



Cuadro de Mando de las Naciones Unidas

Quero Hijano, Óscar Grado en Ingeniería Informática

Andrés Sanz, Humberto

Fecha Entrega



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

FICHA DEL TRABAJO FINAL

Título del trabajo:	Cuadro de Mando de las Naciones Unidas
Nombre del autor:	Quero Hijano, Óscar
Nombre del consultor:	Andrés Sanz, Humberto
Fecha de entrega (mm/aaaa):	06/2023
Área del Trabajo Final:	Business Intelligence
Titulación:	Grado en Ingeniería Informática

Resumen del Trabajo (máximo 250 palabras):

En este documento se detalla todo el proceso de integración de una serie de datos de las Naciones Unidas, detallados en el alcance, y la elaboración de un Cuadro de Mando analítico, para permitir a cualquier usuario entender la realidad de un país, grupo de países, región o el mundo.

A nivel integración, se ha diseñado una estructura de datos a tres capas (*Staging*, Harmonizada y Modelada), con ficheros .QVD y .CSV en cada capa, para poder controlar fácilmente todo el proceso. Además, el poder disponer de los archivos CSV de la capa modelada permite utilizarlos en otra aplicación analítica si se quisiera.

A través de diferentes análisis, tanto a nivel general con *KPIs* globales, como a nivel detallado, con análisis en detalle de cada tipología de información, el usuario podrá "descubrir" la situación actual y las tendencias del área de interés, y poder tomar buenas decisiones.

Los datos cargados se han acotado al periodo 2011-2020, para asegurar que se disponen de datos en este periodo para todas las tipologías de información. Si alguien quisiera ampliar el periodo, bastaría con generar los ficheros origen con el periodo deseado. Además, todos los datos incluidos en el modelo han sido depurados y harmonizados, para poder ofrecer una visión única de la verdad.

Abstract (in English, 250 words or less):

This document details the entire process of integrating United Nations' data, detailed in scope, and the development of an analytical Dashboard, to allow any user to understand the reality of a country, group of countries., region or the world.

At the integration level, a three-layer data structure (Staging, Harmonized and Modeled) has been designed, with .QVD and .CSV files in each layer, in order

to easily control the entire process. In addition, being able to have the CSV files of the modeled layer makes it possible to use them in another analytical application if desired.

Through different analyses, at a general level, with global KPIs, and at a detailed level, with detailed analysis of each type of information, the user will be able to "discover" the current situation and the trends of the area of interest and be able to take good decisions.

The data loaded has been limited to the period 2011-2020, to ensure that data is available in this period for all types of information. If someone would like to extend the period, it would be enough to generate the source files with the desired period. In addition, all the data included in the model have been refined and harmonized, in order to offer a unique vision of the truth.

Palabras clave (entre 4 y 8):

Naciones Unidas, Qlikview, Indicadores, Auto-servicio.

Índice

1. Introducción	
1.1. Contexto y justificación del Trabajo	1
1.2. Objetivos	2
1.2.1. Objetivo Principal	
1.2.2. Objetivos Secundarios	
1.3. Alcance	
1.4. Enfoque y método seguido	
1.5. Planificación del Trabajo	
1.6. Recursos	
1.7. Costes del proyecto	
1.8. Riesgos identificados	
1.9. Breve sumario de productos obtenidos	
1.10. Breve descripción de los otros capítulos de la memoria	
2. La solución analítica	
2.1. Selección de las herramientas	
2.2. Construcción del modelo de datos	
2.2.1. Variables	
2.2.2. Mappings	
2.2.2.1. Map_Countries	
2.2.2.1.1. Descripción	
2.2.2.1.2. Script	
2.2.2.2. Map_Sex	
2.2.2.2.1. Descripción	
2.2.2.2.2. Script	
2.2.2.3. Map_AgeRange	
2.2.2.3.1. Descripcion	
2.2.2.4. Map_AgeTipo	
2.2.2.4.1 Map_Age ripo	
2.2.2.4.1. Descripcion	
2.2.2.5. Map Divisa Euro	
2.2.2.5.1. Descripción	
2.2.2.5.2. Script	
2.2.3. Capa Staging (SA)	22
2.2.3.1. Población	
2.2.3.1.1. Descripción	
2.2.3.1.2. Script	
2.2.3.2. Muertes	
2.2.3.2.1. Descripción	
2.2.3.2.2. Script	
2.2.3.3. Homicidios	
2.2.3.3.1. Descripción	
2.2.3.3.2. Script	
2.2.3.4. Contaminación	
2.2.3.4.1 Descripción	23

2.2.3.4.2.	Script	. 23
2.2.3.5.	Energía	. 23
2.2.3.5.1.	Descripción	. 23
2.2.3.5.2.	Script	. 23
2.2.3.6.	GDP	. 23
2.2.3.6.1.	Descripción	. 23
2.2.3.6.2.	Script	. 23
2.2.3.7.	GNI	. 23
2.2.3.7.1.	Descripción	. 23
2.2.3.7.2.	Script	. 23
2.2.3.8.	Gasto	. 23
2.2.3.8.1.	Descripción	. 24
2.2.3.8.2.	Script	. 24
2.2.3.9.	WDI	. 24
2.2.3.9.1.	Descripción	
2.2.3.9.2.	Script	
2.2.3.10.	Refugiados	
2.2.3.10.1.	Descripción	
2.2.3.10.2.	Script	
	apa Harmonizada (HA)	
2.2.4.1.	Población	
2.2.4.1.1.	Descripción	
2.2.4.1.2.	Script	
2.2.4.2.	Muertes	
2.2.4.2.1.	Descripción	
2.2.4.2.2.	Script	
2.2.4.3.	Homicidios	
2.2.4.3.1.	Descripción	
2.2.4.3.2.	Script	
2.2.4.4.	Contaminación	
2.2.4.4.1.	Descripción	
2.2.4.4.1.	Script	
2.2.4.5.	!	. 20 . 26
2.2.4.5. 2.2.4.5.1.	Descripción	
2.2.4.5.1.	Script	
2.2.4.6.	GDP	
2.2.4.6.1.	Descripción	
2.2.4.6.1.	Script	
2.2.4.0.2. 2.2.4.7.	GNI	
2.2.4.7. 2.2.4.7.1.	Descripción	
2.2.4.7.1. 2.2.4.7.2.	·	
	Script	
2.2.4.8.	Gasto	
2.2.4.8.1.	Descripción	
2.2.4.8.2.	Script	
2.2.4.9.	WDI	
2.2.4.9.1.	Descripción	
2.2.4.9.2.	Script	
2.2.4.10.	Refugiados	
2.2.4.10.1.	Descripción	
2.2.4.10.2.	Script	. 27

2.2.5. Ca	apa Analítica (MA)	27
2.2.5.1.	M_Paises	27
2.2.5.1.1.	Descripción	27
2.2.5.1.2.	Script	27
2.2.5.2.	M_Sexo	27
2.2.5.2.1.	Descripción	28
2.2.5.2.2.	Script	28
2.2.5.3.	F_Datos	28
2.2.5.3.1.	Descripción	28
2.2.5.3.2.	Script	28
2.3. Mode	elo de Datos	28
2.4. Herra	amienta Analítica	29
2.4.1. Dii	mensiones	29
2.4.1.1.	Año	29
	Edad	
	Región	
	País	
	Sexo	
	étricas	
2.4.2.1.	Habitantes	
2.4.2.2.	Muertes	
2.4.2.3.	Homicidios	
_	Refugiados	
	Contaminación	
2.4.2.6.	PIB	
2.4.2.6.1.	PIB total	
2.4.2.6.2.	PIB per cápita	
2.4.2.7.	GNI	
2.4.2.7.1.	GNI total	
2.4.2.7.1. 2.4.2.7.2.	GNI per cápita	
2.4.2.7.2. 2.4.2.8.	Gasto	
_	Indicadores de desarrollo	
2.4.2.9. 2.4.2.9.1.		31
2.4.2.9.1. 2.4.2.9.2.		
2.4.2.9.2. 2.4.2.9.3.	Transporte Aéreo (Pasajeros)	
2.4.2.9.3. 2.4.2.9.4.	Tráfico contenedores (Puertos)Igualdad de genero	
2.4.2.9.4. 2.4.2.9.5.	Calidad Administración Pública	
2.4.2.9.6.	Protección Social	
2.4.2.9.7.	Transparencia Sector Público	
2.4.2.10.	Energía	32
	Plantas de energía	
	erfaz del Cuadro de Mando	
2.4.3.1.	Introducción	
	Disponibilidad información	
2.4.3.3.	KPIs	
	Análisis en Detalle	
2.4.3.4.1.	Población	
2.4.3.4.2.	Muertes	
2.4.3.4.3.	Homicidios	
2.4.3.4.4.	Contaminación	40

	2.4.3.	4.5.	Gasto			 	41
	2.4.3.	4.6.	Energía			 	42
	2.4.3.	4.7.	Indicadores D	Desarrollo: Tra	ansporte	 	43
	2.4.3.	.4.8.	Indicadores D	Desarrollo: Ra	tings	 	43
	2.4.3.	4.9.	Refugiados			 	45
	2.4.3.						
	2.5.	Plan de	pruebas			 	48
_							
			•				
	3.2.	Objetivo	os conseguido	S		 	51
		•					
4.							
5.		_					
6.							
		_					
					Unidas.pod		
			_ ,				
	6.4. 6.5.	Archivo			Qlikview		
					QIIKVIEW		
	6.6.				le datos		
	6.7.						
	6.8.				ebas.pdf)		
	6.9.				io.pdf)		
	6.10.						

Lista de figuras

Figura 1 - Algunos Datasets disponibles en el repositorio de las	Naciones
Unidas	
Figura 2 - Análisis inicial de las fuentes de datos	
Figura 3 - Análisis en profundidad y selección final	
Figura 4 - Estructura de Descomposición del Trabajo (EDT)	7
Figura 5 - Planificación con Projectlibre (1/2)	
Figura 6 - Planificación con Projectlibre (2/2)	11
Figura 7 - Diagrama de Gantt (1/3)	
Figura 8 - Diagrama de Gantt (2/3)	
Figura 9 - Diagrama de Gantt (3/3)	
Figura 10 - Hitos del proyecto	
Figura 11 - Dedicación y coste por roles	13
Figura 12 - Tabla detallada riesgos	
Figura 13 - Mapa de riesgos visual	
Figura 14 - Cuadrante Mágico de Gartner de herramientas de Integrac	
Figura 15 - Análisis herramientas integración	
Figura 16 - Cuadrante Mágico de Gartner de herramientas de	Business
Intelligence	
Figura 17 - Análisis herramientas Business Intelligence	
Figura 18 - Modelo de Integración	
Figura 19 - Modelo Analítico Cuadro de Mando	28
Figura 20 - Página Introducción del Cuadro de Mando	
Figura 21 - Página Disponibilidad Información del Cuadro de Mando	
Figura 22 - Filtrado tabla disponibilidad información	34
Figura 23 - Página KPIs del Cuadro de Mando	34
Figura 24 - Selectores años	35
Figura 25 - Funcionamiento "No Disponible"	35
Figura 26 – botonera	35
Figura 27 – Filtros	35
Figura 28 - Selectores años	36
Figura 29 - Población: Indicadores	36
Figura 30 - Población: Evolución	36
Figura 31 - Población: Distribución por sexo	37
Figura 32 - Población: Distribución por sexo y año	37
Figura 33 - Muertes: Indicadores	38
Figura 34 - Muertes: Evolución	38
Figura 35 - Muertes: Distribución por sexo	38
Figura 36 - Muertes: Distribución por sexo y año	39
Figura 37 - Homicidios: indicadores	39
Figura 38 - Homicidios: evolución	39
Figura 39 - Homicidios: comparación homicidios y muertes	40
Figura 40 - Homicidios: comparación homicidios y población	
Figura 41 - Contaminación: indicadores	40
Figura 42 - Contaminación: evolución anual	41
Figura 43 . Contaminación: peso contaminantes	
Figura 44 - Contaminación: peso contaminante respecto al total	

Figura 45 - Gasto: indicadores análisis	. 41
Figura 46 - Gasto: análisis evolutivo	. 42
Figura 47 - Energía: indicadores	. 42
Figura 48 - Energía: evolución anual energías	. 42
Figura 49 - Energía: evolución anual energía concreta	. 42
Figura 50 - Energía: evolución anual capacidad plantas energía	
Figura 51 - Indicadores desarrollo transporte: indicadores	
Figura 52 - Indicadores desarrollo transporte: evolución	. 43
Figura 53 - Indicadores desarrollo (ratings): indicadores	
Figura 54 - Indicadores desarrollo (ratings): Top 10 países	. 44
Figura 55 - Indicadores desarrollo (ratings): Evolución	. 44
Figura 56 - Refugiados: indicadores	. 45
Figura 57 - Refugiados: evolución anual	. 45
Figura 58 - Refugiados: Top 10 países procedencia	. 45
Figura 59 - Refugiados: desglose por país de procedencia (anual)	. 46
Figura 60 - Extractos: dimensiones y métricas	
Figura 61 - Extractos: filtros	. 47
Figura 62 - Extractos: generación tabla	. 47
Figura 63 - Extractos: exportar tabla a Excel	. 48
Figura 64 - Extractos: selecciones realizadas	. 48
Figura 65 - Anexo: Source_files.xlsx	. 55
Figura 66 - Anexo: Plan de Proyecto	. 55
Figura 67 - Archivo Qlikview para generar las diferentes capas de datos	. 56
Figura 68 - Archivo .Zip con las capas de datos	. 56
Figura 69 - Instalador Qlikview	. 57
Figura 70 - Anexo scripts modelo de datos	
Figura 71 - Cuadro de Mando de las Naciones Unidas	. 58
Figura 72 - Plan de Pruebas	. 58
Figura 73 - Manual de usuario	
Figura 74 - Video tour	. 60
Figura 75 - Presentación Defensa	. 60
Figura 76 - Video Defensa	. 61

1. Introducción

1.1. Contexto y justificación del Trabajo

Cada minuto nuevos negocios abren en el mundo y la gran mayoría de estos emprendedores carecen de información "fiable" con la que poder tomar decisiones tanto para la primera apertura como cuando quieren internacionalizar sus servicios.

Las compañías de nueva creación y los emprendedores necesitan conocer de forma ágil información diversa de las posibles regiones y países, de cara a definir su Plan de Negocio. Disponer de un sistema ágil de consulta de esta información les permitirá identificar dónde puede haber una oportunidad de mercado.

Uno de los problemas a los que se enfrentan muchos emprendedores es el coste de obtener esta información y mantenerla actualizada, ya que la adquisición de información y análisis ad-hoc suele ser muy cara y supone un esfuerzo recurrente. Por tanto, el único camino posible para este grupo de personas es acudir a fuentes públicas

En el mundo hay mucha información pública (Open Data) y uno de los mayores repositorios de datos existente con perfil global es la web de las Naciones Unidas (http://data.un.org/Explorer.aspx). En este sitio se concentra información de múltiples fuentes, que tratan sobre diferentes áreas, y con estructuras y periodos de datos no homogéneos. Esta es una captura de la web de Naciones Unidas, en donde se alberga la información:

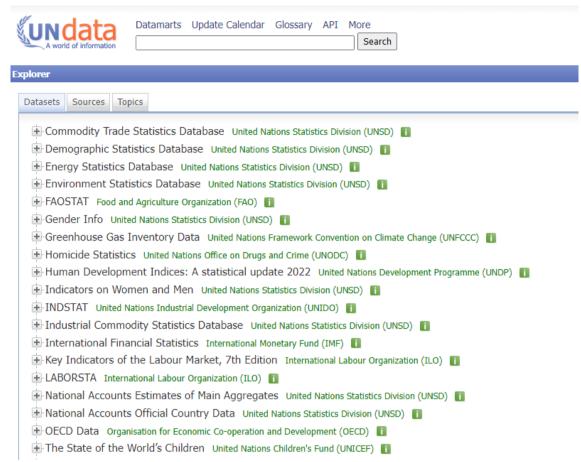


Figura 1 - Algunos Datasets disponibles en el repositorio de las Naciones Unidas

Para dar cobertura a esta necesidad, planteo un proyecto que consista en el tratamiento de los datos seleccionados residentes en este repositorio y la realización de un Cuadro de Mando para explotar la información de una manera eficiente.

1.2. Objetivos

1.2.1. Objetivo Principal

Para este proyecto, he definido el siguiente objetivo principal:

 Crear una solución analítica que permita analizar los diferentes países, tanto desde una visión gerencial, como a nivel específico para cada tipología de información.

Este Cuadro de Mando Analítico estará estructurado con una primera página explicativa de la información contenida, una página donde se muestren los indicadores más importantes en base a la información contenida y un análisis detallado de cada tipología de información, que permita a los usuarios tener una visión lo más realista de la información seleccionada.

Para poder obtener los datos necesarios en cada momento, se incluirán una serie de filtros que permitan acotar la información de forma que se pueda mostrar lo buscado, de forma ágil y eficiente.

Para poder cumplir con este gran objetivo principal, es crítico cumplir con una serie de objetivos secundarios, que se detallan a continuación.

1.2.2. Objetivos Secundarios

Adicionalmente, he definido los siguientes objetivos secundarios:

- Armonizar toda la información bajo un modelo único, accionable y escalable.
 - Uno de los problemas que se debe gestionar es la poca homogeneidad entre las fuentes, dado que los orígenes reales de los datos no son únicos (se puede considerar que es un "agregador de datos"). Por tanto, es importante que a nivel de los "hechos" se establezca una estructura homogénea y que el modelo de datos en su conjunto sea consistente.
- Solucionar los posibles problemas de calidad del dato. Los datos descargados en CSV contienen multitud de información no relevante y que genera problemas de calidad. Además, Es importante excluir todo lo que no se refiera a países, dado que en caso contrario afectará de forma grave a los análisis globales.
- Codificar las dimensiones clave bajo estándares internacionales, para facilitar su integración con otras fuentes. Los diferentes datos que se van a integrar son elementos inicialmente desconectados, pero comparten dimensiones consideradas maestras, como puede ser el país o el sexo. Es importante estandarizarlo todo, incorporando índices y tablas maestras, permitiendo homogeneidad y escalabilidad.
- Homogeneizar ciertas métricas para que los análisis puedan ser comparables.
 En alguna de las fuentes integradas las métricas están basadas en dimensiones con carácter local, como puede ser la moneda de cada país. Para poder disponer de "métricas comparables" entre sí, es importante convertir todas a una unidad común.

1.3. Alcance

El proyecto consistirá en el tratamiento de un conjunto de datos de las Naciones Unidas para el periodo 2011-2020, dado que es un volumen de datos suficiente para realizar un análisis de los países, e implementar un cuadro de mando analítico para poder entender cómo son los diferentes países.

De cara seleccionar las fuentes de datos, dado que el repositorio de Naciones Unidas contiene más de 200 fuentes informacionales, se ha realizado un trabajo de análisis enfocado en dos aspectos:

- Que la fuente de datos ofrezca una cobertura de países amplia.
- Que disponga de datos para el periodo 2011-2020.

Esta es una captura de un extracto de la hoja de Excel donde se recoge la información y se realiza un primer análisis:



Figura 2 - Análisis inicial de las fuentes de datos

Y sobre esta primera selección se realiza un segundo análisis más en detalle, para asegurar que todas las fuentes son confiables. Esta es una captura de las fuentes finales seleccionadas:

Туре	Entidad Entidad			▼ Start Ye ▼	End Y
Demography	Deaths by age, sex and urban/rural residence	ita.aspx?d=PC	Υ	1948	2021
Demography	Population by sex and urban/rural residence	ita.aspx?d=PC	Υ	1995	2021
Demography	Intentional Homicide Victims by counts and rates per 100,000 population	g/DocumentI	Υ	2000	2020
Demography	Employed population by status in employment, industry and sex	a.aspx?d=PO	N	1995	2020
Demography	Native and foreign-born population by age, sex and urban/rural residence	ta.aspx?d=PC	N	1995	2022
Demography	Live births by sex and urban/rural residence	ita.aspx?d=PC	N	1948	2021
Pollution	Methane (CH4) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent	ta.aspx?d=GF	Υ	1990	2020
Pollution	Carbon dioxide (CO2) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent	ta.aspx?d=GF	Υ	1990	2020
Pollution	Nitrous oxide (N2O) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent	ta.aspx?d=GF	Υ	1990	2020
Pollution	Greenhouse Gas (GHGs) Emissions, including Indirect CO2, without LULUCF, in kilotonne CO2 equivalent	ta.aspx?d=GF	N	1990	2020
Pollution	Greenhouse Gas (GHGs) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent	http://data.u	Υ	1990	2020
Pollution	Hydrofluorocarbons (HFCs) Emissions, in kilotonne CO2 equivalent	ta.aspx?d=GF	Υ	1990	2020
Pollution	Unspecified mix of Hydrofluorocarbons (HFCs) and Perfluorocarbons (PFCs) emissions, in kilotonne CO2 equivalent	http://data.u	N	1990	2020
Pollution	Nitrogen trifluoride (NF3) Emissions, in kilotonne CO2 equivalent	ta.aspx?d=GF	N	1990	2020
Pollution	Perfluorocarbons (PFCs) Emissions, in kilotonne CO2 equivalent	ta.aspx?d=Gl	Υ	1990	2020
Pollution	Sulphur hexafluoride (SF6) Emissions, in kilotonne CO2 equivalent	ita.aspx?d=Gl	Υ	1990	2020
Expenditure	Table 1.2 Gross domestic product by expenditures at constant prices	1.aspx?d=SNA	Υ	1946	2021
Expenditure	Table 2.2 Value added by industries at constant prices (ISIC Rev. 3)	a.aspx?d=SNA	N	1946	2021
Expenditure	Table 2.5 Value added by industries at constant prices (ISIC Rev. 4)	1.aspx?d=SNA	Υ	1946	2021
GDP	Per capita GDP at current prices - US dollars	VIA&f=grID%3	Υ	1070	2020
GDP	GDP by Type of Expenditure at current prices - US dollars	VIA&f=grID%3	Υ	1070	2020
GNI	GNI at current prices - US dollars	VIA&f=grID%3	Υ	1070	2020
GNI	Per capita GNI at current prices - US dollars	VIA&f=grID%3	Υ	1070	2020
Refugees	Table with data on Refugees	aspx?d=UNH(Υ	1075	2021
Energy	Electricity, net installed capacity of electric power plants	ata.aspx?d=E	Υ	1090	2021
Energy	Nuclear Electricity	ata.aspx?d=E	Υ	1090	2021
Energy	Solar Electricity	ata.aspx?d=E	Υ	1090	2021
Energy	Thermal Electricity	ata.aspx?d=E	Υ	1090	2021
Energy	Tide, wave and ocean electricity	ata.aspx?d=E	Υ	1090	2021
Energy	Wind Electricity	ata.aspx?d=EI	Υ	1090	2021
World Development Indicators	Air transport, passengers carried	?d=WDI&f=In	Υ	1970	2020
World Development Indicators	Air transport, registered carrier departures worldwide	?d=WDI&f=In	Υ	1970	2020
World Development Indicators	Container port traffic (TEU: 20 foot equivalent units)	=WDI&f=Indic	Υ	1970	2020
World Development Indicators	CPIA social protection rating (1=low to 6=high)	=WDI&f=Indic	Υ	1970	2020
World Development Indicators	CPIA transparency, accountability, and corruption in the public sector rating (1=low to 6=high)	=WDI&f=Indic	Υ	1970	2020
World Development Indicators	CPIA quality of public administration rating (1=low to 6=high)	:WDI&f=Indic	Υ	1970	2020

Figura 3 - Análisis en profundidad y selección final

Se incorporará como anexo el documento con el análisis de las fuentes. Por tanto, las fuentes finales seleccionadas son las siguientes (30 fuentes):

- o Deaths by age, sex and urban/rural residence
- Population by sex and urban/rural residence
- Intentional Homicide Victims by counts and rates per 100,000 population
- Methane (CH4) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent
- Carbon dioxide (CO2) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent
- Nitrous oxide (N2O) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent
- Greenhouse Gas (GHGs) Emissions without Land Use, Land-Use Change and Forestry (LULUCF), in kilotonne CO2 equivalent
- Hydrofluorocarbons (HFCs) Emissions, in kilotonne CO2 equivalent
- o Perfluorocarbons (PFCs) Emissions, in kilotonne CO2 equivalent
- Sulphur hexafluoride (SF6) Emissions, in kilotonne CO2 equivalent
- Table 1.2 Gross domestic product by expenditures at constant prices
- Table 2.2 Value added by industries at constant prices (ISIC Rev. 3)
- Table 2.5 Value added by industries at constant prices (ISIC Rev. 4)
- Per capita GDP at current prices US dollars
- o GDP by Type of Expenditure at current prices US dollars
- o GNI at current prices US dollars
- o Per capita GNI at current prices US dollars
- Table with data on Refugees
- Electricity, net installed capacity of electric power plants
- Nuclear Electricity
- Solar Electricity
- Thermal Electricity
- Tide, wave and ocean electricity
- Wind Electricity
- Air transport, passengers carried
- o Air transport, registered carrier departures worldwide
- Container port traffic (TEU: 20 foot equivalent units)
- CPIA social protection rating (1=low to 6=high)
- CPIA transparency, accountability, and corruption in the public sector rating (1=low to 6=high)
- CPIA quality of public administration rating (1=low to 6=high)

En base a esta selección de datos, se trabajará en diseñar un modelo de datos en capas, que serán las siguientes:

- SA o Staging: Capa donde se recogerá de forma individual cada uno de los datos sin transformar.
- HA o Harmonized: Capa donde se aplicarán reglas de normalización y harmonización para asegurar que todos los datos se pueden analizar como un conjunto.
- MA o Modeled: Capa donde se implementará el modelo analítico, del que se nutrirá en Cuadro de mando.

Respecto al cuadro de Mando, deberá permitir entender cómo es un país concreto en base a diferentes análisis, facilitando la visualización de los datos.

1.4. Enfoque y método seguido

El proyecto se estructurará en estos grandes bloques:

1. Análisis de la información del repositorio.

Lo primero que hay que realizar en este proyecto es navegar por toda la información disponible en el repositorio de las Naciones Unidas e identificar aquella información potencialmente utilizable. Se plantea que se integren los últimos 10 años que estén disponibles para todas las fuentes.

Una vez finalizada esta fase del proyecto, se dispondrá de un catálogo final de datos seleccionados, que poder compartir con personas fuera del proyecto.

2. Elección de las herramientas

Esta fase es crítica ya que puede marcar el éxito o fracaso de las fases posteriores. Es aquí donde se decidirá qué herramienta/s utilizar, buscando siempre la eficiencia en recursos, tanto humanos como económicos.

3. Diseño del modelo de datos.

En esta fase se trabajará el modelo de datos, con un enfoque de capas: Staging, armonizada y modelada. Para cada capa se generarán archivos, para asegurar la trazabilidad de los datos.

4. Diseño del cuadro de Mando.

Es aquí donde se elaborará el Cuadro de Mando Analítico y se implementarán todos los análisis, para asegurar el correcto análisis y la posterior toma de decisiones.

5. Pruebas de Validación

Por último, se tienen que realizar una serie de pruebas de validación del producto, para asegurar que cumple con los objetivos y requerimientos definidos al comienzo.

En base a este enfoque, esta sería la EDT:



Figura 4 - Estructura de Descomposición del Trabajo (EDT)

Código: 1.1	Nombre: Investigación Datos disponibles
Descripción	Análisis del repositorio y descubrimiento de los datos
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	3 días
Resultados	Todas las fuentes informadas en un documento.
Restricciones	N/A
Dependencias	N/A
Criterios de aceptación	Investigación de datos realizada

Código: 1.2	Nombre: Creación de un catálogo de datos potencialmente interesantes	
Descripción	Elección de las fuentes que podrían ser candidatas.	
Responsable	Óscar Quero	
Recursos	Analista de Datos	
Duración	5 días	
Resultados	Fuentes potenciales marcadas como elegibles.	
Restricciones	Los datos deben estar disponibles para las fechas 2011-2020.	
Dependencias	Investigación de datos realizada	
Criterios de aceptación	Todas las fuentes potenciales identificadas e incluidas en un documento.	

Código: 1.3	Nombre: Catálogo de datos a integrar definido
Descripción	Selección final de las fuentes a integrar.
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	3 día
Resultados	Catálogo de fuentes finales finalizado.
Restricciones	Los datos deben estar disponibles para la mayoría de los países.
Dependencias	Catálogo de fuentes potenciales finalizado.
Criterios de aceptación	Fuentes finales seleccionadas y validadas.

Código: 2.1	Nombre: Análisis de las herramientas disponibles
Descripción	Estudio de las opciones disponibles en el mercado.
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	3 días
Resultados	Analista de Datos
Restricciones	N/A
Dependencias	N/A
Criterios de aceptación	Información obtenida.

Código: 2.2	Nombre: Elección del set de herramientas a utilizar		
Descripción	En base a la información obtenida, elección de las herramientas.		
Responsable	Óscar Quero		
Recursos	Project manager		
Duración	3 día		
Resultados	Herramientas elegidas.		
Restricciones	N/A		
Dependencias	N/A		
Criterios de aceptación	Información sobre las herramientas disponible.		

Código: 3.1	Nombre: Diseño capa Staging		
Descripción	Carga inicial de las fuentes originales sin transformaciones.		
Responsable	Óscar Quero		
Recursos	Analista de Datos		
Duración	5 días		
Resultados	Estructura de datos creada.		
Restricciones	N/A		
Dependencias	Herramienta de integración elegida.		
Criterios de aceptación	N/A		

Código: 3.2	Nombre: Diseño capa Armonizada		
Descripción	Estandarización de las fuentes		
Responsable	Óscar Quero		
Recursos	Analista de Datos		
Duración	10 días		
Resultados	Estructura de datos creada.		
Restricciones	Todos los datos deben estar relacionados con países.		
Dependencias	Capa staging finalizada		
Criterios de aceptación	Todos los objetos creados.		

Código: 3.3	Nombre: Diseño capa modelada
Descripción	Diseño de la capa de datos que utilizará el Cuadro de Mando
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	10 días
Resultados	Estructura de datos creada.
Restricciones	La estructura de los hechos debe ser común y deben existir tablas
	maestras.
Dependencias	Capa armonizada finalizada
Criterios de aceptación	Todos los objetos creados.

Código: 4.1	Nombre: Definición dimensiones y métricas
Descripción	Definición de las dimensiones y métricas con las que se trabajará
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	5 días
Resultados	Catálogo finalizado.
Restricciones	N/A
Dependencias	Capa modelada finalizada
Criterios de aceptación	Dimensiones y métricas documentadas.

Código: 4.2	Nombre: Diseño estructura Dashboard
Descripción	Definición y diseño de la estructura del Cuadro de Mando, que
	contendrá los análisis
Responsable	Óscar Quero
Recursos	Analista de Datos
Duración	3 días
Resultados	Estructura finalizada
Restricciones	N/A
Dependencias	N/A
Criterios de aceptación	Estructura finalizada y visualizable

Código: 4.3	Nombre: Implementación análisis	
Descripción	Desarrollo de los análisis visuales	
Responsable	Óscar Quero	
Recursos	Analista de Datos	
Duración	25 días	
Resultados	Cuadro de Mando finalizado	
Restricciones	N/A	
Dependencias	Estructura Cuadro de Mando finalizada	
Criterios de aceptación	Cuadro de Mando funcional	

Código: 5.1	Nombre: Ejecución pruebas
Descripción	Realización de las pruebas que validen el producto
Responsable	Óscar Quero
Recursos	Jefe de Proyecto
Duración	2 días
Resultados	Pruebas realizadas
Restricciones	N/A
Dependencias	Cuadro de Mando finalizado
Criterios de aceptación	Pruebas realizadas

Código: 5.2	Nombre: Documentación pruebas
Descripción	Documentación de todas las pruebas realizadas
Responsable	Óscar Quero
Recursos	Jefe de Proyecto
Duración	2 días
Resultados	Pruebas documentadas
Restricciones	N/A
Dependencias	Pruebas realizadas
Criterios de aceptación	Pruebas documentadas

1.5. Planificación del Trabajo

Para el desarrollo de este proyecto se estima una media de 4 horas diarias de dedicación diaria y requerirá para su finalización un total de 316 horas (79 días). Todo el trabajo será realizado por un único recurso (Óscar Quero), que asumirá diferentes roles durante el desarrollo del mismo (Jefe de Proyecto y Analista de Datos), y se utilizará un ordenador personal para cumplir con todas las tareas

El proyecto será desarrollado entre el 01.03.2023 y el 18.06.2023, y se realizarán una serie de entregas parciales o de seguimiento, donde se irá entregando mayor contenido. Al finalizar el desarrollo, se efectuará una entrega final y se compartirán los documentos adjuntos comprometidos.

Para el correcto seguimiento del proyecto se ha creado un Gantt del proyecto, donde se detallan las diferentes actividades, con sus fechas y dedicación:

	(1)	Nombre	Duracion	Inicio	Terminado	Predecesores
1		Cuadro de Mando Naciones Unidas	99 days	1/03/23 8:00	10/06/23 8:00	
2		Análisis de la información del repositorio	11 days	1/03/23 8:00	13/03/23 17:00	
3		Investigación Datos disponibles	3 days	1/03/23 8:00	3/03/23 17:00	
4		Creación de un catálogo de datos potencialmente i	5 days	6/03/23 8:00	10/03/23 17:00	3
5		Catálogo de datos a integrar definido	3 days	11/03/23 8:00	13/03/23 17:00	4
6		HITO: Catálogo de datos finalizado	0 days	13/03/23 17:00	13/03/23 17:00	5
7		Elección de las herramientas	6 days	14/03/23 8:00	19/03/23 17:00	2
8		Análisis de las herramientas disponibles	3 days	14/03/23 8:00	16/03/23 17:00	
9		Elección del set de herramientas a utilizar	3 days	17/03/23 8:00	19/03/23 17:00	8
10		HITO: Entrega PEC1	0 days	19/03/23 17:00	19/03/23 17:00	9
11		Diseño del modelo de datos	31 days	20/03/23 8:00	19/04/23 17:00	7
12		Diseño capa Staging	5 days	20/03/23 8:00	24/03/23 17:00	
13		Datos: Población	1 day	20/03/23 8:00	20/03/23 17:00	
14		Datos: Muertes	1 day	21/03/23 8:00	21/03/23 17:00	13
15		Datos: Homicidios	0,5 days	22/03/23 8:00	22/03/23 13:00	14
16		Datos: Contaminación	0,5 days	22/03/23 13:00	22/03/23 17:00	15
17		Datos: Energía	0,5 days	23/03/23 8:00	23/03/23 13:00	16
18		Datos: GDP	0,5 days	23/03/23 13:00	23/03/23 17:00	17
19		Datos: GNI	0,5 days	24/03/23 8:00	24/03/23 13:00	18
20		Datos: WDI	0,5 days	24/03/23 13:00	24/03/23 17:00	19
21		HITO: Capa Staging finalizada	0 days	24/03/23 17:00	24/03/23 17:00	20
22		Diseño capa Armonizada	10 days	25/03/23 8:00	3/04/23 17:00	12
23		Datos: Población	2 days	25/03/23 8:00	26/03/23 17:00	
24		Datos: Muertes	2 days	27/03/23 8:00	28/03/23 17:00	23
25		Datos: Homicidios	1 day	29/03/23 8:00	29/03/23 17:00	24
26		Datos: Contaminación	1 day	30/03/23 8:00	30/03/23 17:00	25
27		Datos: Energía	1 day	31/03/23 8:00	31/03/23 17:00	26
28		Datos: GDP	1 day	1/04/23 8:00	1/04/23 17:00	27
29		Datos: GNI	1 day	2/04/23 8:00	2/04/23 17:00	28
30		Datos: WDI	1 day	3/04/23 8:00	3/04/23 17:00	29
31		HITO: Capa Armonizada finalizada	0 days	3/04/23 17:00	3/04/23 17:00	30
32		Diseño capa Armonizada	16 days	4/04/23 8:00	19/04/23 17:00	12;22
33		M_Paises	3 days	4/04/23 8:00	6/04/23 17:00	
34		M_Sexo	2 days	7/04/23 8:00	8/04/23 17:00	33

Figura 5 - Planificación con Projectlibre (1/2)

35		F_Datos	5 days	9/04/23 8:00	13/04/23 17:00	34
36		HITO: Capa Modelada finalizada	0 days	13/04/23 17:00	13/04/23 17:00	35
37	Ö	HITO: Entrega PEC2	0 days	19/04/23 17:00	19/04/23 17:00	36
38		Diseño del Cuadro de Mando	32 days	20/04/23 8:00	21/05/23 17:00	11
39		Definición dimensiones y métricas	5 days	20/04/23 8:00	24/04/23 17:00	
40		Diseño estructura Dashboard	3 days	25/04/23 8:00	27/04/23 17:00	39
41		Implementación análisis	24 days	28/04/23 8:00	21/05/23 17:00	40
42		KPIs	3 days	28/04/23 8:00	30/04/23 17:00	
43		Análisis diferentes áreas de información	22 days	28/04/23 8:00	19/05/23 17:00	
44		HITO: Cuadro de mando implementado	0 days	19/05/23 17:00	19/05/23 17:00	43
45	Ö	HITO: Entrega PEC3	0 days	21/05/23 17:00	21/05/23 17:00	43
46		Pruebas de validación	4 days	22/05/23 8:00	25/05/23 17:00	38
47		Ejecución de las pruebas	2 days	22/05/23 8:00	23/05/23 17:00	
48		Documentación pruebas	2 days	24/05/23 8:00	25/05/23 17:00	47
49		Tareas administrativas		26/05/23 8:00	10/06/23 8:00	46
50		Revisión documento final	5 days	26/05/23 8:00	30/05/23 17:00	
51		Preparación Anexos	5 days	31/05/23 8:00	4/08/23 17:00	50
52		Preparación presentación	5 days	5/06/23 8:00	9/06/23 17:00	51
53		HITO: Entrega final	0 days	10/06/23 8:00	10/06/23 8:00	52

Figura 6 - Planificación con Projectlibre (2/2)

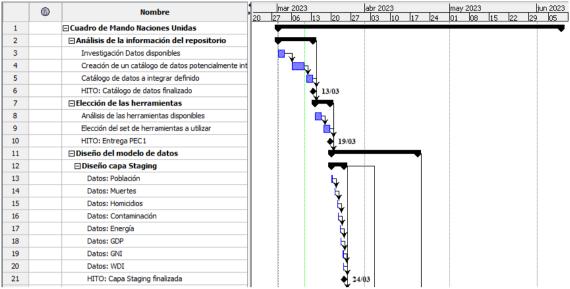


Figura 7 - Diagrama de Gantt (1/3)

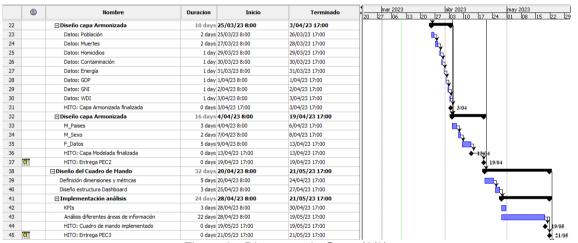


Figura 8 - Diagrama de Gantt (2/3)



Figura 9 - Diagrama de Gantt (3/3)

Se plantean los siguientes hitos del proyecto:

Hito	Fecha Hito	Tipo
Catálogo de datos finalizado	13/03/2023	Ejecución
Entrega PEC1	19/03/2023	Seguimiento
Capa Staging finalizada	24/03/2023	Ejecución
Capa Armonizada finalizada	03/04/2023	Ejecución
Capa Modelada finalizada	13/04/2023	Ejecución
Entrega PEC2	19/04/2023	Seguimiento
Cuadro de mando implementado	19/05/2023	Ejecución
Entrega PEC3	21/05/2023	Seguimiento
Entrega final	09/06/2023	Seguimiento

Figura 10 - Hitos del proyecto

Todos los hitos marcados con verde se consideran que son críticos para lograr que el proyecto sea un éxito, por lo que se hará un seguimiento más exhaustivo de ellos.

1.6. Recursos

Tal y como se comentó en el punto anterior, el proyecto será elaborado en su totalidad y de manera individual por un solo recurso (Óscar Quero), que desempeñará dos roles diferentes:

- Jefe de Proyecto
 Sera el rol decisor y el encargado de realizar las pruebas de validación del producto. Sus funciones son más de coordinación.
- Analista de Datos
 Es el rol encargado del análisis/entendimiento de los datos, la
 integración de estos en un modelo analítico y la implementación
 de un Cuadro de Mandos que permita la toma de decisiones.

1.7. Costes del proyecto

El desarrollo del proyecto no tiene costes fijos ni recurrentes a nivel tecnológico, ya que se utilizará una herramienta gratuita para su desarrollo (Qlikview), pero si hay costes asociados a la dedicación al proyecto. La siguiente tabla los detalla:

Rol	Días	Horas/día	Total Horas	Promedio Coste/Hora (€)	Total (€)
Jefe de Proyecto	7	4	28	100	2.800,00€
Analista de Datos	72	4	288	75	21.600,00€
Total	79	4	316	87,5	24.400,00€

Figura 11 - Dedicación y coste por roles

Por tanto, el coste total del proyecto serán 24.400€.

1.8. Riesgos identificados

He realizado un análisis de los posibles riesgos que pueden impactar negativamente en el desarrollo y finalización exitosa del mismo, además de las acciones correctoras planteadas:

ID	Riesgo	Descripción	Probabilidad	Impacto	Score	Acciones	
R1	Retrasos en tareas	Incorrecta estimación de la dedicación	6	10	60	Volver a replanificar las tareas afectadas	
R2	Problemas con la calidad de datos	Si el dato no tiene una calidad mínima las conclusiones podrían ser erróneas	4 8		<mark>3</mark> 2	Implementar medidas correctoras adicionales en los procesos ETL	
R3	Diseño del modelo de datos incorrecto	Un diseño incorrecto puede afectar a las conclusiones de los análisis	5	9	45	Volver a trabajar el modelo	
R4	Problemas laborales	Necesidad de dedicar más horas de las habituales al trabajo	6	9	54	Dedicar más tiempo por las noches	
R5	Problemas con el equipo	Problemas técnicos con el equipo	3	8	24	Cambiar de equipo	

Figura 12 - Tabla detallada riesgos

Para cada uno de los riesgos se ha identificado la posibilidad de que se produzcan y el posible impacto, ambos con indicadores cuantitativos. En el siguiente gráfico se ilustran todos ellos de forma visual:



Figura 13 - Mapa de riesgos visual

1.9. Breve sumario de productos obtenidos

Una vez finalice el proyecto, se aportarán los siguientes entregables:

- Las entregas parciales y la memoria final del proyecto.
 Se generarán los diferentes entregables en pdf y se subirán a la plataforma.
- Archivo de Qlikview (QVW) que se encarga de generar las diferentes capas de datos.
 Se ha creado un archivo de Qlikview que contiene todos los procesos de creación de los objetos de las diferentes capas de datos.

- Contenido del modelo de datos en las diferentes capas.
 - CSVs y QVDs de la capa 1 de Staging
 Se generarán ambos tipos de archivos para cada fuente de datos.
 - CSVs y QVDs de la capa 2 Harmonizada
 Se generarán ambos tipos de archivos para cada fuente de datos.
 - CSVs y QVDs de la capa 3 Modelada
 Se generarán ambos tipos de archivos para cada estructura analítica.
- o Excel de mapeo de dimensiones.
- Excel de datos maestros.
- Cuadro de Mando Analítico con Qlikview.
- Pruebas realizadas.
- Anexos.

1.10. Breve descripción de los otros capítulos de la memoria

En los próximos capítulos se describirán en detalle los diferentes etapes de construcción del proyecto, que son las siguientes:

- Elección de las herramientas
 Análisis de las diferentes herramientas disponibles y elección de las más adecuadas.
- El Modelo de Datos
 Descripción de las diferentes capas del modelo de datos y cada uno de sus elementos, explicando su estructura y transformaciones.
- El Cuadro de Mando Analítico
 Descripción del Cuadro de Mando y todos los análisis que incluye, además del detalle de las dimensiones y métricas incluidas.
- El Plan de Pruebas
 Documentación de todo el proceso de pruebas de validación.
- Las Conclusiones
 Conclusiones de la realización de este Trabajo.
- El Glosario

Glosario de términos, para facilitar la comprensión por parte de los lectores.

- La Bibliografía
 Referencias bibliográficas usades para la elaboración del proyecto.
- Los Anexos
 Anexos incluidos y que apoyan el desarrollo.

2. La solución analítica

2.1. Selección de las herramientas

El proyecto tiene dos áreas de trabajo claramente diferenciadas: integración y analítica. Por tanto, el enfoque clásico sería utilizar herramientas especializadas para dar solución a las necesidades del proyecto.

Para la parte de integración, Este es el Cuadrante Mágico de Gartner para 2022:



Figura 14 - Cuadrante Mágico de Gartner de herramientas de Integración

Para este análisis se han seleccionado dos herramientas líderes (Informatica y Talend) y una Open Source (Kettle), siendo estas las características evaluadas y su peso, en función de la importancia que tienen:

- Coste (10%)
- Seguridad (20%)
- Facilidad de Implementación/uso (20%)
- Soporte/comunidad (15%)
- Velocidad (5%)
- Calidad (10%)
- Capacidad de integración (20%)

Obteniéndose la siguiente comparativa:

		Talend	Kettle	PowerCenter
Coste	10%	6,3	10,0	0,7
Seguridad	20%	6,7	3,0	9,7
Facilidad de Implementación/uso	20%	6,3	3,3	7,0
Soporte/comunidad	15%	6,3	9,7	3,7
Velocidad	5%	8,0	7,3	9,3
Calidad	10%	7,3	5,7	8,0
Capacidad de integración	20%	7,0	6,3	7,0
		6,7	5,9	6,6

Figura 15 - Análisis herramientas integración

Las valoraciones se han basado en el siguiente análisis:

- Talend: Herramienta con valoraciones medias-altas en todos los aspectos, destacando en la velocidad.
- Kettle: Herramienta Open Source con buenas capacidades y gran soporte de la comunidad, aunque algo compleja en su uso.
- Powercenter: Herramienta muy potente pero muy cara, con alta dependencia del proveedor.

Por tanto, en base al análisis, la más adecuada sería Talend, dado que cubre con garantías todo lo que necesita el proyecto en el corto, medio y largo plazo, y, aunque tiene un coste de licencias, es muy estable y eficiente. Se descarta Powercenter porque, aunque tiene una valoración similar a Talend, el coste es muy alto y desequilibraría el proyecto.

Para la herramienta analítica, este sería el Cuadrante Mágico de Gartner:



Figura 16 - Cuadrante Mágico de Gartner de herramientas de Business Intelligence

Para seleccionar la herramienta, se han analizado las siguientes:

- Power BI
- Pentaho
- Qliksense

En base a unos aspectos críticos para el proyecto, a los que se les ha aplicado un peso en función de la importancia que tienen:

- Coste (30%)
- Seguridad (20%)
- Facilidad de Implementación/uso (20%)
- Self service (15%)
- Soporte/comunidad (10%)
- Capacidad de integración (5%)

Obteniéndose la siguiente comparativa:

9	•	Power BI	Pentaho	Qlik
Coste	30%	8,0	9,0	5,0
Seguridad	20%	7,3	3,3	9,0
Facilidad de Implementación/uso	20%	9,0	1,7	9,0
Self service	15%	8,3	2,7	9,0
Soporte/comunidad	10%	6,0	9,3	10,0
Capacidad de integración	5%	6,0	5,3	8,0
		7,8	5,3	7,9

Figura 17 - Análisis herramientas Business Intelligence

Las valoraciones se han basado en el siguiente análisis:

- Power BI: Herramienta con costes bajos y fácil implementación, con curvas de aprendizaje muy cortas. Buen enfoque self-service.
- Pentaho: Herramienta Open source que, aunque puede ser a coste cero en licencias con la versión comunity, es de implementación compleja y muy rígida. Como punto positivo, hay que resaltar el soporte de la comunidad.
- Qlik: Herramienta con un coste medio y muy buen soporte, que cubre todos los aspectos con garantías.

Por tanto, en base al análisis, mi recomendación es utilizar Qlik, dado que cubre con garantías todo lo que necesita el proyecto en el corto, medio y largo plazo, y, además, tiene una curva de aprendizaje muy baja y costes controlados. Otro factor importante que he considerado para elegir Qlik frente a Power BI es el conocimiento de la herramienta por parte del equipo de proyecto.

El hecho de elegir Qlik como herramienta aporta además un valor adicional: tiene capacidades ETL (de integración), por lo que podemos descartar Talend y que toda la capa de integración la asuma Qlik. Además, a la hora de poder trabajar de forma ágil en local, se plantea utilizar Qlikview en vez de Qliksense, al tener disponible una herramienta gratuita instalable localmente.

2.2. Construcción del modelo de datos

Para poder llegar a un modelo analítico "accionable y escalable" se ha implementado un proceso de integración utilizando Qlikview, que se encarga de tratar los ficheros origen e ir generando las diferentes capes de datos, tanto intermedias (SA, HA) como analítica (MA). Esta sería la representación visual del modelo de integración:

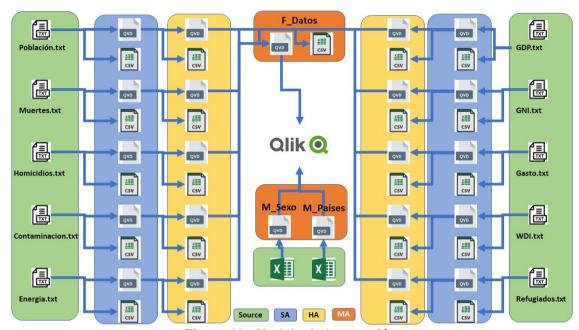


Figura 18 - Modelo de Integración

2.2.1. Variables

Se han implementado las siguientes variables, para poder adaptar elproceso de integración a cualquier entorno:

- Ubicación archivos origen:
 - Let Source_Files = 'C:\oscar\UOC\0 TFG\Desarrollo\Source files';
- Ubicación de los archivos QVD para cada capa de datos:
 - o Let QVDs_Layer1 = 'C:\oscar\UOC\0
 TFG\Desarrollo\QVDs\Layer1';
 - Let QVDs_Layer2 = 'C:\oscar\UOC\0
 TFG\Desarrollo\QVDs\Layer2';
 - c Let QVDs_Layer3 = 'C:\oscar\UOC\0
 TFG\Desarrollo\QVDs\Layer3';
- Ubicación de los archivos CSV para cada capa de datos:

 - Let CSVs_Layer2 = 'C:\oscar\UOC\0
 TFG\Desarrollo\CSVs\Layer2';

2.2.2. Mappings

Para poder armonizar toda la información y poder ofrecer una visión única y comparable de la verdad, se han tenido que implementar una serie de mapeos de información, que detallo a continuación.

2.2.2.1. Map_Countries

2.2.2.1.1. Descripción

Con este mapping se convierte cada descriptivo de país en su código ISO correspondiente.

2.2.2.1.2. Script

Ver anexo scripts.

2.2.2.2. Map_Sex

2.2.2.2.1. Descripción

Con este mapping se eliminarán los descriptives del campo sexo y se sustituirán por un ID numérico.

2.2.2.2. Script

Ver anexo scripts.

2.2.2.3. Map_AgeRange

2.2.2.3.1. Descripción

Este mapping sirve para crear una dimensión "Rango_edad" siempre que el campo edad venga informado.

2.2.2.3.2. Script

Ver anexo scripts.

2.2.2.4. Map_AgeTipo

2.2.2.4.1. Descripción

Con este mapping se crea una nueva dimensión en aquellos casos con la edad informada, para indicar si se refiere a menores de edad o mayores de edad.

2.2.2.4.2. Script

Ver anexo scripts.

2.2.2.5. Map_Divisa_Euro

2.2.2.5.1. Descripción

Con este mapping convertimos todas las cantidades en divises locales a euros.

2.2.2.5.2. Script

Ver anexo scripts.

2.2.3. Capa Staging (SA)

En esta capa es donde se cargan los datos de los ficheros origen, generando un único fichero por Fuente y limpiándolos de aquello que no es válido. Todas las Fuentes se catalogan por tipo (Data_Type), subtipo (Data_Subtype) y, en algunos casos, se incluye un detalle del subtipo (Data_Subtype_Detail).

2.2.3.1. Población

2.2.3.1.1. Descripción

Carga de los datos de población de los países, por año y sexo.

2.2.3.1.2. Script

Ver anexo scripts.

2.2.3.2. Muertes

2.2.3.2.1. Descripción

Carga de los datos de muertes de los países, por año, sexo y edad. En este caso, dado el volumen de registros anual y lo que permite descargar la web, hay un fichero por año.

2.2.3.2.2. Script

Ver anexo scripts.

2.2.3.3. Homicidios

2.2.3.3.1. Descripción

Carga de los datos de homicidios, con la complejidad de que vienen en una tabla con una columna por año.

2.2.3.3.2. Script

Ver anexo scripts.

2.2.3.4. Contaminación

2.2.3.4.1. Descripción

Carga de los datos referentes a contaminación, disponibles por país y año, con un fichero por cada tipo de contaminante.

2.2.3.4.2. Script

Ver anexo scripts.

2.2.3.5. Energía

2.2.3.5.1. Descripción

Carga de los datos de producción de energía de los diferentes países, con un fichero por tipo de energía. Para esta Fuente de datos se cuenta con un nivel más de detalle, que se almacena con propósitos analíticos.

2.2.3.5.2. Script

Ver anexo scripts.

2.2.3.6. GDP

2.2.3.6.1. Descripción

Carga de los datos de GDP (PIB) por año, país y desglose por subtipos, tanto global como per cápita.

2.2.3.6.2. Script

Ver anexo scripts.

2.2.3.7. GNI

2.2.3.7.1. Descripción

Carga de los datos de GNI (Renta Interior Bruta) por año, país y desglose por subtipos, tanto global como per cápita.

2.2.3.7.2. Script

Ver anexo scripts.

2.2.3.8. Gasto

2.2.3.8.1. Descripción

Aquí es donde se cargarán los datos de gestos de los países por año y tipo, cada uno en la moneda de curso legal del país.

2.2.3.8.2. Script

Ver anexo scripts.

2.2.3.9. WDI

2.2.3.9.1. Descripción

Aquí se cargarán diferentes indicadores de desarrollo globales:

- Número de vuelos de transporte
- Número de pasajeros en vuelos
- Tráfico de contenedores de transporte de mercancías
- Indicadores de igualdad de genero
- Indicadores de calidad de la administración Pública
- Indicadores de protección social
- Indicadores de transparencia del sector público

Todos estos datos están a nivel país y año, aunque no están disponibles para la globalidad de países.

2.2.3.9.2. Script

Ver anexo scripts.

2.2.3.10. Refugiados

2.2.3.10.1. Descripción

Por último, se cargan los datos de refugiados, donde se indica, por año y país, el número por país de origen.

2.2.3.10.2. Script

Ver anexo scripts.

2.2.4. Capa Harmonizada (HA)

En esta capa se estandarizan todas las fuentes a la estructura de datos final:

- Data_Type
- Data_SubType
- Data_SubType_Detail
- Country_ID

- Year
- Sex ID
- Age
- Age_Range
- Age_Type
- Value

2.2.4.1. Población

2.2.4.1.1. Descripción

En esta segunda fase del proceso de integración el objetivo es obtener los IDs de país y sexo.

2.2.4.1.2. Script

Ver anexo scripts.

2.2.4.2. Muertes

2.2.4.2.1. Descripción

En el proceso de capa 2 con los datos de muertes, se incluyen los IDs de sexo y país (código ISO), además de calcular los rangos de edad y el tipo de individuo en función de su edad.

2.2.4.2.2. Script

Ver anexo scripts.

2.2.4.3. Homicidios

2.2.4.3.1. Descripción

En los datos de homicidios, aparte de armonizar la estructura, se incluye el código ISO de los países.

2.2.4.3.2. Script

Ver anexo scripts.

2.2.4.4. Contaminación

2.2.4.4.1. Descripción

En los datos de contaminación, aparte de armonizar la estructura, se incluye el código ISO de los países.

2.2.4.4.2. Script

Ver anexo scripts.

2.2.4.5. Energía

2.2.4.5.1. Descripción

En los datos de energía, aparte de armonizar la estructura, se incluye el código ISO de los países.

2.2.4.5.2. Script

Ver anexo scripts.

2.2.4.6. GDP

2.2.4.6.1. Descripción

En los datos de PIB, aparte de armonizar la estructura, se incluye el código ISO de los países y se convierten los valores a euros para que los análisis económicos sean homogéneos en todo el aplicativo.

2.2.4.6.2. Script

Ver anexo scripts.

2.2.4.7. GNI

2.2.4.7.1. Descripción

En los datos de GNI, aparte de armonizar la estructura, se incluye el código ISO de los países y se convierten los valores a euros para que los análisis económicos sean homogéneos en todo el aplicativo.

2.2.4.7.2. Script

Ver anexo scripts.

2.2.4.8. Gasto

2.2.4.8.1. Descripción

En los datos de gasto, aparte de armonizar la estructura, se incluye el código ISO de los países y se convierten los valores a euros para que los análisis económicos sean homogéneos en todo el aplicativo (de origen vienen en monedas locales).

2.2.4.8.2. Script

Ver anexo scripts.

2.2.4.9. WDI

2.2.4.9.1. Descripción

En los datos de WDI (Indicadores de Desarrollo Mundiales), aparte de armonizar la estructura, se incluye el código ISO de los países.

2.2.4.9.2. Script

Ver anexo scripts.

2.2.4.10. Refugiados

2.2.4.10.1. Descripción

En los datos de refugiados, aparte de armonizar la estructura, se incluye el código ISO de los países.

2.2.4.10.2. Script

Ver anexo scripts.

2.2.5. Capa Analítica (MA)

En la capa analítica se ha definido una única tabla de hechos (F_Datos) y dos tablas maestras, con los descriptivos de dos dimensiones maestras: Países (M_Paises) y Sexo (M_Sexo).

2.2.5.1. M Paises

2.2.5.1.1. Descripción

Tabla maestra que contiene la tabla de códigos ISO de los países, junto con los descriptivos y las regiones asignadas. Todos los descriptivos están en inglés.

2.2.5.1.2. Script

Ver anexo scripts.

2.2.5.2. M Sexo

2.2.5.2.1. Descripción

La segunda tabla maestra es la que tiene información sobre los diferentes sexos, y se compone de un ID y un descriptivo.

2.2.5.2.2. Script

Ver anexo scripts.

2.2.5.3. F_Datos

2.2.5.3.1. Descripción

La tabla de hechos es la que contiene todos los datos que se van a analizar. La estructura de datos ya se harmonizó en la capa anterior, y es aquí donde se junta toda la información de las diferentes tipologías de información, asegurando que no hay problemas de calidad.

2.2.5.3.2. Script

Ver anexo scripts.

2.3. Modelo de Datos

Finalmente, el modelo de datos queda conformado por las dos tablas maestras mencionadas anteriormente y la tabla de hechos. Esta sería la estructura que tendría y los campos incluidos:

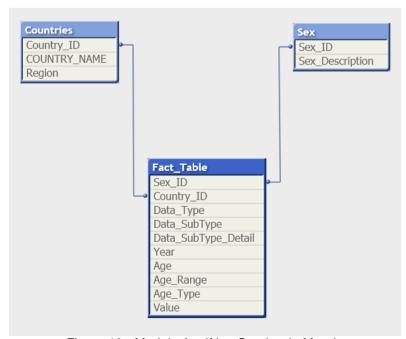


Figura 19 - Modelo Analítico Cuadro de Mando

2.4. Herramienta Analítica

La herramienta analítica la he desarrollado usando Qlikview y la he enfocado como un análisis basado en el proceso normal de descubrimiento, es decir, de datos más generales a más específicos.

En los próximos apartados detallaré todos los aspectos relevantes de la herramienta, desde las dimensiones/métricas a los diferentes análisis implementados.

2.4.1. Dimensiones

2.4.1.1. Año

Esta es la dimensión temporal de los datos, que según el alcance del proyecto contendrá valores de 2011 a 2020.

2.4.1.2. Edad

Es una dimensión que contiene la edad de las persones y solo está disponible para la métrica de las muertes. Cuando no tiene valor numérico, el valor es NA.

2.4.1.3. Región

Todos los países tienen una región asignada, pudiéndose analizar los datos por agrupaciones de países en regiones también.

2.4.1.4. País

Todos los grupos de datos que se han agregado al modelo están a nivel de país, por lo que se dispone de una dimensión concreta con los diferentes descriptivos.

2.4.1.5. Sexo

Algunos grupos de datos, como la población, contiene datos por sexo. Los valores posibles son male, female y unknown.

2.4.2. Métricas

2.4.2.1. Habitantes

Esta métrica se refiere a la población por región, país, año y sexo. Este es el código de la métrica:

num(sum({<Data_Type={'Demography'},Data_SubType={'Population'}, Data_SubType_Detail= ,Sex_Description= >}Value),'###.###.###.###")

2.4.2.2. Muertes

Esta métrica se refiere a las muertes por región, país, año, edad y sexo. Este es el código de la métrica:

num(sum({<Data_Type={'Demography'},Data_SubType={'Deaths'}, Data_SubType_Detail=,
Sex_Description= >}Value),'###.###.###")

2.4.2.3. Homicidios

Esta métrica se refiere a los homicidios por región, país y año. Este es el código de la métrica:

=num(sum({<Data_Type={'Homicides'},Data_SubType={'Homicides'},Data_SubType_Detail=,Sex_Description=>}Value),'###.###.###.###.###')

2.4.2.4. Refugiados

Esta métrica se refiere a los refugiados por región, país y año, teniéndose también el país de origen de los refugiados. Este es el código de la métrica:

=num(sum({<Data_Type={'Refugees'},Data_SubType={'Refugees'},Data_SubType_Detail=,Sex_Description=>}Value),'###.###.###.###.###')

2.4.2.5. Contaminación

Esta métrica se refiere a la contaminación por región, país y año. Este es el código de la métrica para CO2 a modo de ejemplo (el resto se configuran cambiando el valor de Data SubType):

=num(sum({<Data_Type={'Pollution'},Data_SubType={'CO2'}, Data_SubType_Detail=,Sex_Description= >}Value),'###.###.###.###.###.###')

2.4.2.6. PIB

Esta métrica se refiere Producto Interior Bruto por región, país y año. Se disponen tanto del PIB total como el PIB per cápita.

2.4.2.6.1. PIB total

num(avg({<Data_Type={'Expenditure'},Data_SubType={'Gross Domestic Product'},
Data_SubType_Detail={'Equals: GROSS DOMESTIC PRODUCT'},Sex_Description=
>}Value),'###.###.###.###.###')

2.4.2.6.2. PIB per cápita

num(avg({<Data_Type={'GDP'},Data_SubType={'GDP per Capita'}, Data_SubType_Detail={'Gross Domestic Product (GDP)'},Sex_Description= >}Value),'###.###.###.###")

2.4.2.7. GNI

Esta métrica se refiere GNI por región, país y año. Se disponen tanto del GNI total como el GNI per cápita.

2.4.2.7.1. GNI total

num(avg((<Data_Type={'GNI'},Data_SubType={'Renta Current prices'}, Data_SubType_Detail={'Gross
National Income (GNI)'},Sex_Description= >}Value),'###.###.###.###.##")

2.4.2.7.2. GNI per cápita

num(avg({<Data_Type={'GNI'},Data_SubType={'GNI per Capita'}, Data_SubType_Detail={'Gross National Income (GNI)'},Sex_Description= >}Value),'###.###.###.###.###')

2.4.2.8. Gasto

Esta métrica se refiere gasto por región, país y año. Este es el código de la métrica para "Alojamiento y comidas":

num(sum({<Data_Type={'Expenditure'},Data_SubType={'Value Added Industries'},
Data_SubType_Detail={'Accommodation and food service activities'},Sex_Description=
>}Value),'###.###.###.###.###')

2.4.2.9. Indicadores de desarrollo

Hay también diferentes métricas para los indicadores de desarrollo (WDI) por región, país y año. Los códigos están a continuación.

2.4.2.9.1. Transporte Aéreo (Mercancías)

=num(sum({<Data_Type={"World Development Indicators"},Data_SubType={'Air Transport - Carrier Departures'}, Data_SubType_Detail=,Sex_Description= >}Value),'###.###.###.###")

2.4.2.9.2. Transporte Aéreo (Pasajeros)

=num(sum({<Data_Type={"World Development Indicators"},Data_SubType={"Air Transport - Passengers Carried"}, Data_SubType_Detail=,Sex_Description= >}Value), '###.###.###.###.###")

2.4.2.9.3. Tráfico contenedores (Puertos)

=num(sum({<Data_Type={'World Development Indicators'},Data_SubType={'Container - Port Traffic'},Data_SubType_Detail=,Sex_Description= >}Value),'###.###.###.###.###')

2.4.2.9.4. Igualdad de genero

num(avg({<Data_Type={"World Development Indicators"},Data_SubType={"Rating - Gender Equality"},
Data_SubType_Detail=,Sex_Description=,Age=, Age_Range=, Age_Type= >}Value),'#,#')

2.4.2.9.5. Calidad Administración Pública

num(avg({<Data_Type={"World Development Indicators"},Data_SubType={"Rating - Quality Public
Administration"}, Data_SubType_Detail=,Sex_Description=,Age=, Age_Range=, Age_Type=
>}Value), "#,")

2.4.2.9.6. Protección Social

num(avg({<Data_Type={'World Development Indicators'},Data_SubType={'Rating - Social Protection'},
Data_SubType_Detail=,Sex_Description=,Age=, Age_Range=, Age_Type= >}Value),'#,#')

2.4.2.9.7. Transparencia Sector Público

num(avg({<Data_Type={'World Development Indicators'},Data_SubType={'Rating - Transparency Public Sector'}, Data_SubType_Detail=,Sex_Description=,Age=, Age_Range=, Age_Type= >}Value),'#,#')

2.4.2.10. Energía

Esta métrica se refiere a la energía producida por región, país y año. Este es el código de la métrica para la energía solar (a modo de ejemplo):

=num(sum({<Data_Type={'Energy'},Data_SubType={'Solar'}, Data_SubType_Detail=,Sex_Description=>}Value),'###.###.###.###.###.")

2.4.2.11. Plantas de energía

Esta métrica se refiere a la capacidad de producción de las plantas de energía por región, país y año. Este es el código de la métrica para la capacidad instalada (a modo de ejemplo):

=num(sum({<Data_Type={'Energy'},Data_SubType={'Power Plants'}, Data_SubType_Detail={'Electricity - net installed capacity of electric power plants public solar', 'Electricity - net installed capacity of electric power plants, public combustible fuels', 'Electricity - net installed capacity of electric power plants, public hydro', 'Electricity - net installed capacity of electric power plants, public hydro', 'Electricity - net installed capacity of electric power plants, public hydro', 'Electricity - net installed capacity of electric power plants, public wind', 'Electricity - net installed capacity of electric power plants, self-producer combustible fuels', 'Electricity - net installed capacity of electric power plants, self-producer geothermal', 'Electricity - net installed capacity of electric power plants, self-producer solar', 'Electricity - net installed capacity of electric power plants, self-producer tide, wave, marine', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind', 'Electricity - net installed capacity of electric power plants, self-producer wind',

2.4.3. Interfaz del Cuadro de Mando

El cuadro de Mando analítico, implementado con Qlikview, se estructura en cinco partes claramente diferenciades, con el objetivo de ayudar al usuario a analizar la información de punto a punto, pero además ofreciendo la posibilidad de adaptar los análisis con dinamismo en los indicadores.

2.4.3.1. Introducción

La primera pantalla que un usuario se encontrará es la de introducción, que será el lugar donde se le ofrezca una primera descripción de la solución. El objetivo de esta página es ayudar al usuario, por lo que se incorporará toda la información relevante, incluyendo el acceso al manual de usuario y a un video tour en mp4.



Figura 20 - Página Introducción del Cuadro de Mando

2.4.3.2. Disponibilidad información

En esta segunda página, el usuario podrá comprobar la información disponible para los diferentes países (verde disponible y rojo no disponible).



Figura 21 - Página Disponibilidad Información del Cuadro de Mando

El usuario podrá filtrar por región y/o país, y esta selección se aplicará al resto de hojas del Cuadro de Mando.



Figura 22 - Filtrado tabla disponibilidad información

2.4.3.3. KPIs

Esta es la primera hoja de análisis, donde se han definido una serie de objetos de análisis de tipo velocímetro para analizar las diferentes métricas. Se mostrará con la fecha la situación del dato respecto al total de todos los países, indicándose el valor numérico y el crecimiento/decrecimiento respecto al año de comparación.

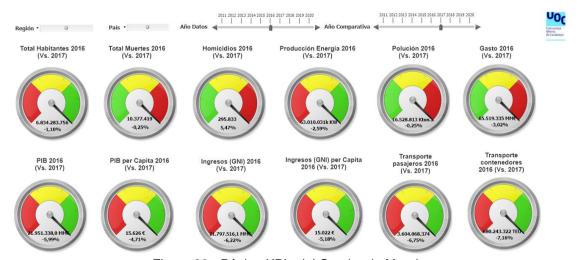


Figura 23 - Página KPIs del Cuadro de Mando

El elemento diferencial de esta hoja, y que aporta dinamismo desde el análisis, es que, además de los filtros, se incorporan dos selectores de años que permiten cambiar la elección del año que te gustaría usar como base del análisis (Año Datos) y el año con el que lo quieres comparar (Año Comparativa).



Figura 24 - Selectores años

En caso que no se dispongan datos para una selección concreta, el sistema lo indicará en los análisis que aplique.

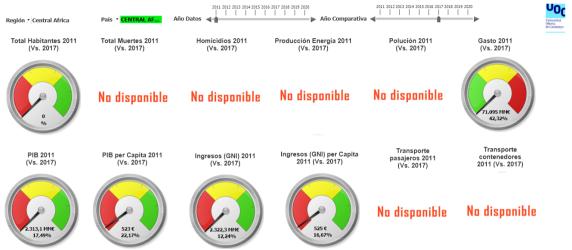


Figura 25 - Funcionamiento "No Disponible"

2.4.3.4. Análisis en Detalle

En esta hoja es donde el usuario podrá analizar con más profundidad los datos. En la parte superior encontrará una botonera, que le permitirá navegar por las diferentes opciones de análisis.



Además, el usuario dispone de filtros:



Figura 27 - Filtros

Y los selectores de años:



Figura 28 - Selectores años

2.4.3.4.1. Población

En este primer análisis trabajaremos con la métrica de población, donde se profundizará en los datos disponibles de varios modos:

Indicadores de análisis

En estos primeros indicadores se muestra el dato del año elegido en el selector y se compara el dato con el año a comparar, tanto a nivel total como por cada sexo.

Población 2011 Evolución Vs. 2017
6.801.590.452 -1,57%
Evolución Hombres Vs. 2017
-1,61% Evolución Mujeres Vs. 2017
-1,53%

Figura 29 - Población: Indicadores

• Evolución anual de la población (incluye línea de tendencia) En este gráfico se muestra la suma de la población para la selección realizada y para cada uno de los años con datos, además de la tendencia lineal de los mismos.

Evolución población



Figura 30 - Población: Evolución

Distribución de la población por sexo

En este gráfico de tarta se muestra el porcentaje de los diferentes tipos de sexo para la selección realizada.

Distribución por sexo

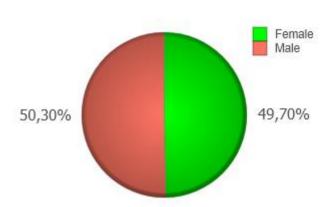
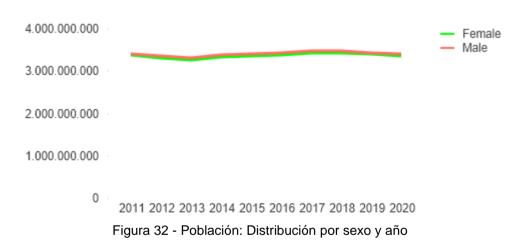


Figura 31 - Población: Distribución por sexo

Distribución de la población por sexo y año
 Por último, en este último análisis se puede ver la evolución en función del sexo.

Evolución Distribución por sexo



2.4.3.4.2. Muertes

En este segundo análisis trabajaremos con la métrica de muertes, donde se profundizará en los datos disponibles de varios modos:

Indicadores de análisis

En estos primeros indicadores se muestra el dato de muertes del año seleccionado y las comparaciones con el año que el usuario eligió.

Muertes 2012 10.142.435

Evolución Hombres Vs. 2017 -3,02%

Evolución Vs. 2017 -2,51%

Evolución Mujeres Vs. 2017 -1,96%

Figura 33 - Muertes: Indicadores

• Evolución anual de la población (incluye línea de tendencia)

En este primer análisis se muestras los datos de muertes por año, además de la tendencia de estos.

Evolución muertes



Figura 34 - Muertes: Evolución

Distribución de la población por sexo
 Con este análisis de puede ver la proporción de muertes por sexo.

Distribución por sexo

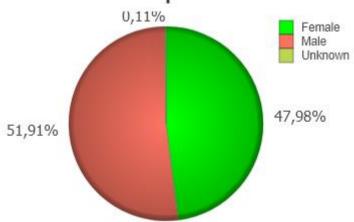
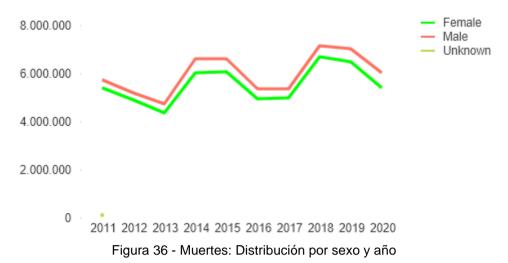


Figura 35 - Muertes: Distribución por sexo

Distribución de la población por sexo y año
 Por último, también se muestra un gráfico evolutivo, pero con los datos por sexo (anuales).

Evolución Distribución por sexo



2.4.3.4.3. Homicidios

En este análisis trabajaremos con la métrica de muertes, donde se profundizará en los datos disponibles de varios modos:

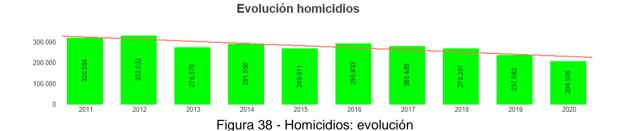
Indicadores de análisis

Homicidios 2011 **320.584**

Evolución Vs. 2017 14,29%

Figura 37 - Homicidios: indicadores

• Evolución anual de los homicidios



Comparación entre homicidios y muertes

Homicidios vs. Muertes

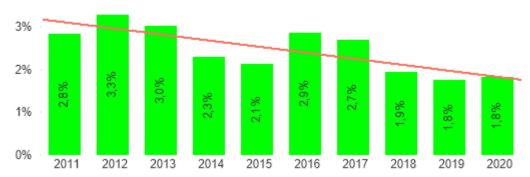


Figura 39 - Homicidios: comparación homicidios y muertes

Comparación entre homicidios y población

Homicidios vs. Población



Figura 40 - Homicidios: comparación homicidios y población

2.4.3.4.4. Contaminación

En este análisis trabajaremos con la métrica de contaminación, donde se profundizará en los datos disponibles de varios modos:

Indicadores de análisis

Contaminación 2011 17.284.618

Evolución Vs. 2017 4,31%

Figura 41 - Contaminación: indicadores

Evolución anual de la contaminación

Evolución contaminación

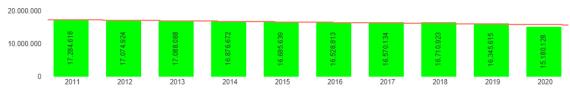


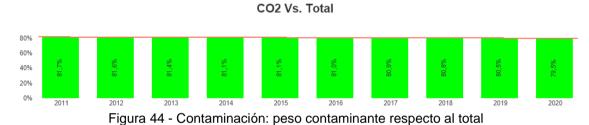
Figura 42 - Contaminación: evolución anual

 Peso de los diferentes contaminantes cada año (solo disponible en el análisis total)



Figura 43. Contaminación: peso contaminantes

 Peso individual de cada contaminante respecto al total cada año (solo disponible para el análisis de cada contaminante)



2.4.3.4.5. Gasto

En esta sección se puede analizar los datos de PIB (total y por tipo de industria) y GNI. Se incluyen los siguientes análisis:

Indicadores de análisis

Valor - 2011 Evolución Vs. 2017 **60.328.447.241.525,3€** -10,70%

Figura 45 - Gasto: indicadores análisis

Análisis evolutivo

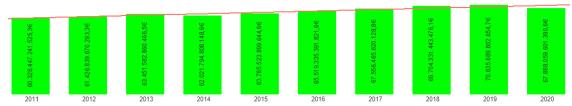


Figura 46 - Gasto: análisis evolutivo

Es importante resaltar que los datos per cápita que se muestran en todos los análisis es una media (en vez de la suma que aplica al resto), para asegurar el correcto análisis.

2.4.3.4.6. Energía

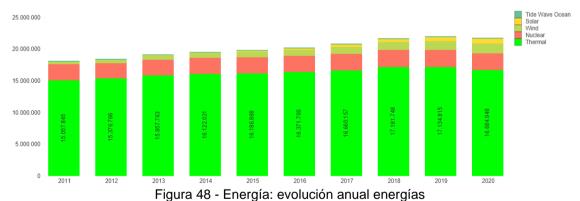
En esta sección se puede analizar los datos de energía (total, por tipo de energía y de las plantas de energía). Se incluyen los siguientes análisis:

Indicadores de análisis

Energía Total 2011 18.091.479 Evolución Vs. 2017 -13,01%

Figura 47 - Energía: indicadores

 Evolución anual por tipo de energía con el peso de las energías para cada año



Evolución anual de una energía concreta



Figura 49 - Energía: evolución anual energía concreta

Evolución de la capacidad de energía de las plantas



Figura 50 - Energía: evolución anual capacidad plantas energía

2.4.3.4.7. Indicadores Desarrollo: Transporte

En esta sección se puede analizar los datos de transporte (mercancías, pasajeros y contenedores). Se incluyen los siguientes análisis:

Indicadores

Valor - 2011 Evolución Vs. 2017 29.955.855 -14,28%

Figura 51 - Indicadores desarrollo transporte: indicadores

Evolución anual



Figura 52 - Indicadores desarrollo transporte: evolución

2.4.3.4.8. Indicadores Desarrollo: Ratings

En esta sección se puede analizar los datos de indicadores variados (igualdad de género, calidad de la Administración Pública, protección social y transparencia del sector público). Estos indicadores se miden de 1 a 5, y en todos los análisis se muestra una media.

Se incluyen los siguientes análisis:

Indicadores

Valor - 2011 **3,4** **5,99%**

Figura 53 - Indicadores desarrollo (ratings): indicadores

• Top 10 países

TOP 10 Países

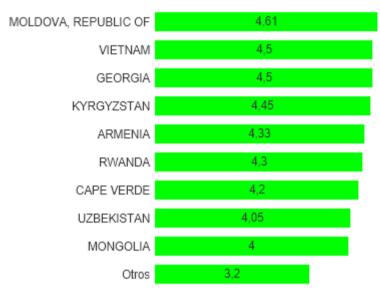


Figura 54 - Indicadores desarrollo (ratings): Top 10 países

Evolución anual

Evolución



Figura 55 - Indicadores desarrollo (ratings): Evolución

2.4.3.4.9. Refugiados

En esta última sección se analizan los datos de refugiados, tanto en base al país al que llegan como a su país de procedencia.

Se incluyen los siguientes análisis:

Indicadores

Refugiados 2011 9.689.275 Evolución Vs. 2017 -50,50%

Figura 56 - Refugiados: indicadores

Evolución anual

Evolución Refugiados



Figura 57 - Refugiados: evolución anual

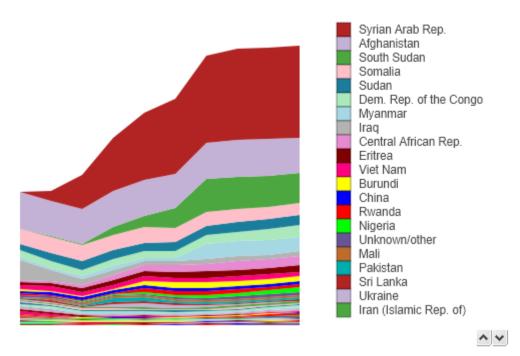
• Top 10 países de procedencia



Figura 58 - Refugiados: Top 10 países procedencia

Desglose por país de procedencia

Desglose por país de procedencia



2011 2012 2013 2014 2015 2016 2017 2018 2019 2020

Figura 59 - Refugiados: desglose por país de procedencia (anual)

2.4.3.5. Extractos

En esta página un usuario podrá configurarse una tabla con los campos que necesita y exportársela a Excel.

En la parte izquierda se muestran todas las dimensiones y métricas disponibles, que se pueden añadir a la tabla seleccionándolas:

□ □ Dimensiones □ Año □ Edad □ Pais □ Region □ Sexo	□ Indicadores □ □ Demografía □ Habitantes □ Homicidios □ Muertes □ Refugiados
□ Indicadores □ □ Gasto □ GNI □ Industrias □ PIB □ □ Indicadores □ □ Contaminación □ CO2	□ Indicadores □ □ Energia □ Calorífica □ Mareas □ Nuclear □ Plantas Energía □ Solar □ Viento
Hexafluoruro Azufre Hidrofluorocarbonos Metano N2O Perfluorocarbonos	□ Indicadores □ Indicadores Desarrollo □ Tráfico contenedores (Puertos) □ Transporte Aéreo (Mercancías) □ Transporte Aéreo (Pasajeros)
□ Indicadores □ Indicadores Desarrollo □ Calidad Administración Pública (Rating) □ Igualdad Género (Rating) □ Protección Social (Rating) □ Transparencia Sector Público (Rating)	

Figura 60 - Extractos: dimensiones y métricas

Y en la parte superior se encuentran los filtros posibles:



Por último, según se seleccionan las dimensiones y métricas, se irán mostrando en la tabla:

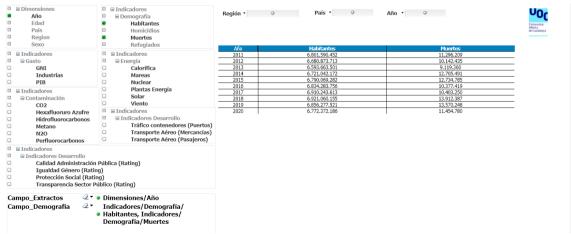


Figura 62 - Extractos: generación tabla

Y se puede exportar la tabla a Excel:

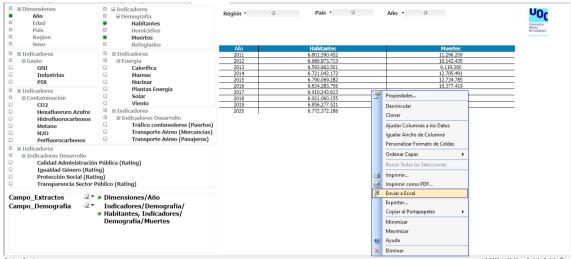


Figura 63 - Extractos: exportar tabla a Excel

Por último, es importante mencionar que todas las selecciones aparecerán en una tabla en la parte inferior izquierda de la pantalla:



Figura 64 - Extractos: selecciones realizadas

2.5. Plan de pruebas

Una vez que el Cuadro de Mando está desarrollado, es importante probar su funcionamiento. Para poder verificarlo, he diseñado un plan de pruebas, con el objetivo de asegurar su correcto funcionamiento.

Aunque todo el detalle del plan de pruebas se encuentra en un anexo, a continuación, detallaré la información principal, la estrategia de pruebas y las pruebas a realizar, con el resultado obtenido:

Identificador: PPC1

Elemento de prueba: Cuadro de Mando de las Naciones Unidas

Riesgos identificados: Ninguno

Características que no se probarán: Ninguna

Estrategia de pruebas:

- Se abrirá el Cuadro de Mando.
- Se seguirá paso a paso con lo definido en cada prueba.
- Se realizarán capturas de los resultados obtenidos, para que sirvan como elemento probatorio.

Pruebas a realizar:

- E1. Análisis de KPIs para unos años concretos
 - Descripción de la prueba: Con esta prueba se persigue probar la correcta funcionalidad de los KPIs. Para ello, se tendrá que

seleccionar el año a analizar y el que servirá de comparador, para después comprobar que los análisis de los diferentes KPIs se ajustan a la selección.

- Pasos a seguir para su realización:
 - Abrir Cuadro de Mando
 - Entrar a la página de KPIs
 - Seleccionar en "Año Datos" el año 2017
 - Seleccionar en "Año Comparativa" el año 2015
 - Visualizar los análisis y comprobar que todo sale correcto
- E2. Extraer datos concretos de un país a excel
 - Descripción de la prueba: Con esta segunda prueba lo que se persigue es probar que la página de extractos funciona correctamente, realizando un ejemplo sencillo de extracción de información.
 - Pasos a seguir para su realización:
 - Abrir Cuadro de Mando.
 - Entrar a la página de Extractos.
 - Filtrar por el país "Germany".
 - Seleccionar la dimensión "Año".
 - Seleccionar los indicadores "Nuclear", "Habitantes" y "Muertes".
 - Pulsar con el botón derecho del ratón sobre la tabla y seleccionar "Enviar a excel".
- E3. Funcionamiento de los análisis tras filtrar datos por sexo
 - Descripción de la prueba: No todos los análisis ofrecen detalle por sexo y se debe asegurar que filtrar por esta dimensión no afecta a esos análisis.
 - Pasos a seguir para su realización:
 - Abrir Cuadro de Mando.
 - Entrar a la página de Extractos.
 - Seleccionar la dimensión "Sexo".
 - Filtrar por "male".
 - Ir a la página de "Disponibilidad de información" y comprobar que todos se muestran correctamente.
 - Ir a la página de "KPIs" y comprobar que todos se muestran correctamente.
 - Ir a la página de "Análisis en detalle" y comprobar que los análisis de estas pantallas se muestran correctamente:
 - Homicidios
 - Contaminación
 - Gasto
 - Energía
 - Indicadores Desarrollo
 - Refugiados
- E4. Funcionamiento de los análisis tras filtrar datos por un grupo de edades

- Abrir Cuadro de Mando.
- Entrar a la página de Extractos.
- Seleccionar la dimensión "Edad".
- Filtrar por "3".
- Ir a la página de "Disponibilidad de información" y comprobar que todos se muestran correctamente.
- Ir a la página de "KPIs" y comprobar que todos se muestran correctamente.
- Ir a la página de "Análisis en detalle" y comprobar que los análisis de estas pantallas se muestran correctamente:
 - Homicidios
 - Contaminación
 - Gasto
 - Energía
 - Indicadores Desarrollo
 - Refugiados
- E5. Correcto funcionamiento del aviso "No Disponible"
 - Descripción de la prueba: Por último, hay que comprobar que aparece "No Disponible" cuando no hay datos para un análisis en el que no se disponga de información.
 - Pasos a seguir para su realización:
 - Abrir Cuadro de Mando.
 - Entrar a la página de "Disponibilidad Información".
 - Filtrar por "Hong Kong"
 - Ir a la página de KPIs
 - Comprobar que aparece como "No Disponible" el KPI de Contaminación.
 - Ir a la página de Análisis en detalle
 - Comprobar que los análisis de Contaminación aparecen como "No Disponible"

Entregables:

- El plan de pruebas
- Para cada caso del plan de pruebas, se entregará una captura de pantalla con el resultado que se ha obtenido, y se indicará si el resultado de la prueba es correcto o error.

Resultado de la ejecución (las evidencias se encuentran en el anexo del Plan de Pruebas):

- E1. Análisis de KPIs para unos años concretos (Correcto)
- E2. Extraer datos concretos de un país a Excel (Correcto)
- E3. Funcionamiento de los análisis tras filtrar datos por sexo (Correcto)
- E4. Funcionamiento de los análisis tras filtrar datos por un grupo de edades (Correcto)
- E5. Correcto funcionamiento del aviso "No Disponible" (Correcto)

3. Conclusiones

3.1. Lecciones aprendidas

Este proyecto me ha permitido poner en práctica muchos conocimientos aprendidos durante la carrera, desde la correcta planificación de un proyecto (desde la EDT al plan de proyecto con sus hitos) hasta la fase de pruebas.

Un aspecto en el que este proyecto me ha permitido profundizar es en la parte de la generación del modelo de datos. Cuando se trabaja en un sistema analítico, el foco suele estar en la herramienta y se suele dar poca importancia a las tareas de preparación de los datos. Gracias a la realización de este proyecto y a trabajar con datos reales, he podido experimentar los diferentes problemas que surgen durante esta etapa y encontrar maneras, a veces imaginativas, para solventarlos.

Por último, el trabajo en el cuadro de mando me ha permitido poner en práctica todo lo aprendido en temas de visualización y explotación de la información, para poder dar información de forma visual y atractiva, facilitando la toma de decisiones.

3.2. Objetivos conseguidos

En cuanto a los objetivos que planteé inicialmente para este trabajo, se han conseguido de forma satisfactoria, a pesar de ser extremadamente ambiciosos.

A base de una estricta planificación y aprovechar cada minuto de mi tiempo libre, he conseguido ir cumpliendo con los hitos comprometidos, incluso adelantándome a la planificación en algunos de ellos.

Un aspecto clave bajo mi punto de vista para lograr alcanzar lo que me propuse es la definición del alcance. Conseguir limitar el alcance a algo lograble en 3 meses de trabajo era clave para el éxito del proyecto y, a la vista de los resultados, he estado muy acertado.

3.3. Planificación y metodología

Tanto la planificación como la metodología que he seguido han sido correctas, ya que he podido organizarme para lograr cumplir con todos los hitos del proyecto.

Si bien es verdad que he tenido que realizar un esfuerzo alto para cumplir con los compromisos ajustados de fechas que planifiqué, siento que ha merecido la pena y estoy muy satisfecho con los resultados.

3.4. Trabajos futuros

Uno de los aspectos que me gustaría trabajar en el futuro es convertirlo en un sistema adaptable al idioma del usuario.

Para lograrlo, mi idea es implementar cambios en el modelo de datos y en el Cuadro de Mando:

- A nivel del modelo de datos, se deberá crear:
 - Una tabla de traducciones para todos los textos del Cuadro de Mandos, con una columna por idioma.
 - o Incluir una tabla maestra de dimensiones, para poder traducir todos los valores de las dimensiones incluidas.
 - Una columna en todas las tablas maestras ya existentes por cada idioma incluido.
- A nivel Cuadro de Mando, se deberá implementar:
 - Unas imágenes para que los usuarios seleccionen el idioma. La selección se almacenará en una variable.
 - Modificar todas las dimensiones para que se construyan de forma dinámica, consiguiendo de este modo que con una selección del usuario todo cambie.

Esta mejora permitirá que la solución sea más internacional, lo que ampliará los potenciales usuarios de la herramienta.

4. Glosario

Open Data: Datos disponibles en abierto para que cualquier persona pueda consumirlos de forma gratuita.

Dataset: Conjunto de datos relativo a una información concreta y definida.

Cuadro de Mando: Aplicación que permite el análisis de datos de forma interactiva por parte de los usuarios.

CSV: Formato de fichero de texto muy utilizado en intercambios de información e integraciones.

QVD: Formato de fichero propietario de Qlik que almacena los datos de un objeto y permite cargas rápidas y eficientes.

5. Bibliografía

Gartner | Información práctica y objetiva para ejecutivos y sus equipos. (s. f.). Gartner. https://www.gartner.es/es (2023, 15 marzo)

Pentaho Download - Lumada Data Integration and Analytics. (s. f.). Hitachi Vantara. https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/data-integration-analytics/download-pentaho.html?ecid=ps_emea_dx_en_sssem340">https://www.hitachivantara.com/en-us/products/dataops-software/dataops-softwa

Cree informes sofisticados y comparta conocimientos que impulsen los resultados | Power Bl. (s. f.). https://powerbi.microsoft.com/es-es/landing/free-account/? (2023, 15 marzo)

Qlik. (s. f.). https://www.qlik.com/es-es/ (2023, 15 marzo)

Talend.Talend | A Complete, Scalable Data Management Solution. Talend - A Leader in Data Integration & Data Integrity. https://www.talend.com/ (2023, 15 marzo)

Informatica. (s. f.). PowerCenter: Enterprise Data Integration Platform | Informatica España. https://www.informatica.com/es/products/data-integration/powercenter.html (2023, 15 marzo)

Pentaho Data Integration. (2023, 31 enero). Hitachi Vantara Lumada and Pentaho Documentation. https://help.hitachivantara.com/Documentation/Pentaho/8.3/Products/Pentaho_Data_Integration (2023, 15 marzo)

UNdata | explorer. (s. f.). http://data.un.org/Explorer.aspx (2023, 15 marzo)

6. Anexos

6.1. Source files.xlsx

Documento de Excel donde se encuentran las fuentes potencialmente interesantes y cuáles son las fuentes finalmente seleccionadas. Esta es una captura de las fuentes elegidas (verde) y descartadas (amarillas) en el análisis final:



Figura 65 - Anexo: Source_files.xlsx

6.2. Cuadro de Mando de las Naciones Unidas.pod

Plan de proyecto diseñado con la herramienta gratuita ProjectLibre.

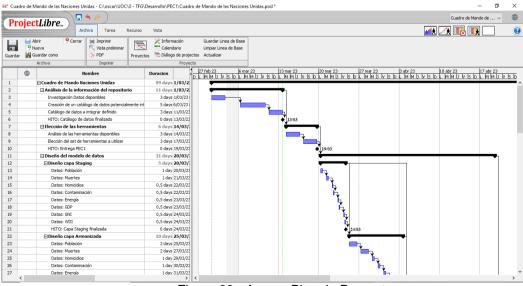


Figura 66 - Anexo: Plan de Proyecto

6.3. TFG_OQ_Layers.QVW

Archivo de Qlikview que se encarga de construir las tres capes del modelo de datos, generando los ficheros en cada una de ellas.

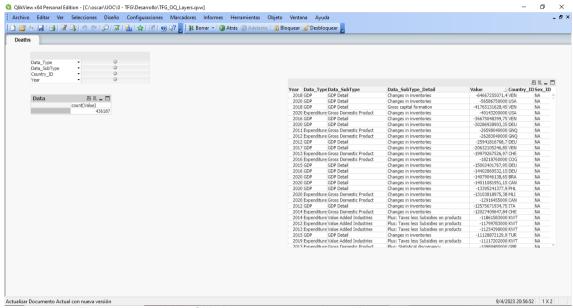


Figura 67 - Archivo Qlikview para generar las diferentes capas de datos

6.4. Ficheros del modelo de datos

Archivo .Zip con los archivos generados para cada una de las capas de datos.



Figura 68 - Archivo .Zip con las capas de datos

6.5. Archivo para instalar Qlikview en el PC (QlikViewDesktop_x64Setup.exe)

Instalador de la herramienta (64 bits).

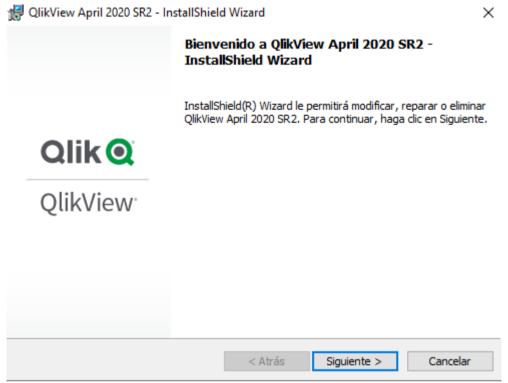


Figura 69 - Instalador Qlikview

6.6. Anexo con los scripts del modelo de datos

Documentación de los scripts utilizados para generar el modelo de datos.

TFG INGENIERÍA INFORMÁTICA Business Intelligence



Cuadro de Mando de las Naciones Unidas: Scripts Modelo de Datos

Figura 70 - Anexo scripts modelo de datos

6.7. Cuadro de Mando

Archivo del Cuadro de Mando de las Naciones Unidas.



Figura 71 - Cuadro de Mando de las Naciones Unidas

6.8. Plan de Pruebas (Anexo PlanPruebas.pdf)

Plan de pruebas completo.

TFG INGENIERÍA INFORMÁTICA Business Intelligence



Cuadro de Mando de las Naciones Unidas: Plan de Pruebas PPC1

Óscar Quero Hijano

Figura 72 - Plan de Pruebas

6.9. Manual de usuario (Manual_Usuario.pdf)

Manual de usuario

TFG INGENIERÍA INFORMÁTICA Business Intelligence



Cuadro de Mando de las Naciones Unidas: Manual de Usuario

Óscar Quero Hijano

Figura 73 - Manual de usuario

6.10. Video tour (Video_Tour.mp4)

Video con un tour comentado por la herramienta.



Figura 74 - Video tour

6.11. Presentación defensa (pptx y pdf)

Archivos con la presentación.

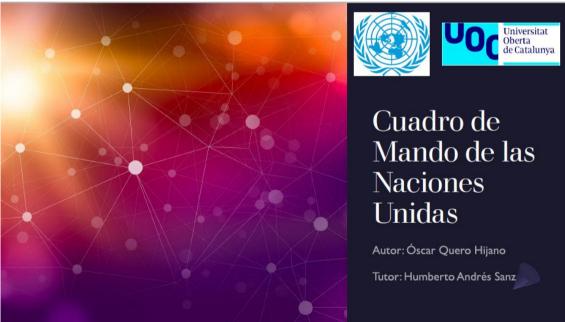


Figura 75 - Presentación Defensa

6.12. Video defensa (mp4)

Video con la presentación.



Figura 76 - Video Defensa