilizan la DW y BI. Se entrevistaron a Gerentes de Inteligencia de Negocios de empresas que desarrollan el sistema de DW y BI.

Se utilizó el modelo del éxito de IS de DeLone y McLean de 2003, con las variables verificadas en el estudio cualitativo. Éste tiene los constructos independientes, Calidad de la Información, Calidad del Sistema y Calidad del Servicio. Para la resolución del modelo se utilizaron las Ecuaciones Estructurales, las cuales son una herramienta de análisis multivariable de última generación que permite trabajar con varios ítems para cada constructo y solucionar el modelo con todos los constructos a la vez.

De esta manera se establecieron las variables o constructos más relevantes, por grado de importancia, y sus componentes, y luego se determinó la significancia de cada uno de ellos, respondiéndose de esta manera las preguntas de investigación y confirmándose las hipótesis más relevantes del estudio cuantitativo.

- d) En el trabajo de Investigación de (Loyo, 2017); tuvo como objetivo de estudio implementar una solución de Inteligencia de Negocios (BI) para el manejo de Indicadores claves de desempeño (KPI) en ventas de los Retails de Farmacias en Farmaenlace Cía. Ltda. Para esto, su fundamento investigativo es de carácter cualitativo, de tipo descriptivo con diseño de transversal. La técnica aplicada fue el grupo focal y entrevista al personal de comercialización, como de TI. Los resultados

obtenidos identifican que las farmacias no manejan indicadores y presentan dificultades en la toma de decisiones comerciales.

Las herramientas de Inteligencia de Negocios (BI) pretenden mejorar los servicios y cadena de valor para los clientes, por medio de Indicadores claves de desempeño (KPI), encargados de identificar, seguir y controlar el desempeño de los procesos.

El BI permite manejar de mejor manera los indicadores del retail de farmacias de los que se consideró: situación general, productividad de la farmacia, productividad de la farmacia por producto, ventas – transacciones, composición de ventas, ventas – consumidor final, ventas – días de la semana – fin de semana, ventas – horas, medicación frecuente, ventas por dispensador (cajero), cumplimiento de top cajeros. El BI ágil es una solución indispensable que permita manejar e interpretar de mejor manera los indicadores KPI de comercialización de los Retails de farmacias, previo a una definición apropiada. Qlik Sense es una opción muy acertada, por acoplamiento y sencillez, con una metodología de las mismas características como QlikView Project Methodology (QPM).

## **2.2 Fundamentos Teóricos**

### **2.2.1. Perfil financiero**

Es un estudio que permite realizar un análisis del índice de la población económicamente activa, utilizando características como salario, sexo, nivel profesional, edad etc. Para realizar una evaluación adecuada y clasificar en que porcentajes podrían ayudar a la toma de decisiones en préstamos a entidades bancarias, instituciones educativas, privadas, etc.

### **2.2.2. Inteligencia de Negocios (Business Intelligence)**

La Inteligencia de Negocios (proceso de recolectar, consolidar y analizar múltiples fuentes de datos para la toma de decisiones estratégicas. En una sola palabra: "Análisis"). Es la habilidad de una empresa para estudiar sus comportamientos y acciones históricas con el objetivo de entender donde la organización ha estado, su situación actual, y predecir qué sucederá en el futuro.

La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio:

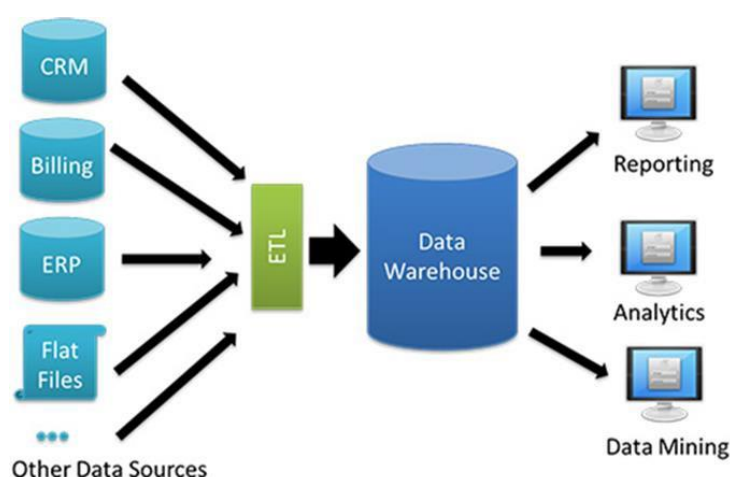
entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de la producción, análisis de perfiles de clientes, rentabilidad de un producto concreto.

Los principales productos de Business Intelligence que existen hoy en día son:

- Cuadro de Mando Integrales (CMI)
- Sistemas de Soporte a la Decision (DSS)
- Sistema de Informacion Ejecutiva (EIS)

Por otro lado, los principales componentes de orígenes de datos en el Business Intelligence que existen en la actualidad son:

Los sistemas y componentes del BI se diferencian de los sistemas operacionales en que están optimizados para preguntar y divulgar sobre datos. Esto significa típicamente que, en un data warehouse, los datos están desnormalizados para apoyar consultas de alto rendimiento, mientras que en los sistemas operacionales suelen encontrarse normalizados para apoyar operaciones continuas de inserción, modificación y borrado de datos. En este sentido, los procesos **ETL** (extracción, transformación y carga), que nutren los sistemas BI, tienen que traducir de uno o varios sistemas operacionales normalizados e independientes a un único sistema desnormalizado, cuyos datos estén completamente integrados.



*Figura NRO. 2: Business Intelligence*

*Fuente: [ (Williams S, 2010)]*

### 2.2.2.1 Beneficios de la implantación del BI

Los primeros sistemas de información aportaban a la dirección básicamente información económica y financiera, con lo que quedaban muchos espacios por cubrir respecto a la organización de la empresa. Los nuevos sistemas de BI abarcan todas las áreas funcionales de una organización (recursos humanos, marketing, logística, etc.) y no sólo los económicos.

Los beneficios que pueden aportar la implantación de sistemas BI son: incremento de la eficiencia en la toma de decisiones, mejora de comunicación entre las diferentes áreas de la organización, mejora de rentabilidad, atracción de nuevos clientes, reducción del fraude, identificación de riesgos, presentación de soluciones basadas en el conocimiento del negocio, etc. La mayor parte de los beneficios de la implantación de un sistema de BI son intangibles, ya que derivan en la mejora de la gestión de la compañía. Esto dificulta la tarea de calcular su ROI (Return On Investment) para obtener beneficios cuantificables.

### 2.2.3. Metadatos

Los metadatos, literalmente «sobre datos», son datos que describen otros datos. En general, un grupo de metadatos se refiere a un grupo de datos que describen el contenido informativo de un objeto al que se denomina recurso.



*Figura NRO. 3: Metadatos*

*Fuente: [Lack E. Myers (1969)]*

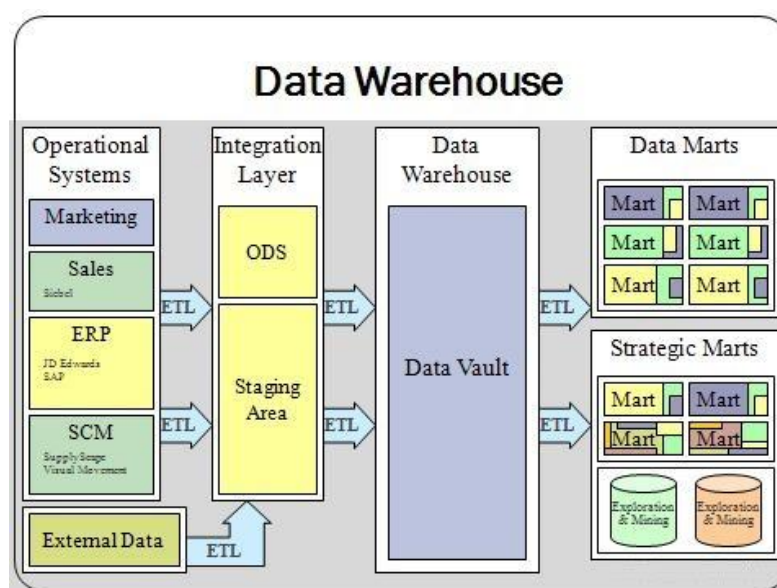


#### 2.2.4. Data Warehouse

El concepto de data warehouse se originó en 1988 con el trabajo de los investigadores de IBM, Barry Devlin y Paul Murphy aunque el término data warehouse fue acuñado por William H. Inmon, el cual es conocido como el padre de Data Warehousing. Inmon describió un data warehouse como una colección de datos orientada a un tema específico, integrado, variante en el tiempo y no volátil, que soporta el proceso de toma de decisiones.

Un Data Warehouse es un almacén electrónico donde generalmente una empresa u organización mantiene una gran cantidad de información. Los datos de un data warehouse deben almacenarse de forma segura, fiable, fácil de recuperar y fácil de administrar.

Un Data Warehouse es un repositorio unificado para todos los datos que recogen los diversos sistemas de una empresa. El repositorio puede ser físico o lógico y hace hincapié en la captura de datos de diversas fuentes sobre todo para fines analíticos y de acceso.



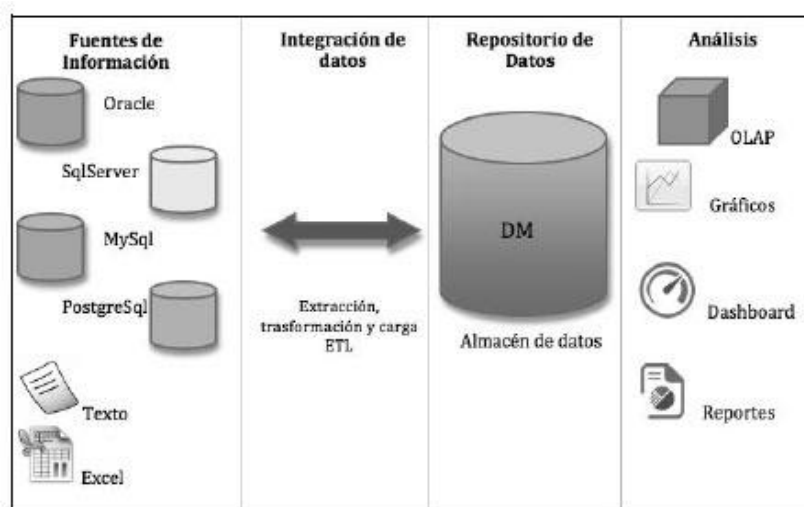
*Figura NRO. 4: Representación del Data Warehouse*

*Fuente: [Bill Inmon 1992]*

### 2.2.5. DataMart

Según la definición de Oracle Corporation “Un Data Mart es una forma simple de un Data Warehouse que está enfocada en una área funcional de la empresa como puede ser ventas, finanzas, marketing, etc.”

**DataMart** es un modelo multidimensional basado en **tecnología OLAP** que representa a un área específica de la empresa, incluyendo las variables claves y los indicadores para el proceso de toma de decisiones.



*Figura NRO. 5: Representación del DataMart*

*Fuente: [Bill Inmon 1992]*

De acuerdo a sinnexus (2012), Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un datamart puede ser alimentado desde los datos de un DataWareHouse, o integrar por sí mismo un compendio de distintas fuentes de información.

### 2.2.6. Datamart OLAP

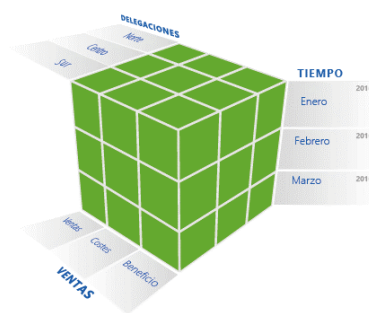
De acuerdo a sinnexus (2012), Datamart OLAP Se basan en los populares cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional. El modo de creación, explotación y mantenimiento de los cubos OLAP es muy heterogéneo, en función de la herramienta final que se utilice.

### 2.2.7. Datamart OLTP

De acuerdo a sinnexus (2012), Datamart OLTP Pueden basarse en un simple extracto del datawarehouse, no obstante, lo común es introducir mejoras en su rendimiento (las agregaciones y los filtrados suelen ser las operaciones más usuales) aprovechando las características particulares de cada área de la empresa. Las estructuras más comunes en este sentido son las tablas report, que vienen a ser fact-tables reducidas (que agregan las dimensiones oportunas), y las vistas materializadas, que se construyen con la misma estructura que las anteriores, pero con el objetivo de explotar la reescritura de queries.

Los datamarts que están dotados con estas estructuras óptimas de análisis presentan las siguientes ventajas:

- Poco volumen de datos
- Mayor rapidez de consulta
- Consultas SQL y/o MDX sencillas
- Validación directa de la información
- Facilidad para la historización de los datos



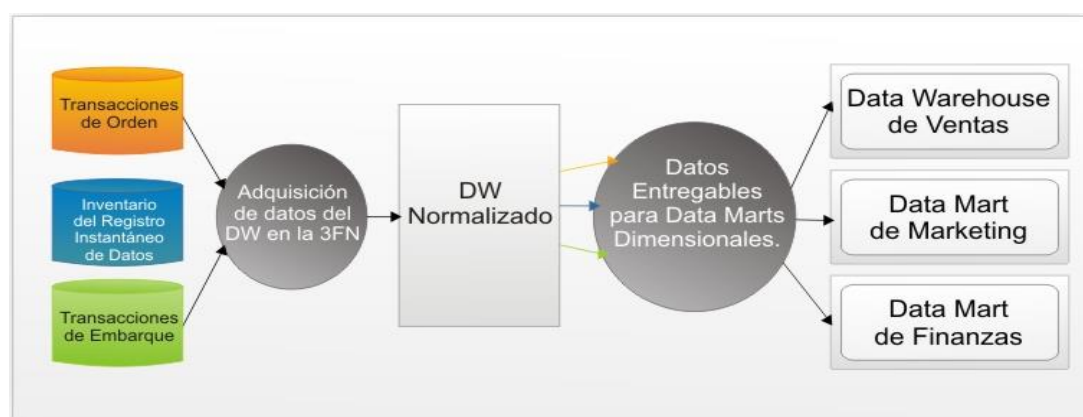
**Figura NRO. 6: DataMart**  
**Fuente: [Bill Inmon 1992]**

### **2.2.8. Metodología de Bill Inmon**

Bill Inmon ve la necesidad de transferir la información de los diferentes OLTP (Sistemas Transaccionales) de las organizaciones a un lugar centralizado donde los datos puedan ser utilizados para el análisis (sería el CIF o Corporate Information Factory). Insiste además en que ha de tener 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.
- **No volátil:** La información no se modifica ni se elimina, una vez almacenado un dato, éste se convierte en información de sólo 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.

La información ha de estar a los máximos niveles de detalle. Los Data warehouses departamentales o datamarts son tratados como subconjuntos de este Data warehouse corporativo, que son construidos para cubrir las necesidades individuales de análisis de cada departamento, y siempre a partir de este Data warehouse central (del que también se pueden construir los ODS (Operational Data Stores ) o similares).



*Figura NRO. 7: Metodología Bill Inmon*

*Fuente: [Bill Inmon 1992]*

El enfoque Inmon también se referencia normalmente como Top-down. Los datos son extraídos de los sistemas operacionales por los procesos ETL y cargados en las áreas stage, donde son validados y consolidados en el DW corporativo, donde además existen los llamados metadatos que documentan de una forma clara y precisa el contenido del DW. Una vez realizado este proceso, los procesos de refresco de los Datamart departamentales obtienen la información de este, y con las consiguientes transformaciones, organizan los datos en las estructuras particulares requeridas por cada uno de ellos, refrescando su contenido.

La metodología para la construcción de un sistema de este tipo es la habitual para construir un sistema de información, utilizando las herramientas habituales (esquema Entidad Relación, DIS (Data Item Sets, etc). Para el tratamiento de los cambios en los datos, usa la Gestión de las dimensiones continuas y discretas (inserta fechas en los datos para determinar su validez para la dimensión continua o bien mediante el concepto de snapshot o foto para la dimensión discreta).

Al tener este enfoque global, es más difícil de desarrollar en un proyecto sencillo (pues estamos intentando abordar el “todo”, a partir del cual luego iremos al “detalle”).

### **2.2.9. Metodología de Ralph Kimball**

La metodología de Kimball, llamada Modelo Dimensional (Dimensional Modeling), se basa en lo que se denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Esta metodología es considerada una de las técnicas favoritas a la hora de construir un Data Warehouse.

En el Modelo Dimensional se constituyen modelos de tablas y relaciones con el propósito de optimizar la toma de decisiones, con base en las consultas hechas en una base de datos relacional que están ligadas con la medición o un conjunto de mediciones de los resultados de los procesos de negocio.

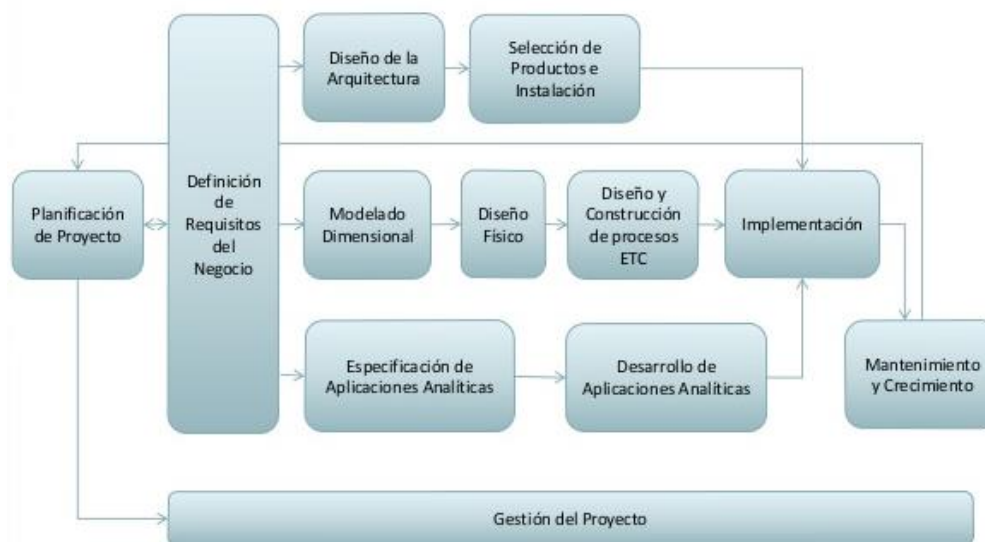
El Modelo Dimensional es una técnica de diseño lógico que tiene como objetivo presentar los datos dentro de un marco de trabajo estándar e intuitivo, para permitir su acceso con un alto rendimiento. Cada Modelo Dimensional está compuesta por una tabla con una llave combinada, llamada tabla de hechos, y con un conjunto de tablas más pequeñas llamadas tablas de dimensiones. 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 contexto de los hechos.

Una base de datos dimensional se puede concebir 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.

Dado que es muy común representar a un modelo dimensional como una tabla de hechos rodeada por las tablas de dimensiones, frecuentemente se le denomina también modelo estrella o esquema de estrella-únión

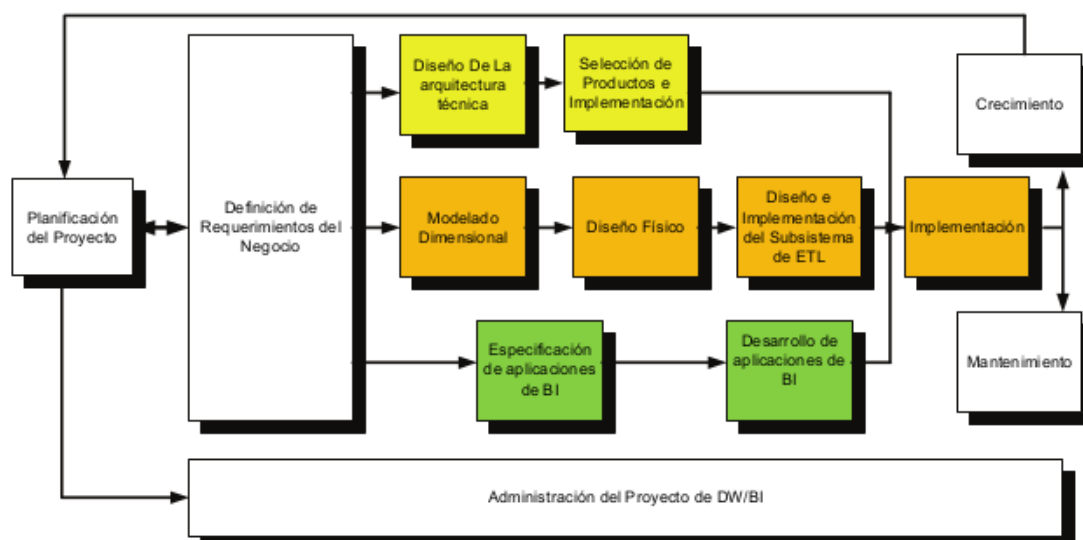


**Figura NRO. 8: Marco Metodológico - Metodología de Ralph Kimball**

*Fuente: [Kimball et al 98, 08, Mundy & Thornthwaite 06]*

#### **2.2.9.1. Ciclo de Vida Ralph Kimball**

- Ilustra el flujo general de implementación de un DW.
- Identifica secuencia de tareas ordenadas y actividades principales.
- No todos los detalles de las tareas del ciclo de vida deben ser ejecutados en todos los proyectos.



*Figura NRO. 9: Metodología de Ralph Kimball*

*Fuente: [Kimball et al 98, 08, Mundy & Thornthwaite 06]*

Como se puede apreciar en la figura, los Requerimientos del Negocio son el soporte inicial de las tareas subsiguientes. También tiene influencia en el plan de proyecto (puede notar la doble fecha entre la caja de definición de requerimientos y la de planificación). Podemos también ver tres rutas o caminos que se enfocan en tres diferentes áreas:

- **Tecnología (Camino Superior):** Implica tareas relacionadas con software específico, por ejemplo, Microsoft SQL Analysis Services.
- **Datos (Camino del medio):** En la misma diseñaremos e implementaremos el modelo dimensional, y desarrollaremos el subsistema de Extracción, Transformación y Carga (Extract, Transformation, and Load - ETL) para cargar el DW.
- **Aplicaciones de Inteligencia de Negocios (Camino Inferior):** En esta ruta se encuentran tareas en las que diseñamos y desarrollamos las aplicaciones de negocios para los usuarios finales.



Estas rutas se combinan cuando se instala finalmente el sistema. En la parte de debajo de la figura se muestra la actividad general de administración del proyecto. A continuación describiremos cada una de las tareas:

**a) Planificación:** En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

En la visión de programas y proyectos de Kimball, Proyecto, se refiere a una iteración simple del Ciclo de Vida de Kimball, desde el lanzamiento hasta el despliegue.

Esta tarea incluye las siguientes acciones típicas de un plan de proyecto:

- Definir el alcance (Entender los Requerimientos del Negocio)
- Identificar las tareas
- Programar las tareas
- Planificar el uso de los recursos
- Asignar la carga de trabajo a los recursos
- Elaboración de un documento final que representa un plan del proyecto

**b) Análisis de Requerimientos:** La definición de los requerimientos es en gran medida un proceso de entrevistar al personal de negocio y técnico, pero siempre conviene tener un poco de preparación previa. Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo. Hay que leer todos los informes posibles de la organización; rastrear los documentos de estrategia interna; entrevistar a los empleados, analizar lo que se dice en la prensa acerca de la organización, la competencia y la industria. Se deben conocer los términos y la terminología del negocio.

- c) **Modelado Dimensional:** El proceso de diseño comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados de la matriz descrita en el punto anterior. El proceso iterativo consiste en cuatro pasos:
- Elegir el Proceso de Negocio.
  - Establecer el Nivel de Granularidad.
  - Elegir las Dimensiones.
  - Identificar medidas y las tablas de hechos.
- d) **Diseño Físico:** Se centra en la selección de las estructuras necesarias para soportar el diseño lógico. Un elemento principal de este proceso es la definición de estándares del entorno de la base de datos. La indexación y las estrategias de particionamiento se determinan en esta etapa.
- e) **Diseño del Sistema de Extracción, Transformación y Cargar (ETL):** Es la base sobre la cual se alimenta el Datawarehouse. Si el sistema ETL se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar (grabar) la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis.
- f) **Especificación y Desarrollo de Aplicaciones BI:** Las aplicaciones de BI son la cara visible de la inteligencia de negocios: los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama informes estándar y aplicaciones analíticas.

### 2.2.9.2. Comparación y selección de la Metodología

Para la elección de la metodología tenemos en cuenta el siguiente cuadro comparativo de acuerdo a los dos enfoques.

*Tabla 1: Análisis comparativo de las Metodologías*

	<b>INMON</b>	<b>KIMBALL</b>
<b>Enfoque</b>	<b>Top-down</b>	<b>Bottom up</b>
Complejidad del método	Bastante complejo	Complejidad bastante simple
Orientación de la data	Maneja data o áreas	Orientado a procesos
Herramientas	Tradicional (ER y DIS)	Modelación Dimensional
Accesibilidad usuario final	Bajo	Alto
Audiencia primaria	IT	Usuarios finales
Objetivos	Ofrecer una buena solución técnica basada en métodos probados	Ofrecer una solución que hace que sea fácil para los usuarios finales para consultar los datos y tener una tasa de repuesta razonable

*Fuente: [ Kimball et al 98 & Bill Inmon 1992]*

De acuerdo al cuadro comparativo y al trabajo que se realizó en este proyecto, de implementar un modelo análisis de sensibilidad para determinar los perfiles financieros de los estudiantes de Posgrado de la Universidad Nacional Pedro Ruiz Gallo en los diferentes programas de maestrías y doctorados basado en Business Intelligence (BI), es más conveniente utilizar el enfoque de Kimball donde podamos ofrecer un solución que sea fácil de utilizar para los usuarios finales y que puedan consultar los datos y obtener respuestas razonables para una toma de decisiones acertada en beneficio de la Universidad.

### **2.2.10. Gestor de base de datos MySQL**

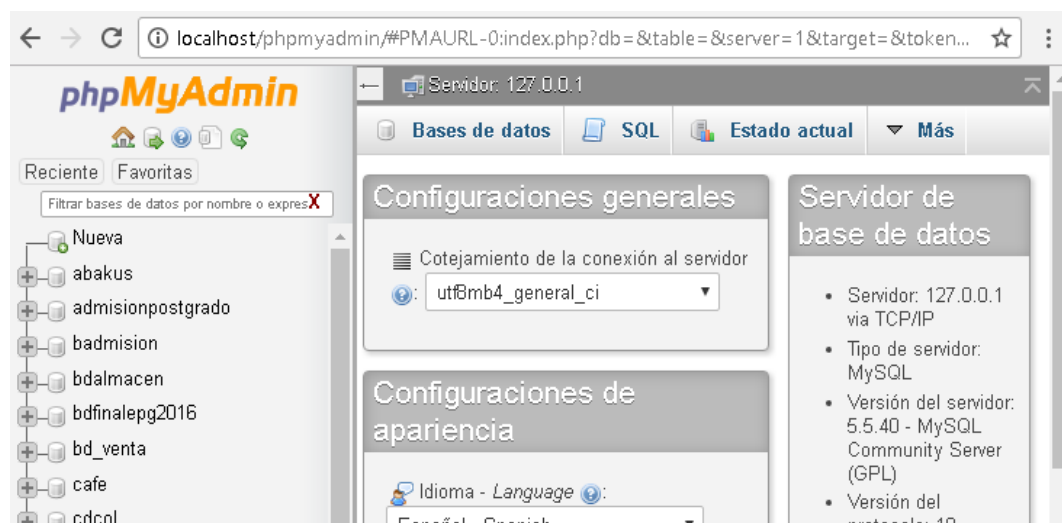
Según Ian Gilfillan (2003). MySQL es un sistema de administración de bases de datos relational (RDBMS). Se trata de un programa capaz de almacenar una enorme cantidad de datos de gran variedad y de distribuirlos para cubrir las necesidades de cualquier tipo de organización, desde pequeños establecimientos comerciales a grandes empresas y organismos administrativos. MySQL compite con sistemas RDBMS propietarios conocidos, como Oracle, SQL Server y DB2.

MySQL utiliza el lenguaje de consulta estructurado (SQL). Se trata del lenguaje utilizado por todas las bases de relacionales. Este lenguaje permite crear bases de datos, así como agregar, manipular y recuperar datos en función de criterios específicos.

#### **2.2.10.1. Herramienta de Administración phpMyAdmin**

phpMyAdmin es una de las herramientas más conocidas para la administración del servidor de base de datos MySQL, está escrita en php y dispone de una interfaz gráfica permitiendo realizar todo tipo de operaciones sobre la base de datos:

- Crear, borrar y modificar tablas.
- Consultar, insertar, modificar y eliminar datos.
- Definir usuarios y asignar permisos.
- Realizar copias de seguridad, etc.



**Figura NRO. 10: Herramienta de Administración phpMyAdmin**

**Fuente: [ phpmyadmin 2018]**

## **2.2.11. Ventajas y Desventajas de MySQL**

### **a) Ventajas**

- MySQL software es Open Source
- Velocidad al realizar las operaciones, lo que le hace uno de los gestores con mejor rendimiento.
- Bajo costo en requerimientos para la elaboración de bases de datos, ya que debido a su bajo consumo puede ser ejecutado en una máquina con escasos recursos sin ningún problema.
- Facilidad de configuración e instalación. Soporta gran variedad de Sistemas Operativos.
- Baja probabilidad de corromper datos, incluso si los errores no se producen en el propio gestor, sino en el sistema en el que está.
- Su conectividad, velocidad, y seguridad hacen de MySQL Server altamente apropiado para acceder bases de datos en Internet.
- Software MySQL usa la licencia GPL.

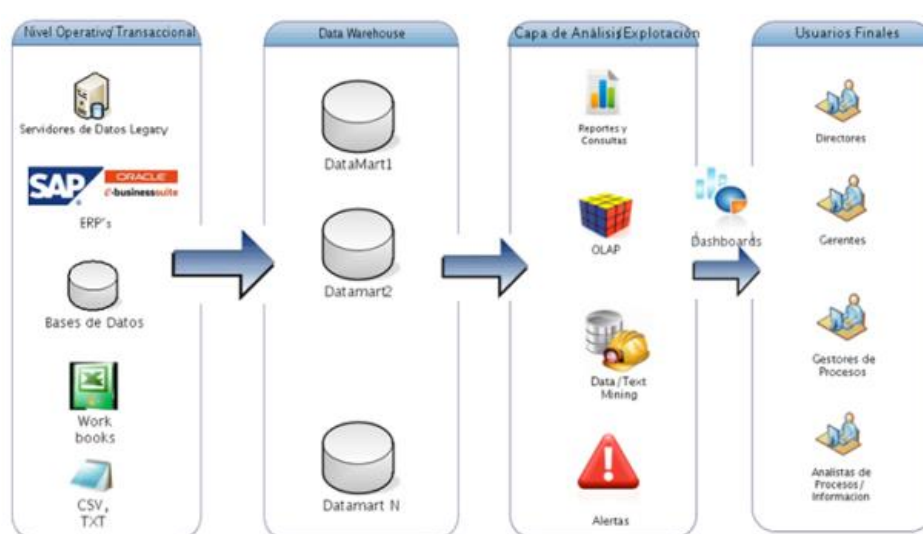
### **b) Desventajas**

- Un gran porcentaje de las utilidades de MySQL no están documentadas.
- No es intuitivo, como otros programas (ACCESS).

### 2.2.12. Pentaho BI Suite

La visibilidad que nos está dando pentaho, la solución de software libre de BI de sus componentes, se ha convertido en un interesante ejercicio de análisis de la arquitectura de una suite de BI.

En las siguientes líneas vamos a comentar en qué estado actual se encuentra la solución y que con que componentes cuenta, para llevar a cabo un proyecto BI, como se indica en la siguiente gráfica:



*Figura NRO. 11: Arquitectura de una solución Business Intelligence*

*Fuente: [Pentaho 2017]*

Según Nestic-todobi (2013). Pentaho se define a sí mismo como una plataforma de BI “orientada a la solución” y “centrada en procesos” que incluye todos los principales componentes requeridos para implementar soluciones basados en procesos y ha sido concebido desde el principio para estar basada en procesos.

Las soluciones que Pentaho pretende ofrecer se componen fundamentalmente de una infraestructura de herramientas de análisis e informes integrados con un motor de workflow de procesos de negocio. La plataforma será capaz de ejecutar las reglas de negocio necesarias, expresadas en forma de procesos y actividades y de presentar y entregar la información adecuada en el momento adecuado.

Su modelo de ingresos parece estar orientado a los servicios (soporte, formación, consultoría y soporte a ISVs y distribuciones OEM) aunque en alguno de los

documentos y páginas que hemos examinado aparece mencionado algunas funcionalidades “Premium” que hacen pensar en ingresos por futuras versiones o funcionalidades de pago.

Pentaho presenta informes en los formatos habituales (html, excel, pdf...) mediante JfreeReport, proyecto incorporado recientemente a Pentaho junto con su responsable Thomas Morgner, u otras plataformas como BIRT o JasperReports. Para la generación de PDFs utilizan, como podría ser previsible, el conocidísimo Apache FOP. Asimismo incorpora la librería JPivot, gracias a la cual podemos ver tablas OLAP a través de un browser y realizar las aplicaciones típicas de análisis OLAP.



Figura NRO. 12: Pentaho Reporting

Fuente: [Pentaho 2017]

Los dashboards son un desarrollo propio de Pentaho. Recogen información de todos los componentes de la plataforma incluyendo aplicaciones externas, feeds RSS y páginas web. Incluyen gestión y filtrado del contenido, seguridad basada en roles y drill down. Pueden ser integrados en terceras aplicaciones, en portales o dentro de la plataforma Pentaho.

Para generar gráficos se apoyan en JFreeChart, una librería para generar los gráficos más comunes (2D, 3D, barras, líneas series temporales, Gantt...), interfaces para acceder

a diferentes fuentes de datos, exportación a PNG, JPEG y PDF y soporte para servlets, JSPs, applets y aplicaciones clientes.



*Figura NRO. 13: Pentaho Dashboards*

*Fuente: [Pentaho 2017]*

Los WebServices son una característica fundamental de Pentaho. Las acciones, que son las tareas más sencillas que constituyen una solución de Pentaho, pueden publicarse como WebServices. Pentaho utiliza como motor de WebServices Apache Axis, quedando los servicios descritos en el lenguaje de definición de servicios web WSDL.

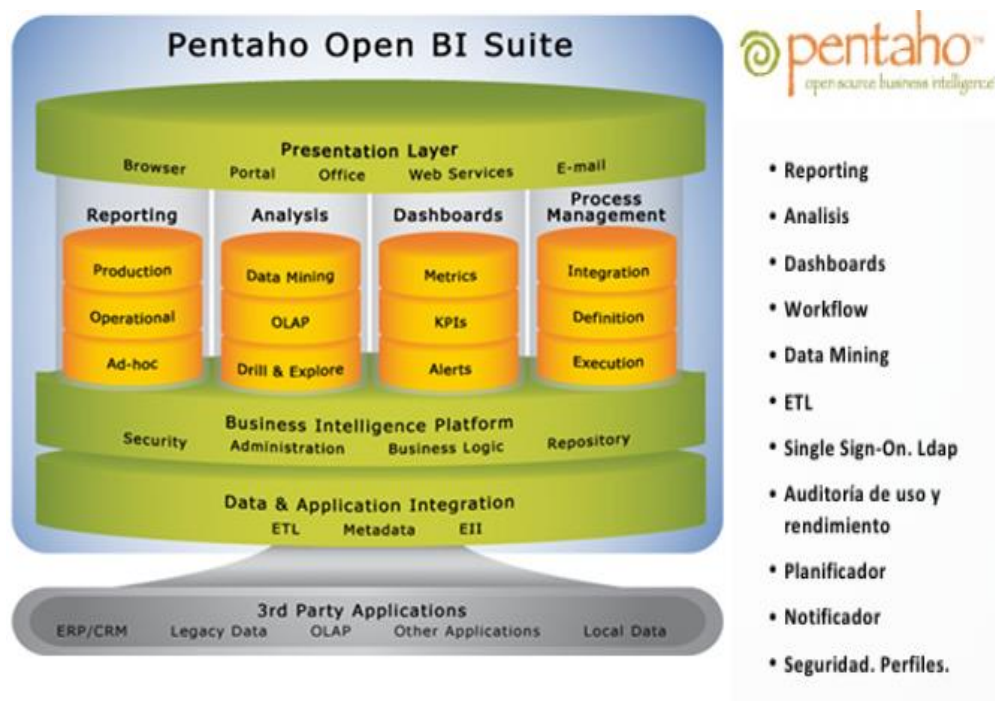
Para entregar la información Pentaho se apoya en una infraestructura. JMS para enviar correos electrónicos y Quartz, un scheduler opensource integrable en aplicaciones J2EE (de hecho necesita ser instanciado).

De acuerdo a bievolutivo (2011). Pentaho Open BI Suite proporciona reporting intuitivo, análisis OLAP, cuadros de mando, integración de datos, minería de datos y Plataforma BI. Esta suite se ha convertido en la líder mundial y la más ampliamente utilizada como herramientas BI de código libre.



Además de poder analizar tu empresa, serás capaz de:

- Informar: Acceder a los datos y suministrar información a toda la empresa.
- Analizar: Explorar y analizar los datos interactivamente y de forma muy rápida.
- Sintetizar: Consigue inmediata visibilidad con medidas y ratios a través de cuadros de Mando.
- Integrar: Pule e integra datos: estén donde estén y desde múltiples fuentes.
- Investigar: "Hacer minería" a través de tus datos para descubrir patrones ocultos e indicadores de tendencias futuras.



*Figura NRO. 14: Pentaho Open BI Suite*

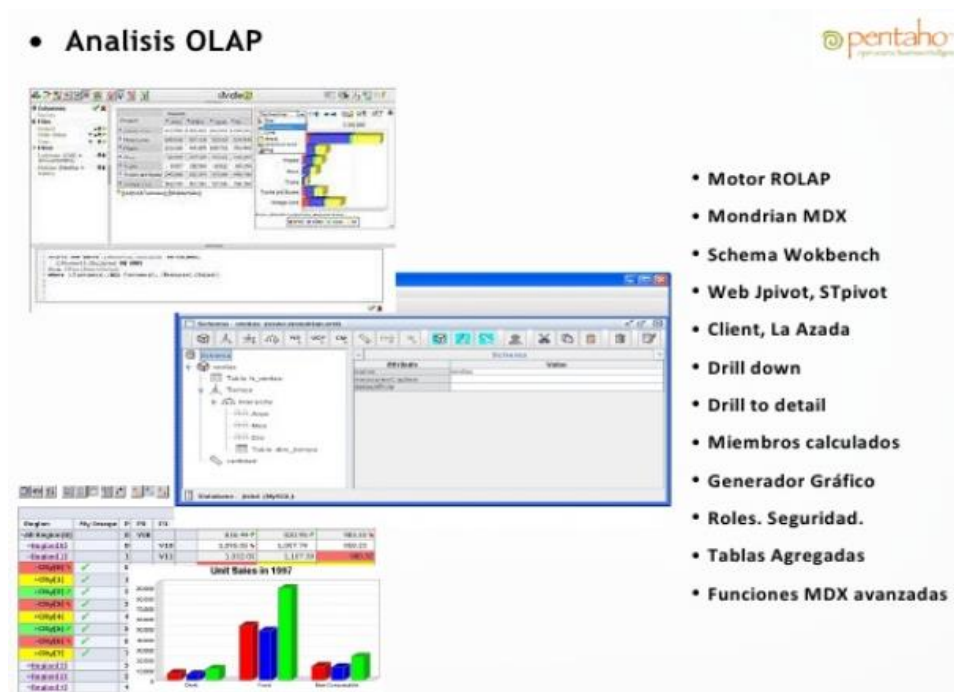
*Fuente: [Pentaho 2017]*

### 2.2.13. Servidor OLAP

Para obtener la funcionalidad de procesamiento analítico en línea (OLAP) se utilizan otras dos aplicaciones: el servidor OLAP Mondrian, que combinado con Jpivot, permiten realizar queries a Datamarts, que los resultados sean presentados mediante un browser y que el usuario pueda realizar drill down y el resto de las navegaciones típicas.

Algunos puntos interesantes sobre Mondrian:

- Mondrian utiliza MDX como lenguaje de consulta, que no tendría nada de raro si no fuera porque fue un lenguaje propuesto por Microsoft.
- En noviembre de 2005 Mondrian se incorporó al proyecto Pentaho, incluyendo a su fundador Julian Hyde.
- Mondrian funciona sobre las bases de datos estándar del mercado: Oracle, DB2, SQL-Server, MySQL... lo cual habilita y facilita el desarrollo de negocio basado en la plataforma Pentaho.
- JPivot está considerado con un “proyecto hermano” de Mondrian. Al incorporarse Mondrian también lo ha hecho JPivot

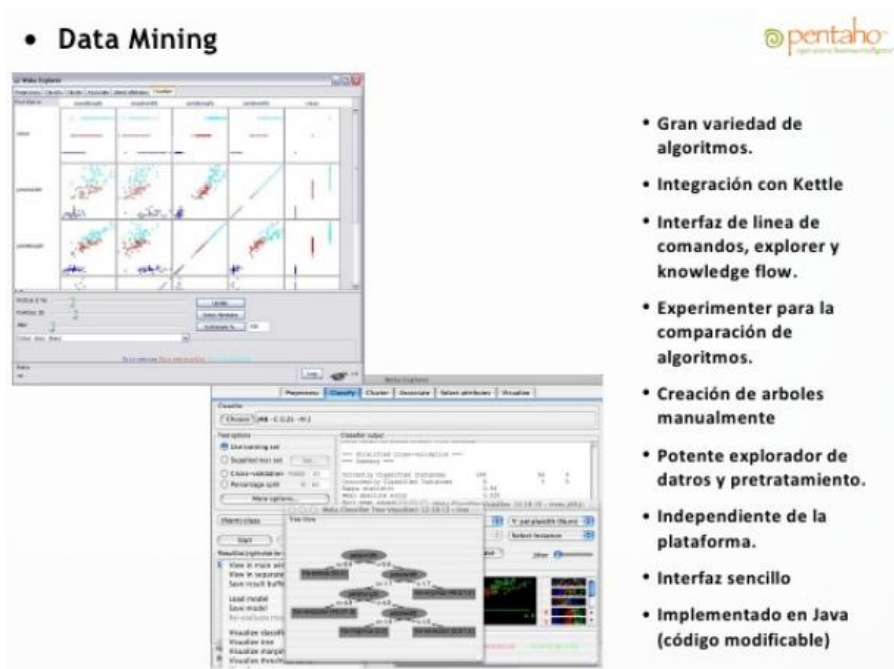


*Figura NRO. 15: Análisis OLAP*

*Fuente: [Pentaho 2017]*

## 2.2.14. Data Mining

Qué sería de una aplicación de BI sin Data Mining. Para tal fin, Pentaho está incorporando la tecnología WeKa. Decimos está incorporando, porque a día de hoy el roadmap de Pentaho indica que están en ello. Weka es una herramienta extensible e integrable que incluye herramientas para realizar transformaciones sobre los datos, tareas de clasificación, regresión, clustering, asociación y visualización.



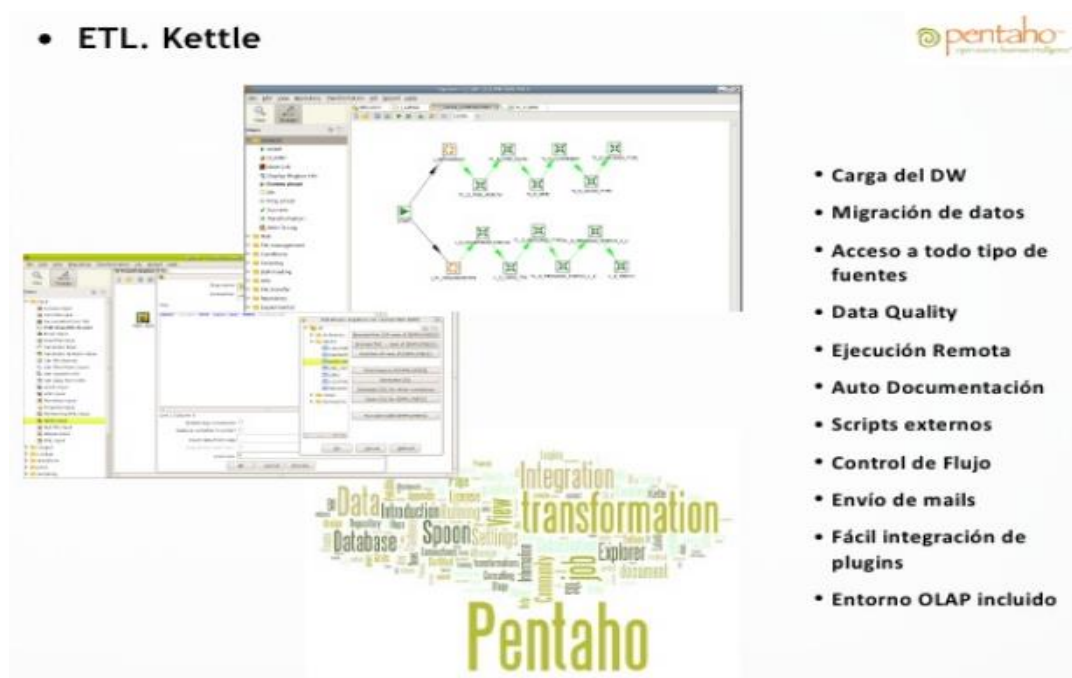
*Figura NRO. 16: Data Mining*

*Fuente: [Pentaho 2017]*

## 2.2.15. ETL

Kettle es un proyecto belga que incluye un conjunto de herramientas para realizar ETL. Uno de sus objetivos es que el proyecto ETL sea fácil de generar, mantener y desplegar. Se compone de 4 herramientas:

- SPOON: permite diseñar de forma gráfica la transformación ETL.
- PAN: ejecuta las transformaciones diseñadas con SPOON.
- CHEF: permite, mediante una interfaz gráfica, diseñar la carga de datos incluyendo un control de estado de los trabajos.
- KITCHEN permite ejecutar los trabajos batch diseñados con Chef.



*Figura NRO. 17: ETL Kettle*

*Fuente: [Pentaho 2017]*

## **2.2.16. Pentaho Report Designer**

Según Gorman (2009). Pentaho Reporting es un código abierto y fácil de usar, la suite de peso ligero de proyectos Java construido para una generación propósito informe. En este libro, usted descubrirá cómo fácil que es integrar Pentaho Reporting en sus proyectos de Java, o utilizarlo como una plataforma de informes independiente. De código abierto de Pentaho Reporting licencia de la GNU Lesser General Public License (LGPL), ofrece a los desarrolladores la libertad para incrustar Pentaho Reporting en su código abierto y aplicaciones propietarias sin costo alguno.

Una comunidad activa participa en el desarrollo y uso de Pentaho Reporting, responder a las preguntas del foro, corrigiendo errores, y la aplicación de nuevas características. Mientras muchas opciones de informes patentados están disponibles, ninguno puede ofrecer la apertura y flexibilidad que Pentaho Reporting ofrece a sus usuarios.

Al igual que con la mayoría de los proyectos de código abierto con éxito, Pentaho Reporting tiene una trayectoria comprobada grabar, junto con una larga lista de características. La mayor parte de esta historia se ha documentado en foros abiertos y en las discusiones de correo electrónico, que aún están disponibles para la gente a navegar a través de recoger ideas. Empezó como un hobby y se convirtió en una generación de informes empresariales Suite en el transcurso de siete años, el Pentaho Reporting Engine y su conjunto de herramientas como el Diseñador de informes, Asistente para informes de diseño, y Ad basado en la web de Pentaho Interfaz de usuario Hoc Reporting, se utilizan como componentes críticos en un sinnúmero de empresas, educativo, gubernamental, y soluciones de tecnología de información de la comunidad.

En la mayoría de las aplicaciones de software de negocios, un componente de informes es necesario, ya sea para resumir los datos, lo que genera un gran número de documentos personalizados, o simplemente para hacer que sea más fácil para imprimir información que sería útil en diversos formatos de salida. Con un conjunto completo de características, incluyendo PDF, Excel, HTML, y la generación de reportes RTF, junto con las capacidades avanzadas de informes como sub-informes y las fichas cruz, Pentaho Reporting puede romper el más simple de los problemas rápidamente, junto con la solución de los retos más avanzados en el diseño, la generación y el despliegue de informes.

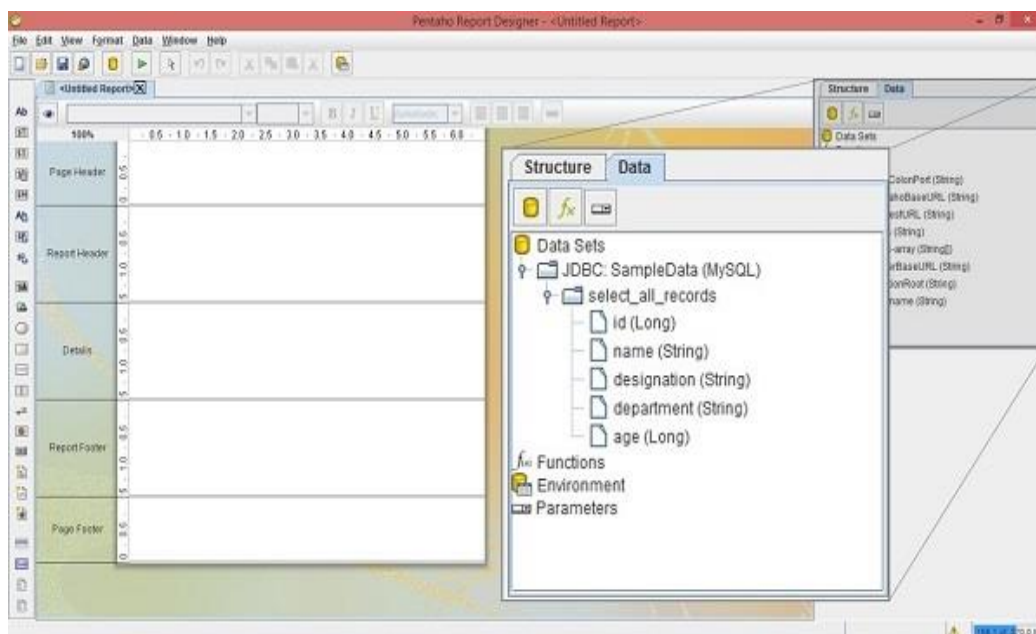


**Figura NRO. 18: Pentaho Report Designer**

**Fuente: [Pentaho 2017]**

De acuerdo a Espinosa (2010). Pentaho Report Designer es una herramienta de reporting que nos permite crear nuestros propios informes, bien para ejecutarlos directamente o para publicarlos en la plataforma BI y que desde allí puedan ser utilizados por los usuarios. La herramienta es independiente de la plataforma y forma parte del conjunto de herramientas de la suite de Pentaho.

Pentaho Report Designer nos permite trabajar con múltiples orígenes de datos (JDBC, Olap4J, Pentaho Analysis, Pentaho Data Integration, XML) incluido el metadata que tengamos definido en nuestro sistema. También nos permite modificar los informes ad-hoc que hayamos creado utilizando WAQR (de hecho es la única forma de modificarlos). El motor de reporting de Pentaho está basado en JFreeReports y ha sido totalmente rediseñado en lo que llaman PRD (Pentaho Report Designer). PRD es un generador de informes del tipo Banded (en contraposición de los generadores orientados a flujo), en los que un informe se divide en secciones o grupos de datos en los que los elementos del informe pueden ser posicionados. Esta forma de trabajar tiene algunas limitaciones, que se pueden superar con el uso de sub informes. El resultado de los informes que vamos diseñado se puede ver con las opciones de pre visualización, y nos permite la salida de resultados en diferentes formatos como PDF, HTML, XLS, RTF y CSV.



**Figura NRO. 19: Ejemplo Pentaho Report Designer**

**Fuente: [Pentaho 2017]**

## **CAPÍTULO III: MARCO METODOLÓGICO**

### **3.1 Tipo y Diseño de la Investigación**

El presente proyecto de investigación se caracteriza por ser cuantitativo.

Área de investigación: Desarrollo de Tecnologías e Innovación

Línea de investigación: Gobierno de Tecnologías de Información

#### **3.1.1 Tipo de Investigación**

La presente investigación es de tipo **Descriptiva** porque muestra una propuesta propositiva, que describe los componentes que la conforman, analizando y evaluando críticamente los sucesos, sin afectar procedimientos actuales, como parte de un proyecto de prueba, generando soluciones, analizando las variables y los potenciales resultados para poder lograr un cambio o mejora en una situación puntual o varias.

### **3.2 Variables**

#### **3.2.1 Variable Independiente**

Un Modelo de análisis de sensibilidad basado en Business Intelligence

#### **3.2.2 Variable Dependiente**

Determinar los perfiles financieros de los estudiantes en los diferentes programas de Maestría y Doctorados de la Escuela de Postgrado.



## **CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS**

## **4.1 Metodología**

La metodología utilizada en este proyecto se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio. Especializándose en la construcción de Datawarehouse (almacenes de datos de una manera modular) pero integrada para la toma de decisiones. Esta metodología ofrece un marco de trabajo que consolida las actividades propias de los proyectos de BI de una manera estructurada, complementando actividades de análisis de requerimientos con los niveles estratégicos e intermedios de la organización, esta actividad tiene como finalidad, conocer las diferentes decisiones que toma cada responsable de área.

Cabe mencionar como puntos resaltantes, la metodología el modelado dimensional y la construcción de los procesos de extracción, transformación y carga (ETL).

### **4.1.1. Diseño de la Arquitectura**

#### **4.1.1.1. Requerimientos del negocio**

Para el desarrollo de este proyecto se han identificado los siguientes requerimientos del negocio que han sido validados con usuarios de la Dirección Administrativa y Académica de la EPG. Las mismas que serán implementadas en el desarrollo del proyecto para que sirvan también como indicadores.

A continuación se muestra la lista de requerimientos de la EPG.

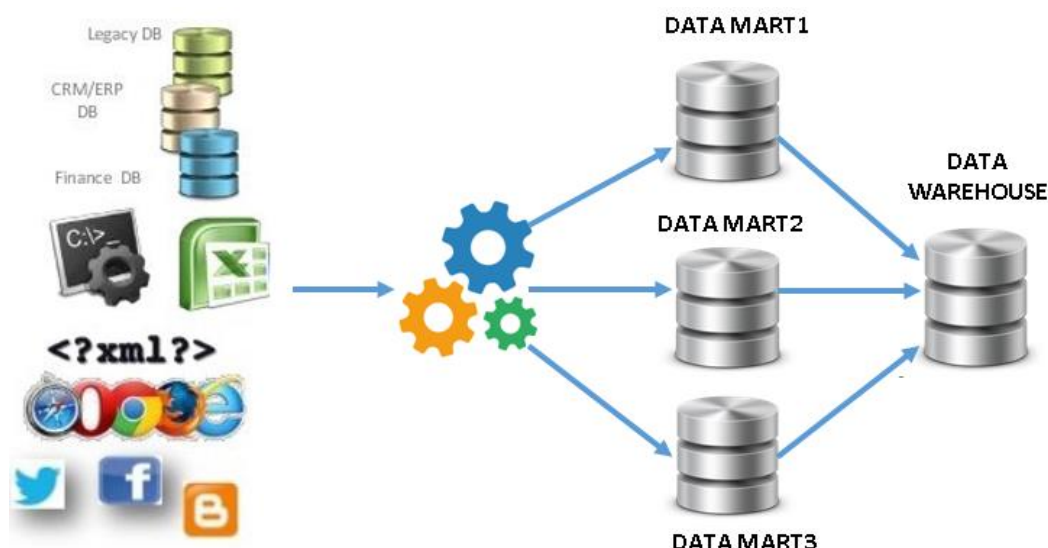
**Tabla 2: Requerimientos del Negocio**

<b>NRO.</b>	<b>REQUERIMIENTO</b>
RQ-01	Se extraerá de forma segura y adecuada según procedimiento, la información de la base de datos del sistema SINEP de la EPG, según corresponda la naturaleza de los archivos (archivos CSV, MYSQL, SQLSERVER, ETC).
RQ-02	Se creará un Data Mart para la dirección académica y administrativa
RQ-03	Reporte de ingresantes por procesos de admisión de los últimos 5 años
RQ-04	Reporte de alumnos por sexo de los últimos 3 años por procesos de admisión
RQ-05	Reporte de ingresos por forma de pago de los últimos 2 años
RQ-06	Reporte de ingresos por mes de los últimos 4 años
RQ-07	Reporte de ingresos por año de los últimos 4 años
RQ-08	Reporte de ingresos por mes de las partidas Derecho de Inscripción, Grados y Títulos, Matrícula, Pensión Enseñanza de los últimos 4 años
RQ-09	Reporte de ingresos por año de las partidas Derecho de Inscripción, Grados y Títulos, Matrícula, Pensión Enseñanza de los últimos 4 años
RQ-10	Reporte de ingresos por meses de la partida Pensión de Enseñanza de los últimos 4 años
RQ-11	Reporte de ingresos por programas de los procesos de admisión 2017-I / 2017-II
RQ-12	Reporte cantidad de alumnos por sexo y programa de los procesos de admisión 2017-I / 2017-II
RQ-13	Reporte de ingresos por programas y forma de pago de proceso de admisión 2017-II
RQ-14	Reporte de ingresos de los programas y conceptos de Matrícula, Pensión de Enseñanza del proceso de admisión 2017-II
RQ-15	Reporte de ingresos mensuales de los programas por concepto de Matrícula del proceso de admisión 2017-II del año 2018
RQ-16	Reporte cantidad de alumnos por rango de edades de los procesos de admisión 2017-I / 2017-II
RQ-17	Se permitirá que el usuario elija los criterios que desea para sus reportes. Podrá realizar las operaciones de Drill Down, Drill Up.
RQ-18	Se proporcionará una plataforma de BI Web para la presentación de los reportes.

*Fuente: [Elaboración Propia]*

#### 4.1.1.2. Diseño de la Arquitectura del Modelo

Para el presente proyecto se presentó una arquitectura donde se ha definido previamente los Data Mart y luego se integran en un Data Warehouse centralizados.



*Figura NRO. 20: Diseño de la Arquitectura*

En donde los Data Mart se cargan a través de procesos ETL, los cuales suministrarán la información adecuada a cada uno de ellos. En muchas ocasiones, los Data Mart son implementados sin que exista el Data Warehouse, ya que tienen sus mismas características pero con la particularidad de que están enfocados en un tema específico. Luego de que hayan sido creados y cargados todos los Data Mart, se procederá a su integración con el depósito.

Dentro de las ventajas de implementar un Data Mart a un negocio se han seleccionado las siguientes:

- Son simples de implementar.
- Se utiliza poco tiempo de construcción y puesta en marcha.
- Permiten manejar información confidencial.
- Reflejan rápidamente sus beneficios y cualidades.
- Reducen la demanda del depósito de datos.

#### 4.1.1.3. Instalación y Configuración del Producto

Para el desarrollo de este proyecto se han seleccionado e instalado las siguientes herramientas de la línea de Pentaho. Por las siguientes comparaciones con otras herramientas.

Para la selección de la plataforma BI se comparó Pentaho BI Open Source y Qlikview BI, Orange.

*Tabla 3: Descripción comparativa entre diferentes Tecnologías*

Pentaho BI Open	Source Qlikview BI	Orange
Replica el modelo de Suite BI completa ofrecida tradicionalmente por Business Objects, Cognos, Microstrategy, etc. Pero basado en el Open Source.	Es muy útil para la elaboración de prototipos rápidos, con un Modelo de Datos sencillo y enfoque de Cuadro de Mando pero pierde eficiencia cuando se desea pasar a un sistema BI más completo con Informes, análisis OLAP, Data Mining, etc.	Es un software open source en machine learning y visualización de los datos para novatos y expertos.
Su versión Community, la más extendida y utilizada, no tiene coste.	Su precio es sensiblemente inferior respecto a Business Objects, Cognos, Microstrategy, etc.	Su gran particularidad es que puedes crear interesantes análisis y visualización de los datos arrastrando cajas.
Utiliza una aproximación de metadatos y un multidimensional centralizado.	Sus Cuadros de Mando son elegantes y sencillos de usar, sin embargo, carece de la variedad y cantidad de módulos que posee Pentaho. Carece de metadatos centralizados.	Es muy interesante este software sino quieres programar, pero puedes realizar análisis estadísticos y gráficos interesantes.
Dispone de cubos OLAP, que permiten manejar fácilmente grandes volúmenes de información.	No dispone de cubos OLAP y ofrece a cambio su tecnología asociativa que permite integrar nuevas fuentes de datos fácilmente.	Es capaz de realizar desde visualización de los datos de manera simple, histogramas, boxplots, heatmaps. Hasta hacer análisis un poco más avanzados como árboles de decisión o métodos de clustering. Pasando por análisis de regresión, pruebas estadísticas paramétricas y no paramétricas.

Dispone de una completa suite ETL, llamada PentahoData Integration	Al usar lógica asociativa, no dispone de una suite ETL.	Uno de los puntos principales es que puedes explorar los datos sin entrar en el código. Es fácil y rápido de trabajar. Una buena opción si deseas realizar rápidamente un BI.
--	---	---

*Fuente: [Pentaho 2017, orange.birolab.si, Source Qlikview BI]*

Teniendo en cuenta la comparación del cuadro anterior se optó por considerar las herramientas de Pentaho como:

- Pentaho Open BI Suite
- Schema Workbench Pentaho
- Pentaho Designer
- Pentaho Dashboard Editor
- Pentaho Data Integration

## **4.2. Diseño de modelamiento dimensional**

### **4.2.1 Modelamiento Dimensional**

El modelo dimensional se basa en hechos (Facts) y dimensiones presentando un modelo lógico para el Data Mart que representa los diferentes requerimientos que requiere la Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo.

#### 4.2.1.1 Modelo Estrella del registro de transacciones

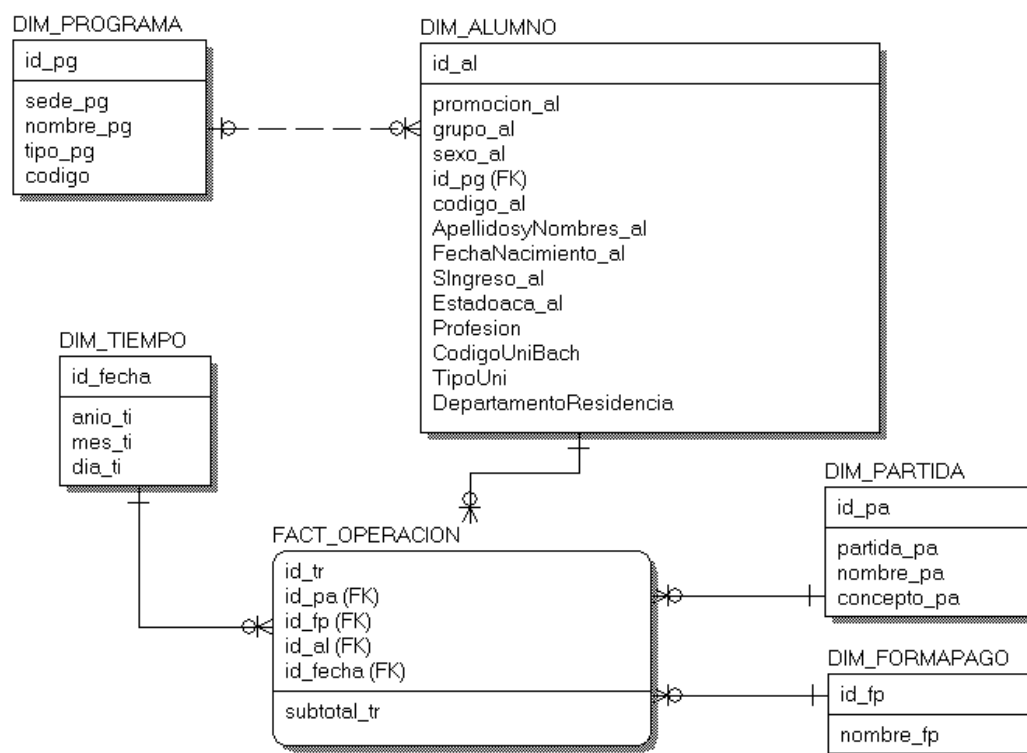


Figura NRO. 21: Modelo estrella usado en el registro de transacciones

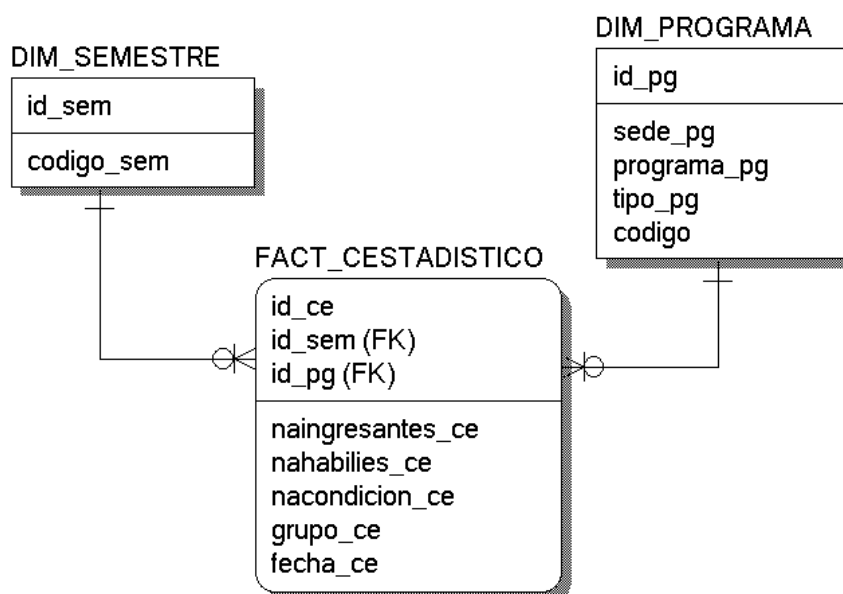


Figura NRO. 22: Modelo estrella usando como cuadro estadístico semanal

Para poder obtener los reportes, indicadores, o perfiles deseados a partir del registro de transacciones se determinó 01 tabla de hechos. Con el nombre FACT\_OPERACION, La siguiente tabla es necesaria para cumplir requerimientos funcionales indicados.

### **FACT\_OPERACION**

Permite mostrar las transacciones que realizan los alumnos en cada uno de los programas por tipo, semestre de ingreso, promociones y diversas formas de pago según el tiempo de fueron realizados.

#### **Métricas asociadas**

**Importe:** Métrica que se obtiene sumando los pagos realizados por los alumnos de acuerdo al programa, semestre de ingreso, promoción y grupo.

**Cantidad de alumnos:** Métrica que se obtiene del conteo de alumnos ingresantes que se han inscrito por programa.

#### **Dimensiones asociadas**

Alumno, tiempo, programa, promoción, grupo, forma de pago.

#### **Cálculos:**

**Importe:** es la sumatoria de las transacciones realizadas por los alumnos de acuerdo al programa de ingreso.

**Cantidad de alumnos:** métrica que se genera a partir de un conteo de los alumnos que se han inscrito en uno de los muchos procesos de admisión.

#### **Periodicidad de carga**

Se deberá realizar, la actualización de la base de datos, cada semana a fin de conocer la cantidad de transacciones realizadas por programa para iniciar las labores académicas y tomar decisiones de tipo financiera.

#### **Granularidad de los datos**

El nivel de granularidad se define con el usuario final de negocio que realizará el trabajo de análisis y toma de decisiones en función a la información que se le presente con la solución. Tener esta información por programa, año, semestre, trimestre hasta niveles



de día o incluso hora, genera valor, puesto que permite realizar comparaciones entre períodos de tiempo y determinar los niveles de crecimiento, posicionamiento, eficiencia, eficacia, rendimiento, entre otros indicadores que permiten a las organizaciones poder controlar sus recursos humanos, recursos de TI y procesos de trabajo alineados a sus políticas, procedimientos y objetivos estratégicos.

#### **4.2.1.2 Dimensiones del modelo de datos lógico**

- **Dimensión Tiempo**

Representa a cada una de las operaciones en línea que realizan los alumnos de la UNPRG por diversos conceptos en el Banco de la Nación, Atributos asociados:

Año, mes, día.

Mes, es tiempo en el que se realizan el pago de la mensualidad de alumnos.

Día, es el tiempo en el que se realizan diversos pagos por conceptos de matrícula, constancias, certificados, etc.

#### **Proceso de carga**

Se deberá realizar, la actualización de la dimensión tiempo, cada fin de semana con la finalidad de poder iniciar las clases los sábados y domingo todos los fines de semana.

- **Dimensión Alumno**

Representa a un alumno ingresante por un proceso de admisión. Contiene información muy general de cada persona con datos, tales como, IdAlumno, programa, sexo, fecha de nacimiento, etc. Atributos asociados:

IdAlumno, es el identificador único del alumno.

Promoción, es el nombre del semestre de ingreso a cual postulo.

Grupo, si los alumnos superar la cantidad mínima por programa se crear mas grupo en diversas aulas como siempre sucede.

EAcademico, es el estado académico actual en el que se encuentra.

SIngreso, indica el semestre de ingreso por el Proceso de Admisión realizado.

Universidad, indica la universidad de procedencia donde obtuvo el grado académico.

TipoUniversidad, si la Universidad es Nacional o Particular.

- **Dimensión Partida**

Representa a los diferentes conceptos que los alumnos pagan en el banco de la nación, como matrícula, pensión de enseñanza, constancia, etc. correspondiente a la Escuela de Posgrado.

- **Dimensión Forma de Pago**

Representa a los diferentes ingresos monetarias que tiene la EPG, desde pagos en efectivo, hasta pagos a través del banco de la nación, con tarjeta de crédito o débito, etc. correspondiente a la Escuela de Posgrado.

**Proceso de carga**

Se deberá realizar, la actualización de la dimensión forma de pago cada semestre por año.

- **Dimensión Programa**

Representa a los diferentes Programas que ofrece la EPG, para diferentes profesionales que requieran profundizar y adquirir nuevos conocimientos y obtener un grado más en su curriculum profesional.

**Proceso de carga**

Se deberá realizar, la actualización de la dimensión programa una vez al año.

#### **4.2.2 Diseño Físico**

En el siguiente diseño se puede apreciar como las tablas (dimensiones y tabla de hechos) del modelo propuesto constituyen el Data Mart del registro de transacciones de la Escuela de Posgrado de la Universidad Nacional Pedro Ruiz Gallo.

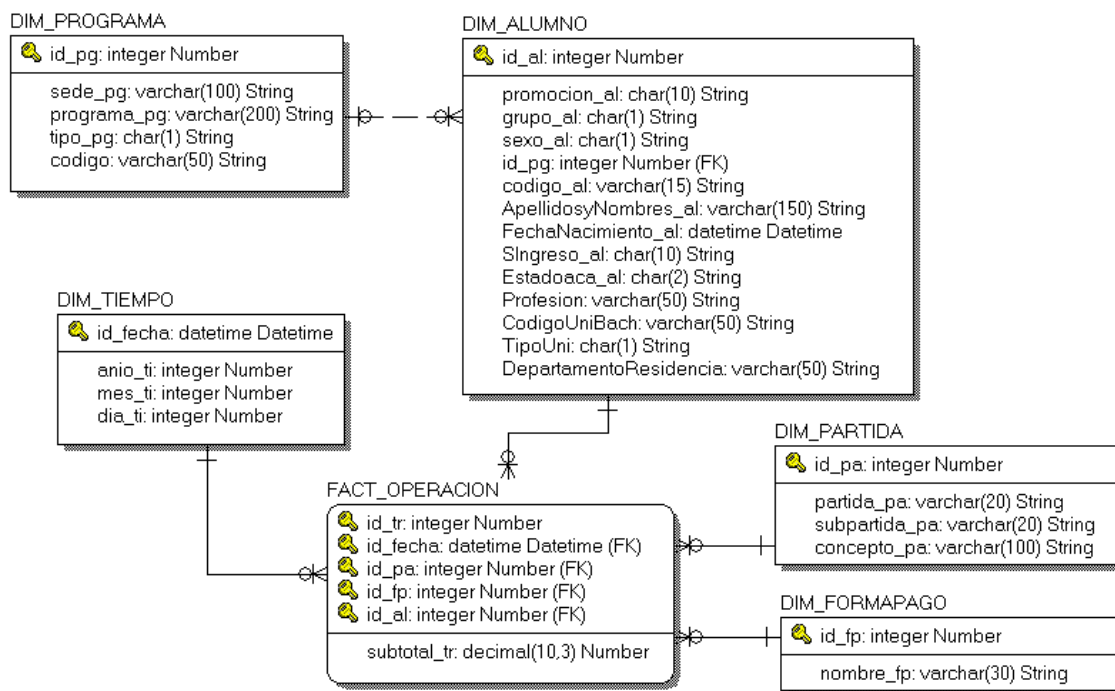


Figura NRO. 23: Diseño físico de Base de Datos transaccional

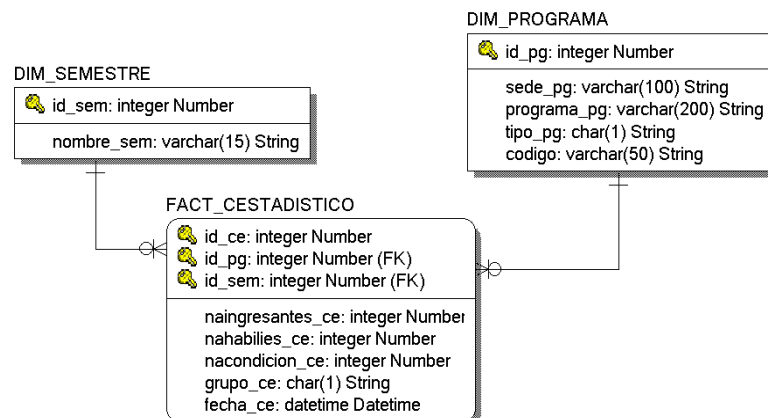


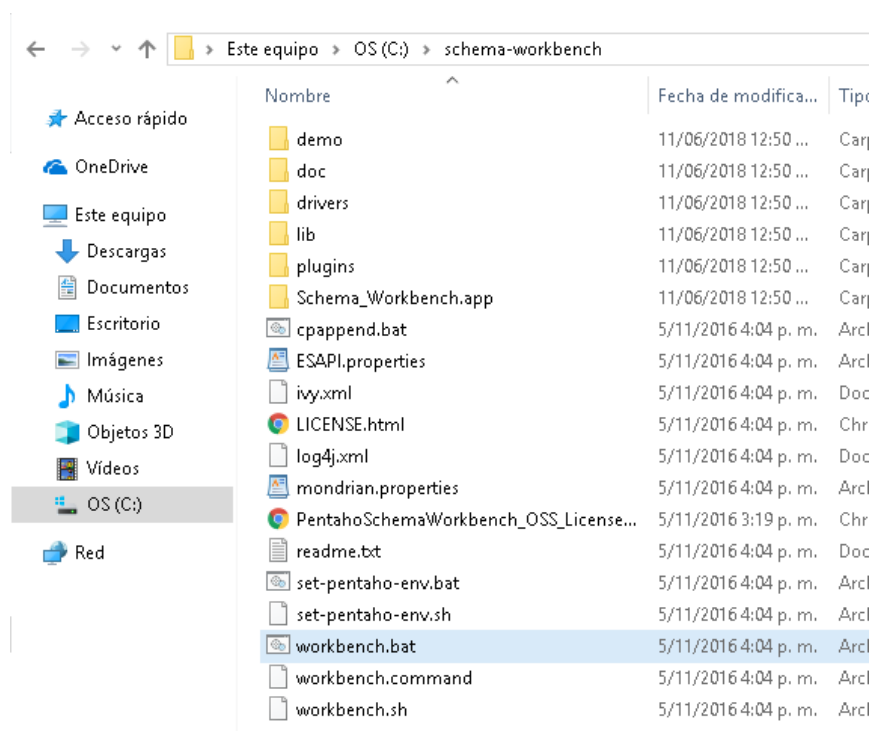
Figura NRO. 24: Diseño físico de Base de Datos estadístico

### 4.2.3 Creación de Cubos

Para la creación del cubo se utilizó la herramienta Schema Workbench Pentaho, Con esta herramienta crearemos un cubo dentro del esquema y le daremos un nombre llamado perfil\_financiero y posteriormente asignaremos al cubo dimensiones, jerarquías, niveles y miembros.

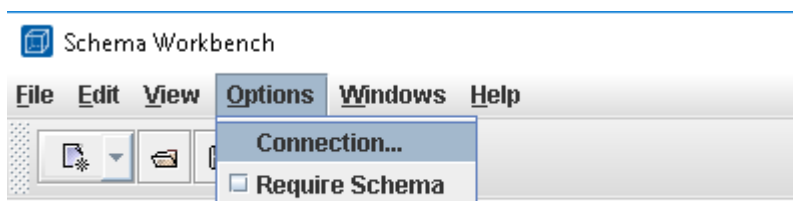
### Procedimientos para crear el cubo Perfil Financiero

**Paso 1:** Ejecutar el aplicativo **Schema Workbench** a través del archivo workbench.bat



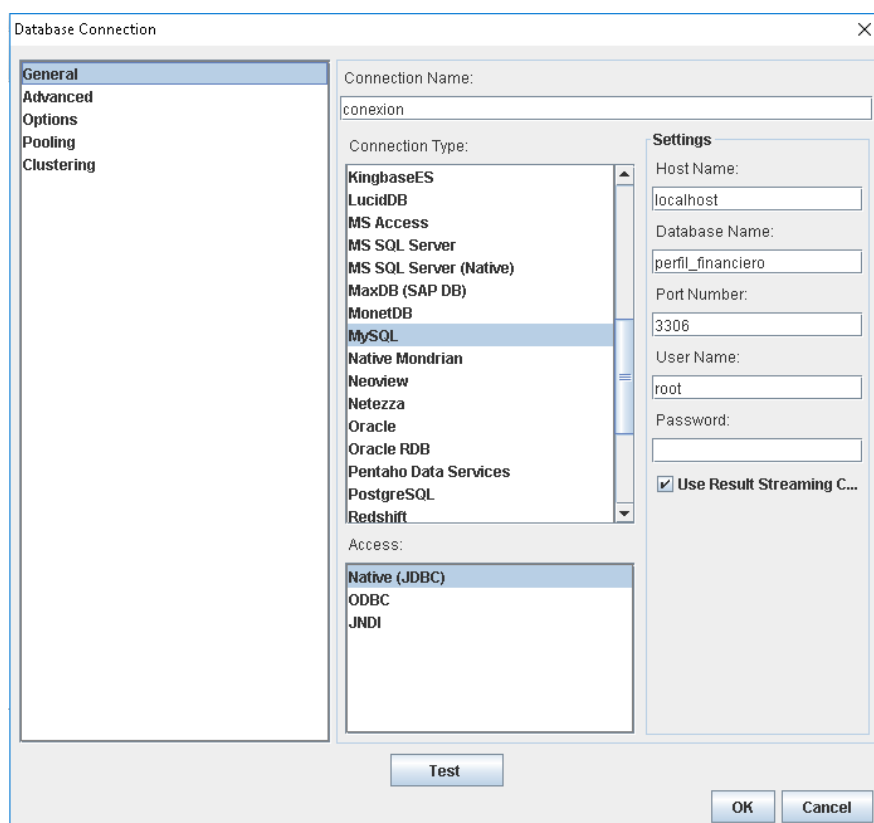
**Figura NRO. 25:** Ejecutando la herramienta schema Workbench portable a través del archivo workbench.bat, descargado de la pagina de pentaho

**Paso 2:** Crear una conexión con la base de datos de MySQL, con los parámetros deseados



*Figura NRO. 26: Menú opción para la conexión con la base de datos MySQL*

**Paso 3:** Configurar la conexión con los parámetros deseados



*Figura NRO. 27: Configurando conexión con MySQL*

#### Paso 4: Creando un nuevo esquema para crear el cubo

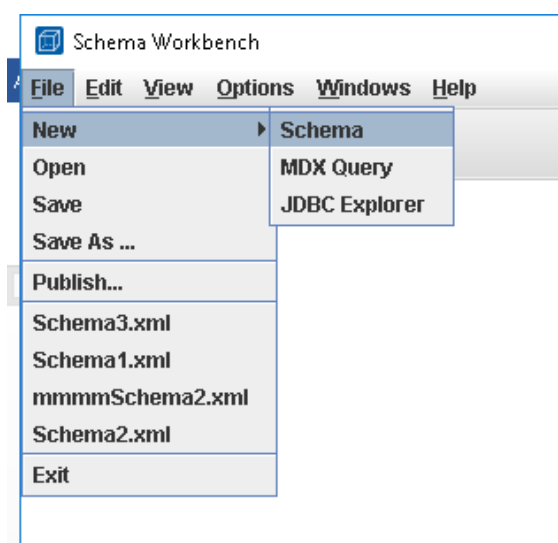


Figura NRO. 28: Crear nuevo esquema a través del menú File

#### Paso 5: Creando el cubo para el desarrollo del proyecto

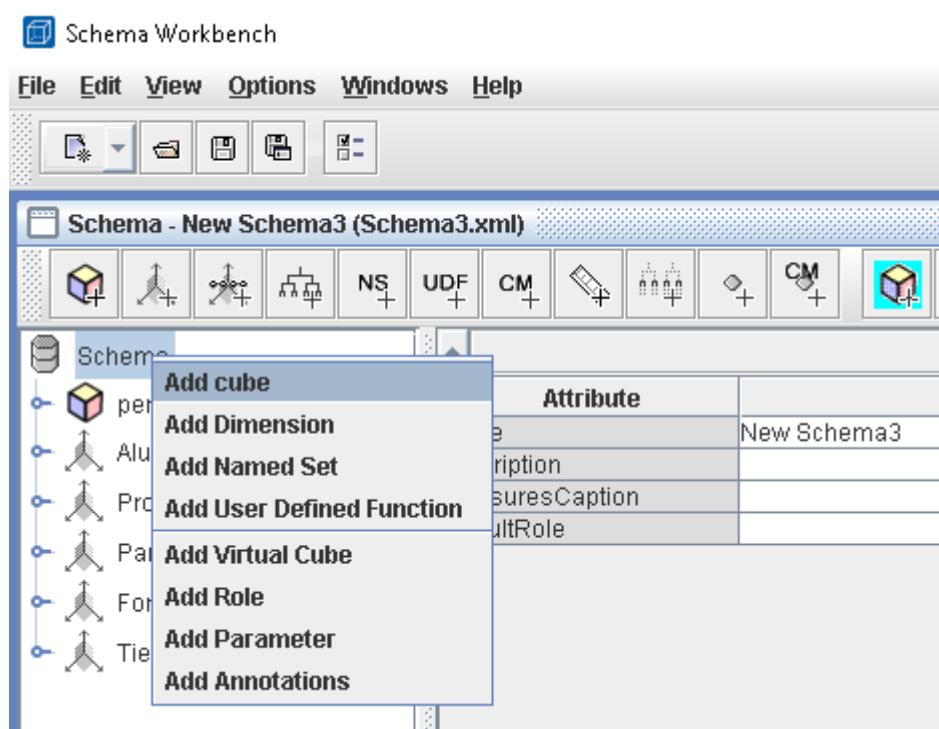


Figura NRO. 29: Creación del nuevo cubo

**Paso 6:** Asignaremos un nombre al cubo que utilizaremos para nuestro proyecto

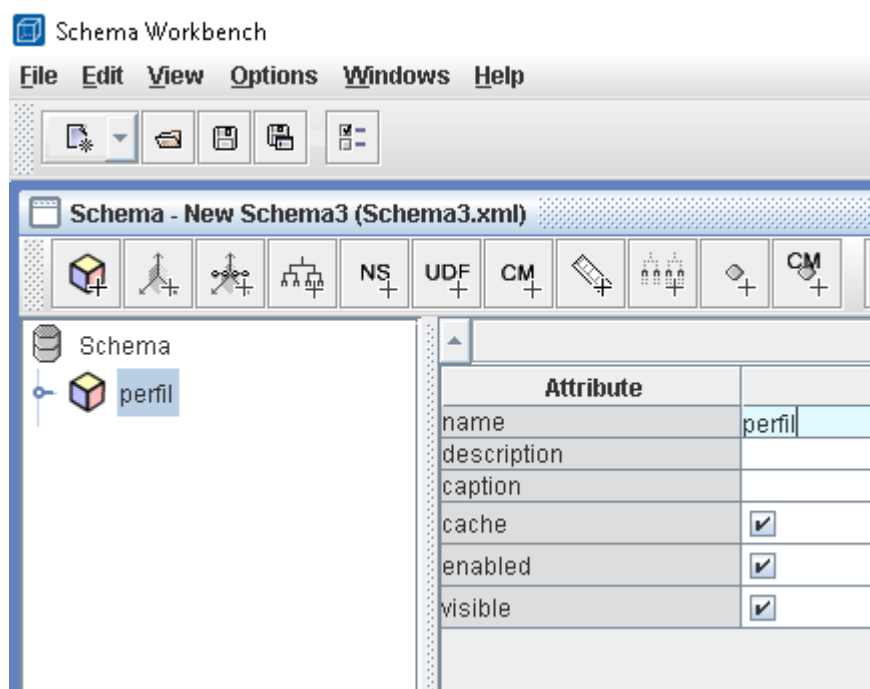


Figura NRO. 30: Aquí se realiza la asignando de nombre a nuestro cubo

**Paso 7:** Agregaremos una tabla a nuestro cubo, esta tabla será la tabla de hechos llamada **FACT\_OPERACION**, tabla principal del cual dependen nuestras dimensiones

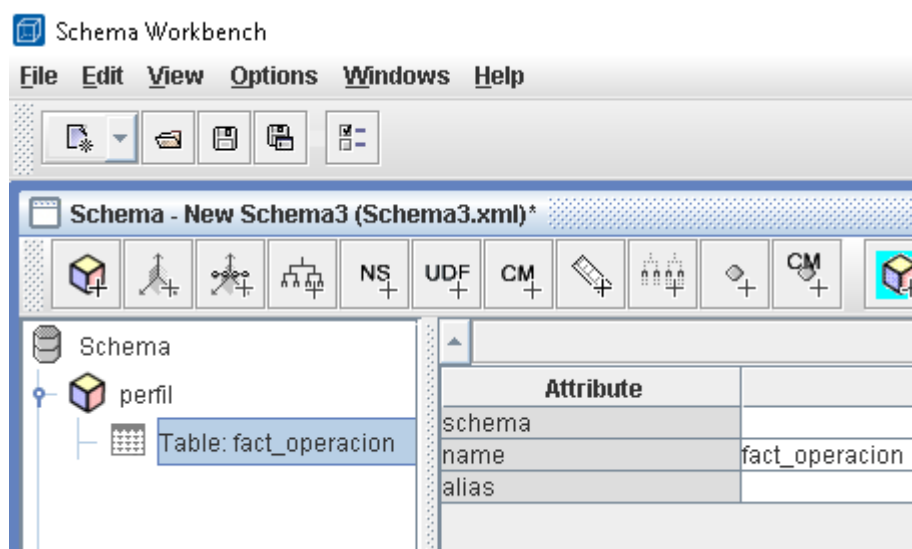


Figura NRO. 31: Agregar nueva tabla de hechos fact\_operacion a nuestro cubo

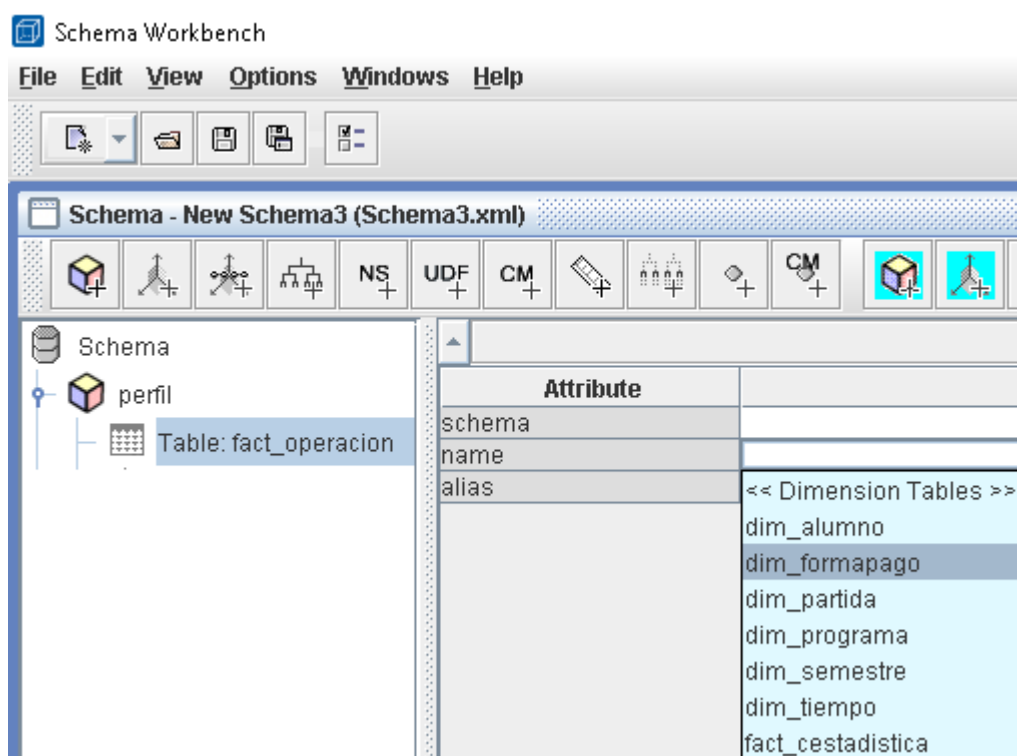


Figura NRO. 32: Seleccionar el nombre de la tabla de hechos, de una tabla de hechos creada

### Paso 8: Agregando las dimensiones a nuestro cubo perfil

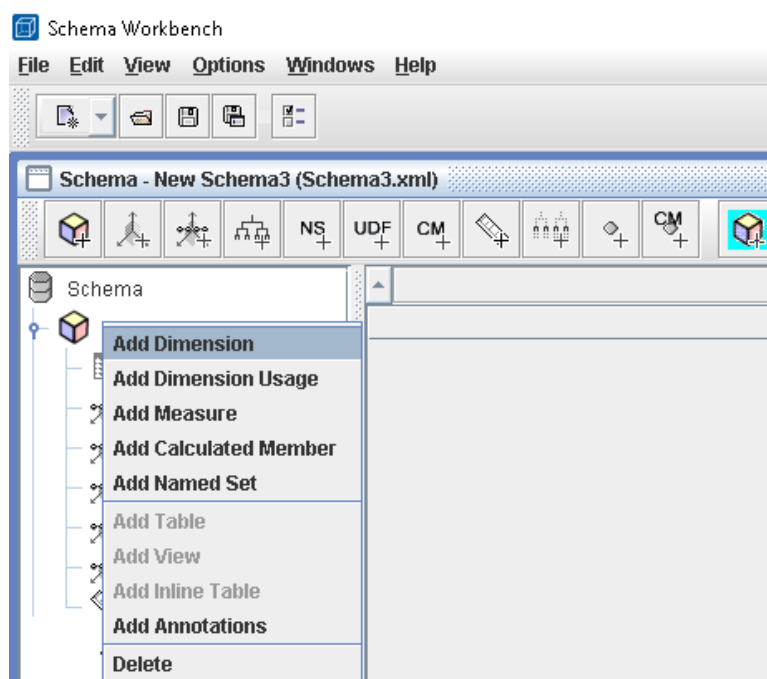


Figura NRO. 33: Agregando la dimensión a nuestro cubo perfil



## Paso 9: Asignando un nombre a la dimensión

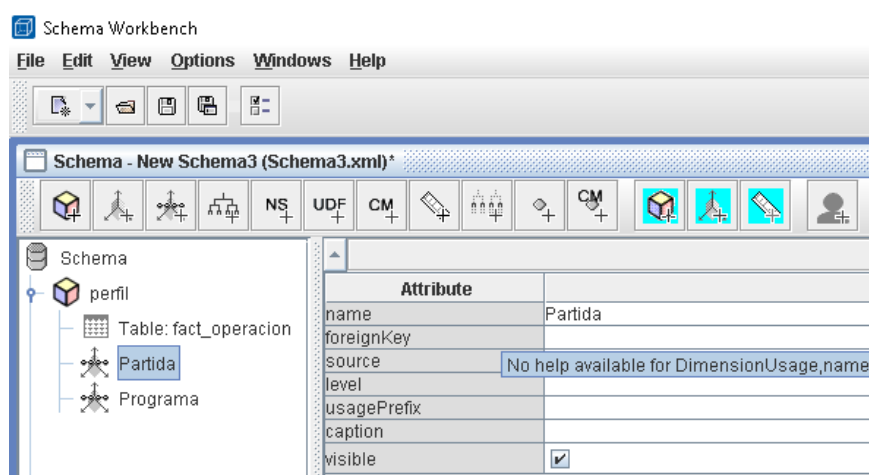


Figura NRO. 34: Asignar nombre a la dimensión

**Paso 10:** Desplegamos en la dimensión Alumno y nos ubicamos debajo de dibujo Jerarquía y le asignamos un nombre para nuestro caso se llamará ApellidosyNombres y agregamos la tabla que corresponda a la dimensión.

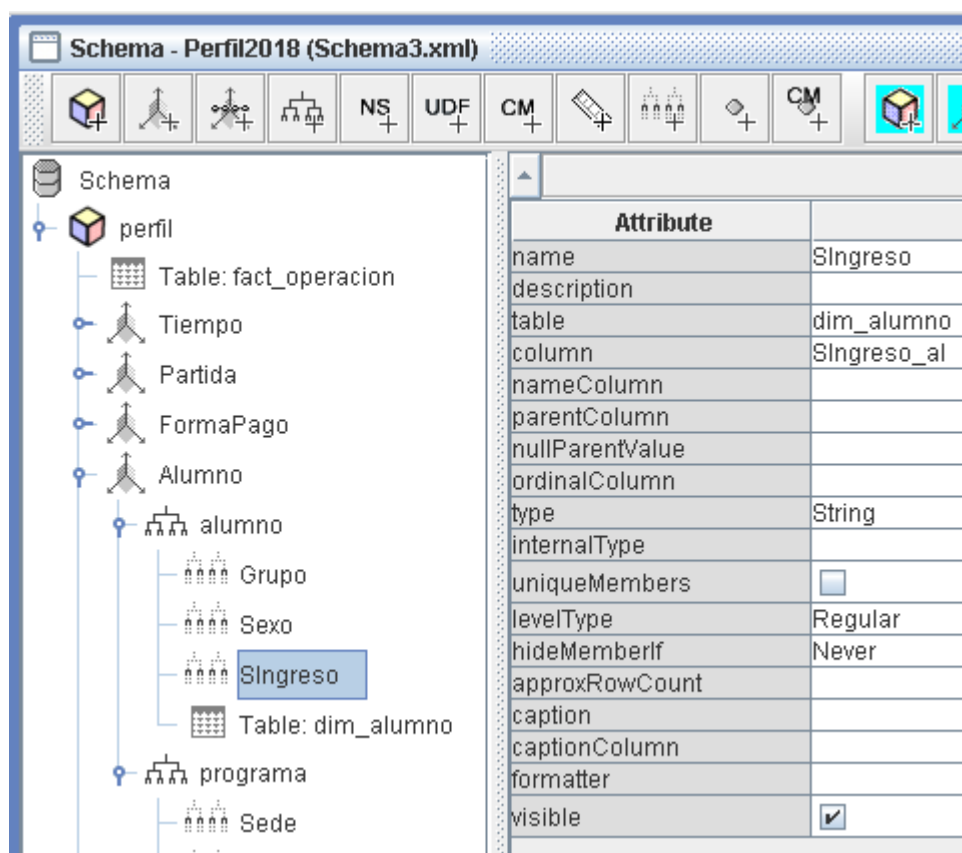


Figura NRO. 35: Agregando una tabla a la dimensión correspondiente

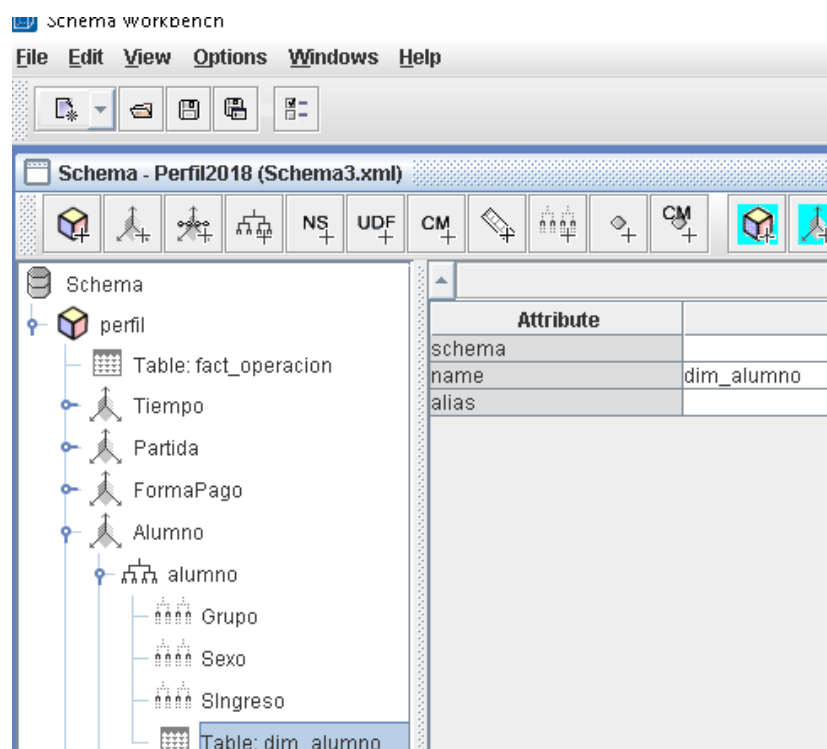


Figura NRO. 36: Seleccionando nombre a la tabla asignada a la dimensión Alumno

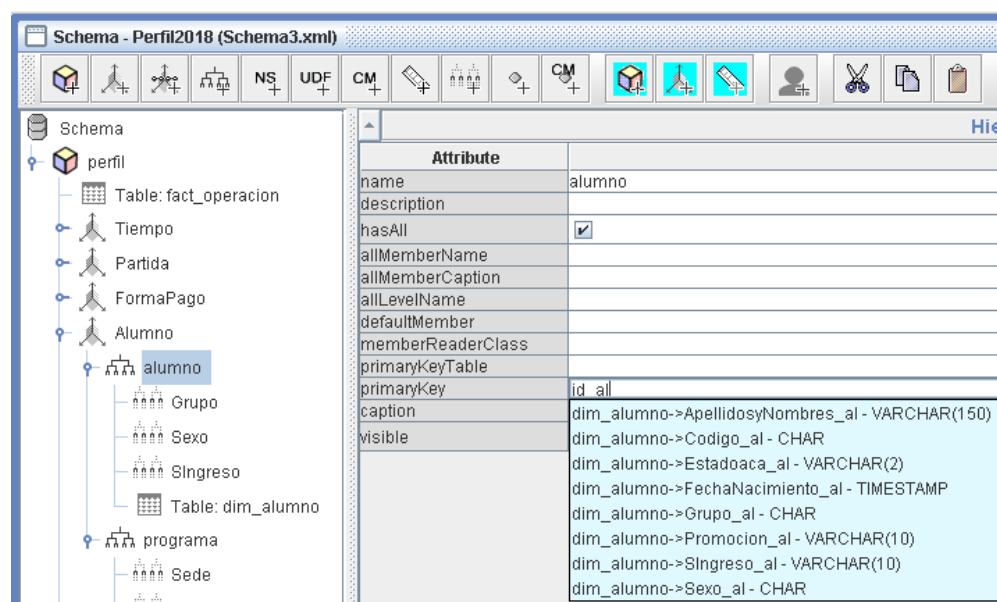


Figura NRO. 37: Asignar PRIMARY KEY en la jerarquía

## Paso 12: Asignamos la **ForeignKey** en la dimensión según corresponda

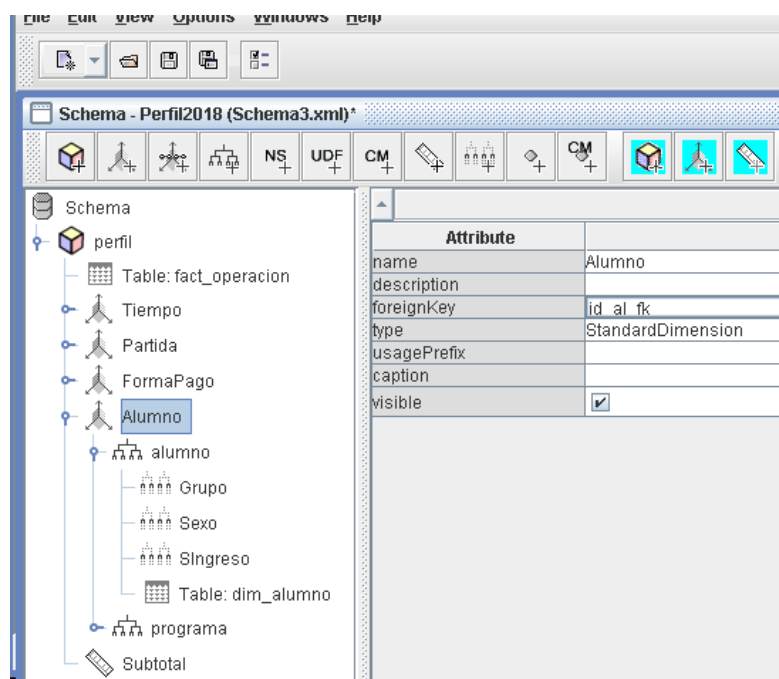


Figura NRO. 38: Asignamos la **ForeignKey** en la dimensión Alumno

## Paso 13: Agregamos los niveles para la Jerarquía

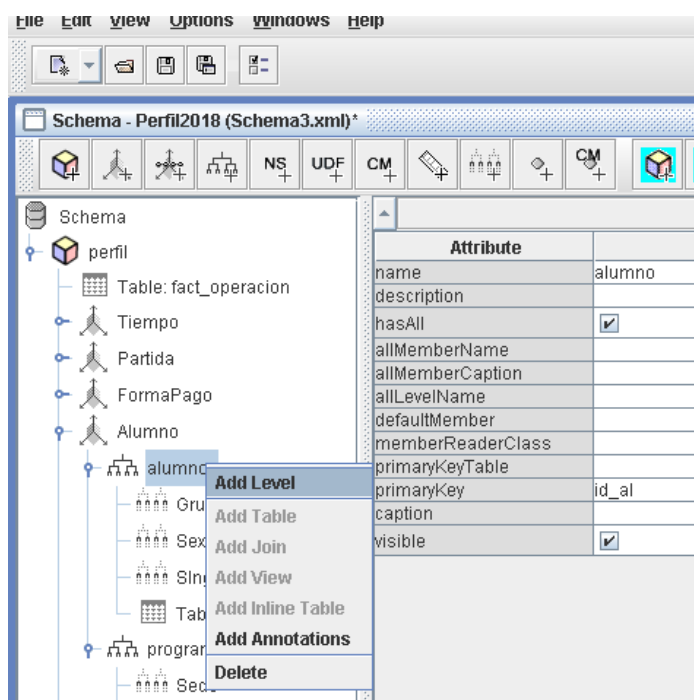


Figura NRO. 39: Agregamos los niveles a la Jerarquía

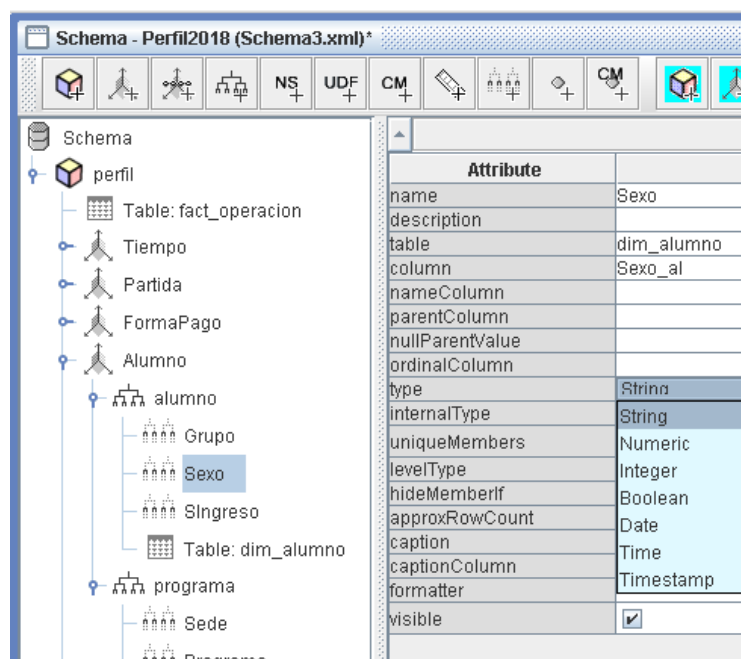


Figura NRO. 40: Asignar el tipo de dato

**Paso 14:** Agregamos las medidas que se van a utilizar en el cubo para hacer los cálculos.

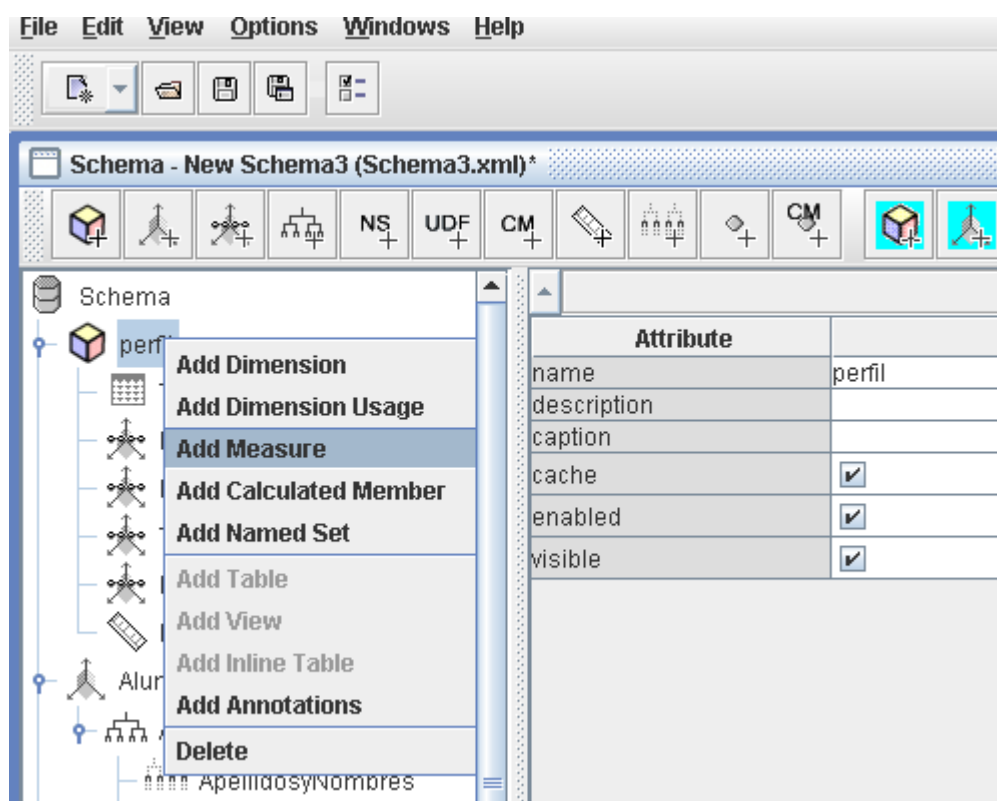


Figura NRO. 41: Agregar las medidas que se utilizaran en los cálculos del cubo

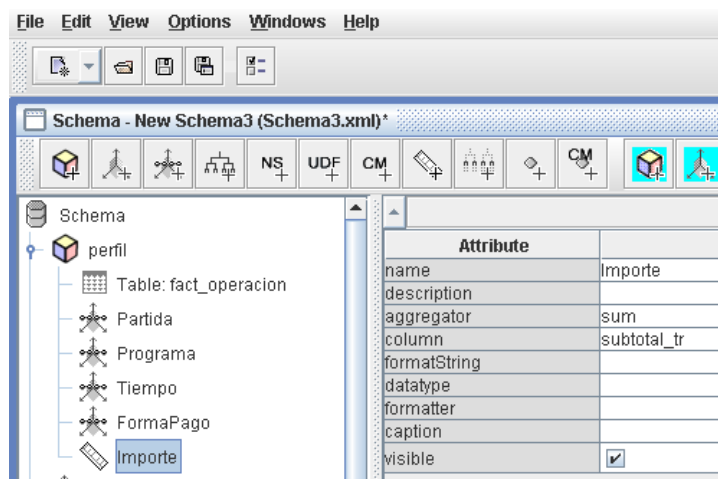


Figura NRO. 42: Asignar un nombre a la medida

#### 4.2.4 Diseño de ETL

El diseño ETL se realiza a través de Kettle que es una herramienta, dentro del paquete Pentaho Data Integration, de gran potencial para la realización de procesos de Extracción, Transformación y Carga de datos (ETL). Asimismo

Dentro de Spoon podemos ver en la parte superior una barra de herramientas donde están los ejecutables de las Transformaciones y los Trabajos. En la parte de la izquierda tendremos dos pestañas (“View” y “Design”), desde Design podremos diseñar nuestros Trabajos o Transformaciones. La pestaña View nos permite ver todos los objetos que hayamos utilizados en esa Transformación o Trabajo, tales como conexiones, pasos, saltos, particiones de esquemas, etc.

#### Paso 1: Creamos una nueva Transformación

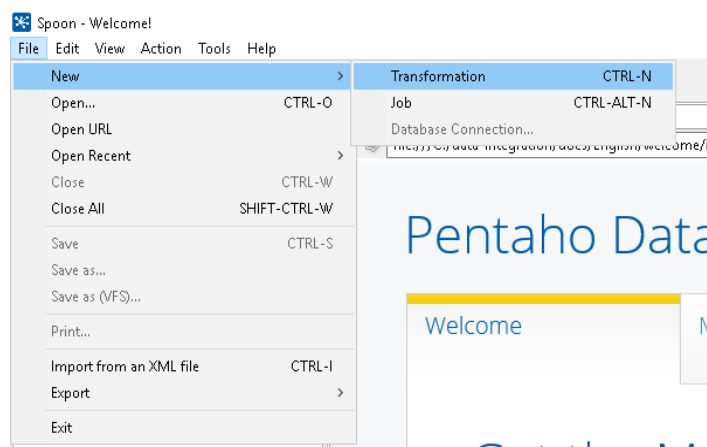


Figura NRO. 43: Creación de una nueva transformación

## Paso 2: Creamos una nueva conexión a la base de datos MySQL en Database connection

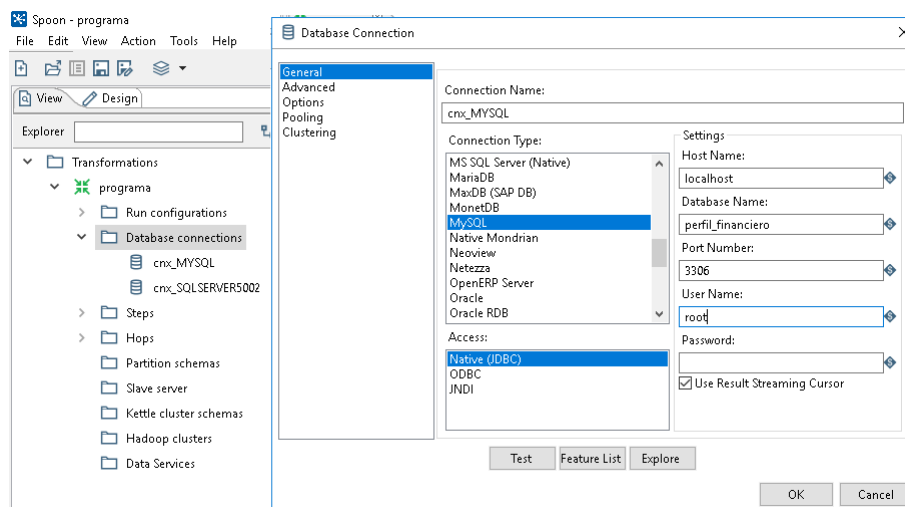


Figura NRO. 44: Creación de una nueva conexión a la base

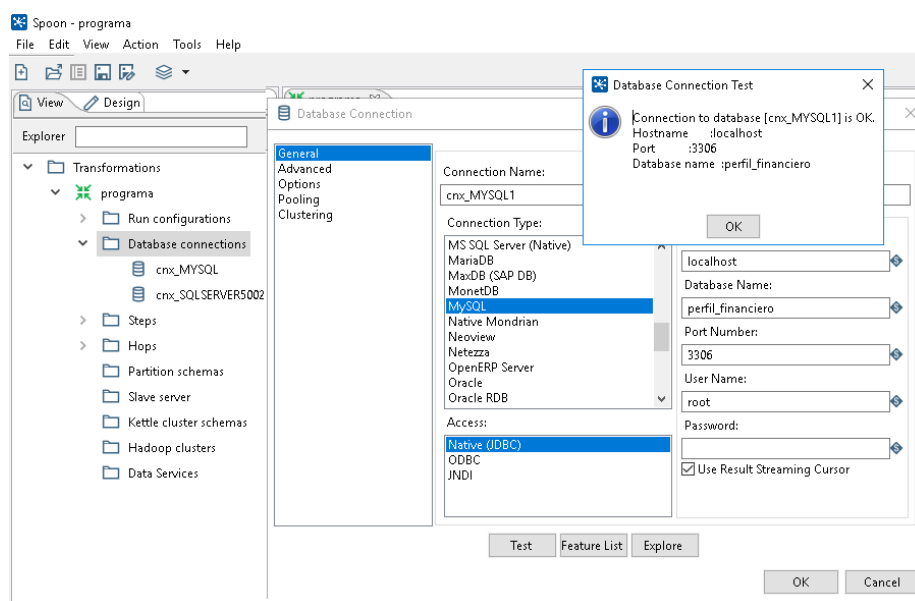


Figura NRO. 45: Prueba de la conexión a la base de datos MySQL

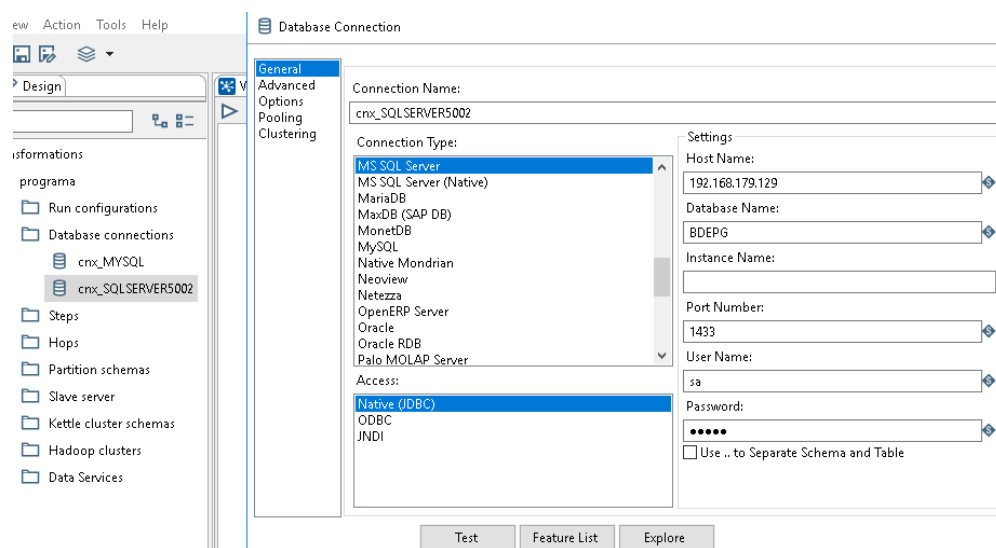


Figura NRO. 46: Prueba de la conexión a la base de datos MSSQL

**Paso 3:** Hacemos clic en la carpeta Input para seleccionar y arrastrar CSV file input

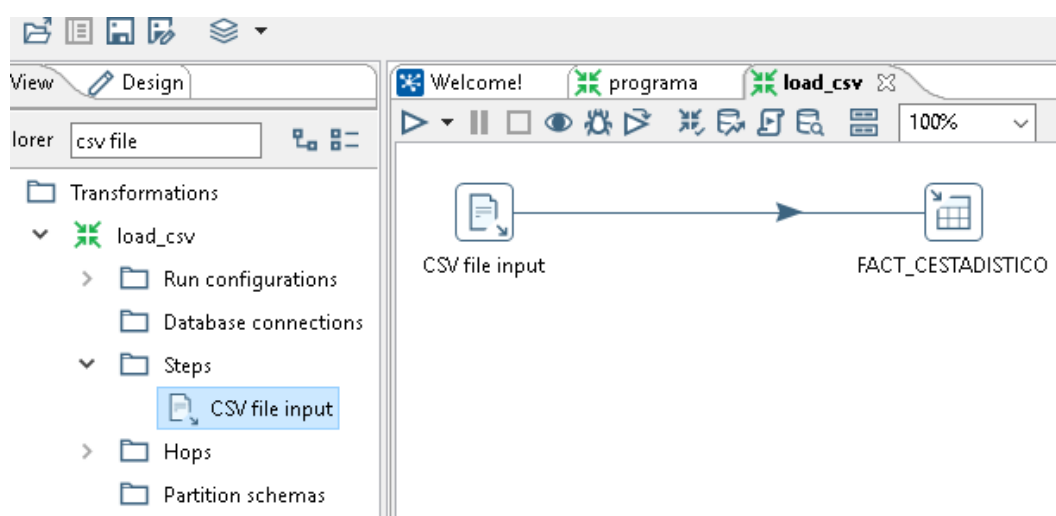


Figura NRO. 47: CSV file input

**Paso 4:** Hacemos doble clic en CSV file input y seleccionamos el archivo con extensión “.csv” que es brindado por el área académica de la EPG.

Step name: CSV file input

Filename: C:\Users\ASUS\Desktop\TESIS 2018\proyecto\Estadistica.csv

Delimiter: ;

Enclosure: "

NIO buffer size: 50000

Lazy conversion? ☒

Header row present? ☒

Add filename to result ☐

The row number field name (optional):

Running in parallel? ☐

New line possible in fields? ☐

File encoding: UTF-8

#	Name	Type	Format	Length	Precision	Currency	Decimal	Group
1	id_ce	Integer	#	15	0	S/.	,	.
2	naingresantes_ce	Integer	#	15	0	S/.	,	.
3	nahabiles_ce	Integer	#	15	0	S/.	,	.
4	nacondicion_ce	Integer	#	15	0	S/.	,	.
5	id_sem_fk	Integer	#	15	0	S/.	,	.
6	id_pg_fk	Integer	#	15	0	S/.	,	.
7	grupo_ce	String		1		S/.	,	.
8	fecha_ce	Date	yyyy-MM-dd			S/.	,	.

Buttons: Help, OK, Get Fields, Preview, Cancel

**Figura NRO. 48:** Selección del archivo CSV



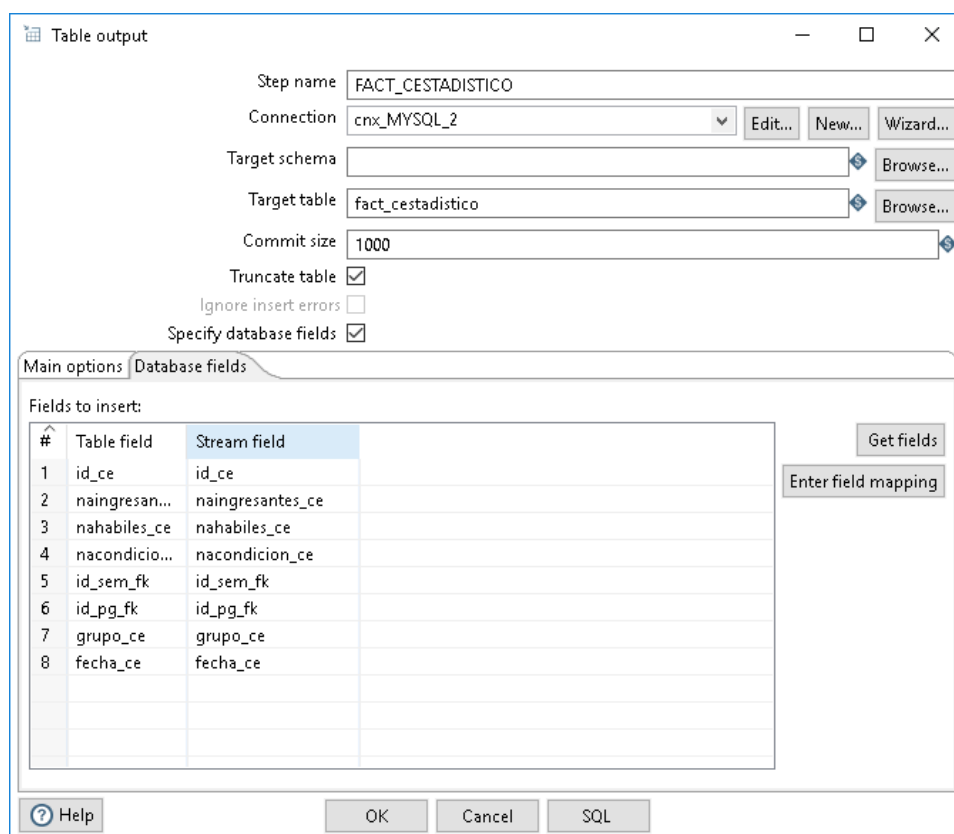


Figura NRO. 49: En este paso se configura la tabla de destino, hacia donde irán los datos que se importaran

### ETL: Carga de datos hacia la tabla Fact\_CEstadistico

Este proceso muestra la relación para la extracción de los datos desde el archivo .CSV llamado Estadística hacia su respectiva tabla Fact\_CEstadistico en la Base de Datos de MySQL que sería el Data Mart (Perfil\_Financiero).

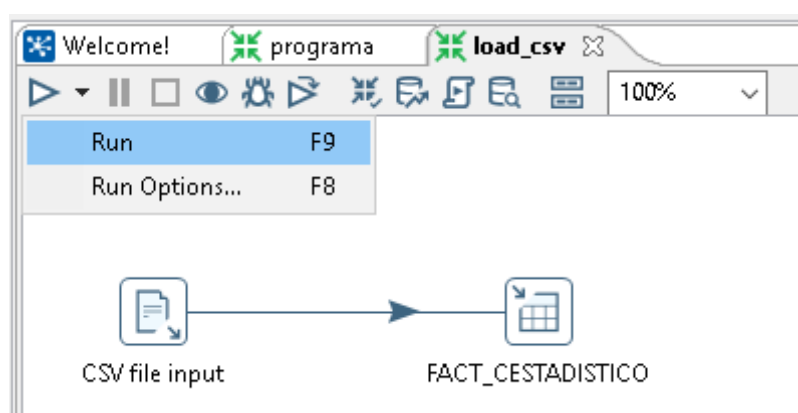


Figura NRO. 50: A continuación se muestra como se importa la información hacia su tabla de destino

## ETL: Carga de datos hacia la Dimensión Alumno

Este proceso muestra cómo se enviara la información del motor de base de datos de SQLServer de la tabla SQL\_ALUMNO hacia MySQL donde se encuentra la tabla de destino dimension\_alumno de la base de datos perfil\_financiero(Data Mart).

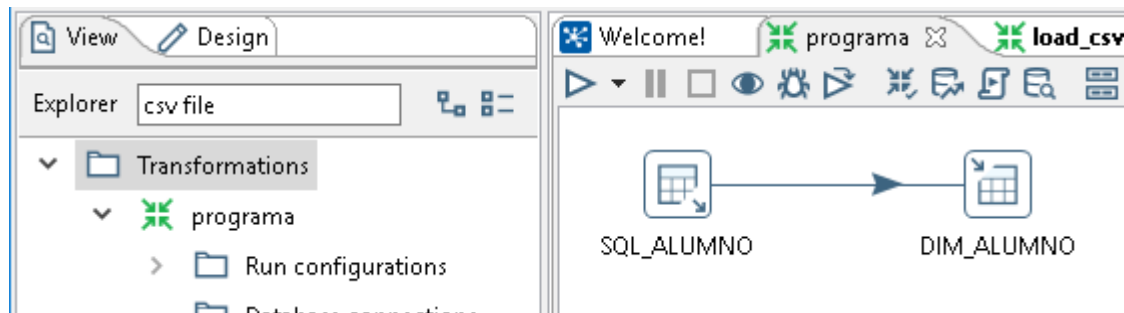


Figura NRO. 51: Diseño de la forma como se importara la información deseada

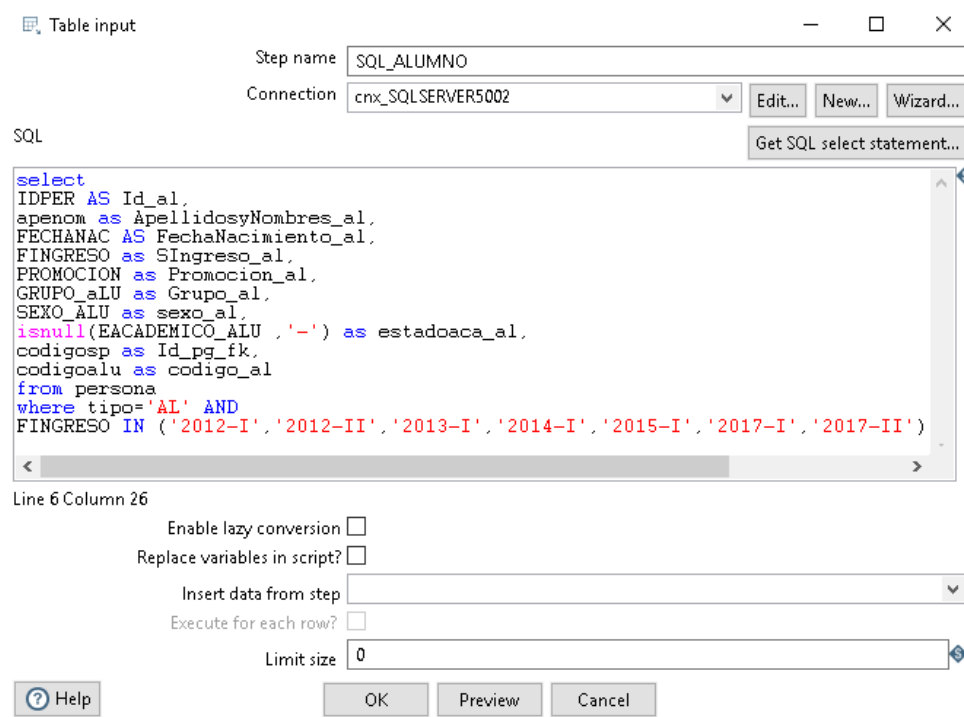


Figura NRO. 52: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración

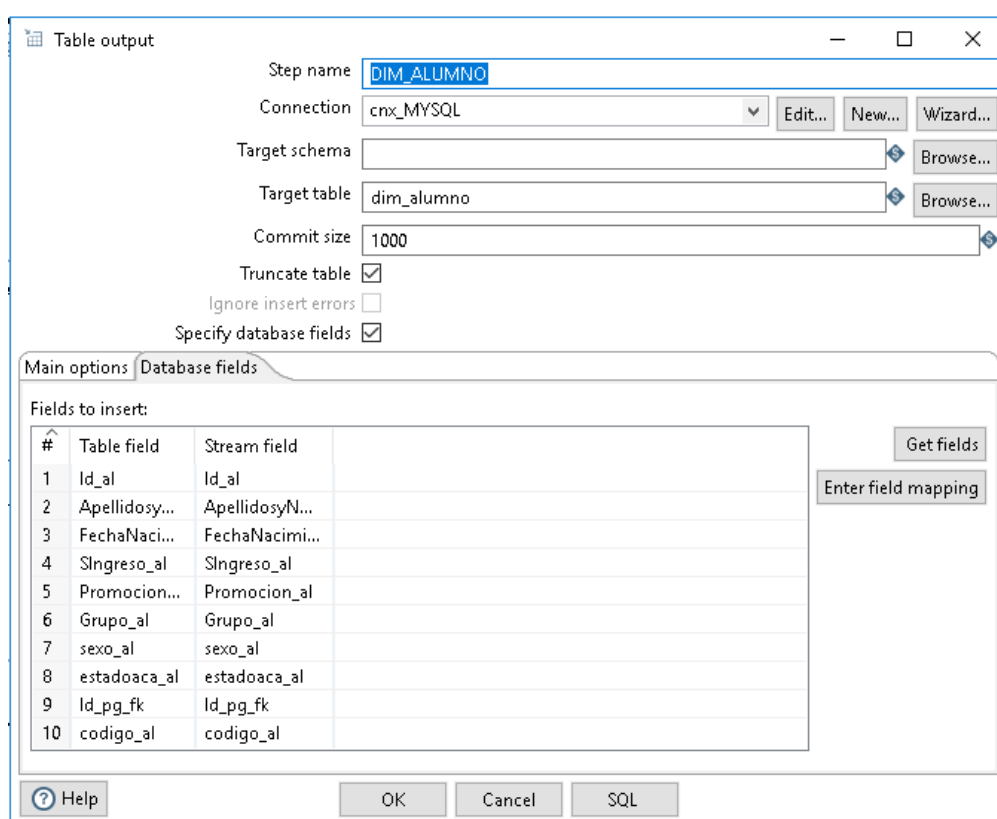


Figura NRO. 53: Tabla de destino lista y preparada para recibir la data deseada

### ETL: Carga de datos hacia la Dimensión Partida

Este proceso muestra cómo se enviara la información del motor de base de datos de SQLServer de la tabla SQL\_PARTIDA hacia MySQL donde se encuentra la tabla de destino dimension\_partida de la base de datos perfil\_financiero(Data Mart).

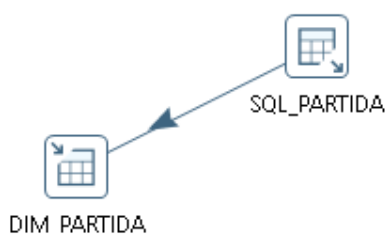


Figura NRO. 54: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim\_partida

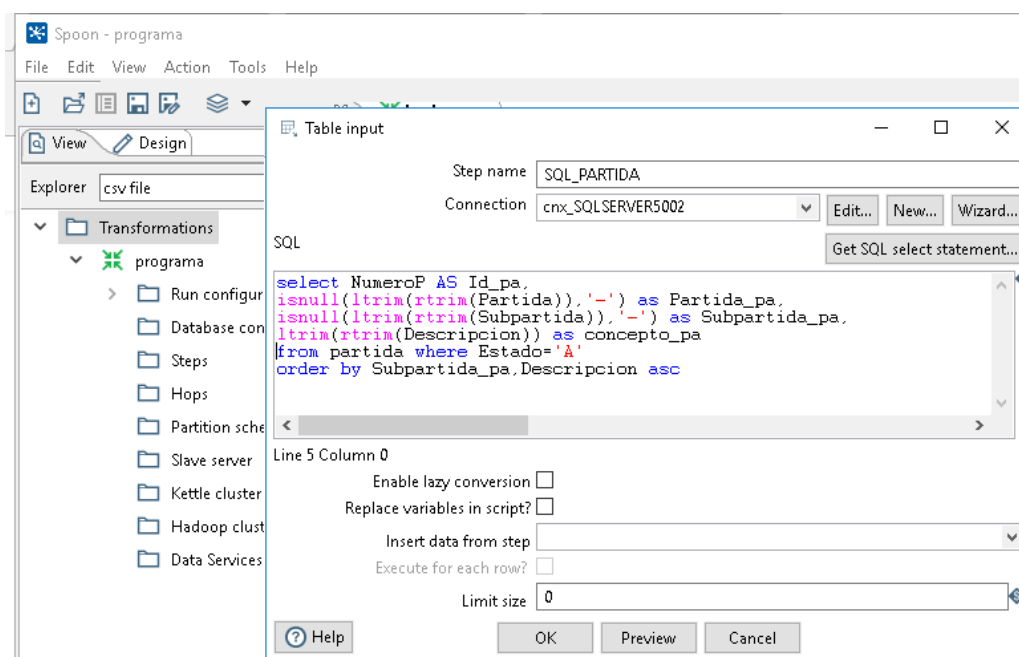


Figura NRO. 55: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración

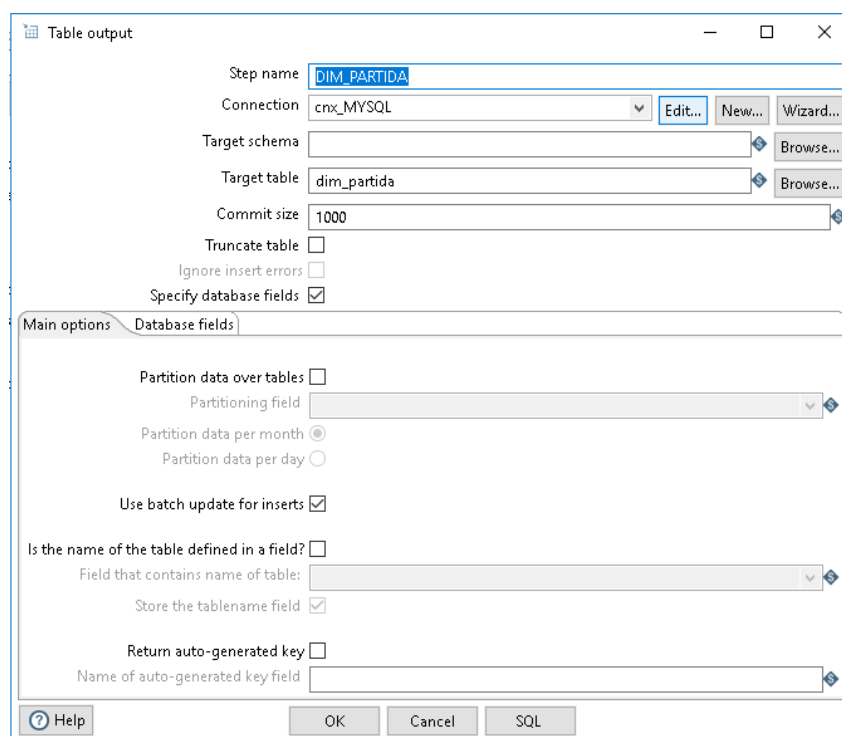
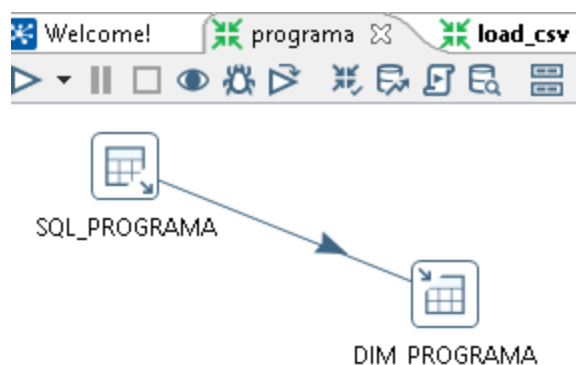


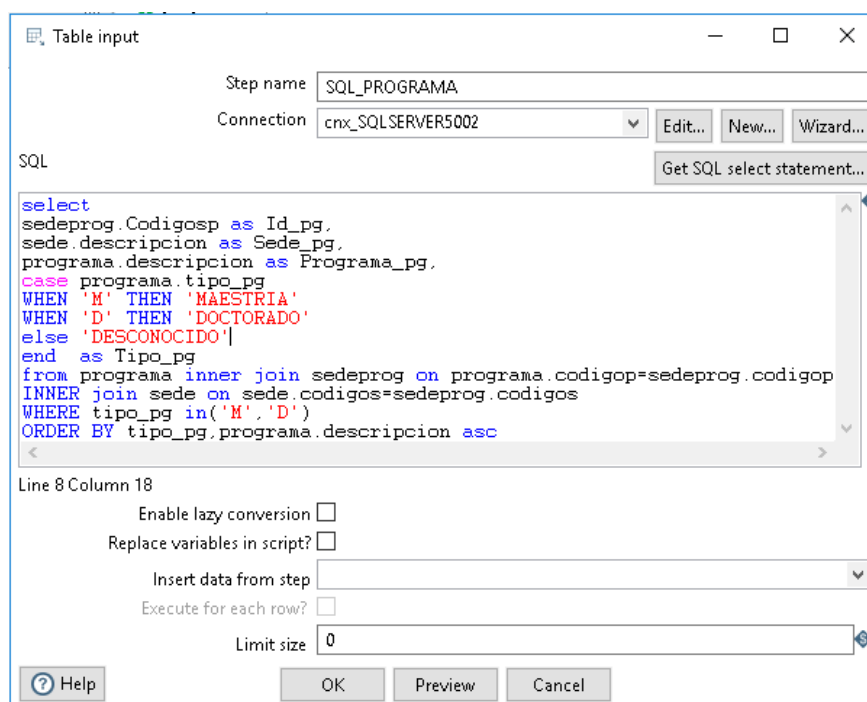
Figura NRO. 56: Tabla de destino lista y preparada para recibir la data deseada

## ETL: Carga de datos hacia la Dimensión Programa

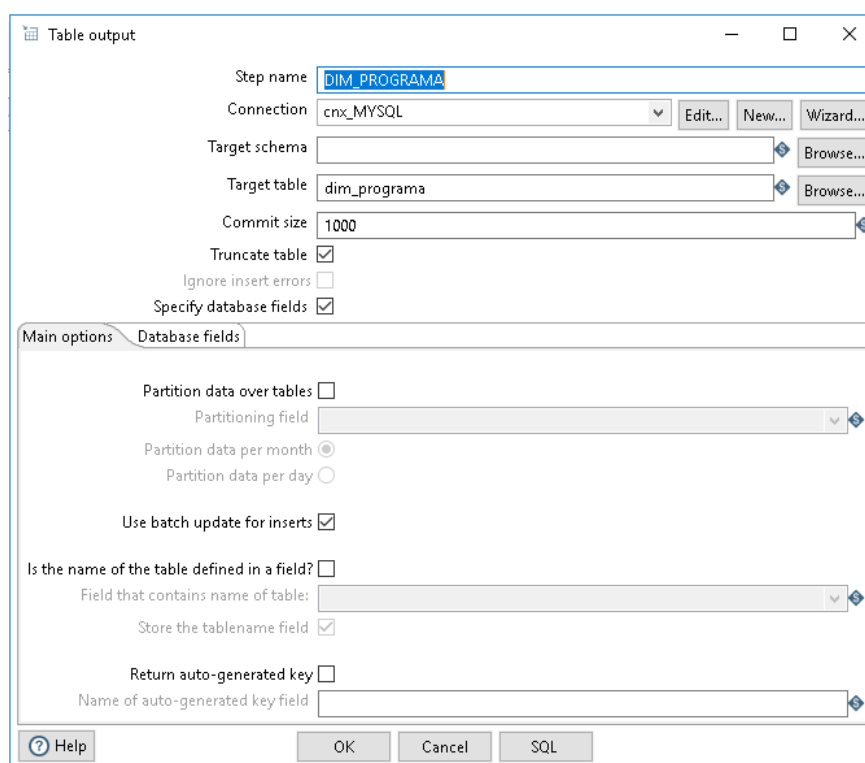
Este proceso muestra cómo se enviara la información del motor de base de datos de SQLServer de la tabla SQL\_PROGRAMA hacia MySQL donde se encuentra la tabla de destino dimension\_programa de la base de datos perfil\_financiero(Data Mart).



*Figura NRO. 57: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim\_programa*



*Figura NRO. 58: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración*



*Figura NRO. 59: Tabla de destino lista y preparada para recibir la data deseada*

### ETL: Carga de datos hacia la Dimensión Semestre

Este proceso muestra cómo se enviara la información del motor de base de datos de SQLServer de la tabla SQL\_SEMESTRE hacia MySQL donde se encuentra la tabla de destino dimension\_semestre de la base de datos perfil\_financiero(Data Mart).



*Figura NRO. 60: A continuación se muestra el diseño y procedimiento de cómo se importa la información hacia su tabla de destino dim\_semestre*

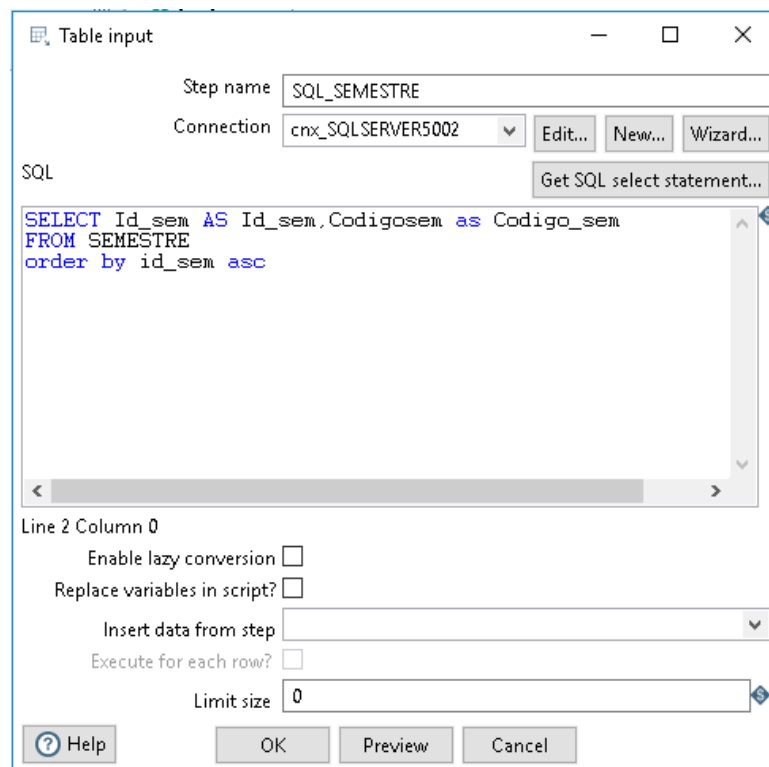


Figura NRO. 61: Tabla donde se encuentra la información de origen deseada, lista y preparada para su migración

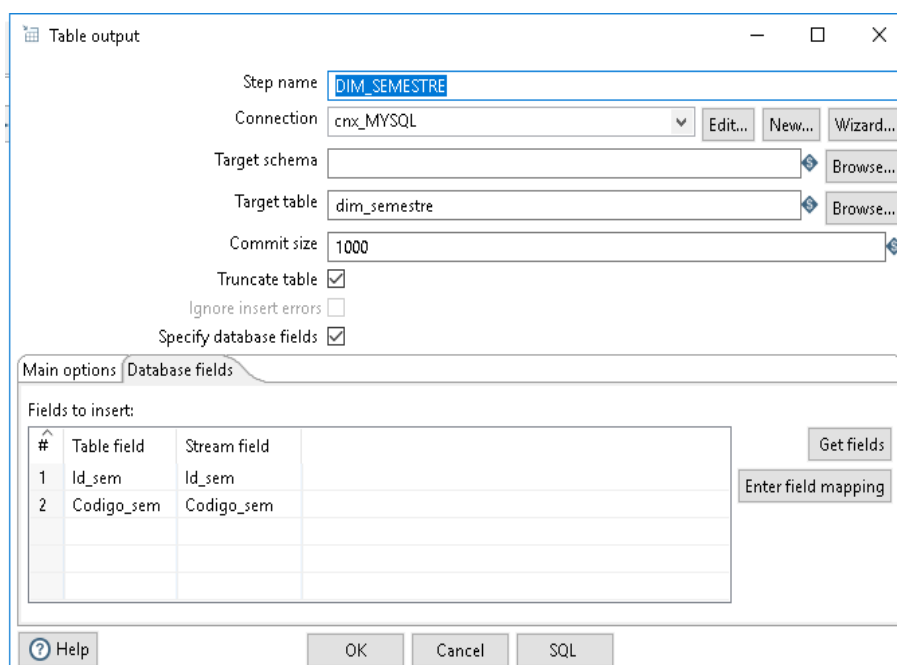
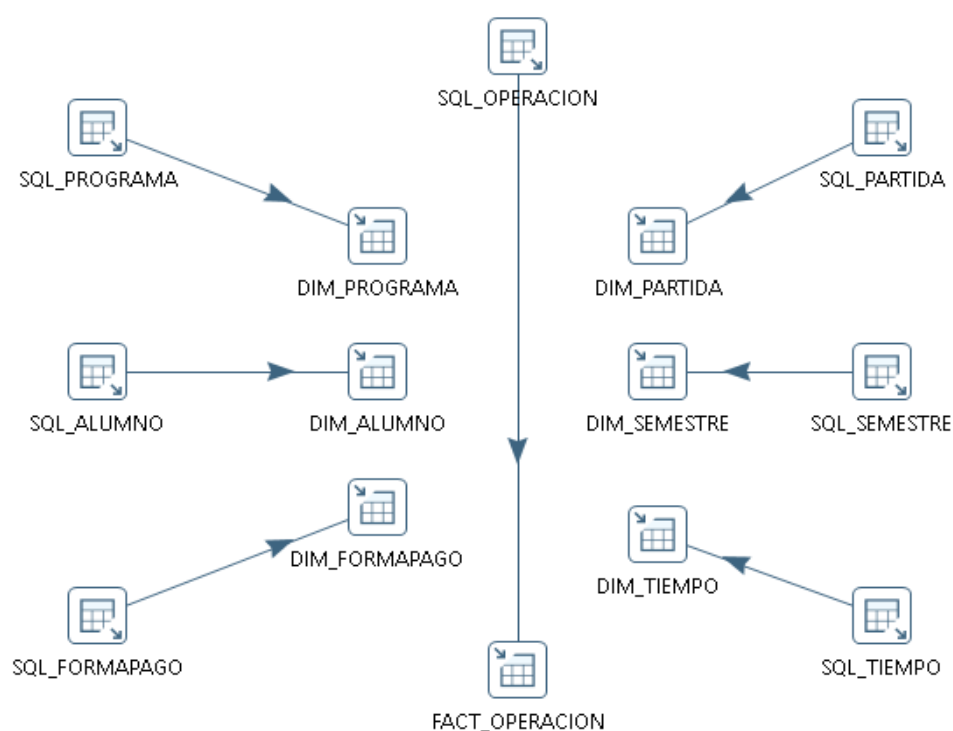


Figura NRO. 62: Tabla de destino lista y preparada para recibir la data deseada



*Figura NRO. 63: Diseño del Data Mart para la migración de Base de Datos operativa de la información deseada*

### 4.3. Definición de los procedimientos de mantenimiento y crecimiento del Data Mart

En esta etapa del proyecto se ha tomado en consideración la realización de las pruebas funcionales del proceso de migración y automatización, así como de usabilidad de los datos y reportes estadísticos del software Pentaho.

#### 4.3.1. Despliegue

Para verificar la consistencia del proceso de migración desarrollado se procedió a realizar un conjunto de pruebas específicas que se detallaran a continuación.

#### Plan de Pruebas

Son las investigaciones empíricas y técnicas cuyo objetivo es proporcionar información objetiva e independiente sobre la calidad del producto a la parte interesada o stakeholder. Es una actividad más en el proceso de control de calidad.



Las pruebas son básicamente un conjunto de actividades dentro del desarrollo de software. Dependiendo del tipo de pruebas, estas actividades podrán ser implementadas en cualquier momento de dicho proceso de desarrollo. Existen distintos modelos de pruebas. A cada uno corresponde un nivel distinto de involucramiento en las actividades de desarrollo.

### **Tipos de Pruebas**

Las pruebas que se han considerado son las siguientes:

- Pruebas de Funcionalidad
  - Pruebas de Unidad
  - Pruebas de Seguridad
  - Pruebas de Sistema
- Pruebas de Rendimiento

Para la implementación de este proyecto se ha considerado las herramientas de Pentaho, por su administración en el volumen de información, ya que conocemos de anticipación las características mínimas que deberían tener los equipos o servidores que alojen la suite de Pentaho BI.

- Requerimientos de hardware

*Tabla 4: Requerimientos de Hardware*

<b>Hardware</b>	<b>Descripción</b>
Memoria RAM	Como mínimo 4 GB
Espacio de almacenamiento en Disco Duro	Mínimo 500GB
Procesador	Intel Pentium Dual Core / AMD64 or EM64 T

*Fuente: [Elaboración Propia]*

- Requerimientos de software

*Tabla 5: Requerimientos de Software*

Software	Descripción
JRE	Java Runtime Environment Version 1.6 o superior / jdk-8u161-windows-x64
Sistemas Operativos y distribuciones soportadas	Windows 7 / Windows 8/10 86x y 64x CentOS 6 SUSE Linux Enterprise Red Hat Enterprise Linux 5 Ubuntu Server versión moderna de Linux

*Fuente: [Elaboración Propia]*

### **Análisis de Recursos:**

La máquina elegida donde funcionara la suite de Pentaho BI deberá cumplir con las características mínimas recomendadas.

### **Cantidad de Usuarios:**

La cantidad de usuarios que se conectaran de forma concurrente a la suite de Pentaho BI dependerá de los requisitos que posea el equipo elegido y la cantidad de operaciones que realice.

### **Cantidad de Información:**

La cantidad de información que se almacene es muy importante para verificar las estadísticas mensuales y anuales, ya que los datos almacenados son muy ligeros el acceso será rápido y preciso.

### **Ejecución de Pruebas**

- **Pruebas Funcionales**

Las pruebas funcionales se realizaran tomando en cuenta un conjunto de datos reales y después de realizar la población a las tablas de destino el sistema revisara si los datos se ven reflejados en el Data Mart.

## Pruebas de Unidad

Para las pruebas de unidad se realizan casos de pruebas que nos permitirá determinar el funcionamiento individual de los componentes desarrollados y si estos cumplen con los requerimientos del sistema, Para los casos de pruebas se utilizará la siguiente plantilla.

*Tabla 6: Plantilla del Caso de Prueba*

<b>Identificador de caso de prueba:</b> <Numero>	<b>Nombre del Sistema:</b> <Nombre>			
<b>Nivel de Prueba:</b> <Nivel>	<b>Tipo de Prueba:</b> <Tipo>			
<b>Nombre Flujo:</b> <Flujo>	<b>Autor del Caso de Prueba:</b> <Autor>			
<b>Nombre Caso de Prueba:</b> <Nombre>	<b>Nombre del Probador:</b> <Probador>			
<b>Versión del Caso de Prueba:</b> <Versión>	<b>Fecha de Creación:</b> <Fecha>		<b>Fecha de Ejecución:</b> <Fecha>	
Condición(es) para que se ejecute el Caso de Prueba:				
<Precondición 1>				
<Precondición 2>				
Para la Ejecución del Caso de Prueba:				
<b>Nro. Paso</b>	<b>Condición</b>	<b>Valor(es)</b>	<b>Resultado Esperado</b>	<b>Resultado Obtenido</b>
Flujo				
<#1>	<Condición>	<Valores Ingresados>	<Resultado Esperado>	<Resultado btenido>
<#2>	<Condición>	<Valores Ingresados>	<Resultado Esperado>	<Resultado Obtenido>
<b>Criterios de Aprobación del Caso de Prueba:</b> <Criterios de aprobación>				
<b>Decisión de Aprobación del Caso de Prueba:</b> Aprobó:_Falló:_ <Aprobó o Falló>				

**Tabla 7: Caso de uso de prueba manipulación de la información**

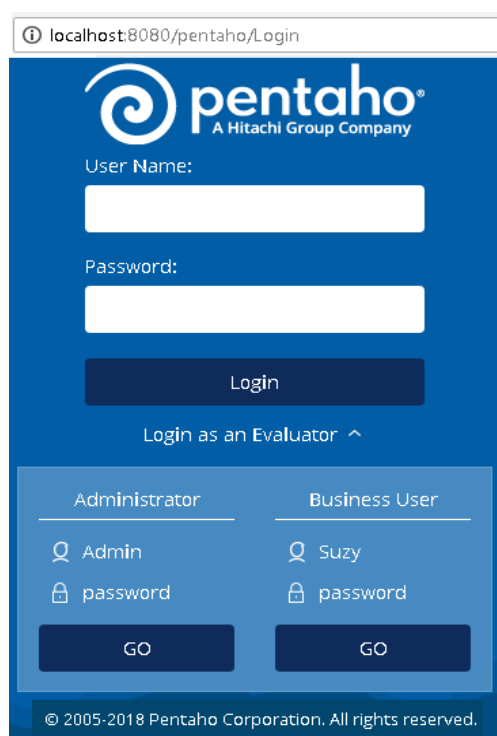
Identificador de caso de prueba: <b>CUP-001</b>		<b>Nombre del Sistema:</b> Data Mart Pentaho BI		
<b>Nivel de Prueba:</b> Normal		<b>Tipo de Prueba:</b> Normal		
<b>Nombre Flujo:</b> Principal		<b>Autor del Caso de Prueba:</b> Luis Alberto Mija Camargo		
<b>Nombre Caso de Prueba:</b> Manipulación de la Información		<b>Nombre del Probador:</b> Luis Alberto Mija Camargo		
<b>Versión del Caso de Prueba:</b> 1.0		<b>Fecha de Creación:</b> 04/07/2018		<b>Fecha de Ejecución:</b> 04/07/2018
<b>Condición(es) para que se ejecute el Caso de Prueba:</b>				
El usuario debe estar registrado y activo en el pentaho-server 7.0.1				
<b>Para la Ejecución del Caso de Prueba:</b>				
<b>Nro. Paso</b>	<b>Condición</b>	<b>Valor(es)</b>	<b>Resultado</b>	<b>Resultado</b>
<b>Flujo</b>			<b>Esperado</b>	<b>Obtenido</b>
1	Haber ingresado al sistema y seleccionar el cubo perfil_financiero a presentar	-	Pentaho-Server deberá permitir filtrar dimensiones y medidas	Pentaho-Server permite el filtrado de las dimensiones y medidas
<b>Criterios de Aprobación del Caso de Prueba:</b> El usuario puede filtrar dimensiones y medidas para analizar la información obtenida.				
<b>Decisión de Aprobación del Caso de Prueba:</b> Aprobó: X    Falló: _				

## Análisis de Resultados

De los casos de prueba realizados se puede inferir que el Data Mart fue correctamente implementado y cumplen con los requerimientos funcionales del mismo.

## Pruebas de Seguridad

Mediante las pruebas de seguridad se nos permitirá proteger la información de usuarios no autorizados o de protocolos no seguros a través de la red. Es por ello que la herramienta suite de Pentaho BI nos permite darle seguridad mediante la aplicación de roles que son asignados a los diferentes usuarios, esta administración se puede realizar gracias a que Pentaho tiene una consola de administración, además maneja conexiones JNDI, que utilizan las seguridades que provee Java, además se tiene archivos de configuración para acceso a las bases de datos Hibernate y Quartz que usa Pentaho como se muestra a continuación.



*Figura NRO. 64: Aquí podemos ver la consola de administración y la url de la herramienta Pentaho Server desde donde se ingresaran los datos necesarios para iniciar sesión <http://localhost:8080/Login>*

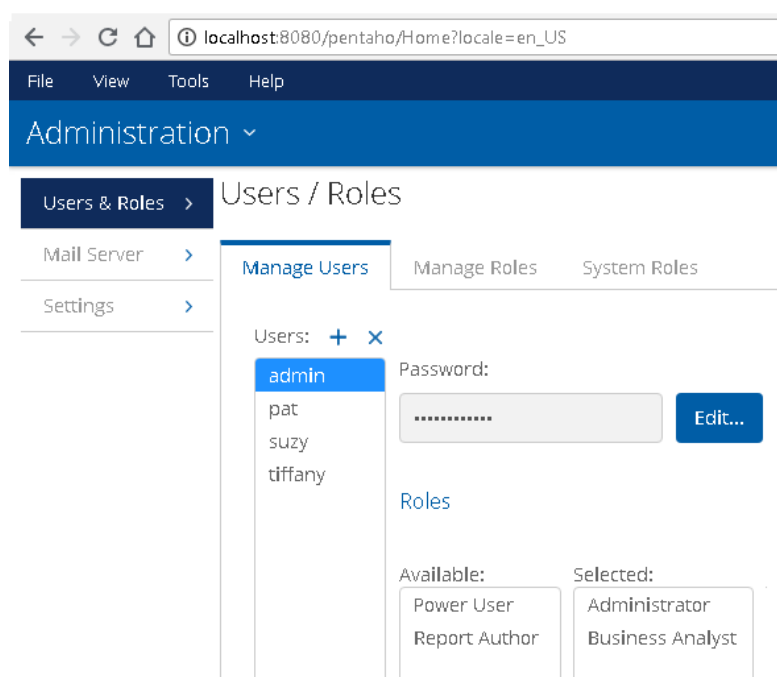


Figura NRO. 65: Lista de usuarios que administraran el aplicativo Web

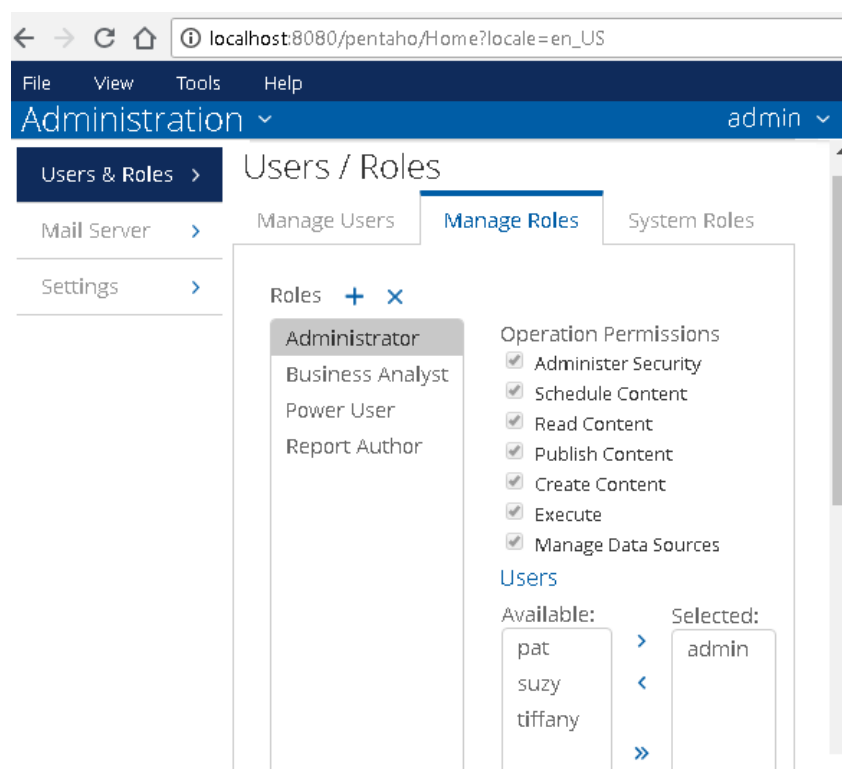


Figura NRO. 66: Roles que serán asignados según los privilegios a cada usuario

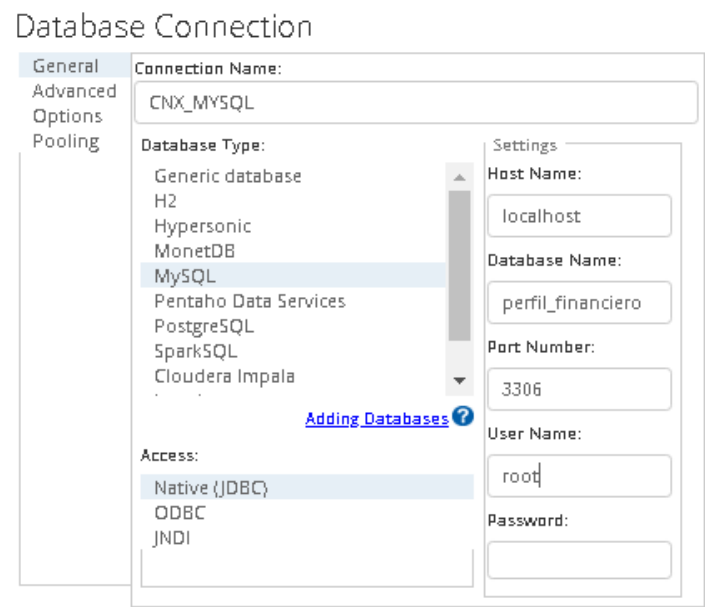


Figura NRO. 67: Conexión a la Base de datos del Data Mart desde donde se crear el cubo el los reportes

En la siguiente tabla detallaremos las pruebas de seguridad realizadas para comprobar que el Sistema Web de Pentaho cuente con mecanismos de seguridad que evita que los usuarios no autorizados puedan acceder al sistema y que los datos enviados a través de una red utilicen un protocolo de red seguro.

**Tabla 8: Pruebas de seguridad**

Propósito: Comprobar que el Sistema Web de Pentaho cuente con mecanismos de seguridad para evitar el acceso a usuarios no autorizados y que los datos enviados utilicen un protocolo de red seguro.	
Caso de Prueba	Resultado
<p>1. Se intenta acceder al Sistema Web a través del protocolo inseguro HTTP.</p> <p>2. Se intenta acceder a una página restringida del Sistema Web sin previa autenticación del usuario.</p>	<p>1. El Sistema Web usa protocolo HTTP debido a que no se manejan datos críticos en donde no es necesario la implementación de protocolos seguros como https</p> <p>2. El sistema redirecciona al usuario a la página de autenticación para que pueda acceder al Sistema Web.</p>
Conclusión: El Sistema Web de Pentaho no controla en forma total el acceso a la información ya que no se vigila que se utilice el protocolo seguro HTTPS, pero se considera que es seguro ya que la información que se maneja no es crítica al enviar datos a través de la red y controla el acceso de usuarios utilizando autenticación.	

### Pruebas de Sistema

Las pruebas de Sistemas nos permitirán verificar el funcionamiento global del sistema, en donde se han definido todos los casos de pruebas de las diferentes dimensiones para el Data Mart para ello se corrió el proceso ETL.

A continuación se muestra información del modelo estrella de eventos y los resultados.

Eventos:

Una vez corrido los procesos ETL, se verificaron los datos en todas las dimensiones:

- Dimensión Tiempo
- Dimensión Programa
- Dimensión Alumno
- Dimensión Semestre
- Dimensión FormaPago
- Dimensión Partida
- Fact\_Operacion
- Fact\_Cestadistico



Servidor: 127.0.0.1 > Base de datos: perfil\_financiero >

Examinar Estructura SQL Buscar

Mostrando filas 0 - 24 (total de 456, La consulta tardó 0.000)

SELECT \* FROM `dim\_tiempo` ORDER BY `anio\_tm` ASC

1 Mostrar todo > >> Número de filas: 2

+ Opciones

id_fecha_tm	anio_tm	mes_tm	día_tm
2017-06-03	2017	6	3
2017-08-24	2017	8	24
2017-05-11	2017	5	11
2017-03-23	2017	3	23
2017-11-07	2017	11	7
2017-07-21	2017	7	21
2017-08-10	2017	8	10
2017-07-15	2017	7	15
2017-11-09	2017	11	9
2017-07-03	2017	7	3
2017-05-20	2017	5	20

Figura NRO. 68: Consulta de datos en la dimensión tiempo

Mostrando filas 0 - 217 (total de 218, La consulta tardó 0.0010 segundos.) [Codigo: -]

SELECT \* FROM `dim\_programa` ORDER BY `dim\_programa`.`codigo` DESC

+ Opciones

	id_pg	Sede_pg	Programa_pg	Tipo_pg
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	223	CAJAMARCA - SAN MARCOS	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	196	UCHIZA	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	194	TRUJILLO	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	193	TARAPOTO	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	192	JULIACA	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	191	CHACHAPOYAS	MAESTRIA EN CIENCIAS DE LA EDUCACION CON MENCION E...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	187	LAMBAYEQUE	MAESTRIA EN CIENCIAS CON MENCION EN CONTROL Y AUDI...	MAESTRIA
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	186	LAMBAYEQUE	DOCTORADO EN MEDICINA	DOCTORADO
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	185	LAMBAYEQUE	MAESTRIA EN DERECHO DEL TRABAJO Y DE LA SEGURIDAD	MAESTRIA

Figura NRO. 69: Consulta de datos de la dimensión programa

Mostrando filas 0 - 24 (total de 252, La consulta tardó 0.0003 segundos.)

SELECT \* FROM `dim\_alumno`

1 > >> Número de filas: 25 Filtrar filas: Buscar en esta tabla

+ Opciones

	id_al	Promocion_al	Grupo_al	Sexo_al	id_pg_fk	Codigo_al	ApellidosyNombres_al	FechaNacimiento_al	Ingresa_al	Estadoal
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	21835	IV	U	M	173	416611B	MUJA CAMARGO, LUIS ALBERTO	1980-03-30 00:00:00	2015-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24886	I	U	F	187	420048G	ACUÑA VASQUEZ, GLORIA ARAMINDA	1982-04-04 00:00:00	2012-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24887	I	U	M	187	420049C	ALDANA MONJA, TULIO CESAR	1975-11-28 00:00:00	2012-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24888	I	U	F	187	420050A	BERNAL LLONTOP, KARINA	1980-03-20 00:00:00	2012-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24889	I	U	F	187	420051H	BOCANEGRA PINEDO, KELLY JANETH	1985-04-25 00:00:00	2012-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24890	I	U	M	187	420052D	BRACO MEJIA, ROGERS EDMUNDO	1956-08-18 00:00:00	2012-I	-
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	24891	I	U	F	187	420053K	BRAVO DAVILA, YOVANE EDITH	1983-09-27 00:00:00	2012-I	-

Figura NRO. 70: Consulta de datos de la dimensión alumno

Servidor: 127.0.0.1 > Base de datos: perfil\_financiero > Tabla: dim\_semestre

Examinar Estructura SQL Buscar Insertar Exportar In

⚠ La selección actual no contiene una columna única. La edición de la grilla y los enlaces de copiado, eli

✓ Mostrando filas 0 - 24 (total de 51, La consulta tardó 0.0004 segundos.) [codigo\_sem: 1993-I - 2005-II]

`SELECT * FROM `dim_semestre` ORDER BY `codigo_sem` ASC`

1 ▼ Mostrar todo > >> Número de filas: 25 ▼ Filtrar filas: Buscar en esta tab

+ Opciones

id_sem	codigo_sem
19931	1993-I
19932	1993-II
19941	1994-I
19951	1995-I
19952	1995-II
19961	1996-I
19962	1996-II
19972	1997-II
19971	1997-I
19981	1998-I
19982	1998-II

Figura NRO. 71: Consulta de datos de la dimensión semestre

Servidor: 127.0.0.1 > Base de datos: perfil\_financiero > Tabla: dim\_formapago

Examinar Estructura SQL Buscar Insertar Ex

✓ Mostrando filas 0 - 4 (total de 5, La consulta tardó 0.0004 segundos.)

`SELECT * FROM `dim_formapago``

Número de filas: 25 ▼ Filtrar filas: Buscar en esta tabla

Ordenar según la clave: Ninguna ▼

+ Opciones

	id_fp	Descripcion_fp
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	2	VISA RECURRENTE
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	3	PAGO EN EFECTIVO
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	4	PAGO EN BANCO DE LA NACION
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	5	VISA POS
<input type="checkbox"/> Editar <input type="checkbox"/> Copiar <input type="checkbox"/> Borrar	6	PAGO CON CHEQUE

Figura NRO. 72: Consulta de datos de la dimensión forma de pago

✓ Mostrando filas 0 - 24 (total de 32, La consulta tardó 0.0005 segundos.) [id\_pa: 8 - 44]

`SELECT * FROM `dim_partida` ORDER BY `id_pa` ASC`

+ Opciones

		id_pa	Partida_pa	Subpartida_pa	Concepto_pa
<input type="checkbox"/>	Editar Copiar Borrar	8	1.3.2.3.1.4	1201.0302.3.1.4.1	CONSTANCIA
<input type="checkbox"/>	Editar Copiar Borrar	9	1.3.2.3.1.4	1201.0302.3.1.4.2	CERTIFICADO
<input type="checkbox"/>	Editar Copiar Borrar	10	1.3.2.3.1.99	1201.0302.3.1.99.1	SUSTENT. DE TESIS
<input type="checkbox"/>	Editar Copiar Borrar	11	1.3.3.3.1.5	1201.03.03.1.1.5.1.1	DERECHO INSCRIPC.
<input type="checkbox"/>	Editar Copiar Borrar	12	1.3.3.3.1.5	1201.03.03.1.1.5.1.3	PENS. ENSEÑANZA
<input type="checkbox"/>	Editar Copiar Borrar	13	1.3.3.3.1.5	1201.03.03.1.1.5.1.2	MATRICULA
<input type="checkbox"/>	Editar Copiar Borrar	14	-	-	EXAM. ESPECIAL
<input type="checkbox"/>	Editar Copiar Borrar	15	1.3.1.5.1.99	1201.0301.5.1.99.1	PROSPECTO
<input type="checkbox"/>	Editar Copiar Borrar	17	1.3.2.3.1.99	1201.0302.3.1.99.6	PROYECTO TESIS
<input type="checkbox"/>	Editar Copiar Borrar	18	1.3.3.3.1.5	1201.0303.1.1.5.3	CURSO

Figura NRO. 73: Consulta de datos de la dimensión partida

⚠ La selección actual no contiene una columna única. La edición de la grilla y los enlaces de copiado, elir

✓ Mostrando filas 0 - 24 (total de 52629, La consulta tardó 0.0374 segundos.) [id\_fp\_fk: 6 - 6]

`SELECT * FROM `fact_operacion` ORDER BY `fact_operacion`.`id_fp_fk` DESC`

+ Opciones

id_tr	id_pa_fk	id_fecha_fk	id_fp_fk	id_al_fk	subtotal_tr
173376	12	2016-09-09	6	33830	299.700
173377	12	2016-09-09	6	36782	333.000
173378	12	2016-09-09	6	37054	299.700
173379	12	2016-09-09	6	37056	333.000
173380	12	2016-09-09	6	33231	299.700
173381	12	2016-09-09	6	33583	333.000
173384	12	2016-09-09	6	36739	299.700
173385	12	2016-09-09	6	36795	299.700
173386	12	2016-09-09	6	36746	299.700
173387	12	2016-09-09	6	36944	299.700
173388	12	2016-09-09	6	33472	299.700

Figura NRO. 74: Consulta de datos de la tabla de hechos operación

⚠ La selección actual no contiene una columna única. La edición de la grilla y los enlaces de copiado, eliminación y edición no están disponibles.

✓ Mostrando filas 0 - 12 (total de 13, La consulta tardó 0.0004 segundos.)

```
SELECT * FROM `fact_cestadistico`
```

+ Opciones

id_ce	naingresantes_ce	nahabiles_ce	nacondicion_ce	id_sem_fk	id_pg_fk	grupo_ce	fecha_ce
1	37	36	13	20171	44	U	2018-06-27
2	68	67	24	20171	13	U	2018-06-27
3	50	48	6	20171	25	U	2018-06-27
4	22	18	14	20171	187	U	2018-06-27
5	24	14	8	20171	16	U	2018-06-27
6	33	31	5	20171	269	U	2018-06-27
7	35	33	16	20171	176	U	2018-06-27
8	48	47	14	20171	12	U	2018-06-27
9	142	139	42	20171	21	U	2018-06-27
10	51	49	17	20171	79	U	2018-06-27
11	31	31	12	20171	185	U	2018-06-27

Figura NRO. 75: Consulta de datos de la tabla de hechos cuadro estadístico

## Despliegue de la información a través del JPivot

### Database Connection

General  
Advanced  
Options  
Pooling

Connection Name: cnx21

Database Type: Generic database, H2, Hypersonic, MonetDB, **MySQL**, Pentaho Data Services, PostgreSQL, SparkSQL, Cloudera Impala, Adding Databases?

Access: Native (JDBC), ODBC, JNDI

Settings

Host Name: localhost

Database Name: perfil\_financiero

Port Number: 3306

User Name: root

Password:

Test OK Cancel

Figura NRO. 76: Establecer parámetros de conexión

### Data Source Wizard

Select Source Type

Data Source Name:

Select Tables

Define Joins

Source Type: Database Table(s)

Select a database connection and click Next to choose from a list of the available database tables.

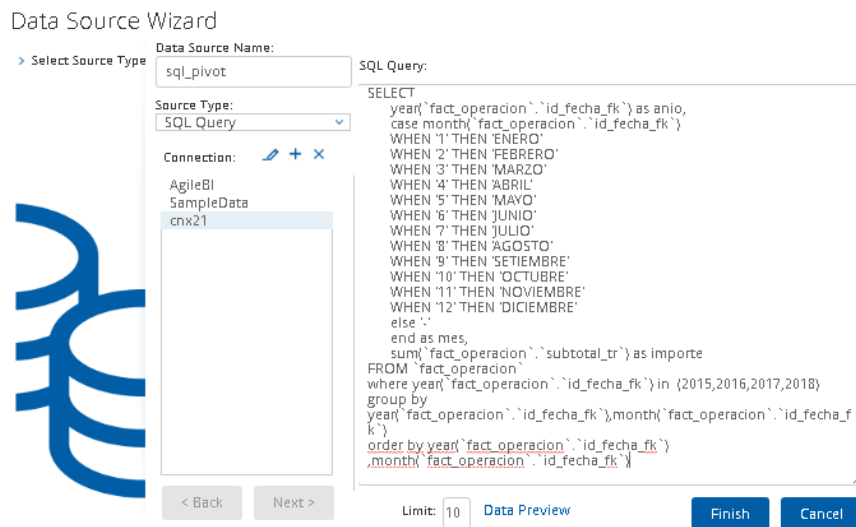
Connection: AgileBI, SampleData

Create data source for:

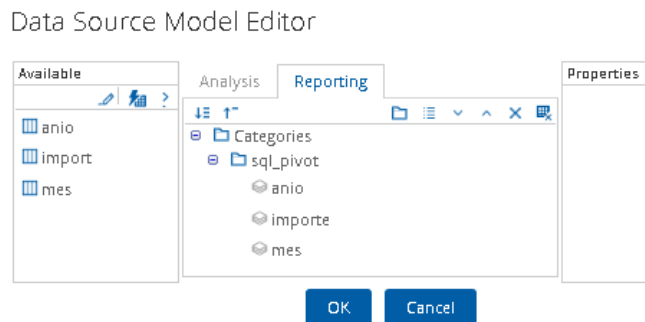
Reporting only  
Reporting and Analysis (Requires Star Schema)

< Back Next > Finish Cancel

Figura NRO. 77: Establecer el origen de datos de la conexión



**Figura NRO. 78:** Establecer la consulta SQL para obtener el reporte deseado (listado de importes por mes de los últimos 4 años)



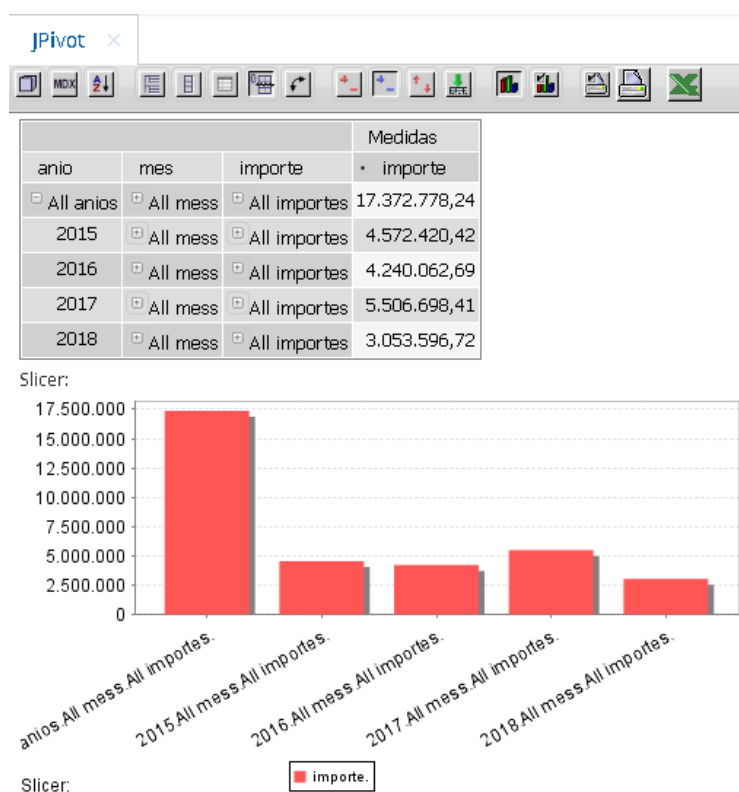
**Figura NRO. 79:** Manipulación de los campos para su respectiva visualización en el JPivot

#### New JPivot View

Schema

Cube

**Figura NRO. 80:** Visualizar nuestro reporte



*Figura NRO. 81: Visualizar los datos elegidos en la creación del reporte (listar de importes por meses de los últimos 4 años)*

### 4.3.2. Mantenimiento

El mantenimiento del Data Mart se debe realizar de forma semanal y mensual para verificar y revisar la información útil para la tomar decisiones.

Debido a que los Data Mart por su utilidad e importancia llegan a ser uno de los sistemas más usados dentro de la organización, el mantenimiento dependerá de factores como:

- Crecimiento de la información
- Calidad de los datos que son ingresados al Data Mart

El apoyo permanente la capacitación y comunicación con los usuarios finales es de vital importancia para la ejecución de sus operaciones y toma de decisiones día a día.

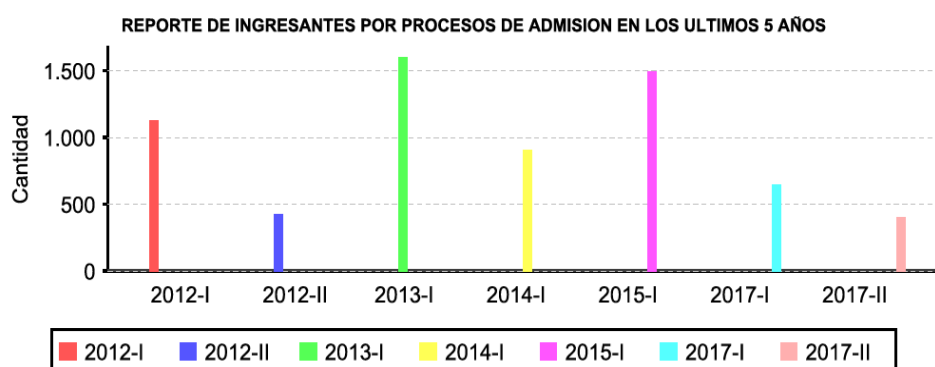
#### 4.4. Interpretación de los resultados obtenidos del Data Mart para encontrar el perfil financiero adecuado

Con el uso de la plataforma Pentaho Open BI Suite podemos decir que una de las etapas más importantes es la recolección de los datos a través del sistema de inscripción, si la información no fuera ingresada correctamente los datos procesados no cumplirían el objetivo deseado, y las decisiones que se ejecuten nos llevaran a un desastre.

Los resultados obtenidos en este proyecto se van a representar mediante gráficos para su interpretación de manera sencilla.

*Tabla 9: Cantidad de postulantes por proceso de admisión*

Nro	S. Ingreso	Cant.
1	2012-I	1127
2	2012-II	422
3	2013-I	1596
4	2014-I	907
5	2015-I	1496
6	2017-I	642
7	2017-II	402

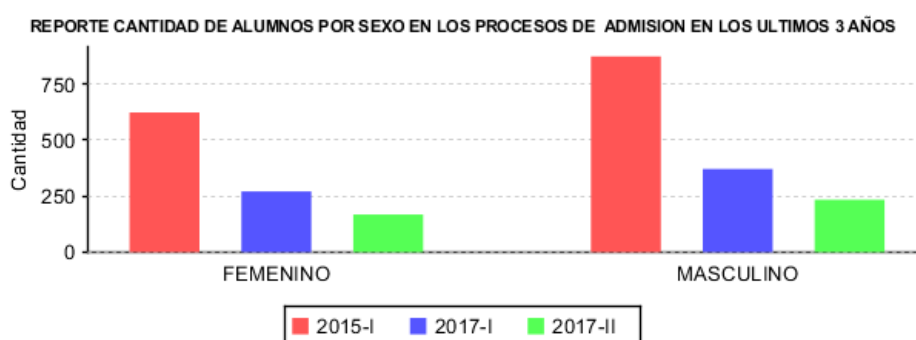


*Figura NRO. 82: Postulantes por procesos de admisión*

Podemos observar que en el proceso de admisión de la EPG del 2013-I, la mayor demanda de inscripción se realizó en el mes de marzo con 1596 postulantes, asimismo podemos observar que en el proceso de admisión 2015-I la mayor demanda de inscripción de postulantes se realizó en el mes de marzo con 1496 postulantes.

**Tabla 10: Resumen de proporciones por sexo de los procesos admisión 2015-I, 2017-I y 2017-II**

Nro	S. Ingreso	Sexo	Cant.	
1	2015-I	Femenino	623	41.64%
2	2015-I	Masculino	873	58.36%
<b>Total</b>			1496	
1	2017-I	Femenino	271	41.21%
2	2017-I	Masculino	371	57.79%
<b>Total</b>			642	
1	2017-II	Femenino	168	41.49%
2	2017-II	Masculino	234	58.21%
<b>Total</b>			402	

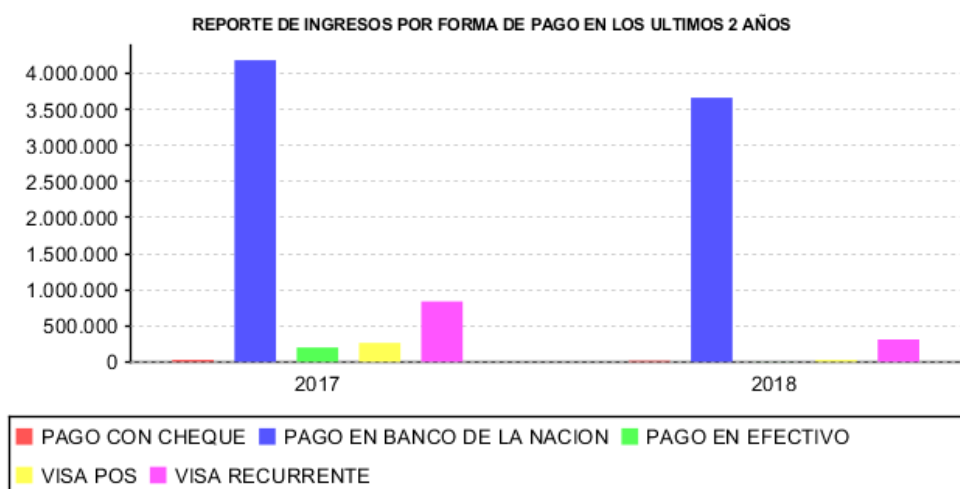


**Figura NRO. 83: Cantidad de alumnos por sexo de los proceso de admisión 2015-I, 2017-I y 2017-II**

**Tabla 11: Importes por medio de pago del año 2017 y 2018**

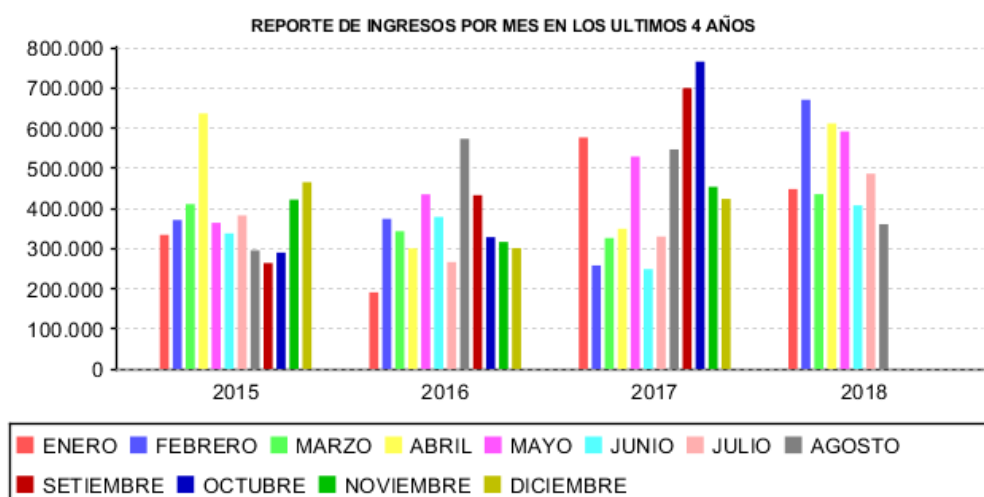
Nro	S. Ingreso	Forma de pago	Importe
1	2017	Pago con Cheque	S/. 26.564,09
2	2017	Pago en Banco de la Nación	S/. 4.181.004,02
3	2017	Pago en Efectivo	S/. 198.921,80
4	2017	Visa Pos	S/. 262.641,00
5	2017	Visa Recurrente	S/. 837.567,50
6	2018	Pago con Cheque	S/. 15.061,80
7	2018	Pago en Banco de la Nación	S/. 3.660.383,17
8	2018	Pago en Efectivo	S/. 1.100,00
9	2018	Visa Pos	S/. 23.330,00
10	2018	Visa Recurrente	S/. 310.345,00





*Figura NRO. 84: Reporte de ingresos por medio de pago del año 2017 y 2018*

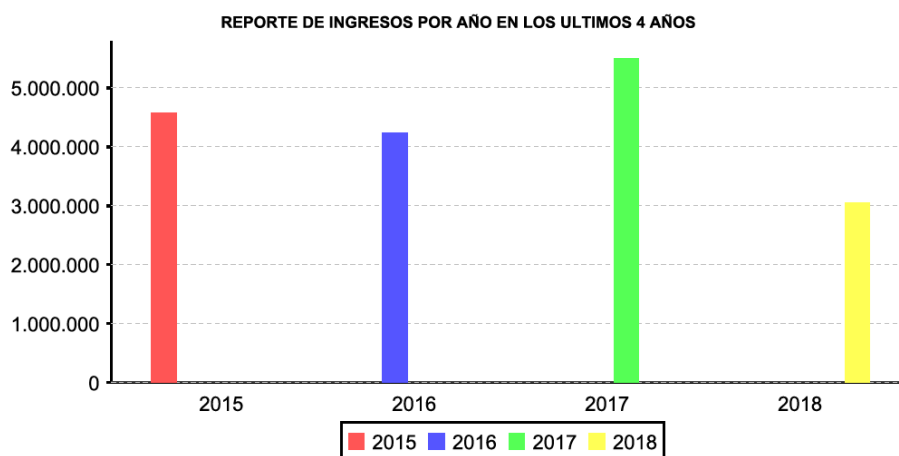
Podemos observar que en la figura nro. 84 los ingresos por banco de la nación superan en gran proporción a los medios de pago como Visa POS, Visa Recurrente y pago en efectivo, se podría descongestionar y recaudar mucho más por cobros en efectivo y tarjeta.



*Figura NRO. 85: Reporte de ingresos por mes de los últimos 4 años*

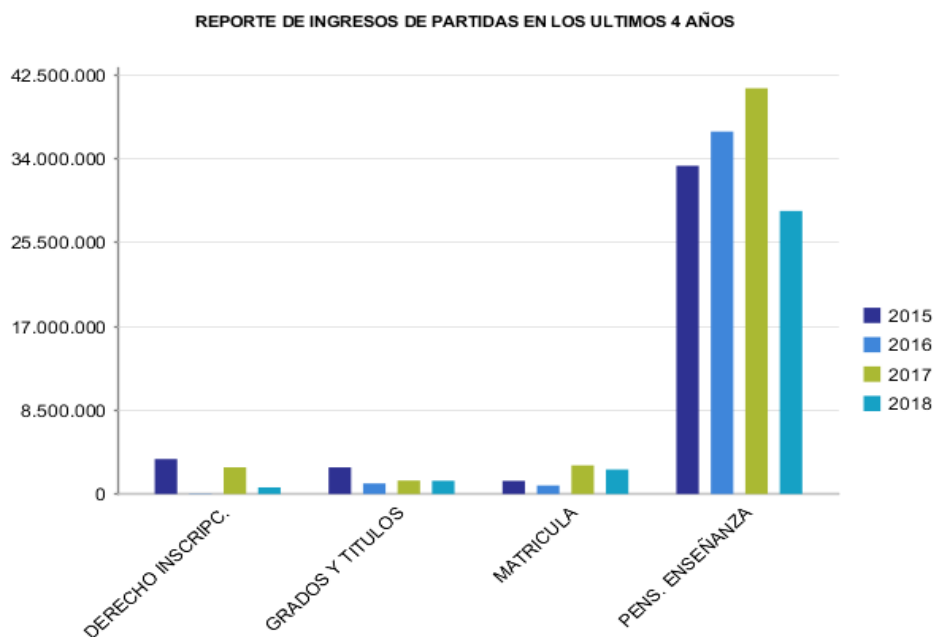
**Tabla 12:** Lista donde se muestran los ingresos por año

Nro	S. Ingreso	Importe
1	2015	S/. 4.572.420,42
2	2016	S/. 4.240.062,69
3	2017	S/. 5.506.698,41
4	2018	S/. 4.012.219,97



**Figura NRO. 86:** Reporte de ingresos de los últimos 4 años

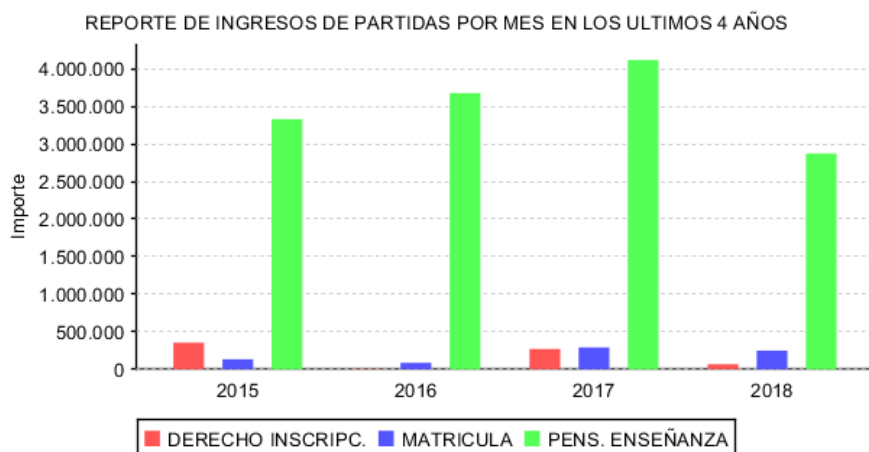
Podemos observar que en la figura nro. 86 los ingresos aumentaron cada año y que a medio año del 2018 ya se ha superado considerablemente al año 2017, eso significa que se espera una proyección mayor en lo recaudado este año.



**Figura NRO. 87:** Reporte de ingresos por partidas de los últimos 4 años

**Tabla 13:** Aquí se muestran los importes por partidas en los últimos 4 años

Nro	Año	Partida	Importe
1	2015	Derecho Inscripc.	S/. 354.870,00
2	2015	Matricula	S/. 133.580,00
3	2015	Pens. Enseñanza	S/. 3.330.109,42
4	2016	Derecho Inscripc.	S/. 4.330,00
5	2016	Matricula	S/. 86.462,50
6	2016	Pens. Enseñanza	S/. 3.677.208,29
7	2017	Derecho Inscripc.	S/. 270.880,00
8	2017	Matricula	S/. 290.712,50
9	2017	Pens. Enseñanza	S/. 4.116.437,81
10	2018	Derecho Inscripc.	S/. 67.250,00
11	2018	Matricula	S/. 248.655,00
12	2018	Pens. Enseñanza	S/. 2.872.048,17



**Figura NRO. 88:** Reporte de ingresos por partidas en los últimos 4 años

Podemos observar que en la figura nro. 88, Los ingresos por pensiones de enseñanza son mucho mayores que otros años, los ingresos aumentaron por algunas mejoras en las políticas de gobierno, políticas educativas, mejoras en los procesos académicos de la UNPRG-EPG, todos estos factores han permitido un mejor crecimiento en lo recaudado este año.

Tabla 14: Listado de importes por programas de admision 2017-I

ADMISION 2017-I				
Nro	Grupo	Programa	Importe	
1	U	MA_AUDITORIA_GUBERNAMENTAL	S/	101,435.00
2	U	MA_CONS_GOVERNABILIDAD	S/	194,200.20
3	U	MA_DER_CIVIL_COMERCIAL	S/	149,488.10
4	A	MA_DER_PENALES	S/	188,575.00
5	B	MA_DER_PENALES	S/	172,015.00
6	C	MA_DER_PENALES	S/	161,415.00
7	U	MA_DER_SEGURIDAD_SOCIAL	S/	113,856.94
8	A	MA_GER_EMPRESARIAL	S/	157,670.00
9	B	MA_GER_EMPRESARIAL	S/	90,591.00
10	U	MA_GER_PUBLICA	S/	117,845.00
11	U	MA_GESTION_DEL_SOFTWARE	S/	111,130.00
12	U	MA_INGENIERIA_AMBIENTAL	S/	73,840.00
13	A	MA_OBRAS_CONSTRUCCION	S/	153,960.00
14	B	MA_OBRAS_CONSTRUCCION	S/	135,590.00
15	U	MA_PROYECTOS_INVERSION	S/	151,215.00
16	U	MA_TERRITORIAL	S/	102,130.00
17	U	MA_TRIBUTACION_ASESORIA	S/	135,350.00

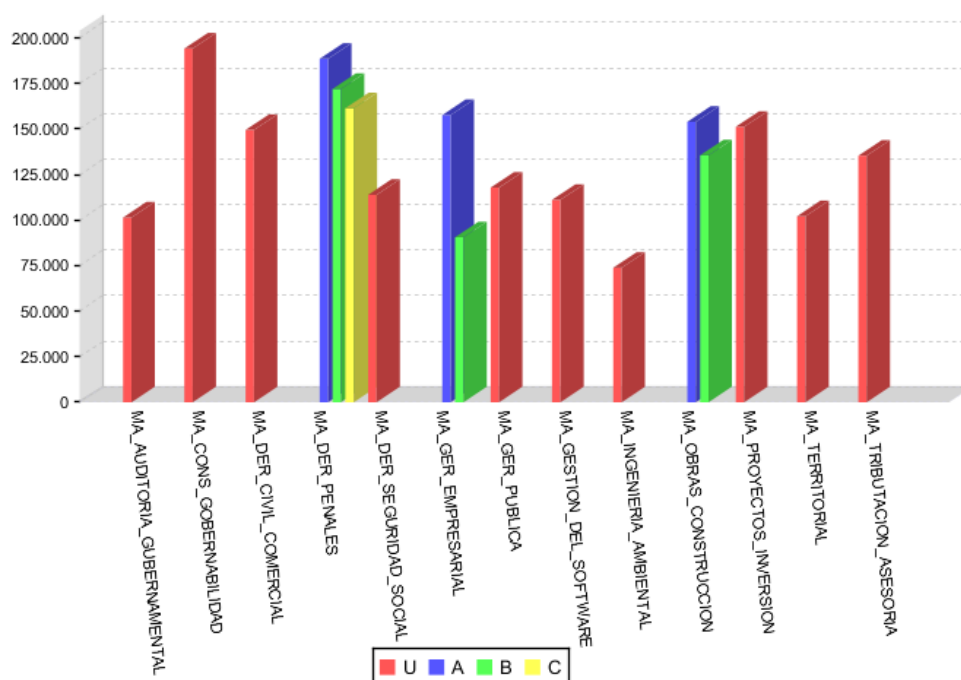


Figura NRO. 89: Reporte de ingresos por programas del admisión 2017-I

Podemos observar que en la figura nro. 89, Los programas con mayores ingresos se encuentran en los programas de Derecho y ciencias administrativas como por ejemplo Proyectos de Inversión, Gerencia Publica ingresos, ingresos que los pilares fundamentales para que los procesos académicos y administrativos sigan funcionando y permitan crecer y mejorar cada año.

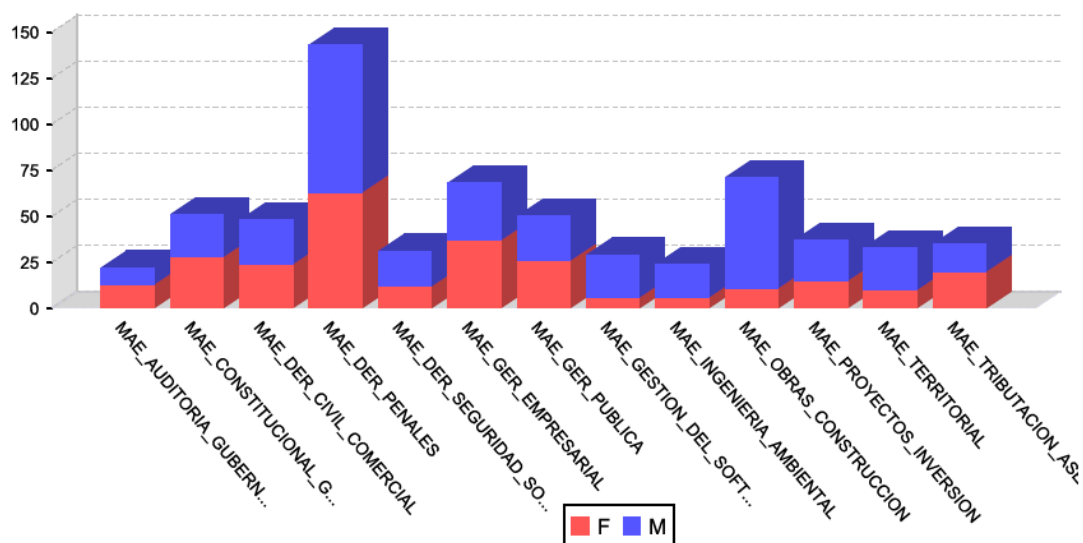


Figura NRO. 90: Cantidad de alumnos por sexo de los programas de admisión 2017-I

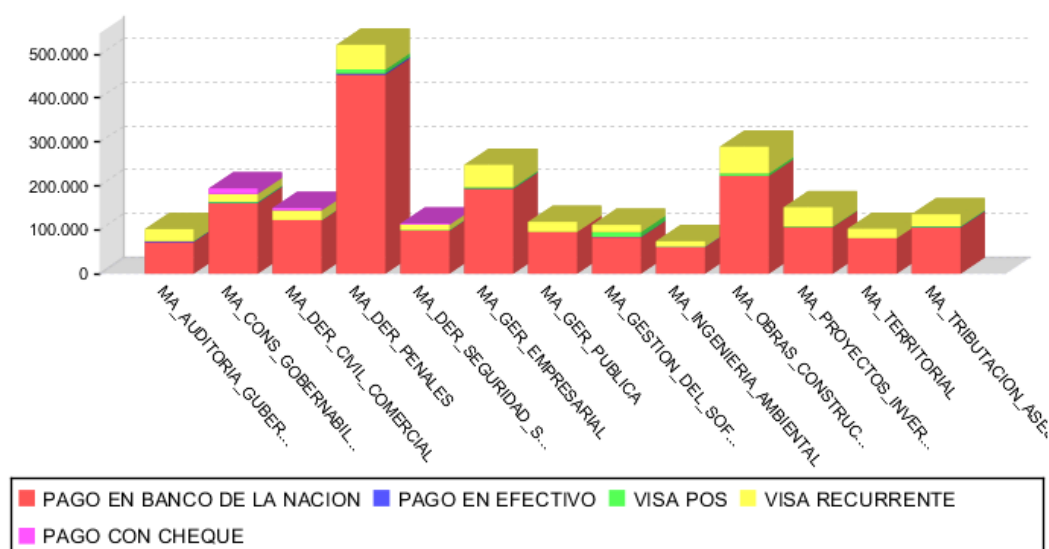


Figura NRO. 91: Reporte de ingresos por forma de pago de los programas de admisión 2017-I

**Tabla 15: Aquí se muestran los ingresos por forma de pago de los programas de admisión 2017-II**

<b>ADMISION 2017-II</b>			
<b>Nro</b>	<b>Programa</b>	<b>Concepto</b>	<b>Importe</b>
1	DOC_CIENCIA_POLITICA	PAGO EN BANCO DE LA NACION	S/ 96,170.00
2	DOC_CIENCIA_POLITICA	VISA RECURRENTE	S/ 3,300.00
3	MA_GER_EMPRESARIAL	PAGO EN BANCO DE LA NACION	S/ 119,720.00
4	MA_GER_EMPRESARIAL	PAGO EN EFECTIVO	S/ 650.00
5	MA_GER_EMPRESARIAL	VISA RECURRENTE	S/ 11,250.00
6	MA_GER_PUBLICA	PAGO EN BANCO DE LA NACION	S/ 46,710.00
7	MA_GER_PUBLICA	VISA RECURRENTE	S/ 3,150.00
8	MA_INGENIERIA_AMBIENTAL	PAGO EN BANCO DE LA NACION	S/ 18,010.00
9	MA_INGENIERIA_AMBIENTAL	VISA RECURRENTE	S/ 2,700.00
10	MA_PROYECTOS_INVERSION	PAGO EN BANCO DE LA NACION	S/ 42,770.00
11	MA_PROYECTOS_INVERSION	VISA RECURRENTE	S/ 3,150.00
12	MA_TRIBUTACION_ASESORIA	PAGO EN BANCO DE LA NACION	S/ 71,100.00
13	MA_TRIBUTACION_ASESORIA	PAGO EN EFECTIVO	S/ 450.00
14	MA_TRIBUTACION_ASESORIA	VISA RECURRENTE	S/ 4,950.00
15	MAE_DOCENCIA_UNIVERSITARIA_IN VESTIGACION	PAGO EN BANCO DE LA NACION	S/ 55,200.00
16	MAE_DOCENCIA_UNIVERSITARIA_IN VESTIGACION	VISA RECURRENTE	S/ 400.00
17	MAE_INGENIERIA_HIDRAULICA	PAGO EN BANCO DE LA NACION	S/ 46,700.00
18	MAE_INGENIERIA_HIDRAULICA	VISA RECURRENTE	S/ 2,250.00
19	MAE_MECANICA_ELECTRICA	PAGO EN BANCO DE LA NACION	S/ 66,120.00
20	MAE_RECURSOS_HIDRICOS	PAGO EN BANCO DE LA NACION	S/ 26,920.00
21	MAE_RECURSOS_HIDRICOS	VISA RECURRENTE	S/ 9,000.00
22	MAE_SALUD_PUBLICA	PAGO EN BANCO DE LA NACION	S/ 58,820.00
23	MAE_SALUD_PUBLICA	VISA RECURRENTE	S/ 2,700.00
24	MAE_SERVICIOS_SALUD	PAGO EN BANCO DE LA NACION	S/ 30,320.00
25	MAE_SERVICIOS_SALUD	VISA RECURRENTE	S/ 3,600.00

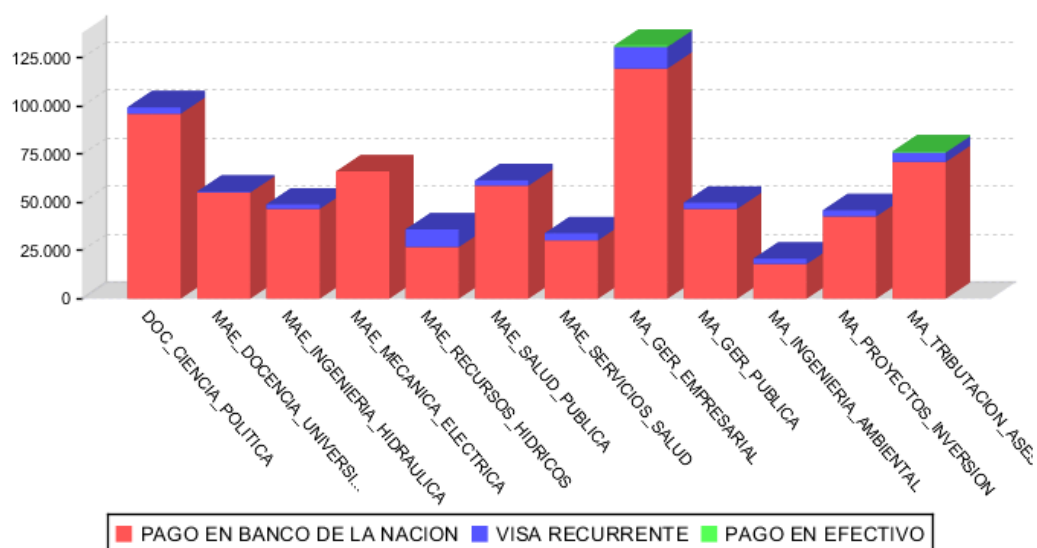


Figura NRO. 92: Reporte de ingresos por forma de pago de los programas de admisión 2017-II

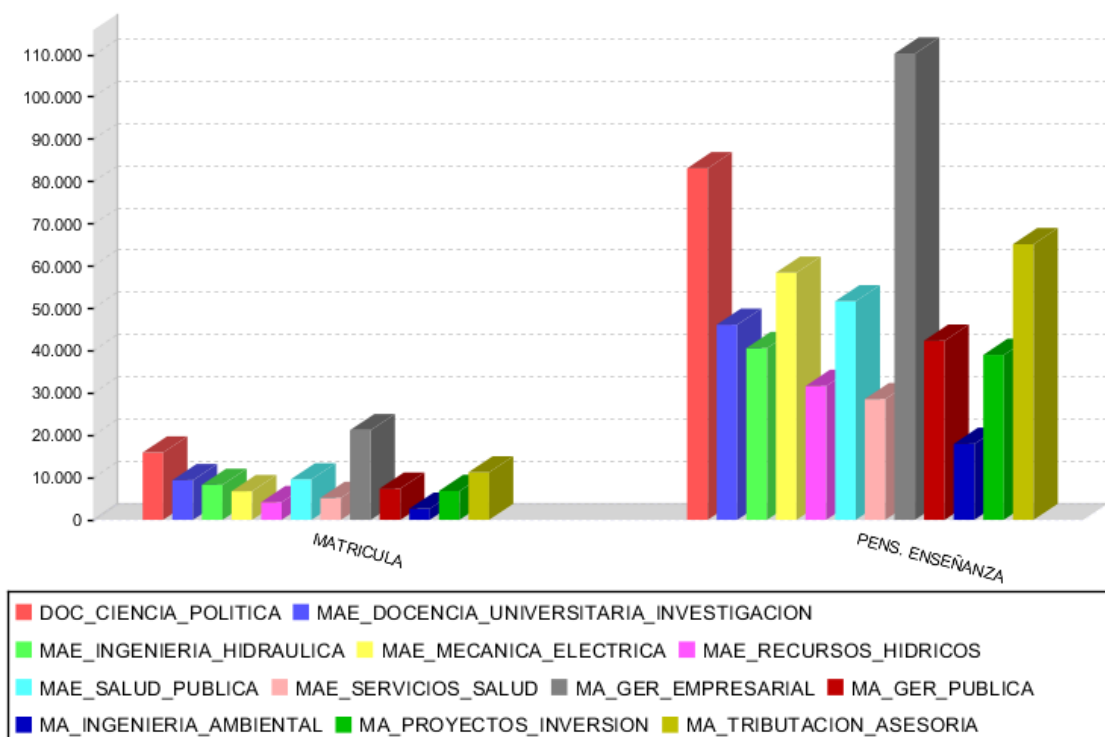


Figura NRO. 93: Reporte de ingresos por programas de admisión 2017-II de los conceptos matrícula y pensión de enseñanza

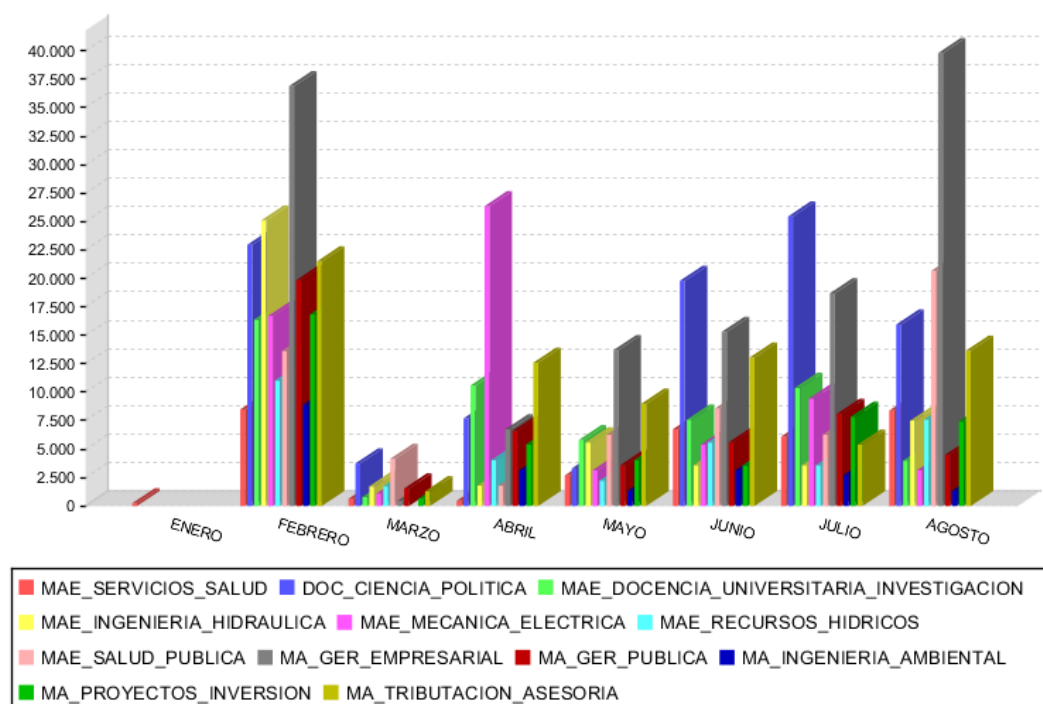


Figura NRO. 94: Reporte de ingresos por mes de los programas de admisión 2017-II por conceptos matrícula y pensión de enseñanza

Tabla 16: Rango de las edades de los alumnos del proceso de admisión 2017-I y 2017-II

ADMISION 2017-I / 2017-II			
Nro	S. Ingreso	Rango	Cant.
1	2017-I	[ 20 - 30 ]	292
2	2017-I	[ 31 - 40 ]	238
3	2017-I	[ 41 - 50 ]	75
4	2017-I	[ 51 - 60 ]	31
5	2017-I	[ 61 - 70 ]	1
6	2017-II	[ 20 - 30 ]	172
7	2017-II	[ 31 - 40 ]	142
8	2017-II	[ 41 - 50 ]	47
9	2017-II	[ 51 - 60 ]	22
10	2017-II	[ 61 - 70 ]	4



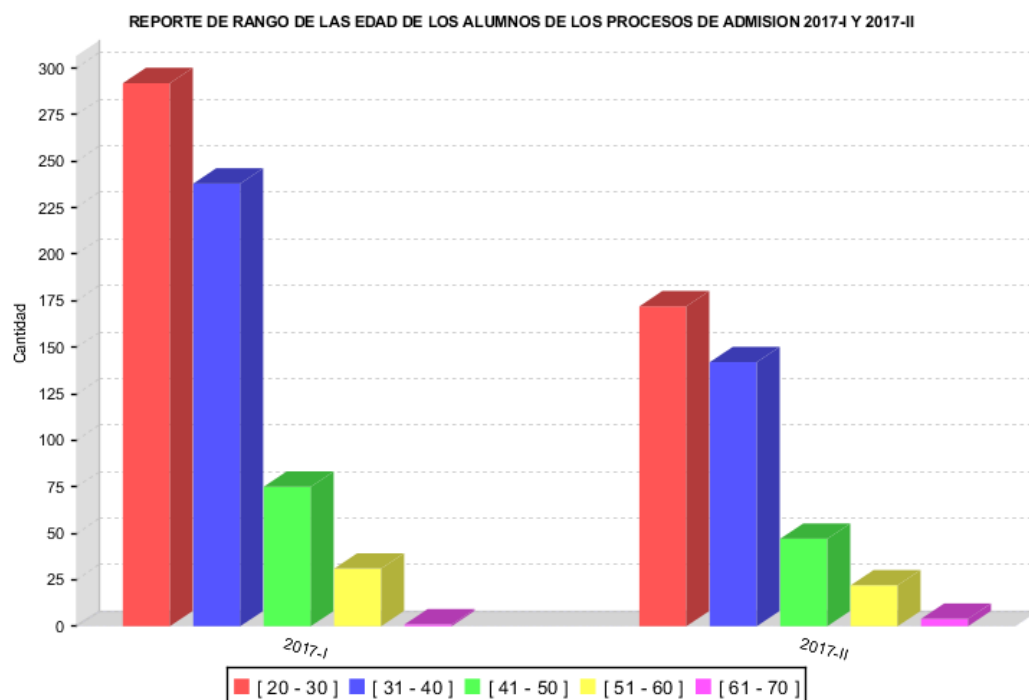


Figura NRO. 95: Reporte del rango de edades de los alumnos del proceso de admisión 2017-I Y 2017-II

Tabla 17: Cantidad de alumnos por universidad de procedencia del proceso de admisión 2017-I y 2017-II

ADMISION 2017-I / 2017-II				
UNIVERSIDAD DE PROCEDENCIA				
Nro	S. Ingreso	Tipo Univ.	Cant.	%
1	2017-I	NACIONAL	343	53.51
2	2017-I	OTROS	1	
3	2017-I	PARTICULAR	298	46.49
4	2017-II	NACIONAL	276	68.66
5	2017-II	PARTICULAR	126	31.34

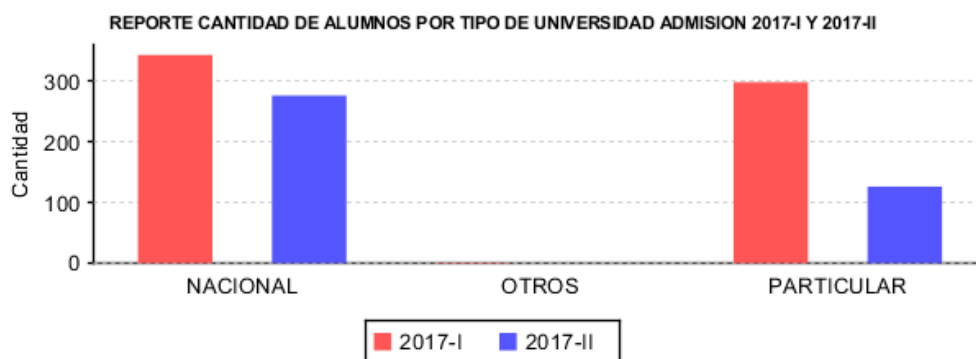


Figura NRO. 96: Cantidad de alumnos por universidad de procedencia del proceso de admisión 2017-I y 2017-II

**Tabla 18: Cantidad de profesiones por tipo de universidad del proceso de admisión 2017-I**

ADMISION 2017-I			
NRO	PROFESION	NAC.	PAR.
1	ABOGADO	105	178
2	ADMINISTRADOR	19	27
3	ARQUITECTO	19	15
4	BIOLOGO	3	0
5	CIENCIAS DE LA COMUNICACION	1	2
6	CONTADOR	34	25
7	DOCENTE	1	0
8	ECONOMISTA	33	6
9	ESTADISTICO	3	0
10	INGENIERO AGRICOLA	10	2
11	INGENIERO AGROINDUSTRIAL	2	0
12	INGENIERO AGRONOMO	11	0
13	INGENIERO AMBIENTAL	2	0
14	INGENIERO CIVIL	45	19
15	INGENIERO COMERCIAL	0	2
16	INGENIERO DE SISTEMAS	17	17
17	INGENIERO ELECTRONICO	1	0
18	INGENIERO EN INDUSTRIAS ALIMENTARIAS	3	0
19	INGENIERO INDUSTRIAL	2	2
20	INGENIERO MECANICO Y ELECTRICO	6	2
21	INGENIERO QUIMICO	10	0
22	INGENIERO ZOOTECNISTA	3	0
23	LIC. EN NEGOCIOS INTERNACIONALES	6	0
24	MEDICO VETERINARIO	1	0
25	MICROBIOLOGO	1	0
26	SOCIOLOGO	3	0
27	TURISMO	2	1

**Tabla 19: Cantidad de profesiones por tipo de universidad del proceso de admisión 2017-II**

<b>ADMISION 2017-II</b>			
<b>NRO</b>	<b>PROFESION</b>	<b>NAC.</b>	<b>PAR.</b>
1	ABOGADO	23	20
2	ADMINISTRADOR	11	29
3	ARQUITECTO	2	2
4	BIOLOGO	9	0
5	CIENCIAS DE LA COMUNICACION	1	1
6	CONTADOR	26	24
7	DOCENTE	17	1
8	ECONOMISTA	21	3
9	ENFERMERA	4	5
10	ESTADISTICO	2	0
11	ESTOMOTOLOGIA	0	1
12	FISICO	3	0
13	INGENIERIA EN ENERGIA	1	0
14	INGENIERO AGRICOLA	44	0
15	INGENIERO AGROINDUSTRIAL	0	1
16	INGENIERO AGRONOMO	6	0
17	INGENIERO AMBIENTAL	1	1
18	INGENIERO CIVIL	14	7
19	INGENIERO COMERCIAL	0	1
20	INGENIERO DE SISTEMAS	5	4
21	INGENIERO ELECTRONICO	3	0
22	INGENIERO EN INDUSTRIAS ALIMENTARIAS	2	0
23	INGENIERO GEOGRAFO	1	0
24	INGENIERO HIDRAULICO	1	0
25	INGENIERO INDUSTRIAL	0	3
26	INGENIERO MECANICO Y ELECTRICO	29	8
27	INGENIERO QUIMICO	7	0
28	INGENIERO ZOOTECNISTA	6	0
29	LIC. EN NEGOCIOS INTERNACIONALES	11	0
30	MARKETING	0	1
31	MATEMATICO	5	0
32	MEDICO CIRUJANO	5	5
33	MEDICO VETERINARIO	12	0
34	OBSTETRA	0	1
35	ODONTOLOGO	1	5
36	OTROS	2	1
37	PSICOLOGO	0	1
38	SOCIOLOGO	1	0
39	TEGNOLOGIA MEDICA - RADIOLOGIA	0	1

**Tabla 20: Cantidad de alumnos aptos admisión 2017-I**

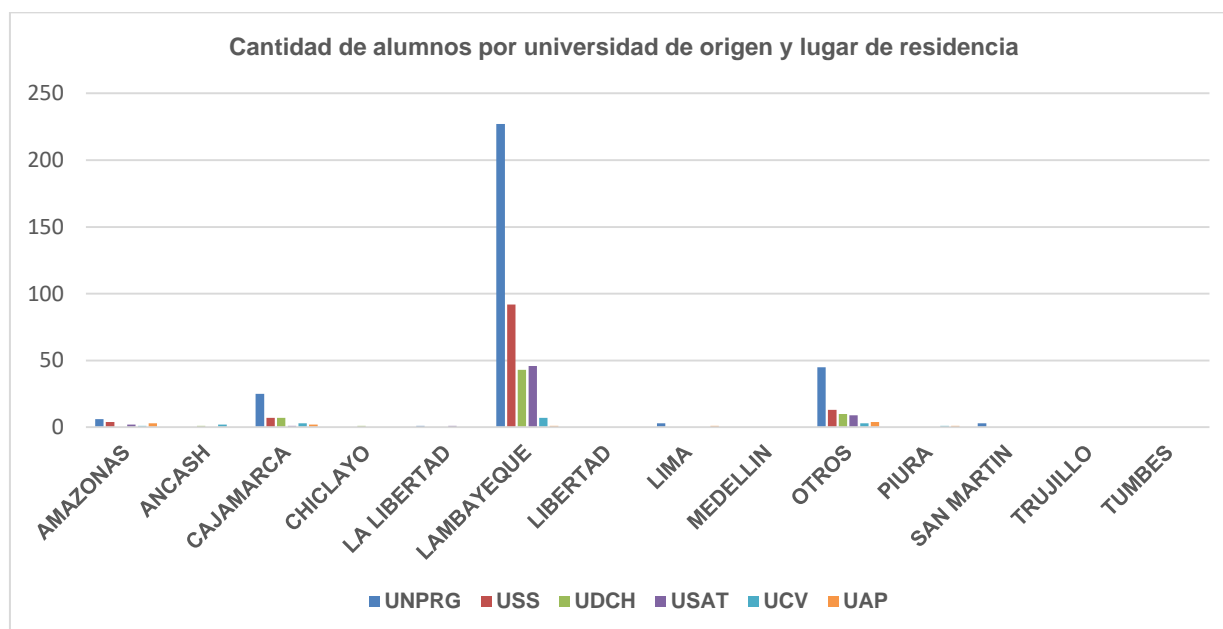
ADMISION 2017-I			
NRO	PROGRAMAS DE MAESTRIAS Y DOCTORADOS	TOTAL ALUMNOS INGRESANTES	TOTAL ALUMNOS HABILIES
1	MA EN PROYECTOS DE INVERSION	37	36
2	MA EN GERENCIA EMPRESARIAL	68	67
3	MA EN GERENCIA PUBLICA	50	37
4	MA EN AUDITORIA DE LA GESTION GUBERNAMENTAL	22	18
5	MA EN INGENIERIA AMBIENTAL	24	14
6	MA EN TERRITORIAL Y DESARROLLO URBANO	33	31
7	MA EN TRIBUTACION Y ASESORIA FISCAL	35	33
8	MA EN DERECHO CIVIL Y COMERCIAL	48	47
9	MA EN DERECHO CIENCIAS PENALES	142	139
10	MA EN DERECHO CONSTITUCIONAL Y GOBERNABILIDAD	51	49
11	MA EN DERECHO DEL TRABAJO	31	31
12	MA EN GERENCIA DE OBRAS Y CONSTRUCCION	72	70
13	MA EN GESTION DEL SOFTWARE	29	27
POBLACION TOTAL		<b>642</b>	<b>599</b>
PORCENTAJE DE ALUMNOS RETIRADOS		<b>43</b>	<b>7%</b>

**Tabla 21: Cantidad de alumnos aptos 2017-II**

ADMISION 2017-II			
NRO	PROGRAMAS DE MAESTRIAS Y DOCTORADOS	TOTAL ALUMNOS INGRESANTES	TOTAL ALUMNOS HABILIES
1	DO EN DERECHO Y CIENCIA POLITICA	36	36
2	MA EN GERENCIA EMPRESARIAL	61	61
3	MA EN GERENCIA PUBLICA	50	49
4	MA EN TRIBUTACION Y ASESORIA FISCAL	38	38
5	MA EN GERENCIA DE SERVICIOS DE SALUD	18	18
6	MA EN DOCENCIA UNIVERSITARIA	30	30
7	MA EN INGENIERIA AMBIENTAL	18	10
8	MA EN INGENIERIA HIDRAULICA	44	44
9	MA EN PROYECTOS DE INVERSION	29	26
10	MA EN MECANICA Y ELECTRICA	31	31
11	MA EN RECURSOS HIDRICOS	20	20
12	MA EN SALUD PUBLICA	27	26
POBLACION TOTAL		<b>402</b>	<b>389</b>
PORCENTAJE DE RETIRADOS		<b>13</b>	<b>3%</b>

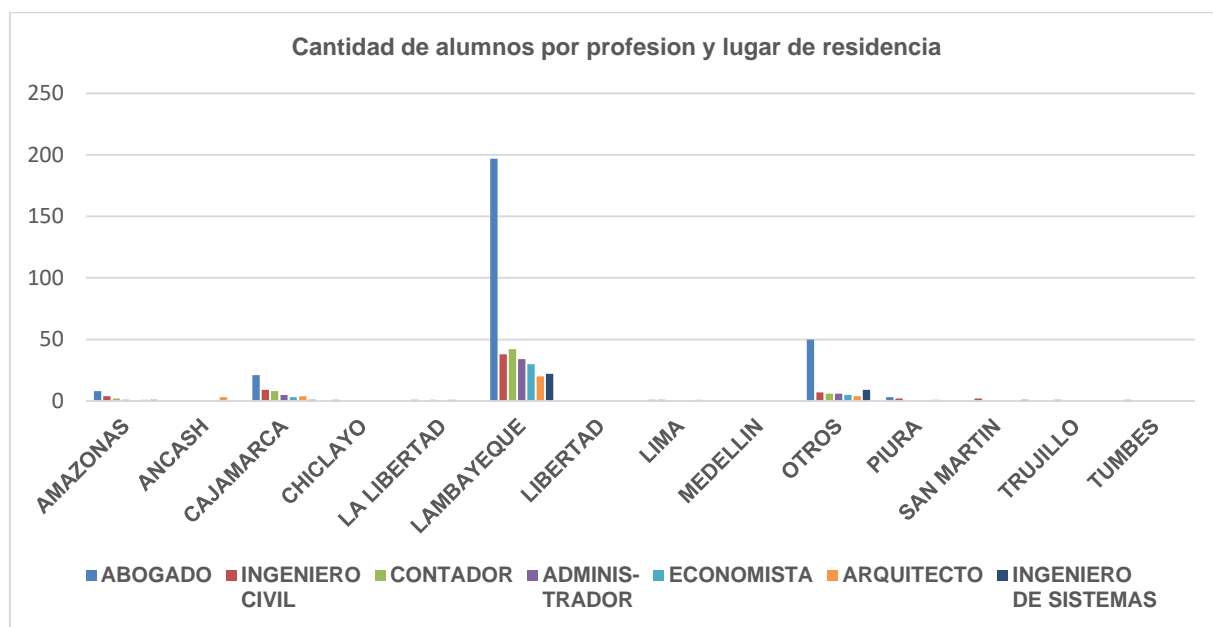
Tabla 22: Cantidad de postulantes de diferentes universidades y departamentos que estudian en la EPG

ADMISION 2017-I NRO DE ALUMNOS POR DEPARTAMENTO Y UNIVERSIDAD						
LUGAR DE RESIDENCIA	UNPRG	USS	UDCH	USAT	UCV	UAP
AMAZONAS	6	4	0	2	1	3
ANCASH	0	0	1	0	2	0
CAJAMARCA	25	7	7	1	3	2
CHICLAYO	0	0	1	0	0	0
LA LIBERTAD	1	0	0	1	0	0
LAMBAYEQUE	227	92	43	46	7	1
LIBERTAD	0	0	0	0	0	0
LIMA	3	0	0	0	0	1
MEDELLIN	0	0	0	0	0	0
OTROS	45	13	10	9	3	4
PIURA	0	0	0	0	1	1
SAN MARTIN	3	0	0	0	0	0
TRUJILLO	0	0	0	0	0	0
TUMBES	0	0	0	0	0	0



**Tabla 23: Cantidad de profesionales que postulan a la EPG por departamento**

ADMISION 2017-I NRO DE ALUMNOS POR DEPARTAMENTO Y PROFESIÓN							
LUGAR DE RESIDENCIA	ABOGADO	ING. CIVIL	CONTADOR	ADMINIS-TRADOR	ECONOMISTA	ARQUITECTO	ING. DE SISTEMAS
AMAZONAS	8	4	2	1	0	1	1
ANCASH	0	0	0	0	0	3	0
CAJAMARCA	21	9	8	5	3	4	1
CHICLAYO	1	0	0	0	0	0	0
LA LIBERTAD	1	0	1	0	1	0	0
LAMBAYEQUE	197	38	42	34	30	20	22
LIBERTAD	0	0	0	0	0	0	0
LIMA	1	1	0	0	0	1	0
MEDELLIN	0	0	0	0	0	0	0
OTROS	50	7	6	6	5	4	9
PIURA	3	2	0	0	0	1	0
SAN MARTIN	0	2	0	0	0	0	1
TRUJILLO	0	1	0	0	0	0	0
TUMBES	1	0	0	0	0	0	0



Después de ver los reportes obtenidos del DataMart, a partir de los ingresos recaudados por diversos conceptos y datos personales como matrícula, pensión de enseñanza, profesión, Lugar de residencia, fecha de nacimiento, etc. Y mediante la forma de pago a través de, Banco de la Nación, Visa Recurrente, Visa Pos etc. Se obtuvieron 5 perfiles financieros de los siguientes sectores.

- Sector Estratégico (Alta dirección, Dirección General, Directores, Gerentes, Recursos Humanos, etc.)
- Sector Económico (Económico- Financiero, gestores comerciales, etc.)
- Sector Ciencia, Tecnología e Ingeniería (Ingeniería Civil, Ingeniería Informática, etc.)
- Sector Justicia (Ministerio Público, Poder Judicial, etc.)
- Sector Salud (Médicos, Enfermeras, Obstetras, etc.)

Se llegó a la conclusión que el perfil financiero idóneo para el alumno se encuentra entre una edad de 20 – 40, de sexo masculino o femenino y que pertenezca a los programas de Derecho, Administración, Proyectos de Inversión, Gerencia de Obras y Construcción, Ingeniería Ambiental, que sea dependiente y del departamento de Lambayeque de profesión Abogado, Contador, Ingeniero Civil.

A continuación se detallan los 5 perfiles obtenidos de la información recopilada a través del DataMart.

<b>Perfil Nro.: 1</b>			
<b>Áreas relacionadas</b>	Derecho, Ciencias Administrativas y Contables, Ingeniería		
<b>Programas de mayor demanda</b>	Maestrías en Derecho, Maestrías en Administración, Maestría En Ciencias Con Mención En Ingeniería Ambiental.		
<b>Edad del alumno</b>	20 – 40	<b>Sexo</b>	masculino o femenino
<b>Profesión</b>	Abogado		
<b>Trabajo</b>	Dependiente / Fijo		
<b>Departamento</b>	Lambayeque, Cajamarca, Amazonas, Piura		
<b>Universidad</b>	Universidad Nacional Pedro Ruiz Gallo(UNPRG) Universidad Señor de Sipán(USS) Universidad Particular de Chiclayo(UDCH)		
<b>Observación</b>	Perfil idóneo, alumnos jóvenes, mayor demanda, retorno de inversión alto		
<b>Perfil de riesgo</b>	Bajo		

<b>Perfil Nro.: 2</b>			
<b>Áreas relacionadas</b>	Ciencias Administrativas y Contables		
<b>Programas de mayor demanda</b>	Maestrías en Administración, Maestría en Proyectos de Inversión		
<b>Edad del alumno</b>	20 – 40	<b>Sexo</b>	masculino o femenino
<b>Profesión</b>	Contador		
<b>Trabajo</b>	Dependiente / Fijo		
<b>Departamento</b>	Lambayeque, Cajamarca, Amazonas, Piura		
<b>Universidad</b>	Universidad Nacional Pedro Ruiz Gallo(UNPRG) Universidad Señor de Sipán(USS)		
<b>Observación</b>	Perfil idóneo, alumnos jóvenes, mayor demanda, retorno de inversión alto		
<b>Perfil de riesgo</b>	Bajo		



Perfil Nro.: 3			
<b>Áreas relacionadas</b>	Ingenierías		
<b>Programas de demanda media</b>	Maestría en Gerencia de Obras, Maestría en Proyectos de Inversión, Maestría En Ciencias Con Mención En Ingeniería Ambiental.		
<b>Edad del alumno</b>	20 – 40	<b>Sexo</b>	masculino o femenino
<b>Profesión</b>	Ing. Civil		
<b>Trabajo</b>	Dependiente / Fijo		
<b>Departamento</b>	Lambayeque, Cajamarca, Amazonas, Piura		
<b>Universidad</b>	Universidad Nacional Pedro Ruiz Gallo(UNPRG), Universidad Señor de Sipán(USS)		
<b>Observación</b>	Perfil idóneo, alumnos jóvenes, mayor demanda, retorno de inversión alto		
<b>Perfil de riesgo</b>	Bajo		

Perfil Nro.: 4			
<b>Áreas relacionadas</b>	Gestión Pública, Salud, médicos enfermeros, etc.		
<b>Programas de menor demanda</b>	Maestría en Salud Pública, Maestría en Ciencias con mención en Gerencia de Servicios de Salud		
<b>Edad del alumno</b>	20 – 40	<b>Sexo</b>	masculino o femenino
<b>Profesión</b>	Enfermera, Matemático, Físico, Odontólogo, Médico Veterinario		
<b>Trabajo</b>	Independiente		
<b>Departamento</b>	Lambayeque		
<b>Universidad</b>	Universidad Nacional Pedro Ruiz Gallo(UNPRG) Universidad Señor de Sipán(USS) Universidad Particular de Chiclayo(UDCH)		
<b>Observación</b>	Menor demanda idóneo, retorno de inversión medio		
<b>Perfil de riesgo</b>	Medio		

<b>Perfil Nro.: 5</b>			
<b>Áreas relacionadas</b>	Arquitectos, Ingenieros sanitarios, etc.		
<b>Programas de menor demanda</b>	Maestría En Ciencias Con Mención En Ordenamiento Territorial Y Desarrollo Urbano, Maestría en ciencias con mención en Docencia Universitaria e Investigación Educativa		
<b>Edad del alumno</b>	20 – 40	<b>Sexo</b>	masculino o femenino
<b>Profesión</b>	Economista, Contador , Administrador Ingeniero de sistemas Lic. En negocios internacionales Ciencias de la Comunicación Ingeniero Industrial , Ingeniero Agrícola Médico Veterinario, Ingeniero Mecánico y Eléctrico		
<b>Trabajo</b>	Independiente		
<b>Departamento</b>	Lambayeque		
<b>Universidad</b>	Universidad Nacional Pedro Ruiz Gallo(UNPRG) Universidad Señor de Sipán(USS) Universidad Particular de Chiclayo(UDCH) Universidad Santo Toribio de Mogrovejo(USAT)		
<b>Observación</b>	Menor demanda no es idóneo, retorno de inversión bajo		
<b>Perfil de riesgo</b>	Alto		

#### **4.5. Discusión de Resultados**

El análisis de los reportes obtenidos del cubo elaborado e implementado a través de la metodología de Ralph kimball pretende ser útil para la toma de decisiones de la EPG.

De este modo, a partir de este análisis se va a llevar a cabo un conjunto de decisiones que se podrían tomar a partir de la información que este informe refleja.

Estas decisiones son más fáciles de tomar gracias al análisis proporcionado por las diferentes técnicas de Business Intelligence llevadas a cabo anteriormente. Además, son decisiones importantes y que aportan beneficios a las empresas de diferentes rubros.

El propósito de la investigación fue determinar los perfiles financieros bajo la influencia del uso de Business Intelligence en la toma de decisiones, los resultados tomaron en consideración distintos tipos de variables que permitieron tomar decisiones que posteriormente afecten e influyan en sus procesos internos y externos.

Los resultados de la investigación son congruentes con los de (Olavarria Paz, 2014); el cual determinó que la implementación de herramientas de Business Intelligence en su proyecto permitió a la UNPRG incrementar la cantidad de postulantes y tomar mejores decisiones e incrementar la seguridad de las mismas. En el caso del presente estudio, se enfocó en la toma de decisiones desde sus resultados, encontrando que el uso de herramientas de BI, tenía una influencia positiva en los resultados de la toma de decisiones. Por su parte, (Lopez, 2008); implemento herramientas de BI en empresas peruanas, facilitando el acceso a información actualizada y estructurada en sus respectivas organizaciones. Logrando determinar que su implementación tuvo influencia en la velocidad y la metodología con las que se tomaban las decisiones dentro de su organización. Por otro lado (Loyo, 2017); tuvo como objetivo de estudio implementar una solución de Inteligencia de Negocios (BI) para el manejo de Indicadores claves de desempeño (KPI) en ventas de los Retails de Farmacias.

De igual manera, el presente trabajo de investigación pudo determinar que, el BI se debe aplicar en diferentes sectores de un país como salud, economía, medicina, etc. La implementación de herramientas de toma de decisiones en base a los datos tiene un efecto positivo en los resultados de dichas decisiones.

## **GLOSARIO DE MNEMONICOS**

**BI:** Business Intelligence

**DW:** Data Warehouse

**DWOF:** Data Warehouse Open Free

**OLAP:** Procesamiento Analítico en Linea (eng: Online Analytical Processing)

**ROLAP:** Procesamiento Analítico en Linea Relacional (eng: Relational Online Analytical Processing)

**MOLAP:** Procesamiento Analítico en Linea Multidimensional (eng: Multidimensional Online Analytical Processing)

**HOLAP:** Procesamiento Analítico en Linea Híbrido (eng: Hybrid Online Analytical Processing)

**FOSS:** Free/Open Source Software

**ETL:** Extracción, Transformación y Carga (eng: Extraction-Transformation-Load)

**DS:** Data Source, origen de datos.

**RDBMS:** Relational Database Management System

**Quartz:** Quartz es un framework de scheduling open source que provee funcionalidad avanzada para la calendarización de tareas en Java

## **CAPITULO V: CONCLUSIONES Y RECOMENDACIONES**

## **5.1. Conclusiones**

- a) Se logró realizar un análisis del funcionamiento de los procedimientos académicos y financieros actuales del proceso de admisión de la EPG, para conocer la problemática y determinar lo que se desea implementar cumpliendo con el objetivo deseado.
- b) Se logró realizar el análisis y diseño del modelo de minería de datos histórico de las transacciones de tipo financiero por diversos conceptos, utilizando las herramientas Pentaho BI open source, que permitieron con mayor facilidad poblar nuestro modelo de datos de forma eficiente y sin errores, bajo el enfoque Ralph Kimball.
- c) Con ayuda del modelo propuesto se logró generar reportes analíticos que permitieron conocer diversos patrones de comportamiento del negocio, para una toma de decisiones acertada, que permita beneficiar no solo económicamente sino también académicamente a los usuarios de la EPG.
- d) Se logró interpretar y obtener perfiles financieros deseados de los alumnos, a través de reportes resultantes del análisis realizado y obtenido, producto de las operaciones en línea, además de otras características como sexo, fecha de nacimiento, profesión, etc., al finalizar el trabajo realizado se identificaron 5 perfiles financieros que nos ayudaron a determinar que tan buenos o malos pagadores tiene la EPG durante cada proceso de admisión.
- e) Se logró evaluar e identificar que para un buen funcionamiento del modelo datos, se debe depurar los datos de la base de datos relacional, del Sistema de Inscripción de postulantes para los procesos de admisión de la EPG, antes de cargarlos al Data Mart a través de los procesos ETL definidos. La metodología de Kimball fue de gran utilidad para el análisis, diseño e implementación del Data Mart perfil financiero, permitiendo concretar las actividades planificadas en los requerimientos descritos anteriormente para la EPG, a través de este trabajo de investigación descriptivo, se llegó a la conclusión que del modelo de datos y reportes. se obtuvo información para definir varios perfiles financieros idóneos del postulante, que más adelante se convertirá en alumno de la EPG.

## **5.2. Recomendaciones**

- a) Se sugiere utilizar la metodología de Ralph Kimball en los procesos académicos de la EPG y UNPRG para mantener actualizado no solo la información de tipo financiera sino académica y conocer en tiempo real el rendimiento académico de los alumnos, porque esta metodología y herramientas ofrecen una solución rápida de implementar y fácil de comprender para todos los usuarios.
- b) Del modelo financiero propuesto en esta investigación, se debería complementar con los procesos académicos que se llevan a cabo, cuando el potencial postulante se convierte en alumno, ya que ambos son complementarios y a su vez los datos obtenidos de la base de datos relacional del sistema de inscripción nos permitirá hacer un nuevo análisis en función a las necesidades que se presenten.
- c) Se sugiere crear y conectar un modelo de análisis a la Escuela de Posgrado creado en esta investigación hacia la oficina de grados y títulos para conocer información académica en línea de los grados de bachiller, títulos profesionales, grados de maestro y doctor que cada año son solicitados por la EPG y alta dirección para enviar a SUNEDU.

## **CAPÍTULO VI: REFERENCIAS BIBLIOGRÁFICAS**



## BIBLIOGRAFÍA

Dean Browne, B. D. (2010). *IBM Cognos Business Intelligence V10.1 Handbook*. Armonk, Nueva York.: IBM Corp. 2010.

Elizabeth Vitt, M. L. (2003). *Business intelligence*. España: McGraw-Hill Interamericana de España.

Falcón, N. A. (2012). Desarrollo de una solución de Inteligencia de Negocios en el manejo de estadísticas de control en la venta de repuestos de la empresa Talleres Ambamazda S.A. de la ciudad de Ambato. 152.

GAUCHET, T. (2015). *SQL Server 2014. Implementación De Una Solución De Business Intelligence (SQL Server, Analysis Services, Power BI...)*. New Jersey: ENI.

Lopez, R. A. (2008). Impacto de la Data WareHouse e Inteligencia de negocios en el desempeño de las Empresas: Investigación Empírica en Perú, como país en vías de desarrollo. 276.

Loyo, I. R. (2017). Modelo de Inteligencia de negocios (BI), para el manejo de indicadores clave de desempeño (KPI) en ventas para la toma de decisiones en los RETAILS de Farmacias de la Empresa Farmaenlace CÍA. Ltda. 117.

Marques, M. P. (2014). *Business Intelligence: Técnicas, Herramientas y Aplicaciones*. España: RC LIBROS (SC LIBRO).

Michael Minelli, M. C. (2012). *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses (Wiley CIO)*. Hoboken, New Jersey: Jhon Wiley & Sons.

Olavarria Paz, J. B. (2014). Modelo de Explotación de datos históricos de los registros de Inscripción a los Procesos de Admisión, basado en BI y CRM, para el incremento de postulantes en el proceso de Admisión de la UNPRG – Lambayeque – Perú. 136.

Rio, L. M. (2000). *Mas alla del Business Intelligence: 16 Experiencias de Exito*. España: Ediciones Gestión 2000.

Rodríguez, I. N. (2012). Desarrollo de una solución de Inteligencia de Negocios en el manejo de estadísticas de control en la venta de repuestos de la empresa Talleres Ambamazda S.A. de la ciudad de Ambato. 152.

Rodríguez, I. N. (18 de Septiembre de 2012). [http://repositorio.uta.edu.ec](http://repositorio.uta.edu.ec/bitstream/123456789/3008/1/Tesis_t781mbd.pdf). Obtenido de [http://repositorio.uta.edu.ec/bitstream/123456789/3008/1/Tesis\\_t781mbd.pdf](http://repositorio.uta.edu.ec/bitstream/123456789/3008/1/Tesis_t781mbd.pdf)

Roland Bouman, J. v. (2010). *Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL*. Indianápolis: Wiley Publishing Inc.

Rud, O. P. (2009). *Business Intelligence Success Factors: Tools for Aligning Your Business in the Global Economy*. New Jersey: Jhon Wiley & Sons, Inc., Hoboken.

Scheps, S. (2008). *Business Intelligence For Dummies*. Indianapolis, Indiana: Wiley Publishing, Inc.

Thomas H. Davenport, J. G. (2007). *Competing on Analytics: The New Science of Winning*. Boston, Massachusetts: Harvard Business School Press.

Vitt, E., Luckevich, M., & Misner, S. (2003). *Business Intelligece: Tecnicas de Analisis para la toma de decisiones Importantes*. España: S.A. MCGRAW-HILL / Interamericana de España.

Williams S, W. N. (2010). *The profit impact of business intelligence*. Morgan Kaufmann.

# ANEXOS

## Anexo 1: Diseño de Extracción

### Tabla de destino: DIM\_ALUMNO

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción
1	id_al	int(11)			No	Ninguna		
2	Promocion_al	varchar(10)	latin1_swedish_ci		No	Ninguna		
3	Grupo_al	char(1)	latin1_swedish_ci		No	Ninguna		
4	Sexo_al	char(1)	latin1_swedish_ci		No	Ninguna		
5	id_pg_fk	int(11)			No	Ninguna		
6	Codigo_al	char(15)	latin1_swedish_ci		No	Ninguna		
7	ApellidosyNombres_al	varchar(150)	latin1_swedish_ci		No	Ninguna		
8	FechaNacimiento_al	datetime			No	Ninguna		
9	Singreso_al	varchar(10)	latin1_swedish_ci		No	Ninguna		
10	Estadoaca_al	varchar(2)	latin1_swedish_ci		Sí	NULL		

Para los elementos que están marcados:

### Tabla de destino: DIM\_FORMAPAGO

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción
1	id_fp	int(11)			No	Ninguna		
2	Description_fp	varchar(30)	latin1_swedish_ci		No	Ninguna		

Para los elementos que están marcados:

### Tabla de destino: DIM\_PARTIDA

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción
1	id_pa	int(11)			No	Ninguna		
2	Partida_pa	varchar(20)	latin1_swedish_ci		Sí	NULL		
3	Subpartida_pa	varchar(20)	latin1_swedish_ci		Sí	NULL		
4	Concepto_pa	varchar(100)	latin1_swedish_ci		Sí	NULL		

Para los elementos que están marcados:

### Tabla de destino: DIM\_PROGRAMA

#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción
1	id_pg	int(11)			No	Ninguna		
2	Sede_pg	varchar(50)	latin1_swedish_ci		Sí	NULL		
3	Programa_pg	varchar(220)	latin1_swedish_ci		Sí	NULL		
4	Tipo_pg	varchar(11)	latin1_swedish_ci		Sí	NULL		
5	Codigo	varchar(50)	latin1_swedish_ci		Sí	NULL		

Para los elementos que están marcados:

## Tabla de destino: DIM\_SEMESTRE

Servidor: 127.0.0.1 > Base de datos: perfil_financiero > Tabla: dim_semestre									
Examinar Estructura SQL Buscar Insertar Exportar Importar Privilegios Operaciones Más									
#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción	
1	id_sem	int(11)			No	Ninguna		Cambiar	Eliminar
2	codigo_sem	varchar(15)	latin1_swedish_ci		No	Ninguna		Cambiar	Eliminar

☐ Marcar todos Para los elementos que están marcados: Examinar Cambiar Eliminar Primaria Único Índice

## Tabla de destino: DIM\_TIEMPO

Servidor: 127.0.0.1 > Base de datos: perfil_financiero > Tabla: dim_tiempo									
Examinar Estructura SQL Buscar Insertar Exportar Importar Privilegios Operaciones									
#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción	
1	id_fecha_tm	date			No	Ninguna		Cambiar	Eliminar
2	anio_tm	int(11)			No	Ninguna		Cambiar	Eliminar
3	mes_tm	int(11)			No	Ninguna		Cambiar	Eliminar
4	dia_tm	int(11)			No	Ninguna		Cambiar	Eliminar

☐ Marcar todos Para los elementos que están marcados: Examinar Cambiar Eliminar Primaria Único Índice

## Tabla de destino: FACT\_CESTADISTICO

Servidor: 127.0.0.1 > Base de datos: perfil_financiero > Tabla: fact_cestadistico									
Examinar Estructura SQL Buscar Insertar Exportar Importar Privilegios Operaciones									
#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción	
1	id_ce	int(11)			No	Ninguna		Cambiar	Eliminar
2	naingresantes_ce	int(11)			No	Ninguna		Cambiar	Eliminar
3	nahabiles_ce	int(11)			No	Ninguna		Cambiar	Eliminar
4	nacondicion_ce	int(11)			No	Ninguna		Cambiar	Eliminar
5	id_sem_fk	int(11)			No	Ninguna		Cambiar	Eliminar
6	id_pg_fk	int(11)			No	Ninguna		Cambiar	Eliminar
7	grupo_ce	char(1)	latin1_swedish_ci		No	Ninguna		Cambiar	Eliminar
8	fecha_ce	date			No	Ninguna		Cambiar	Eliminar

☐ Marcar todos Para los elementos que están marcados: Examinar Cambiar Eliminar Primaria Único Índice

## Tabla de destino: FACT\_OPERACION

Servidor: 127.0.0.1 > Base de datos: perfil_financiero > Tabla: fact_operacion									
Examinar Estructura SQL Buscar Insertar Exportar Importar Privilegios Operaciones									
#	Nombre	Tipo	Cotejamiento	Atributos	Nulo	Predeterminado	Extra	Acción	
1	id_tr	int(11)			No	Ninguna		Cambiar	Eliminar
2	id_pa_fk	int(11)			No	Ninguna		Cambiar	Eliminar
3	id_fecha_fk	date			No	Ninguna		Cambiar	Eliminar
4	id_fp_fk	int(11)			No	Ninguna		Cambiar	Eliminar
5	id_al_fk	int(11)			No	Ninguna		Cambiar	Eliminar
6	subtotal_tr	decimal(10,3)			No	Ninguna		Cambiar	Eliminar

☐ Marcar todos Para los elementos que están marcados: Examinar Cambiar Eliminar Primaria Único Índice

## Anexo 2: Instalación de la Plataforma Pentaho 7.1.0.0-12

**Paso 1:** Se debe descargar Pentaho BI Platform Data Integration

<https://sourceforge.net/projects/pentaho/files/Data%20Integration/7.1/>

Home / [Data Integration](#) / 7.1

Name	Modified	Size	Downloads / Week
Parent folder			
<a href="#">pdi-ce-7.1.0.0-12.zip.sum</a>	2017-05-17	46 Bytes	29
<a href="#">pdi-ce-7.1.0.0-12.zip</a>	2017-05-17	903.9 MB	2,942
Totals: 2 Items		903.9 MB	2,971

**Paso 2:** Una vez realizada la descarga del Data Integration 7.1 se procede a descomprimir el archivo zip, para posteriormente ejecutar el archivo Spoon.bat

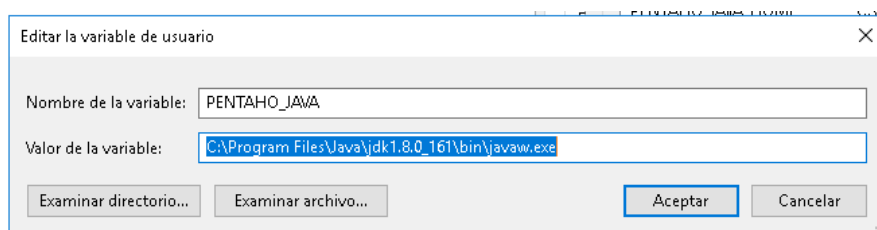
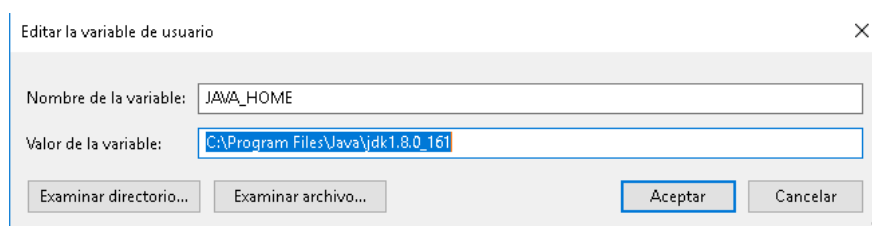
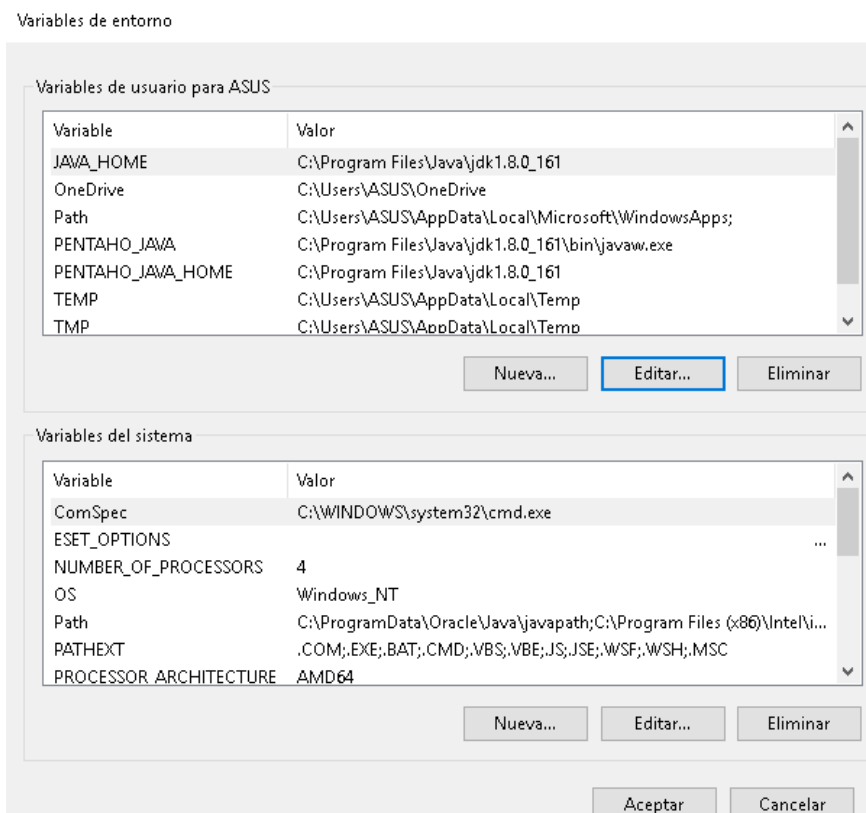
Este equipo > OS (C:) >	ste equipo > OS (C:) > data-integration >
Nombre	Nombre
data-integration	Fecha de modifica...
	Tipo
	spoon.command 16/05/2017 8:08 p.... Archi
	spoon.png 16/05/2017 8:08 p.... Archi
	Carte.bat 16/05/2017 8:08 p.... Archi
	Encr.bat 16/05/2017 8:08 p.... Archi
	Import.bat 16/05/2017 8:08 p.... Archi
	Kitchen.bat 16/05/2017 8:08 p.... Archi
	Pan.bat 16/05/2017 8:08 p.... Archi
	purge-utility.bat 16/05/2017 8:08 p.... Archi
	runSamples.bat 16/05/2017 8:08 p. m. Archi
	set-pentaho-env.bat 16/05/2017 8:08 p.... Archi
	Spark-app-builder.bat 16/05/2017 8:08 p.... Archi
	Spoon.bat 16/05/2017 8:08 p.... Archi
	SpoonConsole.bat 16/05/2017 8:08 p.... Archi
	SpoonDebug.bat 16/05/2017 8:08 p.... Archi
	carte.sh 16/05/2017 8:08 p.... Archi
	encr.sh 16/05/2017 8:08 p.... Archi
	import.sh 16/05/2017 8:08 p.... Archi
	kitchen.sh 16/05/2017 8:08 p.... Archi
	pan.sh 16/05/2017 8:08 p.... Archi
	purge-utility.sh 16/05/2017 8:08 p.... Archi
	runSamples.sh 16/05/2017 8:08 p.... Archi
	set-pentaho-env.sh 16/05/2017 8:08 p.... Archi
	snark-ann-builder.sh 16/05/2017 8:08 n.... Archi

**Paso 3:** Ahora configuraremos las variables de entorno de java, para trabajar con el software que hemos descargado.

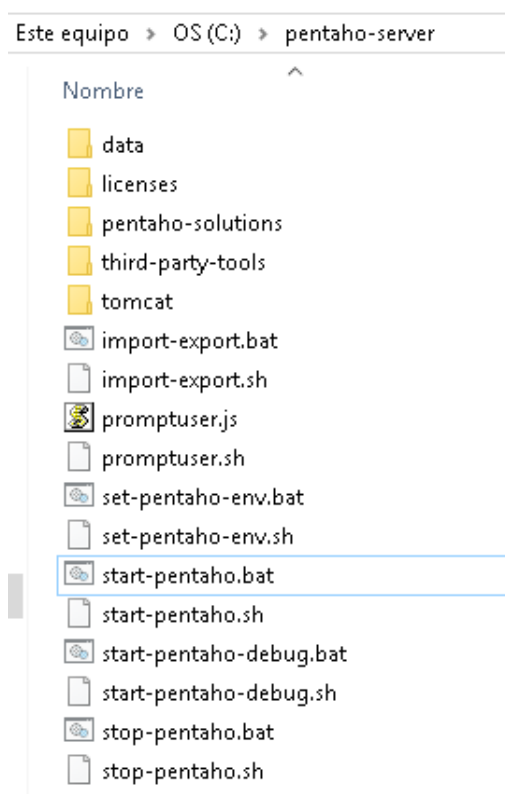
- Búsqueda de Windows escribir mi pc click derecho, propiedades, configuración avanzada del sistema, variables de entorno.

Nombre de la variable: JAVA\_HOME

Valor de la variable: JDK de Java “C:\Program Files\Java\jdk1.8.0\_161”



**Paso 4:** Después de haber configurado las variables de entorno de JAVA ahora ejecutaremos **pentaho-server** para ello debemos dirigirnos a la carpeta pentaho-server que se encuentra en la unidad C y ejecutamos **start-pentaho.bat**





**Paso 5:** Abrir una ventana de navegación y digitamos la siguiente url:  
<http://localhost:8080/pentaho/Login>

localhost:8080/pentaho/Login

**pentaho**  
A Hitachi Group Company

User Name:

Password:

Login

Login as an Evaluator ^

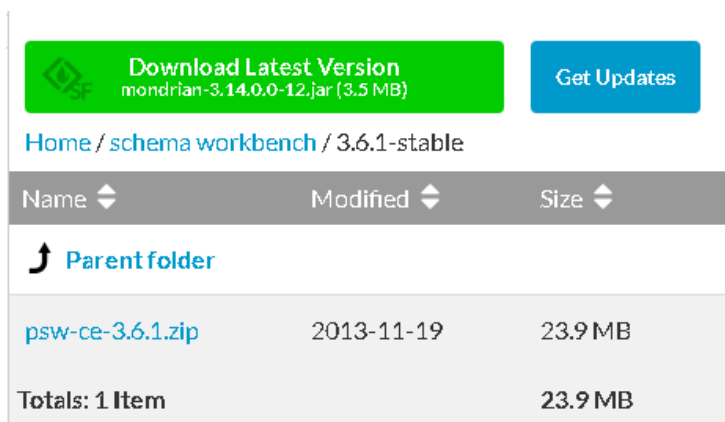
Administrator	Business User
<input type="text" value="Admin"/>	<input type="text" value="Suzy"/>
<input type="password" value="password"/>	<input type="password" value="password"/>
GO	GO

© 2005-2018 Pentaho Corporation. All rights reserved.

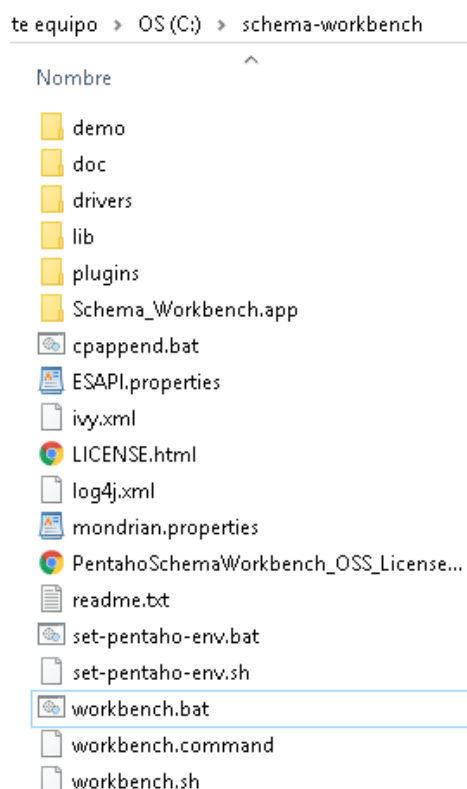
### Anexo 3: Instalación de Schema Workbench Pentaho

**Paso 1:** Se debe descargar Schema Workbench Pentaho

<https://sourceforge.net/projects/mondrian/files/schema%20workbench/3.6.1-stable/>



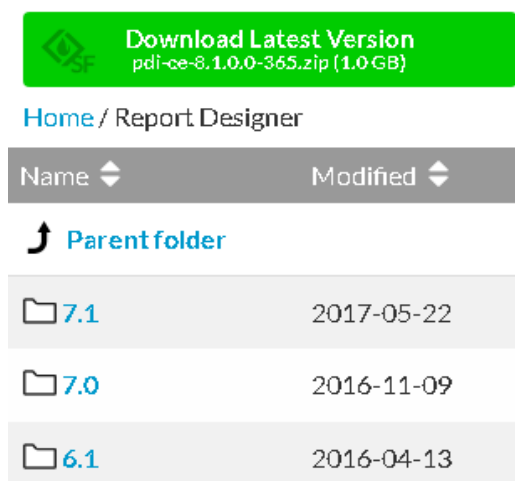
**Paso 2:** Una vez realizada la descarga se procede a descomprimir el archivo para su ejecución a través del archivo **workbench.bat**.



## Anexo 4: Instalación de Pentaho Designer

**Paso 1:** Se debe descargar Pentaho Designer

<https://sourceforge.net/projects/pentaho/files/Report%20Designer/>



**Paso 2:** Una vez realizada la descarga se procede a descomprimir el archivo y ejecutar el **report-designer.bat** que se encuentra en la carpeta respectiva.

