

Solución Business Intelligence para la empresa Nutribox

Nicolás Miguel Ayut González

Grado de Ingeniería Informática 75.650 - TFG - Business Intelligence

Xavier Martinez Fontes Atanasi Daradoumis Haralabus

06/2019



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada <u>3.0 España de Creative</u> <u>Commons</u>

FICHA DEL TRABAJO FINAL

Título del trabajo:	Solución Business Intelligence para la empresa Nutribox	
Nombre del autor:	Nicolás Miguel Ayut González	
Nombre del consultor/a:	Xavier Martinez Fontes	
Nombre del PRA:	Atanasi Daradoumis Haralabus	
Fecha de entrega (mm/aaaa):	06/2019	
Titulación:	Grado de Ingeniería Informática	
Área del Trabajo Final:	75.650 - TFG - Business Intelligence	
Idioma del trabajo:	Castellano.	
Palabras clave	Cupones, Healthy discounts, Pentaho, ETL, Cuadro de mando, OLAP, Business Intelligence	

Resumen del Trabajo (máximo 250 palabras): Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.

Este trabajo de fin de grado trata sobre el diseño e implementación de un prototipo inicial de BI para la empresa Nutribox, una empresa creada para gestionar la actividad de una plataforma (aplicación móvil y web service) basada en "healthy discounts".

Esta organización surge a partir del programa piloto "TFG emprendedor interestudios", por medio del cual, se propone crear una iniciativa emprendedora, a partir, de los trabajos de fin de grado de los estudiantes de diferentes áreas.

Así este prototipo pretende dotar a la dirección ejecutiva de esta entidad de información relativa a la consecución de objetivos establecidos inicialmente, por medio de cuadros de mando e informes que permitirán valorar la evaluación del negocio y una toma de decisiones en base a la información obtenida.

En este trabajo se detallan las fases de análisis, diseño e implantación de la solución. El prototipo creado consta de procesos de ETL, sistema de almacenamiento y gestión de datos (data warehouse), cubos OLAP, así como cuadros de mando e informes que permiten valorar los indicadores definidos en la fase de análisis.

Finalmente, en el capítulo "Líneas de futuro" se exponen una serie de mejoras o ampliaciones a instaurar en la solución que permitirán estudiar nuevos indicadores y realizar nuevos análisis que dotarán a la empresa de un mayor conocimiento de su actividad.

En definitiva, se puede considerar que se ha construido un prototipo de solución Bl para la empresa Nutribox que cumple los objetivos propuestos.

Abstract (in English, 250 words or less):

This final project deals with the design and implementation of an initial BI prototype for Nutribox, a company created to manage the activity of a platform (mobile application and web service) based on "healthy discounts".

This organization arises from the program "TFG emprendedor interestudios", by means of which it is proposed to create an enterprising initiative, from the final projects of students from different areas.

Therefore, this prototype aims to provide the executive management of this company the information related to the achievement of initially established objectives, through dashboards and reports, which will allow evaluating the business and making decisions based on the information obtained.

In this work the phases of analysis, design and implementation of the solution are detailed. The created prototype consists of ETL processes, storage system and data management (data warehouse), OLAP cubes, as well as dashboards and reports that allow to assess the indicators defined in the analysis phase.

Finally, in the chapter "Líneas de futuro" a series of improvements or extensions are exposed to be introduced in the solution that will allow studying new indicators and carry out new analyzes that will give the company a greater knowledge of its activity.

In conclusion, it can be considered that a prototype of BI solution has been built for the Nutribox company that agrees with the proposed objectives.

Índice

1. Introducción	1
1.1. Contexto y justificación del Trabajo	1
1.2. Objetivos del Trabajo	3
1.3. Planificación del Trabajo	4
1.4. Breve sumario de productos obtenidos	7
1.5. Breve descripción de los siguientes capítulos de la memoria	7
2. Análisis	7
2.1. Definición de fuentes y tipos de datos	8
2.2. Definición de KPI	11
3. Diseño	21
3.1. Base de datos operativa Nutribox y datos de prueba	22
3.2. Componentes de la solución	22
3.2.1. Data wharehouse	23
3.2.2. Procesos ETL	25
3.2.3. Cubo OLAP	25
3.2.4. Diseño de dashboard o cuadro de mandos	27
3.2.4.1. CM1: Cuadro de mando Administrativo Nutribox	27
3.2.4.2. CM2: Cuadro de mando para empresa colaboradora	31
3.2.5. Reports	34
3.2.5.1. informe de canjes por categoría y año	35
3.2.5.2. Informe top colaboradores del año	36
3.2.5.3. Informe Productos canjeados del año	37
3.3. Selección del software	38
3.3.1. Selección del Software de virtualización	38
3.3.2. Selección del sistema operativo utilizado en el prototipo	38
3.3.3. Selección del sistema gestor de bases de datos	38
3.3.3.1. MySQL	39
3.3.3.2. PostgreSQL	39
3.3.3.3. Elección de sistema gestor de base de datos	40
3.3.4. Selección de Suite Business Intelligence	40
4. Implantación	41
4.1. Descarga e instalación de componentes necesarios para desarro	llar la
solución.	41
4.1.1. Software de virtualización y sistema operativo	41
4.1.2. Sistema gestor de bases de datos	42
4.1.3. Instalación de Java	42
4.1.4. Descarga e instalación de los componentes necesarios de la	suite
BI Pentaho.	42
4.2. Creación de las bases de datos Nutribox y Datawharehouse	45
4.3. Creación de la data set de pruebas	45
4.4. Construcción de ETL	45
4.4.1. Transformación HechosCanje.ktr	47
4.4.2. Transformación DimTiempo.ktr	47
4.4.3. Transformación DimProductos.ktr	48
4.4.4. Transformación DimColaboradores.ktr	49

	4	4.4.5.	Transformación DimUsuarios.ktr	50
	4.5.	Cubo	os OLAP	51
	4.6.	Crea	ción de cuadros de mando	53
	4	1.6.1.	CM1: Cuadro de mando para la dirección ejecutiva de Nutribox	53
	4	1.6.2.	CM2: Cuadro de mando empresa colaboradora	62
	4.7.	Crea	ción de informes dinámicos en Pentaho Report Designer	69
	4	1.7.1.	Informe de canjes por categoría y año	70
	4	1.7.2.	Informe Top 10 colaboradores del año	71
	4	1.7.3.	Informe Productos canjeados del año	73
5.	L	_íneas	de futuro	74
6.	C	Conclu	siones	76
7.	E	Bibliogr	afía	77
8.	A	Anexos	5	78
	8.1.	Desc	carga y uso del prototipo	78
	8.2.	Uso	de generatedata.com para crear los datos de prueba	83
	8.3.	Proc	esos ETL con Pentaho Data Integration	86
	8.4.	Defir	nición del esquema OLAP	93
	8.5.	Anál	isis de los datos del cubo OLAP con jpivot 1	00
	8.6.	Anál	isis de los datos del cubo OLAP con saiku 1	03
	8.7.	Crea	ción de cuadros de mando con Pentaho CDE Dashboard 1	05
	8.8.	Crea	ción de informes con Pentaho Report Designer1	80

Lista de Imágenes

Imagen 1: Diagrama de Gantt del proyecto	6
Imagen 2: Estructura de la base de datos "Nutribox"	8
Imagen 3: Diseño de la infraestructura BI Nutribox	. 22
Imagen 4: Diagrama del Data wharehouse de Nutribox	. 24
Imagen 5: Ejemplo de Cubo OLAP	. 26
Imagen 6: Diseño cuadro de mando para la dirección ejecutiva de Nutribox	. 27
Imagen 7: Diseño cuadro de mando para empresas colaboradoras	. 31
Imagen 8: Iniciar pentaho al encender la máquina virtual.	. 44
Imagen 9: Transformación HechosCanje.ktr	. 47
Imagen 10: Transformación DimTiempo.ktr	. 47
Imagen 11: Transformación DimProductos.ktr	. 48
Imagen 12: Transformación DimColaboradores.ktr	. 49
Imagen 13: Transformación DimUsuarios.ktr	. 50
Imagen 14: Schema OLAP Nutribox	. 51
Imagen 15: Cuadro de mandos Nutribox para el año 2019	. 53
Imagen 16: Imagen cuadro de mandos Nutribox 2020	. 54
Imagen 17: Cuadro de mando para la empresa Papa Arrugada S.L	. 62
Imagen 18: Cuadro de mando para la empresa Cancajos S.L. del año 2020.	. 63
Imagen 19: Importar servicio virtualizado.	. 78
Imagen 20: Indicar la ruta del archivo.	. 79
Imagen 21: Sistema virtualizado.	. 79
Imagen 22: Proceso de importación	. 80
Imagen 23: Iniciar máquina virtual	. 80
Imagen 24: Acceder al servidor Pentaho Business Analytics	. 81
Imagen 25: Detalle Pentaho Business Analytics.	. 81
Imagen 26: Browse files en Pentaho Business Analytics.	. 82
Imagen 27: Generar datos de usuarios.	. 83
Imagen 28: Generar datos de cupones.	. 84
Imagen 29: Generar transacciones.	. 85
Imagen 30: Nueva Transformación PDI.	. 86
Imagen 31: Otra Nueva transformación.	. 86
Imagen 32: Área de trabajo PDI.	. 87
Imagen 33: Detalle del objeto table input para la tabla couponUse	. 87
Imagen 34: Detalle del objeto table input para la tabla coupons.	. 88
Imagen 35: Detalle del objeto Sort rows	. 89
Imagen 36: Detalle del objeto Merge join.	. 89
Imagen 37: Detalle objeto Calculator	. 90
Imagen 38: Detalle objeto Split fields.	. 90
Imagen 39: Detalle objeto String operations.	. 91
Imagen 40: Detalle objeto Calculator2.	. 91
Imagen 41: Detalle objeto Combination lookup/update	. 92
Imagen 42: Trabajo HechosCanjes.	. 93
Imagen 43: Área de trabajo Pentaho Schema Workbench	. 93
Imagen 44: Detalle conexión con la base de datos	. 94
Imagen 45: Área de trabajo Pentaho Schema Workbench 2	. 94
Imagen 46: Dimensiones del cubo.	. 95
Imagen 47: Detalle nivel colaborador.	. 96
-	

Imagen	48: Detalle nivel Año	96
Imagen	49: Detalle nivel Mes.	97
Imagen	50: Añadir uso de dimensión al cubo	97
Imagen	51: Añadir medida al cubo	98
Imagen	52: Detalle del esquema final.	98
Imagen	53: Publicar el esquema	99
Imagen	54: Crear nuevo	00
Imagen	55: Crear nueva vista JPivot1	00
Imagen	56: Medidas y jerarquías 1	01
Imagen	57: Ejemplo Jpivot tabla 1	01
Imagen	58: Gráfico generado con Jpivot1	02
Imagen	59: Crear nuevo1	03
Imagen	60: Seleccionar cubo1	03
Imagen	61: Área de trabajo Saiku 1	04
Imagen	62: Vista en tabla de consulta con Saiku1	04
Imagen	63: Gráfica generada con Saiku 1	05
Imagen	64: Crear nuevo CDE Dashboard1	05
Imagen	65: CDE Dashboard, detalle "Layout Panel" 1	06
Imagen	66: CDE Dashboard, detalle " Datasources Panel " 1	06
Imagen	67: CDE Dashboard, detalle "Component Panel" 1	07
Imagen	68: CDE Dashboard, guardar cuadro de mando 1	07
Imagen	69: Nuevo reporte 1	80
Imagen	70: Crear nuevo reporte con el asistente 1	09
Imagen	71: Seleccionar apariencia 1	09
Imagen	72: Fuentes de datos y consultas1	10
Imagen	73: Tipo de conexión. 1	10
Imagen	74: Pentaho Analysis Data Source1	11
Imagen	75: Consultas	11
Imagen	76: Detalle final de la conexión1	12
Imagen	77: Previsualización de la consulta1	12
Imagen	78: Conexión creada en el asistente 1	13
Imagen	79: Selector de campos1	13
Imagen	80: Previsualización del informe1	14
Imagen	81: Ajuste de formato del informe 1	15
Imagen	82: Segunda previsualización del informe 1	15
Imagen	83: Edición desde el área del trabajo 1	16
Imagen	84: Tercera Previsualización 1	16
Imagen	85: Insertar gráfico1	17
Imagen	86: Otra previsualización 1	17
Imagen	87: Añadir consulta 1	18
Imagen	88: Añadir parámetro1	18
Imagen	89: Editar parámetro1	19
Imagen	90: Modificar consulta anterior 1	19
Imagen	91: Se ha modificado la cabecera 1	20
Imagen	92: Informe final Vista del año 2018 1	20
Imagen	93: Informe final Vista del año 2019 1	21
Imagen	94: Publicar el informe 1	21
Imagen	95: Detalles de la conexión 1	22
Imagen	96: Datos de la publicación 1	22
Imagen	97: Confirmación reporte publicado1	22

Lista de tablas

Tabla 1: Planificación inicial	5
Tabla 2: Tabla users (base de datos operativa Nutribox)	9
Tabla 3: Tabla collaborators (base de datos operativa Nutribox)	9
Tabla 4: Tabla products (base de datos operativa Nutribox)	0
Tabla 5: Tabla coupons (base de datos operativa Nutribox)	0
Tabla 6: Tabla couponUse (base de datos operativa Nutribox)10	0
Tabla 7: KPI 1 Popularidad del producto1	1
Tabla 8: KPI 2 Tendencia Producto	2
Tabla 9: KPI 3 Porcentaie de canies por categoría de productos. 13	3
Tabla 10 [°] KP I4 Popularidad del Colaborador	3
Tabla 11:KPI 5 Tendencia Colaborador	4
Tabla 12 ⁻ KPI 6 Tendencia Canies anuales	4
Tabla 13: KPI 7 Total canjes	5
Tabla 14: KPI 8 Usuarios registrados	5
Tabla 15: KPI 9 Usuarios activos	6
Tabla 16: KPI 10 Tendencia Lisuarios activos	7
Tabla 17:KP 11 Tondoncia do clientos del colaborador	7
Tabla 17. KP 11 Tendencia de clientes del colaborador.	γ Q
Table 10: KPI 12 Activided del cliente	0
Table 20: KDI 14 Llouerice actives per caleborader	0
Tabla 20. KPT 14 Usuallos activos por colaborador.	9
Tabla 21: KPI 15 Popularidad del producto de colaborador	0
Tabla 22: KPI 16 Tendencia de canjes de producto por colaborador	0
Tabla 23: KPI 17 Tendencia Canjes anuales por colaborador2	1
Tabla 24: KPI 18 Total canjes anuales por colaborador	1
Tabla 25: Diseño CM1 Componente1 "Número de usuarios registrados"	8
Tabla 26:Diseño CM1 Componente 2 "Número de usuarios activos"	8
Tabla 27: Diseño CM1 Componente 3 "Actividad mensual de los usuarios"	8
Tabla 28: Diseño CM1 Componente 4: "Top Colaboradores"	9
Tabla 29: Diseño CM1 Componente 5 "Por categoría"	9
Tabla 30: Diseño CM1 Componente "Top Productos"	0
Tabla 31: Diseño CM1 Componente 7 "Canjes anuales" 30	0
Tabla 32: Diseño CM1 Componente 8 "Número de canjes realizados por mes"	1
Tabla 33: Diseño CM2 Componente 1 "Clientes en el ejercicio".	2
Tabla 34: Diseño CM2 Componente 2 "Actividad mensual de nuestros clientes" 3	2
Tabla 35: Diseño CM2 Componente 3: "Top Clientes"	3
Tabla 36: Diseño CM2 Componente "Top Productos"	3
Tabla 37: Diseño CM2 Componente 5 "Canjes anuales"	4
Tabla 38:Diseño CM2 Componente 6 "Número de canjes realizados por mes" 34	4
Tabla 39: Descripción Informe Canjes por categoría	5
Tabla 40: Informe Top colaboradores del año	6
Tabla 41: Informe Productos canjeados del año	7
Tabla 42: Procesos ETL generados con Pentaho Data Integration44	6
Tabla 43: CM1 Componente 1 "Selector de año"5	5
Tabla 44: Implantación CM1 Componente 2 "Actividad mensual de usuarios"	6
Tabla 45: Implantación CM1 Componente 3 "Usuarios totales"	6
Tabla 46: Implantación CM1 Componente 4 "Usuarios activos"	7
Tabla 47: Implantación CM1 Componente 5 "Tabla empresas colaboradoras"	8
Tabla 48: Implantación CM1 Componente 6 "Canjes por categoría"	9
Tabla 49: Implantación CM1 Componente 7 "Tabla de productos"	0
Tabla 50: Implantación CM1 Componente 8 "Canjeados al año"	1
Tabla 51: Implantación CM1 Componente 9 "Canjes por mes"	1
Tabla 52: Implantación CM2 Componente 1 "Selector de año". 6	4
Tabla 53: Implantación CM2 Componente 2 "Selector de empresa colaboradora"6	4

Tabla 54: Implantación CM2 Componente 3 "Clientes en el ejercicio"	65
Tabla 55: Implantación CM2 Componente 4 "Actividad mensual de los clientes	s"66
Tabla 56: Implantación CM2 Componente 5 "Top Clientes"	67
Tabla 57: Implantación CM2 Componente 6 "Top productos"	68
Tabla 58: Implantación CM2 Componente 7 "Canjes anuales"	68
Tabla 59: Implantación CM2 Componente 8 "Número de canjes realizados al r	mes."69
Tabla 60: Informe de canjes por categoría y año	71
Tabla 61: Informe Top 10 colaboradores del año	72
Tabla 62: Informe Productos canjeados del año	74

1. Introducción

1.1. Contexto y justificación del Trabajo

Nutribox es una empresa creada para gestionar la actividad de una plataforma (aplicación móvil y web service) basada en "healthy discounts". Esta empresa, trata de conseguir acuerdos con mercados locales (para promover la compra en estos pequeños comercios), centrándose en el descuento en comida saludable.

De esta forma las empresas colaboradoras se aprovecharán de este servicio para ofrecer sus productos a un mayor número de clientes y obtener una mayor visibilidad en internet, y por otro lado los usuarios tendrán a su disposición de una gran cantidad de ofertas en un único portal.

Esta organización surge a partir del programa piloto "TFG emprendedor interestudios", por medio del cual, se propone de crear una iniciativa emprendedora, en base a los trabajos de fin de grado de los estudiantes de diferentes áreas, en este caso dos estudiantes del grado en administración y dirección de empresas y un estudiante del grado de ingeniería informática.

La participación en este programa "TFG emprendedor interestudios" supone una mayor complejidad a este proyecto debido a que no sólo se deberá tratar el área del trabajo de finde grado, sino que será necesario un proceso de definición empresarial y un trabajo conjunto del equipo de emprendedor para desarrollar esta iniciativa.

Del mismo modo, hay que tener en cuenta que cada integrante del grupo deberá llevar a cabo un trabajo individual sobre su área. Para el caso de este trabajo sobre Business Intelligence al tratarse de una iniciativa a desarrollar no se tienen referencias sobre posibles necesidades, oportunidades u objetivos a valorar en este trabajo.

Así pues, con la definición de servicio o plataforma, se ha planteado la necesidad de integrar soluciones Business Intelligence (BI), para tratar la información generada por la propia plataforma a través de su funcionamiento y de otros datos externos (como podrían ser los extraídos de redes sociales, IoT, Open Data...), con la finalidad de extraer y compartir el conocimiento que pueda aportar valor a la solución.

Este TFG pretende cubrir esta necesidad, diseñando e implantando un prototipo inicial de BI para esta empresa; que permita dotar a la dirección ejecutiva de información relativa a la consecución de objetivos establecidos inicialmente por medio de cuadros de mando, informes y/o

reports que permitirán valorar la evaluación del negocio y la toma de decisiones en base a la información obtenida.

Con el objetivo de obtener cuadros de mando e informes que representen esta información, este prototipo inicial necesita la definición y producción de procesos de ETL, sistema de almacenamiento y gestión de datos (data warehouse), y un cubo OLAP que permita realizar valoraciones a los indicadores inicialmente establecidos.

Estos indicadores o KPI (del inglés, key performance indicator) se definen durante la realización de este trabajo, ya que actualmente, la plataforma Nutribox se encuentra en una fase idea, sin ningún prototipo desarrollado.

Por esta razón, en este TFG, se deberá considerar qué datos dispondrá la futura plataforma y cómo se tratarán, estudiando posibles técnicas y herramientas, con la finalidad de obtener una solución adecuada; y, por lo tanto; la definición de los KPI dependerá del modelado de datos inicial de la plataforma.

En consecuencia, el conjunto de indicadores se incrementará en posteriores ampliaciones del sistema, en base al análisis de las nuevas dimensiones del negocio, fuentes de datos, objetivos corporativos, etc; ya que este prototipo se pretende que sirva, como base para la solución BI de la empresa y su plataforma; que irá evolucionando, ampliando sus características y funcionalidades, en base a las necesidades del negocio.

Teniendo en cuenta que la plataforma de la empresa Nutribox se encuentra en un estado inicial o de definición empresarial, para la realización del prototipo se deberá generar un dataset ficticio con los que mostrar la funcionalidad de este prototipo.

Debido a la falta de financiación inicial del proyecto, se establece, el uso de soluciones basadas en Open Source necesarias para la implantación de un modelo conceptual, con la finalidad de ahorrar en el coste de licencias. Para futuras actualizaciones se podrá valorar el uso de otras soluciones, en base a nuevas necesidades o carencias que detectadas en el sistema.

1.2. Objetivos del Trabajo

Se han establecido los principales objetivos de este TFG:

- Valorar las posibles fuentes de datos del servicio/plataforma. En este trabajo se describe la base de datos operacional de Nutribox sobre la que se construye el prototipo inicial y posteriormente en el apartado de líneas de futuro se propondrán nuevas fuentes de datos, que podrá utilizar la solución BI de la organización.
- Definir KPI iniciales, que servirán para evaluar los resultados de la empresa. En esta fase inicial del servicio estos indicadores estarán destinados principalmente a la medición de aspectos como podrían ser el número de usuarios, popularidad de los diferentes productos ofrecidos desde la plataforma o diferentes tendencias.
- Diseñar e implantar un prototipo inicial de BI para esta plataforma, cuya finalidad será obtener reports/informes y cuadro de mando; con los que valorar los resultados de la empresa a partir de los KPI definidos en esta fase inicial del prototipo.
- Exponer las posibles mejoras, ampliaciones y nuevas funcionalidades a implementar en el prototipo. En el apartado de Líneas de futuro se expondrán posibles opciones como el uso de nuevas fuentes de datos o herramientas analíticas que mejoren y amplíen las capacidades del sistema.

Para cumplir los objetivos establecidos se definen los siguientes aspectos en el alcance del proyecto:

 Modelado de datos: Puesto que, como se ha explicado en la descripción del proyecto, la plataforma se encuentra en la fase de idea, sin siquiera, un prototipo, se hace necesario realizar un modelado de datos.

Para el prototipo inicial, que se tratará en este TFG, sólo se considerará la base de datos del servicio o plataforma; no se estudiarán otras fuentes de datos que se podrán plantear para futuras evoluciones de la solución BI de la organización.

 El proyecto incluye la definición y construcción de los sistemas de extracción, transformación, carga (ETL), construcción de un sistema de almacenamiento y gestión de datos (data warehouse), diseño de cubos OLAP, así como el diseño y construcción de cuadros de mando e informes que presenten los indicadores (KPI) que se definen en fases anteriores.

1.3. Planificación del Trabajo

En este proyecto se pueden describir los siguientes hitos:

- Definición y planificación del proyecto.
- Análisis y diseño conceptual.
 - Selección de técnicas y herramientas a utilizar para la creación de la solución.
 - o Definición de fuentes y tipos de datos a tratar.
 - Definición de KPI.
 - Diseño de visualizaciones.
- Construcción del prototipo.
 - Creación de dataset de pruebas.
 - Construcción de los sistemas de extracción, transformación, carga (ETL).
 - Construcción de un sistema de almacenamiento y gestión de datos (data warehouse).
 - Diseño y creación de cubos OLAP (online analytical processing).
 - Creación de prototipo de cuadros de mando e informes.
- Cierre del proyecto
 - Propuesta de posibles evoluciones mejoras, ampliaciones y nuevas funcionalidades a implementar en el prototipo.
 - Redactar de informe de cierre.
 - Finalizar entregables (memoria del proyecto, producto, video presentación, informe de autoevaluación...).

A partir de los hitos que se han definido anteriormente y teniendo en cuenta las fechas de las entregas de las diferentes PEC se establece la siguiente planificación para desarrollar este trabajo:

Hito	Duración	Comienzo	Fin
PEC1	16 días	lun 25/02/19	lun 18/03/19
Definición y planificación del proyecto.	16 días	lun 25/02/19	lun 18/03/19
PEC2	20 días	mar 19/03/19	lun 15/04/19
Selección de técnicas y herramientas a utilizar	2 días	mar 19/03/19	mié 20/03/19
Definición de fuentes y tipos de datos	5 días	jue 21/03/19	mié 27/03/19
Definición de KPI	3 días	jue 28/03/19	lun 01/04/19
Creación de dataset de pruebas	5 días	mar 02/04/19	lun 08/04/19
Construcción de ETL	5 días	mar 09/04/19	lun 15/04/19
PEC3	25 días	mar 16/04/19	lun 20/05/19
Construcción de datawharehouse	4 días	mar 16/04/19	vie 19/04/19
Diseño y creación cubos OLAP	6 días	sáb 20/04/19	vie 26/04/19
Diseño de visualizaciones	5 días	sáb 27/04/19	jue 02/05/19
Creación de prototipo cuadro de mando	8 días	vie 03/05/19	mar 14/05/19
Creación de reports	4 días	mié 15/05/19	lun 20/05/19
Entrega Final	27 días	mar 21/05/19	lun 17/06/19
Propuesta de mejoras, ampliaciones y nuevas funcionalidades a implementar	9 días	mar 21/05/19	jue 30/05/19
Finalizar entregables	17 días	jue 30/05/19	lun 17/06/19

Tabla 1: Planificación inicial

Dicha planificación se puede representar en el siguiente diagrama de Gantt:



Imagen 1: Diagrama de Gantt del proyecto

1.4. Breve sumario de productos obtenidos

Se han obtenido los siguientes productos:

- Archivo VirtualBox
 - Prototipo.ova
- Archivos con los script sql:
 - Nutribox.sql
 - insertarDatos.sql
 - creaDatawharehouse.sql
- Archivos de Pentaho Data Integration:
 - DimColaboradores.ktr
 - DimProductos.ktr
 - DimTiempo.ktr
 - DimUsuarios.ktr
 - HechosCanje.ktr
- Archivos de Pentaho Report Designer:
 - PorCat.prpt
 - PorProducto.prpt
 - TopColaborador.prpt

1.5. Breve descripción de los siguientes capítulos de la memoria

En esta memoria se detalla todo el proyecto y se estructura en las fases de análisis, diseño e implantación, así como los capítulos de líneas de futuro y conclusión.

2. Análisis

En esta fase de análisis, como la empresa Nutribox se encuentra en una fase inicial o de idea, se realiza un trabajo de definición de las fuentes de datos de la plataforma, así como los KPI o principales indicadores que se valorarán en este prototipo de Business Intelligence.

De esta forma, es necesario que la empresa cuestione cómo el tratamiento de los datos que dispone puede ser usado para alcanzar los objetivos y metas de la organización, es decir, como los datos existentes se convierten en información de utilidad para la toma de decisiones.

Del mismo modo la definición de los indicadores debe ir alineada con las necesidades y cumplimiento de objetivos de la empresa, de forma que estos puedan medir las diferentes acciones definidas y la evolución de la entidad.

2.1. Definición de fuentes y tipos de datos

Como se ha descrito en el alcance del proyecto, para este prototipo inicial sólo se considerará la base de datos del servicio o plataforma como fuente de datos. Esta base de datos la podemos representar con el siguiente diagrama:



ProductDescription : varchar(250)

ProductCategory : varchar(25)

Imagen 2: Estructura de la base de datos "Nutribox"

La tabla users almacena la información de los usuarios de la plataforma y se compone de los siguientes campos:

Tabla users	
UserID	Se trata del identificador de usuario que se generará de
	forma automática.
UserEmail	En este campo se registra la dirección de correo electrónico
	del usuario.
UserPassword	En este campo se registra la contraseña del usuario.
UserFirstName	En este campo se registra el nombre del usuario.
UserLastName	En este campo se registra el/los apellidos del usuario.
UserCity	En este campo se registra la ciudad del usuario.
UserState	En este campo se registra la provincia del usuario.
UserZip	En este campo se registra el código postal del usuario.
UserPhone	En este campo se registra el número de teléfono del usuario.
UserCountry	En este campo se registra el País del usuario.
UserAddress	En este campo se registra la dirección del usuario.

Tabla 2: Tabla users (base de datos operativa Nutribox)

La tabla colaborators almacena la información de las empresas colaboradoras de la plataforma y consta de los siguientes campos:

Tabla collaborators	
CollaboratorID	Se trata del identificador de colaborador que se generará de forma automática.
CollaboratorEmail	En este campo se registra el correo electrónico de la empresa colaboradora.
CollaboratorPasswor d	En este campo se registra la contraseña de la empresa colaboradora en la plataforma.
CollaboratorName	En este campo se registra el nombre de la empresa colaboradora.
CollaboratorCity	En este campo se registra la ciudad de la empresa colaboradora.
CollaboratorState	En este campo se registra la provincia de la empresa colaboradora.
CollaboratorZip	En este campo se registra el código postal de la empresa colaboradora.
CollaboratorPhone	En este campo se registra el número de teléfono de la empresa colaboradora.
CollaboratorCountry	En este campo se registra el país de la empresa colaboradora.
CollaboratorAddress	En este campo se registra la dirección de la empresa colaboradora.

Tabla 3: Tabla collaborators (base de datos operativa Nutribox)

La tabla products almacena la información de los productos que las empresas colaboradoras promocionan a través de la plataforma y consta de los siguientes campos:

Tabla products	
ProductID	Identificador de producto que se generará de forma automática
ProductName	En este campo se almacena el nombre del producto
TTOUUCINAITIE	En este campo se amacena el nombre del producio.
ProductPrice	En este campo se almacena el precio unitario del
	producto.
ProductDescription	En este campo se registra una breve descripción del
	producto.
ProductCategory	En este campo se registra la categoría del producto.

Tabla 4: Tabla products (base de datos operativa Nutribox)

La tabla coupons almacena la información acerca de los cupones/ofertas que se ofrecen en la plataforma y consta de los siguientes campos:

Tabla coupons			
CouponID	Identificador del cupón que se generará de forma automática.		
CollaboratorID	Identificador del colaborador que registra el cupón.		
ProductID	Identificador del producto que se oferta.		
Discount	Descuento que se ofrece en este cupón.		
BeginDate	Se registra la fecha en la que empieza la oferta.		
FinishDate	Se registra la fecha en la que finaliza la oferta.		

Tabla 5: Tabla coupons (base de datos operativa Nutribox)

La tabla couponUse almacena los datos de las diferentes transacciones o canjes de cupones y consta de los siguientes campos:

Tabla couponUse	
TransactionID	En este campo se almacena el identificador de la
	transaccion.
CouponID	En este campo se almacenará el id del cupón a canjear.
UserID	En este campo se registra el identificador de usuario que utiliza el cupón
UseDate	En este campo se registra la fecha en la que se ha utilizado el cupón.

Tabla 6: Tabla couponUse (base de datos operativa Nutribox)

2.2. Definición de KPI

Un KPI (key performance indicator), conocido también como indicador clave, medidor de desempeño o indicador clave de rendimiento, es una medida del nivel del rendimiento de un proceso. El valor del indicador está directamente relacionado con un objetivo fijado previamente.¹

Para la consecución de este proyecto, se definen KPI que permitan valorar el rendimiento de la plataforma Nutribox, relativos a diferentes aspectos como usuarios, empresas colaboradoras, productos o canjes de cupones. Estos indicadores deberán ser actualizados de acuerdo a las necesidades y estrategia de la empresa, así como al desarrollo de la plataforma BI. Inicialmente se definen los siguientes indicadores:

Nombre	KPI 1 Popularidad del producto
Finalidad	Mide la popularidad de un producto de acuerdo al número de canjes de cupones de un producto determinado.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	№ de canjes de un producto por año.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	20/100. Se ha establecido que, en la fase inicial de la plataforma, cada producto ofertado en la plataforma debería tener entre 20 y 100 canies anuales
Iniciativas	-Crear campañas específicas para publicitar este producto. -Búsqueda de colaboradores que ofrezcan este tipo de producto.

Tabla 7: KPI 1 Popularidad del producto

¹ Indicador clave de rendimiento (2019) [En línea] de Wikipedia

[[]Consulta: 1 de Junio de 2019] <https://es.wikipedia.org/wiki/Indicador_clave_de_rendimiento>

Nombre	KPI 2 Tendencia Producto
Finalidad	Conocer la tendencia de los canjes de cupones de un producto determinado con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	(Canjes del producto/año)-(Canjes del producto/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	 -Crear campañas específicas para publicitar este producto. -Búsqueda de colaboradores que ofrezcan este tipo de producto.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 8: KPI 2 Tendencia Producto

Nombre	KPI 3 Porcentaje de canjes por Categoría de productos
Finalidad	Conocer el porcentaje canjes de cupones de una categoría determinada de productos con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	(Canjes categoría/canjes totales) *100
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual / Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0

Iniciativas	-Crear campañas específicas para publicitar esta categoría. -Búsqueda de colaboradores que ofrezcan este tipo de productos.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 9: KPI 3 Porcentaje de canjes por categoría de productos.

Nombre	KPI 4 Popularidad del Colaborador
Finalidad	Mide la popularidad de un colaborador de acuerdo al número de canjes de cupones ofertados.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	№ de canjes de cupones publicados por el colaborador en el año.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	200/20. Se ha establecido que, en la fase inicial de la plataforma, cada colaborador consiga al menos 20 canjes de sus productos para garantizar su continuidad en la plataforma. Se plantea un objetivo inicial de 200 canjes anuales.
Iniciativas	-Crear campañas específicas para publicitar colaboradores que no han cumplido el objetivo mínimo de 20 canjes.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 10: KP I4 Popularidad del Colaborador

Nombre	KPI 5 Tendencia Colaborador
Finalidad	Conocer la tendencia de los canjes de cupones por colaborador con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	(Canjes categoría/año)-(Canjes categoría/año anterior)

Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	 -Crear campañas específicas para publicitar a cada colaborador. -Ofrecer beneficios específicos a los colaboradores que cumplen objetivos. -Establecimiento de unos requisitos mínimos para los colaboradores de la plataforma.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 11:KPI 5 Tendencia Colaborador

Nombre	KPI 6 Tendencia Canjes anuales
Finalidad	Conocer la tendencia del total de canjes efectuados en la plataforma con respecto al ejercicio anterior (si se han producido más, menos o las mismas transacciones) y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	(Canjes categoría/año)-(Canjes categoría/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Promover los productos menos vendidos por medio de promociones específicas.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 12: KPI 6 Tendencia Canjes anuales.

Nombre	KPI 7 Total canjes
Finalidad	Conocer el total de canjes efectuados en la plataforma en un ejercicio para valorar el volumen de negocio en ese periodo.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	Contar todos los canjes realizados en el periodo a valorar.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual / Mensual / Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Promover los productos menos vendidos por medio de promociones específicas.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 13: KPI 7 Total canjes.

Nombre	KPI 8 Usuarios registrados
Finalidad	El número de usuarios registrados es un indicador que mide el volumen de negocio de la compañía.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de operaciones
Definición/fórmula	Count-distict(UserID)
Proceso relacionado	Registro de clientes, Satisfacción y fidelización de clientes, Atención al cliente.
Frecuencia	Anual
Quien lo mide	Dirección de operaciones
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	2000/ 1500 / 1000
Iniciativas	-Crear campañas de marketing para publicitar la plataforma. -Programa de afiliados.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 14: KPI 8 Usuarios registrados.

Nombre	KPI 9 Usuarios activos
Finalidad	Conocer el número de usuarios únicos activos en el año, este un indicador ayuda a medir el volumen de negocio de la compañía.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox, Departamento de marketing de Nutribox y departamento de operaciones.
Definición/fórmula	Cuenta usuarios que hayan canjeado algún cupón
Proceso relacionado	Registro de clientes, Satisfacción y fidelización de clientes, Atención al cliente.
Frecuencia	Anual / Mensual / Semanal
Quien lo mide	Departamento de operaciones
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	2000/ 1500 / 1000
Iniciativas	-Crear campañas de marketing para publicitar la plataforma. -Programa de afiliados -Ofrecer beneficios específicos a los usuarios con mayor actividad.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 15: KPI 9 Usuarios activos.

Nombre	KPI 10 Tendencia Usuarios activos
Finalidad	Conocer la tendencia de actividad de los usuarios de la plataforma y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva Nutribox / Departamento de marketing de Nutribox
Definición/fórmula	(usuariosActivos/año)-(usuariosActivos/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Crear campañas de marketing para publicitar la plataforma.

	 Programa de afiliados Ofrecer beneficios específicos a los usuarios más activos.
Quien lo recibe	Dirección de Marketing, Dirección de finanzas, Dirección ejecutiva de Nutribox.

Tabla 16: KPI 10 Tendencia Usuarios activos

Nombre	KPI 11 Tendencia de clientes del Colaborador
Finalidad	Conocer la tendencia del número de clientes por colaborador con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Operativo
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	(Clientes/año)-(Clientes/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	 -Crear campañas específicas para publicitar a cada colaborador. -Ofrecer beneficios específicos a los colaboradores que cumplen objetivos. -Establecimiento de unos requisitos mínimos para los colaboradores de la plataforma.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 17:KP 11 Tendencia de clientes del colaborador.

Nombre	KPI 12 Actividad del cliente.
Finalidad	Mide la actividad del cliente de acuerdo al número de canjes de cupones ofertados por un colaborador determinado.
Categoría	Operativo
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	Nº de canjes de productos ofertado por un colaborador.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	2/0.
Iniciativas	-Crear campañas para perfiles específicos de clientes.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 18: KPI 12 Actividad del cliente.

Nombre	KPI 13 Tendencia en la actividad del cliente.
Finalidad	Conocer la tendencia de los canjes de cupones ofrecidos por un colaborador, realizados de un cliente determinado con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	(Canjes del cliente/año)-(Canjes del cliente/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Crear campañas para perfiles específicos de clientes.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 19: KPI 13 Tendencia en la actividad del cliente.

Nombre	KPI 14 Usuarios activos por colaborador
Finalidad	Conocer el número de usuarios únicos activos que realizan canjes de los productos del colaborador en un mes concreto, este un indicador ayuda a medir el volumen de negocio del colaborador.
Categoría	Operativo
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	Cuenta usuarios que hayan canjeado algún cupón ofertado por la empresa colaboradora.
Proceso relacionado	Registro de clientes, Satisfacción y fidelización de clientes, Atención al cliente.
Frecuencia	Anual / Mensual/ Semanal
Quien lo mide	Departamento de operaciones
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	15/ 3
Iniciativas	-Crear campañas de marketing para publicitar la plataforma. -Programa de afiliados -Ofrecer beneficios específicos a los usuarios con mayor actividad.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 20: KPI 14 Usuarios activos por colaborador.

Nombre	KPI 15 Popularidad del producto de colaborador X
Finalidad	Mide la popularidad de un producto ofertado por un colaborador de acuerdo al número de canjes de cupones de un producto determinado ofrecidos por este colaborador.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	Nº de canjes de un producto ofrecido por un colaborador determinado al año.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual /Mensual /Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.

Meta/Umbral	20/100. Se ha establecido que, en la fase inicial de la plataforma, cada producto ofertado en la plataforma debería tener entre 20 y 100 canjes anuales.
Iniciativas	-Crear campañas específicas para publicitar este producto.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 21: KPI 15 Popularidad del producto de colaborador

Nombre	KPI 16 Tendencia de canjes de producto por colaborador
Finalidad	Conocer la tendencia de los canjes de cupones de un producto determinado ofrecidos por un colaborador con respecto al ejercicio anterior y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	(Canjes del producto/año)-(Canjes del producto/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes.
Frecuencia	Anual/ Mensual/ Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Crear campañas específicas para publicitar este producto.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 22: KPI 16 Tendencia de canjes de producto por colaborador

Nombre	KPI 17 Tendencia Canjes anuales por colaborador
Finalidad	Conocer la tendencia del total de canjes efectuados en la plataforma sobre productos ofrecidos por el colaborador con respecto al ejercicio anterior (si se han producido más, menos o las mismas transacciones) y actuar en consecuencia.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	(Canjes categoría/año)-(Canjes categoría/año anterior)
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores

Frecuencia	Anual
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	 Promover los productos menos vendidos por medio de promociones específicas.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 23: KPI 17 Tendencia Canjes anuales por colaborador

Nombre	KPI 18 Total canjes por colaborador
Finalidad	Conocer el total de canjes efectuados en la plataforma sobre las ofertas de un colaborador en un ejercicio para valorar el volumen de negocio del colaborador en ese periodo.
Categoría	Estratégico
Responsabilidad	Dirección ejecutiva empresa colaboradora.
Definición/fórmula	Contar todos los canjes realizados en el año de productos de este colaborador.
Proceso relacionado	Promoción de productos, Satisfacción y fidelización de clientes, Atención a colaboradores/proveedores
Frecuencia	Anual/Mensual/Semanal
Quien lo mide	Dirección de Marketing
Fuente de los datos	Base de datos operacional de Nutribox.
Meta/Umbral	<0/0
Iniciativas	-Promover los productos menos vendidos por medio de promociones específicas.
Quien lo recibe	Dirección ejecutiva empresa colaboradora.

Tabla 24: KPI 18 Total canjes anuales por colaborador.

3. Diseño

En esta fase del proyecto se definen las características de la solución a realizar. Para ello es necesario determinar los datos que usará el prototipo en base al análisis realizado en la fase anterior, describir los diferentes componentes de la solución y seleccionar el software necesario para construir el producto final.

3.1. Base de datos operativa Nutribox y datos de prueba.

Como ya se ha indicado en la fase de análisis la empresa Nutribox no está operativa actualmente por lo que es necesario generar unos datos iniciales o de prueba, que permitan mostrar la utilidad del prototipo.

Con este fin en el prototipo se crea una base de datos llamada Nutribox que simulará la base de datos operativa de la plataforma y teniendo en cuenta la estructura de la base de datos descrita en la fase de análisis, se generarán:

- 20 registros en la tabla de empresas colaboradoras.
- 20 registros en la tabla de productos
- 200 registros de cupones
- 1000 registros en la tabla usuarios
- 5000 registros de transacciones o canjes de cupones en la base de datos Nutribox.

3.2. Componentes de la solución.

Para obtener los resultados previstos es necesario diseñar e implantar una infraestructura capaz de extraer los datos de la base de datos operacional de Nutribox y que tras su procesamiento permita obtener unos reportes y cuadros de mando o dashboards, capaces de realizar un seguimiento de los principales KPI. Esta infraestructura se describe en las siguientes secciones de esta memoria y la podemos representar en el siguiente diagrama:



Imagen 3: Diseño de la infraestructura BI Nutribox.

3.2.1. Data wharehouse

Un data wharehouse, es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se usa para realizar informes (reports) y análisis de datos, siendo un componente fundamental de la inteligencia empresarial.² Los data wharehouse se estructuran principalmente siguiendo los esquemas en estrella o en copo de nieve.

Para esta solución se ha escogido el esquema en estrella³. Este esquema separa los datos del proceso de negocios en: hechos y dimensiones. Los hechos contienen datos medibles, cuantitativos, relacionados a la transacción del negocio, y las dimensiones son atributos que describen los datos indicados en los hechos.

Una razón para para utilizar los esquemas en estrella es su simplicidad desde el punto de vista del usuario final. Las consultas no son complicadas, ya que las condiciones y las uniones necesarias sólo involucran a la tabla de hechos y a las de dimensiones, no haciendo falta que se encadenen uniones y condiciones a dos o más niveles como ocurriría en un esquema en copo de nieve.

Además, es la opción con mejor rendimiento y velocidad pues permite indexar las dimensiones de forma individualizada sin que repercuta en el rendimiento de la base de datos en su conjunto.

Este datawharehouse se estructurará en una tabla de hechos HechosCanje y las tablas de dimensiones DimProductos, DimUsuarios, DimColaboradores y DimFecha. Como se puede ver representada en el siguiente diagrama:

²_Almacén de datos (2019) [En línea] de Wikipedia

[[]Consulta: 1 de Junio de 2019] <https://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos>

³ Esquema en estrella (2019) [En línea] de Wikipedia

[[]Consulta: 1 de Junio de 2019] < https://es.wikipedia.org/wiki/Esquema_en_estrella>

🗸 👩 datawharehouse DimColaboradores

- g collaboratorSku : bigint(20)
- # version : int(11)
- date_from : datetime
- date_to : datetime
- # CollaboratorID : int(11)
- CollaboratorEmail : text
- CollaboratorName : varchar(50)
- CollaboratorCity : varchar(90)
- CollaboratorState : varchar(20)
- CollaboratorZip : varchar(12)
- CollaboratorPhone : varchar(20)
- CollaboratorCountry : varchar(20)
- CollaboratorAddress : varchar(50)

🗸 👩 datawharehouse DimTiempo

- fechalnicial : datetime
- # suma1Dia : int(11)
- fecha_id : datetime
- fechaTexto : tinytext
- # diames : int(11)
- # mes : int(11)
- # año : int(11)
- # diaSemana : int(11)
- @ fechasku : int(11)
- diaText : varchar(9)
- mesText : varchar(10)





🗸 👌 datawharehouse DimUsuarios UserSku : bigint(20) # version : int(11) date_from : datetime date_to : datetime # UserID : int(11) UserEmail : text UserPassword : text primerNombre : tinytext segundoNombre : tinytext primerApellido : tinytext segundoApellido : tinytext UserCity : varchar(90) UserState : varchar(50) UserZip : varchar(12) UserPhone : varchar(20) UserCountry : varchar(20) UserAddress : varchar(50)

3.2.2. Procesos ETL

Una vez definida la estructura del data wharehouse, serán necesarios definir los procesos ETL que permitan extraer la información a partir de la base de datos operativa, volcándolos en el data wharehouse de la Compañía, tras realizar las operaciones de formato y limpieza que sean necesarios. Para ello, será necesario implementar los siguientes trabajos:

- A partir de los datos incluidos en las tablas CouponUse y coupons de la base de datos operativa Nutribox poblar la tabla HechosCange, tras haber realizado el tratamiento necesario de los datos.
- A partir de los datos incluidos en la tabla products de la base de datos operativa Nutribox poblar la tabla DimProducts, tras haber realizado el tratamiento necesario de los datos.
- A partir de los datos incluidos en la tabla users de la base de datos operativa Nutribox poblar la tabla DimUsuarios, tras haber realizado el tratamiento necesario de los datos.
- A partir de los datos incluidos en la tabla collaborators de la base de datos operativa Nutribox poblar la tabla DimColaboradores, tras haber realizado el tratamiento necesario de los datos.
- Genere los datos necesarios para crear la dimensión tiempo y su carga en la tabla DimTiempo del data wharehouse.

3.2.3. Cubo OLAP

OLAP es el acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Su objetivo es agilizar la consulta de grandes cantidades de datos. Para ello utiliza estructuras de datos diversas, normalmente multidimensionales (o Cubos OLAP), que contienen datos resumidos de grandes Bases de datos o Sistemas Transaccionales (OLTP).⁴

En este caso se podrán definir diferentes cubos incluyendo las dimensiones colaboradores, usuarios, tiempo y productos, y diferentes métricas que ayuden a valorar los diferentes indicadores que se han establecido.

⁴ OLAP (2019) [En línea] de Wikipedia

[[]Consulta: 1 de Junio de 2019] <https://es.wikipedia.org/wiki/OLAP>

Por ejemplo, se podría generar un cubo con las dimensiones tiempo, usuarios y producto de forma que cada celda contendrá los canjes de ofertas del productoX realizados por el UsuarioX en un mes indicado. Este cubo se representa en el siguiente diagrama:



Imagen 5: Ejemplo de Cubo OLAP

Para el diseño del prototipo inicial se plantea la definición de un único cubo OLAP con las dimensiones colaboradores, usuarios, tiempo y productos, para valorar las métricas:

- cuponesCanjeados: Cuenta todos los identificadores de canje (transactionSkU) distintos de una selección.
- Número de colaboradores: Cuenta todos los identificadores de colaborador (CollaboratorID) distintos de una selección.
- Número de usuarios: Cuenta todos los identificadores de usuario (UserID) distintos de una selección.
- Número de productos: Cuenta todos los identificadores de producto (ProductID) distintos de una selección.
- cuponesUnicos: Cuenta todos los identificadores de cupón (CouponID) distintos de una selección.

3.2.4. Diseño de dashboard o cuadro de mandos.

Con la finalidad de representar gráficamente en valor de los diferentes indicadores que se han establecido en la fase de análisis, se diseñan diferentes cuadros de mando que permiten su visualización de una forma rápida, sencilla y accesible.

Para la realización de los bocetos se ha usado la herramienta online Balsamiq Cloud disponible a través de la URL: <u>https://balsamiq.cloud/</u>.

3.2.4.1. CM1: Cuadro de mando Administrativo Nutribox

Este cuadro de mando se ha diseñado para que la gerencia de la entidad pueda obtener una representación accesible y simplificada de los principales indicadores que afectan a la evolución anual del proceso de canjes de cupones en la plataforma:



Cuadro de mando para la dirección ejecutiva de Nutribox



A continuación, se detalla cada elemento:

Diseño CM1 Componente 1: Número de usuarios registrados		
Número de usuarios registrados: 1000		
Descripción: Muestra el número total de usuarios registrado		
en la plataforma.		
Indicadores relacionados:	KPI 9 Usuarios registrados	

Tabla 25: Diseño CM1 Componente1 "Número de usuarios registrados".

Diseño CM1 Componente 2: Número de usuarios activos		
Número de usuarios activos: 800 up		
Descripción:	Muestra el número total de usuarios activos (que han realizado algún canje de cupón en la plataforma) en el año indicado, así como su tendencia con respecto al ejercicio anterior.	
Indicadores relacionados:	KPI 10 Usuarios activos al año. KPI 12 Tendencia Usuarios activos.	

Tabla 26:Diseño CM1 Componente 2 "Número de usuarios activos".



Tabla 27: Diseño CM1 Componente 3 "Actividad mensual de los usuarios".

Diseño CM1 Componente 4: Top Colaboradores					
		Top Colabo	oradores	1	
	Nombre de	l colaborador	Canjes	Tendencia	1
	Colaborador1		80	up	
	Colaborador1		80	up	
	Colaborador1		80	up	
	Colaborador1 Colaborador1		80	up	
			80	up	
Descripción: Muestra e que más base al n ofertas. M año actua		n forma éxito ha úmero c uestra a , su tenc	to de tabla an tenido de cupone además de dencia resp	los 5 colaboradores en la plataforma en s canjeados de sus el valor numérico del pecto al año anterior.	
Indicadores relacionados: KPI 4 Pop KPI 5 Ten		ularidad dencia (l del Colab Colaborado	orador or.	

Tabla 28: Diseño CM1 Componente 4: "Top Colaboradores".



Tabla 29: Diseño CM1 Componente 5 "Por categoría".

Diseño CM1 Componente 6	: Top Produ	ictos.		
	Top Produ	uctos		
Nombre	del producto	Canjes	Tendencia	1
Product	Producto1		up	
Product	Producto2		up	
Product	Producto3		up	
Product	Producto4		up	
Product	05	50	up]
Descripción:	Muestra e más éxito número d Muestra a actual, su	n forma han ter e cupoi además tenden	to de tabla ido en la p nes canjea del valor cia respect	los 5 productos que lataforma en base al dos de sus ofertas. numérico del año o al año anterior.
Indicadores relacionados: Popularid		ad del p	roducto, Te	endencia Producto.

Tabla 30: Diseño CM1 Componente "Top Productos".

Diseño CM1 Componente 7	Diseño CM1 Componente 7: Canjes anuales		
Canjes anuales: 1000 down			
Descripción:	Muestra el número de canjes de cupones total a través de la plataforma en el año a analizar, así		
	como su tendencia con respecto al ejercicio		
Indicadores relacionados:	KPI 7 Total canies anuales		
	KPI 6 Tendencia Canjes anuales		

Tabla 31: Diseño CM1 Componente 7 "Canjes anuales"

Diseño CM1 Componente 8:	Número de canjes realizados por mes	
Número de canjes realizados por mes		
Descripción:	Muestra el número total de canjes por mes, del	
Indicadores relacionados:	KPI 8 Total canjes mes.	



3.2.4.2. CM2: Cuadro de mando para empresa colaboradora

Este cuadro de mando se ha diseñado para que cada empresa colaboradora de la compañía pueda obtener una representación accesible y simplificada de los principales indicadores que afectan a la evolución anual del proceso de canjes de cupones de los productos que han ofertado en la plataforma:



Cuadro de mando para empresa colaboradora año 2019

Imagen 7: Diseño cuadro de mando para empresas colaboradoras.

A continuación, se detalla cada elemento:

Diseño CM2 Componente 1: Clientes en el ejercicio		
Clientes en el ejercicio: 800 up		
Descripción:	Muestra el número total de usuarios que han realizado algún canje de cupones del colaborador, en el año indicado, así como su tendencia con respecto al ejercicio anterior.	
Indicadores relacionados:	KPI 13 Tendencia clientes del colaborador.	

Tabla 33: Diseño CM2 Componente 1 "Clientes en el ejercicio".

Diseño CM2 Componente 2: Ac	tividad mensual de nuestros clientes.
A	actividad mensual de nuestros clientes:
Descripción:	Muestra el número total de usuarios activos que han realizado algún canje de cupón ofertado por este colaborador en la plataforma, por mes.
Indicadores relacionados:	KPI 16 Usuarios activos por colaborador y mes

Tabla 34: Diseño CM2 Componente 2 "Actividad mensual de nuestros clientes".

Тор	Clientes	
Nombre del cliente	Canjes	Tendencia
Cliente1	80	up
Cliente2	80	up
Cliente3	80	up
Cliente4	80	up
Cliente5	80	up

Descripción:	Muestra en formato de tabla los 5 clientes que han realizado más canjes de cupones ofertados por este colaborador, que más éxito han tenido en la plataforma en base al número de cupones canjeados de sus ofertas. Muestra además del valor numérico del año actual, su tendencia respecto al año anterior.
Indicadores relacionados:	KPI 14: Actividad del cliente.
	KPI 15: Tendencia en la actividad del cliente.

Tabla 35: Diseño CM2 Componente 3: "Top Clientes".

Diseño CM2 Componente 4: Top Productos					
		Top Pro	ductos		
	Nombre del pro	ducto	Canjes	Tendencia	ר
	Producto1		50	up	
	Producto2		50	up	
	Producto3		50	up	
	Producto4		50	up	
	Producto5		50	up	
					-
Descripción:		Muestra en formato de tabla los 5 producto ofertados por este colaborador, que más éxit han tenido en la plataforma en base al númer de cupones canjeados de sus ofertas. Muestr además del valor numérico del año actual, s tendencia respecto al año anterior.		5 productos ue más éxito se al número rtas. Muestra ño actual, su	
Indicadores relacionados: X. KPI 17 Popularidad del producto de colal X. KPI 18 Tendencia de canjes de produ colaborador.		e colaborador producto por			

Tabla 36: Diseño CM2 Componente "Top Productos".

Diseño CM2 Componente 5: Canjes anuales		
Canjes anuales: 1000 down		
Descripción:	Muestra el número de canjes de cupones total que ha obtenido este colaborador en el año a analizar, así como su tendencia con respecto al ejercicio anterior.	
Indicadores relacionados:	KPI 19 Tendencia Canjes anuales por colaborador.KPI 20 Total canjes anuales por colaborador.	

Tabla 37: Diseño CM2 Componente 5 "Canjes anuales"



Tabla 38:Diseño CM2 Componente 6 "Número de canjes realizados por mes"

3.2.5. Reports

A partir de Pentaho Report Designer se crearán diferentes informes dinámicos que servirán para representar la información obtenida a partir de los datos de la organización. Estos informes podrán ser impresos o exportarse en diversos formatos.

Para la realización de los bocetos se ha usado la herramienta online Balsamiq Cloud disponible a través de la URL: <u>https://balsamiq.cloud/</u>. A continuación, se define el diseño de varios posibles informes generados en este trabajo:



3.2.5.1. informe de canjes por categoría y año

Tabla 39: Descripción Informe Canjes por categoría.

3.2.5.2. Informe top colaboradores del año



Tabla 40: Informe Top colaboradores del año.



3.2.5.3. Informe Productos canjeados del año

Tabla 41: Informe Productos canjeados del año.

3.3. Selección del software

En este apartado se justificará la elección de las diferentes herramientas y soluciones software que componen la arquitectura del prototipo en base al diseño planteado en el apartado anterior.

3.3.1. Selección del Software de virtualización

Para cumplir los objetivos propuestos, se creará una máquina virtual sobre la que se instalará todo el software necesario. Para este proyecto se ha seleccionado el software Virtualbox, debido principalmente a su facilidad de uso y experiencia previa con esta herramienta.

Aunque es propiedad de Oracle Corporation, es desarrollado bajo un modelo de código abierto, licenciado bajo GNU General Public License (GPL) Version 2 y su uso es totalmente gratuito, tanto para uso personal como professional⁵, por lo que cumple con los requisitos de este proyecto.

3.3.2. Selección del sistema operativo utilizado en el prototipo

Ubuntu es una distribución de Linux basada en la arquitectura de Debian y està orientada al usuario promedio, con un fuerte enfoque en la facilidad de uso y en mejorar la experiencia del usuario. Estadísticas web sugieren que la cuota de mercado de Ubuntu dentro de las distribuciones Linux es, aproximadamente, del 52% y con una tendencia a aumentar como servidor web.

Se ha seleccionado el sistema operativo Ubuntu 18.04 ya que es una de las distribuciones Linux más utilizadas en la actualidad, por lo que cuenta con una amplia comunidad, así como infinidad de documentación, manuales y otros recursos gratuitos que facilitan su uso.

Además, este sistema operativo es gratuito y de código abierto; cumpliendo los requisitos impuestos para este prototipo. Por otro lado, esta versión (la 18.04 LTS) cuenta con soporte extendido de 5 años, un aspecto a tener en cuenta a la hora de implantar una solución definitiva.

3.3.3. Selección del sistema gestor de bases de datos.

Para la creación de este prototipo es necesaria la instalación de un sistema gestor de bases de datos que servirá, por un lado, para simular la base de datos operacional de Nutribox que será tomada como fuente

⁵ Oracle VM VirtualBox (2019) [En línea] de ORACLE

[[]Consulta: 1 de Junio de 2019] <https://www.oracle.com/es/virtualization/virtualbox/>

de datos en este proyecto y por otro lado como almacén para el datawharehouse de la compañía.

Entre los sistemas gestores de bases de datos Open Source destacan MySQL y PostgreSQL. Ambos sistemas se consideran válidos para este proyecto, por lo que se realiza una comparativa de ambas soluciones, para poder realizar una correcta elección.

3.3.3.1. MySQL

MySQL es la base de datos de código abierto más popular del mercado. Gracias a su rendimiento probado, a su fiabilidad y a su facilidad de uso, MySQL se ha convertido en la base de datos líder elegida para las aplicaciones basadas en web, así como una elección muy popular como base de datos integrada.⁶

Ventajas:

• Velocidad a la hora de realizar las operaciones, lo que le hace uno de los gestores que ofrecen mayor rendimiento.

• Su bajo consumo lo hace apto para ser ejecutado en una máquina con escasos recursos sin ningún problema.

• Gran facilidad de configuración e instalación.

• Tiene una probabilidad muy reducida de corromper los datos, incluso en los casos en los que los errores no se produzcan en el propio gestor, sino en el sistema en el que está.

• El conjunto de aplicaciones Apache-PHP-MySQL es uno de los más utilizados en Internet.

Inconvenientes:

- Carece de soporte para transacciones, rollback's y subconsultas.
- No maneja la integridad referencial.
- Peor rendimiento con grandes bases de datos.⁷

3.3.3.2. PostgreSQL

PostgreSQL se ha enfocado más que nada a la fiabilidad e integridad de los datos, así como también a mejorar como es visto del lado del desarrollador. Actualmente cuenta con planificador de consultas extremadamente sofisticado, que nos permite unir tablas sin importar lo grande que sean de una forma sencilla y eficiente; algo muy importante para aquellos sistemas donde se trabaja con gran cantidad de datos.⁸

⁶ Oracle MySQL (2019) [En línea] de ORACLE

[[]Consulta: 1 de Junio de 2019] <https://www.oracle.com/es/mysql/>

⁷ PostGreSQL vs. MySQL (2016) [En línea] de Daniel Pecos Martínez [Consulta: 1 de Junio de 2019] https://danielpecos.com/documents/postgresql-vs-mysql/>

⁸ Infografía MySQL vs PostgreSQL (2016) [En línea] de infranetworking

[[]Consulta: 1 de Junio de 2019] <https://blog.infranetworking.com/infografia-mysql-vspostgresql/>

Ventajas:

• Posee una gran escalabilidad. Es capaz de ajustarse al número de CPUs y a la cantidad de memoria que posee el sistema de forma óptima, haciéndole capaz de soportar una mayor cantidad de peticiones simultáneas de manera correcta.

• Implementa el uso de rollback's, subconsultas y transacciones, haciendo su funcionamiento mucho más eficaz, y ofreciendo soluciones en campos en las que MySQL no podría.

• Tiene la capacidad de comprobar la integridad referencial, así como también la de almacenar procedimientos en la propia base de datos. Inconvenientes:

• Consume gran cantidad de recursos.

• Tiene un límite de 8K por fila, aunque se puede aumentar a 32K, con una disminución considerable del rendimiento.

Más lento que MySQL.9

3.3.3.3. Elección de sistema gestor de base de datos

Teniendo en cuenta las ventajas e inconvenientes de ambas alternativas, se escoge para este proyecto el uso de MySQL, debido principalmente a la menor cantidad de recursos que requiere y su mayor rendimiento; ya que la construcción de este prototipo sobre una máquina virtual con recursos limitados podría ser un factor limitante.

Además, su mayor popularidad y uso favorecen la obtención de manuales, tutoriales y otros recursos didácticos necesarios para poder obtener unos conocimientos adecuados para desarrollar el proyecto adecuadamente.

3.3.4. Selección de Suite Business Intelligence

Para cumplir con los objetivos previstos es necesario disponer de una solución que permita diseñar y ejecutar procesos ETL que, con la finalidad de extraer los datos de la base de datos operativa, tratarlos y almacenarlos en el data warehouse de la compañía; definir de cubos OLAP para su posterior tratamiento, así como, la creación de cuadros de mandos y reportes.

Para este proyecto la solución de BI escogida es Pentaho Community Edition. Es considerada líder en el mercado actualmente, por lo que existe una gran cantidad de recursos como manuales, tutoriales, videos y otros materiales, que facilitan el aprendizaje.

La suite Pentaho se compone de las siguientes herramientas necesarias para cumplir los objetivos del proyecto:

• Pentaho Business Analytics Platform: desde este servidor de la plataforma podremos gestionar las conexiones con las distintas fuentes

⁹ PostGreSQL vs. MySQL (2016) [En línea] de Daniel Pecos Martínez

[[]Consulta: 1 de Junio de 2019] < https://danielpecos.com/documents/postgresql-vs-mysql/>

de datos, integración con otras herramientas como cde para crear los cuadros de mando, jpivot o saiku para hacer análisis OLAP o acceder a un Marketplace que permite añadir otras funcionalidades

- Pentaho Data Integration: Es la herramienta de ETL de la Suite.
- Pentaho Report Designer (PRD): Herramienta para crear informes.
- Pentaho Schema Workbench (PSW): en este caso se trata de una interfaz gráfica de diseño para crear y probar esquemas para los cubos OLAP de Mondrian.

4. Implantación

En este apartado se describirá la implantación de la solución, de acuerdo a las valoraciones realizadas en fases anteriores y su uso para obtener los distintos entregables y objetivos planteados.

4.1. Descarga e instalación de componentes necesarios para desarrollar la solución.

Con la finalidad de cumplir los objetivos descritos en apartados anteriores es necesario descargar e instalar los diferentes componentes de la solución.

4.1.1. Software de virtualización y sistema operativo.

El primer paso para la creación de la solución escogida será la instalación del software de virtualización y el sistema operativo sobre el que se construye la solución.

Se ha descargado Virtualbox desde el enlace:

https://www.virtualbox.org/wiki/Downloads

Tras instalar este software en el equipo anfitrión, se crea una nueva máquina con la configuración por defecto para Ubuntu 18.04. y se instala este sistema operativo que previamente ha sido descargado desde el siguiente enlace:

https://www.ubuntu.com/download/desktop

Se ha generado el usuario "nico" con contraseña "123456"

4.1.2. Sistema gestor de bases de datos

Para instalar el gestor de bases de datos mysql, se ejecuta el siguiente comando desde terminal:

sudo apt-get install mysql-server

Posteriormente se configura la seguridad del servidor por medio del comando:

sudo mysql_secure_installation

Tras seguir el asistente ya se ha configurado mysql creando un usuario administrador "admin" con la contraseña "12345678". Este será el usuario usado en el resto del desarrollo del proyecto.

4.1.3. Instalación de Java.

La Suite Pentaho se ejecuta sobre java por lo que será un requisito indispensable para la solución, la instalación este software. Con este fin se ejecuta desde terminal los siguientes comandos:

sudo add-apt-repository ppa:webupd8team/java sudo apt update sudo apt-get install oracle-java8-set-default

Una vez instalado se debe configurar la variable del sistema JAVA_HOME, para lo cual se edita el archivo del sistema /etc/environment Introduciendo la siguiente línea:

JAVA_HOME="/usr/lib/jvm/java-8-oracle/"

Tras estos cambios se actualiza esta configuración del sistema mediante el comando:

sudo source /etc/environment

4.1.4. Descarga e instalación de los componentes necesarios de la suite Bl Pentaho.

Será necesaria la descarga de los siguientes componentes: Servidor Pentaho BI server: <u>https://sourceforge.net/projects/pentaho/files/Pentaho%208.2/server/pent</u> <u>aho-server-ce-8.2.0.0-342.zip/download</u>

Pentaho Data Integration:

https://sourceforge.net/projects/pentaho/files/Pentaho%208.2/client-tools/

Pentaho schema workbench:

https://sourceforge.net/projects/mondrian/files/schema%20workbench/3. 14.0/

Pentaho Report Designer:

https://sourceforge.net/projects/pentaho/files/Report%20Designer/7.1/prd -ce-7.1.0.0-12.zip/download

Se descarga el plugin Saiku ce desde el siguiente enlace: <u>https://github.com/OSBI/saiku</u>

Tras descargar los archivos se han descomprimido en la carpeta /home/nico/Downloads/ para poder ser usados. El plugin de saiku tras descomprimirse se mueve a la carpeta /home/nico/Downloads/pentaho-server-ce-8.2.0.0-342/pentaho-server/pentaho-solutions/system/saiku/

Para realizar las conexiones con las bases de datos es necesario descargar el archivo mysql-connector-java-5.1.23.jar desde el siguiente enlace:

https://dev.mysql.com/downloads/connector/j/5.1.html

Para copiarlo en las siguientes carpetas:

- /home/nico/Downloads/pentaho-server-ce-8.2.0.0-342/pentahoserver/tomcat/lib/
- /home/nico/Downloads/psw-ce-3.14.0.0-12/schema-workbench/lib/
- /home/nico/Downloads/pdi-ce-8.2.0.0-342/data-integration/lib/
- /home/nico/Downloads/prd-ce-7.0.1.0-12/report-designer/lib/

Además, es necesrio modificar los permisos del contenido de la carpeta /home/nico/Downloads/pentaho-server-ce-8.2.0.0-342/ con el siguiente comando:

sudo chmod 777 -R /home/nico/Downloads/pentaho-server-ce-8.2.0.0-34 2/

Para iniciar el servidor cuando encienda la máquina virtual se añade un inicio automático a través del "Startup Appications Preferences" de Ubuntu indicando el siguiente comando:

Helpe	r program for	launching snap applications that are configured t	Add
SSH M	Key Agent	Add Startup Program 🔗	Remove
	Name: Command: Comment:	IniciarPentaho /pentaho-server/start-pentaho.sh Browse Iniciar el servidor Pentaho Cancel Add	

Imagen 8: Iniciar pentaho al encender la máquina virtual.

sh /home/nico/Downloads/pentaho-server-ce-8.2.0.0-342/pentaho-server/ start-pentaho.sh

Con esto ya se encuentra preparado en el equipo todo el software necesario para este prototipo.

Para acceder al servidor de Pentaho hay que introducir la siguiente url <u>http://localhost:8080</u> desde el navegador y validarse con el usuario administrador por defecto:

usuario: Admin Contraseña: password

4.2. Creación de las bases de datos Nutribox y Datawharehouse

Se crea el script "crearDatawharehouse.sql" y "Nutribox.sql" que a través de código sql genera las bases de datos, las diferentes tablas, así como sus relaciones.

Estos script se ejecutarán desde terminal del servidor utilizando los comandos:

mysql -u admin -p < Nutribox.sql

mysql -u admin -p Nutribox < crearDatawharehouse.sql

4.3. Creación de la data set de pruebas

Como ya se ha indicado anteriormente, la empresa Nutribox no está operativa actualmente por lo que se generan unos datos iniciales o de prueba, que servirán para mostrar la utilidad del prototipo.

Para introducir los datos de prueba en la base de datos se ha generado un script SQL que inserta 20 registros en la tabla de empresas colaboradoras, 20 registros en la tabla de productos, 200 registros de cupones, 1000 registros en la tabla usuarios y 5000 registros de transacciones o canjes de cupones en la base de datos Nutribox.

Para generar los 20 registros de empresas y los 20 registros de productos se ha escrito el código sql correspondiente, mientras que para el resto de los registros se utiliza la web http://generatedata.com/ para simplificar el proceso. Se ha descrito en el anexo "Uso de generatedata.com para crear los datos de prueba" que valores se han introducido en la web para generar estas tablas.

Tanto el código generado como el obtenido de generatedata.com se copia a un archivo nombrado insertarDatos.sql. Este script se ejecutará desde terminal para introducir los datos utilizando el siguiente comando:

mysql -u admin -p Nutribox < insertarDatos.sql

4.4. Construcción de ETL

Extract, Transform and Load («extraer, transformar y cargar», frecuentemente abreviado ETL) es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, data mart, o data warehouse

para analizar, o en otro sistema operacional para apoyar un proceso de negocio.¹⁰

Para la creación de estos porcesos ETL en este proyecto se utilizan transformaciones de Pentaho Data Integration, que extraen los datos de la base de datos operativa de la plataforma, realiza las transformaciones y cambios necesarios para posteriormente volcarlos en el datawharehouse.

En la siguiente tabla se detallan los procesos generados con Pentaho Data Integration:

Proceso	Descripción
HechosCanje.ktr	A partir de los datos incluidos en las tablas CouponUse y coupons de la base de datos operativa Nutribox poblar la tabla HechosCange, tras haber realizado el tratamiento necesario de los datos.
DimTiempo.ktr	Genera los datos necesarios para crear la dimensión tiempo y su carga en la tabla DimTiempo del data wharehouse.
DimProductos.ktr	A partir de los datos incluidos en la tabla products de la base de datos operativa Nutribox poblar la tabla DimProducts, tras haber realizado el tratamiento necesario de los datos.
DimColaboradores.ktr	A partir de los datos incluidos en la tabla collaborators de la base de datos operativa Nutribox poblar la tabla DimColaboradores, tras haber realizado el tratamiento necesario de los datos.
DimUsuarios.ktr	A partir de los datos incluidos en la tabla users de la base de datos operativa Nutribox poblar la tabla DimUsuarios, tras haber realizado el tratamiento necesario de los datos.

Tabla 42: Procesos ETL generados con Pentaho Data Integration

En el anexo "Procesos ETL con Pentaho Data Integration" se detalla cómo construir uno de estos trabajos. Además, los archivos .ktr generados serán entregados con el prototipo.

¹⁰ Extract, transform and load (2019) [En línea] de Wikipedia

[[]Consulta: 1 de Junio de 2019]

< https://es.wikipedia.org/wiki/Extract,_transform_and_load>

4.4.1. Transformación HechosCanje.ktr



Imagen 9: Transformación HechosCanje.ktr

Esta transformación realiza las siguientes acciones:

- Conecta con la base de datos operativa de Nutibox para extraer los datos de las tablas couponUse y coupons.
- Crea dos nuevos campos tipo string que identifican a la fecha (fechaid) y la hora (horaid) del canje.
- Realiza transformaciones para normalizar los datos (eliminando puntos y otros caracteres especiales, dejando sólo los números de estos dos campos fechaid y horaid)
- Finalmente vuelca incrementalmente los datos, en la tabla HechosCanje del datawharehouse por medio del operador Combination lookup/update.



4.4.2. Transformación DimTiempo.ktr

Imagen 10: Transformación DimTiempo.ktr

Este trabajo sólo se ejecutará una vez para cargar los datos de esta dimensión puesto que no dependerá de las operaciones de la plataforma. Al ejecutarlo realiza las siguientes acciones:

- Genera un dataset con 10000 registros tipo fecha con el valor 20180101
- Genera una variable para incrementar el valor de los registros de fecha en un día.
- A partir de los campos generados anteriormente se generan otros campos como la fecha en tipo texto; día del mes, mes, año y día de la semana en formato numérico.
- A partir del campo diaSemana (valor numérico que representa los días de la semana) se genera otro campo indicando el día de la semana en formato texto (Lunes, Martes...)
- A partir del campo mes (valor numérico que representa los días el mes) se genera otro campo indicando el mes en formato texto (Enero, Febrero,...)
- Se seleccionan los valores, descartando los valores auxiliares generados para crear nuevos atributos.
- Finalmente vuelcan los datos en la tabla DimTiempo los datos en el datawharehouse.

4.4.3. Transformación DimProductos.ktr



Imagen 11: Transformación DimProductos.ktr

Esta transformación realiza las siguientes acciones:

- Conecta con la base de datos operativa de Nutibox para extraer los datos de la tabla products.
- Crea modificaciones en algunos campos para estandarizar su formato.

- Comprueba si hay registros con el valor NULL, y en su caso completándolo con la cadena "aaaaaaa".
- Finalmente vuelca incrementalmente los datos en la tabla DimProductos del datawharehouse por medio del operador Dimension lookup/update, el cual sirve además para actualizar las dimensiones lentamente cambiantes, ya que crea claves subrogadas que permiten mantener una historia tanto de los hechos como de la propia dimensión.

4.4.4. Transformación DimColaboradores.ktr



Imagen 12: Transformación DimColaboradores.ktr

Esta transformación realiza las siguientes acciones:

- Conecta con la base de datos operativa de Nutibox para extraer los datos de la tabla collaborators.
- Crea modificaciones en algunos campos para estandarizar su formato.
- Comprueba si hay registros con el valor NULL, y en su caso completándolo con la cadena vacía.
- Finalmente vuelca incrementalmente los datos en la tabla DimColaboradores del datawharehouse por medio del operador Dimension lookup/update, el cual sirve además para actualizar las dimensiones lentamente cambiantes, ya que crea claves subrogadas que permiten mantener una historia tanto de los hechos como de la propia dimensión.

4.4.5. Transformación DimUsuarios.ktr



Imagen 13: Transformación DimUsuarios.ktr

Esta transformación realiza las siguientes acciones:

- Conecta con la base de datos operativa de Nutibox para extraer los datos de la tabla users.
- Crea modificaciones en algunos campos para estandarizar su formato.
- Divide el campo que contiene el nombre en dos campos (para nombres compuestos).
- Divide el campo que contiene los apellidos en dos campos.
- Comprueba si hay registros con el valor NULL, y en su caso completándolo con la cadena vacía.
- Finalmente vuelca incrementalmente los datos en la tabla DimUsuarios del datawharehouse por medio del operador Dimension lookup/update, el cual sirve además para actualizar las dimensiones lentamente cambiantes, ya que crea claves subrogadas que permiten mantener una historia tanto de los hechos como de la propia dimensión.

4.5. Cubos OLAP.

•

Para la definición de los cubos es necesario crear un equema xml a través de la aplicación pentaho schema workbench. En el anexo definición del esquema OLAP se detallan los pasos realizados para crear y publicar el esquema NutriboxSchema.xml utilizado en este proyecto.

En base a lo descrito en este anexo se ha creado el cubo Olap Nutribox con la estructura que se muestra en la siguiente imagen:

🛅 Schema - Olap Nutribox (NutriboxSchema.xr	nl)
🕅 🍂 📌 հ山 NS UDF CM 💊 🖮	
🛢 Schema	
🔶 🅎 Nutribox	
— 🎹 Table: HechosCanje	
— 💏 Usuario	
— 👾 Tiempo	
Producto	
Colaborador	
- 🚫 cuponesCanjeados	
 – Número de colaboradores 	
 Número de usuarios 	
 Número de productos 	
- 📎 cuponesUnicos	
Maxcanjes Acalebaradar	
• 杰 Usuario	
Table: DimUsuarios	
🛉 ភ្នំភ្ន Tiempo	
- ditte Mes	
Table: DimTiempo	
🕈 🍌 Producto	
🕈 र्र्न्नेस Producto	
— 🗄 🖞 Categoria	
- main Producto	
Table: DimProductos	
Database - datawharehouse (MySOL)	

Imagen 14: Schema OLAP Nutribox

Este esquema define las dimensiones:

- Colaborador, que usa como fuente de datos la tabla Dimcolaboradores de la base de datos datawharehouse y tiene el nivel colaboradores.
- Usuario, que usa como fuente de datos la tabla DimUsuarios de la base de datos datawharehouse y tiene el nivel Usuario.
- Tiempo, que usa como fuente de datos la tabla DimTiempo de la base de datos datawharehouse y tiene los niveles Año y Mes.
- Producto, que usa como fuente de datos la tabla DimProductos de la base de datos datawharehouse y tiene el nivel Producto.

Se ha definido el cubo Nutribox, que usa las dimensiones descritas anteriormente, la tabla de hechos HechosCanje y se han definido las métricas:

- cuponesCanjeados: Cuenta todos los identificadores de canje (transactionSkU) distintos de una selección.
- Número de colaboradores: Cuenta todos los identificadores de colaborador (CollaboratorID) distintos de una selección.
- Número de usuarios: Cuenta todos los identificadores de usuario (UserID) distintos de una selección.
- Número de productos: Cuenta todos los identificadores de producto (ProductID) distintos de una selección.
- cuponesUnicos: Cuenta todos los identificadores de cupón (CouponID) distintos de una selección.

Este cubo servirá como fuente de datos estructurada para crear diferentes componentes de cuadros de mando, o reports y permitirá mostrar diferentes análisis a través de herramientas como jpivot o saiku (instaladas en el prototipo del proyecto.)

En los anexos "Análisis de los datos del cubo con jpivot" y "Análisis de los datos del cubo con saiku" se muestran ejemplos del uso de estas herramientas.

4.6. Creación de cuadros de mando.

En base a los diseños planteados en la fase de diseño de la solución, se han definen un conjunto de cuadros de mando que representan la información de la organización y permitiendo valorar los indicadores. de una forma rápida, sencilla y accesible.

4.6.1. CM1: Cuadro de mando para la dirección ejecutiva de Nutribox

Se ha generado el siguiente cuadro de mando para la dirección de la empresa Nutribox en base a los diseños realizados:



Imagen 15: Cuadro de mandos Nutribox para el año 2019



Imagen 16: Imagen cuadro de mandos Nutribox 2020

Para su creación se ha utilizado pentaho CDE que se encuentra integrado en el servidor de pentaho business Analytics del prototipo. A continuación, se describe cada componente.



Tabla 43: CM1 Componente 1 "Selector de año"



Para obtener estos datos se utiliza la siguiente consulta MDX al cubo OLAP Nutribox:

select NON EMPTY({Descendants([Tiempo].[\${Tiempo}] ,[Tiempo].[Mes])}) on ROWS, NON EMPTY({[Measures].[Número de usuarios]}) on Columns from [Nutribox]

Indicadores	KPI 9 Usuarios activos.	
relacionados		
Table 44. Implementación CM4 Componente 2 "Actividad meneval de veverias"		

Tabla 44: Implantación CM1 Componente 2 "Actividad mensual de usuarios"

Implantación CM1 Componente 3: Usuarios totales.		
Número total de usuarios registrados: 1000		
Descripción:	Este componente muestra el número total de usuarios registrados en el sistema.	
Consultas nece	esarias para la creación del componente:	
Para obtener e datawharehouse	ste dato se utiliza la siguiente consulta SQL a la base de datos e:	
SELECT Count(DISTINCT UserID) As "Usuarios totales" From DimUsuarios		
Indicadores relacionados	KPI 8 Usuarios registrados	

Tabla 45: Implantación CM1 Componente 3 "Usuarios totales"

Implantación CM1 Componente 4: Usuarios activos.			
Ususarios			
activos			
F20			
	228		
Descripción:	Este componente muestra el número de usuarios que han canieado algún cunón el año a evaluar y su tendencia con		
	respecto al año anterior. Para obtener estos años se utiliza la		
	siguiente consulta MDX al cubo OLAP Nutribox:		

Consultas necesarias para la creación del componente:

Para obtener estos datos se utiliza la siguiente consulta MDX al cubo OLAP Nutribox:

WITH MEMBER [Tiempo].[Tendencia] as [Tiempo].[\${Tiempo}] -[Tiempo].[\${Tiempo}].PREVMEMBER SET [~C] AS {[Tiempo].[\${Tiempo}], [Tiempo].[\${Tiempo}].PREVMEMBER, [Tiempo].[Tendencia]} SELECT NON EMPTY [~C] ON COLUMNS, NON EMPTY [~C] ON COLUMNS, NON EMPTY {[Measures].[cuponesCanjeados]} ON ROWS FROM [Nutribox] Indicadores relacionados KPI 9 Usuarios activos. KPI 10 Tendencia Usuarios activos.

Tabla 46: Implantación CM1 Componente 4 "Usuarios activos"

Implantación C	CM1 Componente 5: T	abla empres	as colaboradoras	
	Top col	aborado	res	
	Empresa colaboradora	Canjes	Tendencia	
	Agricultores De La Palma S.L.	201		
	Cooperativa Palmera	172		
	Higopico S.L.	163		
	Bimbache S.L.	162		
	Alguito S.L.	156	2	
		Pre	evious Next	
Descripción:	Este componente n Empresas colabora cupones canjeados se indica el número como su tendencia e	nuestra una adoras orde para adquiri de canjes to en comparac	tabla con los v enados por el sus productos. E otales en el año i ión con el ejercic	alores de 5 número de En esta tabla ndicado, así io anterior.

	Aunque inicialmente sólo muestra el top 5 en orden de canjes realizados pueden verse el resto de los registros utilizando las etiquetas de Previous y Next.
Consultas nece	esarias para la creación del componente:
Para obtener e Nutribox:	estos años se utiliza la siguiente consulta MDX al cubo OLAP
WITH MEMBE [Tiempo SET [~C {[Tien	R [Tiempo].[Tendencia] as [Tiempo].[\${Tiempo}] -].[\${Tiempo}].PREVMEMBER COLUMNS] AS npo].[\${Tiempo}],
[Tiempo SET [~R].[\${Tiempo}].PREVMEMBER,[Tiempo].[Tendencia]} COWS] AS
	lboradorj.[Colaboradoresj.Members}
NON EN	/IPTY CrossJoin([~COLUMNS],
{[Measu	res].[cuponesCanjeados]}) ON COLUMNS,
NON EN	IPTY [~ROWS] ON ROWS
FROM	Nutribox]
Indicadores	KPI 4 Popularidad del Colaborador
relacionados	KPI 5 Tendencia Colaborador.

Tabla 47: Implantación CM1 Componente 5 "Tabla empresas colaboradoras"



Consultas nece	esarias para la creación del componente:	
Para obtener estos valores se utiliza la siguiente consulta MDX al cubo OLAP Nutribox:		
select NON EMPTY({Descendants([Producto].[All Productos], [Producto].[Categoria])}) on ROWS, NON EMPTY({[Measures].[cuponesCanjeados]}) on Columns from [Nutribox] WHERE {[Tiempo].[\${Tiempo}]}		
Indicadores	KPI 3 Porcentaje de canjes por Categoría de productos.	
Indicadores relacionados	KPI 3 Porcentaje de canjes por Categoría de productos.	



Implantación C	M1 Componente	/: Tabla de p	oroductos	
	т	op product	OS	
	Producto	Canjes	Tendencia	
	Café	218	2	
	Pulpo	183		
	Tomate Ensalada	156		
	Ron De Caña	155	A.	
	Platano	154	2	
			Previous Next	
Descripción:Este componente muestra una tabla con los valores de 5 productos ordenados por el número de cupones canjeados para adquirir ese producto en el año a estudiar. En esta tabla se indica el número de canjes totales en el año indicado, así 				
Consultas nece	esarias para la cro	eación del co	mponente:	
Consultas necesarias para la creación del componente:				

Para obtener e Nutribox:	estos valores se utiliza la siguiente consulta MDX al cubo OLAP
Para obt	ener estos años se utiliza la siguiente consulta MDX al cubo
OLAP Ni	utribox:
WITH MEMBEI [Tiempo]	R [Tiempo].[Tendencia] as [Tiempo].[\${Tiempo}] - .[\${Tiempo}].PREVMEMBER
SET [~C	OLUMNS] AS
{[]	[iempo].[\${Tiempo}], [Tiempo].[\${Tiempo}].PREVMEMBER,
[Tiempo]	.[Tendencia]}
SET [~R	OWS] AS
{[F	Producto].[Producto].Members}
SELECT	
NON EM	IPTY CrossJoin([~COLUMNS],
{[Measur	res].[cuponesCanjeados]}) ON COLUMNS,
NON EM	IPTY [~ROWS] ON ROWS
FROM [N	Nutribox]
Indicadores	KPI 1 Popularidad del producto.
relacionados	KPI 2 Tendencia Producto.

Tabla 49: Implantación CM1 Componente 7 "Tabla de productos"

Implantación CM1 Componente 8: Canjeados al año	
	Anuales 2548
Descripción:	Indica el número de cupones canjeados el año indicado.
Consultas nece	esarias para la creación del componente:
Para obtener e Nutribox:	stos valores se utiliza la siguiente consulta MDX al cubo OLAP
WITH MEMBEI [Tiempo] SET [~C [Tiempo] SELECT NON EM	R [Tiempo].[Tendencia] as [Tiempo].[\${Tiempo}] - .[\${Tiempo}].PREVMEMBER] AS Tiempo].[\${Tiempo}], [Tiempo].[\${Tiempo}].PREVMEMBER, .[Tendencia]} PTY [~C] ON COLUMNS.

NON EMPTY {[Measures].[cuponesCanjeados]} ON ROWS FROM [Nutribox]	
Indicadores	KPI 7 Total canjes
relacionados	KPI 6 Tendencia Canjes anuales

Tabla 50: Implantación CM1 Componente 8 "Canjeados al año"



Tabla 51: Implantación CM1 Componente 9 "Canjes por mes"

4.6.2. CM2: Cuadro de mando empresa colaboradora.

Se ha generado el siguiente cuadro de mando para la dirección ejecutiva de las diferentes empresas colaboradoras, en base a los diseños realizados:



Imagen 17: Cuadro de mando para la empresa Papa Arrugada S.L.


Imagen 18: Cuadro de mando para la empresa Cancajos S.L. del año 2020.

Para su creación se ha utilizado pentaho CDE que se encuentra integrado en el servidor de pentaho business Analytics del prototipo. A continuación, se describe cada componente.



	con transacciones registradas. El año selecionado se almacena en un parámetro "Tiempo" que se utilizará para actualizar el resto de los elementos.
Consultas nece	esarias para la creación del componente:
Para el selecto Nutribox que re with mer select To ,[Tiempo {[Measu from [Nu	or del año se ha usado una consulta MDX sobre el cubo OLAP ecoge los años en los que ha habido actividad en la plataforma: nber [Measures].[Name] as '[Tiempo].CurrentMember.UniqueName' opCount(filter({Descendants([Tiempo].[All Tiempos]].[Año])}, not isempty(([Tiempo].CurrentMember))), 50) on ROWS, ures].[Name]} on Columns utribox]
A partir de este	selector se crea el parámetro Tiempo.





Tabla 53: Implantación CM2 Componente 2 "Selector de empresa colaboradora".

Implantación C	Implantación CM2 Componente 3: Clientes en el ejercicio.		
	Clientes en el ejercicio: 57		
Descripción:	Muestra el número total de usuarios que han realizado algún canje de cupones del colaborador, en el año indicado, así como su tendencia con respecto al ejercicio anterior.		
Consultas nece	esarias para la creación del componente:		
WITH MEMBER [Tiem [Tiempo].[\${Tien SET [~C] AS {[Tiempo].[Tende SELECT NON EMPTY [~ NON EMPTY [[~ FROM [Nutribox where [Colabora	po].[Tendencia] as [Tiempo].[\${Tiempo}] - npo}].PREVMEMBER].[\${Tiempo}], [Tiempo].[\${Tiempo}].PREVMEMBER, encia]} C] ON COLUMNS, Measures].[Número de usuarios]} ON ROWS] ador].[\${Col}]		
Indicadores relacionados	KPI 11 Tendencia clientes del colaborador.		

 Tabla 54: Implantación CM2 Componente 3 "Clientes en el ejercicio".



Consultas necesarias para la creación del componente:			
select NON EMPTY({Descendants([Tiempo].[\${Tiempo}], [Tiempo].[Mes])}) on			
ROWS,			
NON EMPTY({[Measures].[Número de usuarios]}) on Columns			
from [Nutribox]			
where [Colaborador].[\${Col}]			
Indicadores	KPI 14 Usuarios activos por colaborador y mes		
relacionados			

Tabla 55: Implantación CM2 Componente 4 "Actividad mensual de los clientes"

Implantac	ción C	M2 Componente 5	: Top Clientes.	
		Τομ	o Clientes	
	Usua	rio	Canjes	Tendencia
	Jonah Sweet		2	
2	Reese Hooper		1	
	Adria F	Reese	1	
1	Colton Andrews		1	
1	Jelani	Gilmore	1	e.
				Previous Next
Descripción: Muestra en formato de tabla los 5 clientes que han realizado n canjes de cupones ofertados por este colaborador, que más és han tenido en la plataforma en base al número de cupor canjeados de sus ofertas. Muestra además del valor numérico año actual, su tendencia respecto al año anterior.				ientes que han realizado más e colaborador, que más éxito ase al número de cupones demás del valor numérico del año anterior.
Consultas	s nece	esarias para la crea	ción del compon	ente:
WITH MEMBER [Tiempo].[Tendencia] as [Tiempo].[\${Tiempo}] - [Tiempo].[\${Tiempo}].PREVMEMBER				
SET [~COLUMNS] AS {[Tiempo].[\${Tiempo].[\${Tiempo}].PREVMEMBER, [Tiempo].[Tendencia]} SET [~ROWS] AS {[Usuario].[Usuario].Members} SELECT				

NON EMPTY CrossJoin([~COLUMNS], {[Measures].[cuponesCanjeados]}) ON COLUMNS, NON EMPTY [~ROWS] ON ROWS FROM [Nutribox] where [Colaborador].[\${Col}]

indicadores	KPT 12. Actividad del cliente.
relacionados	KPI 13: Tendencia en la actividad del cliente.

Tabla 56: Implantación CM2 Componente 5 "Top Clientes".

Produc	to	Canjes	Tendencia
Tomate I	Ensalada	9	
Pepino E	ico	9	1 % .
Piña Tro	pical	8	•
Aguacat	e	6	•
Pulpo		6	(*)
	Mussing on fo		Previous Next
1 1 X X	iviuestra en to	ormato de tabla los 5 lue más éxito han tenid	productos ofertados por e lo en la plataforma en base
escripción:	colaborador, c número de cuj valor numérico	pones canjeados de sus o del año actual, su tend	s ofertas. Muestra además encia respecto al año anter
escripción: consultas nec	colaborador, c número de cu valor numérico esarias para la	pones canjeados de sus o del año actual, su tend creación del compon	s ofertas. Muestra además encia respecto al año anter nente:

NON EMPTY CrossJoin([~COLUMNS], {[Measures].[cuponesCanjeados]}) ON COLUMNS, NON EMPTY [~ROWS] ON ROWS FROM [Nutribox] where [Colaborador].[\${Col}]	
Indicadores	KPI 15 Popularidad del producto de colaborador X.
relacionados	KPI 16 Tendencia de canjes de producto por colaborador.

Tabla 57: Implantación CM2 Componente 6 "Top productos".

Implantación CM2 Componente 7: Canjes anuales.			
	Anuales		
	201 🧖		
Descripción:	Muestra el número de canjes de cupones total que ha obtenido este colaborador en el año a analizar, así como su tendencia con respecto al ejercicio anterior.		
Consultas nece	esarias para la creación del componente:		
WITH			
MEMBER [Tiem	po].[Tendencia] as [Tiempo].[\${Tiempo}] -		
[TIEMPO].[\${TIEMPO}].PREVMEMBER			
{[Tiempo].[\${Tiempo}]. [Tiempo].[\${Tiempo}].PREVMEMBER.			
[Tiempo].[Tendencia]}			
SELECT			
	NON EMPTY [~C] ON COLUMNS,		
NON EIVIPTT {[IVIeasures].[CUPONESCANJEADOS]} ON KOVVS			
where [Colaborador].[\${Col}]			
	31. C)3		
Indicadores	KPI 17 Tendencia Canjes anuales por colaborador.		
relacionados	KPI 18 Total canjes anuales por colaborador.		

Tabla 58: Implantación CM2 Componente 7 "Canjes anuales".



Tabla 59: Implantación CM2 Componente 8 "Número de canjes realizados al mes.".

4.7. Creación de informes dinámicos en Pentaho Report Designer.

En base a los diseños planteados en la fase de diseño de la solución, se han definen un conjunto de informes dinámicos que representan información de la organización y permiten valorar los indicadores establecidos en la fase de análisis. Se puede ver un ejemplo de su construcción en el anexo Creación de informes con Pentaho Report Designer. A continuación, se detallan algunos de estos informes:

4.7.1. Informe de canjes por categoría y año

Informe de canjes por catego	oría y año		
Canies realiz	ados por categoría y año - Prir	at Preview	
Report Export View Help			
B & < > > < <	100% 🔻		
Tiemno 2019			-
Auto-Update on selection			Update
Canjes realizados por	categoría el año 20	19	
	2		
(169)	1 796		
10%	4%		
10%			
			=
12%	13%		
11%	10%		
🐞 Fruta Eco 🔅 Ha	ninas 🦲 Salsas 🔵 Infusiones		
Verduras Eco	Bebidas Alcoholicas 🕚 Cosmeticos		
Carnes O Pesc	aderia		
Categoría		Cantidad	
Fruta Eco		154	
Harinas		52	
Salsas		56	
Infusiones		218	
Verduras Eco		119	
Bebidas Alcoholicas		127	
Cosmeticos		133	
Carnes		114	
Pescadena		183	
L			
•			1/1
Descripción:	En este informe se	representa en un	n gráfico
	tipo pastel, el porce	ntaje de canjes rea	alizados
	sobre una categori	ía determinada y	en una
	tabla se presenta ur	n listado con las di	ferentes
	categorías y el núm	nero de canjes rea	lizados.
	En la parte superior	se ha incluido un	selector
	para seleccionar el	año a representa	ar en el
	informe.		
Consultas necesarias para la	a creación del informe:		
Para el selector del año se h	a usado una consulta l	MDX sobre el cub	
Nutriboy que recordo los años	a usado una consulta i	actividad on la nlat	taforma
Tradition que recoye los allos	5 cm 103 que ha habitu d	actividad en la plat	aionna.

with member [Measures].[Name] as '[Tiempo].CurrentMember.UniqueName' select TopCount(filter({Descendants([Tiempo].[All Tiempos],[Tiempo].[Año])}, not isempty(([Tiempo].CurrentMember))), 50) on ROWS, {[Measures].[Name]} on Columns from [Nutribox]

A partir de este selector se crea el parámetro tiempo.

Para obtener los datos que se representan en el gráfico y la tabla del informe se crea la siguiente consulta MDX al cubo OLAP Nutribox:

select NON EMPTY({Descendants([Producto].[All Productos] ,[Producto].[Categoria])}) on ROWS, NON EMPTY({[Measures].[cuponesCanjeados]}) on Columns from [Nutribox] WHERE {[Tiempo].[\${Tiempo}]}

Indicadores relacionados:	KPI 3 Porcentaje de canjes por Categoría de
	productos.

Tabla 60: Informe de canjes por categoría y año

4.7.2. Informe Top 10 colaboradores del año



Descripción:	En este informe se representa en un gráfico de barras y en una tabla el número de canjes realizados a los productos ofertados por los 10 colaboradores más demandados de un año concreto. En la parte superior se ha incluido un selector para seleccionar el año a representar en el informe.			
Consultas necesarias para la	a creación del informe:			
Para el selector del año se ha usado una consulta MDX sobre el cubo OLAP Nutribox que recoge los años en los que ha habido actividad en la plataforma.				
with member [Measures].[Name] as '[Tiempo].CurrentMember.UniqueName' select TopCount(filter({Descendants([Tiempo].[All Tiempos] ,[Tiempo].[Año])}, not isempty(([Tiempo].CurrentMember))) , 50) on ROWS, {[Measures].[Name]} on Columns from [Nutribox]				
A partir de este selector se crea	a el parámetro tiempo.			
Para obtener los datos que se crea la siguiente consulta MDX	representan en el gráfico y la tabla del informe se al cubo OLAP Nutribox:			
WITH SET [~ROWS] AS {[Colaborador].[Colaboradores].Members} SELECT NON EMPTY {[Measures].[cuponesCanjeados]} ON COLUMNS, NON EMPTY TopCount([~ROWS], 10, [Measures].[cuponesCanjeados]) ON ROWS FROM [Nutribox]				
Indicadores relacionados:	KPI 4 Popularidad del colaborador.			

Tabla 61: Informe Top 10 colaboradores del año

4.7.3. Informe Productos canjeados del año

Informe Productos canjeados de	el año	
Productos c	anjeados del año - Print <u>Preview</u>	
Report Export View Help		
	9% 💌	
Tiempo 2019		
Auto-Update on selection		Update
Productos canjeados de	el año 2019	4
3% 5% 5% 5% 5% 5% 5% 5% 5% 5%	6% 2% 2% 9% 5% 5% 6%	Ē
 Platano Gofio Mojo Picón Café Piña Trop Tomate Ensalada Calabaza Eco Pepino Eco Alfonsiño Calamar Vieja 	vical 单 Lechuga 🌒 Vino Bodeguero 🌑 Ron De Caña 🌒 Crema Aloe Vera Chorizo Eco 🔵 Cochino Negro 🌑 Pollo De Corral 🕘 Aguacate 🕘 Pulpo	
Categoría del producto Producto	Cantidad	
Harinas Gofio	52	
Salsas Mojo Picón	56	
Infusiones Café	218	
•		1/2
Descripción:	En este informe se representa en un de pastel el porcentaje de canje producto y en una tabla el nomb categoría y la cantidad de canjes de producto.	gráfico es por re, su e cada

Consultas necesarias para la creación del informe:

Para el selector del año se ha usado una consulta MDX sobre el cubo OLAP Nutribox que recoge los años en los que ha habido actividad en la plataforma.

with member [Measures].[Name] as '[Tiempo].CurrentMember.UniqueName' select TopCount(filter({Descendants([Tiempo].[All Tiempos],[Tiempo].[Año])}, not isempty(([Tiempo].CurrentMember))), 50) on ROWS, {[Measures].[Name]} on Columns from [Nutribox]

A partir de este selector se crea el parámetro tiempo.

Para obtener los datos que se representan en el gráfico y la tabla del informe se crea la siguiente consulta MDX al cubo OLAP Nutribox:

WITH

SET [~ROWS] AS {[Producto].[Producto].Members} SELECT NON EMPTY {[Measures].[cuponesCanjeados]} ON COLUMNS, NON EMPTY [~ROWS] ON ROWS FROM [Nutribox] WHERE ([Tiempo].[\${Tiempo}])

Indicadores relacionados: KPI1 Popularidad del producto.

Tabla 62: Informe Productos canjeados del año

5. Líneas de futuro

La solución BI de la empresa Nutribox deberá adaptarse a las necesidades de la organización de forma que sea útil para obtener conocimiento a través de los datos disponibles. Para ello se seguirá un modelo de desarrollo iterativo de la solución donde se continuarán las fases de análisis, diseño e implantación de los cambios de las diferentes versiones de esta solución.

Así una vez implantada una versión se deberá iniciar una fase de análisis donde, como se ha hecho en este proyecto, se deberán plantear los indicadores y fuentes de datos con el objetivo de valorar el cumplimiento de los nuevos objetivos establecidos.

Ampliación del cubo OLAP existente definiendo nuevas dimensiones, ampliar las existentes, o la creación de nuevos cubos. Así, por ejemplo, se propone la creación de los niveles Semana y Hora, que permitan realizar nuevos análisis a partir de los datos disponibles.

Para siguientes iteraciones o versiones de la solución se propone el registro y tratamiento de otros datos obtenidos a partir de la actividad de

la plataforma que no han sido valorados en este prototipo inicial, como pueden ser la fecha de registro de los usuarios, que permitirá estudiar la evolución de registros y nuevos usuarios de la plataforma.

Además, con la finalidad de ofrecer un mejor servicio de los usuarios, se hace necesario un conocimiento mayor de los mismos por lo que se contempla que en el registro se obtengan de forma voluntaria otros datos personales y que estos puedan ser tratados en la solución de business intelligence de la compañía.

Del mismo modo, en posteriores evoluciones de la solución se plantea el uso de técnicas de web scraping a través de la herramienta Pentaho Data Integration usada en este proyecto, para obtener datos que puedan dar lugar a información relevante para la compañía.

Por ejemplo, el uso de estas técnicas sobre diferentes redes sociales puede suministrar un mayor conocimiento sobre los posibles usuarios de la compañía. También pueden servir para obtener datos disponibles en diferentes repositorios de Open Data que pueden aportar información acerca del entorno o necesidades de estos usuarios.

Por otro lado, el uso de datos geográficos o de localización pueden servir para proveer soluciones a determinados problemas. Así el tratamiento de estos datos y representación de estos a través de la solución propuesta, podrá ser otra posibilidad a tener en cuenta en evoluciones futuras.

Así, por ejemplo, a través de datos de localización de usuarios se podrá medir la distancia desde estos a los diferentes establecimientos de las empresas colaboradoras de Nutribox y representarlos por medio de componentes de mapa (Map component) en los cuadros de mando realizados por la solución BI a través de Pentaho, de forma que esta información resulte accesible y fácilmente interpretada.

Con el objeto de dotar a la plataforma de una mayor capacidad analítica se podrán integrar otras herramientas como Weka, que se encuentra disponible como plugin de Pentaho y permite aplicar algoritmos de machine learning (aprendizaje automático) y minería de datos.

Así, por ejemplo, se podrían utilizar algoritmos de agrupamiento con la finalidad de clasificar los diferentes perfiles de usuarios o el uso de árboles de decisión que permitirán observar patrones de compra de nuestros usuarios y poder realizar sugerencias de compras.

Finalmente, si el volumen de datos a tratar por la solución excediera a las capacidades de la base de datos actual, se podrá plantear el uso de Hadoop, que es actualmente soportado por Pentaho y permite el tratamiento escalable y distribuido de los datos aumentando la capacidad del sistema.

6. Conclusiones

En primer lugar, cabe destacar la dificultad que ha supuesto basar este trabajo en una idea de negocio propuesta a partir de la iniciativa de la UOC "Emprendedores interestudios" a la hora de definir, planificar y ejecutar este proyecto.

Como integrantes de esta iniciativa hemos trabajado sobre temáticas muy diferenciadas, las compañeras de Administración de empresas en el ámbito empresarial y yo sobre una solución de BI para esta empresa, por lo que finalmente la mayoría del trabajo realizado lo hemos llevado a cabo de forma autónoma.

Desde mi punto de vista, para que surgieran las sinergias esperadas entre los diferentes perfiles, era necesario un perfil de desarrollo o ingeniería de software que tratara el caso de la plataforma sobre la que se basa el servicio de Nutribox y sirviera de nexo entre los demás trabajos.

Por otro lado, puesto que esta empresa se encuentra en fase de idea, no se dispone de una fuente de datos real sobre la que trabajar por lo que ha sido necesario el diseño de una base de datos inicial sobre la que se basa esta solución BI y la creación de unos datos ficticios sobre los que poder mostrar los resultados.

Además, como esta empresa se encuentra en proceso de definición, no se tienen referencias sobre los posibles problemas o necesidades de esta organización, lo que ha dificultado en gran medida la definición de objetivos, alcance e indicadores a tratar por esta solución.

Tenido estos hechos en cuenta, se ha definido y simulado la base de datos operativa de la plataforma introduciendo un conjunto de datos o dataset de prueba. Posteriormente se han definido los principales indicadores o KPI a valorar por él prototipo.

Utilizando herramientas Open Source se ha procedido a la definición y construcción de los procesos de extracción, transformación, carga (ETL), construcción de un sistema de almacenamiento y gestión de datos (data warehouse), diseño de cubos OLAP, así como el diseño y construcción de cuadros de mando e informes que presenten los indicadores (KPI) definidos en la fase de análisis.

Hay que destacar que el uso de diferentes herramientas, como Pentaho Data Integration, Pentaho Schema Workbench, Pentaho report designer o pentaho cde, ha supuesto una gran dedicación para comprender su funcionamiento y configuración; y así poder obtener los resultados esperados.

Finalmente, se han descrito una serie de líneas de futuro que pueden servir como base para las futuras versiones de la solución implementada,

por lo que se puede considerar que se han alcanzado los objetivos planteados en la definición del trabajo.

7. Bibliografía

- Indicador clave de rendimiento (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 https://es.wikipedia.org/wiki/Indicador_clave_de_rendimiento
- Almacén de datos (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 ">https://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos>
- Almacén de datos (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 ">https://es.wikipedia.org/wiki/Almac%C3%A9n_de_datos>
- Esquema en estrella (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 https://es.wikipedia.org/wiki/Esquema_en_estrella
- OLAP (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 https://es.wikipedia.org/wiki/OLAP
- Oracle VM VirtualBox (2019) [En línea] de ORACLE [Consulta: 1 de Junio de 2019]
 https://www.oracle.com/es/virtualization/virtualbox/>
- Oracle MySQL (2019) [En línea] de ORACLE [Consulta: 1 de Junio de 2019] <https://www.oracle.com/es/mysql/>
- PostGreSQL vs. MySQL (2016) [En línea] de Daniel Pecos Martínez [Consulta: 1 de Junio de 2019]
 https://danielpecos.com/documents/postgresql-vs-mysql/
- Infografía MySQL vs PostgreSQL (2016) [En línea] de infranetworking [Consulta: 1 de Junio de 2019]
 https://blog.infranetworking.com/infografia-mysql-vs-postgresql/
- Extract, transform and load (2019) [En línea] de Wikipedia [Consulta: 1 de Junio de 2019]
 < https://es.wikipedia.org/wiki/Extract,_transform_and_load>

8. Anexos

8.1. Descarga y uso del prototipo.

Se ha alojado el archivo Prototipo.ova en OneDrive para que pueda ser descargado a través del siguiente enlace:

https://uoc0-

my.sharepoint.com/:u:/g/personal/nayut_uoc_edu/EapS0JNmHsJLsM8nPfDe JzMBBGjO6mM2fg03fBFECDnrdw?e=TmS6b2

Para poder utilizar este archivo que contiene el servicio virtualizado, es necesario tener instalado el software VirtualBox que puede ser descargado desde:

https://www.virtualbox.org/wiki/Downloads

Desde esta aplicación se deberá importar el servicio virtualizado a través del menú Archivo, seleccionar la opción "Importar el servicio virtualizado" como se muestra en la siguiente imagen:



Imagen 19: Importar servicio virtualizado.

Se mostrará una ventana donde se deberá indicar la ruta del archivo Prototipo.ova:

			2	ä
Importar servicio virtualizado				
Servicio a importar				
Virtualeox actualmente soporta importar servicos guardados en Open Virtualization P Importar abajo.	ormat (DVP), Para com	inuar, seleccion	e el archiv	10.1
D: OneDrive - Universitat Oberta de Catalunya (TFG/Entrega final Prototipo.ova]
				4.5
			-	

Imagen 20: Indicar la ruta del archivo.

Tras pulsar en siguiente o "Next" se mostrará la información del archivo y se podrá indicar la carpeta que contendrá las máquinas virtuales:

			?	×
Import	ar servicio virtualizado			
Prefere	encias de servicio			
Estas sor Puede ca	n las máquinas virtuales contenidas en el se ambiar algunas de las propiedades mostrada	rvicio y las preferencias sugeridas de las máquinas virtuales impor Is haciendo doble dic en los elemenos y deshabilitar otras usando	tadas de Virtua las casillas de a	lBox. bajo.
Sistema	virtual 1			^
- 😓	Nombre	prototipoTFG		
P	Producto	Prototipo Solución Bl para la empresa Nutribox		
F	Vendedor	Nicolás Miguel Ayut González		
P	Descripción	Prototipo Solución Bl para la empresa Nutribox		
	Tipo de SO invitado	칼 Ubuntu (64-bit)		
	CPU	1		
	RAM	5185 MB		~
Puede mo (por mág	odificar la carpeta base que contendrá toda juina virtual). tanuinasVirtuales	is las máquinas virtuales. Las carpetas de inicio pueden ser modifi	cadas individual	Imente
Política de	dirección MAC: Induir solo las direcciones	NAT de adaptador de red		
Oncion	nes adicionales: 🔽 Importar discos como l			
Servicio vi	irtualizado no firmado			
		Restaurar valores predeterminados Impor	tar Cano	celar

Imagen 21: Sistema virtualizado.

Tras pulsar sobre la opción Importar en la ventana anterior comenzará el proceso de importación:



Imagen 22: Proceso de importación

Una vez finalizado el proceso de importación y tras seleccionar la máquina virtual prototipoTFG como se ve en la imagen, se podrá iniciar la máquina usando el botón de Iniciar:



Imagen 23: Iniciar máquina virtual.

Se iniciará la máquina virtual y automáticamente cargará la sesión del usuario por defecto (usuario: nico, contraseña: 123456). Se ha configurado este servicio para que tras la carga del sistema operativo inicie el servidor Pentaho Business Analytics.

Se recomienda esperar unos minutos antes de empezar a utilizarlo ya que los servicios del servidor Pentaho tardan en estar listos para su uso. Una vez el servidor esté operativo se podrá acceder a el desde el navegador introduciendo la url: localhost:8080/:

🙋 Pentaho User Console - 🛛 🗙	.+
€ → ℃ @	Iocalhost:8080/pentaho/Login
👲 Getting Started	
	HIIACHI
	Pentaho User Console
	(I mar Mannar
	Admin
	Password:
	Login
	Login as an Evaluator 🛩

Imagen 24: Acceder al servidor Pentaho Business Analytics.

Tras validar con el usuario administrador por defecto (Usuario: Admin, contraseña: password). Será posible acceder a todos los módulos utilizados en este prototipo, así como a los cuadros de mando e informes publicados en este prototipo. Así para poder acceder a estos archivos, se debe pulsar sobre la opción Browse Files del servidor Pentaho Business Analytics:

 Pentaho User Console × + ← → C û i localhost:8080/pe Getting Started 	Pentaho User Console - Mozilla Firefox
File View Tools Help Home ~	
Browse Files	Pentaho Business Analytics
Create New	Get help and contribute with
Manage Data Sources	your knowledge.
Imagen 25: Detallo	e Pentaho Business Analytics.

Al abrir, esta opción se muestra una sección de carpetas donde se puede observar todo el árbol de directorios del servidor y una sección de archivos, donde se muestra el contenido de la carpeta seleccionada en la sección anterior. Los archivos publicados en este trabajo se han ubicado en la carpeta Public, tras crear los directorios CuadrosDeMando e Informes como puede verse en la siguiente imagen:

Folders O	Files	File Actions
 Home admin pat suzy tiffany Public 	 Colaboradores.cda Colaboradores.cdfde CuadroDE.cda CuadroDE.cdfde Cuadro de mando para empresas Colab Cuadro de mando para la dirección ejec 	Open Open in a new window Edit Cut Copy Move to Trash Rename
 Informes: Trash 		Download Share Add to Favorites Properties

Imagen 26: Browse files en Pentaho Business Analytics.

Así se podrá realizar diferentes acciones sobre los archivos como abrirlo, abrirlo en una nueva ventana o editarlo, usando las acciones de la columna derecha.

Para utilizar las herramientas pentaho report designer, pentaho data integration y pentaho schema workbench, se deberán ejecutar desde terminal los siguientes comandos:

Para Pentaho data integration:

sudo sh /home/nico/Downloads/pdi-ce-8.2.0.0-342/data-integration/spoon.sh

Para Pentaho report designer:

sudo sh /home/nico/Downloads/prd-ce-7.1.0.0-12/report-designer/report-designer.sh

Para Pentaho schema workbench:

sudo sh /home/nico/Downloads/psw-ce-3.14.0.0-12/schema-workbench/workb ench.sh

8.2. Uso de generatedata.com para crear los datos de prueba

En este anexo se detallan los pasos seguidos para obtener los distintos datos de pruebas a partir de esta web. Así, para obtener el código que generará los usuarios se introduce los siguientes valores en la web:

Ejempios Opciones Ayuda Elim. * No hay ejemplos disponibles. No hay opciones disponibles. ? ? * eZg29gdF5K1 (Contraseña) LLLxxLLxLL ? ? * Alex (cualquier sexo) * Name ? ? * Smith (apellido) * Surname ? ? ? * No hay ejemplos disponibles. No hay opciones disponibles. ? ? ? * No hay ejemplos disponibles. No hay opciones disponibles. ? ? ? * No hay ejemplos disponibles. Image: Spanish Autonomies & Comp. ? ? ? * No hay ejemplos disponibles. Image: Spanish Autonomies ? ? ?
Ejempios Opciones Ayuda Elim No hay ejemplos disponibles. No hay opciones disponibles. eZg29gdFSK1 (Contraseña) LLLxxLLxLL ? Alex (cualquier sexo) Name ? Smith (apellido) Surname ? No hay ejemplos disponibles. No hay ejemplos disponibles. No hay ejemplos disponibles. No hay ejemplos disponibles. ? ? ? ? ? No hay ejemplos disponibles. No hay ejemplos disponibles. ? ? ? No hay ejemplos disponibles. ? ? ? No hay ejemplos disponibles. ? ?
Ejemplos Opciones Ayuda Elim No hay ejemplos disponibles. No hay opciones disponibles. eZg29gdF5K1 (Contraseña) LLLxXLLXLL Image: Contraseña LLLxXLLXLL Image: Contraseña LLxxLLXLL Image: Contraseña Sumame Image: Contraseña Image: Co
Ejempios Opciones Ayuda Elim No hay ejempios disponibles. No hay opciones disponibles. Zg292gdF5K1 (Contraseña) LLLxxLLxLL Z Alex (cualquier sexo) Name Smith (apellido) Surname Simith (apellido) Surname No hay ejemplos disponibles. No hay ejemplos disponibles. No hay ejemplos disponibles. No hay ejemplos disponibles. Spanish Autonomies No hay ejemplos disponibles.
Ejemplos Opciones Ayuda Elim No hay ejemplos disponibles. No hay opciones disponibles. eZg29gdf5K1 (Contraseña) LLLxxLL ? Alex (cualquier sexo) Name Smith (apellido) Surname ? No hay ejemplos disponibles. Spanish Autonomies ? ?
V No hay ejemplos disponibles, ? V eZg29gdf5K1 (Contraseña) LLLxLLxLL ? V Alex (cualquier sexo) V Name ? V Smith (apellido) V Surname ? V No hay ejemplos disponibles. No hay opciones disponibles. ? V No hay ejemplos disponibles. % ? V No hay ejemplos disponibles. ? ?
• eZg29gdF5K1 (Contraseña) LLLxxLLxLL ? • Alex (cualquier sexo) • Name ? • Smith (apellido) • Surname ? • No hay ejemplos disponibles. No hay opciones disponibles. ? • No hay ejemplos disponibles. Image: Spanish Autonomies Comp. Monte 2 ? • No hay ejemplos disponibles. ? ? • No hay ejemplos disponibles. ? ?
 Alex (cualquier sexo) Name Smith (apellido) Surname No hay ejemplos disponibles. No hay ejemplos disponibles. Spanish Autonomies Completion No hay ejemplos disponibles. Spanish Autonomies Organish Autonomies
• Smith (apellido) • Surname ? • No hay ejemplos disponibles. No hay opciones disponibles. ? • No hay ejemplos disponibles. • Spanish Autonomies Computer to
No hay ejemplos disponibles. ?
No hay ejemplos disponibles. Spanish Autonomics Comduttore ? No hay ejemplos disponibles. Spanish Autonomics ?
No hay ejemplos disponibles. Spanish Autonomies
Pormatos diferentes
🔻 No hay ejemplos disponibles. 🔲 Limitar los países a los elegidos más arriba 👔 🛄
No hay ejemplos disponibles. No hay opciones disponibles.
Ejemplos Opciones Ayuda Elim
Vo hay ejemplos disponibles. Vo hay ejemplos disponibles. Vo hay ejemplos disponibles. Vo hay ejemplos disponibles. Vo hay opciones disponibles. Ejemplos Opciones. Ay

Imagen 27: Generar datos de usuarios.

Para obtener el código que generará los cupones se introduce los siguientes valores:

Generate Abo	out News	Donate			
coupons			SAVE		ð
OUNTRY-SPECIFIC DATA	0				
Spain 🗙					
ATA SET 🐶 Order Table Column	Data Type	Examples	Options	Help	De
1 CollaboratorID	Number Range	 No examples available. 	Between 1 and 20	?	E
	Number Range	 No examples available. 	Between 1 and 20	?	E
2 ProductID	Homber Hunge				
2 ProductID 3 Discount	Number Range	 No examples available. 	Between 1 and 100	?	E
2 ProductID 3 Discount Order Table Column 4dd 1 Row(s)	Number Range Data Type	 No examples available. Examples 	Between 1 and 100 Options	? Help	De
2 ProductID 3 Discount Order Table Column Add 1 Row(s) XPORT TYPES CSV Excel HTM Database table name	Data Type Data Type JSON LDIF Progra	No examples available. Examples amming Language SQL XML Statement Type	Between 1 and 100 Options	 Help hide data format UPDATE 	Oeti
2 ProductID 3 Discount Drder Table Column Add 1 Row(s) XPORT TYPES CSV Excel HTM Database table name Database Type	Data Type Data Type UL JSON LDIF Progra Coupons MySQL	No examples available. Examples amming Language SQL XML Statement Type INSERT batch size	Between 1 and 100 Options Options INSERT IGNORE 100	 Help hide data format UPDATE 	De
2 ProductID 3 Discount Drder Table Column Add 1 Row(s) XPORT TYPES CSV Excel HTM Database table name Database Type Misc Options	Number Range Number Range Data Type IL JSON LDIF Progra coupons MySQL Include CREATE TABLE query	 No examples available. Examples amming Language SQL XML Statement Type INSERT batch size Primary Key 	Between 1 and 100 Options	 Help hide data format UPDATE 	opti
2 Productib 3 Discount Order Table Column Add 1 Row(s) XPORT TYPES CSV Excel HTM Database table name Database Type Misc Options	Number Range Number Range Data Type IL JSON LDIF Progra coupons MySQL Include CREATE TABLE query Include DROP TABLE query	No examples available. Examples amming Language SQL XML Statement Type INSERT batch size Primary Key	Between 1 and 100 Options Options INSERT IGNORE 100 None Add default auto-increment colum	- hide data format	opti

Imagen 28: Generar datos de cupones.

	1.4	1 / 11	,		•	• • •		• • •	
Para	ohtener		alle denerara	lae t	rancarcionae e	a introd	IIICA IA	s signinentes	valores.
i aia	ODICITO	ci couigo	que generara	103 1				3 Siguicrites	values.

Scherere	About News	Donate		
couponUse			SAVE	
DUNTRY-SPECIFIC D	DATA 💿			
Spain 🗙				
ATA SET 🔞				
Irder Table Colu	mn Data Type	Examples	Options	Help De
1 CouponID	Number Range	 No examples available. 	Between 1 and 200	?
2 UserID	Number Range	 No examples available. 	Between 1 and 1000	?
3 UseDate	Date	▼ MySQL datetime ▼	From: 04/19/2018 To: 04/19/202	0 7
			Format code: Y-m-d H:i:s	
4	Select Data Type	•		0
rder Table Colu	mn Data Type	Examples	Options	Help De
rder Table Colu dd 1 Row(s) KPORT TYPES O CSV Excel	mn Data Type	Examples Programming Language SQL XML	Options	- hide data format opti
Add 1 Row(s) Add 1 Row(s) CSV Excel Database table name	mn Data Type HTML JSON LDIF F couponUse	Examples Programming Language SQL XML Statement Type	Options	- hide data format opti UPDATE
Add 1 Row(s) Add 1 Row(s) CSV Excel Database table name Database Type	mn Data Type HTML JSON LDIF F couponUse MySQL	Examples Programming Language SQL XML Statement Type INSERT batch size	Options • INSERT O INSERT IGNORE (100	Help De
Add 1 Table Column Add 1 Row(s) CSV Excel Database table name Database Type Alisc Options	MN Data Type	Programming Language SQL XML Statement Type INSERT batch size uery Primary Key	Options Options INSERT IGNORE None	Help De
Add 1 Table Colu Add 1 Row(s) CSV Excel Database table name Database Type Aisc Options	MMN Data Type	Examples Programming Language SQL XML Statement Type INSERT batch size uery rry	Options Options INSERT IGNORE Add default auto-increment column	Help De

Imagen 29: Generar transacciones.

Tanto el código generado como el obtenido de generatedata.com se copia a un archivo nombrado insertarDatos.sql.

8.3. Procesos ETL con Pentaho Data Integration

En este anexo se indica a modo de ejemplo cómo se ha creado el archivo HechosCanje.ktr en la aplicación Pentaho Data Integration. El primer paso será crear la nueva transformación, esto lo podemos hacer desde el menú file, new, transformation:

File	Edit	View	Action	Tools	Help		
N	ew.				6	Transformation	CTRL-N
O	oen				CTRL-O	doL	CTRL-ALT-N
0	oen UR	L				Database Connection	AR I
0	ben Re	cent					
cl	ose All			SHIF	T-CTRL-W		
Im	port f	rom an	XML File		CTRL-I		
Ex	port						
Ex	it						WORK
							THO THE

Imagen 30: Nueva Transformación PDI.

O bien desde la pantalla de bienvenida, pulsando sobre la opción "New Transformation":



Imagen 31: Otra Nueva transformación.

En el área de trabajo iremos arrastrando los diferentes elementos desde la columna de elementos o design:



Imagen 32: Área de trabajo PDI.

Como vamos a obtener como origen de datos para este trabajo dos tablas de la base de datos Nutribox (coupons y couponUse) se arrastran dos objetos table input que configuraremos haciendo doble clic sobre ellos:

	_	Table in	put	
	Step name	couponUse ing	put	
	Connection	Nutribox	Edit New	Wizard
SQL			Get SQL select s	tatement
SELECT Transaction , CouponID , UserID , UseDate FROM couponU	nID se			
Line 1 Column 0		-		2.6
Enable lazy	conversion			
Replace variabl	es in script?	0		
insert dat	a from step			
Execute fo	r each row!			
	Limit size	0		
0		a Dreasans	and the second second	

Imagen 33: Detalle del objeto table input para la tabla couponUse.

8	Table in	put			8
Step name	coupons				
Connection	Nutribox		Edit	New	Wizard
SQL			Get SQL	select st	atement
SELECT CouponID , CollaboratorID , ProductID , Discount FROM coupons					
((C	1111				ac
Enable lazy conversion					
Replace variables in script?					
Insert data from step					
Execute for each row?					
Limit size	0				•
() Help Ок	Previe	w Ca	ancel		

Imagen 34: Detalle del objeto table input para la tabla coupons.

Para poder unir estas tablas deberemos previamente ordenar sus registros por el atributo CouponID. Para ello cada uno de los dos elementos Table Imput creados anteriormente se conectarán a elementos Sort rows que se configurarán de la siguiente forma:

			Sort rows		
		Step na	ame Sort rows		
		Sort direct	tory %%java.io.tmpdir%%		• Browse
		TMP-file pr	efix out		
	Sorts	ize (rows in mem	ory) [1000000		•
	Free me	mory threshold (i	n %)		0
Onl Fiel	y pass unique ro	Compress TMP Fi ows? (verifies key:	les? onl		0
•	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strengt
	Leogene				
10	Help		OK Cancel	Get Fields	

Imagen 35: Detalle del objeto Sort rows.

Cada uno de los elementos Sort rows (que han ordenado las tablas por CouponID se conectarán a un elemento Merge join que unirá las dos tablas (coupons y couponUse), para lo cual habrá que configurarlo de la siguiente forma:

		Merg	je jo	in	8
SI	tep name	Merg	ge jo	in	
F	irst Step:	Sort	rows	i	•
Sec	ond Step:	Sort	rows	; 2	•
2	Join Typ <mark>e</mark> :	RIGH	TOL	ITER	
Key	s for 1 <mark>st</mark> sl	tep:	Key	s for 2nd st	ep:
Ψ.	Key field		-	Key field	
0	Get key fie	lds		Get key fiel	ds
0	Help C	Ж		Cancel	

Imagen 36: Detalle del objeto Merge join.

Con la finalidad de crear dos campos, fechasku y horasku a partir del atributo Usedate, se usa un elemento Calculator que crea un campo UseDateString que será una copia del atributo Usedate:

Step name			11.40	1499	2-11-21				
Genera USeDateStri	ng								
g Throw an error o	n non existing files								
* New field	Calculation	Field A	. 1	Fi	Value type	Lengt	Pri	Remove	Conversion mask
1 Usebatestring	Lifeate a copy of field A	Usebate			Jering	1		14	ууууччы-од планта

Imagen 37: Detalle objeto Calculator

Y posteriormente un Split fields que genera fechasku y horasku a partir de UseDateString:

			Split	fields		8
	5	step na	me Separa str	ings de Fecha y ho	ora	
	Fie	d to sp Delimi Enclos	olit UseDateSt	ring		
ielo	ds					
•	New field	ID	Remove ID?	Туре	Length	Pre
1	fechasku		N	String		
2	horasku		N	String		
+ (111				

Imagen 38: Detalle objeto Split fields.

Posteriormente con un elemento String operations, se eliminan los caracteres especiales como ":" y espacios que originalmente tienen los atributos fechasku y horasku:

		21632 2			String op	erations	1022-0000				- (
			Step nam	transform	ndión						
The	fields to process:										
*	In stream field	Out stream field	Trim type	Lower/Upper	Padding	Pad char	Pad Length	InitCap	Escape	Digits	Remove Special character
1	fechasku		both	none	none			N	None	only	none
2	horasku		both	none	none			N	None	only	none

Imagen 39: Detalle objeto String operations.

Se conecta a otro elemento Calculator que a partir de fechasku y horasku, que son de tipo string, los nuevos campos fechaid y horasid del tipo Integer:

	Calcula	ator			3
name					:::
Throw an er	ror on non existing files				
is: New field	Calculation	Field A	Field B	Field C	Value ty
fechaid	Create a copy of field A	fechasku			Integer
horasid	Create a copy of field A	horasku			Integer
	0)				
	name culator Throw an er s: New field fechaid horasid	name culator Throw an error on non existing files s: New field Calculation fechaid Create a copy of field A horasid Create a copy of field A	name sulator Throw an error on non existing files s: New field Calculation Field A fechaid Create a copy of field A fechasku horasid Create a copy of field A horasku	name sulator Throw an error on non existing files s: New field Calculation Field A Field B fechaid Create a copy of field A fechasku horasid Create a copy of field A horasku	name sulator Throw an error on non existing files s: New field Calculation Field A Field B Field C fechaid Create a copy of field A fechasku A field B field C horasid Create a copy of field A horasku A field B field C

Imagen 40: Detalle objeto Calculator2.

Finalmente, con un elemento Combinatio lookup/update almacenará los atributos TransactionID, CouponID, UserID, Discount, fechaid, CollaboratorID, ProductID en la base de datos datawharehouse generando los atributos transactionSkU y versión, de forma que si estos datos se actualizan en la base de datos original en la base de datos datawharehouse se mantendrán los datos originales y se copiará la nueva versión:

		(Combination lookup/update	8
		Step name	Combination lookup/update	
		Connection	datawharehouse 💌 Edit New	Wizard
	Та	rget schema	• · · · · · · · · · · · · · · · · · · ·	Browse
		Target table	HerbosCapie	Proviso
		Commit size	Cache size loopo	DIOWSe
		comme size		
Key	fields (to look up r	ow in table):	Pre-toad the cache?	
	Dimension field	Field in stre	am	
1	TransactionID	Transaction	ND	
2	CouponID	CouponID		
3	UserID	UserID		
4	Discount	Discount		
5	fechaid	fechaid		
6	CollaboratorID	Collaborat	orID	
7	ProductID	ProductID		
	Techn	ical key field	transactionSkU	
		8.2400 8 .2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 - 2400 	Creation of technical key	
			Ose table maximum + 1	
			O Use auto increment field	
	Remove lo	okup fields?		
	Us	e hashcode?		
	Hashcode	field in table		
Dat	e of last update fie	d (optional)		1
6	Hala	OK	Concel Cot Sields 501	
0	Help	OK	Cancel Get Fields SQL	

Imagen 41: Detalle objeto Combination lookup/update.

Con estos elementos conectados como se ha indicado se obtiene el trabajo completo HechosCanje.ktr:





8.4. Definición del esquema OLAP

En este anexo se detalla el proceso de creación del esquema OLAP que se ha usado en el prototipo.

Para la creación del esquema el primer paso después de abrir la aplicación es definir la conexión a los datos. Para ello desde la aplicación vamos al menú "Options" y se pulsa en el apartado "Connection..." Como se muestra en la siguiente ventana:

	Schema Workbench	
<u>File Edit View Options Windows H</u> elp		

Imagen 43: Área de trabajo Pentaho Schema Workbench.

Para este caso será necesario establecer como tipo de conexión MySQL con la base de datos datawharehouse, alojada en localhost, usando el usuario admin, contraseña 12345678 y puerto número 3306; como se puede ver en la siguiente imagen:

	Database Connection		
General	Connection Name:		
Advanced Options	datawharehouse		
Pooling Clustering	Connection Type	Settings	
	MaxDB (SAP DB)	Host Name:	
	MonetDB	locahost	
	Native Mondrian	Database Name	
	Neoview	datawharehouse	
	Netezza	Port Number:	
	Oracle RDB	3308	
	Pentaho Data Services	User Name:	
	PostgreSQL Redebilt	ladmin 1	
	Remedy Action Request Sys	Password.	
	SQLite		
	Access	🕑 Use Result Streaming.	
	Mative (008C) ODBC JHOR		
	Test	OK Cancel	

Imagen 44: Detalle conexión con la base de datos.

Tras pulsar ok, volveremos a la ventana inicial y seleccionaremos crear un nuevo esququema a través de "file", "new", "schema" en el menú superior. Si pulsamos sobre el icono del schema podremos añadir cubos y dimensiones en el siguiente menú:

e Edit View Options Wi	idows <u>H</u> elp							
12								
Schema - New Schema3 (Se	hema3.xmD						1	d D
Q 人 ☆ ☆ wig un	e av 🔗	114 °4	₫.	9		2 %	1	Ey/
11				S	Dema		 	
Add cube Add Dimension	Attribu	ite				Value		
Add Named Set	teription		lew sche	maa				
Add User Defined Function	asuresCast	ien						
Add Virtual Cube	antumb							
Add Role Add Parameter								
Add Annotations								
Database - datawharehou	w (MySOL)							

Imagen 45: Área de trabajo Pentaho Schema Workbench 2.

Las dimensiones se pueden generar internamente en los cubos o como se ha hecho en este caso en la raíz del esquema de forma que puedan ser reutilizadas por diferentes cubos. En el esquema creado para este proyecto se crean las dimensiones Colaborador, Usuario, Tiempo y Producto, como se puede ver en la siguiente imagen:



Imagen 46: Dimensiones del cubo.

Para cada dimensión se debe definir las jerarquías, que se componen de los niveles y las tablas de datos. Así para la dimensión Colaborador, se crean el nivel colaborador y se define la tabla DimColaboradores. Como se puede observar en el diagrama anterior se pueden definir varios niveles de jerarquía; en este caso se han definido para tiempo los niveles Año y Mes, o en la dimensión producto los niveles Categoría y Producto. Esto permitirá filtrar la información por año y mes o agrupar por categorías y productos.

Para definir los niveles se debe introducir los valores en su tabla de atributos indicando el nombre del nivel, la columna donde se encuentra el índice que relaciona con la tabla de hechos el nombre de la columna, el tipo de datos devueltos, si son únicos los datos, el tipo de nivel o si es o no un nivel visible. La siguiente imagen es un ejemplo de como se ha definido el nivel Colaboradores:

 Lev 	vel for 'Colaborador' Hierarchy	٣
Attribute	Value]
name	Colaboradores	
description		
table		
column	CollaboratorID	
nameColumn	CollaboratorName	
parentColumn		
nullParentValue		
ordinalColumn		
type	String	
internalType		
uniqueMembers		
levelType	Regular	
hideMemberlf	Never	
approxRowCount		
caption		
captionColumn		
formatter		
visible		

Imagen 47: Detalle nivel colaborador.

Los niveles año y mes de la dimensión Tiempo se crean con un tipo TimeYears y TimeMonths respectivamente; mientras que el resto de los niveles definidos en este esquema, serán del tipo Regular.

Attribute	Value
name	Año
description	
table	
column	año
nameColumn	
parentColumn	
nullParentValue	
ordinalColumn	
type	Integer
internalType	
uniqueMembers	
levelType	TimeYears
hideMemberIf	Never
approxRowCount	
caption	
captionColumn	
formatter	
visible	

Imagen 48: Detalle nivel Año.

Level for 'Tiempo' Hierarchy		
Attribute	Value	
name	Mes	
description		
table		
column	mes	
nameColumn	mesText	
parentColumn		
nullParentValue		
ordinalColumn		
type	String	
internalType		
uniqueMembers		
levelType	TimeMonths	
hideMemberlf	Never	
approxRowCount		
caption		
captionColumn	mesText	
formatter		
visible		

Imagen 49: Detalle nivel Mes.

Tras definir las dimensiones, se define el cubo, que en este caso se denomina Nutribox para lo cual, sobre el schema se pulse el botón derecho del ratón y se selecciona la opción add cube. Tras nombrarlo añadiendo el valor en su tabla de atributos, hay que indicar su tabla de hechos, en este caso HechosCanje para lo cual hay que pulsar sobre el cubo con el botón derecho y pulsar sobre "ad table".

Además, se deberán indicar sus dimensiones para lo cual usaremos la opción "add dimensión usage" e introduciremos los atributos correspondientes:

Schema	Dimension Usage for 'Nutribox' Cube		
Nutribox	Attribute	Value	
	name	Usuario	
– 🏢 Table: HechosCanje	foreignKey	UserID	
aña Hauaria	source	Usuario	
- Usuanu	level		
– 📌 Tiempo	usagePrefix		
and Braducta	caption		
1	visible		

Imagen 50: Añadir uso de dimensión al cubo.

Por ejemplo, en el caso de la dimensión usuario (como se ve en la imagen anterior) indicaremos el nombre usuario, la foreignKey UserID, y source el nombre de la dimensión definida anteriormente, en este caso también "usuario". Una vez definidas las dimensiones del cubo, habrá que definir las medidas que se incluirán en el cubo, para ello sobre el cubo, tras hacer clic con el botón derecho del ratón, pulsaremos sobre la opción "Add measure" e introducir sus valores en su tabla de definición. Se puede ver en la siguiente imagen como se ha definido la medida "Número de colaboradores":

Measure for 'Nutribox' Cube		
Attribute	Value	
name	Número de colaboradores	
description		
aggregator	distinct-count	
column	CollaboratorID	
formatString		
datatype		
formatter		
caption		
visible	v	

Imagen 51: Añadir medida al cubo.

En esta tabla de definición se ha asignado un nombre en el atributo name, se indica el tipo de operación a realizar en el campo "aggregator", que en este caso es distinct-count y se indica la columna sobre la que se aplica la operación, en este caso Collaborator ID.

Para nombrar el esquema, sobre el icono del schema se puede asignar un nombre, en este caso Olap Nutribox a través de su tabla de atributos:




Una vez creado el esquema es necesario guardar el archivo en la opción "save as..."el menú file.

Para poder ser utilizado por el resto de las herramientas de la suite pentaho, este esquema hay que publicarlo, a través de la opción "Publish" del menú "file" e introduciendo la información del servidor pentaho en la ventana que aparece:

Publish Schema 🛛 😣
Pentaho Credentials Server URL:
http://localhost:8080/pentaho/
User:
admin
Password:
••••••
Publish Settings
Pentaho or JNDI Data Source:
datawharehouse
Register XMLA Data Source
☑ Remember these Settings
Publish Cancel

Imagen 53: Publicar el esquema.

Una vez introducidos los datos del servidor hay que pulsar en el botón "Publish" y una vez realizado el proceso se desplegará una ventana indicando que se ha publicado con éxito el esquema.

8.5. Análisis de los datos del cubo OLAP con jpivot

En este anexo se ejemplifica un análisis OLAP con jpivot a partir del cubo Olap Nutribox.

Browse File	es	Penta	ho Business Ar	alytics
Create Nev	v	Saiku Analytics]	
Manage Data Sc	ources	CDE Dashboard	and contribut	te with
Documentati	ion	JPivot View	e some ways ł do it.	NOW
Recents		Data Source		
CDE Otra2	\$			
COE Prueba1	Å			
home_dashbo	ard ☆	Doc	umentation	Forums

Imagen 54: Crear nuevo.

Tras validar en el servidor de pentaho business Analytics, entrando en <u>http://localhost:8080/pentaho</u> desde el navegador en el prototipo, al pulsar sobre el botón "Create New" y seleccionando JPivot View permitirá crear un análisis de los datos del cubo, para lo cual será necesario seleccionar el esquema y cubo entre los disponibles en el servidor:

Schema	
Olap Nutribo	ox 🗸
Cube	
Nutribox ~	1

Imagen 55: Crear nueva vista JPivot.

Pulsando sobre el botón del cubo 🔟 se pueden definir las medidas y las jerarquías o agrupaciones que se quieren analizar:

🗏 Columns 🛛 🕹
Measures
Rows
⊙ ⊙ ⊙ Usuario
© ⊙ ⊙⊙ Tiempo
OK Cancel

Imagen 56: Medidas y jerarquías.

Pulsando sobre el botón se desplegará un editor de MDX donde podremos observar o editar el código MDX de la consulta. Por ejemplo, en nuestro prototipo, para obtener el número de cupones canjeados por empresa colaboradora bastará con introducir la siguiente consulta:

SELECT

NON EMPTY {[Measures].[cuponesCanjeados]} ON COLUMNS, NON EMPTY {[Colaborador].[Colaboradores].Members} ON ROWS FROM [Nutribox]

Mostrará los siguientes resultados:

	Measures
Colaborador	⊙ cuponesCanjeados
Platanocanario S.L.	105
Agricultores De La Palma S.L.	396
Guanchito S.L.	217
Mojopicon S.L.	201
Cooperativa Palmera	318
Platanito S.L.	200
Papa Arrugada S. L.	259
Bienmesabe S.L.	197
Higopico S.L.	336
Buen Clima S.L.	217
Cancajos S.L.	308
Powerguanche S.L.	254
Benahorita S.L.	285
Tanausú S.L.	230
La Pedrada S.L.	286
Alguito S.L.	323
Galguen S.L.	123
Garafiano S.L.	251
Breñusco S.L.	279

Imagon	57.	Fiom	nla l	nivot	tahla
imayen	57.	Lieu	pio J	ρινοι	lavia.

Además, será posible obtener su gráfico al pulsar el botón 📧:

Slicer:



Imagen 58: Gráfico generado con Jpivot.

8.6. Análisis de los datos del cubo OLAP con saiku

En este anexo se ejemplifica un análisis OLAP con Saiku a partir del cubo Olap Nutribox.

File View Tools Help Home ~	
Browse Files	Pentaho Business Analytics
Create New	Saiku Analytics
Manage Data Sources	CDE Dashboard wiledge.
Documentation	JPIvot View e some ways how do it.
Recents	Data Source
Otra2	
🚥 Prueba1 😽	
Cor home_dashboard	Documentation Forums

Imagen 59: Crear nuevo.

Tras validar en el servidor de pentaho business Analytics, entrando en <u>http://localhost:8080/pentaho</u> desde el navegador en el prototipo, al pulsar sobre el botón "Create New" y seleccionando Saiku Analytics permitirá crear un análisis de los datos del cubo, para lo cual será necesario seleccionar el esquema y cubo entre los disponibles en el servidor:

Saiku Analytics 🛛 🛛	
Cubes	2
Nutribox	
Select a cube	
Olap Nutribox (Olap N	utribox)
Nutribox	
SampleData (SampleD	ata)
Quadrant Analysis	
SteelWheels (SteelWh	ieels)
SteelWheelsSales	

Imagen 60: Seleccionar cubo.

Una vez seleccionado el cubo sobre el que trabajar en la columna de la derecha aparecerán ordenadas todas las medidas y dimensiones del cubo.

Cubes 😤	👅 🛛 🖉	1 10		9	11	9	.8.	-14		2	10	(9)	R	12	0
Nutrios	Measures	٠	You ne	ed to include	at least	t one me	asure or	a level o	Colum	n for as	valid gue	ey. You r	eed to	include	at least to
Measures Add															
copones Carjendos Número de costorisidores Número de ossários	Columna	•													
Número de productos coporent/victos Mos/Corpo	Rows	٠													
Dimensions															
Colaborador (Al) Dolatoradore Profixito Profixito Tempe Chusen	Filter	•													

Imagen 61: Área de trabajo Saiku.

Arrastrando estos parámetros a las cajas Measures, Columns, Rows y Filter del área de trabajo, se generará una consulta que se mostrará los datos indicados. Por ejemplo, si queremos obtener el número de cupones canjeados por empresa colaboradora (ejemplo realizado anteriormente con jpivot) bastará con arrastrar cuponesCanjeados a la caja de Measures y Colaboradores a Rows:

Saiku Analytics			
Cubes 💈	😑 🗃 🗅 📈 🕨 💵	31 🖸 🗢 🕦	Q + #
Nutribox	Measures	Coluboradores	cupones Carienados
	cupontsCargendos	Pletanocanano S.L.	105
Measones Add		Agricultores De La Palma S.L.	396
*		Guanchito S.L.	217
cupones/Conjections		Mojopicon S.L.	201
Número de colaboradores	Columns	Cooperativa Patnera	318
Nomero de usuarios		Platanto S.L.	200
Nomero de productos		Papa Arrugada S. L.	259
Expenses Unices		Benmesabe S.L.	197
Millio Carriera	Rows	Higopico S.L.	336
Dimensions	Colaborador	Buen Clima S.L.	217
	Colaboradores	Cancajos S.L.	306
 Colaborador 		Powerguanche S.L.	254
(AI)		Benahorita S.L.	285
Coaboratores Directures	Filter	Tamausiú S.L.	230
Tiempo		La Pedrada S.L.	296
⊁ Usuario		Alguito S.L.	323
		Gaiguen S.L.	123
		Garatiano S.L.	251
		Breflusco S.L.	279
		Bimbache S.L.	315-

Imagen 62: Vista en tabla de consulta con Saiku.

Además, en la esquina superior izquierda, podemos alternar entre la



đb

:

vista de tabla o la gráfica en los botones

Imagen 63: Gráfica generada con Saiku.

8.7. Creación de cuadros de mando con Pentaho CDE Dashboard.

Tras acceder a Pentaho BI server se deberá pulsar sobre la opción Create New, CDE Dashboard:



Imagen 64: Crear nuevo CDE Dashboard.

Una vez dentro del área de trabajo hay que configurar la distribución del cuadro de mando. Esta distribución se debe definir en forma de tabla, añadiendo filas y columnas desde el "Layout Panel".



Imagen 65: CDE Dashboard, detalle "Layout Panel"

Desde este panel es posible añadir código html o javascript y se define el nombre del área que se usará como referencia para ubicar los componentes disponibles.

Desde el panel "Datasources Panel" se deben indicar las distintas fuentes de datos, que se usarán en los distintos componentes del cuadro de mando:

Witarib	Gatagournes	9049	Properties	
Community Data Access	tor.	Lare .	in the second se	1014
(egacy Datamarces	* Enug vals over mendrurgett	MOX.Queries conjectionQuery	Access Level	Fullic
NuSQL Detaiourcee	And over the other paid	Partial and a farmy	prob	dataw/tarehouse
DATASERVICES Quertes	Hate liver monthing of	Print in Annual Conny Print a Managery Connection of the Print	Alexandrum umama Query Recented	Olap Hearthese WiThis Millial SER (Two up of 1)
MDX Quartes	reds over mandriar(ndf	SelectorProDucy	Banded Weite	Compact
Ores drawns	robcover mondmergraft	CargeshonHGGuery	Califidated Columns	0
Compound Querres	indu siver mendmargo at	Tertifierclat/suentraNiCuely	Galaren	0
SCRIPTING Queries	+ Drug spinsraphit	TetalUsumer	Output Columns: Output Minde	0 Indute
RETTLE Queries			Cache Neyle	0
MOL Queries			CatheStanton	340
SQL Queries			20105	
VPATH Clusters				

Imagen 66: CDE Dashboard, detalle " Datasources Panel "

Una vez definidos los orígenes de los datos a representar en el cuadro de mandos (en el caso de este trabajo, diferentes consultas MDX al cubo OLAP Nutribox y SQL a la base de datos datawharehouse), habrá que definir los componentes desde el "Component Panel":

· Parameters	Components	1 7 0 R	Properties / Advanced	huperties	
. Educer	Tabe -	No.	Projecty.		
5. 2000 C	- Group	Chieth	Name:	CargesPorMESENant	
 Standard 	CCC Line Chieft	CarpioPorMESChart	Tile	1.	
Charts	CCC Pie Chart	carginsEartOvint	Anteners:	["Dempio"]	
	CCC Line Chart	Tendencial/suancoMChart	Parteneoers	[["Tempo", "Tempo"]]	
CENH'S	+ Grmg	Standard	Dutacourse	Canjin/ForMESQuary	
Legecy	Tuble Component	TablaColubiosador	insight.	300	
	Table Component	GouartouPregnoradore	With	900	
Scripti	Table Component	UnimosAttivos	Htm/Object	AbayoBuena	
Community Contributions	Table Component	Max	Base Avits Title		
	Table Component	TablaProductors	Ockable	∓alse	
	+ finnig	Parametors	Ork Action		
	Simple Parameter	Therepo	Compatibility Version	1	
	+ Group	Salarzy	Crossitab Motile	True	
	Select Component	SelemonArtiobelecture	Legend	∓akse	
			Ortho Aets Title		
			Series In Rows	∓ølse	
			Time Series	Faite	

Imagen 67: CDE Dashboard, detalle "Component Panel"

Una vez definidos los diferentes componentes, se ubican en las distintas áreas generadas en el "Layout Panel" con la propiedad HtmlObject.

Para finalizar es necesario guardar el trabajo en el servidor desde la opción Save as:

Save as	Ok	Cancel
Choose Format		
 Dashboard 	 Widget 	
Choose Folder		
🛅 home		
📮 public		
🗎 etc		
File Name*		
Insert Text		
insere lextin		
Title		
Insert Text		
Description		
Insert Text		

Imagen 68: CDE Dashboard, guardar cuadro de mando.

En esta nueva ventana indicaremos la ruta donde se guardará el archivo, su nombre, el titulo y una descripción.

8.8. Creación de informes con Pentaho Report Designer

En este anexo se describe la creación de un reporte usando la herramienta Pentaho Report Designer.



Imagen 69: Nuevo reporte.

Al ejecutar el programa, se muestra una pantalla de bienvenida que nos permitirá crear un nuevo archivo de reporting utilizando un asistente (Report Wizard) o desde cero (New Report), así como ver los archivos de ejemplo que contiene la aplicación y que se encuentran en la columna "Samples de la derecha":

En este caso se partirá del asistente, por lo que será necesario pulsar la opción "Report Wizard" o bien desde la barra de menú ir a file, Report Design Wizard ..:

		Pentaho Report Designer
<u>File Edit View Insert I</u>	ormat <u>D</u> ata	a <u>E</u> xtras <u>W</u> indow <u>H</u> elp
New	Ctrl-N	K 🖻 🛎 🗙 🔒
Report Design Wizard	Ctrl+Shift-N	
Open	Ctrl-O	
Open From Repository	. Ctrl+Shift-NumP	Pad/
Open <u>R</u> ecent		
Close Report	Ctrl-F	Welcome
Close all tabs		
Save	Ctrl-S	
Save As		
Publish	Ctrl+Shift-P	bentaho [°]
Export		A Hitachi Group Company
Preview		· · · · · · · · · · · · · · · · · · ·
Print		
Page Setup		
Print Preview		
Report Properties		
Configuration		
Resources		
Validate Report		
🗆 Work Offline		
Quit Report-Designer		eport New

Imagen 70: Crear nuevo reporte con el asistente.

El primer paso de este asistente será el denominado Look and Feel donde asignaremos una apariencia, a partir de una plantilla predefinida al seleccionar la opción Template o bien de un documento existente al elegir la opción Report Document. En este ejemplo se ha utilizado la apariencia predefinida Jade:

	Report Design W	Vizard 🛞
Look and Feel Data Source and Query Layout Step	Select Look and Feel Template CReport Document	
Format Step	Cobalt Nickel Jade Maple	
	Back	t Finish Cancel

Imagen 71: Seleccionar apariencia.

Tras pulsar en el botón Next se mostrará el paso siguiente donde se deberá indicar las fuentes de datos a partir de los cuales se generará el informe. Por defecto, se encuentran asignadas unas fuentes de datos de pruebas como se ve en la siguiente imagen:

	Report Design Wizard	8
Look and Feel	Data Source	
🕏 Data Source and Query	P-JDBC	
Layout Step	└─ Sample Query	
Format Step		
A		
~~~		
53		
	Back Next	Finish Cancel

Imagen 72: Fuentes de datos y consultas.

Se deberán eliminar estas fuentes de datos para lo cual hay que pulsar sobre su nombre y posteriormente sobre el botón an la esquina superior derecha. Tras este paso, será necesario añadir las nuevas

fuentes de datos, pulsando sobre el botón 😐

En la siguiente ventana será necesario escoger el tipo de origen de datos. En este caso se utilizará el cubo OLAP que ha sido previamente definido a través de Pentaho schema workbench y publicado en el servidor Pentaho Business Analytics, seleccionando la opción "Pentaho Analysis":

				1
Met	adata	(Cust a	m)	Ē
Mur	RGobi			
OLA	P4	800.020		
OLA	P4J (Cu	istam)		L
OLA	P4] (Cu	istam, I	Denarm	į
OLA	P4J (Cu	istum,	Legacy]	
OLA	P4] (Do	norma	lized)	
OLA	Phille	dech)		
Ope	nERP L	Data Ac	cess	ł
Pen	taho A	nalysis		ŧ
Pen	taho A	nalysis	Custo	٩.
Pen	taho A	nalysis	Custo	P
Pen	taho A	rialysis	Custo	ţ.
Pen	taho A	nalysis	[Denor	ĩ
Pen	taho A	nalysis	Tredac	1
Pen	taho D	ata int	egratio	ŧ.
Scri	ptable			
Seq	uence	Genera	ator	
Tab	e			
XML		17.0		3
4	1	-	1.	Ľ

Imagen 73: Tipo de conexión.

En la siguiente imagen se deberá indicar la ruta del esquema NutriboxSchema.xml a través de la opción "Pentaho Analysis Schema File:" y seleccionar la conección a la base de datos "datawharehouse" en el apartado "Connections:" cómo se puede ver en la siguiente imagen:

	Pen	taho Analysis Data I	iourra		
Data Source Global Scripting Pentaho Analysis Schema File:				2	
/home/hics/Documents/TFG/Nutribor5chema.xml				Browse	Properties
Logical Schema Name on Server:					
Olep Mutribde				Update Se	chema Name
Connections: Edit Security 🥢 🥥 😫	Available Que	ries			08
SampleData SampleData Diypersonic) SampleData (Locol) SampleData (Memory) SampleData (MySQL)					
dalawharobouse	Query Name:				
	Static Query	Query Scripting			11-
	Query:				
			Max Preview Rows		E Prestew
					OK Cancel

Imagen 74: Pentaho Analysis Data Source.

Sobre esta ventana, es necesario indicar las consultas para obtener los

datos para lo cual será necesario pulsar sobre el botón ¹ y modificar su nombre en "Query Name:". En este caso se le ha llamado porCat.

Available Queries:	• 🛛
porCat	
Query Name:	
porCat	
Static Quary Scripting	

Imagen 75: Consultas.

En el espacio asignado introduciremos la consulta necesaria para obtener los datos, como puede verse en la siguiente ventana:

	Pentaho Analysis Data Source	
Data Source Global Scripting Pentaho Analysis Schema File: Permahea Qarumenta/Po/Nutribo/Schema ami Logical Schema Name on Server: Olap Nutribo: Connections: Edit Security 2008 SampleData SampleData (Hypersonic)	Available Queries: perCat	Browse Properties Update Schema Name © 3
SampleData D.exal) SampleData (Memory) SampleData (MySQL) datawharehouse	Query Name:	
	Static Query         Query Scripting           guery:         select NON EMPTYORSTINCTY (Descendents)(Producte), (Al Productes), (P NON EMPTYORSTINCTY (Descendents))) on Columns           MONE EMPTYORSTINCTY (Measures), (cuponeoConjesdoe))) on Columns           from (Mutribox)	raclucta(.(Cobegona(J)))) on ROWS.
	🔛 Max Preview Rows	OK Cancel

Imagen 76: Detalle final de la conexión.

Pulsando sobre el botón "Preview", en la parte inferior derecha de la ventana podemos tener una vista previa de los datos que nos devuelve la consulta, que en este ejemplo se trata del número de canjes realizados por categoría de productos:

8	Preview	۲
[Producto]-[(AII)]	(Producto).[Categoria]	[Measures]. [cuponesCanjeados]
All Productos	Fruta Eco	346
All Productos	Harinas	116
All Productos	Salaas	119
All Productos	Infusiones	442
All Productos	Verduras Eco	221
All Productos	Bebidas Alcoholicas	244
All Productos	Cosmeticos	259
All Productos	Carnes	244
All Fraductos	Pescaderia	405
11 77		Close

Imagen 77: Previsualización de la consulta

Una vez introducidas las consultas necesarias (en este caso inicialmente sólo es necesaria esta) serán visibles en el "Data Source" del asistente.

	Report Design Wizard	8
Look and Feel	Data Source	/ 🕂 🛛
🕏 Data Source and Query	Pentaho Analysis datawharehouse - Olap Nutribox (/home/nico,	/Documents/TFG/N
Layout Step	porCat	
rumai step		
5		
\$		
A		
	•	► ►
	Back Next Finis	h Cancel

Imagen 78: Conexión creada en el asistente.

Tras pulsar en Next agregaremos los campos a mostrar en la tabla del informe. En este caso las categorías ([Producto].[Categoria]) y las cantidades ([Measures].[cuponesCanjeados]), sin ningún tipo de agrupación:



Imagen 79: Selector de campos.

Report Preview 😣			
Report Export View Help			
	<b>^</b>		
	May 29, 2019 @ 04:18		
Report Title			
Sub Title 1			
SubTitle 2			
[Producto].[Categoria]	[Measures].[cuponesCanieados]		
Fruta Eco	346		
Harinas	116		
Salsas	119		
Infusiones	442		
Verduras Eco	221		
Bebidas Alcoholicas	244		
Cosmeticos	259		
Carnes	244		
Pescadería	405		
Report	Footer		
•	1/1		

### Pulsando en "Preview" podemos obtener una vista previa del informe:

#### Imagen 80: Previsualización del informe.

En el siguiente paso podemos indicar el formato, en este caso damos al nombre de categoría un 80% del espacio y alineación centrada para las cantidades:

	Report Design Wizard	8
Look and Feel Data Source and Query Layout Step Format Step	Groups:  Details:  [Producto].[Categoria] [Measures].[cuponesCanje ]	Format Display Name: [Producto].[Categoria] Alignment: Data Format: Data Format: None Width %: 80 Auto Width Aggregation: None Distinct Only Preview
	Back Next	Finish Cancel

Imagen 81: Ajuste de formato del informe.

Tras pulsar en "Preview" se muestra el siguiente resultado:

ort Export View Help	
B ≪ < > >> 0, 9, 100% ▼	
	Hay 29, 2014 (6 OF JR
Report Title	
Sub Title 1	
Sub Title 2	
Producto).(Categoria)	(Measures][cupon esCanieados]
inda Kon	346
Harinas	110
lation	119
nhialanes	442
Antiburas Essi	221
bibidas Aktoholitai.	244
Cournelieus	259
ares .	244
Festivatie fill	405
Report Feater	

Imagen 82: Segunda previsualización del informe.

Tras cerrar la previsualización y pulsar en "Finish" en el asistente se cargarán estos datos en el área de trabajo de la aplicación:



Imagen 83: Edición desde el área del trabajo.

Se edita el título del documento, así como las cabeceras de las tablas y

podemos visualizar una previsualización pulsando sobre el botón by luego la opción Print Preview:

Print P	review
ort Export View Help	
윤 《《 > 3》 역, 역, 100% -	
	1
	Hay 29, 2019 (0.05-31
Canies realizados por categorí	a
ategoria	Cantidad
nda Eco	316
larinas.	316
alsas	119
fusiones	442
erduras Esp	223
ebidas Alcoholicas	244
asmeticas	259
ames	244
escade na	405

Imagen 84: Tercera Previsualización.

Para una mayor comprensión de los datos se añadirá un gráfico de tipo pastel, para lo cual se selecciona la posición donde irá ubicado (en este

caso Group Header) y se pulsará sobre el botón . Se incluirá un gráfico en esta posición que se podrá configurar haciendo doble clic sobre él:

				Edit Chart		
		18) (AT		) 💽 🔛 🚛 🔀		1
H				Primary DataSource Socondar	y barasburge	
Pie Chart (Imag	ie)					
Name	Value	Espt	P	ie DataSet Collector (Chart Data	10	
Required	and a state of the	0000		Name	Value	_
gnore-nulls	True	0	4 1 1	Common		
gnore-seros	True		- 00	lue-column	(Measures). (cuponesCanieados)	
ng-data-messa	CHART,USER N			Series		
1000		1000	- Sa	mes-by-field	[[Producto].]Categoria]]	_
thart-title			- 35	uto-generate-series		
risert-ticke-neid	Record and	<u> </u>	- 10	Group		
tile-ront	Sanssent-B0		9	oup-by		_
e options	Collegent and the		1.12	set-group		
sice-colors	Cartoout area		-	Optional	24	
show-tacters	(deradit)	<u> </u>	-	osstab-column-filter		_
aberront	20052018-0		-			
etete disclosite	Train	<u>×</u>				
malade close	True	- X-				
and and and						
- Constant	and the second second	No.				
3-D	False					
a-color		1				
ad-image	1.00	1				
show-border	False	6				
parder-color	LOUGH CAL	6				
ants-slins	True	0				
slot-bg-calor		0				
plot-fg-alpha	1.0	0				
slot-bg-alpha	1.0	0				
training with the		10.0				

Imagen 85: Insertar gráfico.

Se obtiene el siguiente resultado:

port Export View Help	
Canjes realizados por catego	pría
Street State	Cartillol
Foreitze	944
Te Date	124
Notice -	-128
The second	
Sector of the	
The second	
Canada	
Prove and the	



Para que este informe se genere para cada año y poder analizar cada ejercicio individualmente es necesario crear un parámetro, que se basará en los diferentes años con canjes registrados en nuestro cubo, por lo que será necesario añadir otra query a las fuentes de datos:



Imagen 87: Añadir consulta.

Posteriormente, para crear el parámetro vamos a la pestaña data al lado derecho del área de trabajo, pulsar en "Add parameter...":



Imagen 88: Añadir parámetro.

Introduciendo los siguientes valores:

	6	dit Parameter	۹
DataSources 🛛 🗸 🗐			
Pentaho Analysis (datawharehouse - Olapi	Name	Tiempo	
porCut	Label	Tiempo	
- D Selectoráño	Label formula		(m)
	Value Type	String	
	Data Format		
	Data format formula		107
	Default Value	2019	
	Default Value Formula		
	Post-Processing Formula		and a
		🔲 Mandatory 🛄 Hidden	
	Prompt		
	Display Type	Drop Down	
	Query Value	SelectorAño	-
		[Tiempo].[Año]	•
	Display Name	[Tiempo].[Año]	
	Display Value Formula		(en)
		Validate Values     Reset to default value on invalid selections     Use first value if default value formula results in NA.	
• <u> </u>			OK Cancel

Imagen 89: Editar parámetro.

Finalmente hay que cambiar la consulta porCat que servía como fuente de datos para que sólo muestre los datos del año seleccionado:

Data Source       Global Scripting         Pentaha Analysis Schema IIIe:       Browse       Proper         Analysis Schema IIIe:       Browse       Proper         Olso Nuribos       Update Schema Na       Update Schema Na         Connections:       Idit Security			Pentaho Analysis Data Source	
Pentaho Analysis Schema File: home/ricu/Datumenta/TEGNuthes/chema.uml Logical Schema frame on Server: Update Schema Nar Connections: Idit Security A wailable Queries: SampleData Otypersonici SampleData (Nypersonici SampleData (Memory) SampleData (MySQL) datawharehouse Query Name: Static Query (Query Scripting Decry: Select ONI EMPTY( (Descendants/(Producto).(Categorial))) on ROWS, NOTE EMPTY( (Descendants/(Producto).(Categorial))) on ROWS, NOTE EMPTY( (Messure). [cuponesCarepedos])) on Columns MeRPE ((Thempa).(1(Thempa).))	Data Source Global Scripting			
Anomenical Dacumenta TréAllubritos-Schema.ami       Browse       Proper         Logical Schema Name on Server:       Update Schema Name         Connections:       Idit Security       Idit Security       Idit Security       Idit Security         SampleData       SampleData (hypersonic)       SampleData (hypersonic)       SelectorAño       Idit Security       Idit Security         SampleData (hypersonic)       SampleData (hypersonic)       SelectorAño       Idit Security       Idit Securi	Pentaho Analysis Schema File:			
Logical Schema Name on Server: Update Schema Name Connections: Idit Security	home/nicg/Documents/TFG/NutribakSc	hema.xml		Browse Properties
Olap Murrison       Update Schema Nar         Connections:       Idit Security       Image: Connections:       Idit Security       Image: Connections:         SampleData       SimpleData (trypersonic)       SampleData (tool)       SelectorMo         SampleData (tool)       SampleData (tool)       SelectorMo         SampleData (tool)       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:       Connections:       Connections:         SampleData (tool)       Connections:       Connections:       Connections:       Connections:       Connections:       Connections:         Connections:       Connections:       Connections:       Connections:       Connections:       Connections:         Connections:       Connections:       Connections:       Connections:       Connections:       C	Logical Schema Name on Server:			
Connections: Idit Security V V V V V V V V V V V V V V V V V V V	Olap Nutribox			Update Schema Name
SampleData SampleData (hypersonic) SampleData (boca) SampleData (Memory) SampleData (Memory) dotawharehouse Query Name: Static Query Query Scripting Defet NON EMPTY: [Nessural, [cuppresCargeados]) on Columns NUM EMPTY: [Nessural, [cuppresCargeados]) on Co	Connections: Edit Security	/ 🔍 🖬	Available Queries:	08
Static Query Query Scripting Query: Solicit MON EMPTY( (Descendants()Producto). (All Producto). (Categona())}) on ROWS, WOR Mathema MeERE ((Tiempo). [4(Tiempo).]) MeERE ((Tiempo). [4(Tiempo).])	SampleData SampleData (Hypersonic) SampleData (Local) SampleData (Memory) SampleData (MySQL) datawharehouse		perCat SelectorAño Ouery Name:	
Static Query Query Scripting  Ouery:  Select NON EMPTY( [Descendants(]Producto].[All Producto].[Categoina])}) on ROWS,  Non. Mutadons  WeERE ([Tempo].[4[Tempo]])  Metere (Tempo].[4[Tempo]])  Provide Rows			A CONTRACTOR OF	
Query: select NON-EMPTY({Descendants(}Producto].(Al Producto].(Categona())}) on ROWS, Non-Building MHERE ((Tiempo).[+{Tiempo}])) MHERE ((Tiempo).[+{Tiempo}])			Static Query Query Scripting	
🗌 Max Preview Rows 👘 Prev			select NON EMPTY (Descendants(Producto), (AF Fraductos), (Produ NON EMPTY ([Nessuran], [cuponesCargesdas]]) on Columns Trans (Nutritional NesERE ((Tiempo), [4(Tiempo)]))	cto].(Categoria())) on ROWS,
			🗌 Max Preview Rows	Proview

Imagen 90: Modificar consulta anterior.



Se modifica la cabecera para que indique el año a valorar a partir del parámetro creado:

Imagen 91: Se ha modificado la cabecera.



Así se incluirá un selector y mostrará los datos del año indicado:

Imagen 92: Informe final Vista del año 2018.

	por categoria y año - Print Preview	
port Export View Help		
B << < > > < < 100%	•	
empo 2019 👻		
Auto-Update on selection		Update
		1
Canies realizados por cal	regoría el año 2019	
10%		
12%)- 11%		
12% 11% 9 Fruita Eco @ Harinas @ 9 Venturas Eco @ Beisicia 6 Carnex @ Pescateria	Satives	
E3%) E1% • fride too @ merries @ • Venturas too @ technia • Carries • Pescateria	Safuar O Informer Anchoncar O Conneteon Cantidad	
Estegoría Fruta Eco	Safkas @ Intereses Archerican @ Connetecon Caritidad 134	
Estegoria Fruta Eco Haritas	Safkas @ Intropres Archoloxal @ Ceantidad 154 52	
Entergonia Prota Eco Prota Eco Prota Eco Prota Eco Harrinas Salsa s	Carticlad 15-4 5-2 15-4 5-2 5-6	
Eatesportia Pruta Eco Pruta Eco Pruta Eco Pruta Eco Pruta Eco Harinas Salzais Infusiones	Satistica de la constitución Alcohorcas e Interves Alcohorcas e Constitución 154 52 56 239	
Categoria Prida Eco Prida Eco Categoria Prida Eco Prida Eco Harinas Salsas Missiones Venduras Eco	Salase © Rhysiones Arcohorcas © Connetican Arcohorcas © Connetican 154 52 56 219 139	
Categoria Frida Eco © Meimas © Carees © Pescateria Categoria Frida Eco Harinas Salsas Infusiones Verduras Eco Bebidas Alcoholicas	Salas: © Interiore: Alcohorcia: © Caritidad ID4 52 56 219 119 127	
Categoria Frida Eco © Merinas Categoria Frida Eco Printa Eco Harinas Salsas Inflationes Verduras Eco Bobidas Alcoholicas Cosmeticos	Sature ● Interiore           Sature ● Interiore           Arcohorciar ● Connections           ID-4           S2           56           21.9           11.9           12.7           13.3	
Categoria Frida Eco © Merres © Verduras Eco © biolota Carnes Precaderia Verduras Eco Harinas Salsas Infusiones Verduras Eco Rebidas Alcoholicas Cosmeticos Carnes	Cantidad           3.4km         Mitronesc.           3.4km/dxxxx         Constidad           10.4         52           56         21.9           11.9         32.7           13.3         13.4	

Imagen 93: Informe final Vista del año 2019.

Finalmente, sólo queda guardarlo, desde el menú "file", "save as…" o publicarlo en el servidor BI de pentaho desde la opción "Publish…":

<u>File E</u> dit <u>V</u> iew Insert F	<u>o</u> rmat <u>D</u> ata	Extras	<u>W</u> indow	<u>H</u> elp
New	Ctrl-N	6	à 🛝 🗙	
Report Design Wizard	Ctrl+Shift-N			
Open	Ctrl-O			
Open From Repository	Ctrl+Shift-NumPa	d/ 🗾 18	3 🔽	B I 1
Open <u>R</u> ecent		• 1.5	5 2.0	2.5
Close Report	Ctrl-F			
Close all tabs				
Save	Ctrl-S	P a	lizad	dos r
Save As		- 4	1124	1001
Publish	Ctrl+Shift-P			
Export				
Preview		•		
Print				
Page Setup Dript Droview				
Print Preview				
Report Properties				
Resources			_	
Validata Banart				
				Pa
Work Offline				
Quit Report-Designer				

Imagen 94: Publicar el informe.

En este caso se publicará indicando los siguientes datos de conexión:

Login		8
Server URL:		
http://localhost:8080/pental	ho	-
Timeout:		
	3	0 -
User: admin		
Password:		
Remember These Settings		
	ок с	Cancel
Imagan 05: Datallaa da l		vián

Imagen 95: Detalles de la conexión.

Tras esto el servidor solicitará los datos de publicación:

Sie Namo:		
PorCial.prpt		
fitle:		
Carges reakindor	ppr tategoria y año	
lepart Descript	ion	
Canjes realizados Location:	por tategoria y año	1
public		
Title	File Name Date Hodfied	Description
dugn-sampl. Steel Wheels	. plugin-samples Steel Wheels	Reporting, Anal
🖌 Show Hidder Output Type:	1 Files	
HTML (Stream)	- Lock	
		OK Cancel

Imagen 96: Datos de la publicación.

Y una vez publicado nos mostrará su confirmación:



Imagen 97: Confirmación reporte publicado.