UNIVERSIDAD INTERNACIONAL DE LA RIOJA unir

Universidad Internacional de La Rioja (UNIR)

Escuela de Ingeniería

Máster en Visual Analytics y Big Data

Democratización de los Presupuestos Generales del Estado

Trabajo Fin de Máster

Presentado por: Cobos Raposo, Sandra

Director/a: de la Fuente Valentín, Luis

Ciudad: Madrid
Fecha: Abril, 2016

Resumen

Este trabajo es un primer acercamiento gráfico de las cuentas del Gobierno a la población. El objetivo consiste en analizar y visualizar de una manera atractiva los resúmenes presupuestarios de los ingresos y gastos que ha realizado el Gobierno desde 2007 hasta 2016. De esta manera, el ciudadano puede entender mejor la distribución de los impuestos y el gasto que supone cada cartera ministerial, otorgando al lector una visión financiera básica de la situación económica en España.

Para llevarlo a cabo, se emplean técnicas avanzadas en el tratamiento de datos, análisis de contenido y visualización interactiva de la información.

Para medir si se ha logrado el propósito, se realiza un estudio final que valora positiva o negativamente el uso de estas técnicas en el ámbito de la macro economía.

Palabras Clave: presupuestos, Big Data, visualización, datos abiertos

Abstract

This is an attempt to create a first join between The General State Budget and the citizen. The objective consists in an attractive layout analazing and showing the revenue and expenditure. The summary budget that you can find is from the years 2007 to 2016. That way, the citizen gets a better understanding on the taxes' distribution and the expense accounts of each ministerial portfolio. So, the result is a well-informed citizen with minimal effort.

For the realization of this project Big Data techniques, related to the extraction of data from web sides, data analysis and data visualization, were used.

The project's success is measured by a final study with independent people, having to fill in a survey about their different impressions when comparing the original reports and the visual graphics created for this project.

Keywords: State Budget, Big Data, Visualization, open data

Índice de Contenido

Resumen ii

Abstract ii

1. INTRODUCCIÓN 1

1.1. Motivación 1

1.2. Problemática 2

1.1. Estructura del documento 3

2. OBJETIVOS 4

3. EXPOSICIÓN DEL CONTEXTO 5

4. ESTADO DEL ARTE 6

Análisis de la Información en prensa 8

Portugal 9

Francia 11

Otros países: Italia e Inglaterra 12

Alemania, un país pionero 14

5. DESCRIPCIÓN DEL PROBLEMA 17

5.1. Alcance de los datos 18

5.2. Estructura de la implementación 18

6. TECNOLOGÍAS 20

6.1. Motivación 20

6.2. Herramientas utilizadas en cada fase 20

6.3. Detalle de las herramientas 23

1. HTML 23

2. Datos abiertos 24

3. XML 24

4. Microsoft Excel 25

5. MongoDB 26

6. Robomongo 27

7. Tableau Public 27
8. Google Technology 28
8. METODOLOGÍA 29
1. Extracción orígenes de datos 30
1.1. Buscar orígenes de datos 30
1.2. ¿Se adaptan al propósito? 31
2. Coherencia y Análisis de datos 32
2.1. ¿Los conceptos orgánicos son consistentes? 36
2.2. Añadir al repositorio 36
3. Transformación de datos 37
3.1. Seleccionar la estructura de consolidación de los presupuestos 37
3.2. Normalizar los presupuestos 37
4. Almacenamiento de datos 39
4.1. Normalizar los presupuestos 39
5. Visualización de la información 39
5.1. Visualización de los datos en gráficos 40
6.1. Creación de la encuesta de satisfacción 41
9. RESULTADOS 43
9.1. Apariencia de la herramienta 43
9.2. Valoración de la herramienta 49
9.1. Otros datos de interés 50
10. CONCLUSIONES 52
10.1. Sobre la herramienta 52
10.2. Líneas futuras 52
BIBLIOGRAFÍA 54
ANEXOS 1
CORREOS ELECTRÓNICOS A LA DIRECCIÓN OFICIAL DEL MINISTERIO 1

Índice de Ilustraciones

Ilustración 1 - Revertir el valor añadido aportado por el análisis 1
Ilustración 2 - Fragmento de los pptos generales del estado 7
Ilustración 3 - Página de la secretaría de estado de presupuestos y gastos 8
Ilustración 4 - Visualización de los presupuestos de 2012 según El País 9
Ilustración 5 – Presupuestos Generales del Estado portugués 10
Ilustración 6 - Gráficos explicativos de los presupuesto generales del estado en Portugal 11
Ilustración 7 - Portal web de los presupuestos generales del estado francés 12
Ilustración 8 - Detalle gráfico de una comparativa en los presupuestos generales de Inglaterra 13
Ilustración 9 - Portal web de los presupuestos generales italianos 14
Ilustración 10 - Presupuestos generales del estado alemán 15
Ilustración 11 - Web privada con los presupuestos generales del estado alemán 16
Ilustración 12 - Fases de la implementación 19
Ilustración 13 - Relación de herramientas y tecnologías 20
Ilustración 14 - Fases del proyecto vs Tecnología aplicada 21
Ilustración 15 - Fragmento de código html 23
Ilustración 16 - Extracto de los Presupuestos Generales del Estado 2016 25
Ilustración 17 - Modificaciones contextuales 26
Ilustración 18 -Interfaz de Robomongo 27
Ilustración 19 - Diagrama de Proceso 29
Ilustración 20 - Aspecto del Portal Maestro 31
Ilustración 21 - Comparativa de los resúmenes de ingresos y gastos 34
Ilustración 22 - Incoherencias de los presupuestos referentes a los gastos 36
Ilustración 23 - Proceso de normalización esquemático 38
Ilustración 24 - Rueda de tensión de "El Arte Funcional" 39
Ilustración 25 - Encuesta de satisfacción 42
Ilustración 26 - Página Evolución Totales Ingresos y Gastos 44
Ilustración 27 - Página Desglose Ingresos y Gastos 45
Ilustración 28 - Detalle de la página Desglose Ingresos y Gastos 46
Ilustración 29 - Página Evolución por órgano ministerial 47
Ilustración 30 - Página Anexo: Evolución del Gasto 48
Ilustración 31 - Resultados de la encuesta 50
Ilustración 32 - Tiempos empleados 51

A mi madre,

y al resto de mujeres que luchan contra gigantes.

1. INTRODUCCIÓN

1.1. Motivación

Ante la dificultad que supone la divulgación de la utilidad del Big Data y la Visualización en todos los ámbitos sociales, me planteo como objetivo acercarla al mayor número de población posible, abarcando el ámbito laboral, estudiantil o, simplemente, de interés público. Por eso, cuando se me presenta la oportunidad de elegir un tema, me decanto por uno donde haya grandes lagunas que puedan resolverse dentro del marco del Proyecto y que esté fundamentado sobre una base social.

Por tanto, siguiendo la filosofía del Business Intelligence y del Big Data, me propongo diseñar un Proyecto cuyo último fin sea utilizar los datos para beneficio de la compañía que los provee. He elegido los datos proporcionados por el gobierno central referidos a los Presupuestos Generales del Estado. Como el organismo seleccionado es de ámbito público, **el beneficio debe ser útil y repercutir directamente sobre el propio ciudadano**.

Ilustración 1 - Revertir el valor añadido aportado por el análisis

También, me he fijado el reto de realizar un proyecto novedoso y gráfico para que el usuario entienda **el potencial de una la información bien estructurada** y tratada como un arma potente de análisis.

Finalmente, también me marco como propósito dar a entender el concepto de Big Data y Visualización de la información de manera didáctica, ya que esta noción es difícil de transmitir cuando se presenta de una manera teórica.

1.2. Problemática

Según Marta Beltrán y Fernando Sevillano en su publicación Cloud Computing, se estima que el 90% de la información vertida en Internet ha sido creada en los dos últimos años. Esto provoca que las estructuras clásicas de información se han colapsado debido al crecimiento exponencial de los datos públicos. Por tanto, el propio incremento de información en la red que las tecnologías también haya evolucionado, resultando nuevas estructuras que rompen con las ideas clásicas de almacenamiento de información.

Este proyecto se engloba en el marco de estas nuevas tecnologías que tienen el desafío de absorber, transformar y visualizar grandes conjuntos de datos y volverlos manejables para el usuario.

Realizando un análisis detallado de la economía del Ayuntamiento de Madrid, descubrimos el estado embrionario en el que se encuentran los portales de datos abiertos de las entidades locales del país. Los ficheros vertidos estaban mal estructurados, con información incompleta y, en algunos casos, manifestaban incoherencia en los datos.

La preocupación se incrementó al observar la falta de gráficos visuales que permitieran al usuario poder manejar la información pública. Ambos factores unidos provocaron el interés por generar y desarrollar una estructura basada en nuevas tecnologías que facilitara el entendimiento y tratamiento de la información disponible en el Ayuntamiento.

Sin embargo, una explosión de movimientos políticos y noticias presupuestarias contradictorias, nos animó a permutar el marco municipal por el maro estatal y así desarrollar este proyecto de forma que nos facilitara una información neutra de los presupuestos, sin partidismos ni especulaciones, convirtiéndolo en una herramienta útil y transparente para el ciudadano.

En los capítulos siguientes se desgranará cómo se ha realizado esta puesta en marcha, haciendo especial hincapié en las distintas fases del proyecto. También se prestará particular atención el flujo de trabajo y las decisiones técnicas que han llevado la preferencia de ciertas tecnologías frente a otras.

1.1. Estructura del documento

En este apartado se va a detallar la estructura sigue este documento. Así pues, el lector puede encontrar en el capítulo 2 los objetivos que se ha marcado la herramienta "Democratización de los Presupuestos Generales de Estado". En relación al contexto donde se enmarca, en el capítulo 3 se abarca el movimiento social que la germinó. Para comprender cómo los distintos gobiernos europeos aplican los conceptos de transparencia y democratización está en el capítulo 4.

A partir este punto ya se encuentran las soluciones específicas que se han tomado en relación a la herramienta como son las acotaciones en los datos tratadas en el capítulo 5 y las decisiones tecnológicas del capítulo 6.

Para comprender la estructura y la relación de las fases se debe consultar el capítulo 8 pero si el interés está relacionado con la discusión de resultados y las conclusiones, se debe visitar los capítulos 9 y 10 respectivamente.

2. OBJETIVOS

Como se ha comentado en el apartado anterior, **el objetivo principal de este proyecto es crear una estructura basada en las nuevas tecnologías de tratamiento de la información aplicada a los Presupuestos Generales del Estado**.

Además de este objetivo prioritario, se fijan las siguientes metas:

1. Estructurar los orígenes de datos obtenidos desde la web del Ministerio de Hacienda y Administraciones Públicas.
2. Normalizar los datos según la última versión de cada órgano o ministerio. Esto quiere decir que si un ente va cambiando de ministerio, dependa éste según la estructura del último año presupuestario.
3. Almacenar los datos para su posterior tratamiento.
4. Comparar los distintos ejercicios fiscales que se han aprobado desde 2007 hasta 2016 utilizando una herramienta de visualización.
5. Evaluar la herramienta de visualización según los siguientes criterios:
    - Grado de comprensibilidad de la herramienta.
    - Establecimiento de comparativas anuales y por concepto.
    - Implantar al menos 3 niveles de profundidad en el análisis

3. EXPOSICIÓN DEL CONTEXTO

Los últimos años han supuesto tiempos convulsos para la economía y la política de España ya que cada semana o cada mes, han aparecido nuevos casos de corrupción relacionados con el blanqueo de capital. Toda esta serie de escándalos ha hecho reflexionar al ciudadano común sobre esta situación; por tanto, ahora se solicita una mayor transparencia en las cuentas de los ayuntamientos, localidades e, incluso, del propio Estado.

Dada esta demanda, las distintas administraciones han comenzado a mostrar los presupuestos de una manera estandarizada, utilizando los portales digitales asignados para estos fines. Sin embargo, estos presupuestos se ponen a disposición utilizando una metodología contable especializada, por lo que el ciudadano debe tener grandes conocimientos financieros o grandes dosis de tiempo libre por mor de la complejidad de los mismos. Es en este punto donde nace el proyecto "*Democratización de los Presupuestos Generales del Estado*".

Analizando el contenido de la web oficial del Ministerio de Hacienda, se puede observar que la información que vierte al ciudadano es muy extensa pero poco útil, ya que ofrece gran cantidad de ficheros donde se albergan los presupuestos con un sistema de carpetas interrelacionado y poco intuitivo.

Por tanto, el objetivo de este proyecto es estructurar la información vertida en la página oficial del estado y realizar un primer acercamiento gráfico de los presupuestos al ciudadano común que no maneje conceptos contables. Para lograr esta meta, el usuario tiene que ser capaz de contestar a una serie de preguntas que le permitan descubrir si esta forma de representación de la información ayuda a un mejor entendimiento de la economía del país.

4.ESTADO DEL ARTE

Las plataformas digitales con datos abiertos son una nueva línea de información al alcance de todo el mundo. Siguiendo la estela de otras comunidades que tienen como filosofía la libre distribución (código abierto, licencia GPL, etc.), esta nueva ramificación aparece como complemento adicional a las existentes.

Esta vertiente surge de la necesidad de explotar los datos que las empresas generan en grandes cantidades y que no consiguen analizar debido a este volumen, perdiéndose así mucha información útil. Si en vez de poner la mira en las empresas privadas apuntamos a las entidades públicas, el dilema planteado sigue siendo el mismo.

Aprovechando estos pasos hacia el análisis masivo, los distintos gobiernos europeos han empezado a publicar en sus webs los datos abiertos que no vulneran las leyes de protección de datos y que son de interés público. Entre todos estos ficheros, están disponibles los presupuestos generales del estado.

Los presupuestos del estado son una serie de apuntes contables que reflejan los ingresos y los gastos que se realizarán a lo largo de un año contable. Estos movimientos pueden ir desde la compra de material de oficina hasta la venta de un edificio histórico. Aunque el detalle tan pormenorizado no está publicado debido a su falta de trazabilidad (eso sería otro proyecto), el sumatorio de todos estos gastos (o ingresos) sí debe preverse y estar publicado. Por tanto, el sumatorio de todos estos apuntes suele derivar en una tarea pesada que sólo el personal experto suele estar interesado en analizar o revisar.

El gobierno español publica cada año los presupuestos del estado de manera contable, es decir, una serie de ficheros diseñados en forma de tabla donde cada fila representa un movimiento contable. En el caso de la ilustración que tenemos a continuación, cada línea es un ingreso que se ha realizado en las arcas del estado.

PRESUPUESTOS GENERALES DEL ESTADO

ESTADO

Resumen general por capítulos del presupuesto de ingresos

EJERCICIO PRESUPUESTARIO
2016

Sección: 98 INGRESOS DEL ESTADO

(Miles de euros)

Económica	Explicación	Total
1	IMPUESTOS DIRECTOS Y COTIZACIONES SOCIALES	69.404.000,00
2	IMPUESTOS INDIRECTOS	43.476.000,00
3	TASAS, PRECIOS PUBLICOS Y OTROS INGRESOS	6.857.871,00
4	TRANSFERENCIAS CORRIENTES	8.267.478,28
5	INGRESOS PATRIMONIALES	5.179.636,00
	TOTAL OPERACIONES CORRIENTES	133.184.985,28
6	ENAJENACION DE INVERSIONES REALES	198.710,00
7	TRANSFERENCIAS DE CAPITAL	1.389.304,72
	TOTAL OPERACIONES DE CAPITAL	1.588.014,72
	TOTAL OPERACIONES NO FINANCIERAS	**134.773.000,00**
8	ACTIVOS FINANCIEROS	2.773.731,00
	TOTAL OPERACIONES FINANCIERAS	**2.773.731,00**
	TOTAL	**137.546.731,00**

Ilustración 2 - Fragmento de los pptos generales del estado

Previo a la temporada 1996, no existen registros públicos de los presupuestos generales del estado. Sin embargo, desde ese año se empiezan a publicar con un formato .pdf orientado esta publicación a una lectura más que a su análisis. Escasea también una representación gráfica que ayude al entendimiento de estos apuntes contables y permita al ciudadano una visión global de la economía española.

Sin embargo, a partir de 2007 se empiezan a publicar los presupuestos en ficheros con formatos más adaptados a la filosofía de datos abiertos. Entre estos formatos, el que más extendido está es el formato web .html, que incluye tablas con el desglose según su apartado. Este modelo de fichero ha sido el utilizado desde entonces gracias a su popularidad en el mundo digital. Como novedad, en los presupuestos publicados en 2016, se incorpora un nuevo formado a los ya publicados anteriormente. Este formato es el .xml (e**X**tensible **M**arkup **L**anguage). Se trata de un formato abierto, guiado por marcas y desarrollado por la comunidad W3C orientado a almacenar grandes volúmenes de información.

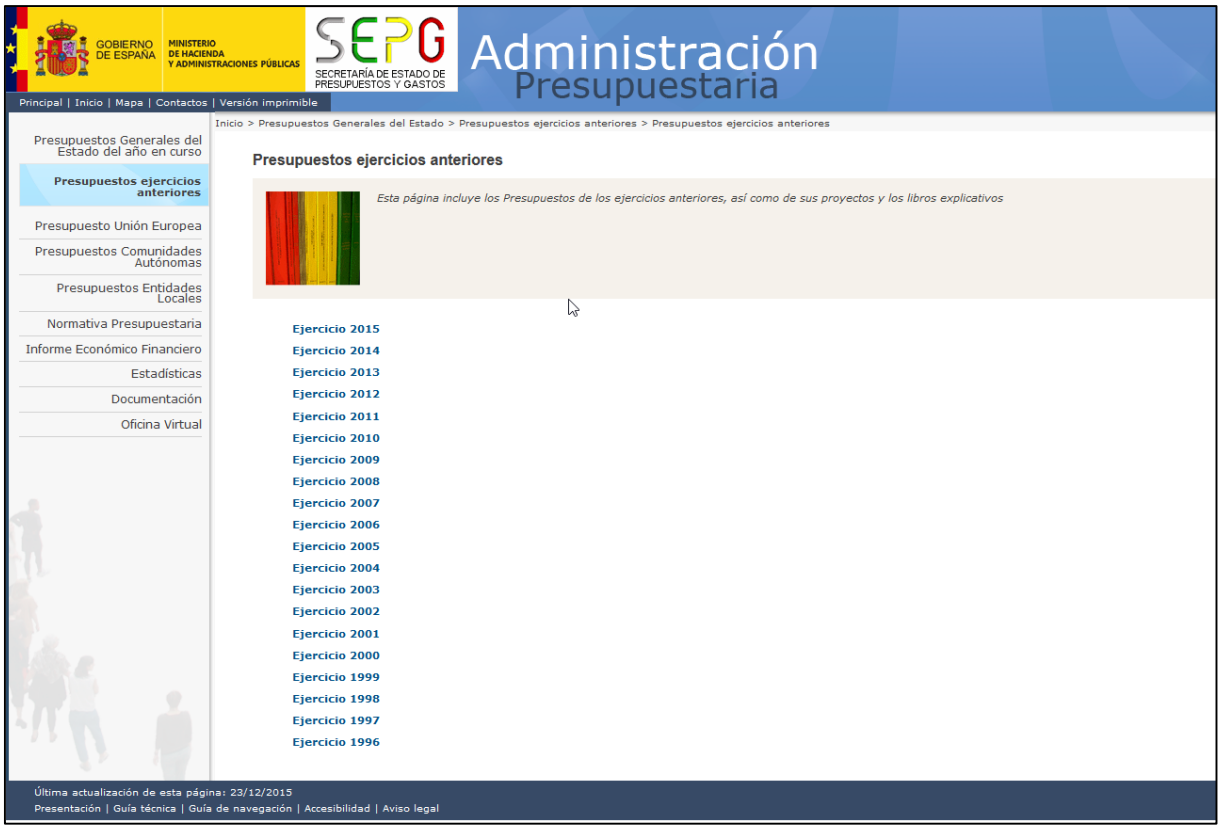


Ilustración 3 - Página de la secretaría de estado de presupuestos y gastos

También es importante destacar que los ejercicios presupuestarios se ofrecen en un formato comprimido para que el usuario lo descargue en forma de aplicación local. Si se abre esta aplicación, aparecen una serie de enlaces a los ficheros que componen los presupuestos. Estos ficheros estructurados por carpetas no son los únicos almacenados, ya que se incluye una gran documentación presupuestaria pesada de entender. En esta estructura de carpetas se entremezclan los distintos formatos (pdf y html) con los libros que recogen el detalle presupuestario de cada serie.

Sin embargo, queda en entredicho la eficacia de estas publicaciones ya que sólo una persona cualificada puede obtener algún análisis al respecto.

Análisis de la Información en prensa

El periódico de tirada nacional El País realizó una aproximación al desarrollo de este proyecto en 2012. Si se analiza la web donde publicó sus resultados, se puede observar un desglose presupuestario diferenciado entre ingresos, gastos y una comparativa con los años 2010 y 2011.

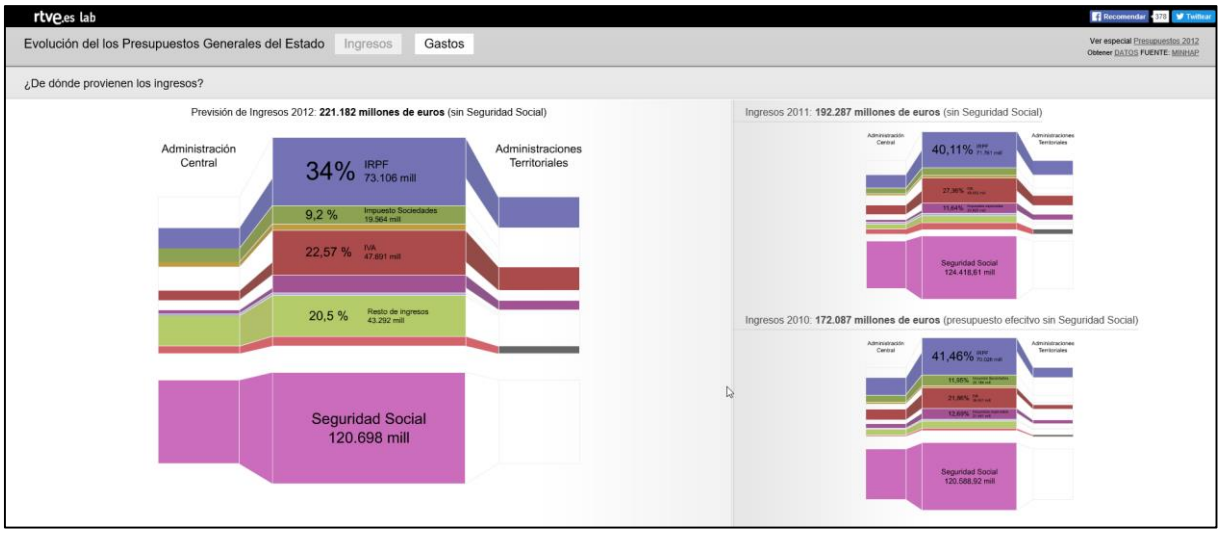


Ilustración 4 - Visualización de los presupuestos de 2012 según El País

La creación de este desglose vía web fue un gran acercamiento a los presupuestos pero que se quedó en una mera aproximación dada a la falta de continuidad en el tiempo. Si se compara el análisis realizado en 2012 con el que este mismo periódico publicó el de los presupuestos del 2016, se puede comprobar que ha optado por ofrecer la información manera seccionada y fuera de un contexto global.

A continuación, analizaremos el estado del arte en países europeos que guardan similitudes de desarrollo tecnológico con España. Cabe mencionar que se han analizado primordialmente las soluciones oficiales dado que la búsqueda de información presupuestaria en empresas privadas o medios de comunicación utilizando otros idiomas supone una barrera lingüista.

Portugal

Si se pone el punto de vista en los vecinos portugueses, el resultado es bastante parecido a la web oficial del ministerio español. En la página web oficial tienen publicado los presupuestos de manera contable, con un selector para elegir el año del ejercicio financiero.

Ilustración 5 – Presupuestos Generales del Estado portugués

La página web sólo refleja unos ficheros con formato pdf, pero sin ofrecer al ciudadano ficheros con formatos adaptados para un análisis posterior.

Como medida novedosa, incorpora para el año 2014 una serie de ficheros Excel que incorporan unos gráficos resumen. Sin embargo, estos gráficos están realizados de una manera poco rigurosa, ya que algunos de ellos no tienen la unidad de medida y las comparativas se realizan entre distintos años: unas entre 2013 y 2014, otras comparativas desde 2010 hasta 2014 y otras entre 2011 hasta 2014.

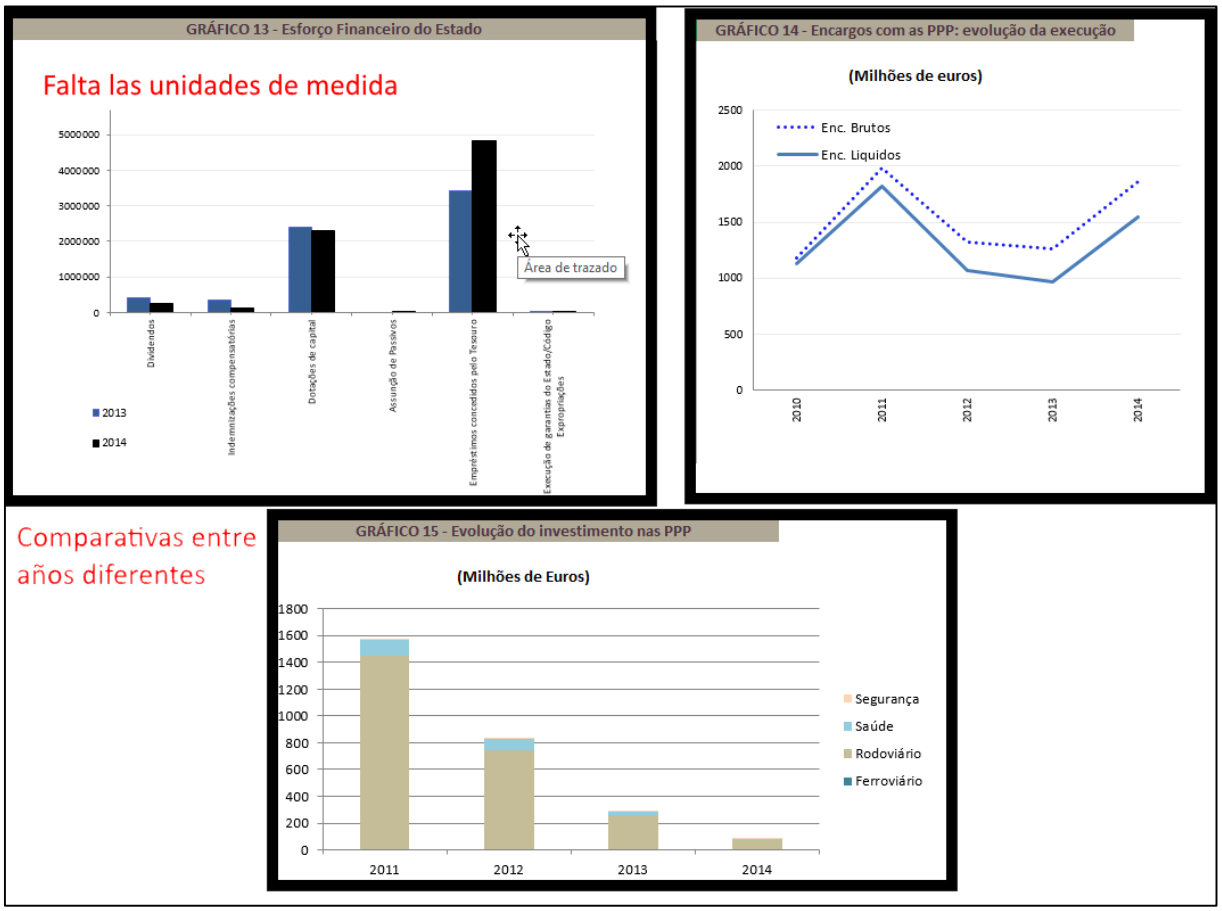


Ilustración 6 - Gráficos explicativos de los presupuesto generales del estado en Portugal

Por tanto, la información que se ofrece la web es poco rigurosa y estática, pudiendo inducir errores en su interpretación.

Francia

En el caso del portal de economía de la república francesa, la información publicada es bastante parecida al caso portugués. Aparecen ficheros explicativos con los movimientos presupuestarios desde 1996 hasta hoy. También se incluyen en estos ficheros algunos gráficos que apoyan la descripción textual.

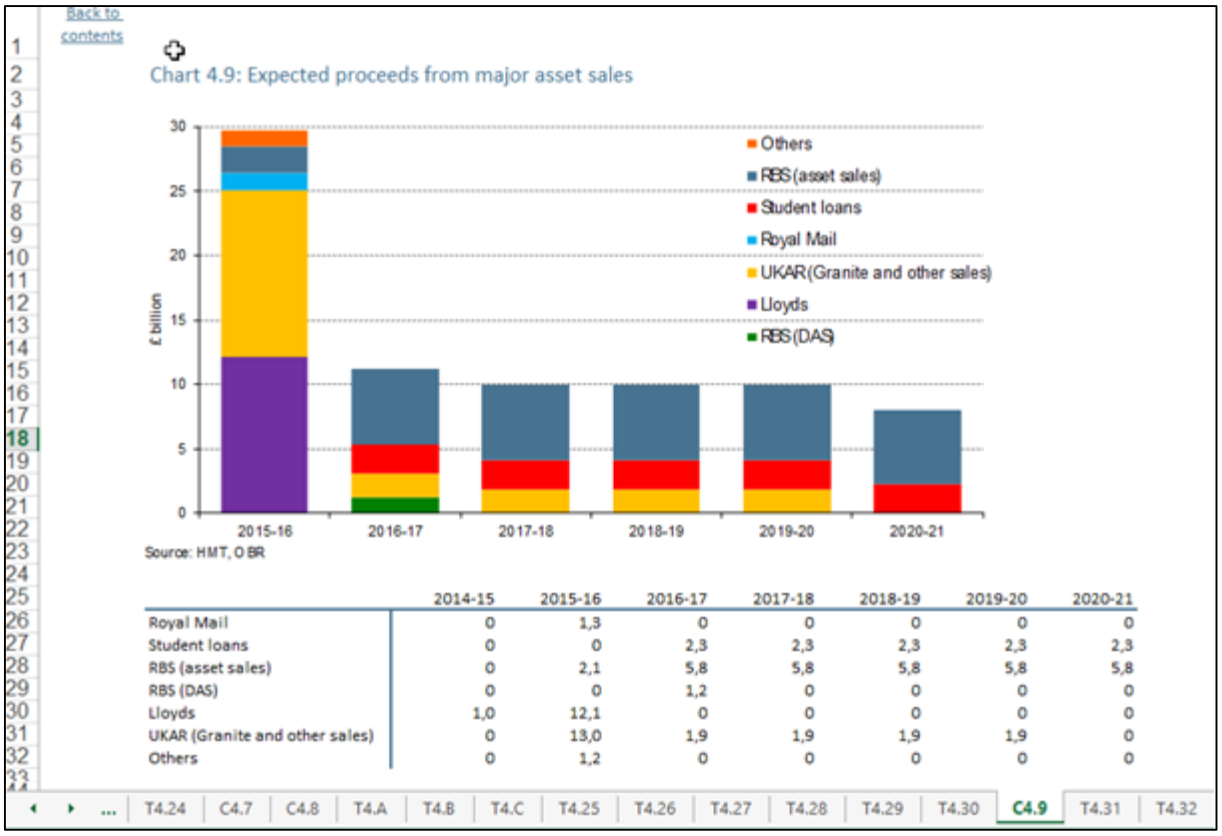
Sin embargo, el gobierno no ofrece al ciudadano la documentación de los presupuestos con formatos asociados a datos abiertos. Esto es que resultaría muy difícil para un usuario de la plataforma tomar sus propias decisiones.

Ilustración 7 - Portal web de los presupuestos generales del estado francés

Otros países: Italia e Inglaterra

Otros países donde se ha consultado esta información son Italia, Holanda e Inglaterra.

Para el caso de Inglaterra, en la página oficial del estado se encuentran los porfolios de los últimos presupuestos y, como pasaba en el caso portugués, se adjunta un fichero Excel bien estructurado con diagramas estáticos para el usuario.



	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21
Royal Mail	0	1,3	0	0	0	0	0
Student loans	0	0	2,3	2,3	2,3	2,3	2,3
RBS (asset sales)	0	2,1	5,8	5,8	5,8	5,8	5,8
RBS (DAS)	0	0	1,2	0	0	0	0
Lloyds	1,0	12,1	0	0	0	0	0
UKAR (Granite and other sales)	0	13,0	1,9	1,9	1,9	1,9	0
Others	0	1,2	0	0	0	0	0

Ilustración 8 - Detalle gráfico de una comparativa en los presupuestos generales de Inglaterra

Esto representa un gran progreso frente a los anteriores países pero con un gran inconveniente: los gráficos son imágenes copias y pegadas en el Excel, por lo que la capacidad de investigación del usuario es nula. También tiene como contrapartida que no puede reutilizarse ya que los orígenes de datos de los gráficos son desconocidos para el usuario.

En el caso italiano, la página web del estado sólo ofrece información en forma de tabla y no se han encontrado gráficos al respecto.

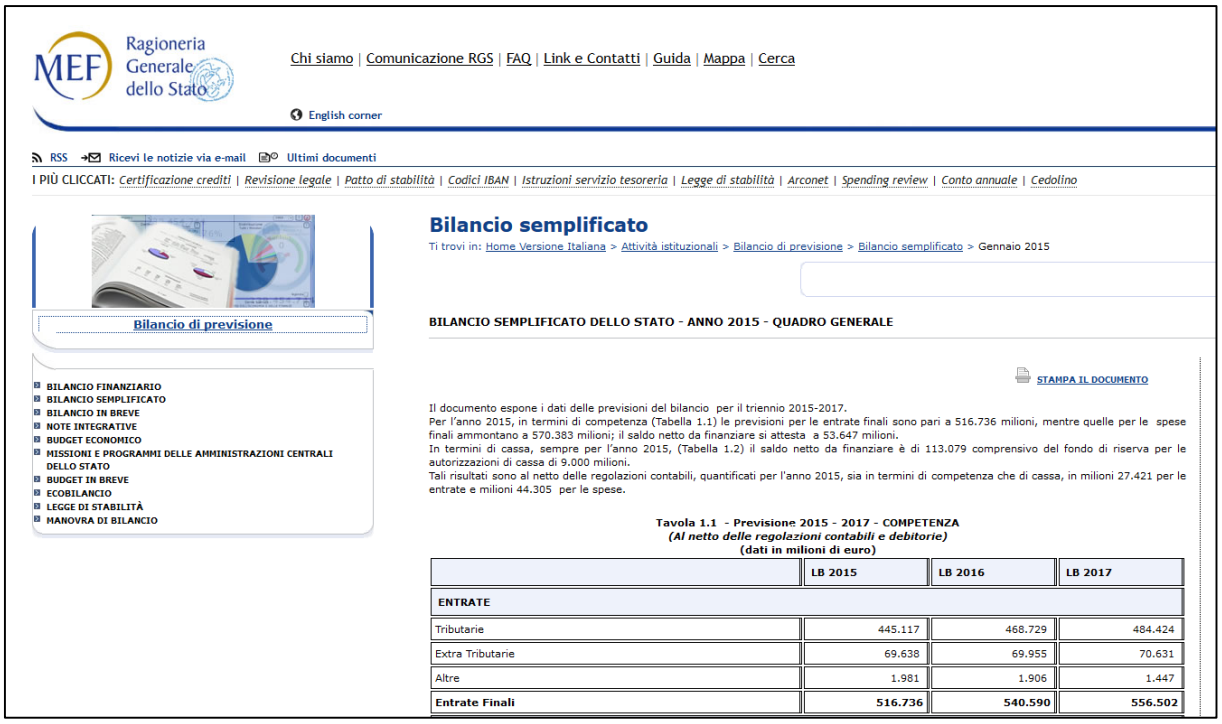


Ilustración 9 - Portal web de los presupuestos generales italianos

Alemania, un país pionero

Sin embargo, en el caso de Alemania los resultados son positivos. Es más, podría decirse que se asemeja mucho a la idea que plantea este proyecto sin haber sido fuente de inspiración.

Existen dos páginas web que ofrecen la información de los presupuestos: un portal realizado por el gobierno y otra perteneciente a una empresa privada.

Si se centra la atención en la página del gobierno, se puede comprobar cómo ofrece dos botones con la indicación de "Ingresos" o "Gastos" y la selección del año fiscal.

Tras una selección entre ellos, se puede ir eligiendo los conceptos del pie chart para ver cómo se han repartido los presupuestos, añadiendo una tabla dinámica en la parte inferior con el porcentaje correspondiente.

De esta manera, el lector puede comprender cómo se reparten los gastos y navegar entre los conceptos más genéricos hasta los más detallados. El nivel de profundidad elegido es de 3 capas, pero es suficiente para que el lector pueda comprobar la repartición.

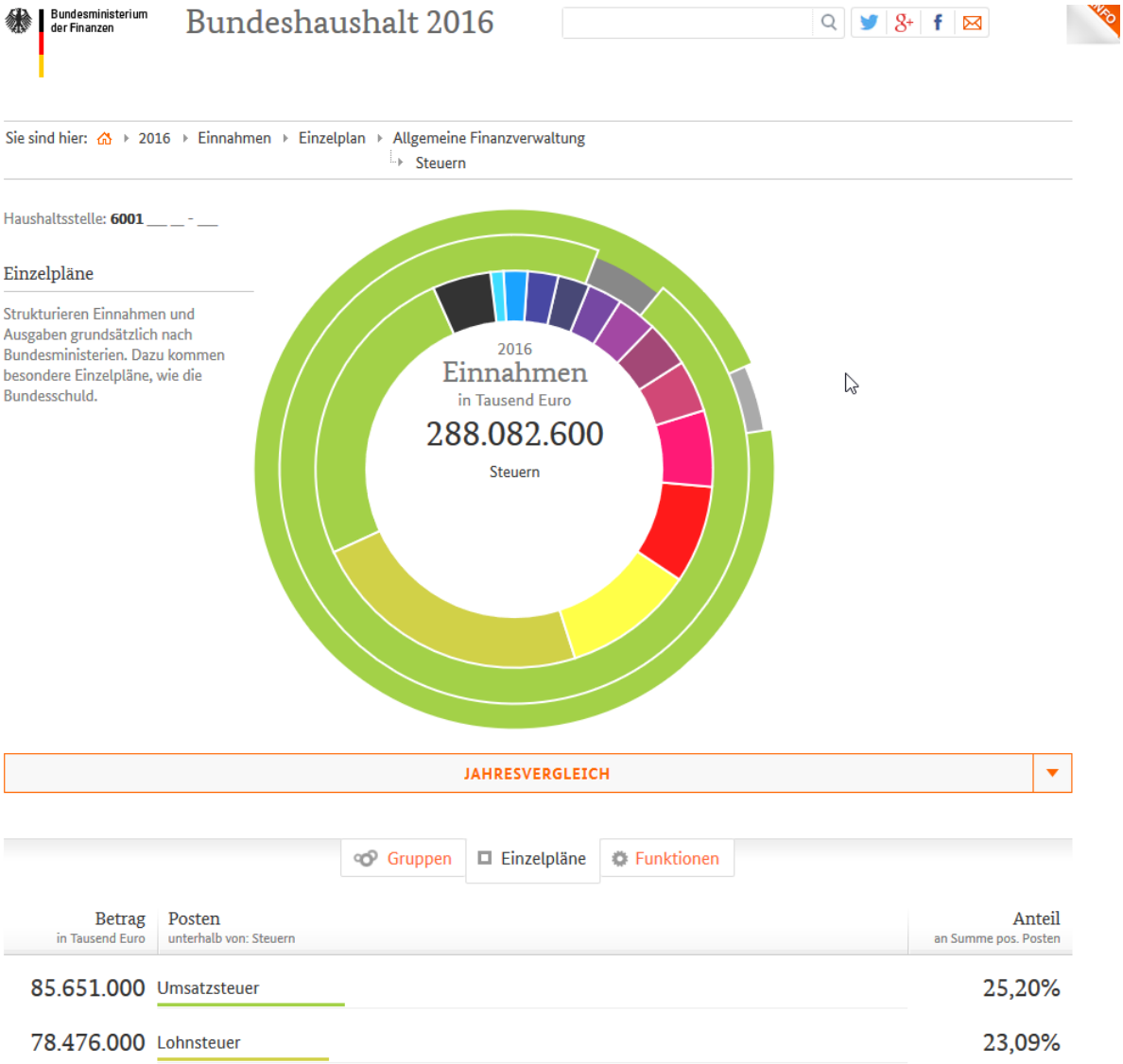


Ilustración 10 - Presupuestos generales del estado alemán

Para el caso de la web desarrollada por la empresa privada, la idea sigue permaneciendo. Lo único que varía es el formato de los gráficos, manteniendo la profundidad de 3 y la tabla dinámica en la parte inferior con el dato en bruto y el porcentaje.

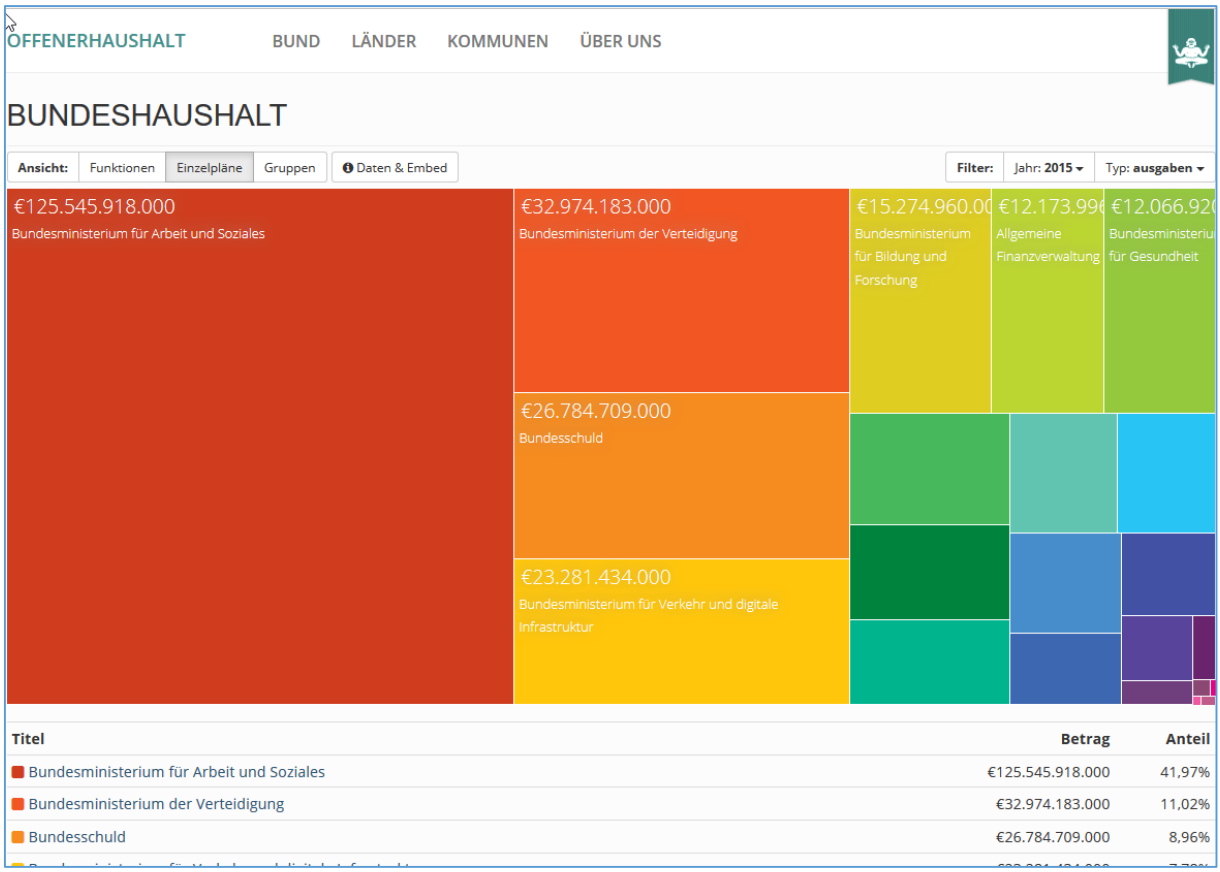


Ilustración 11 - Web privada con los presupuestos generales del estado alemán

Cabe comentar aquí que Alemania es uno de los países europeos que más esfuerzos está realizando por la creación de datos abiertos siendo esta muestra uno de ellos. La información bien organizada y la capacidad de dar poder al lector convierte el caso alemán como un caso de éxito.

Tras las investigaciones que se han llevado a cabo, se puede aseverar que **el tratamiento de los presupuestos es un tema pendiente generalizado en Europa.** Los portales de transparencia que promueve cada gobierno todavía están lejos de cubrir las expectativas que se esperan de ellos.

Como propuesta a esta falta de claridad, se detalla a continuación el proceso para la creación de la herramienta "Democratización de los Presupuestos Generales del Estado".

5. DESCRIPCIÓN DEL PROBLEMA

Si nos apoyamos en la definición de Presupuestos Generales del Estado realizado por la Wiki de Finanzas, encontramos que se denominan así "*al presupuesto público que elabora un Estado a través del organismo competente encargado de la realización de las previsiones y planificaciones, de acuerdo con la normativa que tenga establecido dicho Estado a tal efecto*".

Por tanto, se puede desglosar de la propia definición, que cada órgano competente realiza estimaciones de los ingresos y gastos públicos para el siguiente año contable y que la suma de todas estas previsiones son, en su conjunto, los Presupuestos Generales del Estado. Dicho con otras palabras, son datos contables en formato tabla complicados de seguir la pista.

Como cada organismo dentro del gobierno tiene su fiscalidad independiente. La suma de los presupuestos de todas las entidades supone ya un gran esfuerzo contable y cuando el gobierno publica estos datos, lo hace siguiendo los cánones financieros, con una estructura de datos no enfocada a una persona sin estudios en la materia.

Así que, un ciudadano que busque información respecto a esta materia, puede encontrar mucho material en la página web del Ministerio de Hacienda y Administraciones Públicas pero si el objetivo es entender el flujo económico del país, necesitará días para poderlo analizar. Pongamos un ejemplo, es fácil obtener cuánto fue la partida presupuestaria para el Museo Nacional del Prado, pero es muy complicado saber si esta partida aumentó o disminuyó respecto al año pasado.

Como conclusión a un análisis de la web oficial del estado, se puede decir que el ciudadano tiene a su alcance el material presupuestario necesario pero no aporta las herramientas necesarias para un análisis rápido y obtener conclusiones.

Éste es origen de este TFM: **mostrar la información de los presupuestos generales de una manera visual para que cualquier ciudadano medio pueda convertir los datos en información.**

Para realizarlo, se han empleado las herramientas Big Data y los conocimientos adquiridos durante el Master para dar esta visión más amigable de los presupuestos.

5.1. Alcance de los datos

Dada la inmensa maraña de documentación que existen en los presupuestos y que el tiempo para desarrollar este proyecto es limitado, el alcance de los datos se ha limitado al **Tomo de "Resúmenes de Ingresos y Gastos" de la Serie Roja** publicada cada año en la página oficial del Ministerio de Hacienda y Administraciones Públicas. En este tomo se simplifica los conceptos presupuestarios y permite mostrarlos englobados en carteras ministeriales. Es importante notar que todos los orígenes de datos han sido obtenidos en la web oficial del estado, analizando sólo aquellos que hayan sido aprobados y desechando los proyectos de ley.

Por otro lado, cabe mencionar que para realizar esta labor, se han analizado los 5 agentes económicos en los que se basan tanto los ingresos como los gastos. Estos son:

- Estado
- Organismos Autónomos
- Agencias Estatales
- Otros organismos públicos
- Seguridad Social.

También es importante detallar la profundidad del análisis. Dada la complejidad de los presupuestos, esta primera visión gráfica se centra en los resúmenes de ingresos y gastos con 3 niveles de profundidad:

- 1er nivel: Ingresos o Gastos totalitarios
- 2do nivel: Ministerios
- 3er nivel: Agencias u Organismos

5.2. Estructura de la implementación

El proyecto se basa en el establecimiento de una serie de fases que han servido estructurar los pasos realizados. Cada una de ellas tiene una entidad propia pero necesaria para el desarrollo global del trabajo.

Las fases en las que se ha estructurado el estudio son:

1. Extracción de los orígenes de datos: Consiste en crear un sistema de carpetas con todos los ficheros presupuestarios englobados en los resúmenes de ingresos y gastos.

2. Coherencia y análisis de datos: Analizar cada fichero y comprobar que los datos son coherentes en el ejercicio presupuestario.
3. Transformación de datos: Establecer una estructura que controle las distintas modificaciones que se realizan en los conceptos presupuestarios con el fin de que sean equivalentes con el paso de los años.
4. Almacenamiento de datos: Crear una arquitectura orientada a las bases de datos para almacenar en ella toda la información de los ficheros tratados con la intención de que estos sean trabajados de manera gráfica y perduren en el tiempo.
5. Visualización de la información: Como último paso, crear una visualización gráfica de la información desde distintos puntos de vista para que el lector pueda analizar los datos sin necesidad de manejar ficheros.
6. Encuesta de satisfacción: Para obtener la valoración de la herramienta, se crea una encuesta para que una población aleatoria plasme sus experiencias tras la utilización de la misma.

De una manera gráfica, las fases del proyecto pueden representarse de la siguiente manera:

Ilustración 12 - Fases de la implementación

Para abordar cada fase, se ha escogido una tecnología diferente que se adapte a las necesidades de un proyecto fin de master educativo basado en unos principios de libre distribución y/o de acceso público. Por tanto, tanto las aplicaciones utilizadas para su desarrollo así como los lenguajes escogidos para cada fase se han basado en estos principios.

6.TECNOLOGÍAS

6.1. Motivación

Las tecnologías escogidas para el desarrollo de esta aplicación han sido seleccionadas en base a sus valores técnicos y morales. Como línea técnica, las tecnologías se debían adecuar a un marco educativo donde el uso y la explotación de la herramienta sean didácticos y, en su medida, de código abierto. Esto se traduce a que el proyecto pueda ser objeto de estudio o ampliable por cualquier persona, sin importar su estatus social o económico. Si estas características no se han podido cumplir en 100% de los casos, la decisión de la herramienta escogida se ha basado en su popularidad y su uso extendido entre la población.

6.2. Herramientas utilizadas en cada fase

Las nuevas herramientas relacionadas con el tratamiento masivo de información se caracterizan por su potencial, pero es necesario hacerlas convivir con tecnologías asentadas en el mercado, cuya fiabilidad y arquitectura sean buenos cimientos en la construcción de un nuevo proyecto como este caso. La convivencia entre todas las filosofías aporta una estabilidad y una seguridad que sirve para preparar el proyecto a los devenires del futuro. Por tanto, se han entremezclado tecnologías y herramientas de nueva generación con otras más longevas para aprovechar las ventajas de su unión.

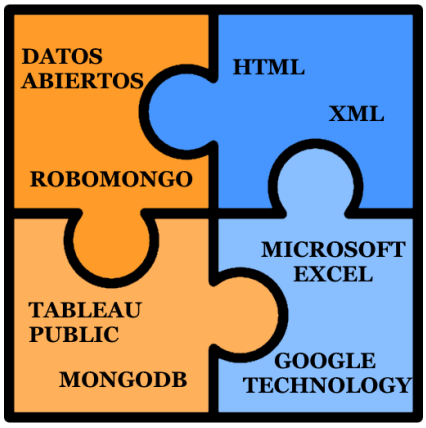


Ilustración 13 - Relación de herramientas y tecnologías

Las fases del desarrollo han marcado los pasos a la hora de trabajar con las distintas herramientas. Al finalizar cada una de ellas, se ha obtenido un producto que ha servido como origen a la siguiente fase. Para ello, se han utilizado las distintas herramientas que se van a ver a continuación. Aunque estas fases se vayan a desarrollar en puntos posteriores, de un primer vistazo las tecnologías utilizadas podrían relacionarse con cada fase de la siguiente manera:

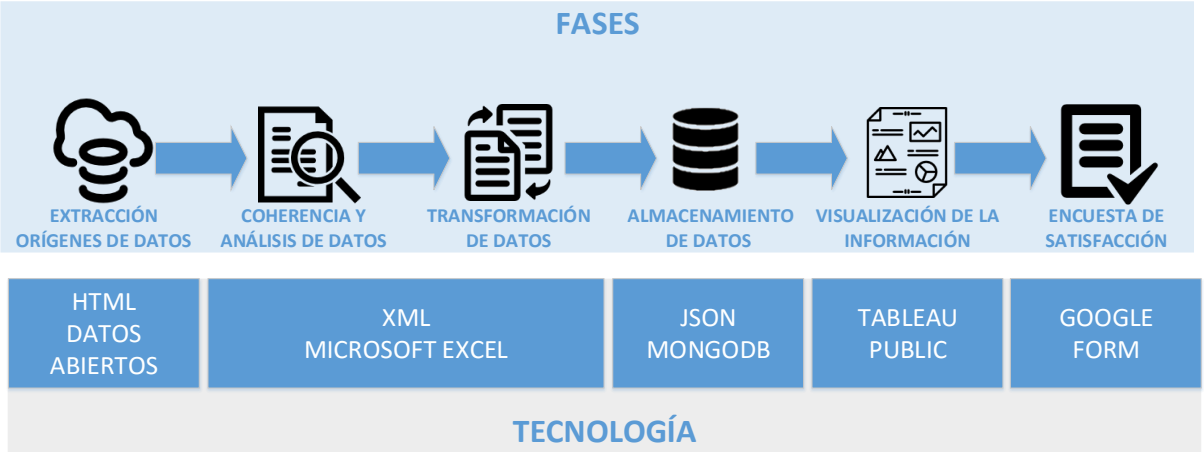


Ilustración 14 - Fases del proyecto vs Tecnología aplicada

En la fase de ***"Extracción de los orígenes de datos"***, los formatos de ficheros que ofrece el Ministerio desde 2007 hasta 2016 son pdf y html. Dada la dificultad de que entraña modificar un formato de visualización como es pdf, se ha optado por utilizar el formato html para los ficheros de origen.

Cabe mencionar que en los presupuestos de 2016 ya aparece como novedad en formato xml. Este formato se adapta mejor a las nuevas tecnologías ya que se trata de un lenguaje basado en marcas y que es legible por las bases de datos. Sin embargo, esta novedad no está implementada para años anteriores, por lo que se ha desechado utilizar este formato en los orígenes. Sin embargo, esta tecnología aparecerá en procesos intermedios de la transformación de datos.

La gran ventaja del formato .html es que permite una visualización cómoda a través del navegador Firefox, ayudando al reconocimiento visual de la información y a la primera criba de orígenes.

En la fase de ***"Coherencia y Análisis de Datos* y *Transformación de datos"***, la tecnología basada en html ya no resulta útil por lo que se decide modificar la extensión de los ficheros a un formato basado en xml. Esto añade una potencia extra al tratarse de un lenguaje basado en etiquetas o marcas que se asemeja en idea y forma a los nuevos formatos de tratamiento de información. Además, al tratarse de un formato entendible desde cualquier tipo de herramienta de información (editores de texto, bases de datos, etc.) su uso permite una gran versatilidad. Como apunte, también es importante destacar que su almacenamiento tipo texto alivia el peso de estos ficheros, convirtiéndose en una gran herramienta de intercambio de información. Otro motivo de peso es que la Administración Pública está haciendo el esfuerzo de crear los nuevos ficheros de datos abiertos en formato xml, por lo que la conversión de html a xml sería innecesaria para sucesivos años.

En la siguiente fase de ***"Transformación de datos"***, se ha seguido la misma filosofía que en la fase anterior. El formato xml se convierte por tanto, en el formato estándar para este proyecto. Sin embargo, no hay que perder de vista que los datos deberán tener un formato .json para alcanzar la base de datos, por lo que se utilizan dos procesos paralelos que transformen los fichechos de xml a json y de html a json. Se verá en la fase de Metodología esta transformación.

En la fase de ***"Almacenamiento de información"*,** se elige como base de datos MongoDB. Esta decisión está tomada en base a su capacidad a la hora de adaptarse a grandes volúmenes de datos y su versatilidad a la hora de introducir todo tipo de información presupuestaria. En este repositorio podrán almacenarte todos los ficheros necesarios para un análisis en profundidad sin tener que realizar modificaciones en las estructuras financieras. Sin embargo, es muy importante prefijar una arquitectura de datos para que la posterior visualización tenga sentido.

Para finalizar, en la última fase correspondiente a la ***"Visualización de la información"***, la herramienta escogida para este proyecto ha sido Tableau Public. La decisión viene tomada tras analizar las diversas plataformas gratuitas de las principales herramientas que hay en el mercado. Esta herramienta ofrece un amplio catálogo de visualizaciones y la capacidad de montar un cuadro de mando profesional con su versión gratuita. También ha sido un punto a su favor la relación de esta aplicación con las nuevas herramientas de explotación masiva de información.

A continuación, vamos a detallar cada herramienta para verla en profundidad.

6.3. Detalle de las herramientas

En este apartado se desgrana cada tecnología y cada herramienta utilizada para entender mejor el posterior desarrollo del proyecto.

1. HTML

HTML es la versión abreviada de "***H**yper**T**ext **M**arkup **L**anguage*" y se refiere a un lenguaje de marcas de hipertexto utilizado en la elaboración de recursos web (páginas web, portales, aplicaciones, etc.). Ampliamente conocido, su fundador es Tim Berners-Lee y fue creado en 1991. Este lenguaje de desarrollo web universal y gratuito utiliza marcas entre los paréntesis angulares <> para referenciar cualquier objeto contenido en su interior. Basando este contenido en una relación "clave – valor", es fácilmente interpretable por los navegadores.

```
<title>
</title>
</head>
<body class = "S2008602935_2143336719_45ESTILO1">
    <h1 class = "S2008602935_2143336719_45ESTILO0">Grupo de informes</h1>
    <div class = "S2008602935_2143336719_45ESTILO1">
    <h2 class = "S2008602935_2143336719_45ESTILO0 S2008602935_2143336719_45ESTILO57">Informe lineal</h2>
    <div class = "S2008602935_2143336719_45ESTILO58"><div>
    <table class = "S2008602935_2143336719_45ESTILO2 S2008602935_2143336719_45ESTILO20">
    <tbody>
    <tr>
    <td class = "S2008602935_2143336719_45ESTILO3 S2008602935_2143336719_45ESTILO4">
    <table class = "S2008602935_2143336719_45ESTILO2 S2008602935_2143336719_45ESTILO20">
    <tbody>
    <tr>
```

Ilustración 15 - Fragmento de código html

En la página web del Ministerio de Hacienda y Administraciones Públicas, los ficheros disponibles con los datos presupuestarios sólo están en formato html o pdf. Por tanto, la elección de los ficheros con extensión .html es prácticamente obligada.

2. Datos abiertos

Los datos abiertos u "open data" forman parte de una filosofía basada en el software libre y el conocimiento público de la información. Aunque ambos términos difieren en su definición, la filosofía que hay por debajo es común a ambos. Esta filosofía reza que la información debe circular libremente para su uso sea democrático y no dependa del estatus social del usuario. Esta concepción del universo digital fue introducida por Richard Stallman en 1985, creador y máximo gurú del software libre, padre de la licencia GNU y fundador de la organización sin ánimo de lucro Free Software Foundation.

Gracias a este movimiento, ha habido en los últimos años un gran apoyo a los portales de transparencia, reivindicando que la información del gobierno sea pública y esté al alcance de cualquier persona. Por tanto, aunque los presupuestos generales del estado llevan siendo públicos muchos años (desde el año 2000 en adelante), los datos hasta entonces se publicaban como un libro electrónico sin que el lector pudiera explotar esos datos de manera directa. Sin embargo, desde el año 2007, la administración pública dispuso los datos en un formato tabla, adecuando su contenido a su posible explotación. Aun así, el esfuerzo se quedaba corto ya que su explotación directa no es posible si no se utilizan herramientas creadas exclusivamente para este fin. A partir de 2016 ya se empiezan a ver los frutos de la presión en esta vertiente y comienzan a publicarse los presupuestos en un formato directamente explotable.

Por tanto, para conseguir analizar los presupuestos generales del estado, era necesario conseguir los datos en un formato adaptado a las siguientes fases, por lo que se optó en transformar todos los ficheros a xml.

3. XML

XML son las siglas correspondientes a "e***X****tensible* ***M****arkup* ***L****anguage*". Este lenguaje también está basado en marcas que identifican cada objeto pero su estructura es próxima a los sistemas de almacenamiento, siendo un formato adecuado para almacenar estructuras de información.

Se optó por este formato debido a que los presupuestos de 2016 ya incorporaban como novedad este formato en los ficheros. Además, el formato XML está universalmente extendido y es

reconocible por todas las herramientas orientadas a almacenar información. En este caso, se ha tratado con los datos en un formato Open XML.

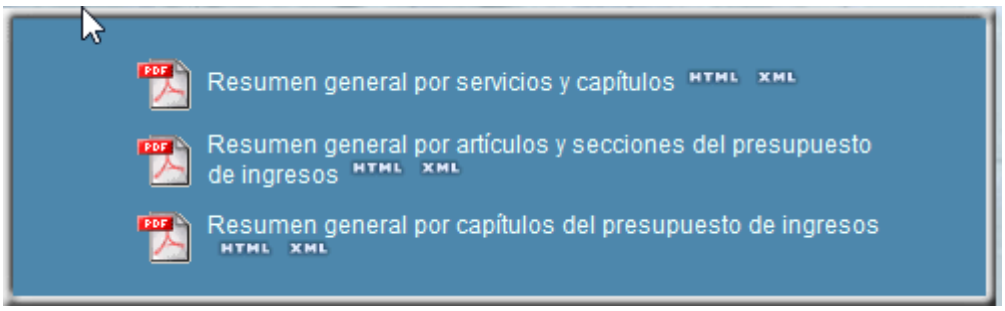


Ilustración 16 - Extracto de los Presupuestos Generales del Estado 2016

Se han tenido que tratar un gran número de datos en la fase de análisis y este tratamiento no hubiera sido posible sin esta tecnología. La visualización de XML en herramientas de uso extendido permite al usuario corregir y normalizar los datos desde un contexto de información. Vemos a continuación esta contextualización.

4. Microsoft Excel

Se trata de una de las herramientas más populares en el mundo de la tecnología por su capacidad de cálculo y su versatilidad a la hora de trabajar con ella. Esta aplicación de Microsoft surge en 1985 y ha ido evolucionando hasta nuestros días con complementos que la sitúan como la aplicación más camaleónica del momento. Esta hoja de cálculo es utilizada tanto por principiantes como expertos informáticos.

En el desarrollo de este proyecto, su uso no sólo se ha limitado al análisis de los datos, ya que sus complementos permiten la transformación de unos formatos a otros.

En la fase de análisis, el contexto de los datos ha sido fundamental a la hora de normalizar los mismos ya que la información vertida estaba poco normalizada, menos de lo esperado. Por ejemplo, existen conceptos que cambian de identificador y de padre dependiendo del año, conceptos que cambian su descripción textual o que se añaden a otros ya existentes.

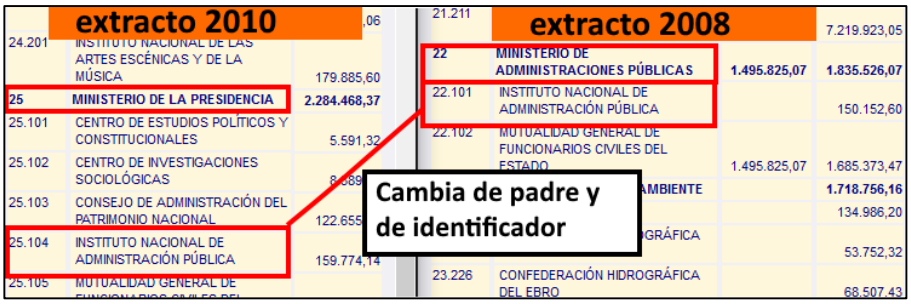


Ilustración 17 - Modificaciones contextuales

Para realizar esta labor, es necesario visualizar la información en formato tabla para adaptar cada concepto al organismo más adecuado y así poderlo estructurar de manera comparativa. Es decir, que a lo largo de los años, se pueda comparar cada ministerio o cada organismo consigo mismo. Al ser una fase no automatizable en la que son necesarias grandes dosis de sentido común y es la primera vez que se realiza esta normalización, ha sido necesaria una herramienta como Excel,

Hay que notar que una vez tratados todo el histórico de datos de los presupuestos generales del estado, esta herramienta es sustituida por la propia herramienta de explotación de base de datos.

5. MongoDB

Se trata de una herramienta de almacenamiento de información muy popular entre los nuevos desarrollos de Big Data. Creada en 2009 por, se caracteriza por ser una base de datos NoSQL orientada a documentos en vez de tablas. Esta estructura es adaptable ya que la información que puede almacenar no tiene por qué tener estructura prefijada como pasa en las bases de datos relacionales. Por otro lado, también destaca MongoDB por utilizar ficheros de información con los datos menos pesados que las bases de datos relacionales ya que no tienen que almacenar la metadata de sus estructuras.

Otra de las grandes ventajas sobre los sistemas de almacenamiento clásicos es que su licencia es GNU, por lo que su distribución y uso es gratuito. Esto permite aventajar en gran medida a sus competidoras clásicas, que tienen licencias de pago en sus productos.

6. Robomongo

Esta potente herramienta es un intérprete de MongoDB. Creada de manera paralela a MongoDB, no se desarrolla su primera versión completa hasta 2012. Siguiento la línea de MongoDB, también se trata de una herramienta de código abierto sostenida económicamente mediante campañas de donación.

Esta herramienta aporta un gran valor añadido a la base de datos, ya que traduce su contenido a una vista clásica, creando un entorno más amigable al usuario. También destacan las distintas visualizaciones de los datos que ofrece la aplicación, su gestión de usuarios y la gestión de documentos.

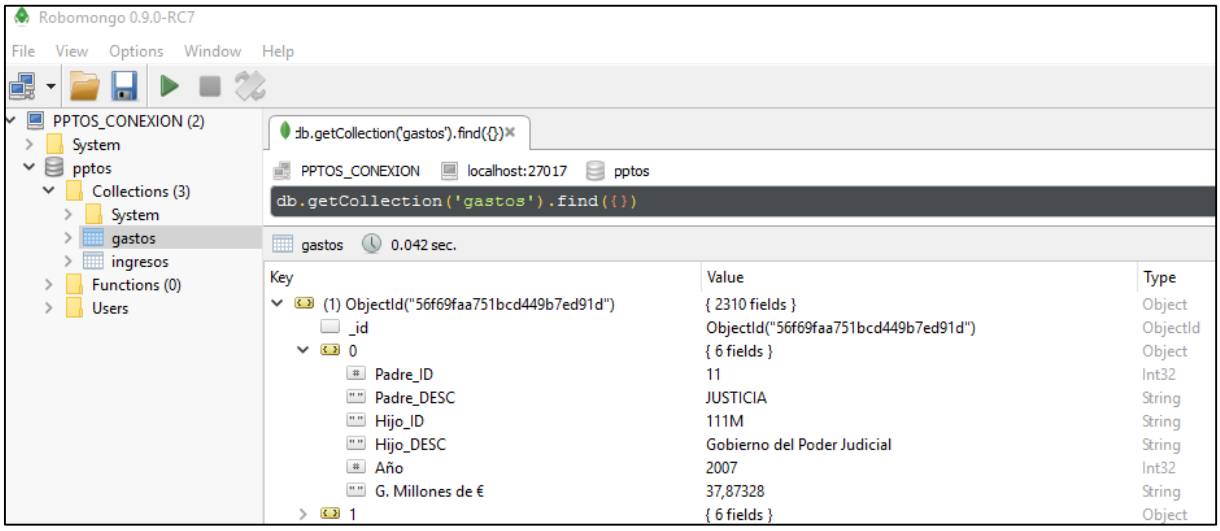


Ilustración 18 -Interfaz de Robomongo

7. Tableau Public

Esta aplicación creada en 2003 está enfocada a la visualización de los datos y, aunque no se trata de un programa basado en software libre, sí es cierto que la capacidad que permite al usuario en su versión de prueba es amplia. Si bien es cierto que los gráficos que pueden escogerse están limitados, esta herramienta ofrece una extensa cartera de personalizaciones.

Además, como ventaja indiscutible sobre su competencia, ofrece la publicación de los trabajos en su web así como su portabilidad o incrustación en otras páginas.

El motivo por el cual ha sido seleccionada esta herramienta es ampliar conocimiento en el campo de la visualización.

8. Google Technology

Si existe una herramienta puntera en el tratamiento de información esa es Google. Nació en 1998 en Estados Unidos con su buscador minimalista, rompiendo todas las reglas de diseño existentes hasta ese momento. A partir de su buscador, la empresa fue desarrollando diversas tecnologías gratuitas para el usuario, entre las que destacan por su interés en relación con este proyecto las siguientes:

- Gmail: cuenta de correo electrónico.

- Drive: almacenamiento en la nube.

- Chrome: navegador web.

- Google: Buscador de información en Internet.

- Google Form: formularios atractivos para evaluaciones.

8. METODOLOGÍA

Este apartado desgrana cómo se ha realizado la democratización de los presupuestos. Como se ha comentado anteriormente, el desarrollo se ha dividido en fases y en cada fase se han realizado una serie de tareas, todas ellas de vital importancia para las siguientes. Por tanto, cada fase tiene una entidad propia con tareas delimitadas e independientes entre ellas, se puede ver en el siguiente diagrama de proceso:

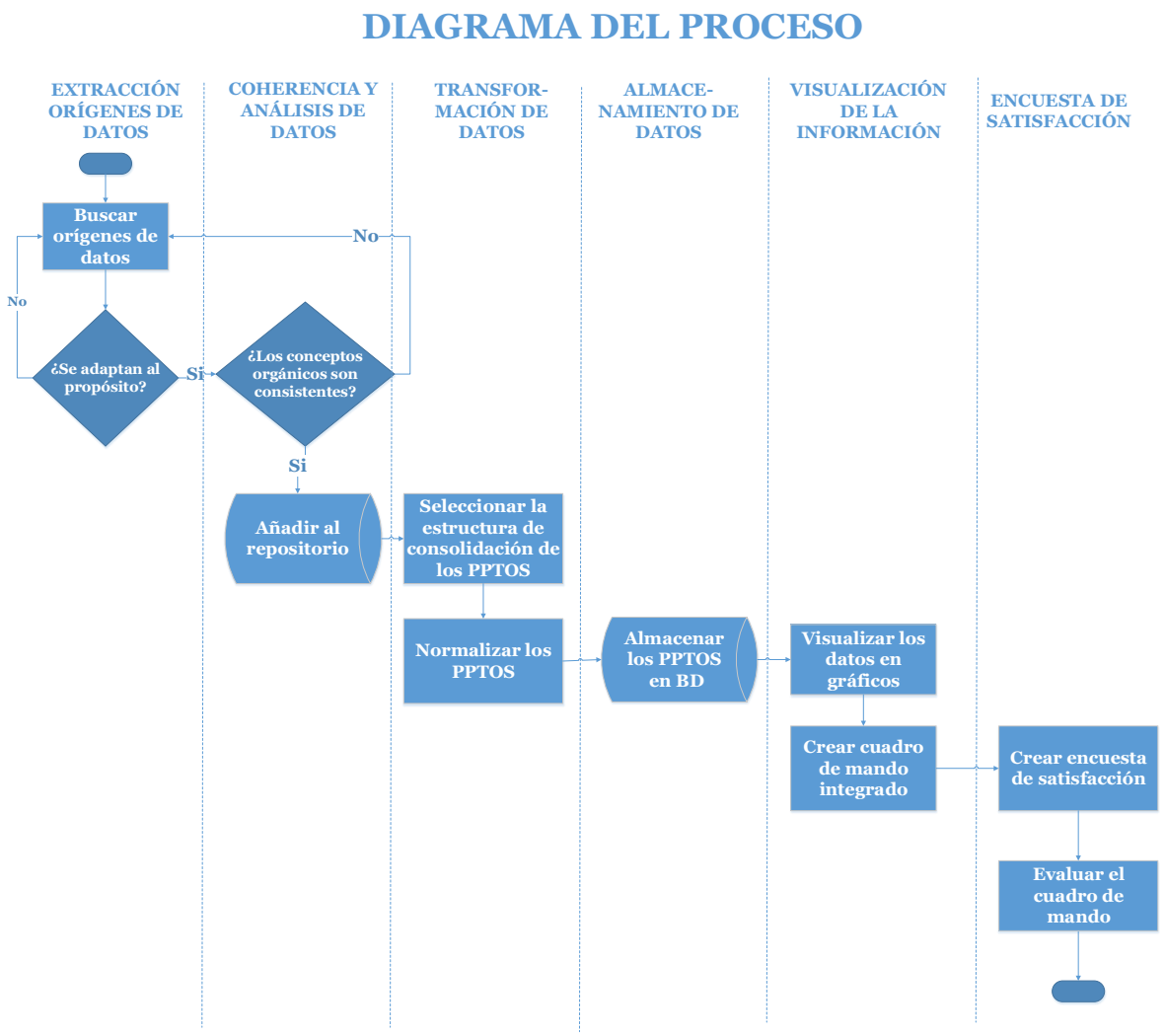


Ilustración 19 - Diagrama de Proceso

A continuación, vemos cada fase y tarea con detalle.

1. Extracción orígenes de datos

El punto de inicio para "La Democratización de los Presupuestos Generales del Estado" es conseguir los presupuestos oficiales en un formato digital y con un mínimo detalle para el análisis. A estos ficheros los llamamos orígenes de datos y son los cimientos del proyecto. Todo el proyecto gira alrededor de esta información y tan importante es que sea oficial como que sea digital para evitar errores de transcripción o de veracidad.

1.1. Buscar orígenes de datos

Cuando se habla de los presupuestos del estado, el mejor sitio donde obtener información es la página web oficial. En ella, se encuentra información no solo relativa a los presupuestos generales, sino también información sobre los presupuestos de la Unión Europea, las Comunidades Autónomas, Entidades Locales, Normativa Presupuestaria, Informes Económico Financiero y algunas estadísticas elaboradas por la propia administración. También se pueden encontrar tablas de los índices que macroeconómicos. Sin embargo, toda esta información no es un uniforme y su estudio es una tarea complicada. Por ejemplo, los datos publicados con el histórico de los índices macroeconómicos, están disponibles desde el año 1954 pero terminan en 2010. Incluso los presupuestos generales que van desde 1996 hasta 2016, cada año está estructurado de una manera diferente.

Tras una primera incursión en los datos presupuestarios, se identifican tres macro estructuras:

- **Desde 1996 hasta el 2001**: los presupuestos están distribuidos en libros que puedes acceder a ellos mediante un documento pdf. Dependiendo el año, se publica uno hasta cuatro libros. Estos libros contienen una extensa narrativa donde justifican el importe de cada concepto. Por tanto, el número de páginas que contiene cada libro suele superar las 250 páginas y cada ejercicio suele componerse de 3 libros, por tanto el número de páginas que hay que leer superan las 750 para cada año.
- **Desde 2002 hasta 2006**: los presupuestos se mantienen en formato libro y además, añaden un vínculo a un portal que le llamaremos Maestro donde se pueden visualizar el articulado de la ley, el estado de los ingresos y los gastos, informes económicos y financieros y las memorias. Todos estos documentos están en formato pdf y su título es identificativo. Este portal supone un buen avance, ya que al menos, evita al lector tener que leer todos los libros pero, sin embargo, estos documentos no tienen una estructura adaptada a poder realizar análisis con ellas, por tanto, son meramente informativas.
- **Desde 2007 hasta 2016**: durante estos años, se añade una novedad en el portal Maestro. Los distintos resúmenes no sólo se visualizan en formato pdf, sino que ya

están disponibles como tablas en un formato html. Este paso añade valor a los presupuestos ya que estas tablas sí son tratables por otras fuentes de explotación de datos. También llama la atención que la estructura es diferente siendo ésta más resumida y, por tanto, más fácil de seguir. Sin embargo, se pierde granularidad en los niveles de detalle.

Desde el 2016, el gobierno pone a disposición en el portal Maestro los ingresos y los gastos con formato xml. Por tanto, los formatos disponibles a partir de este año son: xml, html y pdf.

1.2. ¿Se adaptan al propósito?

Como se comentaba en el apartado anterior, la información disponible en la web del ministerio es amplia y realizar una selección de los orígenes de datos es una tarea complicada y costosa en tiempo. Para unificar los criterios, la primera decisión que se tomó fue utilizar el portal Maestro para los ejercicios desde 2007 hasta 2016. Esta decisión se basa en que la estructura para ese periodo de años es idéntica, facilitando la normalización de los datos finales.

Ilustración 20 - Aspecto del Portal Maestro

Como se ve en la ilustración, este portal está dividido en cuatros series que, cada una, tienen una información diferente sobre los presupuestos. Si nos ceñimos a la descripción que se aporta en la web Finanzas para mortales vemos que:

- La **serie roja** contiene entre otros, los textos aprobados por las Cortes Generales y que conforman los Presupuestos Generales del Estado.
- La **serie verde** contiene la forma de adecuar la ejecución de los presupuestos. Así pues, los resúmenes de la serie roja se pueden encontrar con un mayor desglose en esta serie. Provocar
- La **serie amarilla** contiene la explicación de los presupuestos en formato narrativo. También contiene el presupuesto consolidado de ingresos y gastos dividido por distintas clasificaciones (capítulos, artículos y sub secciones).
- La **serie gris** contiene información adicional sobre las particularidades que han tenido que realizarse.

Cada serie, a su vez, contiene una estructura de carpetas y subcarpetas interconectadas. Para que el lector pueda hacerse una idea del volumen de información que se maneja, **el portal Maestro contiene cada año unos de 10000 archivos y alrededor de 1500 carpetas.** Por tanto, para realizar este análisis se debe emplear una gran cantidad de horas buscando patrones de información que sean comparables entre sí.

Se solicitó ayuda al Ministerio a través del correo electrónico que tiene publicado para cuestiones relacionadas con los presupuestos. La dirección es: InternetDGPresupuestos@sepg.minhap.es. Hasta en tres ocasiones (Ver Anexos) se intentó la resolución de dudas, pero no hubo respuesta por parte del Ministerio. Así que, tras las consultas fallidas, se requirió ayuda de personal financiero cualificado para resolver las mismas.

Tras este análisis contable, se llega a la conclusión que la carpeta de la serie roja "**Resúmenes de Ingresos y Gastos**" es la que mejor se adecúa a las necesidades del proyecto.

2. Coherencia y Análisis de datos

Esta carpeta alberga información desde distintos puntos de vista. Para poder realizar comparativas entre ellos, se toma la decisión de utilizar sólo aquellos que desglosen los conceptos por capítulos ya que cada explicación corresponde a un ente con el que está familiarizado el ciudadano.

INGRESOS

Para el caso de los ingresos, los ficheros que se han tratado han sido:

INGRESOS
ESTADO
Resumen general por servicios y capítulos
ORGANISMOS AUTÓNOMOS
Resumen general por organismos y capítulos del presupuesto de ingresos
AGENCIAS ESTATALES
Resumen general por organismos y capítulos del presupuesto de ingresos
OTROS ORGANISMOS PÚBLICOS
Resumen general por organismos y capítulos del presupuesto de ingresos
SEGURIDAD SOCIAL
Resumen general por servicios y capítulos del presupuesto de ingresos

Tabla 1 –Ficheros de ingresos tratados en la aplicación

Y el formato de los ficheros de origen tiene esta estructura, sólo variando el número de capítulos que aplican, es de la siguiente manera:

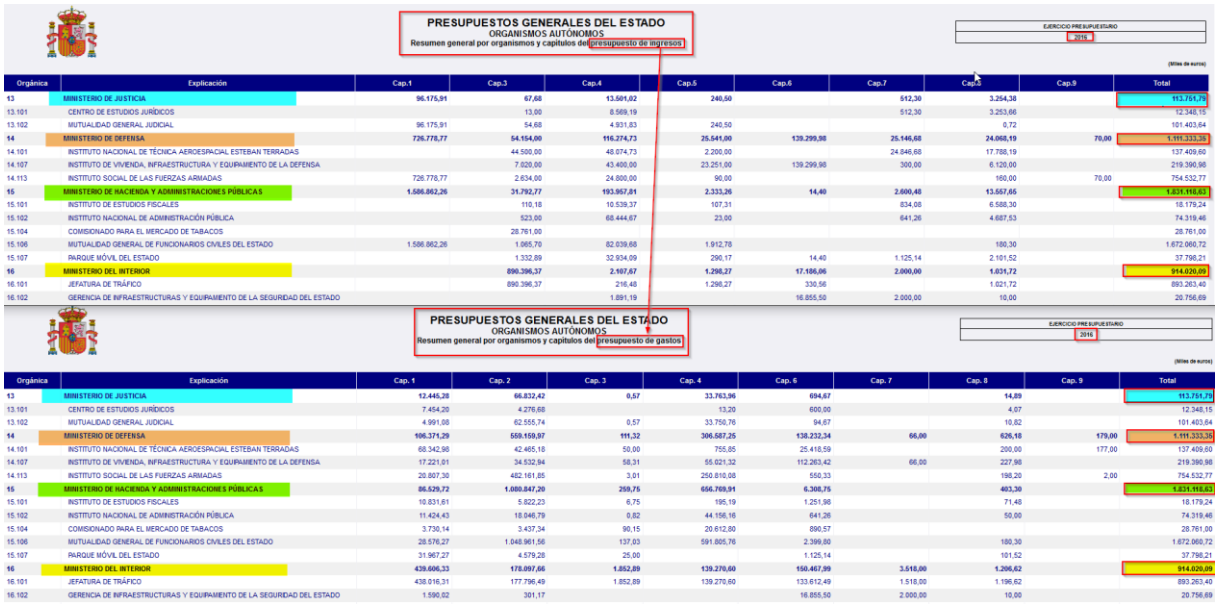
Campo	Significado	Tipos de campo	Unidad de medida (si procede)	Válido para el análisis	Debe ser normalizable
Orgánica	Código de la explicación	Número decimal	-	si	si
Explicación	Descriptor del ente que genera el ingreso	Texto	-	si	si
Cap.1	Impuestos directos	Número decimal	Miles de euros	no	no aplica
Cap.2	Impuestos indirectos	Número decimal	Miles de euros	no	no aplica
Cap.3	Tasas, precios públicos y otros ingresos	Número decimal	Miles de euros	no	no aplica
Cap.4	Transferencias correintes	Número decimal	Miles de euros	no	no aplica
Cap.5	Ingresos patrimoniales	Número decimal	Miles de euros	no	no aplica

Cap.6	Enajenación de inversiones reales	Número decimal	Miles de euros	no	no aplica
Cap.7	Transferencias de capital	Número decimal	Miles de euros	no	no aplica
Cap.8	Activos financieros	Número decimal	Miles de euros	no	no aplica
Cap.9	Pasivos financieros	Número decimal	Miles de euros	no	no aplica
Total	Cuantía total del ente	Número decimal	Miles de euros	si	si

Tabla 2 –Formato de los ficheros de origen para los ingresos

GASTOS

En un principio, también se estaba tratando los resúmenes de gastos que hay asociados en el mismo apartado, pero se comprobó que la calidad del dato ya era muy defectuosa. Si se compara, por ejemplo, el resumen de Organismos Autónomos para los ingresos y los gastos, se evidencia que el presupuesto totalitario coincide al decimal. Esto quiere decir que se presupone que cada organismo (ministerio, instituto, gerencia, jefatura, etc.) va a ingresar lo mismo que va a gastar. El sentido común indica que esta afirmación no es sostenible, por lo que se investigó más y se encontraron los presupuestos consolidados de gastos, cuyo resultado se aproximaba más tanto en coherencia como en valor.



PRESUPUESTOS GENERALES DEL ESTADO
ORGANISMOS AUTÓNOMOS
Resumen general por organismos y capítulos del presupuesto de ingresos

EJERCICIO PRESUPUESTARIO 2016

(Miles de euros)

Orgánica	Explicación	Cap.1	Cap.3	Cap.4	Cap.5	Cap.6	Cap.7	Cap.8	Cap.9	Total
13	MINISTERIO DE JUSTICIA	96.175,91	67,68	13.501,02	240,50		512,30	3.254,36		113.751,79
13.101	CENTRO DE ESTUDIOS JURÍDICOS		13,00	8.569,19			512,30	3.253,66		12.348,15
13.102	MUTUALIDAD GENERAL JUDICIAL	96.175,91	54,68	4.931,83	240,50			0,72		101.403,64
14	MINISTERIO DE DEFENSA	726.778,77	54.154,00	116.274,73	25.541,00	139.299,98	25.146,68	24.068,19	70,00	1.111.333,35
14.101	INSTITUTO NACIONAL DE TÉCNICA AEROESPACIAL ESTEBAN TERRADAS		44.500,00	48.074,73	2.200,00		24.846,68	17.788,19		137.409,60
14.107	INSTITUTO DE VIVIENDA, INFRAESTRUCTURA Y EQUIPAMIENTO DE LA DEFENSA		7.820,00	43.400,00	23.251,00	139.299,98	300,00	6.120,00		219.390,98
14.113	INSTITUTO SOCIAL DE LAS FUERZAS ARMADAS	726.778,77	2.834,00	24.800,00	90,00			160,00	70,00	754.532,77
15	MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS	1.586.862,26	31.792,77	193.957,81	2.333,26	14,40	2.600,48	13.557,65		1.831.116,63
15.101	INSTITUTO DE ESTUDIOS FISCALES		110,18	10.539,37	107,31		834,06	6.588,30		18.179,24
15.102	INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA		523,00	68.444,67	23,00		641,26	4.687,53		74.319,46
15.104	COMISIONADO PARA EL MERCADO DE TABACOS		28.761,00							28.761,00
15.106	MUTUALIDAD GENERAL DE FUNCIONARIOS CIVILES DEL ESTADO	1.586.862,26	1.065,70	82.039,68	1.912,78			180,30		1.672.060,72
15.107	PARQUE MÓVIL DEL ESTADO		1.332,89	32.934,09	290,17	14,40	1.125,14	2.101,52		37.798,21
16	MINISTERIO DEL INTERIOR		890.396,37	2.107,67	1.298,27	17.186,06	2.000,00	1.031,72		914.020,09
16.101	JEFATURA DE TRÁFICO		890.396,37	216,48	1.298,27	330,56		1.021,72		893.263,40
16.102	GERENCIA DE INFRAESTRUCTURAS Y EQUIPAMIENTO DE LA SEGURIDAD DEL ESTADO			1.891,19		16.855,50	2.000,00	10,00		20.756,69

PRESUPUESTOS GENERALES DEL ESTADO
ORGANISMOS AUTÓNOMOS
Resumen general por organismos y capítulos del presupuesto de gastos

EJERCICIO PRESUPUESTARIO 2016

(Miles de euros)

Orgánica	Explicación	Cap. 1	Cap. 2	Cap. 3	Cap. 4	Cap. 6	Cap. 7	Cap. 8	Cap. 9	Total
13	MINISTERIO DE JUSTICIA	12.445,28	66.832,42	0,57	33.763,96	694,67		14,89		113.751,79
13.101	CENTRO DE ESTUDIOS JURÍDICOS	7.454,20	4.276,68		13,20	600,00		4,07		12.348,15
13.102	MUTUALIDAD GENERAL JUDICIAL	4.991,08	62.555,74	0,57	33.750,76	94,67		10,82		101.403,64
14	MINISTERIO DE DEFENSA	106.371,29	559.159,97	111,32	306.587,25	138.232,34	66,00	626,18	179,00	1.111.333,35
14.101	INSTITUTO NACIONAL DE TÉCNICA AEROESPACIAL ESTEBAN TERRADAS	68.342,98	42.465,18	50,00	755,85	25.418,59		200,00	177,00	137.409,60
14.107	INSTITUTO DE VIVIENDA, INFRAESTRUCTURA Y EQUIPAMIENTO DE LA DEFENSA	17.221,01	34.532,94	58,31	55.021,32	112.263,42	66,00	227,98		219.390,98
14.113	INSTITUTO SOCIAL DE LAS FUERZAS ARMADAS	20.807,30	482.161,85	3,01	250.810,08	550,33		198,20	2,00	754.532,77
15	MINISTERIO DE HACIENDA Y ADMINISTRACIONES PÚBLICAS	86.529,72	1.080.847,20	259,75	656.769,91	6.308,75		403,30		1.831.116,63
15.101	INSTITUTO DE ESTUDIOS FISCALES	10.831,61	5.822,23	6,75	195,19	1.251,98		71,48		18.179,24
15.102	INSTITUTO NACIONAL DE ADMINISTRACIÓN PÚBLICA	11.424,43	18.046,79	0,82	44.156,16	641,26		50,00		74.319,46
15.104	COMISIONADO PARA EL MERCADO DE TABACOS	3.730,14	3.437,34	90,15	20.612,80	890,57				28.761,00
15.106	MUTUALIDAD GENERAL DE FUNCIONARIOS CIVILES DEL ESTADO	28.576,27	1.048.961,56	137,03	591.805,76	2.399,80		180,30		1.672.060,72
15.107	PARQUE MÓVIL DEL ESTADO	31.967,27	4.579,28	25,00		1.125,14		101,52		37.798,21
16	MINISTERIO DEL INTERIOR	439.606,33	178.097,66	1.852,89	139.270,60	150.467,99	3.518,00	1.206,62		914.020,09
16.101	JEFATURA DE TRÁFICO	438.016,31	177.796,49	1.852,89	139.270,60	133.612,49	1.518,00	1.196,62		893.263,40
16.102	GERENCIA DE INFRAESTRUCTURAS Y EQUIPAMIENTO DE LA SEGURIDAD DEL ESTADO	1.590,02	301,17			16.855,50	2.000,00	10,00		20.756,69

Ilustración 21 - Comparativa de los resúmenes de ingresos y gastos

Por lo tanto, el fichero que se ha utilizado relativo a los gastos tiene esta estructura:

Campo	Significado	Tipos de campo	Unidad de medida (si procede)	Válido para el análisis	Debe ser normalizable
Programa	Código + Descriptor del ente que genera el gasto	Texto	-	Si	Si
2000	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2001	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2002	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2003	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2004	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2005	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2006	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si (cuando aplique)
2007	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2008	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2009	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2010	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2011	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2012	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2013	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2014	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2015	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si
2016	Valor del gasto para ese ente	Número decimal	Millones de euros	Si	Si

Tabla 3 –Formato de los ficheros de origen para los gastos

2.1. ¿Los conceptos orgánicos son consistentes?

Aunque la fuente de información es un organismo oficial, puede suceder que haya erratas o transcripciones erróneas en los datos vertidos a la red. Es por esto que antes de tratar la información es conveniente contrastarla y tomar decisiones respecto a los resultados obtenidos.

Cuando se trató el programa de gastos, se descubrieron ciertas inexactitudes que provocaban distintas totalidades entre el gasto ministerial y la suma de los entes que lo componen. Así fue, por ejemplo, que la partida presupuestaria del Ministerio de Industria y Energía para el 2008 no coincidía su desglose con su totalización.

PRESUPUESTOS GENERALES DEL ESTADO
GASTOS
ESTADO, OO.AA., AGENCIAS, O.PÚBLICOS Y S.SOCIAL
Programas de gasto (capítulos 1 a 8)
(Millones de euros)

Programas	2007	2008**	2009	2010	2011	2012	2013
41 AGRICULTURA, PESCA Y ALIMENTACIÓN	8.536	8.849	8.861	8.959	8.575	8.455	
421M: Dirección y servicios generales de Industria y Energía	123	110	112	102	93	71	
421N : Regulación y protección de la propiedad industrial	52	60	67	67	60	54	
421O : Calidad y seguridad industrial	8	8					
422A : Incentivos regionales a la localización industrial	281	297					
422B : Desarrollo industrial	45	85					
422M : Reconversión y reindustrialización	509	629					
423M : Desarrollo alternativo de las comarcas mineras del carbón	242	250	280	260	145	41	
423N : Explotación minera	733	741	750	720	658	523	
424M : Seguridad nuclear y protección radiológica	44	45	51	51	48	47	
425A : Normativa y desarollo energético	80	88	80	328	322	30	
42 INDUSTRIA Y ENERGÍA	2.117	2.324	2.875	3.229	2.801	1.897	

El sumatorio es 2.313 y no 2.324

Totalizado por políticas de gas | Programa gasto | Tableau C ...

PROMEDIO: 231 RECUENTO: 10 SUMA: 2.313 100 %

Ilustración 22 - Incoherencias de los presupuestos referentes a los gastos

Así pues, se decidió tomar como válido el concepto de más bajo nivel y eliminar las totalizaciones dadas, para poder continuar con el desarrollo del Proyecto.

2.2. Añadir al repositorio

Tras analizar y realizar unas comprobaciones de consistencia de cada partida presupuestaria, los datos son almacenados en un repositorio para su posterior normalización. Los ficheros no pueden utilizarse tal cual están, ya que la información tiene que estar normalizada. En la partida de ingresos se hablaba de miles de euros y, sin embargo, en la de gastos los conceptos están indicados en millones de euros. Tampoco el formato de cada fichero es común ya que ni los campos ni los conceptos están tratados para una posterior visualización global.

3. Transformación de datos

3.1. Seleccionar la estructura de consolidación de los presupuestos

Como se hablaba en el apartado anterior, la calidad del dato es indispensable para realizar una visualización de la información. Dado que los ficheros no plantean la misma estructura, ha sido indispensable crear una arquitectura común tanto para ingresos como para gastos. Esta estructura viene determinada de la siguiente manera:

Campo	Significado	Tipos de campo
Padre_ID	Código del ministerio o ente que abarca la partida presupuestaria	Entero
Padre_DESC	Descripción del ministerio o ente que abarca la partida presupuestaria	Texto
Hijo_ID	Código del organismo al que aplica el importe del ingreso o gasto	Texto
Hijo_DESC	Descripción del organismo al que aplica el importe del ingreso o gasto	Texto
Año	Año del año presupuestario que aplica	Año(Fecha)
Tipo	Ingreso o gasto presupuestario	Enumerado
Millones de €	Cifra del importe del ingreso o gasto	Decimal

Tabla 4 –Arquitectura de los presupuestos para su almacenamiento

Esta organización permite que la granularidad del estudio se pueda realizar a través de tres atributos:

- Tipo: Donde se la partida entre ingresos o gastos.
- Padre_Desc: Donde una vez seleccionado el tipo, se puede visualizar la cuantía que se lleva cada ministerio.
- Hijo_Desc: Una vez seleccionado el padre, se puede profundizar al siguiente nivel que corresponde a los entes que conforman cada ministerio.

3.2. Normalizar los presupuestos

Partiendo de los ficheros de ingresos y gastos, se debe realizar una normalización para llegar a la estructura común. Esta **normalización** consiste en la curación de los datos, es decir, realizar una **limpieza gramatical y tipológica de los datos así como una estructuración de los entes**. En esta fase, interviene el sentido común del analista ya que la primera vez que se realiza esta curación debe hacerse manualmente. Una vez establecidas las bases, podría realizarse una automatización de esta tarea.

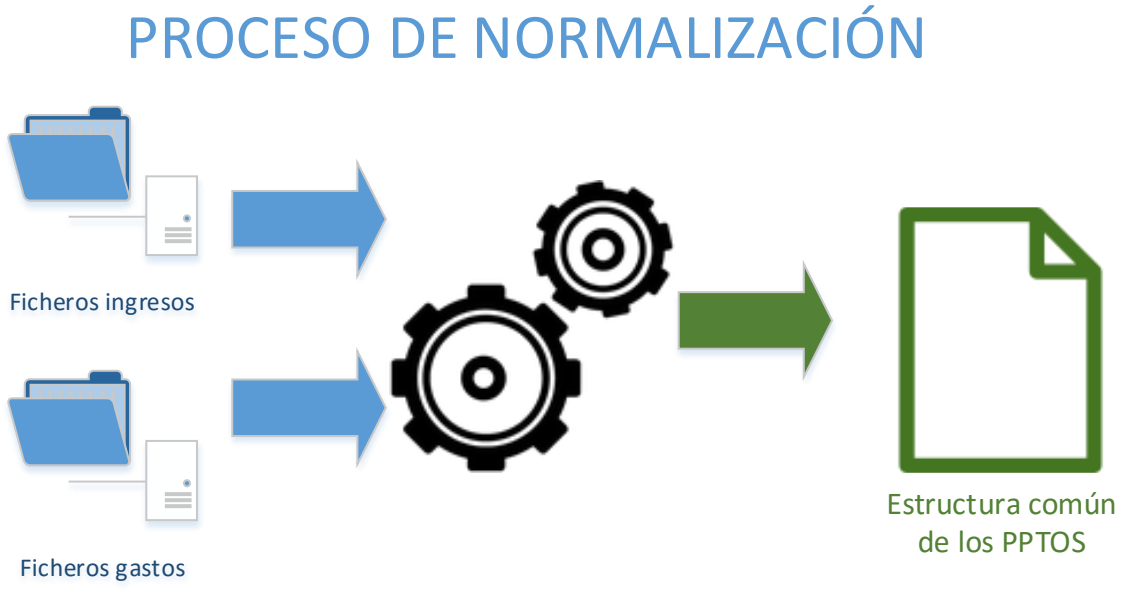


Ilustración 23 - Proceso de normalización esquemático

La curación que se ha llevado a cabo en el proyecto, está basado en los siguientes puntos:

- **Tipo de datos**: Los presupuestos vienen indicados con distintas unidades de medida, por lo que se estandariza la unidad a millones de euros.
- **Identificadores**: En la partida de gastos, los identificadores vienen incluidos dentro de la descripción de los organismos, por lo que se debe separar en dos: identificador y descriptor.
- **Estandarización padre-hijo**: llamamos padre al ministerio e hijo al ente o entes que cuelgan del mismo. A lo largo de la historia, los entes (agencias, organismos, etc.) van dependiendo de uno u otro ministerio según lo haya decidido el gobierno de ese año. También se modifican los nombres de los mismos organismos o se añaden a otros ya existentes. Para poder comparar ministerios de manera equivalente, se toma en consideración el año actual (2016) y la estructura ministerial que actualmente está vigente. Según vamos retrocediendo en los años, cada variación de ente se va ubicando en el ministerio que proceda según la estructura del 2016. Así también, se unifican los nombres para adecuarlo al año en curso.

4. Almacenamiento de datos

4.1. Normalizar los presupuestos

Una vez normalizado el fichero presupuestario, el almacenamiento en MongoDB es casi inmediato. La aplicación Robomongo ofrece una interfaz intuitiva para gestionar los datos de manera gráfica, de una manera similar a como se organizan en una base de datos SQL. Este almacenamiento perdurará en el tiempo y servirá para poder realizar análisis posteriores. Cabe mencionar que la estructura de datos sigue vigente, por lo que no hay que modificar la arquitectura.

5. Visualización de la información

Para poder realizar los gráficos relacionados con los presupuestos, se utiliza la herramienta Tableau Public. Orientada a presentar la información de manera descriptiva, el manejo de la herramienta requiere una cierta destreza de diseño.

Utilizando los conocimientos adquiridos durante el Master, se aplica la rueda de tensión descrita en el libro de Cairo, El Arte Funcional.

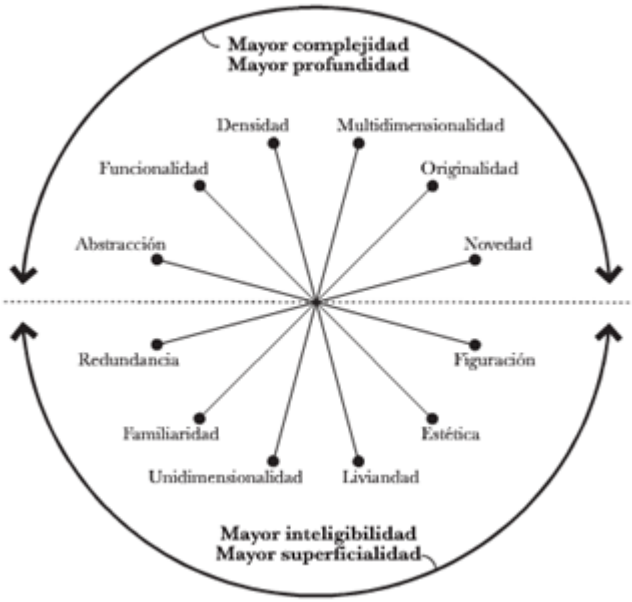


Ilustración 24 - Rueda de tensión de "El Arte Funcional"

Teniendo en cuenta que el público objetivo puede o no tener conocimientos financieros y el objetivo es que los datos que se viertan sean entendibles, se utilizan características que aplican

mayor inteligibilidad aunque también mayor superficialidad. Así, por ejemplo, se introducen redundancia en los datos y se aplican diagramas de barras conocidos por toda la población occidental. Sin embargo, para darle un poco más de análisis al cuadro de mando, se aplica una multidimensionalidad de grado 3 y una navegabilidad que ayuda a agrupar conceptos.

La estética también se mantiene sobria para no distraer las miradas de la información y se crea una pestaña de anexos con información que relaciona el presidente de gobierno con los gastos presupuestados cada año.

5.1. Visualización de los datos en gráficos

Para poder realizar el cuadro de mando, Tableau exige que se cree cada pestaña por separado, por tanto, se realizan 4 visiones diferenciadas:

1. **Evolución Totales Ingresos y Gastos**: Se permite ver la evolución del importe total de ingresos y gastos presupuestados sobre el mismo diagrama. Además, se indica con una etiqueta el valor para ayudar más a la comprensión.
2. **Desglose de Ingresos y Gastos**: Se visualiza el porcentaje de ingresos o gastos que compone ese año utilizando un gráfico tipo TreeMap. Si se pincha en cualquier casilla que lo compone, se desglosará en el detalle inferior la cantidad seccionada por organismo. Si no se pincha ningún concepto, en el detalle aparecerá todo el ejercicio presupuestario.
3. **Evolución por órgano ministerial:** En esta hoja se visualiza la evolución de cada ministerio a través de los años. Por tanto, esta gráfica refleja las tendencias presupuestarias que se ha fijado desde el 2007 hasta el 2016. En la parte inferior del detalle, se observa el desglose del ministerio, pudiendo seguir las evoluciones de los organismos que lo componen. Esto es muy importante ya que ayuda a entender qué factores componen cada organismo en cada año y permite analizar picos presupuestarios.
4. **Anexo. Evolución del Gasto**: Como valor añadido, se relaciona el gasto presupuestario con el presidente y partido político que gobernaba ese año. Sin embargo, esta pestaña es la única que no comienza en 2007 ya que se disponen de datos normalizados desde el año 2000. Como no se enfrenta el dato contra los ingresos, sino contra el presidente del gobierno, se amplía el análisis estos 7 años extra.

6. Encuesta de Satisfacción

6.1. Creación de la encuesta de satisfacción

Para evaluar esta aplicación y comprobar si se cumplen los objetivos marcados, se elabora una encuesta de satisfacción creada con Google Form. El estudio está orientado a una población aleatoria que, voluntariamente, rellene el formulario con la comparativa entre el portal Maestro disponible en la web del Ministerio para el año 2016 y la herramienta "Democratización de los Presupuestos Generales del Estado".

La aplicación de Google para formularios permite capturar las respuestas de manera anónima y analizar las respuestas de manera conjunta o separada. El formato que utiliza para enunciar cada pregunta es fácil e intuitivo, incluyendo enlaces a las web que son objeto de análisis.

6.2. Evaluación del cuadro de mando

Para moldear las impresiones de los encuestados, los voluntarios tendrán que contestar a las siguientes cuestiones.

Como pregunta introductoria que sirve para evaluar el interés previo sobre la materia de los encuestados aparece la pregunta:

1. **¿Ha visitado alguna vez la página web de los presupuestos generales del estado?**

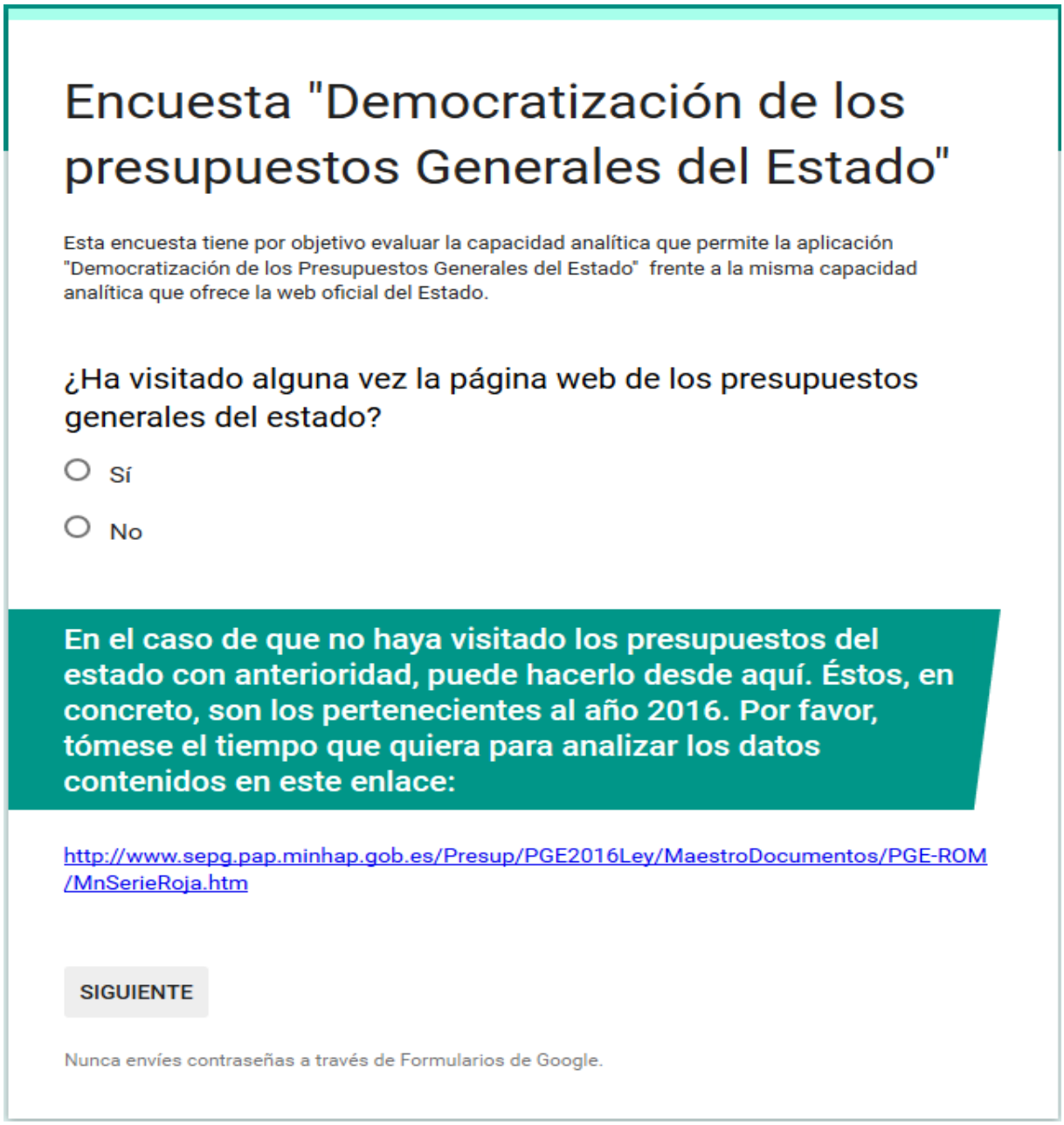
Posteriormente, deberán contestar a los siguientes asuntos relacionados primero con el portal Maestro del Ministerio y posteriormente, con la herramienta "Democratización de los Presupuestos Generales del Estado":

1. **¿Se puede encontrar de una manera sencilla cuánto prevé gastar el gobierno para 2016? En caso afirmativo, indique la cantidad**
2. **¿Se puede comparar los ingresos o los gastos totalitarios con los de años anteriores de manera fácil?**
3. **¿Se puede comparar el ingreso y el gasto de cada ministerio?**
4. **¿Se puede analizar las tendencias presupuestarias, es decir, cómo ha evolucionado la sanidad, la educación, o las pensiones?**
5. **En general, ¿La herramienta web proporcionada por el ministerio permite entender los presupuestos generales del estado?**

A parte de las anteriores preguntas, también se solicita que el usuario valore del 1 al 5 siendo 1 muy poco de acuerdo y 5 muy de acuerdo, la herramienta según las siguientes afirmaciones:

1. **El entorno le resulta amigable.**
2. **Ver la información de manera gráfica ayuda y aporta valor añadido.**
3. **La web le ayuda a entender los presupuestos del estado.**
4. **Recomendaría la web a otras personas**

Por último, se permite al encuestado añadir un texto sobre las sugerencias que tenga sobre la aplicación.



Encuesta "Democratización de los presupuestos Generales del Estado"

Esta encuesta tiene por objetivo evaluar la capacidad analítica que permite la aplicación "Democratización de los Presupuestos Generales del Estado" frente a la misma capacidad analítica que ofrece la web oficial del Estado.

¿Ha visitado alguna vez la página web de los presupuestos generales del estado?

- Sí
- No

En el caso de que no haya visitado los presupuestos del estado con anterioridad, puede hacerlo desde aquí. Éstos, en concreto, son los pertenecientes al año 2016. Por favor, tómese el tiempo que quiera para analizar los datos contenidos en este enlace:

http://www.sepg.pap.minhap.gob.es/Presup/PGE2016Ley/MaestroDocumentos/PGE-ROM/MnSerieRoja.htm

SIGUIENTE

Nunca envíes contraseñas a través de Formularios de Google.

Ilustración 25 - Encuesta de satisfacción

9. RESULTADOS

La herramienta "Democratización de los Presupuestos Generales del Estado" ha sido diseñada para extraer de la página web oficial del Ministerio de Hacienda y Administraciones públicas los datos oficiales de los presupuestos desde 2007 hasta 2016. Además, también trabaja sobre los datos realizando un proceso de tratamiento y normalización de los mismos para su posterior análisis visual. Por último, esta herramienta tiene como misión comprobar (a través de esta encuesta) que una visualización gráfica de los presupuestos ayuda a entender los mismos, aumentando el conocimiento del lector y su capacidad de decisión.

9.1. Apariencia de la herramienta

Como resultado del análisis y posterior tratamiento de los datos, **se ha obtenido un cuadro de mando que resume toda la información contenida en la base de datos**, permitiendo al usuario cambiar el punto de vista y seleccionar los elementos que más le interesen para su análisis. Se ve a continuación la interfaz del proyecto "Democratización de los Presupuestos Generales del Estado" indicando con notas numeradas las partes que componen cada visualización.

9.1.1. Evolución Totales Ingresos y Gastos

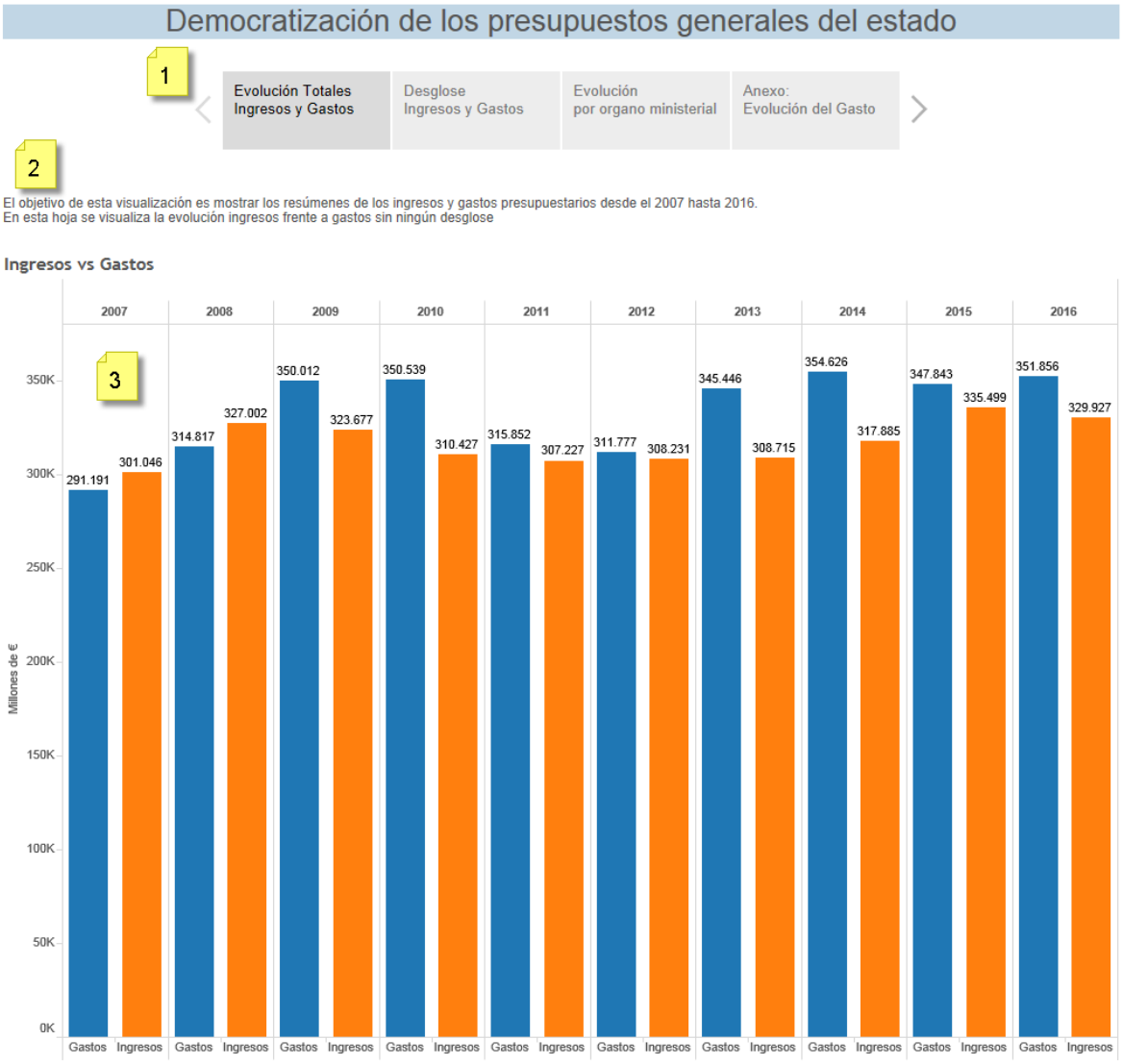


Ilustración 26 - Página Evolución Totales Ingresos y Gastos

Esta es la pantalla de bienvenida que aparece al pinchar sobre el enlace de la aplicación. Se puede comprobar que está compuesta de 3 partes: (1) la barra de navegación, (2) una descripción breve y (3) la gráfica que compara los ingresos y los gastos desde 2007 hasta 2016.

El objetivo de esta página es que un lector novel sea capaz de comprenderla. Por eso se ha elegido un diagrama de barras básico con dos variables (Ingresos y Gatos) donde se mantiene el significado de cada barra para facilitar la lectura. También se incluyen las etiquetas de los valores para una interpretación más sencilla.

9.1.2. Desglose Ingresos y Gastos

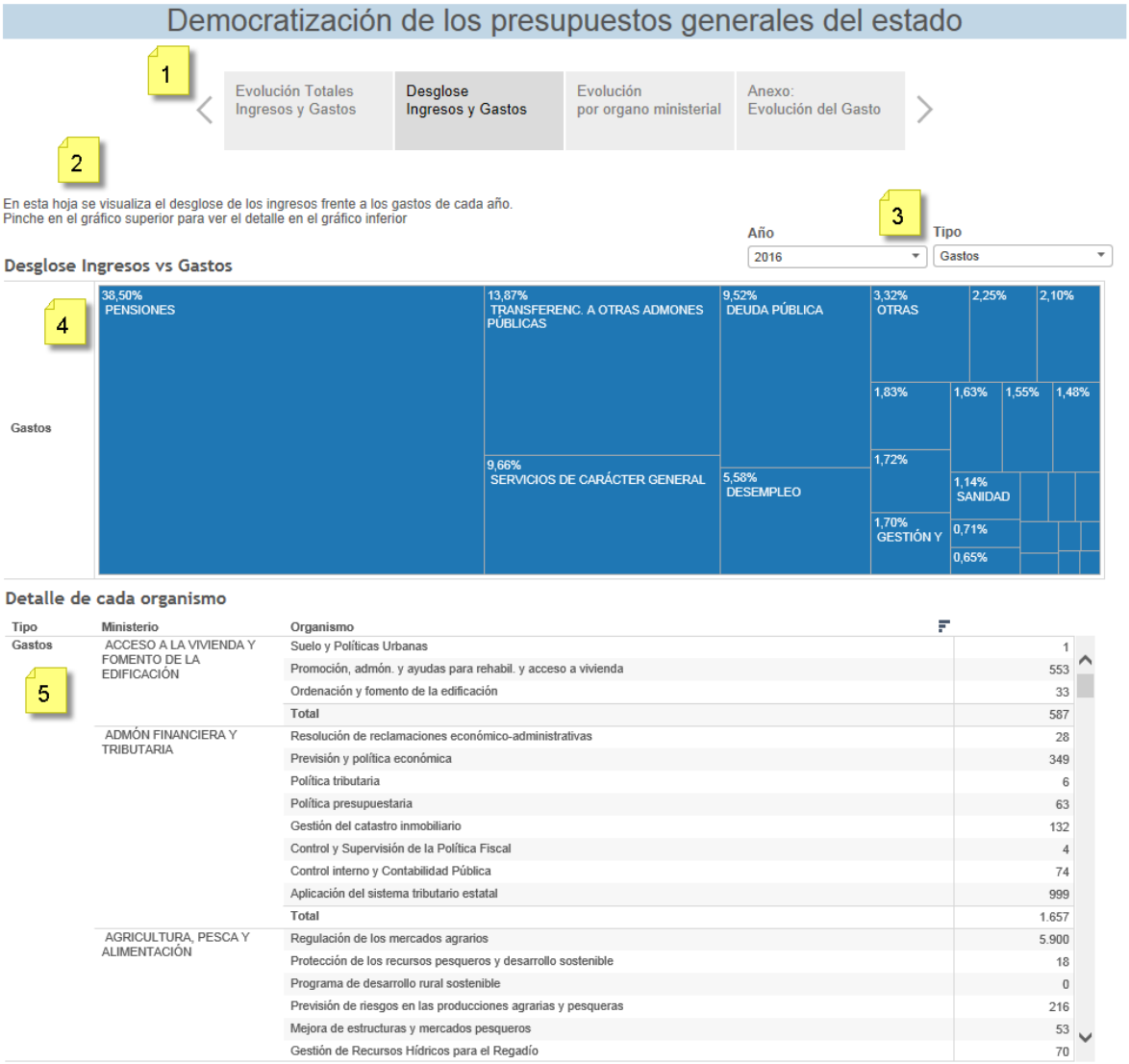


Ilustración 27 - Página Desglose Ingresos y Gastos

Esta pantalla está compuesta de 6 estructuras que aumentan la complejidad. Si bien (1) los botones de navegación y (2) la descripción de la página se mantienen, aparece como novedad dos selectores (3) que permiten al lector seleccionar el año presupuestario que quiere analizar y la tipología (Ingresos o Gastos).

Estos selectores filtran tanto la sección de (4) "Desglose Ingresos vs Gastos" como (5) "Detalle de cada organismo".

Una vez escogido el año y el tipo, en el desglose gráfico aparece un gráfico Treemap donde un usuario avanzado puede sacar más detalle sobre la partida presupuestaria. De manera visual se identifican carteras que afectan más porcentualmente al monto y en el detalle puede verse

exactamente la cifra y su desglose por organismos. Si no existe ningún concepto seleccionado, la tabla inferior mostrará toda la partida presupuestaria de la tipología y año seleccionado.

En la siguiente figura, se puede ver cómo usando los selectores y seleccionando una de las partidas, se desglosa el detalle en la tabla inferior haciendo esta tabla más legible.

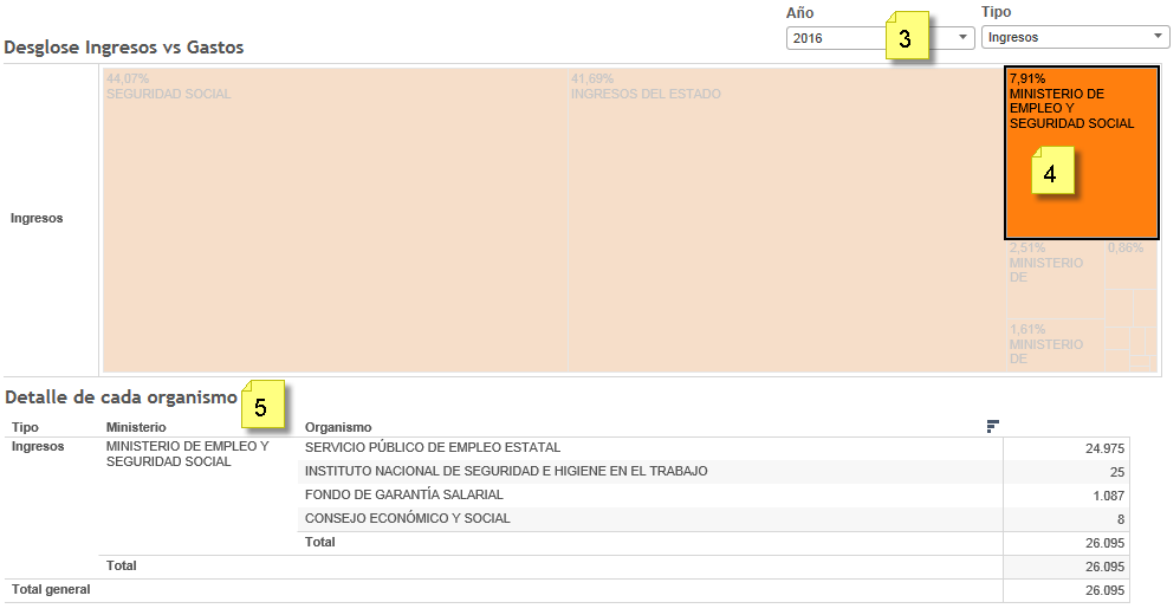


Ilustración 28 - Detalle de la página Desglose Ingresos y Gastos

Para eliminar la selección realizada en el Treemap, basta sólo con hacer clic en el texto "Desglose Ingresos vs Gastos" o en un espacio no graficado.

9.1.3. Evolución por órgano ministerial

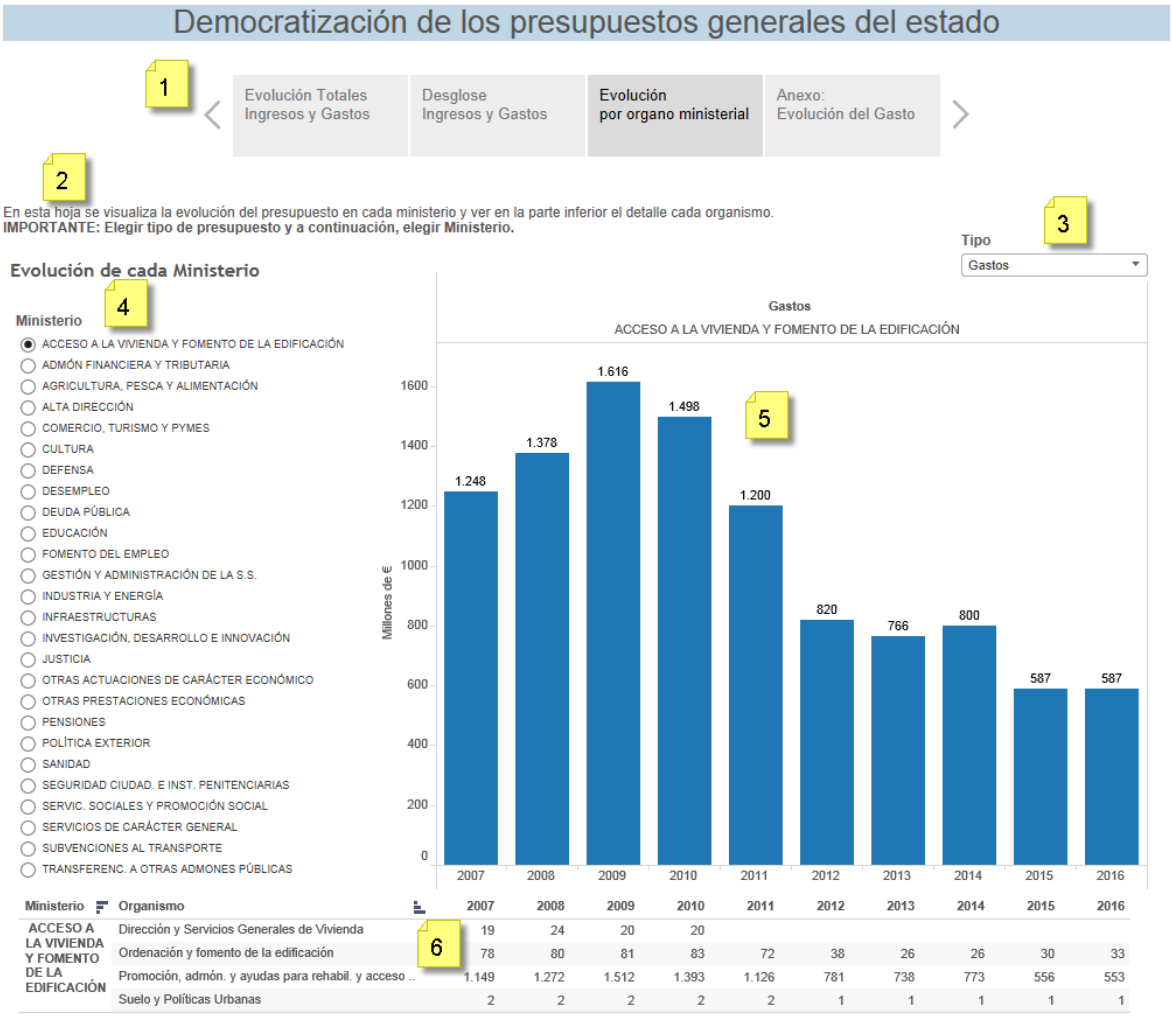


Ilustración 29 - Página Evolución por órgano ministerial

La tercera página ya no requiere un esfuerzo del lector ya que se vuelve a un gráfico de barras que simula la tendencia presupuestaria de cada organismo. Así pues, usando (3) el selector de Tipo y seleccionando en la parte derecha (4) el Ministerio que más le interese, aparece un (5) gráfico que muestra la evolución anual de esa cartera.

Como a lo largo de la historia, los organismos han variado en el tiempo bien por el título otorgado, bien por el ministerio al que rinden cuentas, bien porque ha aparecido o desaparecido, se muestra en la parte inferior (6) una tabla que muestra el detalle del ministerio. Así pues, puede visualizarse de una manera rápida la composición de cada año y analizar posibles picos o disminuciones de cada presupuesto.

9.1.4. Anexo: Evolución del Gasto

Como última página se añade una evolución del gasto. Aparece como anexo porque intervienen dos factores que no aplican en el resto de la herramienta:

1. El gasto se analiza desde el año 2000 hasta el 2016
2. Se relaciona la información con el partido político gobernante y el presidente que estaba a la cabeza del gobierno.

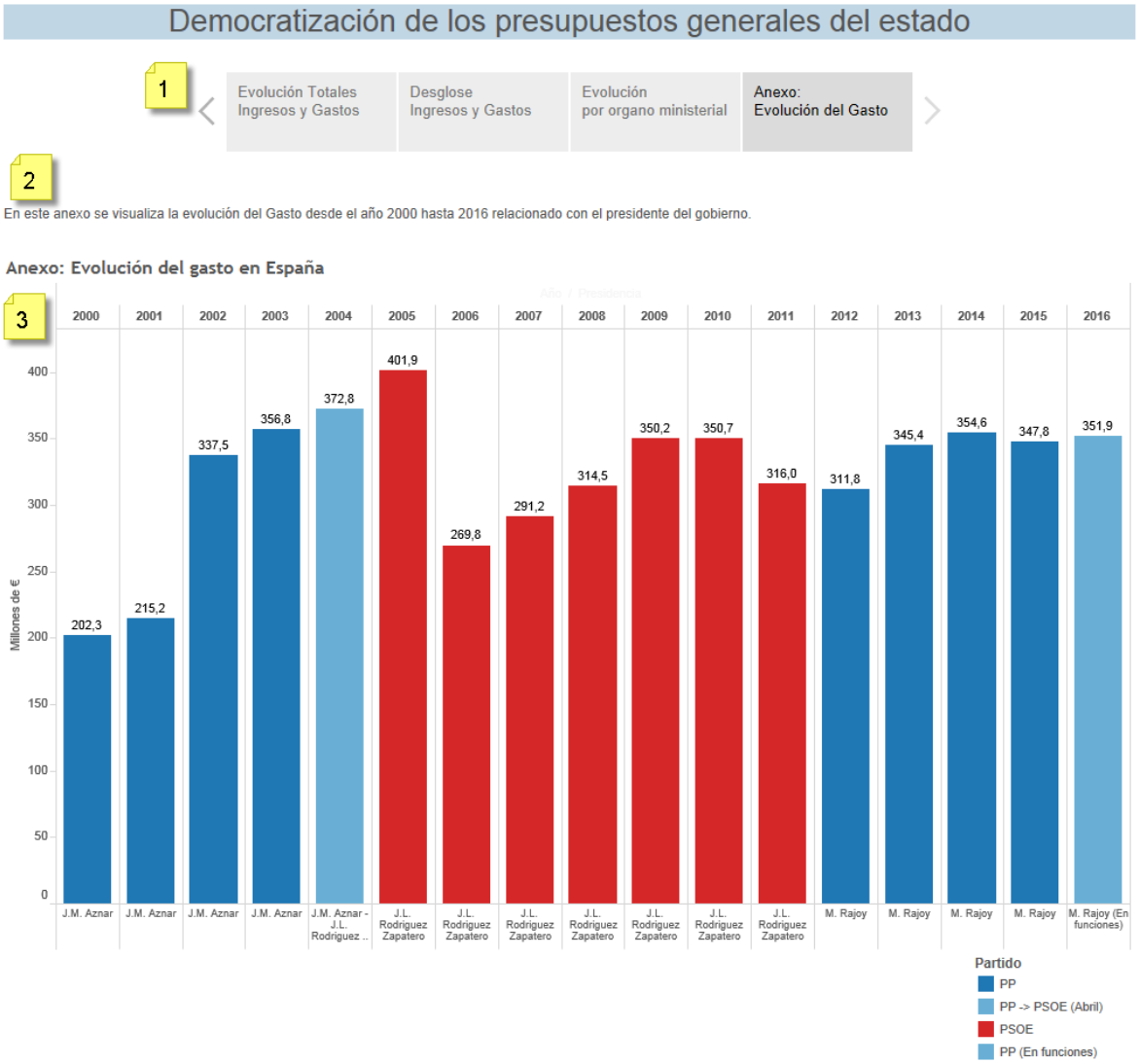


Ilustración 30 - Página Anexo: Evolución del Gasto

Esta última pestaña ayuda a relacionar información externa con los datos puros presupuestarios, siendo muy útil a la hora de formarse sus propias impresiones.

9.2. Valoración de la herramienta

Se ha realizado una encuesta digital para conocer si la herramienta cumple con los objetivos marcados. Para ello, el público que la ha realizado se ha prestado voluntario y los resultado se han guardado de manera anónima.

Aun cuando el tiempo de evaluación ha sido tan ajustado, se han realizado en total 21 encuestas y los resultados obtenidos han sido muy reveladores:

- El 52% de la población había visitado antes la página de los presupuestos generales del estado y, sin embargo, sólo **un 5% ha sabido contestar correctamente a la pregunta sobre la previsión de gasto para 2016**.
- Sin embargo, el **80% de la población consigue contestar correctamente a la pregunta de la previsión de gasto para 2016** sin haber visto previamente la herramienta "Democratización de los presupuestos generales del Estado".

- Más del 95% de la población considera que la web del ministerio no permite comparaciones entre los ingresos y los gastos, evoluciones y tendencias y, en general, que la web no es una herramienta "entendible".
- Sin embargo, **más del 95% de la población sí que considera que se pueda hacer comparativas y tendencias con la herramienta**.

- **El 100% de la población encuestada afirma que la aplicación "Democratización de los presupuestos Generales del Estado" ayuda a entender los presupuestos.**

- **El 98% de la población considera que la información gráfica aporta valor añadido y conceden una puntación de 9 sobre 10 a la herramienta.**

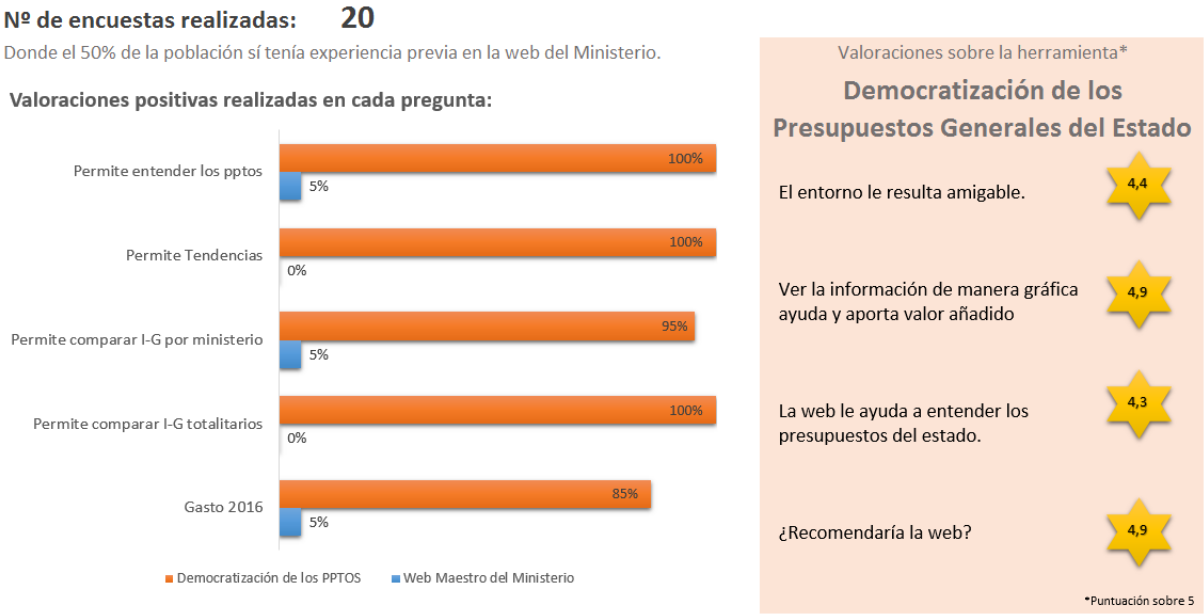


Ilustración 31 - Resultados de la encuesta

9.1. Otros datos de interés

Con un proyecto de esta envergadura, cabe también mencionar el coste que ha supuesto cada fase dentro del marco del proyecto. Así pues, vamos a determinar los datos globales y disgregados para intentar ver el alcance de la información vertida.

- La fecha de inicio para "Democratización de los Presupuestos Generales de Estado" es el 20 de Diciembre de 2016 y la fecha de finalización es el 22 de Abril de 2016. En total, se ha tardado 4 meses en realizar la aplicación desde 0 y su memoria y sólo para la aplicación, 112 días.
- Si sólo tenemos en cuenta la parte técnica, la preparación del dato ha supuesto el 70% del tiempo empleado. Si se desglosa por fases, se ve que:
  .1. La extracción del dato supuso 15%
  .2. La coherencia y análisis, 26%
  .3. La transformación del dato, 30%
  .4. El almacenamiento, 10%
  .5. La visualización, 14%
  .6. La encuesta de satisfacción, 5%.

Se puede ver los distintos análisis temporales en el siguiente diagrama:

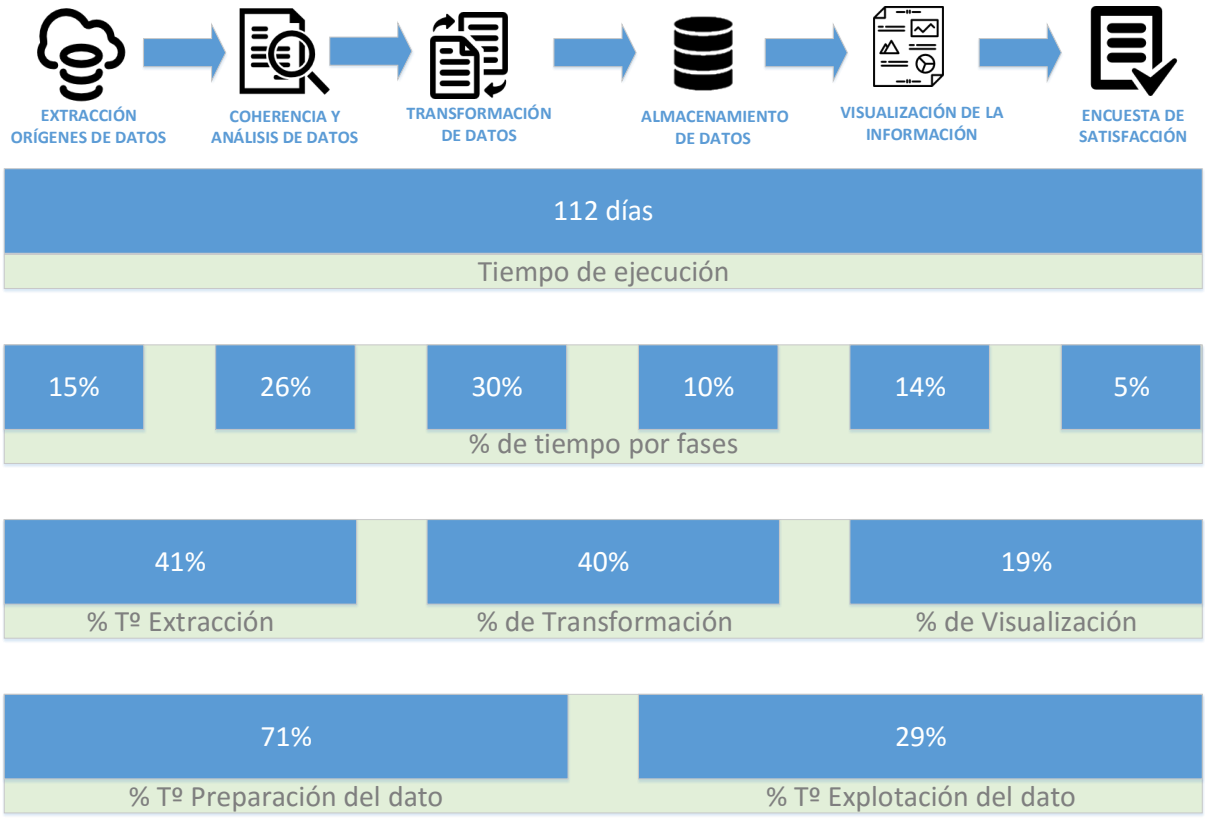


Ilustración 32 - Tiempos empleados

10. CONCLUSIONES

10.1. Sobre la herramienta

Tras la ejecución de este proyecto y basándonos en las puntuaciones vertidas por los encuestados, se puede **afirmar que la visualización gráfica es una herramienta muy potente para el análisis de información**.

La capacidad que el ser humano tiene para procesar información desciende inversamente cuando ésta está suministrada en grandes cantidades. Por tanto, son necesarias herramientas que ayuden a una primera digestión y que permitan ir profundizando según el interés y el conocimiento del analista.

Sin embargo, 2 de cada 3 horas se emplearon en el tratamiento del dato. Esto demuestra que los **procesos de normalización son costosos** y que son el mayor condicionante en un proyecto de análisis.

Analizando las respuestas libres de la encuesta, se evidenció que el **público objetivo** que vaya a hacer uso de la herramienta **condiciona la complejidad de la visualización** Se debe diferenciar entre un usuario novel que necesite redundancia y familiaridad frente a un usuario avanzado que demande novedad y granularidad.

10.2. Líneas futuras

Como procesos de mejora y tras la ejecución del proyecto, se pueden aplicar las siguientes recomendaciones:

1. **Automatización en el proceso de análisis**. Si bien es cierto que es necesario un análisis pormenorizado de los conceptos, su coherencia y veracidad, existe la posibilidad de automatizar ciertos cambios (errores léxicos o consensos) para posibilitar la recarga de información en el caso de necesidad. Esta línea ya se comenzó a explotar aunque no fue finalizada por falta de tiempo.
2. **Ahondar más la visualización**. Tableau Public es una herramienta potente que permite una amplia personalización. Realizar navegaciones dentro de los gráficos Tree Map o Pie Chart reforzaría el concepto de “partida dividida en subpartidas”. También

una visualización propia implementada con D3 también podría aportar la flexibilidad que exige una visualización presupuestaria.

3. **Mayor profundidad**. Para la realización de este prototipo se ha utilizado 3 nieves de granularidad. Sin embargo, los presupuestos oficiales ofrecen hasta 5 niveles, por lo que se podría seguir ahondando en el análisis de los datos.

BIBLIOGRAFÍA

- Ministerio de Hacienda y Administraciones Públicas, **Presupuestos Generales del Estado**. Sitio visitado el 14 de abril de 2016. [Online]. Disponible en: http://www.minhap.gob.es/es-ES/Paginas/Home.aspx
- CAIRO, Alberto. "*El arte Funcional*". Madrid: Alamut, 2011.
- Wiki – Finanzas, **Presupuestos Generales del Estado**. Sitio visitado el 1 de abril de 2016. [Online]. Disponible en: http://www.wiki-finanzas.com/index.php?seccion=Contenido&id=2014C01360
- Sistema Operativo GNU, **Licencia GPL**. Sitio visitado el 28 de marzo de 2016. [Online]. Disponible en: http://www.gnu.org/licenses/gpl-3.0.en.html
- Enciclopedia Libre Wikipedia, **formato XML**. Sitio visitado el 6 de abril de 2016. [Online]. Disponible en: https://es.wikipedia.org/wiki/Extensible_Markup_Language
- Enciclopedia Libre Wikipedia, **Diagrama de flujo**. Sitio visitado el 28 de marzo de 2016. [Online]. Disponible en: https://es.wikipedia.org/wiki/Diagrama_de_flujo
- Aiteco Consultores, **Diagrama de flujo**. Sitio visitado el 31 de marzo de 2016. [Online]. Disponible en: http://www.aiteco.com/que-es-un-diagrama-de-flujo/
- Iconos diseñados por Freepik desde Flaticon con licencia Creative Commons BY **CC 3.0 BY.** [Online]. Disponible en: http://www.Flaticon.com
- Lab RTVE, Evolución de los Presupuestos Generales del Estado, año 2012. Sitio visitado el 6 de abril de 2016. [Online]. Disponible en: http://lab.rtve.es/evolucion-presupuestos-generales-estado/index.html
- Ministerio de Hacienda y Administraciones Públicas, **Presupuestos Generales del Estado**. Sitio visitado la última vez el 1 de abril de 2016. [Online]. Disponible en: http://www.sepg.pap.minhap.gob.es/sitios/sepg/es-ES/Presupuestos/pge2016/Paginas/pge2016.aspx
- Enciclopedia Libre Wikipedia, **Richard Stallman**. Sitio visitado el 13 de abril de 2016. [Online]. Disponible en: https://es.wikipedia.org/wiki/Richard_Stallman

- Página oficial de Richard Stallman. Sitio visitado el 13 de abril de 2016. [Online]. Disponible en: https://stallman.org/
- Marta Beltrán Pardo, Fernando Sevillano Jaén, *Cloud Computing, Tecnología y negocio*. España: Ed. Paraninfo, 2013. [Online]. Disponible en: https://books.google.com/books?isbn=8428335141
- Finanzas para mortales, Presupuestos Generales del Estado. Sitio visitado la última vez el 16 de abril de 2016. [Online]. Disponible en: http://enletragrande.finanzasparamortales.es/visualizador.php?doc=22
- Dirección General de Presupuestos en Portugal, Presupuestos Generales del Estado. Portugal, 2016. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: www.dgo.pt/
- Centro de documentación de economía y finanzas francés. Francia, 2016. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: http://www.economie.gouv.fr/cedef/chiffres-cles-budget-etat
- Página oficial de los presupuestos del estado UK, Inglaterra. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: http://budgetresponsibility.org.uk/efo/economic-and-fiscal-outlook-november-2015/
- Página oficial de los presupuesto del estado italiano, Italia. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: http://www.rgs.mef.gov.it/VERSIONE-I/Attivit--i/Bilancio_di_previsione/Bilancio_semplificato/Gennaio-2015/
- Página oficial de los presupuesto del estado alemán, Alemania 2016. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: http://www.bundeshaushalt-info.de/#
- Open Knowledge Foundation Deutschaland, página no oficial con los Presupuestos del Estado Alemán. Alemania, 2016. Sitio visitado el 28 de Marzo de 2016. [Online]. Disponible en: http://offenerhaushalt.de/haushalt/bund/

ANEXOS

CORREOS ELECTRÓNICOS A LA DIRECCIÓN OFICIAL DEL MINISTERIO

Se solicitó hasta en tres ocasiones la ayuda a la hora de resolver dudas presupuestarias utilizando el correo electrónico que el ministerio tiene para tal fin. Sin embargo, nunca hubo respuesta:

De: Sandra Cobos [mailto:sandra.cobos@provide.es]
Enviado el: miércoles, 20 de enero de 2016 13:40
Para: 'internetdgpresupuestos@sepg.minhap.es' <internetdgpresupuestos@sepg.minhap.es>
Asunto: Presupuestos Generales del Estado 2015

Muy buenas tardes,

Mi nombre es Sandra Cobos y estoy analizando los presupuestos generales del estado del 2015 para un Trabajo Fin de Master de la Universidad de la Rioja (UNIR). Al estudiarlos, me han surgido unas dudas y me gustaría, si son tan amables, consultarlas con ustedes para intentar aclararlas.

En el siguiente vínculo se muestra el "Resumen general por organismos y capítulos del presupuesto de ingresos":

http://www.sepg.pap.minhap.gob.es/Presup/PGE2015Proyecto/MaestroDocumentos/PGE-ROM/doc/HTM/N_15_A_R_2_102_1_A_1.HTM

Al estudiarlo, en un principio pensé que era la cuantía que cada ministerio recibía por cada capítulo, pero viendo que estoy en el apartado de "Ingresos", creo que estaba equivocada y que las cifras que aparecen corresponden a lo que cada Ministerio ha ingresado por el concepto de cada capítulo. Por ejemplo:

Si nos vamos a la Orgánica 13 – **Ministerio de Justicia**, se han ingresado un total de **109.843,36 Miles de Euros**. Esta cuantía se ha recibido de la siguiente manera:

99.346,93 Miles de € mediante la Mutualidad General Judicial (orgánica 13.102)

10.496,43 Miles de € mediante el Centro de Estudios Jurídicos (orgánica 13.101)

Nótese que no estoy viendo la distribución por capítulos para simplificar la pregunta.

Quedo a la espera de su respuesta.

Reciban un cordial saludo

Sandra Cobos

De: Sandra Cobos [mailto:sandra.cobos@provide.es]
Enviado el: miércoles, 20 de enero de 2016 17:59
Para: 'internetdgpresupuestos@sepg.minhap.es' <internetdgpresupuestos@sepg.minhap.es>
Asunto: RE: Presupuestos Generales del Estado 2015

Muy buenas de nuevo,
Sigo encontrándome con algún escoyo a la hora de entender los presupuestos.
En el apartado de "Beneficios fiscales" incluido en el apartado de Ingresos
http://www.sepg.pap.minhap.gob.es/Presup/PGE2015Proyecto/MaestroDocumentos/PGE-ROM/doc/1/2/1/5/N_15_A_R_2_101_1_5_1_1.PDF

Los conceptos son del tipo:
1 IMPUESTOS DIRECTOS

10 Sobre la renta
100 De las personas físicas 15.216.620,00
10000 Reducción por rendimientos del trabajo 6.908.570,00
10001 Reducción por prolongación laboral 26.000,00
10002 Reducción por movilidad geográfica 13.180,00
10003 Reducción por aportación a sistemas de previsión social 1.001.420,00
10004 Reducción por arrendamientos de viviendas 530.510,00
[...]

Me choca la idea de que sean Ingresos cuando el concepto es "Reducción por ..." ya que entiendo que si es una reducción, lo que ha propiciado en las arcas del estado es un gasto. ¿Podrían explicarme el apartado de "Beneficios fiscales" dentro de los Ingresos del estado?

Gracias de antemano y un saludo

Sandra Cobos

De: Sandra Cobos [mailto:sandra.cobos@provide.es]
Enviado el: jueves, 21 de enero de 2016 12:53
Para: 'internetdgpresupuestos@sepg.minhap.es' <internetdgpresupuestos@sepg.minhap.es>
Asunto: Dudas de los pptos generales del estado

Muy buenas,

Mi nombre es Sandra Cobos y soy alumna del Master de Big Data de una Universidad de la Rioja (UNIR). Estoy realizando el TFM y tenía una pregunta general de los presupuestos generales del estado:

En el archivo "Resumen general x organismos y capítulos del ppto de ingresos", estos datos se refieren a los ingresos aportados por cada ministerio o es el capital que el Gobierno como entidad proporciona a cada ministerio? Es decir, en el caso de 2015 (ver adjunto) el ministerio de justicia ingresará 109843,36 M€ o es que el Gobierno le asignará de los ingresos que reciba un capital de 109843,36 M€

Muchas gracias por adelantado

Un saludo

Sandra Cobos