



Universidad de Jaén

Escuela Politécnica Superior de Jaén

VISUALIZACIÓN DE DATOS MEDIANTE TÉCNICAS DE BUSINESS INTELLIGENCE

Autor: Ana Lucía Álvarez Garrido

Grado: Ingeniería Informática

Director: Víctor Manuel Rivas Santos
Departamento del director: Informática

Fecha: 10/09/2024

Licencia CC



CREA



Universidad de Jaén
Escuela Politécnica Superior de Jaén
Departamento de Informática

Don Víctor Manuel Rivas Santos, tutor del Trabajo Fin de Grado titulado: '**Visualización de datos mediante técnicas de Business Intelligence**', que presenta Ana Lucía Álvarez Garrido, autoriza su presentación para defensa y evaluación en la Escuela Politécnica Superior de Jaén.

Jaén, septiembre de 2024

La alumna:

El tutor:

Ana Lucía Álvarez Garrido

Victor Manuel Rivas Santos

AGRADECIMIENTOS

A mi familia (padres y hermana), por todo lo que ha pasado en este tiempo, por aguantar mis rabietas sin saber muy bien lo que pasaba, solo espero que esteis orgullosos aunque haya tardado tanto en llegar el día que parecía que nunca lo iba a hacer, siempre habeis confiado en mi más que yo.

A ti Susana, por aguantar y ponerme las cosas claras, aún cuando ni yo las veía, al fin vamos a poder lanzar el gorro.

A Juan y Esteban, por estar siempre acompañandome y guiandome no solo en este proceso si no en la vida, sois pilar fundamental para mantenerme firme.

En especial a Luis, que, estando donde estás se que me has dado la fuerza interna para poder sacar esto y no dejarlo como tantas veces he pensado.

Gracias a mi tutor, Víctor, porque después de un problema que a ambos se nos escapaba cuando alguien me dijo que él podía ser el tutor perfecto para mí no dudo en tenderme la mano y ayudarme para que todo esto saliera adelante a pesar de todo.

Y por último y en cierto modo gracias a mí, realizar ese trabajo de fin de grado no ha sido fácil, no porque el tema fuese más o menos complicado si no por todo lo que he vivido hasta verme aquí cerrando por fin un ciclo de mi vida que tanto necesitaba. Cuando echo la vista atrás me acuerdo del momento en el que se acercaba “el final” y en el que sentí el miedo que hasta ahora ha estado atormentándome y del que por fin voy a poder liberarme.

Índice

Índice de Ilustraciones.....	4
Índice de Tablas.....	6
MOTIVACION	7
OBJETIVOS	7
1. INTRODUCCIÓN	7
2. BUSINESS INTELLIGENCE	8
2.1. ¿Qué es Business Intelligence?	8
2.2. Evolución a lo largo del tiempo.....	11
2.3. Actualidad	13
3. PROCESO BUSINESS INTELLIGENCE	15
3.1. Orígenes de datos.....	15
3.2. Proceso ETL	16
3.3. Data Warehouse	18
3.3.1. Diseño.....	20
3.3.2. Data Warehouse vs Data mart	21
3.3.3. Dos conceptos de almacenamiento de datos: Kimball vs Inmon.	23
3.3.4. OLAP	26
3.4. Visualización de datos.....	28
3.4.1. Herramientas utilizadas.....	30
4. POWER BI	37
4.1. ¿Qué es Power BI?	37
4.2. Interfaz Power BI (Descubriendo Power BI)	39
5. METODOLOGÍAS ÁGILES	46
5.1. ¿Qué son las metodologías ágiles?	46
5.1.1. Metodología Ágil vs Metodología Tradicional	48
5.1.2. Tipos de Metodologías Ágiles	50
6. EJEMPLOS PRÁCTICOS CON POWER BI	55
Primer ejemplo	56
Segundo ejemplo	59
Tercer ejemplo	61
Cuarto ejemplo	63
Quinto ejemplo	67
7. CONCLUSIONES	69
8. BIBLIOGRAFÍA	70

Índice de Ilustraciones

Ilustración 1. Pasos hacia la toma de decisiones.....	10
Ilustración 2. De los datos al conocimiento.....	11
Ilustración 3. Proceso Business Intelligence.....	15
Ilustración 4. Proceso ETL	17
Ilustración 5. Área Staging	17
Ilustración 6. Proceso Data mart dependiente	19
Ilustración 7. Proceso Data mart Independiente	19
Ilustración 8. Esquema Estrella	20
Ilustración 9. Esquema Copo de nieve	21
Ilustración 10. Arquitectura básica de Kimball	24
Ilustración 11. Arquitectura básica de Inmon	25
Ilustración 12. Proceso OLAP.....	26
Ilustración 13. Cubo OLAP	27
Ilustración 14. Magix Quadrant for Analytics and Business Intelligence Platforms.....	31
Ilustración 15. Matriz cuadrante de Gartner.....	31
Ilustración 16. Logos de herramientas de visualización de datos	32
Ilustración 17. Proceso de desarrollo con Power BI.....	38
Ilustración 18. Power BI Desktop Principal	39
Ilustración 19. Menú Inicio Power BI Desktop.....	40
Ilustración 20. Menú Insertar Power BI Desktop	40
Ilustración 21. Menú Modelado Power BI Desktop	40
Ilustración 22. Menú Ver Power BI Desktop	40
Ilustración 23. Opción Mostrar Paneles Power BI Desktop.....	41
Ilustración 24. Panel lateral derecho Power BI Desktop	42
Ilustración 25. Obtención de datos Power BI Desktop	43
Ilustración 26. Import Query Power BI Desktop	44
Ilustración 27. Direct Query Power BI Desktop.....	44
Ilustración 28. Opciones Avanzadas Origen de datos Power BI Desktop	45
Ilustración 29. Opción para Transformar Datos Power BI Desktop	46
Ilustración 30. Power Query en Power BI Desktop	46
Ilustración 31. Tradicional vs Ágil	49
Ilustración 32. Metodología Lean.....	50
Ilustración 33. Metodología Kanban	51
Ilustración 34. Metodología XP	52
Ilustración 35. Metodología Scrum	53
Ilustración 36. Metodología Ágil (Sprint)	54
Ilustración 37. Microsoft Planner	55
Ilustración 38. Modelo de datos.....	56
Ilustración 39. Mock up Ejemplo 1	57
Ilustración 40. Ventas con promoción por categoría	58
Ilustración 41. Ventas con promoción por subcategoría	58
Ilustración 42. Tabla de Ventas	59
Ilustración 43. Mock up Ejemplo 2.....	60
Ilustración 44. Grafico circular ventas por canal	60
Ilustración 45. Opciones de gráfico circular	61
Ilustración 46. Gráfico de mapa. Ventas por región	62

Ilustración 47. Mapa definitivo de ventas por Región.....	63
Ilustración 48. Opciones para reglas de colores	63
Ilustración 49. Boceto Gráfico de barras.....	64
Ilustración 50. Opciones de filtrado de gráficos para Top N.....	65
Ilustración 51. Gráficos de barras finales.....	65
Ilustración 52. Opciones de Selección, Marcadores y Formato	66
Ilustración 53. Gráficos finales con botones	66
Ilustración 54. Mock up Ejemplo 5.....	67

Índice de Tablas

Tabla 1. Comparativa Data Warehouse y Data mart	22
Tabla 2. Comparativa de las herramientas de visualización	37
Tabla 3. Comparativa Import vs Direct Query	45
Tabla 4. Metodología Ágil vs Metodología Tradicional.....	49

MOTIVACION

Hasta hace relativamente poco, unos tres años, la única idea que tenía del término Business Intelligence (BI) era que significaba Inteligencia de negocio. Cuando comencé a realizar prácticas en la empresa en la que hoy trabajo, me ofrecieron la oportunidad de elegir, aparte del área que ya conocía por la carrera (Full Stack), otras en las que también tenían proyectos, BI, área tecnológica en la que decidí adentrarme y que hasta entonces había sido completamente desconocida para mí.

Fue entonces cuando después de hablar con mi tutor, cambié completamente el rumbo de mi Trabajo de Fin de Grado y decidí realizarlo del área que estaba empezando a descubrir y que tanto me estaba gustando.

OBJETIVOS

La finalidad de este proyecto es dar a conocer la importancia de la visualización de datos en el proceso de Business Intelligence. Para esto, antes analizaremos la naturaleza del término Business Intelligence, su evolución en el tiempo, el proceso que se sigue hasta llegar a esa visualización y lo que supone para la actual situación empresarial.

Identificaremos y analizaremos algunas de las herramientas existentes para la creación de informes poniendo de manifiesto ventajas e inconvenientes de estas y centrándonos en una en concreto, Microsoft Power BI.

Además, desarrollaremos ejemplos de informes realizados con la herramienta Power BI a partir de peticiones simuladas por un cliente final.

1. INTRODUCCIÓN

La visualización de datos tiene sus orígenes en la prehistoria, en esas pinturas rupestres representadas en las cuevas y que narraban las cacerías. Continuó con la representación de datos en mapas y diagramas, en el siglo XVII comenzaron a aparecer gráficos para analizar estadísticas de mortalidad, en el XVIII se introdujeron

gráficos de barras y líneas y fue en el siglo XIX cuando la visualización de datos comenzó a ganar popularidad en ciencia y estadística. En el siglo XX con la llegada de las nuevas tecnologías se desarrollaron herramientas más sofisticadas y se expandieron a otros campos como la psicología o economía. Y ya en el siglo XXI con la revolución digital y el auge de la informática es cuando la visualización de datos ha tomado un papel fundamental.

En la era de la información, las empresas se enfrentan al reto de convertir grandes cantidades de datos en información útil que potencie la toma de decisiones estratégicas. El concepto de Business Intelligence (BI) ha surgido como respuesta esencial para abordar esta situación, comprender de manera más profunda como se comportan los clientes y prever los cambios en el mercado. Este trabajo describirá el concepto BI, analizando su definición, procedimiento y aplicación actual, además de comparar las diferentes herramientas disponibles para lograr una visualización de datos efectiva.

Finalmente se desarrollarán ejemplos prácticos que demostraran la capacidad de la herramienta de Microsoft Power BI, la cual ha ganado notable aceptación en el mercado y presenta características que la hacen destacar entre sus competidores, y será la utilizada para el desarrollo de la visualización de datos. Además, podremos ver un Informe creado con esta a partir de casos de uso específicos.

2. BUSINESS INTELLIGENCE

En este apartado vamos a tratar el término BI, su evolución y actualidad.

2.1. ¿Qué es Business Intelligence?

El término “Inteligencia de Negocios”, conocido habitualmente como Business Intelligence o “BI”, es muy utilizado hoy en día, tanto por expertos, como por las empresas, pero, ¿Qué significa Business Intelligence? No existe una definición concreta del término, por lo que empezaremos a desarrollarlo de la forma más sencilla, buscando en el diccionario de la Real Academia Española los dos términos que componen este concepto.

En primer lugar, “Business”, traducido como “Negocio” el cual tiene varios significados, entre ellos: *“aquello que es objeto o materia de una ocupación lucrativa o de interés”* (Real Academia Española, s.f., definición 3), *“utilidad o interés que se logra en lo que se trata, comercia o pretende”* (Real Academia Española, s.f., definición 5).

En segundo lugar, la palabra “Intelligence”, traducido como “Inteligencia”, que también puede significar varias cosas: *“capacidad de entender o comprender”* (Real Academia Española, s.f., definición 1), *“capacidad de resolver problemas”* (Real Academia Española, s.f., definición 2), *“conocimiento, comprensión, acto de entender”*, (Real Academia Española, s.f., definición 3), *“sentido en que se puede tomar una proposición, un dicho o una expresión”* (Real Academia Española, s.f., definición 4), *“habilidad, destreza y experiencia”* (Real Academia Española, s.f., definición 5).

Como podemos observar ambas palabras poseen diferentes significados, los cuales varían levemente y al unirlos pueden servir para crear una definición del término.

Avanzando un poco más en la definición encontramos varias propuestas para este concepto, Cano, J (2007). “Business Intelligence: Competir con Información” crea a partir del glosario de términos de Gartner¹ la siguiente definición: *“BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un datawarehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones. El proceso de Business Intelligence incluye la comunicación de los descubrimientos y efectuar los cambios. Las áreas incluyen clientes, proveedores, productos, servicios y competidores.”*

La de Curto, J (2010). “Introducción al Business Intelligence”, que entiende por BI *“al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización”*.

¹ Empresa consultora y de investigación de tecnologías de la información.

O la que proponen en “*The datawarehouse Insitute*”: “*Business Intelligence (BI)* es un término paraguas que abarca los procesos, las herramientas, y las tecnologías para convertir datos en información, información en conocimiento y planes para conducir de forma eficaz las actividades de los negocios. BI abarca las tecnologías de datawarehousing los procesos en el ‘back end’, consultas, informes, análisis y las herramientas para mostrar información (estas son las herramientas de BI) y los procesos en el ‘front end’”.

Tras leer algunas de las definiciones de este concepto concluiremos diciendo que el Business Intelligence son todas aquellas metodologías, tecnologías y estrategias que permiten agrupar datos con el objetivo de transformarlos en información de calidad y esa información en conocimiento, mejorando la velocidad y capacidad de las empresas para ayudar a mejorar la competitividad y facilitar así la toma de decisiones.



Ilustración 1. Pasos hacia la toma de decisiones

Fuente: Elaboración Propia

Un aspecto importante tras leer las definiciones es saber comprender y diferenciar entre datos, información y conocimiento.

- Dato: conjunto de caracteres que constituyen la base de todo sistema de codificación, aunque, por sí solos no tienen mucha aportación.
- Información: Dato o conjunto de datos que, en un contexto y lugar específicos, tienen importancia para alguien y comunican un mensaje práctico.
- Conocimiento: Proceso que permite transformar la información en decisiones y acciones.

De esta forma, una vez que los datos agrupados, clasificados y estructurados se entregan al usuario se convierten en información a partir de la cual obtendremos el conocimiento con el que tendremos mejores resultados.



Ilustración 2. De los datos al conocimiento.

Fuente: Ramos, S (2016). BI & Analytics. El arte de convertir datos en conocimiento

2.2. Evolución a lo largo del tiempo

En el día a día van surgiendo nuevas necesidades, nuevas formas de utilizar recursos anteriores y que van mejorando tanto a la hora de la implementación como en el mercado. Son pocas las veces que coincide la primera aparición de un elemento con su auge, por lo que nos remontaremos unos años atrás, para así encontrar la raíz y comprender el camino que ha seguido el término de Business Intelligence hasta posicionarse a la altura e importancia que tiene actualmente como líder de la gestión de negocios.

Ahondando en la historia, describiremos varios hitos durante el desarrollo de este concepto para comprender mejor su evolución a lo largo del tiempo.

Aunque parece reciente, fue el investigador de IBM Hans Peter Luhn, quien, en 1958 en su artículo “A Business Intelligence System” hizo mención al término con la siguiente definición: “*es la habilidad de aprender las relaciones de hechos presentados de forma que guíen las acciones hacia una meta deseada*”. Una definición algo básica, pero gracias a la cual se abrió un gran campo de investigación.

En 1962, el informático teórico y matemático canadiense Kenneth Iverson hizo un gran avance inventando el primer lenguaje de programación multidimensional, conocido como OLAP y que sería la base para el proceso analítico en línea.

A finales de los años sesenta, surgió la idea de base de datos y fue en los años setenta cuando se empezaron a desarrollar grandes bases de datos y aplicaciones empresariales, lo que impulsó el crecimiento del BI. Si bien es cierto que, a pesar de la existencia de potentes sistemas de bases de datos que permitían acceder a la información que se almacenaba en ellas, seguía siendo complicado encontrar información específica y lograr la eficacia correspondiente ya que eran de baja calidad, lentas y poseían un difícil acceso a información específica.

Hubo un avance durante los ochenta, gracias a la creación del concepto Data Warehouse y a una mejora en cuanto al reporting (gestión de documentos de una empresa, con información privilegiada relativa al marketing que lleva a cabo, para solucionar problemas de negocio). Y en 1985 Microsoft lanza al mercado una herramienta que se convirtió en la más popular y ampliamente utilizada, Excel 1.0.

Fue en la década de los noventa, época determinante para las tecnologías que usamos hoy día, cuando el término de BI se popularizó gracias a Howard Dresner, exvicepresidente de Gartner, quien volvió a reincidir en el tema después de varias décadas en el anonimato. Surgieron nuevas tecnologías y herramientas que junto a la creación de cubos de datos multidimensionales y a la generación de informes interactivos dieron paso a soluciones más integradas, pudiendo llegar a crear aplicaciones de BI que ofrecían información estructurada y acceso a las bases de datos de las empresas.

Con la entrada del nuevo milenio (2000s) llega la computación en la nube que revoluciona el BI, permitiendo a las empresas acceder a herramientas y datos desde cualquier lugar. Se empiezan a tener en cuenta información y documentos no estructurados y llega el Business Intelligence 2.0. La capacidad de procesamiento de datos aumenta significativamente, lo que permite el análisis de grandes volúmenes de información. Surgen nuevas tecnologías, como la minería de datos, el análisis predictivo y el descubrimiento de patrones, que amplían las capacidades de BI y ayudan a las organizaciones a obtener información más profunda y valiosa.

En la década de 2010, se observó el auge del Big Data, con la aparición de enormes cantidades de datos generados a velocidades nunca antes vistas y provenientes de múltiples fuentes. Las tecnologías de BI evolucionaron para lidiar esta gran cantidad de datos en cargas masivas, posibilitando análisis en tiempo real. Se produjo un cambio hacia la autosuficiencia y la democratización de datos. Surgieron herramientas de BI self- service que permitieron a los usuarios de negocio crear y acceder a informes y paneles de control sin depender de los equipos de TI. La visualización de datos se convirtió en una parte integral de BI, lo que facilitó la comprensión y la comunicación de la información empresarial de manera más efectiva.

Desde el 2015, la automatización ha tenido un papel fundamental en el desarrollo del término. Las herramientas empezaron a incluir características automatizadas lo que redujo el tiempo y esfuerzo en la realización de análisis complejos permitiendo así centrarse más en tareas estratégicas y mejorando la precisión y consistencia de los informes.

BI se adapta a las demandas del mercado y a las innovaciones tecnológicas aplicadas en los negocios, permitiendo a las empresas aprovechar de manera eficiente la información almacenada para beneficio de sus operaciones. La generación de informes y análisis por parte de esta herramienta son habilidades competitivas fundamentales para la toma de decisiones, lo que hace que el Business Intelligence sea esencial.

2.3. Actualidad

En la actualidad, prácticamente todas las organizaciones podrían utilizar tecnologías analíticas para mejorar la visualización de datos, realizar predicciones más precisas y tomar decisiones más adecuadas a los datos que se les presentan. No obstante, el reto radica no solo en empezar a aplicar esas tecnologías, sino en establecer un sistema de aprendizaje constante para que las empresas obtengan conocimientos al instante en tiempo real.

La rápida evolución de las tecnologías hace necesario reducir la distancia entre formación-tecnología-implementación para que las empresas puedan crecer y

aumentar su valor. En un entorno tan dinámico y cambiante y a la vez tan competitivo como el actual, las empresas enfrentan crecientes demandas del mercado, por lo que si obtienen resultados positivos es a consecuencia de una buena utilización de la información en la toma de decisiones.

Cada vez son más las diferentes fuentes que nos proporcionan datos, como ERP, CRM, WMS, SCM, redes sociales, Internet de las cosas, entre otros, con información más desorganizada y con la participación de clientes internos en las decisiones que se toman en las empresas de las que forman parte. Además, los usuarios tienen más necesidades y el tiempo de respuesta de las áreas de tecnología es menor, mientras que la paciencia de los clientes internos es limitada.

Gracias a las herramientas de visualización de datos, las empresas pueden tomar decisiones más estratégicas y aprovechar al máximo las oportunidades del mercado, al simplificar la gestión y el análisis de grandes volúmenes de datos.

En los últimos dos años, el enfoque del Business Intelligence ha experimentado una transformación significativa. Se ha pasado de un modelo tradicional, altamente tecnológico, a uno moderno que prioriza la flexibilidad y empodera a los ejecutivos para crear sus propios análisis ya que pueden acceder a la información, proporcionándoles datos precisos y confiables. Esta evolución ha sido impulsada por la creciente necesidad de agilidad en la toma de decisiones y por los avances tecnológicos, como la acogida de la nube y el desarrollo de plataformas analíticas más intuitivas.

En la última década, BI ha evolucionado hacia la analítica predictiva y el uso de inteligencia artificial. Las empresas ahora pueden no solo analizar datos históricos, sino también predecir tendencias futuras y comportamientos del cliente. Las herramientas de BI se han vuelto más accesibles y fáciles de usar, permitiendo que más personas dentro de una organización puedan aprovechar los datos.

3. PROCESO BUSINESS INTELLIGENCE

Como ya hemos comentado con anterioridad la finalidad del Business Intelligence es ayudar a las empresas en el proceso de obtener, analizar y utilizar datos en su beneficio para la toma de decisiones. En este apartado me centraré en desarrollar los distintos componentes de un sistema BI.

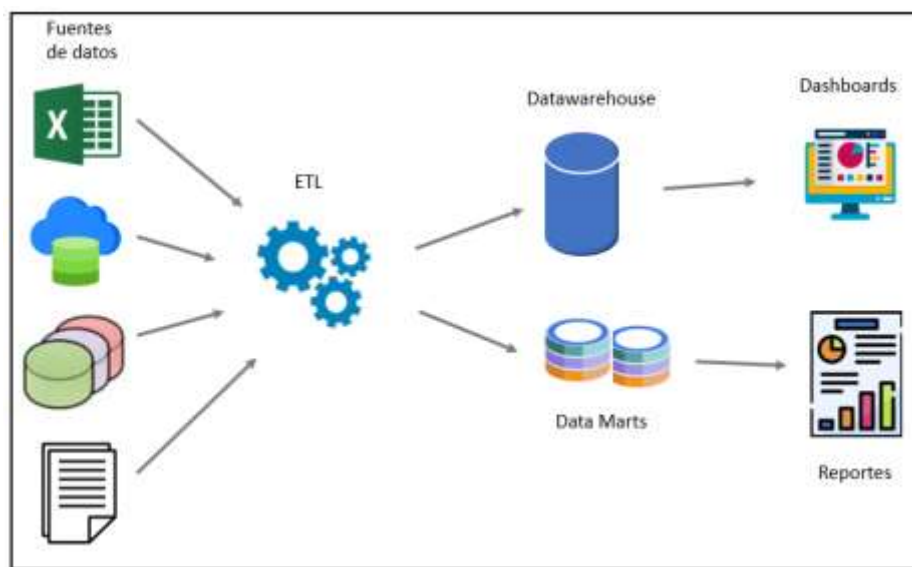


Ilustración 3. Proceso Business Intelligence

Ilustración 3. Pasos BI. Fuente: Elaboración Propia

3.1. Orígenes de datos

Diversas fuentes de información (orígenes de datos), que podemos obtener de varias maneras y en diferentes formatos BBDD transaccionales como ERP, CRM, RRHH, archivos (de texto, libros de Excel etc) páginas y servicios web, fuentes de datos abiertas, etc, y de las cuales procederemos a cargar la información en el datawarehouse. Existen varios factores que causan la complicación de la carga, siendo el principal el número de fuentes de información distintas; acceder a diferentes bases de datos implica tener habilidades y conocimiento de diversas sintaxis de SQL.

Un aspecto clave será identificar las fuentes más apropiadas, comprender el modelo de información transaccional y los elementos que lo componen. Sobre todo, es importante analizar si la información que tenemos es la adecuada. Una vez

elegidas las fuentes debemos verificar la calidad de los datos; como afirmó Bill Inmon en su artículo “Business Intelligence Network”: *“Las organizaciones actúan bajo la suposición de que la información de la que disponen es precisa y válida. Si la información no es válida, entonces no pueden responder de las decisiones basadas en ella.”*. Asumir que la calidad del dato es buena puede llegar a ser un error fatal en el futuro. Existen siete características que debería cumplir un dato para tener buena calidad: precisión, integridad, coherencia, totalidad, validez, disponibilidad y accesibilidad.

3.2. Proceso ETL

Siguiendo el modelo inicial ahora analizaremos el proceso ETL (extracción, transformación y carga), clave en el transcurso de un proyecto ya que ocupa una parte significativa de este y método potente que puede trabajar junto a herramientas de gestión e integración de datos para cumplir los objetivos requeridos. En él se unen los datos de las diferentes fuentes, se realizan los cálculos necesarios, se mejora la calidad de los datos y se adaptan al nuevo modelo físico. Se divide en tres subprocesos:

- **Extracción:** proceso en el que los datos se copian o exportan a una zona de preparación desde las distintas fuentes externas e interna. El principal objetivo es extraer únicamente los datos requeridos y prepararlos para el resto de subprocesos. Una parte sustancial del proceso es realizar un análisis de los datos extraídos y verificar que cumplan la estructura que se estipuló, de no ser así serán rechazados. Un requisito importante es que cause el mínimo impacto en el sistema de origen.
- **Transformación:** fase en la que se aplican una serie de reglas sobre los datos extraídos para convertirlos en datos válidos, íntegros y útiles para la posterior carga. Durante el proceso se incluyen: cambios de formato, valores derivados o agregados, cálculos de totales, división de columnas, eliminación de inconsistencias y duplicidades, etc.
- **Carga:** es la última etapa, cargamos los datos en el destino definido.



Ilustración 4. Proceso ETL

Ilustración4. Proceso ETL. Fuente: <https://www.talend.com/>

En ocasiones la información no pasa directamente del origen al almacenamiento de datos, sino que lo hace a través de unas bases de datos intermedias llamadas área de staging y/o ODS (Operational Data Store), necesarias dada la complejidad y diferencia de los datos en los orígenes.

El área staging es un espacio de almacenamiento temporal durante el proceso de ETL, creado con el objetivo de mejorar la extracción de los datos, realizar limpieza de datos, mejorar la calidad de los datos e incluso almacenando el caché de los datos operacionales.



Ilustración 5. Área Staging

Un almacén operacional de datos (ODS) es un espacio de almacenamiento provisional que da soporte a los sistemas transaccionales, y funciona como base de datos de consulta, a la que se conectan las herramientas de reporting reduciendo así la carga de trabajo a los sistemas transaccionales. Se encuentra normalizado y puede o no existir, independientemente de que exista un almacenamiento de datos, ya que no es componente específico del proceso de BI.

3.3. Data Warehouse

Data Warehouse o almacén de datos, es un repositorio de datos integrados desde una o varias fuentes que se explotan con la finalidad de convertir los datos en información entendible, organizada, actualizada y preparada para realizar análisis y producir informes. No es una base de datos operacional, es decir los datos se almacenan para un uso futuro. Se construye generalmente con un proceso ETL y según Bill Inmon debe cumplir las siguientes características:

Orientado a un área: los datos están relacionados por temas concretos para que así sea más fácil para los usuarios su comprensión.

Integrado: la información se organiza en diferentes niveles de precisión para adaptarse a las necesidades del usuario, debiendo eliminarse las inconsistencias existentes y tener una estructura de datos consistente.

Indexado en el tiempo lo que nos facilita analizar los cambios a lo largo de este desde una perspectiva histórica.

No volátil, una vez introducida la información debe ser solo de lectura, nunca modificarla o eliminarla, debe mantenerse permanente para futuras consultas.

Se pretende almacenar los datos de manera que se aumente su flexibilidad, y accesibilidad, así como su manejo a través de interfaces y operadores, aunque en ocasiones su administración puede resultar compleja y la recopilación de datos requiere mucho tiempo, por lo que se necesita un Data mart para que los usuarios puedan acceder de una forma sencilla y enfocada al análisis.

Según Ralph Kimball, *“Un Data mart es un conjunto de datos flexible, idealmente basado en el nivel de granularidad mayor que sea posible, presentado en un modelo dimensional que es capaz de comportarse bien ante cualquier consulta del usuario. En su definición más sencilla, un Data mart representa un único proceso de negocio”*.

Un Data mart es un subconjunto de datos orientado a un área de negocio específica con el enfoque de satisfacer las necesidades de esa determinada área.

Existen tres tipos de Data mart en función de cómo se alimentan, como obtienen sus datos y como se insertan en este:

- **Data mart dependiente**, generado a partir de un Data Warehouse principal, seleccionando el subconjunto necesario y trasladando una réplica de éstos de manera resumida, alcanzando así un mejor rendimiento ya que esos datos anteriormente pasaron por un proceso ETL. Este caso se conoce como **Top Down**.

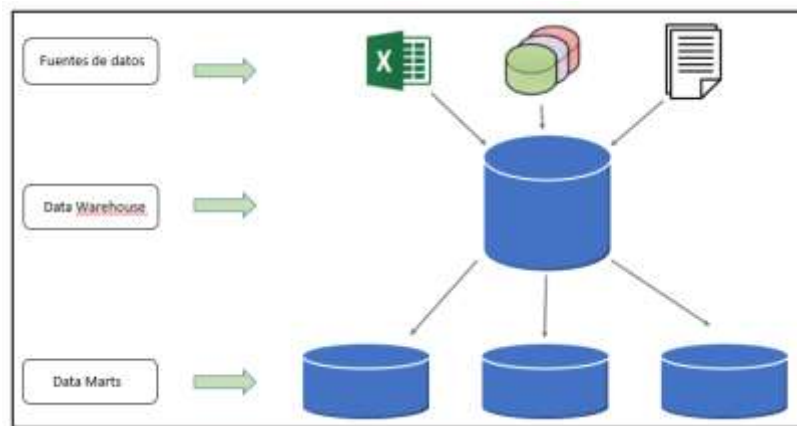


Ilustración 6. Proceso Data mart dependiente

- **Data mart independiente**, creado a partir de fuentes externas y sin depender de un Data Warehouse principal. Se necesita realizar un proceso de ETL aunque este será menor porque la cantidad de datos así lo es. Es beneficioso si necesitamos resultados rápidos, sin embargo, tiene menos capacidad de escalabilidad y la duplicación de datos es inevitable. Tras esto si es requerido por la empresa se incorporan los datos a un Data Warehouse. Este enfoque se conoce como **Bottom Up**.

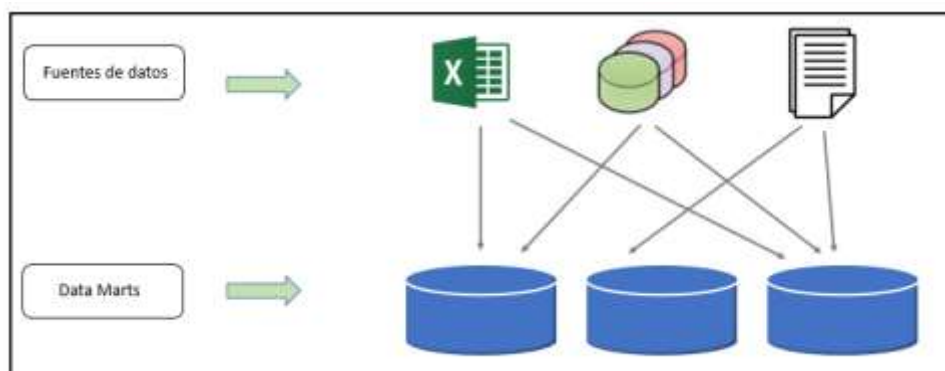


Ilustración 7. Proceso Data mart Independiente Fuente: Elaboración Propia

- **Data mart híbrido**, permite combina las ventajas de los dos anteriores y es útil cuando se necesita una integración **ad hoc**.

3.3.1. Diseño

Al momento de diseñar el Data Warehouse o el Data mart es importante considerar cual es el esquema más adecuado para lograr obtener los resultados requeridos. Existen dos tipos de modelado dimensional, aunque en ocasiones se opta por un modelo híbrido:

- **Esquema Estrella (star schema)**, en el que existe una tabla central, tabla de hechos (fact table), que contiene todas las medidas y tablas de dimensiones (dimension table), una por cada perspectiva que queremos analizar. Las tablas de dimensiones tienen un campo en común con la tabla de hechos lo que les permite la relación entre sí.

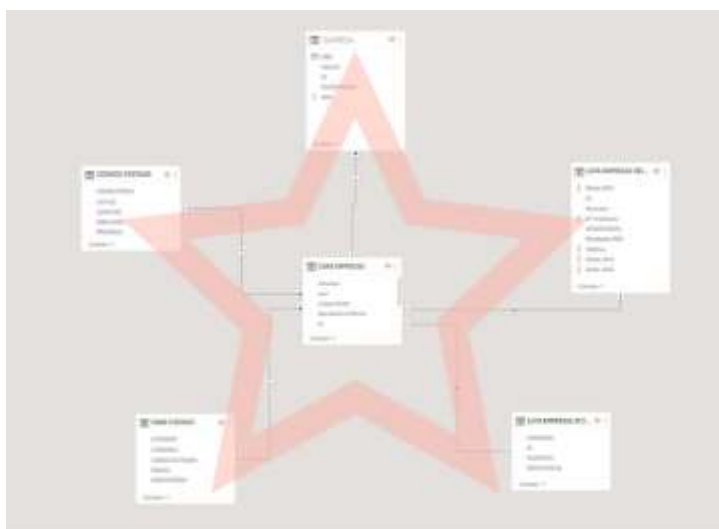


Ilustración 8. Esquema Estrella

Fuente: <https://www.autocontrolempresarial.es/2022/12/23/power-bi-y-las-relaciones-de-tablas/>

- **Esquema Copo de Nieve (snowflake schema)**, es más complejo que el anterior, aunque deriva de este y el más utilizado, su estructura se basa en que algunas de las tablas de dimensiones no están directamente relacionadas con la tabla de hechos si no que se relacionan con ésta a través de otras dimensiones.

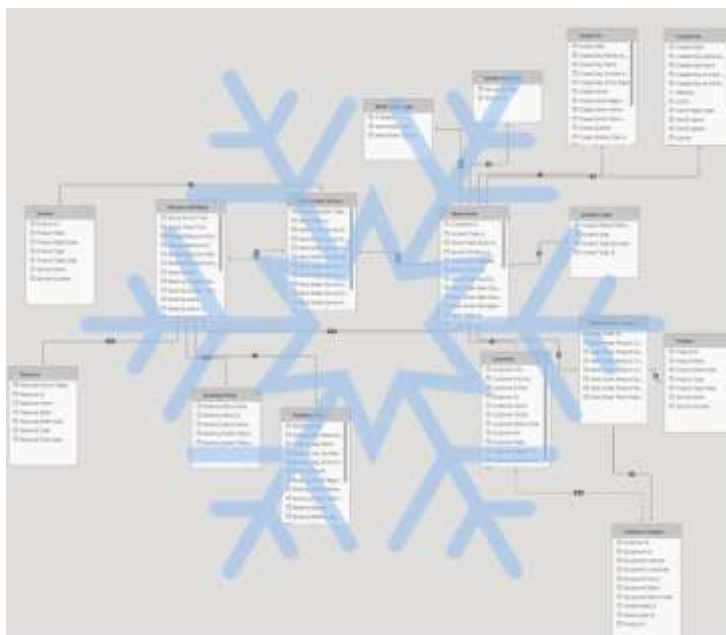


Ilustración 9. Esquema Copo de nieve

Fuente: <https://www.autocontrolempresarial.es/2022/12/23/power-bi-y-las-relaciones-de-tablas/>

Ambos modelos cuentan con sus pros y contras, mientras el modelo estrella es simple y veloz, el esquema copo de nieve evita la redundancia de los datos.

3.3.2. Data Warehouse vs Data mart

Más allá de la definición teórica existen similitudes y diferencias entre estos dos almacenamientos de datos.

Las principales similitudes son, ambos almacenan datos para su posterior análisis, son utilizados para la generación de informes e integran datos de múltiples fuentes. En cuanto a las diferencias, las principales son: el alcance, el uso, las áreas temáticas, las fuentes de datos, el tamaño y el tiempo de alimentación.

Mientras que un Data Warehouse tiene un alcance corporativo, un uso para decisiones estratégicas, abarca distintas áreas, sus fuentes de datos pueden llegar a ser muy numerosas, el tamaño puede superar el terabyte y el tiempo de implementación puede tardar de varios meses a años; el Data Mart está enfocado a una línea de negocio concreta, tiene un uso para decisiones tácticas, se alimenta de

pocas fuentes de datos, el tamaño puede estar entre en los centenares de gigabytes y el tiempo se reduce a algunos meses.

En esta tabla creada por la empresa Bismart, podemos ver más detalladamente las principales distinciones según diversos criterios.

	DATA WAREHOUSE	DATA MART
01. Uso	Facilita la integración de datos y el proceso de toma de decisiones empresariales.	Asiste en la toma de decisiones estratégicas concretas.
02. Objetivo	Proporciona un entorno integrado y coherente para todos los activos de datos de la empresa.	Proporciona un entorno integrado para los datos referentes a un departamento empresarial concreto.
03. Diseño	Proceso de diseño difícil. No tiene por qué estar basado en un modelo dimensional.	Proceso de diseño fácil basado en un modelo dimensional.
04. Enfoque	General. Información relativa a toda la compañía.	Específico. Información relativa a un departamento o área de negocio concreta.
05. Tipos de datos	Datos detallados de estructura no volátil y variantes en el tiempo.	Principalmente datos consolidados preparados para satisfacer las necesidades informativas de los responsables del área temática.
06. Alcance	Almacena datos vinculados a toda la empresa y a cualquier aspecto de la actividad empresarial. Ejerce de fuente de información para cualquier área de la organización.	Almacena datos relativos a un departamento o un aspecto de la actividad empresarial concreto. Cada departamento o área de negocio puede disponer de un data mart propio.
07. Fuentes de origen	Recopila datos procedentes de una gran cantidad de fuentes.	Recopila datos de una cantidad de fuentes reducida, normalmente del propio data warehouse.
08. Tamaño	Habitualmente de 100 GB a 1 TB.	Habitualmente menos de 100 GB.
09. Tiempo de implementación	Su implementación suele durar entre varios meses y varios años.	El tiempo de implementación suele ser de unos pocos meses.

Tabla 1. Comparativa Data Warehouse y Data mart

Fuente: <https://bismart.com/>

Es importante saber que ambos pueden coexistir y que en ocasiones se emplea un enfoque híbrido. La decisión de utilizar un Data mart en lugar de un Data Warehouse se basa en las necesidades específicas de la empresa. Algunas de las situaciones por las que se puede optar a utilizar un Data mart son cuando se necesita: un análisis y enfoque específico, un modelado de datos simplificado, una implementación rápida, agilidad y flexibilidad a la hora de adaptarse a los cambios, seguridad de datos y control de acceso ya que estos pueden diseñarse con unas medidas más específicas, rentabilidad y autonomía de la unidad de negocio.

3.3.3. Dos conceptos de almacenamiento de datos: Kimball vs Inmon.

Ambas metodologías tienen sus pros y sus contras, en este apartado indagaremos un poco más en ellas.

La metodología Kimball sigue un enfoque Bottom Up para el diseño de un almacén de datos en el que los Data marts se crean primero en función de los requisitos. Tras esto se evalúan las fuentes de datos y se crea un proceso ETL para cargar los datos en un área de preparación que generalmente no está normalizada. Este modelo divide los datos en la tabla de hechos y las tablas de dimensiones con las que posteriormente los usuarios crearan varios esquemas de estrella para satisfacer los diversos objetivos.

Para la integración de los datos sugiere utilizar dimensiones de datos consolidadas, las cuales serían una tabla base común entre las diversas tablas de hechos del almacén de datos para asegurar la consistencia en la utilización de un único elemento en todos los hechos. Un elemento clave en esta metodología es el bus empresarial de Kimball, el cual organiza verticalmente los hechos y horizontalmente las dimensiones conformadas.

La estrategia de Kimball para la vida útil de un almacén de datos también se centra en hechos confirmados, es decir, data marts que se desarrollan de manera independiente en conjunto con una arquitectura sólida.

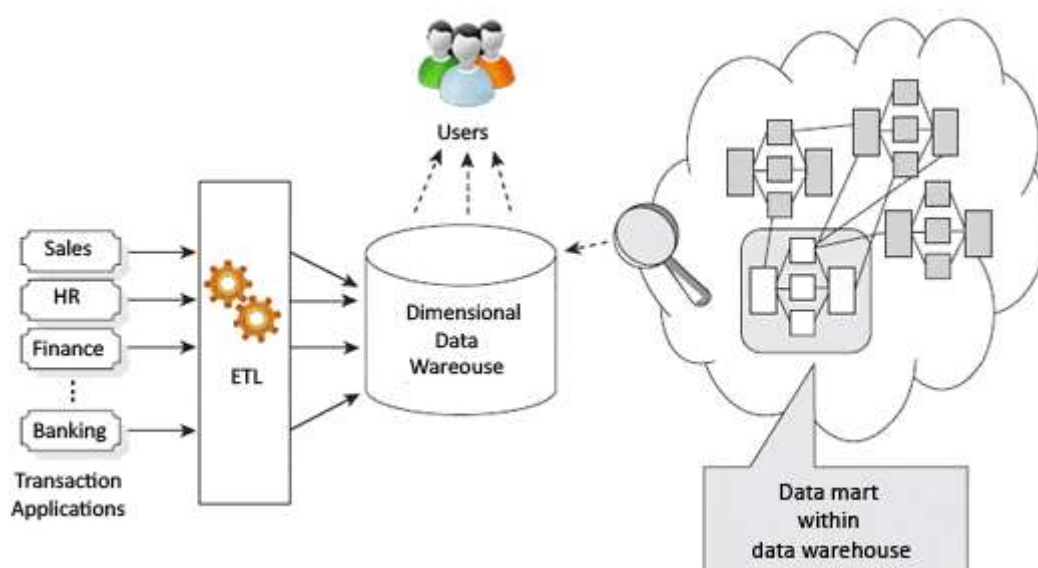


Ilustración 10. Arquitectura básica de Kimball

Fuente: <https://www.zentut.com/data-warehouse/ralph-kimball-data-warehouse-architecture/>

Las ventajas de este enfoque son, es rápido de construir ya que no implica normalización, utilizar el esquema en estrella fácilmente comprensible por la mayoría de usuarios, permite una rápida recuperación de los datos al estar divididos en tablas de hechos y dimensiones, no se necesita un gran equipo de diseñadores y tiene una estructura dimensional conformada para calidad de datos.

En cuanto a las desventajas, pueden surgir anomalías al actualizar los datos debido a la técnica de desnormalización, problemas de rendimiento debido a la agregación de columnas en la tabla de hechos, y complejidades al incorporar grandes volúmenes de datos.

La metodología Inmon, sigue un enfoque Top Down, y emplea la forma normalizada para crear la estructura de la entidad, evitando la redundancia de datos todo lo posible, lo que permite una identificación clara de los objetivos y prevee irregularidades a la hora de actualizar los datos. Se elabora un modelo lógico integral para cada entidad principal y una sola fuente de información veraz para todo el negocio. La simplificación de la carga de datos gracias a la normalización

contrasta con la dificultad de realizar consultas debido a la multitud de tablas y enlaces.

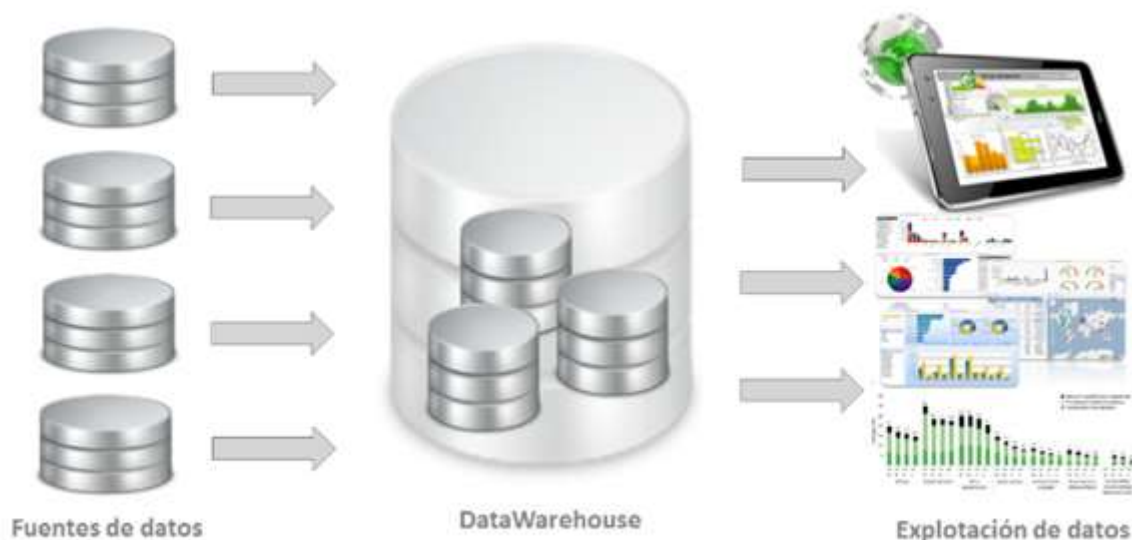


Ilustración 11. Arquitectura básica de Inmon

Fuente: <https://microsys.com.mx/blog/almacenes-de-datos-data-warehouse-y-data-mart/>

Las ventajas del enfoque incluyen la capacidad de abordar varios requisitos de informes en toda la empresa, ofrece mayor flexibilidad, simplifica los procesos, el almacén actúa como fuente unificada donde están todos los datos integrados, tiene una muy baja redundancia de datos lo que hace menos probable la aparición de irregularidades a la hora de actualizar los datos.

En cuanto a desventajas, aumenta la complejidad, se requieren recursos capacitados, una operación de proceso ETL adicional, mucho tiempo a la hora de la configuración y la entrega y sobre todo que se administre de manera efectiva.

Tras analizar ambos métodos diremos que escoger uno u otro va a depender de las necesidades que requiera la empresa. Si se necesitan informes integrados, tenemos un mayor plazo de proyecto, contamos con un equipo grande para el desarrollo, los cambios pueden ser frecuentes y se asumen los gastos elegiríamos el método de Bill Inmon. En cambio, si se necesitan informes centrados en un proceso, el plazo de entrega y el equipo de desarrollo son menores, los cambios poco

frecuentes, y solo se busca la mejora de los informes, elegiríamos el método Kimball.

3.3.4. OLAP

El procesamiento analítico en línea (OLAP), introducido por Edgar F. Codd en 1993, es la tecnología más común para analizar los datos almacenados en un Data Warehouse. A pesar de que sus conceptos básicos tienen raíces más profundas, OLAP se ha convertido en el estándar de la industria debido a su capacidad para gestionar grandes volúmenes de datos y permitir consultas complejas de forma eficiente. Su evolución constante ha dado lugar a una amplia gama de soluciones comerciales que incorporan motores OLAP.

“Se entiende por OLAP, o proceso analítico en línea, al método ágil y flexible para organizar datos, especialmente metadatos, sobre un objeto o jerarquía de objetos como en un sistema u organización multidimensional, y cuyo objetivo es recuperar y manipular datos y combinaciones de los mismos a través de consultas o incluso informes.” (Curto, J. 2007)

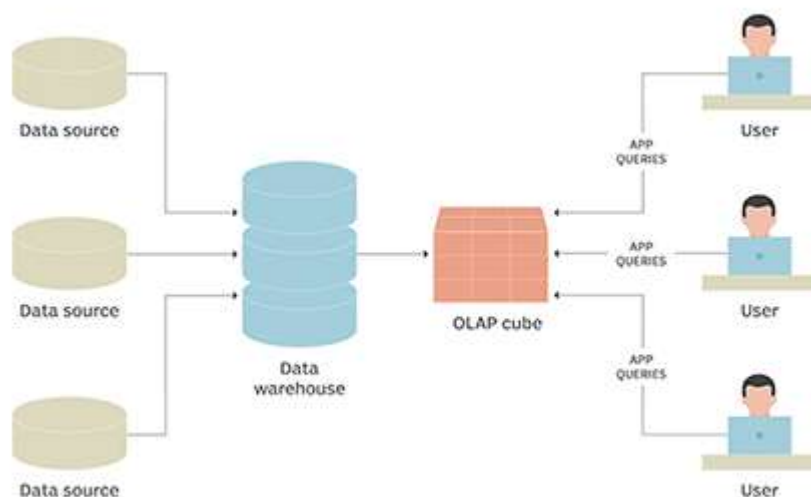


Ilustración 12. Proceso OLAP

Fuente: <https://www.computerweekly.com/es/definicion/OLAP-o-procesamiento-analitico-en-linea>

Una herramienta OLAP está formada por un motor que es lo definido en el párrafo anterior y un visor que es una interfaz que permite consultar, manipular, reordenar y filtrar datos existentes mediante consultas en lenguaje MDX².

La definición se basa en las 12 Reglas que propuso Codd (1993) creador del modelo para definir los requerimientos que ha de cumplir un sistema para ser considerado relacional y que Nigel Pendse (The OLAP Report) resumió en el concepto FASMI (Fast, Analysis, Share, Multidimensional, Information).

El propósito de esta tecnología es impulsar la consulta de grandes volúmenes de información, utilizando los cubos OLAP, estructuras multidimensionales con datos pre calculados y agregados. Teniendo una velocidad de respuesta muy superior a los sistemas OLTP.

Un cubo OLAP es un vector multidimensional, de N dimensiones (perspectivas desde donde analizar la información) en el que la información se almacena de forma ordenada y jerarquizada en cada una de estas dimensiones, contribuyendo de esta manera a llevar a cabo un rápido estudio de su contenido. Una base de datos multidimensional puede contener varios de estos cubos OLAP.

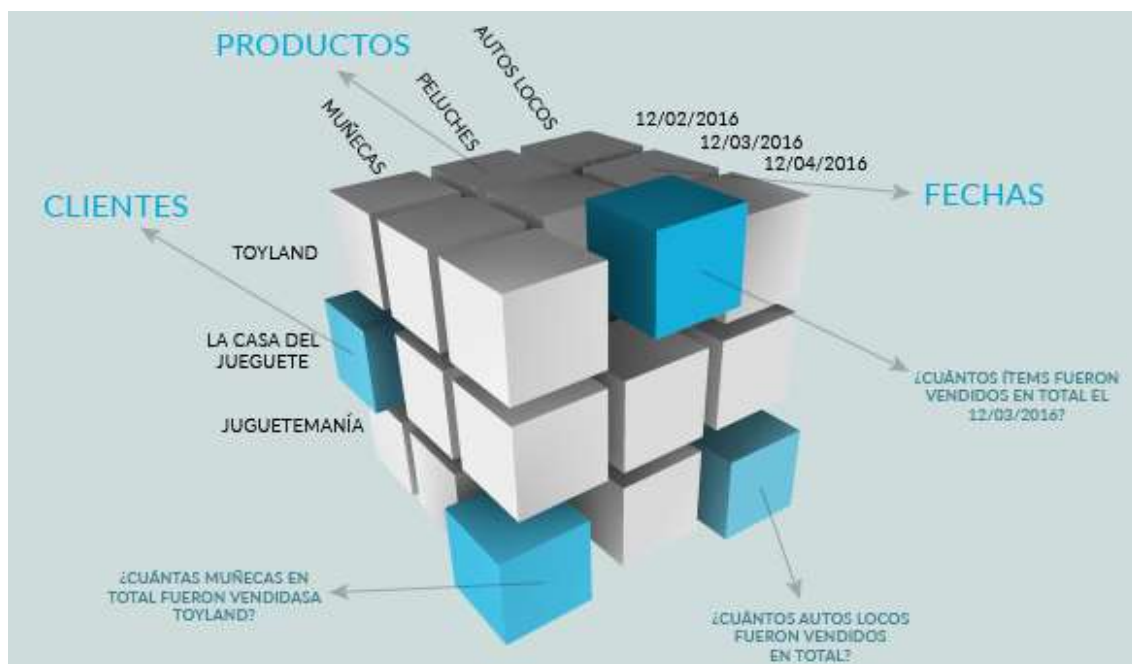


Ilustración 13. Cubo OLAP

² Lenguaje de consulta para soluciones de inteligencia empresarial

Fuente: <https://www.evaluandosoftware.com/>

Existen diferentes tipos de OLAP, diferenciados principalmente por la forma en la que guardan los datos. Son los siguientes:

- **MOLAP (Multidimensional OLAP)**, es lo que se conoce como BBDD multidimensionales.
- **ROLAP (Relational OLAP)**, trabaja directamente con BBDD relacionales.
- **HOLAP (Hybrid OLAP)**, los mayores volúmenes de datos se hayan en las bases de datos relacionales y las agregaciones en multidimensional.
- **DOLAP (Desktop OLAP)**, orientado a equipos de escritorio. La información requerida se extrae de la base de datos relacional y se almacena en el escritorio.
- **In-memory OLAP**, es un enfoque novedoso donde la estructura dimensional se crea únicamente en memoria y el dato original se almacena en un formato que favorece su despliegue.

Por último, decir que existen diferentes elementos comunes a estas tipologías como son: esquema, cubo, tabla de hecho, dimensión y métrica, jerarquía, nivel, miembro, roles y el lenguaje MDX.

3.4. Visualización de datos

Durante los últimos 30 años, el ámbito de la visualización de datos ha experimentado un crecimiento significativo en numerosas áreas de estudio. Los paneles y herramientas para buscar datos, junto con las aplicaciones de tableros de control y análisis, ofrecen a empresas y usuarios la posibilidad de explorar sus datos de manera creativa y novedosa.

Establecer herramientas de BI para empresas no solo consiste en extraer, organizar y almacenar los datos también hay que ofrecer al usuario informes interactivos e inmediatos con toda la información que necesita y cumpliendo con las

características requeridas y esto lo conseguimos a través de la visualización de datos una de las áreas más importantes en los proyectos y con un gran impacto ya que se trata de la presentación de todo el trabajo previo realizado. Si podemos ver algo, lo interiorizamos rápidamente, pero es importante saber transmitir la información que se ha generado a partir de los datos recibidos.

La visualización de datos es el proceso mediante el cual se transforman los números y las estadísticas en objetos visuales (tablas, tarjetas, gráficos, etc.), permitiendo entender de manera más sencilla los patrones, tendencias, relaciones y outliers presentes en conjuntos de datos complejos y ofreciendo diferentes perspectivas para analizar dichos datos a los usuarios que visualizaran el informe en tiempo real, lo que ayudará a la detección de problemas y hará que se puedan tomar las medidas oportunas rápidamente.

La visualización de datos es el proceso mediante el cual se transforman los números (datos) y las estadísticas (información) en objetos visuales (tablas, tarjetas, gráficos, etc.), volviéndose así más fáciles de entender, posibilitando el acceso a gran cantidad de dicha información de manera dinámica, permitiendo un análisis y comprensión de ésta de forma rápida y sencilla e identificando patrones, tendencias, relaciones y outliers presentes en conjuntos de datos complejos, que ofrecen diferentes perspectivas a los usuarios que visualizan el informe en tiempo real, lo que ayudará a la detección de problemas y a la anticipación de cambios para poder ajustar estrategias de manera proactiva y tomar las medidas y decisiones oportunas de una forma más ágil, garantizando también una comunicación efectiva.

Existen tres componentes clave en la visualización de datos: la narrativa, fundamental a la hora de una representación eficaz de los datos, ya que establece el objetivo y el contexto guiando la presentación de la información para construir estrategias y tomar las mejores decisiones; los datos, la selección y análisis de éstos, junto con una limpieza efectiva para eliminar información incorrecta, son fundamentales ya que aseguran el uso de los datos correctos y los componentes visuales (tablas, gráficos, tarjetas, mapas, etc.) que son esenciales para poder representar dichos datos de forma atractiva.

Algunos conceptos fundamentales a tener en cuenta son: tener un objetivo claro, elegir el objeto visual correcto, pensar que menos es más (enfocarse en lo necesario), crear títulos claros y concisos que llamen la atención y usar los colores de forma inteligente y correcta, papel clave en el proceso ya que estos influyen en nuestro cerebro y un uso adecuado en la visualización ayudará a que el mensaje que queremos transmitir sea mucho más claro.

Los pasos a seguir de manera global para llevar a cabo una visualización de datos eficiente y precisa son: **establecer objetivos**, comprendiendo los datos en profundidad lo que nos permitirá hacer un buen análisis y preparación de estos, **saber cuáles son sus orígenes, explorarlos y limpiarlos** es clave para eliminar errores e información incorrecta y finalmente **construir el dashboard**, representando la información importante de manera visualmente atractiva.

Cada proyecto y cada dato necesitan una representación visual que refleje lo que se desea comunicar o la narrativa que se quiere compartir, y para ello disponemos de las herramientas de visualización, que nos permitirán realizar un análisis a gran profundidad sobre nuestro negocio y nos ayudarán a tomar decisiones adecuadas para avanzar y crecer.

3.4.1. Herramientas utilizadas

Como podemos observar en la imagen el líder diferenciado en el sector de la visualización es Microsoft (Power BI), seguido de Salesforce (Tableau). También podemos contemplar a más o menos la misma altura Qlik, Google y Oracle. Y en el lado superior izquierdo como desafiantes mencionamos Amazon Web Services, Domo y MicroStrategy. A destacar SAP en la parte inferior derecha como visionarios.



Ilustración 14. Magix Quadrant for Analytics and Business Intelligence Platforms

Fuente: <https://powerbi.microsoft.com/en-us/blog/microsoft-named-a-leader-in-the-2024-gartner-magic-quadrant-for-analytics-and-bi-platforms/>

Para saber en base a que está estructurada la Ilustración 14, tenemos la siguiente imagen.

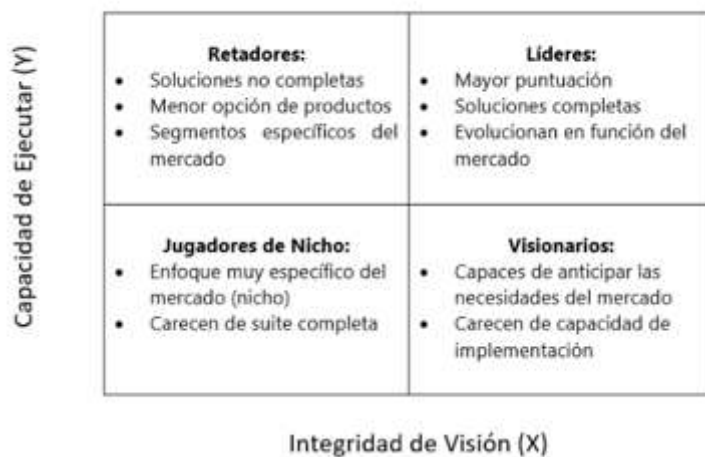


Ilustración 15. Matriz cuadrante de Gartner

Son muchas las herramientas utilizadas para la visualización de datos, a continuación, en esta sección describiré algunas de las más utilizadas, similitudes y diferencias.



Ilustración 16. Logos de herramientas de visualización de datos

Power BI

“Power BI es una colección de servicios de software, aplicaciones y conectores que funcionan conjuntamente para convertir orígenes de datos sin relación entre sí en información coherente, interactiva y atractiva visualmente”. (Microsoft, s.f.).

Dispone de capacidades de análisis, seguridad y cumplimiento avanzadas, una interfaz fácil de usar e intuitiva, actualizaciones que se producen regularmente mejorando la versión con nuevas funciones y características, se integra fácilmente con herramientas de Microsoft Office, permite amplia conexión a fuentes de datos, trabajo con IA y seguimiento en tiempo real junto con una gran variedad de visualizaciones interactivas.

Lo más destacable de las desventajas es que al tener límite de datos, puede ralentizarse cuando se manejan cantidades excesivas de estos, es difícil de migrar fuera del ecosistema de Microsoft.

Tableau

Tableau es una plataforma de análisis visual que permite analizar, visualizar y compartir grandes volúmenes de información de forma rápida, flexible y amigable. Ofrece facilidad al momento de llevar a cabo proyectos complejos, ya que puede operar en un servidor, en un entorno local o en la nube. Está diseñado para que individuos de cualquier sector empresarial puedan gestionar y analizar la información de forma sencilla. Transforma el modo en que aprovechamos los datos para solucionar problemas, posibilitando que las personas y las empresas obtengan el mayor beneficio de la información que poseen.

Entre las características destacadas de esta solución se encuentran: la capacidad de integrar datos procedentes de múltiples fuentes, la generación de visualizaciones interactivas que permiten una exploración profunda de los datos, el análisis en tiempo real con funcionalidades avanzadas para filtrar y segmentar la información, la creación de paneles personalizados y la posibilidad de compartir insights con otros usuarios.

En cuanto a ventajas podríamos destacar: la facilidad de uso (Drag and Drop)³, una visualización avanzada, permitiendo a los usuarios crear visualizaciones altamente personalizadas e interactivas para explorar y presentar datos de manera efectiva, integración con otras herramientas y plataformas (servicios cloud) lo que facilita esa integración, sin olvidar la eficacia, seguridad, escalabilidad y confiabilidad que presenta.

En lo relativo a las desventajas: puede resultar costoso para algunas empresas, aunque es fácil de usar puede llevar tiempo la familiarización con todas las características y funcionalidades y es una herramienta basada en la nube, que no tiene capacidad para ETL y transformación de Datos.

Qlik

Es una de las herramientas más competitivas dentro del Business Intelligence, capaz de proporcionar un exhaustivo análisis a los usuarios en el cual puedan basar

³ Arrastrar y soltar, técnica que favorece la interacción persona- ordenador.

sus decisiones, las cuales conduzcan a mejores resultados. El modelo asociativo que tiene ayuda a los usuarios a explorar los datos de manera más intuitiva ya que al clicar sobre cualquier dato pueden ver como se relaciona con otros facilitando un análisis más profundo.

En lo relativo a las características y ventajas, tiene implementada la automatización del proceso de la preparación de datos lo que aumenta la productividad ya que se puede dedicar más tiempo al resto de procesos, es capaz de gestionar grandes volúmenes de datos y almacenarlos en memoria lo que le permite una rápida recuperación y análisis, el software permite la integración perfecta de diversos orígenes de datos. Posee una IA potente y una plataforma en la nube híbrida.

En cuanto a las contras, necesita una infraestructura de servidor exclusiva para despliegues a gran escala, la curva de aprendizaje es elevada, puede resultar complejo aprender y dominarla especialmente si no estás familiarizado con herramientas de visualización, tiene un coste inicial alto y puede llegar a ser barrera para algunas organizaciones, los paneles de control complejos pueden llegar a requerir una alta potencia de procesamiento lo que ralentiza el rendimiento.

DOMO

DOMO es una compañía de software en la nube especializada en herramientas de inteligencia empresarial y visualización de datos. Sus funciones son integrar, analizar y aprovechar los datos. También posee un análisis de autoservicio.

En lo referente a características destacables, combina integración de datos, visualización, creación de aplicaciones, seguridad y control, exploración de datos a través del chat, una flexible creación de modelos junto con una gestión eficiente, creación de paneles de control personalizables, interactivos y en tiempo real y al estar creada en la nube, se conecta fácilmente a datos de otros sistemas y data warehouses de esta, a través de su API permite enviar notificaciones proactivas “push”. Automatiza tareas repetitivas relacionadas con los datos y así deja tiempo para centrarse en la estrategia y toma de decisiones. Dispone de una potente aplicación de móvil.

Con respecto a las desventajas, tiene precios elevados que para la entrada de clientes puede resultar una barrera reseñable, la curva de aprendizaje para nuevos usuarios es significativamente compleja y depende de la conectividad constante a Internet.

MicroStrategy

MicroStrategy quizás sea el más antiguo, pionera en el espacio de inteligencia comercial y analíticas, creada en origen para analizar y dar forma a los datos relacionales, ofrece un software innovador que fusiona la inteligencia artificial con la empresarial para optimizar la gestión de negocios, guarda, evalúa y genera informes utilizando monitoreo y predicciones en tiempo real a través de aprendizaje automático, PNL y análisis de IA. Las ventajas de implementar esta herramienta son, capacidad de análisis predictivo, toma de decisiones mejorada, gestión e integración de datos simplificada, acceso rápido si estamos conectados a internet desde cualquier dispositivo e implementación flexible. En cuanto a desventajas, sus opciones de exportación son limitadas y la plataforma puede resultar compleja de administrar.

Looker

“Looker Studio es una herramienta sin coste que convierte tus datos en informes y paneles claros, totalmente personalizables y fáciles de consultar y compartir”. (Google, s.f.).

Looker complementa sus capacidades de análisis de datos con un potente motor de visualización, permitiendo a los usuarios generar informes y paneles interactivos personalizados para comunicar de manera efectiva los insights obtenidos.

Permite la exploración de datos mediante filtros, agrupaciones y métricas personalizadas, lo que simplifica la detección de patrones y tendencias, ofrece herramientas avanzadas de modelado de datos, creación, colaboración y compartición de informes y paneles interactivos, una interfaz web sencilla, una gran variedad de plantillas de informes, una única fuente centralizada para todos los

datos e integración con herramientas como Slack, Salesforce, Google Analytics, Amazon Redshift y Google Workspace, su API permite a esta última automatizar la gestión y migración de los recursos. Tiene gran capacidad de escalabilidad y puede procesar conjuntos de datos de cualquier tamaño, ya sean BBDD pequeñas locales o grandes almacenes en la nube.

Podemos decir que, en cuanto a desventajas, depende de internet ya que está basada en la nube, y aunque tiene facilidad de uso puede llevar tiempo que los usuarios se familiaricen con sus funcionalidades y características. Posee un soporte limitado para el análisis de datos en tiempo real.

Google Data Studio

Google Data Studio es una plataforma gratuita de visualización de datos que facilita la creación de informes y paneles de control interactivos a partir de múltiples fuentes de datos. Facilita la exploración y comunicación de insights a través de visualizaciones personalizadas gracias a su interfaz intuitiva y a sus potentes capacidades de visualización.

Entre sus capacidades podemos destacar, una conexión de datos versátil, que facilita la integración de múltiples conjuntos de datos en un solo informe, creación de informes interactivos junto con una personalización avanzada de los mismos y colaboración y compartición de paneles en tiempo real facilitando así la distribución de información de manera eficiente. Posee integración con el ecosistema de Google, ofrece una interfaz intuitiva y fácil de usar que no requiere conocimientos técnicos avanzados y se actualiza regularmente con nuevas características y mejoras.

Como puntos débiles nos encontramos la dependencia de la conectividad a Internet, la creación de visualizaciones avanzadas tiene algo de complejidad si la comparamos con otras herramientas y existen algunos límites en la integración con fuentes de datos locales y privadas.

Características	Power BI	Tableau	Qlik	Domo	Microstrateg y	Looker	Google Data Studio
Facilidad de uso	Muy intuitivo, fácil de aprender	Interfaz amigable con curva de aprendizaje	Interfaz mas compleja, requiere tiempo	Fácil de usar pero con limitaciones	Curva de aprendizaje empinada	Requiere algo de formación	Muy fácil de usar y accesible
Integración de datos	Amplia variedad de conectores	Muchas integraciones disponibles	Buenas integraciones pero mas limitadas	Integraciones sólidas	Buenas integraciones, especialmente con datos empresariales	Integraciones fuertes enfocadas en SQL	Integraciones Con Google y otras APIs
Visualizaciones	Gran variedad	Excelente calidad gráfica	Potente en gráficos y visualizaciones	Visualizaciones dinámicas y atractivas	Muy personalizables y complejas	Focalizado en dashboard y métricas	Opciones básicas, fáciles de crear
Capacidades analíticas	Análisis de datos en tiempo real	Análisis avanzado y profundizaciones	Análisis asociativo	Analítica en tiempo real	Análisis avanzado y predicciones	Modelado de datos por capas	Análisis básico
Coste	Precios asequibles	Precios variados	Precios altos	Relativamente altos	Alto para empresas	Altos para grandes empresas	Gratuito con limitaciones
Despliegue	Nube y local	Cloud y On-Premise	Cloud, On-Premise	Basado en la nube	Cloud, On-Premise	Mayormente en la nube	Basado en la nube
Actualizaciones							
Escalabilidad	Escalable para grandes empresas	Escalable para empresas medianas y grandes	Escalable, ideal para grandes datos	Alta escalabilidad	Escalable y robusto	Muy escalable	Escalabilidad básica
Colaboración	Fuerte con Microsoft Teams	Funciones de colaboración efectivas	Herramientas de colaboración limitadas	Orientado a la colaboración	Herramientas de colaboración	Colaboración integrada	Funciones básicas
Seguridad	Seguridad empresarial incluida	Excelente manejo de permisos	Buenas funciones de seguridad	Enfoque en seguridad corporativa	Seguridad a nivel empresarial	Alta seguridad y control de accesos	Seguridad básica

Tabla 2. Comparativa de las herramientas de visualización

Fuente: Elaboración Propia

4. POWER BI

De todas las mencionadas anteriormente, en este apartado nos centraremos en la herramienta de visualización líder en el mercado y uno de los productos de mayor aceptación, Power BI.

4.1. ¿Qué es Power BI?

“Es un servicio de análisis de datos de Microsoft orientado a proporcionar visualizaciones interactivas y capacidades de inteligencia empresarial (en inglés: business intelligence o BI) con una interfaz lo suficientemente simple como para que los usuarios finales puedan crear por sí mismos sus propios informes y paneles.

Fue concebida originalmente por Thierry D'Hers y Amir Netz del equipo de SQL Server Reporting Services en Microsoft. Diseñado originalmente por Ron George y nombrado Proyecto Crescent. Más tarde renombrado a Power BI, Microsoft lo dio a

conocer en septiembre de 2013 como *Power BI para Office 365*. La primera versión de *Power BI* se basó en complementos de *Microsoft Excel*: *Power Query*⁴, *Power Pivot*⁵ y *Power View*⁶. Con el tiempo, *Microsoft* también agregó muchas características adicionales como preguntas y respuestas, conectividad de datos de nivel empresarial y opciones de seguridad a través de las puertas de enlace de *Power BI*". ("Power BI",2024)



Ilustración 17. Proceso de desarrollo con Power BI

Fuente: Elaboración propia

Power BI comprende múltiples componentes que trabajan en conjunto y forman el ecosistema de esta herramienta (Wikipedia s.f.):

- **Power BI Desktop**, aplicación de Windows para escritorio, principalmente para diseñar informes.
- Servicio de software como servicio (SaaS) en línea denominado **Power BI Service**, en el que podemos ver los informes creados en página web, se pueden crear y editar informes, pero es más limitado. Una característica muy interesante que ofrece es manejar alertas.
- Las aplicaciones **Power BI Mobile**, disponibles para los sistemas operativos Windows, iOS, y Android.

⁴ Tecnología de visualización de datos para la creación de objetos visuales

⁵ Tecnología de modelado de datos que permite crear modelos, establecer relaciones y crear cálculos

⁶ Tecnología utilizada para establecer conexiones con distintas fuentes, pudiendo extraer, transformar y cargar datos desde estas.

- Puertas de enlace, **Power BI Gateway**, se utilizan para sincronizar datos externos tanto dentro como fuera de Power BI.
- **Power BI Embedded**, la REST API⁷, usada para la creación de informes en las aplicaciones personalizadas que sirven para los que son usuarios y los que no de Power BI
- **Power BI Report Server** es un servidor local donde se pueden publicar los informes creados en Power BI Desktop.
- **Power BI Visuals MarketPlace**. Un mercado donde puedes encontrar objetos visuales personalizados que ayudan a aumentar la expresividad ya que ofrecen ventajas que no tienen los que vienen dados por defecto.
- **Power BI Report Builder** es utilizado para generar informes con paginación y luego distribuirlos a través del servicio Power BI. Presenta un formato apto para una página y muestra todos los datos en una tabla.

4.2. Interfaz Power BI (Descubriendo Power BI)

Al abrir un informe en blanco nos encontramos con la siguiente interfaz. Un espacio central donde irían los objetos visuales.

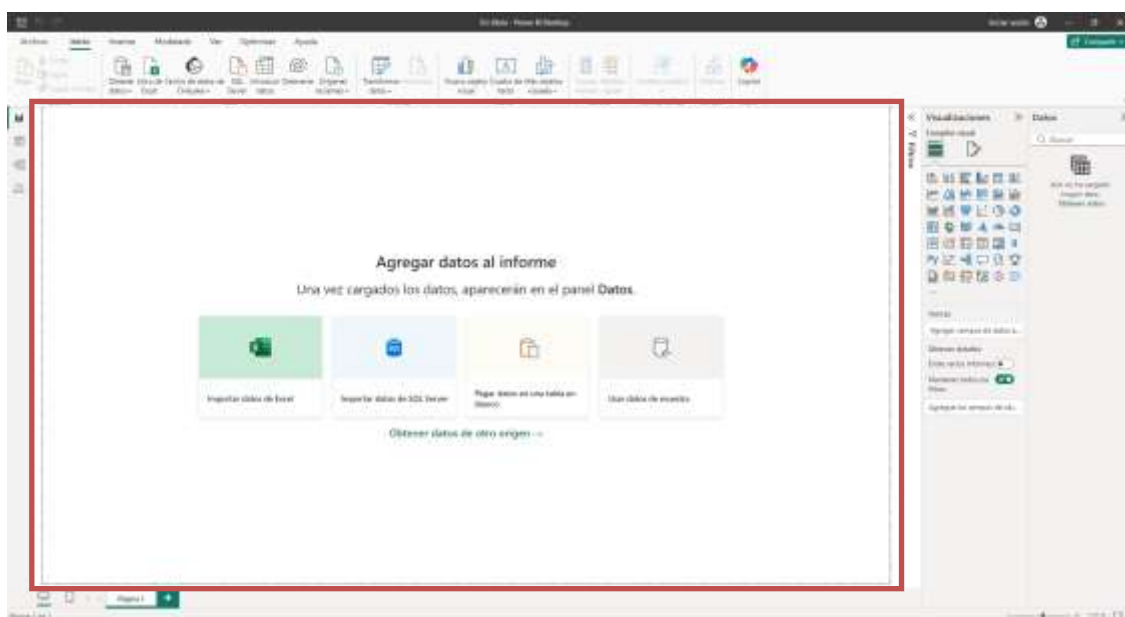


Ilustración 18. Power BI Desktop Principal

⁷ Interfaz de comunicación entre sistemas de información que usa el protocolo de transferencia de hipertexto

Un menú en la parte superior en el que podemos navegar entre, Archivo, Inicio, Insertar, Modelado, Ver, Optimizar, Ayuda, cada uno con sus diferentes opciones. Las más utilizadas durante el desarrollo (ver *Ilustración 19*, *Ilustración 20*, *Ilustración 21* e *Ilustración 22*)



Ilustración 19. Menú Inicio Power BI Desktop



Ilustración 20. Menú Insertar Power BI Desktop

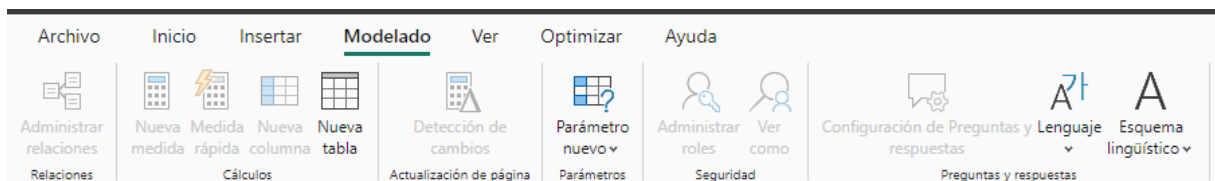


Ilustración 21. Menú Modelado Power BI Desktop



Ilustración 22. Menú Ver Power BI Desktop

De este último hacemos especial mención a la opción **Mostrar Paneles**, que es muy importante en el desarrollo de informes.

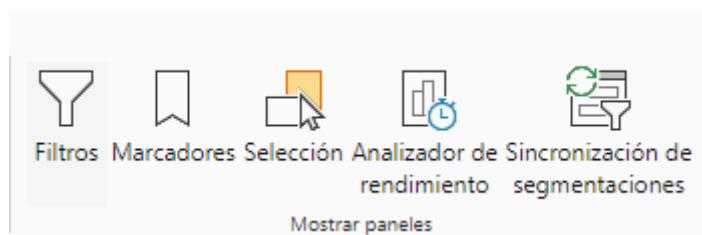


Ilustración 23. Opción Mostrar Paneles Power BI Desktop

Un menú lateral izquierdo compuesto por cuatro botones:

-  Vista de informe
-  Vista de tabla
-  Vista de modelo
-  Vista de consultas **DAX**⁸ (incluido en las últimas versiones)

Y en el lateral derecho, dos espacios fijos, Visualizaciones y Datos a los que se les pueden agregar otros como son Filtros, Marcadores, Selección, Analizador de Rendimiento y Sincronización de segmentaciones.

⁸ “Lenguaje” utilizado en Power BI, recopila funciones, operadores y constantes para devolver

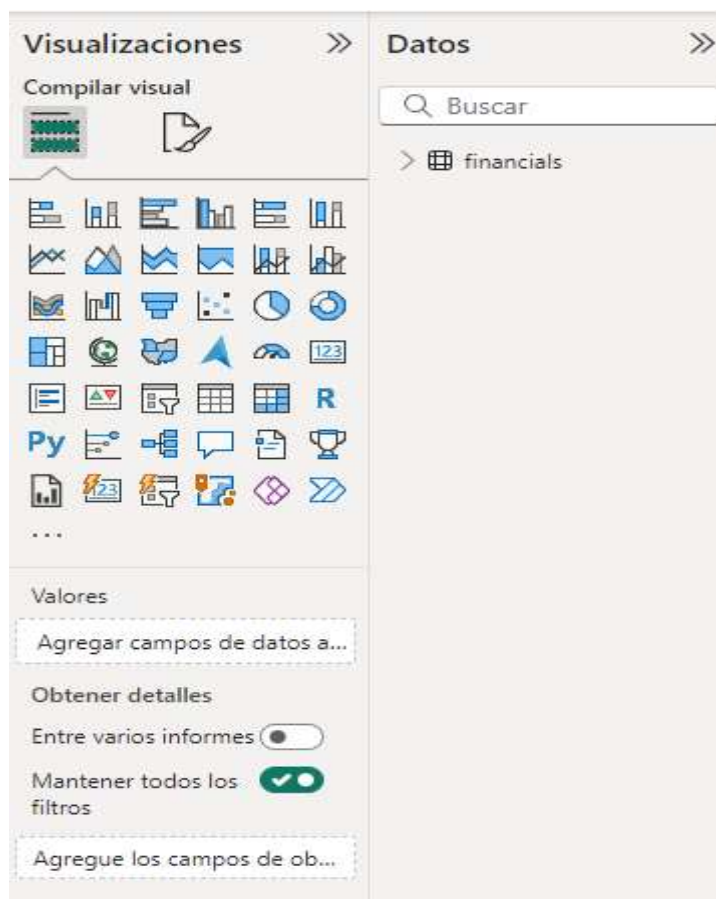


Ilustración 24. Panel lateral derecho Power BI Desktop

Para empezar con la carga de datos clicamos en este icono y si el origen está en la lista predeterminada simplemente pulsamos, si no tenemos la opción de Más que nos ofrece muchas más posibilidades.

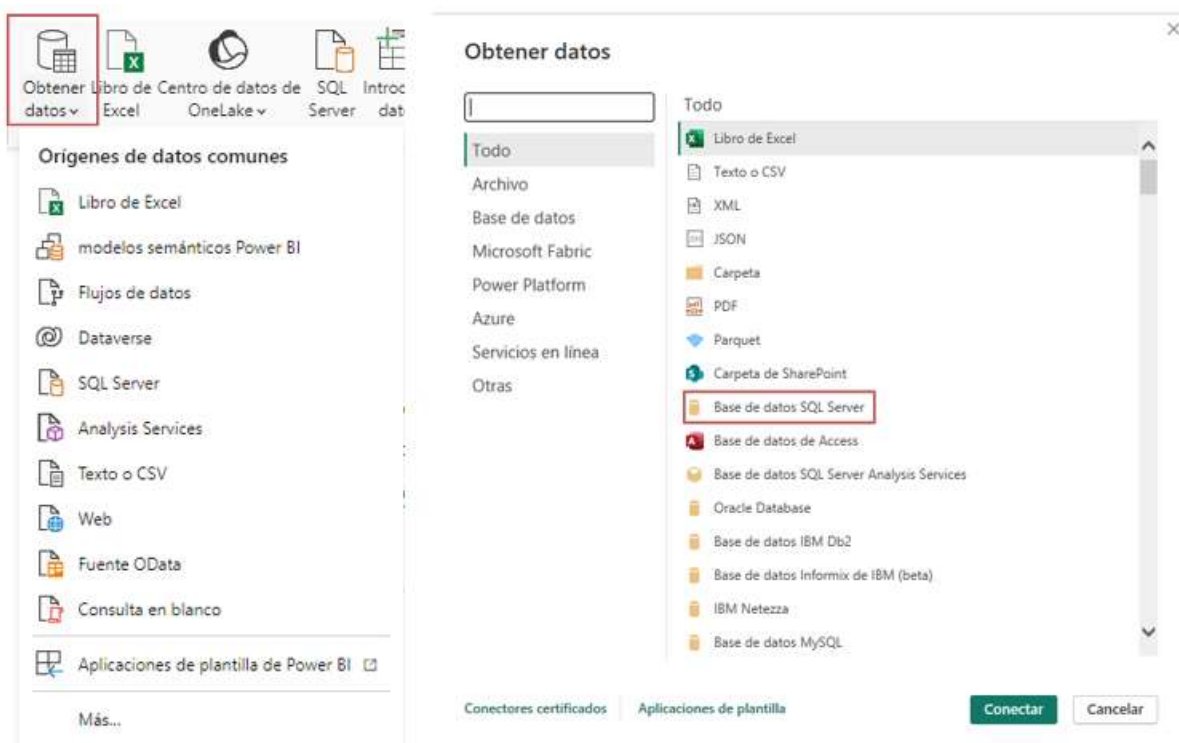
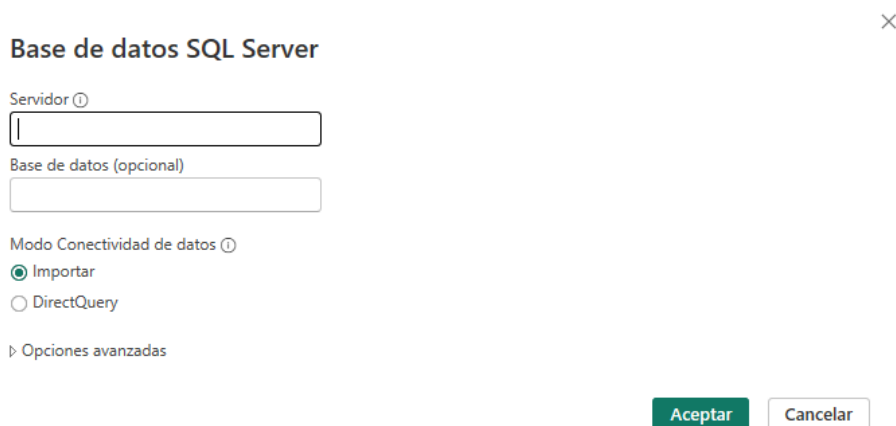


Ilustración 25. Obtención de datos Power BI Desktop

Una de las fuentes de datos más comunes utilizadas en Power BI son las bases de datos relaciones como la de SQL Server. Si clicamos en esa opción, por ejemplo, aunque también se da en las demás BBDD, obtendremos un cuadro como el de la Ilustración (), en el que podemos elegir el modo de conectividad con el que preferimos almacenar los datos de nuestras tablas.

- **Import Query:** extrae la información de la base de datos y la inserta dentro del Power BI, conjunto completo de transformaciones y manipulación de datos dentro del Desktop, limitación de 1GB y la información puede ser actualizada hasta 8x por día.

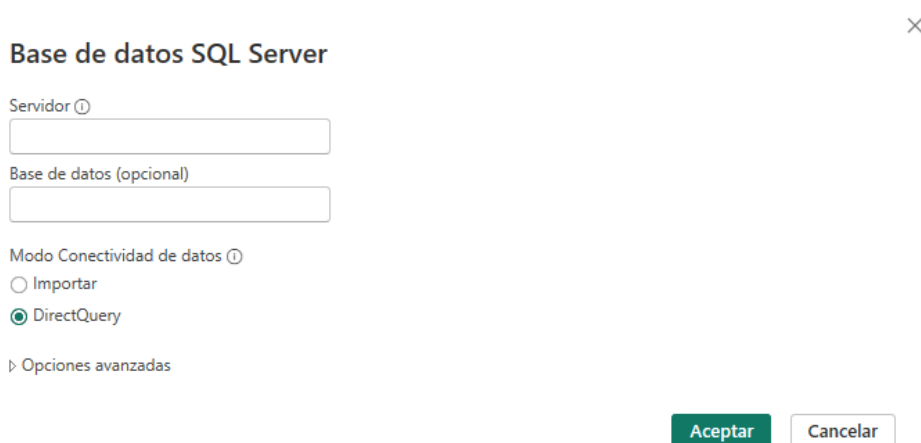


The screenshot shows a dialog box titled 'Base de datos SQL Server' with a close button (X) in the top right corner. It contains the following fields and options:

- Servidor** (with an information icon): A text input field.
- Base de datos (opcional)**: A text input field.
- Modo Conectividad de datos** (with an information icon): Two radio button options:
 - Importar**: Selected (indicated by a filled circle).
 - DirectQuery**: Not selected (indicated by an empty circle).
- Opciones avanzadas**: A link with a right-pointing arrow.
- At the bottom right, there are two buttons: **Aceptar** (green) and **Cancelar** (white with a grey border).

Ilustración 26. Import Query Power BI Desktop

- **Direct Query:** lanza queries para obtener los datos, limitación de 1GB, la información se actualiza en tiempo real, mantiene los datos en la BBDD de origen.



This screenshot is identical in layout to the previous one, but with the 'DirectQuery' radio button selected under the 'Modo Conectividad de datos' section.

Ilustración 27. Direct Query Power BI Desktop

Y como elección para ambas, las Opciones Avanzadas: limitan el tiempo de extracción de datos, optimiza el tamaño de las tablas e incluye relaciones entre columnas.

▲ Opciones avanzadas

Tiempo de espera del comando en minutos (opcional)

Instrucción SQL (opcional, requiere una base de datos)

☒ Incluir columnas de relación

☐ Navegar usando la jerarquía completa

Ilustración 28. Opciones Avanzadas Origen de datos Power BI Desktop

Podemos ver algunas de las diferencias resumidas en esta tabla.

Características	Import Query	Direct Query
Tamaño	Hasta 1GB por conjunto de datos	No limites
Fuente de datos	Importar datos de múltiples fuentes	Los datos deben venir de una sola fuente.
Performance	Motor de consultas de alto rendimiento	Depende de la respuesta de la fuente de datos para cada consulta.
Cambio de datos en los datos subyacentes	No reflejado. Es necesario realizar una actualización manual en el Desktop y volver a publicar el informe.	Power BI almacena en cache los datos para un mejor rendimiento, por lo que es necesario actualizar para asegurar los datos mas recientes.
Programar actualización	Máximo 8x por día	Programar cada 15 minutos
Puerta de enlace Power BI	Solo es necesario obtener los datos más recientes de fuentes de datos locales	Debe requerir obtener datos de fuentes de datos locales
Transformación de datos	Soporta todas las transformaciones	Soporta muchas transformaciones con algunas limitaciones

Tabla 3. Comparativa Import vs Direct Query

Fuente: Elaboración propia

Otra de las funciones más utilizadas e importantes es Transformar datos, clicando este botón se nos abre la pestaña de Power Query (cargar nuevos datos, editar consultas existentes, crear nuevas consultas etc).

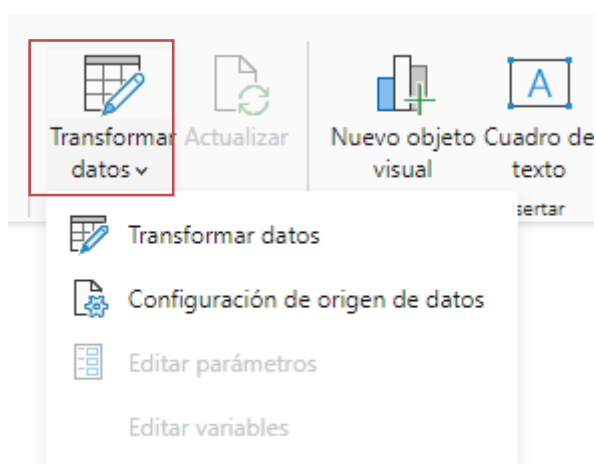


Ilustración 29. Opción para Transformar Datos Power BI Desktop

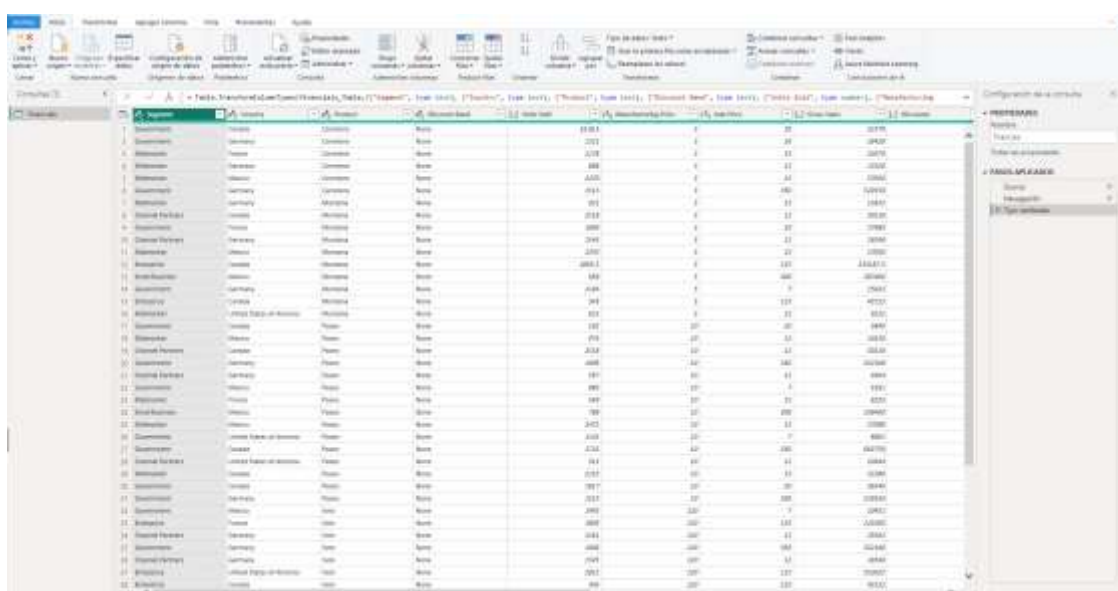


Ilustración 30. Power Query en Power BI Desktop

5. METODOLOGÍAS ÁGILES

En este apartado vamos a explicar que son las metodologías ágiles junto con una breve evolución, diferencias con las metodologías tradicionales, que tipos existen y algún ejemplo práctico de ellas.

5.1. ¿Qué son las metodologías ágiles?

Desde que las empresas se están viendo obligadas a una adaptación a la transformación digital hasta hoy, ha habido una serie de avances que han transformado la manera de trabajar y llevar a cabo numerosos procesos. Las metodologías ágiles para gestionar los proyectos están actualmente en uso entre las organizaciones que buscan una transformación completa y aprovechar sus ventajas.

La aparición de dichas metodologías comenzó a finales de la década de los 70 debido al estancamiento que experimentaron los desarrolladores de software y los problemas a los que tenían que hacer frente. En la llamada crisis del software se implementaron nuevas estrategias y metodologías que se vieron truncadas debido a la complejidad o singularidad de cada proyecto y la dificultad de realizar estimaciones precisas y que contuvieran el menor número de errores.

Fue en 1986, Ikujiro Nonaka y Hirotaka Takeuchi quienes elaboraron un concepto o una forma de trabajo parecida a la mentalidad con la que se juega en rugby (scrum). Y no fue hasta 1995 que esta idea se hace oficial en la industria del software. La estrategia era la adaptación a los cambios que sufrían los proyectos sin tener que descartarlos si no tenidos en cuenta como algo que podía suceder.

En 2001, aparece el [Manifiesto Agile](#), como una forma de documentar la manera de desarrollar proyectos dentro de la industria, y su impacto fue tan significativo que se convirtió en el método de trabajo conocido como agile.

“La metodología ágil no hace referencia a una serie de indicaciones sobre qué hacer exactamente durante el desarrollo de software. Se trata más bien de una forma de pensar en la colaboración y los flujos de trabajo, y define un conjunto de valores que guían nuestras decisiones con respecto a lo que hacemos y a la manera en que lo hacemos.” (Red Hat, 2022)

Por tanto, definiremos las metodologías ágiles como un grupo de técnicas y métodos utilizados en la creación de software y otros proyectos, destacando la

importancia de la adaptabilidad, la colaboración y el avance gradual y valorando la flexibilidad para responder a los cambios del entorno y las necesidades del cliente.

5.1.1. Metodología Ágil vs Metodología Tradicional

El enfoque tradicional (metodología en cascada) lleva usándose más de 70 años y generalmente se representa mediante un gráfico de **GANTT** que muestra todas las tareas, asignaciones de recursos o personas y las dependencias. Cuenta con una metodología establecida, en la que el proyecto pasa por diferentes fases: inicio, planificación, ejecución, seguimiento y cierre.

Se centra en la planificación proactiva y predictiva, donde factores como el costo, alcance y tiempo son importantes. Pretende anticiparse a los cambios y considera la elaboración de un plan completo en la etapa inicial del proyecto, lo que imposibilita cambiar los requisitos o necesidades, esto supone un costo y tiempo variables pero unos requisitos fijos, por lo que a menudo hay problemas de presupuesto y plazos.

En la vida real rara vez un proyecto sigue una secuencia lineal, cualquier error o problema de diseño identificado durante la etapa de prueba implica rediseñar y realizar una nueva implementación del código afectado, lo que conlleva un aumento en los costos y la imposibilidad de continuar debido a que no se puede llevar a cabo una etapa sin haber completado la anterior.

El enfoque ágil se reconoce que en su etapa inicial no se cuenta con suficiente información para realizar un plan detallado y se efectúa una entrega de software útil en lugar de una total. Se planifica en función de estimaciones realizadas por todo el equipo en base a la priorización de tareas con mayor valor para el cliente/usuario.

Por otro lado, en los proyectos ágiles se prioriza el trabajo en equipo, todos juntos elaboran un plan y monitorean el progreso del proyecto de principio a fin, jugando la transparencia un papel clave para mantener un entorno de trabajo productivo y altamente cohesivo. También destaca el nivel de flexibilidad a la hora de realizar cambios, ya que entre los miembros del equipo hay libertad para intercambiar ideas y ayudar así a mejorar aún más el producto.

Es un proceso iterativo, centrado en las necesidades de los clientes por lo que la colaboración con estos es fundamental en el proceso. Las versiones continuas en cada sprint del proyecto, ayudan al equipo a enfocarse en solucionar problemas o inconvenientes que puedan surgir más que en seguir un proceso riguroso.



Ilustración 31. Tradicional vs Ágil

Fuente: <https://todobi.com/>

Características	Metodología Ágil	Metodología Tradicional
Estructura organizativa	Iterativa	Lineal
Requisitos	Dinámicos	Bien definidos antes de empezar
Modelo de desarrollo	Entrega evolutiva	Ciclo de vida
Planificación	De Sprint en Sprint	Todo con gran detalle
Escala de proyectos	Pequeños y medios	Grandes
Estimación del esfuerzo	El Scrum Master facilita las tareas y el equipo hace estimación	El gestor del proyecto estima y obtiene la aprobación del jefe de proyecto.
Preferencias del modelo	Favorece la adaptación	Favorece la anticipación
Gestión de escalado	Cuando ocurren problemas todo el mundo trabaja junto para resolverlos	El problema se escala a los gerentes del proyecto
Implicación del cliente	Alta	Baja
Producto o proceso	Menos enfoque en los procesos formales y directivos	Más enfocado sobre procesos que sobre el producto
Revisiones y aprobaciones	Después de cada iteración	Constantes por parte de los líderes del proyecto

Tabla 4. Metodología Ágil vs Metodología Tradicional

Fuente: <https://www.escueladenegociosfeda.com/>

Resumiendo, podemos decir que cada enfoque tiene sus ventajas, pero que en términos generales las metodologías tradicionales se centran en la planificación, mientras que las ágiles ponen foco en la ejecución.

5.1.2. Tipos de Metodologías Ágiles

Los distintos tipos de metodologías comparten unos mismos principios solo hay que saber elegir el mejor método que se adapte al proyecto. Algunos de los más usados son los siguientes:

- **Metodología Lean**

Filosofía nacida en Japón en los años 80 de la mano del ingeniero de Toyota, Taiichi Ohno, quien se centró en conseguir aumentar la eficiencia en la fabricación de vehículos de la marca. Tiene como fin encontrar problemas en el método de trabajo y resolverlos, principalmente eliminando los desperdicios y basándose en la identificación de procesos que no aportan valor, funcionalidades no solicitadas, cambio de tareas, esperas (hasta tener la mayor cantidad posible de información, para hacer más acertada la decisión), entregas lo más rápidas posibles (los retrasos también son un desperdicio) y defecto en el código o en el despliegue lo que hace que el proceso se ralentice. Su máxima está en reducir la inversión, el tiempo y el esfuerzo.

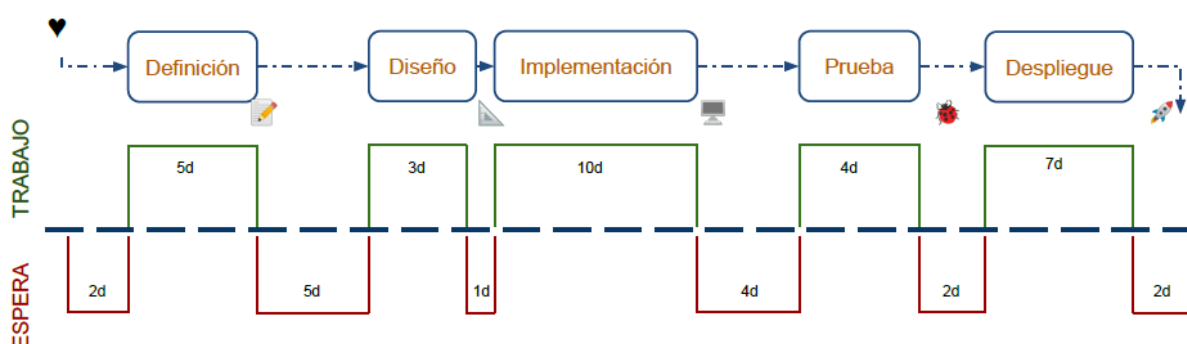


Ilustración 32. Metodología Lean

Fuente: Tema 4. Lean y Kanban. Rivas Santos V

- **Metodología Kanban**

Tiene su origen en el japonés y es un marco de trabajo visual, de hecho, su significado es “tarjetas visuales” y es utilizado en la gestión de proyectos y procesos con el fin de mejorar eficiencia y productividad, ya que permite visualizar el flujo de trabajo a simple vista y de una manera sencilla pudiendo ver en qué punto se encuentra el proceso y detectar posibles cuellos de botella. Fue desarrollado con la metodología de producción “just in time”, donde se produce justo lo necesario y en el momento adecuado. Es muy útil en proyectos donde los cambios suelen aparecer con facilidad.

El método se basa en un tablero donde los integrantes del equipo pueden observar el estado de cada tarea y el progreso del proyecto. Cada tarea se representa en tarjetas o post-it y se van moviendo a través de tres etapas: tareas por hacer, tareas en progreso y tareas realizadas. Implementa la filosofía Lean.



Ilustración 33. Metodología Kanban

Fuente: <https://www.inesdi.com/>

- **Metodología Extreme Programming**

Es un método de desarrollo de software creado por Kent Beck en la década de los noventa y que junto a su amigo Ron Jeffries empezó a enseñar a los miembros de su equipo cuando fue contratado para dirigir el equipo del Sistema de Compensación Integral de Chrysler. En 1999 formalizó las prácticas, principios y

valores en las que se basa esta tecnología en su libro “Extreme Programming Explained: Embrace Change.”

XP tiene como objetivo que equipos de desarrollo compuestos por un nivel pequeño o medio de personas produzcan software de alta calidad de forma constante y se adapten a los requisitos cambiantes y en evolución de una forma veloz y simple. Se basa en 5 valores (simplicidad, comunicación, comentarios, valentía, respeto), 5 reglas (planificación, gestión, diseño, codificación, prueba) y 12 prácticas de programación. La diferencia con las otras es que destaca en aspectos técnicos de desarrollo de software.

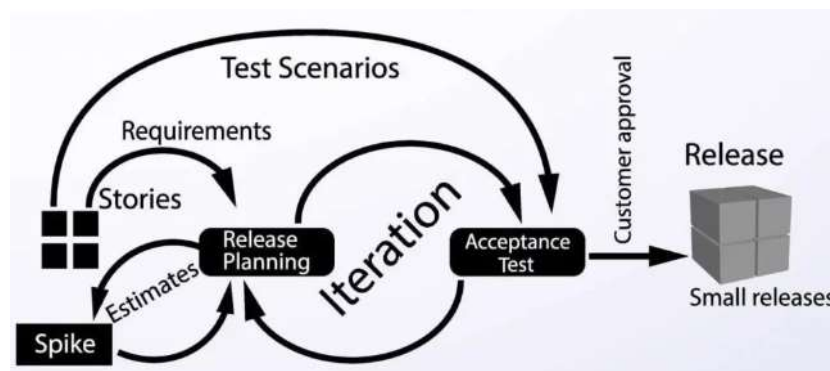


Ilustración 34. Metodología XP

Fuente: <https://www.nimblework.com/es/agile/programacion-extrema-xp/>

- **Metodología Scrum**

Es la más utilizada y un marco para el desarrollo ágil de software de referencia. Se centra más en el producto que pidió el cliente que en los formalismos necesarios para lograrlo, aplicando un conjunto de buenas prácticas, buscando problemas que impidan ser más productivos para trabajar en las soluciones. Scrum es ideal para proyectos de entornos complejos, requisitos cambiantes y poco definidos y donde priman la innovación, productividad, competitividad y flexibilidad y es crucial lograr resultados rápidos. Si a eso le sumamos la entrega continua de resultados, una retroalimentación temprana y frecuente donde los clientes pueden ver el progreso y dar su feedback y optimización del tiempo y recursos asegurando realizar un trabajo de la forma más eficiente posible, tendremos muchas de las ventajas y características de esta metodología.

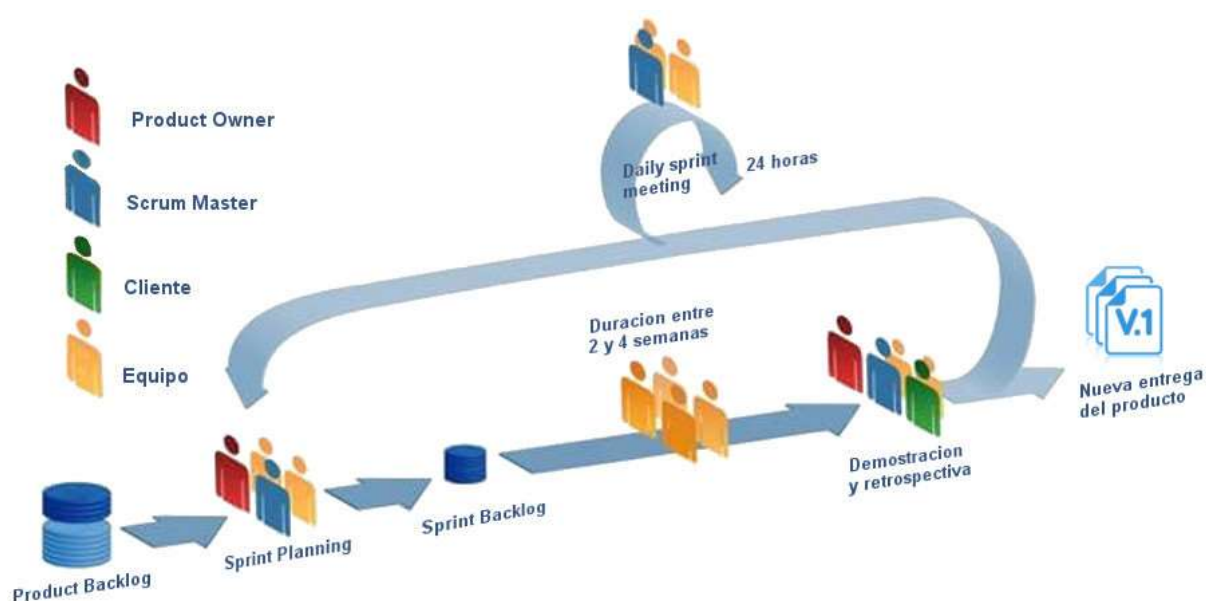


Ilustración 35. Metodología Scrum

Fuente: <https://welcomedevolvers.es/?s=scrum>

A continuación y observando la (Ilustración) procedo a explicar los roles que existen dentro del proceso: **Product Owner (PO)**, en representación de los **stakeholders** y responsable de asegurar la eficiencia del trabajo en el equipo, priorizar las tareas del Product Backlog y mantener comunicación constante con el Scrum Master y el equipo de desarrollo, **Scrum Master (SM)**, figura central encargado de los eventos del Scrum, mantenimiento del flujo de comunicación entre PO y el equipo y de ayudar en cualquier obstáculo que se pueda encontrar durante el desarrollo del proyecto, el **Equipo**, compuesto por profesionales que realizan el trabajo y se encargan de crear y entregar el producto mejorado en cada sprint colaborando entre ellos y con el PO y el **Cliente**, clave en el desarrollo, es prioridad ya que se busca su satisfacción y es quien necesita el producto.

Scrum utiliza un ciclo de desarrollo iterativo e incremental en el que la unidad básica de trabajo se denomina **sprint**, que va muy ligado al objetivo principal, rapidez en la entrega. Un sprint no es más que una parte del proyecto, que se asigna a un equipo, el cual se compromete a terminarla en un plazo concreto. Por lo que se llevarán a cabo varios sprint en el desarrollo de un mismo proyecto.

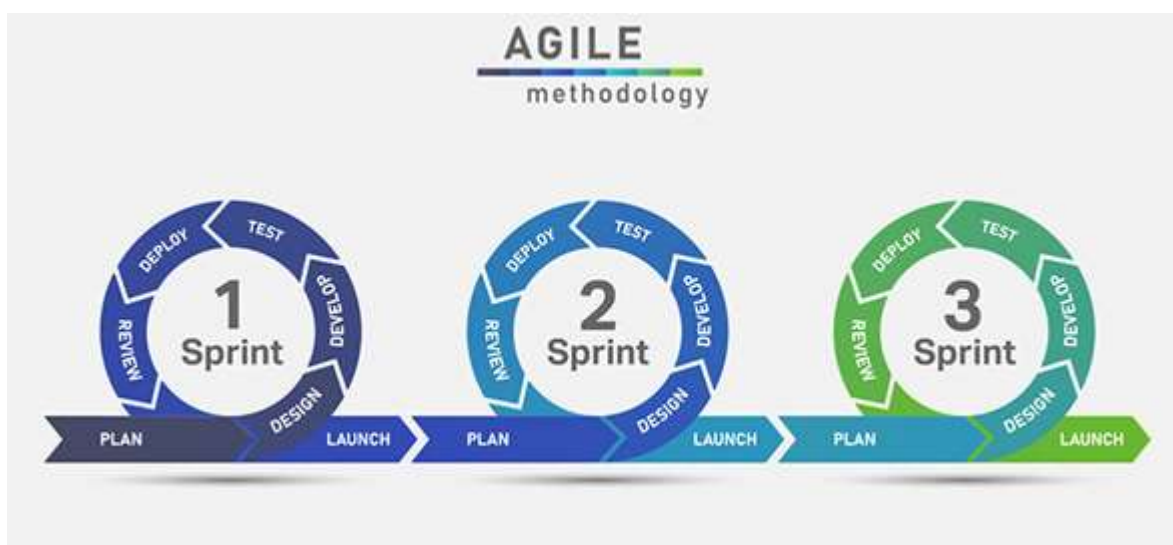


Ilustración 36. Metodología Ágil (Sprint)

Fuente: <https://www.inesdi.com/>

Por su parte, cada sprint se divide en tareas más pequeñas e independientes llamadas Scrum events y que son los siguientes:

- **Sprint Planning:** reunión donde se planifica el trabajo a realizar durante el próximo Sprint, seleccionando y priorizando tareas del product backlog.
- **Daily Sprint Meeting:** reunión diaria breve en la que el equipo conversa sobre los avances, los planes del día y posibles obstáculos.
- **Sprint Review:** al final del sprint, sesión en la que el equipo presenta el trabajo completado y recopila retroalimentación del product owner y otros participantes.
- **Sprint Retrospective:** reunión post-sprint para analizar lo ocurrido, identificar acierto, errores y planificar mejoras.
- **Sprint Backlog:** sesión para revisar, reorganizar y estimar el tiempo de las tareas en el product backlog.
- **Product Backlog:** lista en constante actualización de todas las tareas pendientes. No es un evento scrum, pero es un elemento fundamental en esta metodología.

Por último, pero no menos importante, ya que es parte indispensable del desarrollo, nos encontramos con el panel de tareas o tablero Scrum, donde se

muestran las tareas asignadas a cada miembro del proyecto y el estado en el que se encuentran. Algunas de las más utilizadas son Jira, ClickUp, Zoho Sprints, Miro, Parabol y por parte de Microsoft, Microsoft Planner.

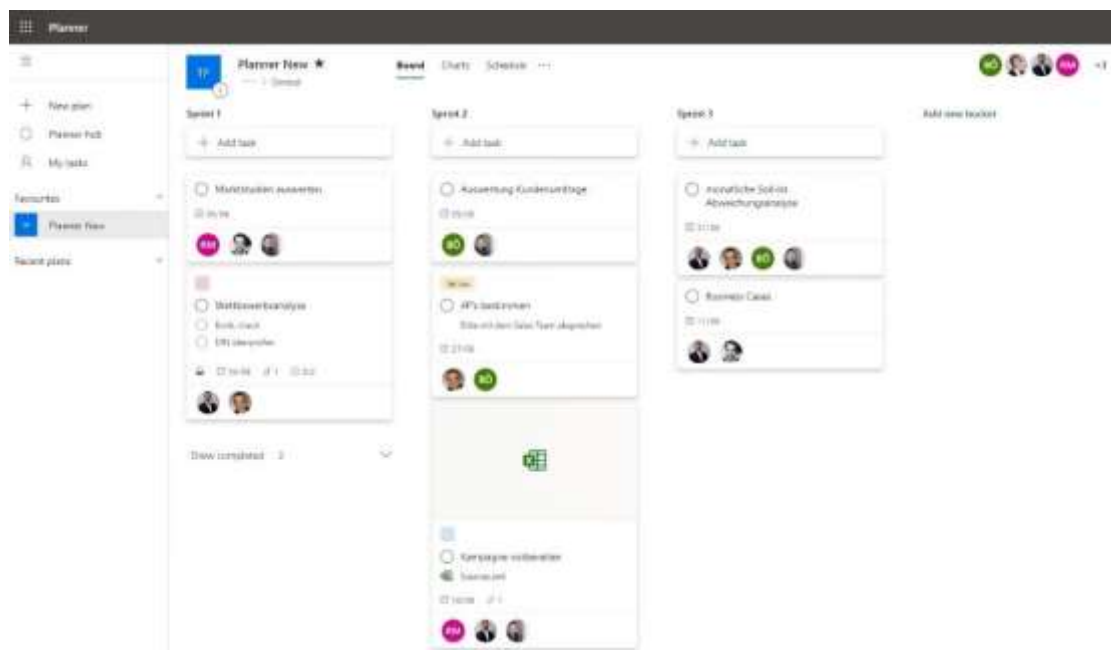


Ilustración 37. Microsoft Planner

Fuente: <https://www.microsoft.com/es-es/microsoft-365/business/task-management-software>

6. EJEMPLOS PRÁCTICOS CON POWER BI

En este apartado realizaremos unos ejemplos de casos de uso pedidos por usuarios y desarrollados a través de la herramienta Power BI. Los datos con los que realizaremos dichos ejemplos han sido extraídos de esta página web: <https://www.kaggle.com/datasets/bhanuthakurr/cleaned-contoso-dataset>. Existen 25 tablas diferentes y hemos escogido las 9 siguientes: FactSales.csv, DimCalendar.csv, DimChannel.csv, DimGeography.csv, DimProduct.csv, DimProductCategory.csv, DimProductSubcategory.csv, DimPromotion.csv, DimStores.csv.

Lo primero será insertar las fuentes de datos en Power BI, mediante la opción (Menú Inicio→Nuevo origen→Texto o CSV). Una vez cargadas, nos dirigimos a la pestaña del modelo para ver si las relaciones se han creado correctamente, ya que

Power BI es capaz de detectar dichas relaciones automáticamente si existe la columna Key correspondiente, aunque esto no significa que no hay que revisar el modelo.

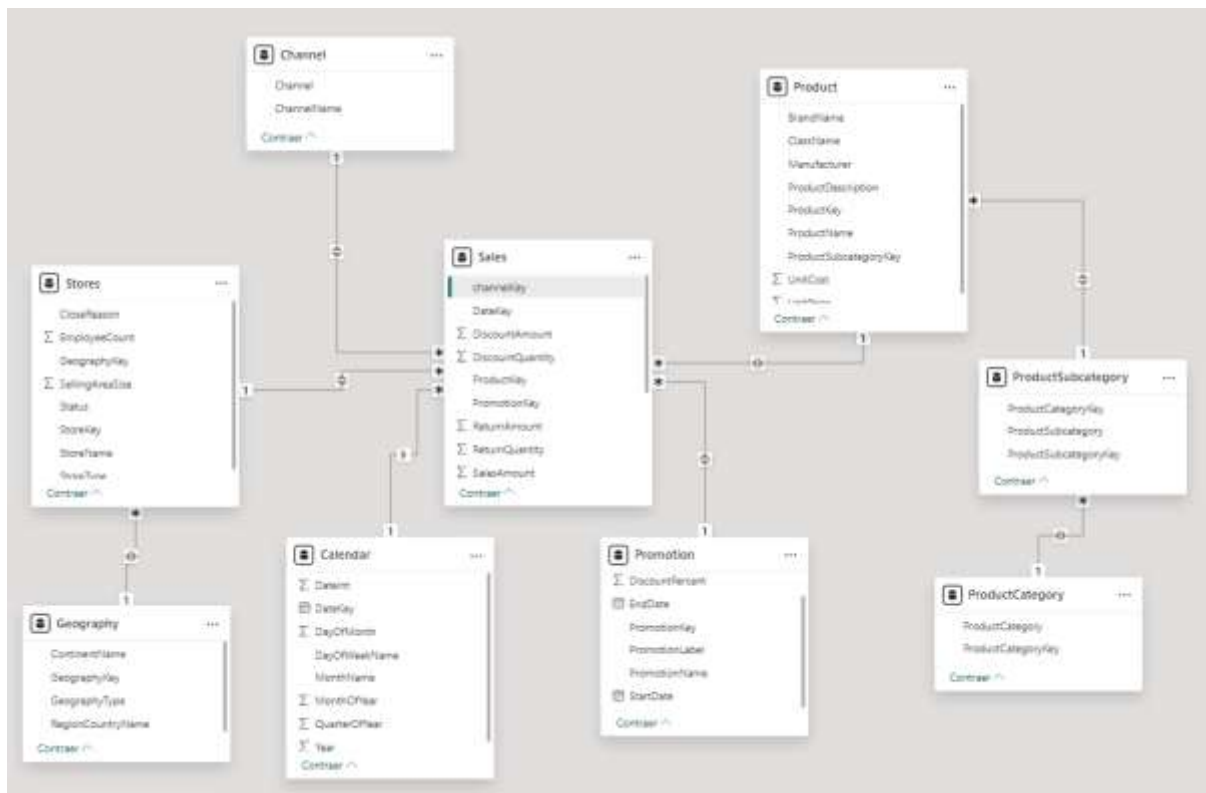


Ilustración 38. Modelo de datos

Fuente Power BI Desktop

Hacemos un pequeño paréntesis, antes del desarrollo, ¿qué es una medida?, ¿y una columna calculada? Esto es algo que hay que tener muy presente a la hora de desarrollar porque son dos conceptos que, aunque se pueden parecer son muy diferentes. Una medida usualmente es el resultado de una operación, no se almacenan y las calcula la CPU mientras que una columna es un cálculo por cada fila, las calcula la RAM y si se almacenan en memoria. Y ahora si comenzamos con el desarrollo de los objetos visuales.

Primer ejemplo

Caso de uso: Distribución de ventas con promoción por Categoría y por Subcategoría

Como usuario: quiero visualizar la distribución por categoría y subcategoría de las ventas que tienen promoción para poder realizar un análisis de ello.

Como desarrollador: lo primero será investigar que datos nos hacen falta para poder mostrar al usuario lo que necesita y en que tablas podemos encontrarlos. En este caso sabemos que tenemos que hacer uso de la tabla ProductCategory y ProductSubcategory y además para calcular las ventas con promoción también tendremos que revisar la tabla de Promotion y Sales.

Una vez tenemos las columnas necesarias pensamos cual sería la mejor forma de representar los datos.

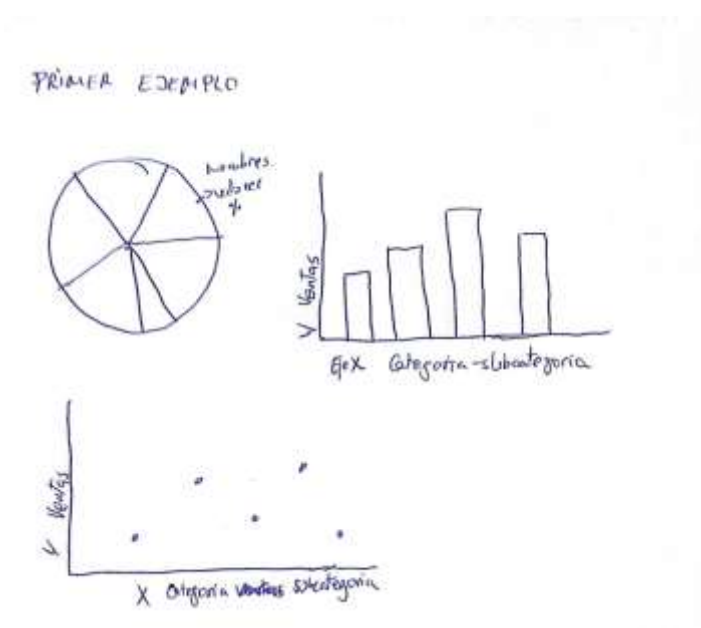


Ilustración 39. Mock up Ejemplo 1

Tenemos tres posibilidades, gráfico circular, gráfico de barras y gráfico de dispersión. El primero podría servir para la categoría porque no hay muchas opciones, pero para la subcategoría podría quedar muy ajustado y no distinguirse bien el nombre. Y entre el segundo y el tercero puede ser que visualmente quede mejor la distribución en uno de dispersión, por lo que esta vez utilizaremos ese.

Tras la elección crearemos la siguiente columna calculada (Tabla Promotion) y la medida para las ventas con promoción.

```
DiscountApplied = IF(Promotion[PromotionName] = "No Discount", "No", "Yes")
```

```
PromotionSales = CALCULATE(SUM(Sales[SalesAmount]), Promotion[DiscountApplied] = "Yes")
```

Como resultado obtenemos lo siguiente: en el eje X insertaremos las columnas ProductCategory y ProductSubcategory y en el eje Y la medida de Ventas con promoción. Para pasar de la Ilustración 40 a la Ilustración 41 solo tenemos que clicar en las flechitas que aparecen en la parte superior izquierda y bajara la jerarquía, lo mismo si queremos subir.

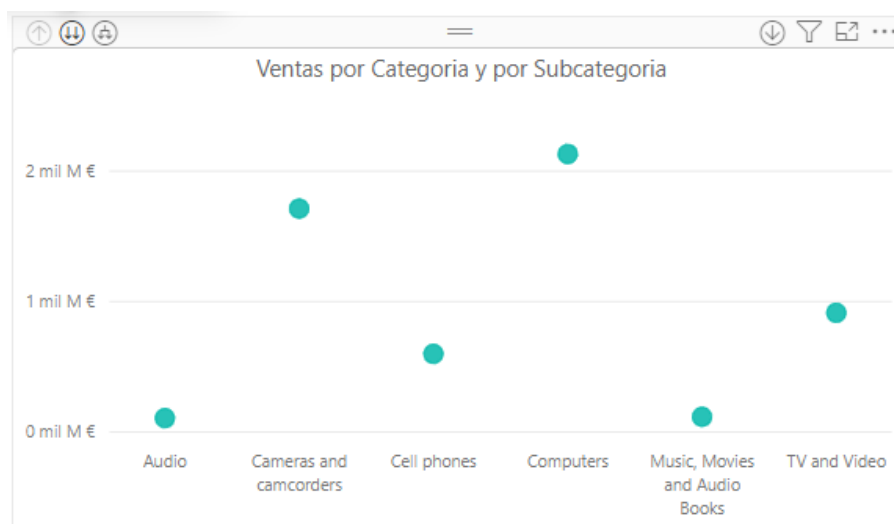


Ilustración 40. Ventas con promoción por categoría



Ilustración 41. Ventas con promoción por subcategoría

A partir de los gráficos podemos observar que en cuanto a categorías las más destacas en cuanto a ventas con promoción son Computers, Cameras and Camcorders y en cuanto a las subcategorías podemos observar que siguen destacando las referentes a las categorías anteriores.

Segundo ejemplo

Caso de uso: Análisis de Ventas por Canal

Como usuario: quiero poder evaluar el rendimiento en los diferentes canales de ventas para comparar y visualizar cuales generan mayor número de ingresos y donde podríamos mejorar. Puntualización: las ventas tienen que ser sin aplicar descuento.

Como desarrollador: tenemos que investigar que datos necesitamos para poder mostrar lo que el usuario necesita y en que tablas se encuentran. En este caso en la tabla de Channel, encontraremos el nombre del mismo y en la tabla de Sales diferentes columnas de precios. Deberemos preguntarnos, ¿qué columna necesito para calcular las ventas totales?

SalesKey	DateKey	channelKey	StoreKey	ProductKey	PromotionKey	UnitCost
537685	23/12/2011 0:00:00	1	54	132	4	\$101,97

UnitPrice	SalesQuantity	ReturnQuantity	ReturnAmount	DiscountQuantity	DiscountAmount	TotalCost	SalesAmount
\$200	6	0	\$0	1	\$40	\$611,82	\$1.160

Ilustración 42. Tabla de Ventas

Viendo esto podemos observar que la columna SalesAmount para este producto es 1.160\$, pero que si multiplicamos la columna UnitPrice*SalesQuantity obtenemos un total de 1200\$. Por tanto, esos 40\$ de diferencia vendrían a ser del DiscountAmount.

Por tanto, necesitamos crear una medida con la multiplicación anterior para tener las ventas que nos pide el usuario. Utilizaremos el SUMX ya que así podemos realizar operaciones en cada fila.

Sales = SUMX(Sales, Sales[UnitPrice]*Sales[SalesQuantity])

A continuación, realizaremos un boceto de cuáles serían las posibles formas de representar los datos.

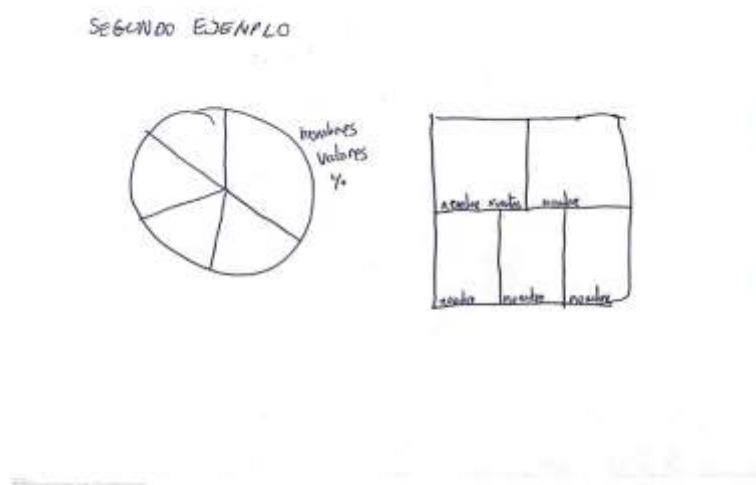


Ilustración 43. Mock up Ejemplo 2

En este caso los dos objetos son muy parecidos, al tener pocas opciones visualmente se va a ver muy bien tanto en el circular como en el treemap, pero optaremos por el primero. E insertando el Channel en la leyenda y la medida de ventas en valores podemos visualizar lo siguiente:

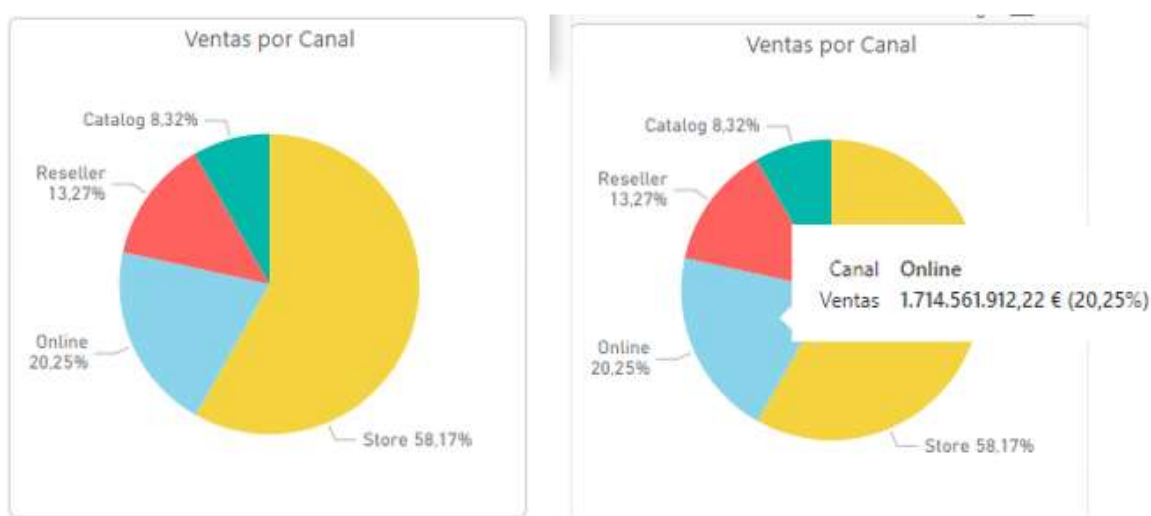


Ilustración 44. Grafico circular ventas por canal

Por defecto Power BI nos mostraría solo la categoría de los canales, pero nos da la opción de elegir diferentes opciones. Nosotros nos quedaremos con la Categoría, porcentaje total. (ver ilustración izquierda). Además, al ser interactivo si pasamos el ratón por encima del objeto visual también podremos ver el número de

ventas de cada canal (ver ilustración derecha), incluso poder añadirle otra columna como información extra.

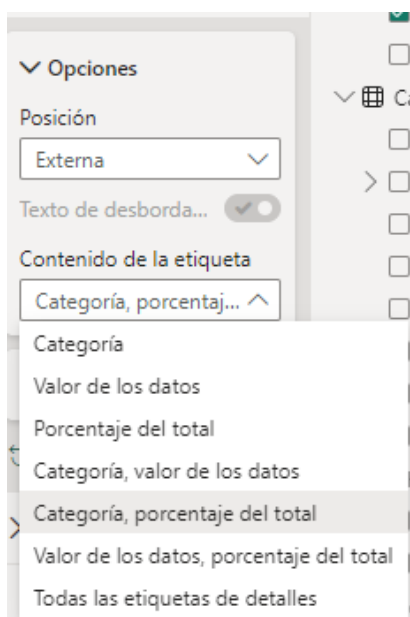


Ilustración 45. Opciones de gráfico circular

Tras visualizar los datos podremos decir que la opción más destacada es Store con casi un 60% y la de Catalog sería la que menos con un 8.32%.

Tercer ejemplo

Caso de uso: Análisis geográfico

Como usuario: quiero poder analizar las ventas en diferentes regiones, indicando áreas con alto o bajo rendimiento de estas.

Como desarrollador: necesitamos saber los datos necesarios para poder llevar a cabo la visualización. Por un lado, tenemos la tabla Geography que contiene 4 campos, la Key, GeograpyType (City, Continent, etc.), ContinentName y RegionCountry Name. Como el usuario quiere verlo por regiones elegiremos la última columna mencionada. Y en cuanto a las ventas usaremos la medida anteriormente calculada.

Para esta representación Power BI nos ofrece dos opciones de mapas: mapa convencional y mapa coropletico. La diferencia entre ambos es que el segundo

utiliza una gama de color para mostrar la cantidad de registros en las áreas y estas se pintan de acuerdo a las variaciones entre sectores.

A continuación, mostramos un ejemplo de ambos:



Ilustración 46. Gráfico de mapa. Ventas por región

Como podemos observar en imágenes para este caso de uso parece que nos da más información el primer mapa ya que el tamaño de las burbujas varía en función de las ventas y en el segundo simplemente colorea la región donde existen ventas. Por consecuente elegiremos el Mapa convencional y además añadiremos diferentes colores para poder visualizar lo que pidió el usuario, identificar áreas con alto y bajo rendimiento en ventas. En las siguientes imágenes podemos ver cómo quedaría nuestro mapa finalmente (ver Ilustración 47) y cómo hemos creado las reglas para la diferenciación de colores (ver Ilustración 48).

Con el mapa podemos obtener información de que las mayores ventas están en Estados Unidos, China, Francia, Reino Unido y Alemania y por el contrario las más bajas entre Suiza, Suecia, Países Bajos, Malta y Portugal.



Ilustración 47. Mapa definitivo de ventas por Región

Color predeterminado - Burbujas - Colores

Estilo de formato

Reglas:

¿En qué campo debemos basar esto?

Sales

Reglas

11 Inversión del orden de los colores + Nueva regla

Si el valor	>=	22500	Número	>	50000	Número	entonces	 	↑ ↓ ×
Si el valor	>	10000	Número	>	22500	Número	entonces	 	↑ ↓ ×
Si el valor	>=	13000	Número	>	15300	Número	entonces	 	↑ ↓ ×

Ilustración 48. Opciones para reglas de colores

Cuarto ejemplo

Caso de uso: Análisis de devoluciones enfocado en tiendas

Como usuario: quiero poder visualizar cuales son las 5 tiendas que más devoluciones tienen y las 5 que menos, pudiendo ver tanto la suma de devolución como las ventas totales que ha tenido cada tienda.

Como desarrollador: tenemos que ver que columna de la tabla de Sales corresponde a devoluciones y que columna de la tabla de Stores necesitamos para que nos muestre el nombre de las tiendas. Después investigaremos la forma de que sean las 5 donde más ventas hay y las 5 donde menos y veremos de qué manera poder representarlo. En la tabla de Sales tenemos una columna llamada “ReturnQuantity” que nos muestra las devoluciones y en la tabla de Store “StoreName”. Podemos representar los datos tanto en tablas como por ejemplo en gráfico de barras y para que sea más visual elegiremos este último. En Power BI si añadimos en el eje X el nombre de la tienda y en el eje Y las devoluciones tendremos lo siguiente:

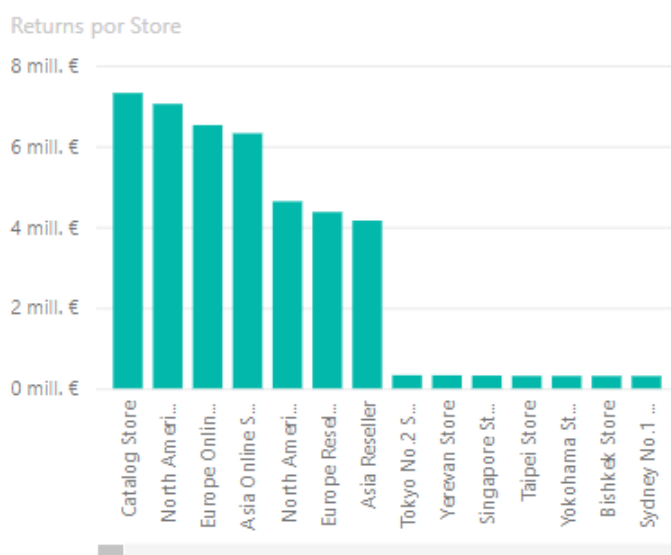


Ilustración 49. Boceto Gráfico de barras

Pero el usuario nos ha pedido que fuese un Top 5 con las máximas y un Top 5 con las mínimas, por lo que tendremos que crear dos gráficos y añadirle un filtro a cada uno. Seleccionamos un gráfico y en filtros del objeto añadimos (ver Ilustracion 50). Para uno será Superior y para el otro Inferior.

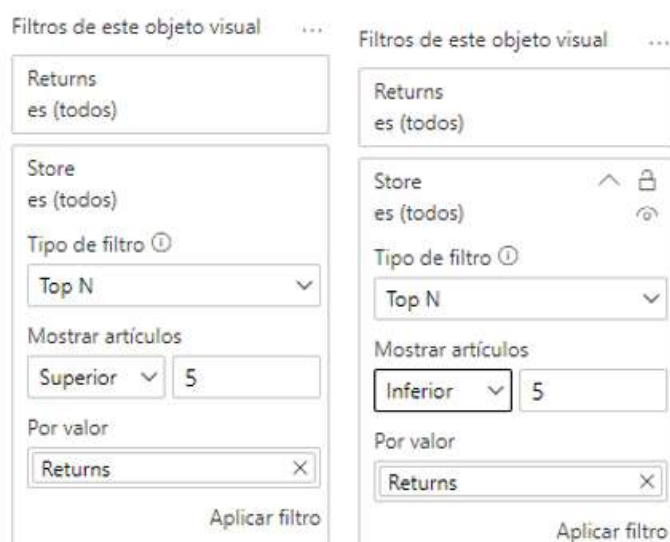


Ilustración 50. Opciones de filtrado de gráficos para Top N

Y obtendríamos lo siguiente, una vez cambiado el título por defecto, le añadamos un borde para que no parezca que está flotando en la página y el 'tooltip' (cuadro que sale informativo cuando pasas con el ratón por encima de una columna) con las ventas totales de cada tienda.



Ilustración 51. Gráficos de barras finales

Podríamos decir que ya está finalizado y representado este caso de uso porque mostramos lo que el usuario nos pide, pero vamos a ir más allá, valorando cuanto espacio ocupan dos gráficos que son muy similares y cómo podríamos reducirlo. En Power BI existen los marcadores que sirven para capturar el estado de una pestaña y junto con un botón nos ayudan bastante en la reducción de espacio sin tener que cambiar a otra pestaña diferente.

El procedimiento a seguir es: crear dos botones, uno para máximas y otro para mínimas, los dos gráficos que ya teníamos, crear también dos marcadores, uno para cada botón y en los botones reflejar la acción correcta.

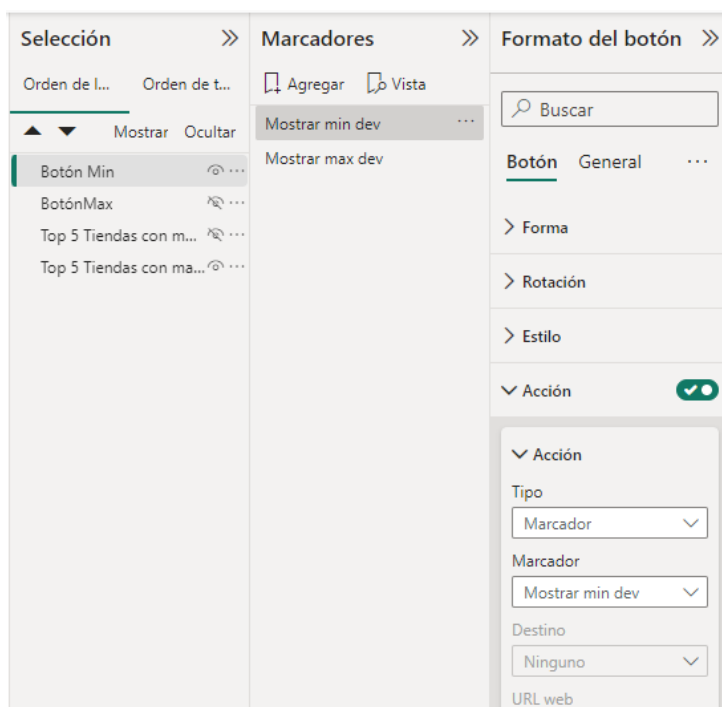


Ilustración 52. Opciones de Selección, Marcadores y Formato

Y nos quedaría algo como lo siguiente, aunque en Power BI se vería solo una opción dependiendo del botón que cliquemos. Max devoluciones o Min devoluciones. Podemos observar que la mínima tiene 19000\$ en devoluciones y la máxima 7.300.000\$, sin entrar en rasgos específicos.



Ilustración 53. Gráficos finales con botones

Quinto ejemplo

Caso de uso: Analizar las tendencias de las ventas con promoción por periodo:

Como usuario: quiero poder identificar si existen tendencias estacionales o patrones que se repiten a través del tiempo en las ventas que tienen promoción.

Como desarrollador: tenemos que revisar las fechas que existen en las tablas y ver si existe la columna que nos sirva para la representación de dicho periodo y también deberemos crear las medidas necesarias para el cálculo de las ventas con promoción.

Tenemos varias opciones para poder representar estos datos y las visualizamos en el siguiente mock-up.

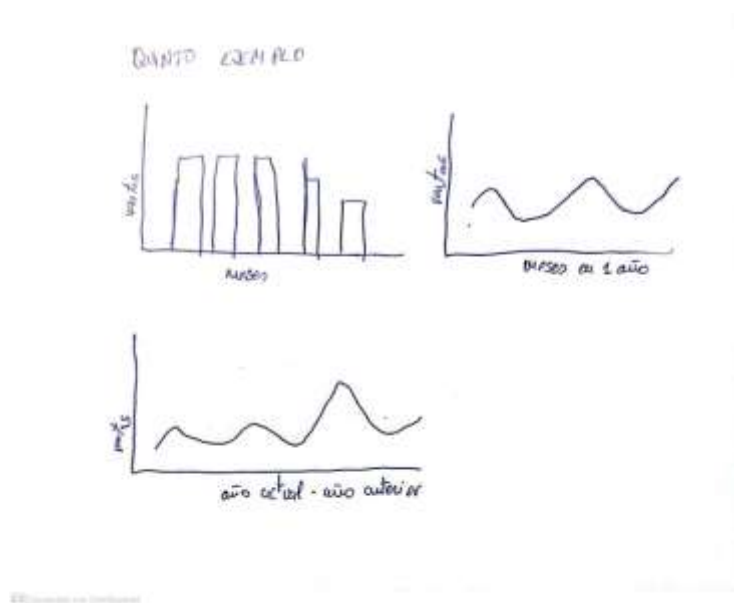


Ilustración 54. Mock up Ejemplo 5

Lo primero es crear las medidas y columnas calculadas que necesitamos y después llevaremos a cabo la representación en Power BI.

Para el tema de fechas es muy importante tener en cuenta la ordenación por columnas, ya que, si insertamos directamente el mes en un gráfico, este nos saldrá ordenado alfabéticamente, por tanto, tenemos que considerar ordenar la columna de MonthName por DayOfMonth.

Esta columna (MonthYear) creada en la tabla Calendar la ordenaremos por MonthName

```
MonthYear = 'Calendar'[MonthName] & " " & 'Calendar'[Year]
```

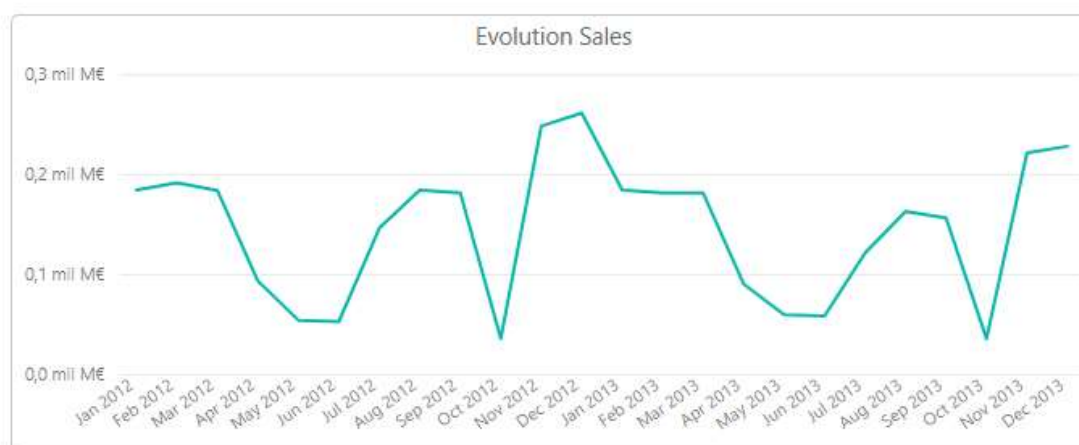
Está otra la ordenaremos por DATESK que contiene la fecha en formato texto (201101).

```
ShortMonthYear = LEFT('Calendar'[MonthYear],3) & " " & RIGHT('Calendar'[MonthYear],4)
```

Y la medida “final” que será la que insertemos en el gráfico de líneas.

```
SalesCY&LY = var cy = 2013  
return CALCULATE(SUM(Sales[SalesAmount]), Promotion[DiscountApplied] = "Yes",  
FILTER('Calendar', YEAR('Calendar'[DateKey]) = cy || YEAR('Calendar'[DateKey]) = cy -1))
```

Este sería el resultado eligiendo el grafico de líneas donde podemos ver el año en curso y el año anterior para así ver la evolución requerida por el usuario.



En la evolución podemos ver como más o menos la tendencia durante el año se mantiene, sobre el mes de mayo, junio bajan las ventas al finalizar el verano en septiembre vuelven a subir y el pico estaría en los meses de diciembre

7. CONCLUSIONES

El concepto de Business Intelligence ha evolucionado significativamente en los últimos años llegándose a convertir en un componente fundamental dentro de las empresas, no solo para la toma de decisiones y el entendimiento de la situación actual en la que se puede encontrar el negocio, sino también para anticiparse a ciertos patrones que se han podido estar cumpliendo que hasta ahora quizás no eran del todo claros y lograr optimizar procedimientos a través de insights basados en datos. Cualquier empresa importante que no invierta en este sector estará en clara desventaja frente a sus competidores lo que puede suponer unos malos resultados.

En cuanto a las herramientas analizadas, hemos podido comprobar que hay una gran variedad en el mercado, cada una con sus ventajas e inconvenientes. La mayoría poseen las mismas capacidades o cualidades a rangos generales, la elección dependerá de las necesidades específicas de cada empresa, de la complejidad que tengan sus datos y de que recursos disponibles tienen para su implementación y mantenimiento. Particularmente, Power BI sigue siendo líder en este ámbito (cuarto año consecutivo, Gartner), como plataforma más completa y con una mayor capacidad de ejecución en el mercado gracias a su interfaz amigable y versátil y a la integración con productos del ecosistema de Microsoft que la convierten en una opción atractiva para muchas empresas.

Finalmente, en el último apartado de casos prácticos, hemos podido observar en gran medida, ya que Power BI es una herramienta muy interactiva, la aplicabilidad de esta en la elaboración de los objetos visuales desarrollados. Como hemos dicho a partir del apartado de herramientas, Power BI dispone de una interfaz muy amigable y nos da la posibilidad de realizar diferentes variedades de gráficos, habiendo incluso varias opciones para representar lo mismo. Con el desarrollo hemos podido adentrarnos un poco más en la interfaz y descubrir con la última versión algunas novedades.

En resumen, después de todo lo descrito durante el trabajo y de todo lo investigado es un hecho que Business Intelligence y la visualización de datos son parte esencial de la inteligencia empresarial, ambos han tenido una evolución muy

significativa desde que se descubrieron y con el avance de la tecnología y el aumento de los datos que consumimos seguirán creciendo a lo largo de los años. Pero hay algo que me ha llamado mucho más la atención, que de cierta manera ya sabía pero que después de esto me queda aún más claro, lo poderosa que es la mente humana y como con el avance de la psicología cognitiva y la investigación de la percepción humana se ha mejorado la interpretación de información y el reconocimiento de patrones por parte del cerebro, la facilidad que esta tiene para comprender imágenes y elementos visuales de forma rápida en lugar de textos y números simples todos dispuestos en tablas.

En un mundo donde todo va tan deprisa es muy importante que cuando llegue el momento de tomar decisiones estratégicas el informe que se presente se haya realizado bajo unas buenas prácticas y visualmente de un solo vistazo el usuario pueda ver dónde están los problemas y donde seguir como hasta el momento.

8. BIBLIOGRAFÍA

Agile manifiesto. (s.f.). <http://agilemanifesto.org/iso/es/manifesto.html>

APD (s.f.) <https://www.apd.es/metodologia-lean-que-es/>

Asana (s.f.) <https://asana.com/es/resources/extreme-programming-xp>

Astera (s.f.) <https://www.astera.com/es/type/blog/types-of-data-marts/>

Bismart (s.f) <https://blog.bismart.com/herramientas-de-visualizacion-de-datos>

Cano J.LL. (2007). *Business Intelligence. Competir con Información*.
https://itemsweb.esade.edu/biblioteca/archivo/Business_Intelligence_competir_con_informacion.pdf

Cloud Foundation (s.f.). https://cloudfoundation.com/blog/what-is-micro-strategy/#Overview_and_Introduction_to_MicroStrategy

Curto,J. (2010). *Introducción al business intelligence*. https://cursos.yura.website/wp-content/uploads/2020/03/Introduccion_al_Business_Intelligence.pdf

Designthinking (s.f.). <https://designthinking.gal/kanban-que-es-y-para-que-sirve/>

DOMO. (s.f.). <https://www.domo.com/es>

Escuela de negocios (s.f.). <https://www.escueladenegociosfeda.com/blog/50-la-huella-de-nuestros-docentes/471-gestion-agil-vs-gestion-tradicional-de-proyectos-como-elegir>

Flaticon (s.f.). <https://www.flaticon.es/>

Google Cloud. (s.f.). <https://cloud.google.com/looker-studio?hl=es>

IBM (s.f.). <https://www.ibm.com/es-es/topics/data-mart>

INESDI (s.f.). <https://www.inesdi.com/>

INESDI (s.f.). <https://www.inesdi.com/blog/que-son-las-metodologias-agiles-tipos-y-ejemplos/#:~:text=Las%20metodolog%C3%ADas%20%C3%A1giles%20promueven%20una,en%20curso%20y%20dar%20feedback>

Inmon, W. (2005). *Building the Data Warehouse*. (3ª ed., pp.81-98)
<https://fit.hcmute.edu.vn/Resources/Docs/SubDomain/fit/ThayTuan/DataWH/Bulding%20the%20Data%20Warehouse%204%20Edition.pdf>

Kimball, R (2013). *The Data Warehouse Toolkit*. (2º ed., pp.78-88)
<http://160592857366.free.fr/joe/ebooks/ShareData/The%20Data%20Warehouse%20Toolkit.pdf>

Luhn, H. (1958). *A Business Intelligence System*.
<https://www.businessintelligence.info/definiciones/business-intelligence-system-1958.html>

Marqués,M. (2015). *Business Intelligence. Tecnicas, herramientas y aplicaciones*.
<https://archive.org/details/perez-marques-maria-business-intelligence-alfaomega/page/3/mode/2up>

MediaCloud (s.f.). <https://mdcloud.es/que-es-etl-extraccion-transformacion-y-carga/>

Microsoft. (s.f.). <https://powerbi.microsoft.com/es-es/why-power-bi/>

Microsoft Learn. (s.f.). <https://learn.microsoft.com/es-es/power-bi/fundamentals/power-bi-overview>

OBS (s.f.). <https://www.obsbusiness.school/blog/las-5-etapas-en-los-sprints-de-un-desarrollo-scrum>

Power BI. (5 de febrero de 2024). En *Wikipedia*. https://es.wikipedia.org/wiki/Power_BI

Qlik. (s.f.). <https://www.qlik.com/es-es/products/qlikview>

Ramos, S. (2016). *BI & Analytics: El arte de convertir los datos en conocimiento*

Real Academia Española. (s.f.). Inteligencia. En *Diccionario de la lengua española*.
<https://dle.rae.es/inteligencia?m=form>

Real Academia Española. (s.f.). Negocio. En *Diccionario de la lengua española*.
<https://dle.rae.es/negocio>

Reddesign Systems. (s.f.). <https://reddesignsystems.com/2024/05/30/historia-inteligencia-negocios/>

Roberth G. Figueroa, Camilo J. Soils, Armando A. Cabrera. (2007) *METODOLOGÍAS TRADICIONALES VS. METODOLOGÍAS ÁGILES*.

https://www.researchgate.net/publication/299506242_METODOLOGIAS_TRADICIONALES_VS_METODOLOGIAS_AGILES

Scheps, S. (2008). *Business Intelligence for Dummies*.
https://www.academia.edu/36753857/Business_intelligence_for_dummies_1_

Tableau (s.f.). <https://www.tableau.com/es-es/why-tableau/what-is-tableau>

UNIR. (s.f.). <https://www.unir.net/ingenieria/revista/herramientas-visualizacion-datos/>

Welcomedevolvers_ (s.f.). <https://welcomedevolvers.es/>