

**HENRY**



# Visualización de datos





# Agenda



- Data Analytics
- Evolución del proceso analítico
- El Camino de los datos
- Visualización de Datos
- Herramientas



# **OBJETIVOS DE LA CLASE**

***Al finalizar esta lecture estarás en la capacidad de...***

- Comprender en qué consiste la analítica de datos y cuáles son las tareas que se realizan
- Conocer los distintos Tipos de Análisis
- Entender en qué consiste la Visualización de Datos, Patrones visuales
- Aplicar herramientas de visualización (Introducción a Power BI y Tableau)



# Data Analytics





# ¿Qué es?

Es el conjunto de métodos y técnicas de medición, que permiten gestionar la información en tres grandes etapas:

- **Recolección**
- **Transformación**
- **Visualización**





# ¿Qué hace un analista de datos?

Los analistas de datos buscan determinar cómo se pueden **usar los datos para responder preguntas y resolver problemas.**

Estudian lo que está sucediendo ahora para **identificar tendencias y hacer predicciones** sobre el futuro.



# Tareas de Un Data Analyst

- Trabajar con equipos de tecnología, administración y/o data scientists para **establecer metas**.
- **Minería de datos** de fuentes primarias y secundarias.
- **Limpieza y disección** de datos.
- **Analizar** e **interpretar** resultados utilizando herramientas y técnicas estadísticas.
- Identificar **tendencias** y **patrones** en conjuntos de datos.
- Identificar **nuevas oportunidades** para la mejora de procesos.
- Proporcionar **informes de datos** para la gestión.
- Diseñar, crear y mantener **bases de datos** y **sistemas de datos**.
- **Solucionar problemas** de código y problemas relacionados con los datos.





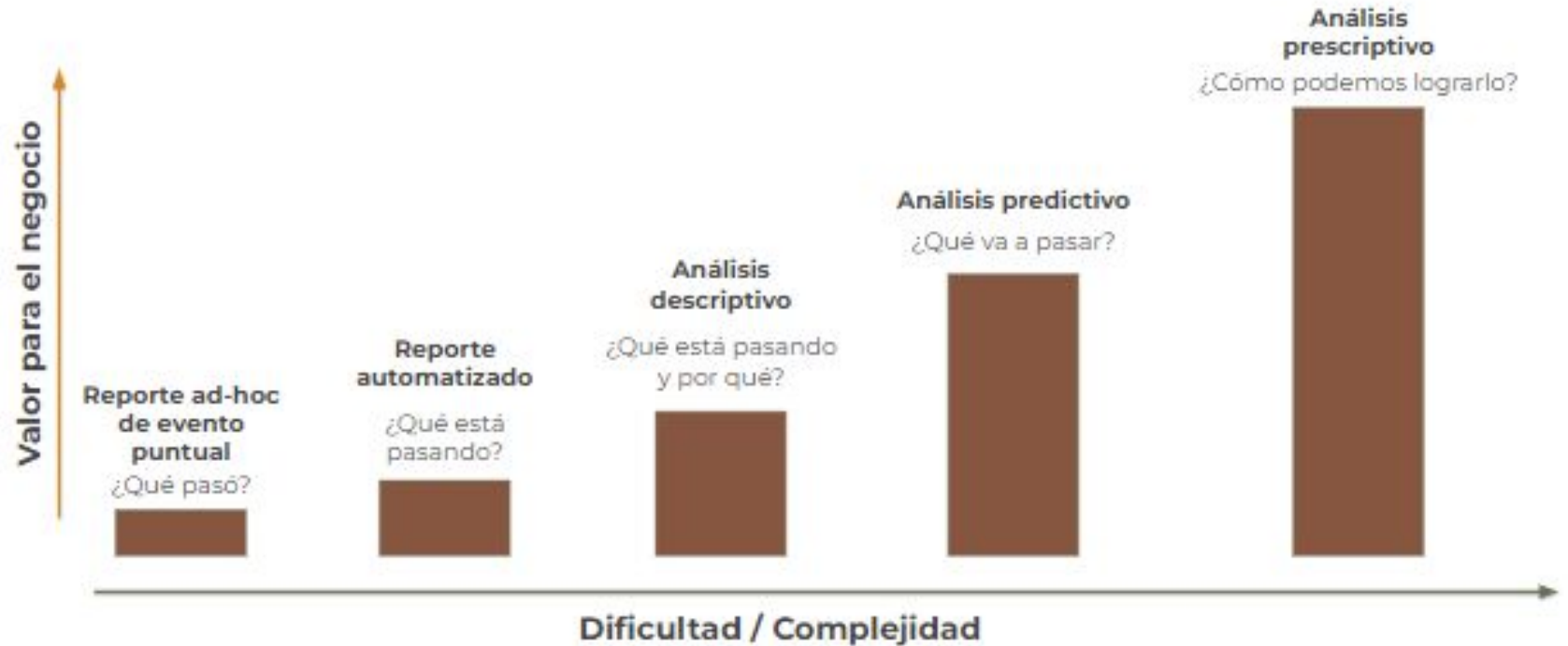
# Evolución del **Proceso Analítico**







# Evolución





# Tipos de Análisis

## Descriptivo



¿Qué pasó?

Describe **patrones claves** en los datos existentes y permite observar **situaciones y comportamientos habituales**.

Reportes

## Diagnóstico



¿Por qué pasó?

Se definen **conclusiones** basadas en los datos históricos identificados en el **análisis descriptivo**.

Dashboards interactivos

## Predictivo

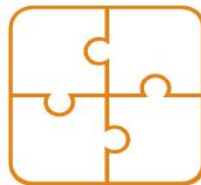


¿Qué pasará?

Se identifican **comportamientos futuros**, en base a información histórica.

Modelos

## Prescriptivo



¿Qué debería hacer?

Con simulaciones y optimizaciones se **compara** la **información** resultante del análisis predictivo, para **implementar** una **acción**.

Recomendaciones de automation

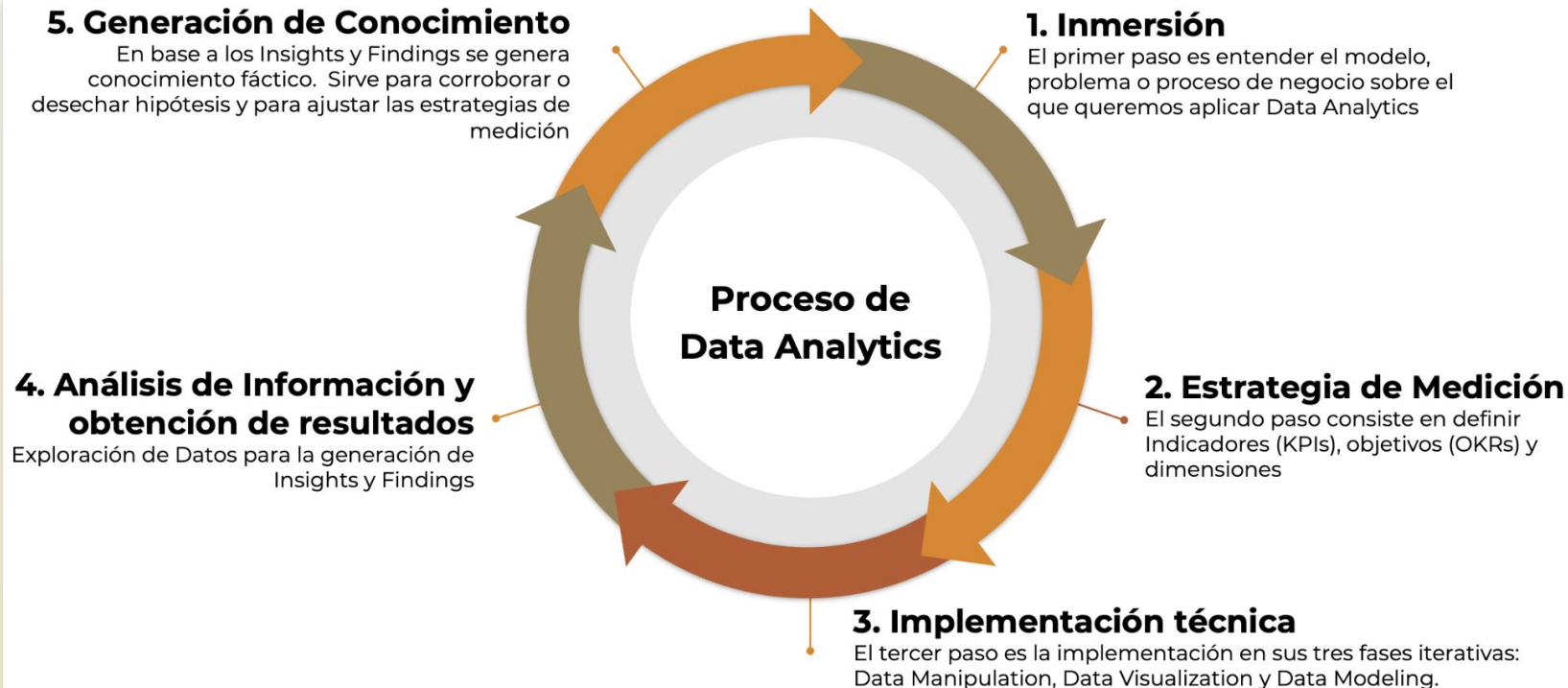


# Data Analytics Journey





# El Camino de los datos





# Visualización de datos



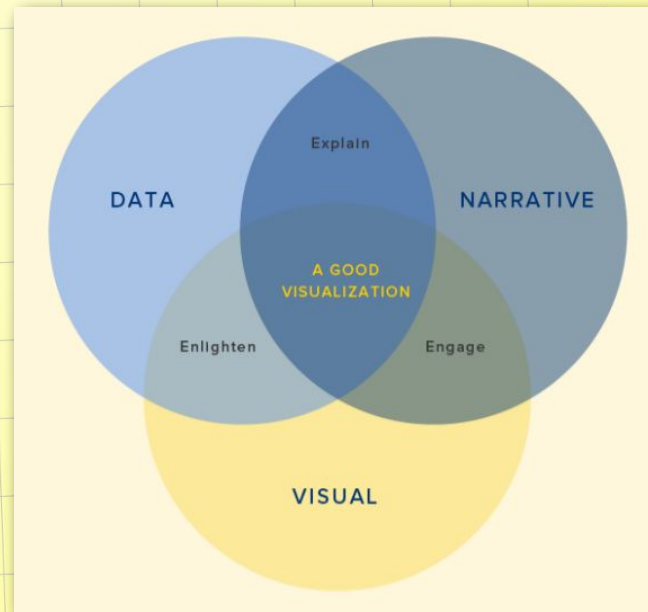


# ¿Qué es?

La visualización de datos es otra forma de arte visual que capta nuestro interés y mantiene nuestros ojos en el mensaje.

Vemos rápidamente las tendencias y los valores atípicos. Si podemos ver algo, lo interiorizamos rápidamente.

Es contar historias con un propósito.



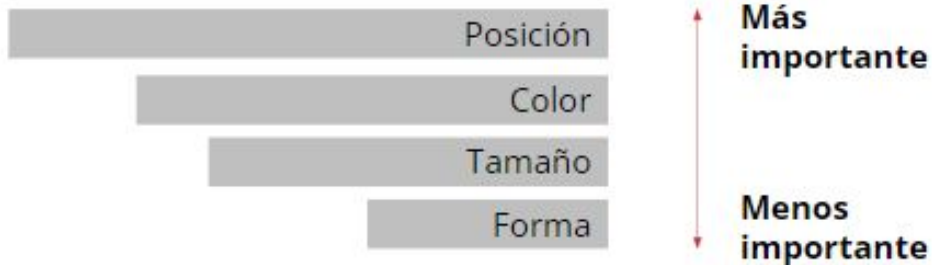




# Principios de Visualización de datos

## Percepción visual

Algunos atributos generan un impacto mayor en nuestro cerebro:







# Principios de Visualización de datos

## Percepción visual

¿Cuántos cuadrados hay? ¿Cuántos círculos? ¿Qué imagen transmite mejor la información?





# Principios de Visualización de datos

## Color

Los usos del color permiten:

1. **SECUENCIA**
2. **DIVERGENCIA**
3. **CATEGORÍA**



# Principios de Visualización de datos

## Color

COLOR: **SECUENCIA**

Los colores **secuenciales** se utilizan para mostrar valores ordenados de mayor a menor

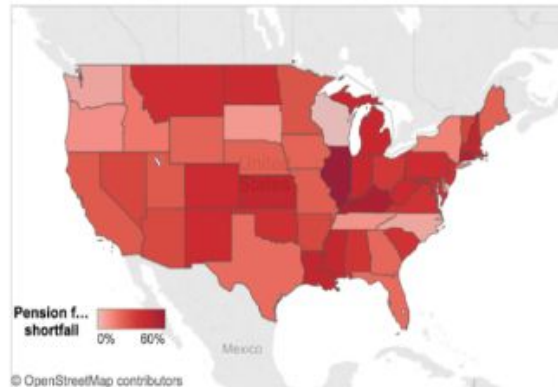
### Pensions in Peril

Despite recent stock market gains, states continue to shortchange their pension plans, leaving many of them badly underfunded. (SOURCE: Pew Charitable Trusts)

CNBC

(Dropdown for AK, HI)

Contiguous US



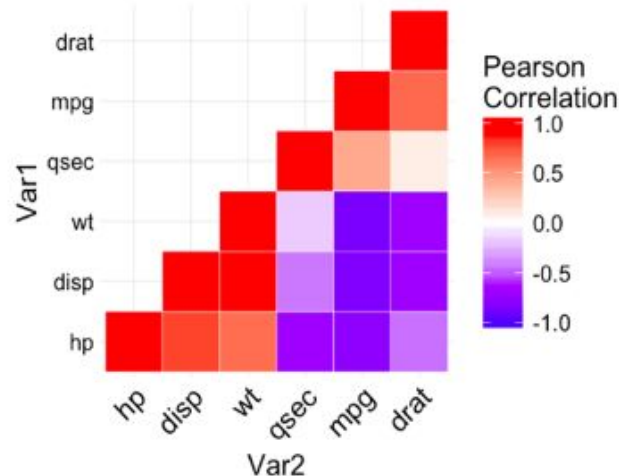


# Principios de Visualización de datos

## Color

COLOR: DIVERGENCIA

Los colores **divergentes** se utilizan para mostrar valores ordenados que tienen un valor crítico, tales como un promedio o cero





# Principios de Visualización de datos

## Color

COLOR: CATEGORÍA

Los colores **categoricos** se utilizan para distinguir datos pertenecientes a diferentes grupos. En particular, se asocia a la representación de **variables categóricas**.

### Body Camera Laws

Ten states have passed laws that control the public's access to footage from police body cameras. Hover over each state for more information.



Source: Reporters Committee for Freedom of the Press



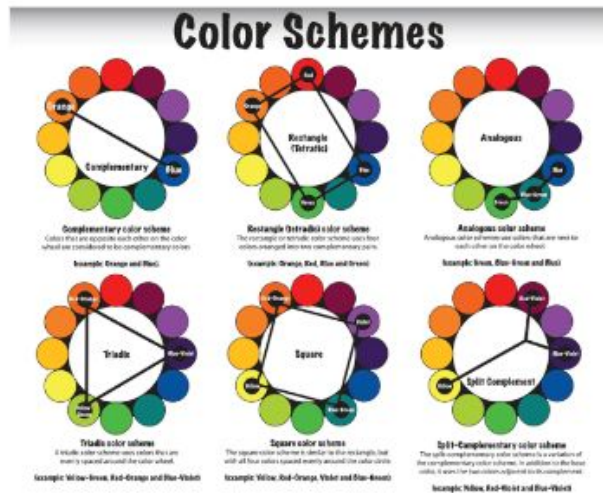
# Principios de Visualización de datos

## Color

### COLOR: CATEGORÍA

Es importante identificar contrastes posibles para diferenciar distintos grupos o categorías:

- 1) **Complementarios**
  - Par
  - Split
- 2) **Triada**
- 3) **Pares complementarios**
  - Rectangular
  - Cuadrangular



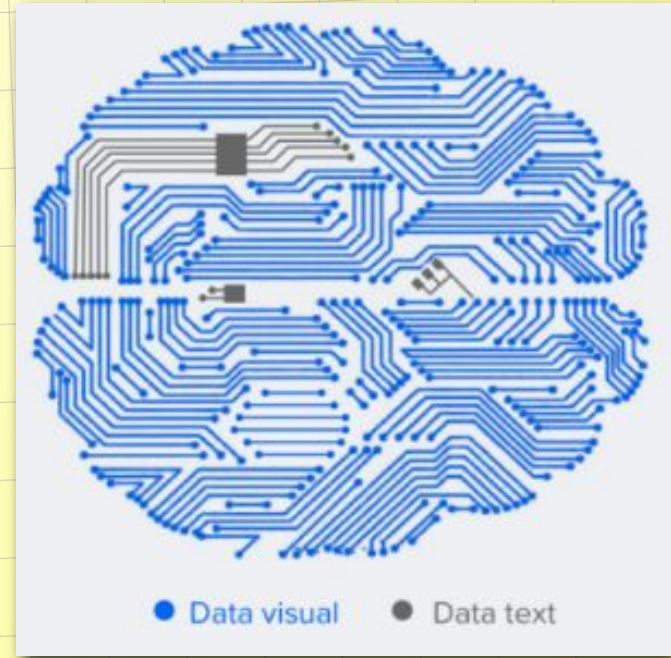




# Diseño

Ver (**percepción**) y comprender (**cognición**) ocurren todo el tiempo.

Cuanto más corto sea el tiempo entre ver y comprender, más efectiva será la visualización.







# Diseño

¿Dónde se pueden observar mejor los números 5?

How many 5s are there?

1021397481027408376037438784937509709167498023780918095097887870212  
0134897834789757939178783939039311891010384887704101993378398370112  
4748790129087348738784890713783580127081273907430740938747801908340  
0239850293809830840980328493982374937807137087571987398621238943221

Try again... how many 5s are there?

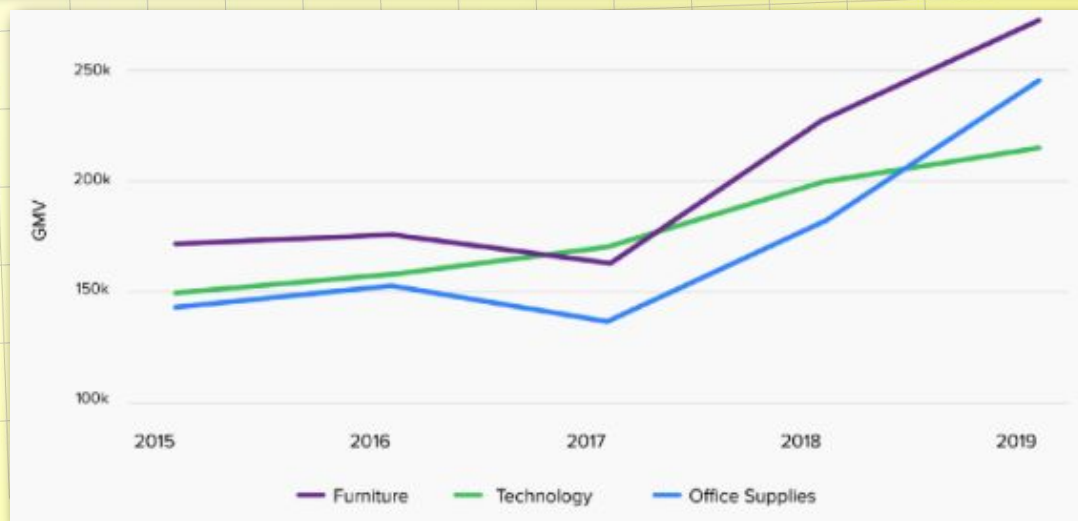
1021397481027408376037438784937**5**0970916749802378091809**5**097887870212  
01348978347897**5**7939178783939039311891010384887704101993378398370112  
4748790129087348738784890713783**5**80127081273907430740938747801908340  
02398**5**0293809830840980328493982374937807137087**5**71987398621238943221



# Diseño

¿Qué categoría ha crecido más rápido?

CATEGORY	2015	2016	2017	2018	2019
Furniture	\$148.624,83	\$157.192,85	\$170.518,24	\$198.901,44	\$215.378,27
Office Supplies	\$143.227,15	\$157.776,41	\$137.233,46	\$183.510,59	\$246.526,57
Technology	\$170.738,67	\$175.278,23	\$162.780,81	\$226.061,80	\$272.033,19





# ¡Importante!

Los procesos cognitivos complejos, consumen gran capacidad de atención.

Esto genera un agotamiento que repercute en la toma de decisiones.

El trabajo de cualquier visualización, es ayudar a percibir rápidamente patrones en la información, que apoyen el proceso de toma de decisiones.

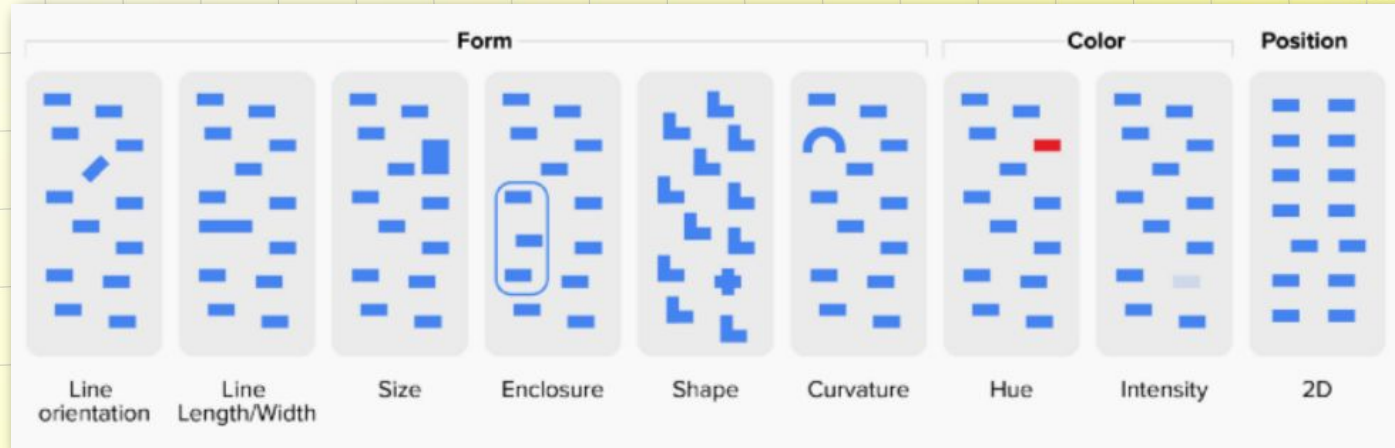




# Patrones visuales

Los usuarios finales no miran los informes, los escanean.

Los patrones visuales, o jerarquías visuales, nos permiten saber qué datos involucran primero a nuestros usuarios, con qué visualizaciones interactúan y en qué orden, mientras están en el tablero.

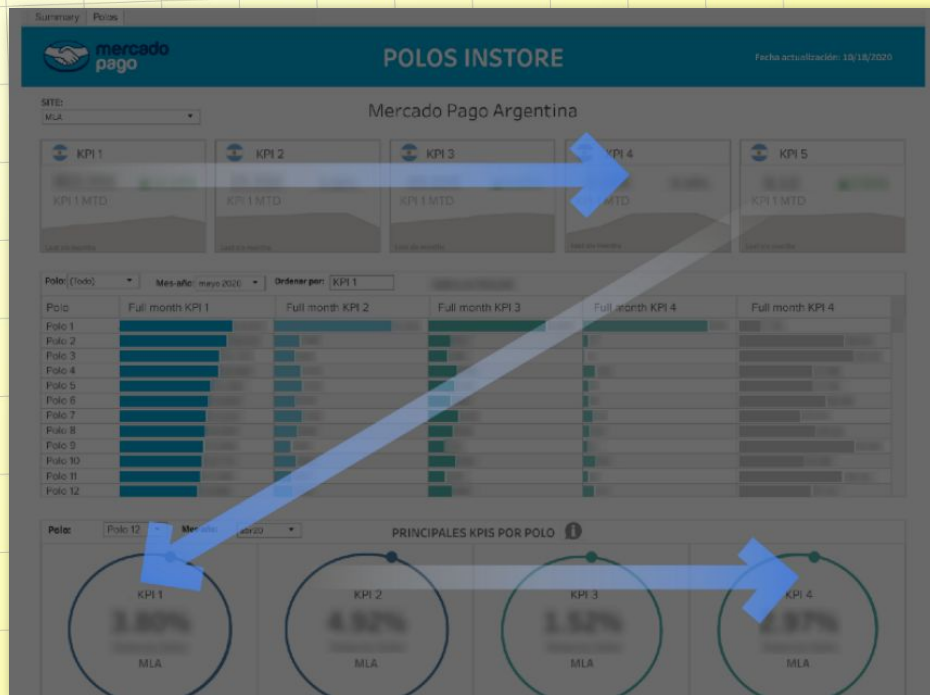




# Patrón Z

Es un concepto de diseño que considera que los usuarios tienden a ver información altamente visual en un Z-Pattern. Se tiende a:

- 1: parte superior izquierda y luego horizontalmente hacia la parte superior derecha.
- 2: corre la vista en diagonal hacia la parte inferior izquierda.
- 3: movimiento horizontal hacia la parte inferior derecha.





# Dashboard: ¿qué es?

Un dashboard es como el tablero del auto.

Se trata de una herramienta esencial para obtener información relevante para la toma de decisiones y centralizar los KPI que sean necesarios para saber qué está pasando realmente con el negocio.







# Dashboard: checklist

1. Entender el contexto.
2. Objetivos a perseguir con nuestro tablero.
3. Lista de preguntas que buscamos responder en el dashboard.
4. Revisión de datos con los que contamos, fuentes, estado de las mismas, limpieza y organización de la data.
5. Dibujo un mock-up/maqueta del tablero, buscando perseguir el objetivo planteado en el punto 1 y verificando que visualizaciones responden a las preguntas del punto 2.
6. Creación del reporte en la herramienta, por ej: Power Bi / Tableau/ Looker/Python
7. Verificar Titulos, formatos, filtros. Contextualizar la data, glosario de datos.
8. Validar con alguien que sea claro el mensaje a transmitir





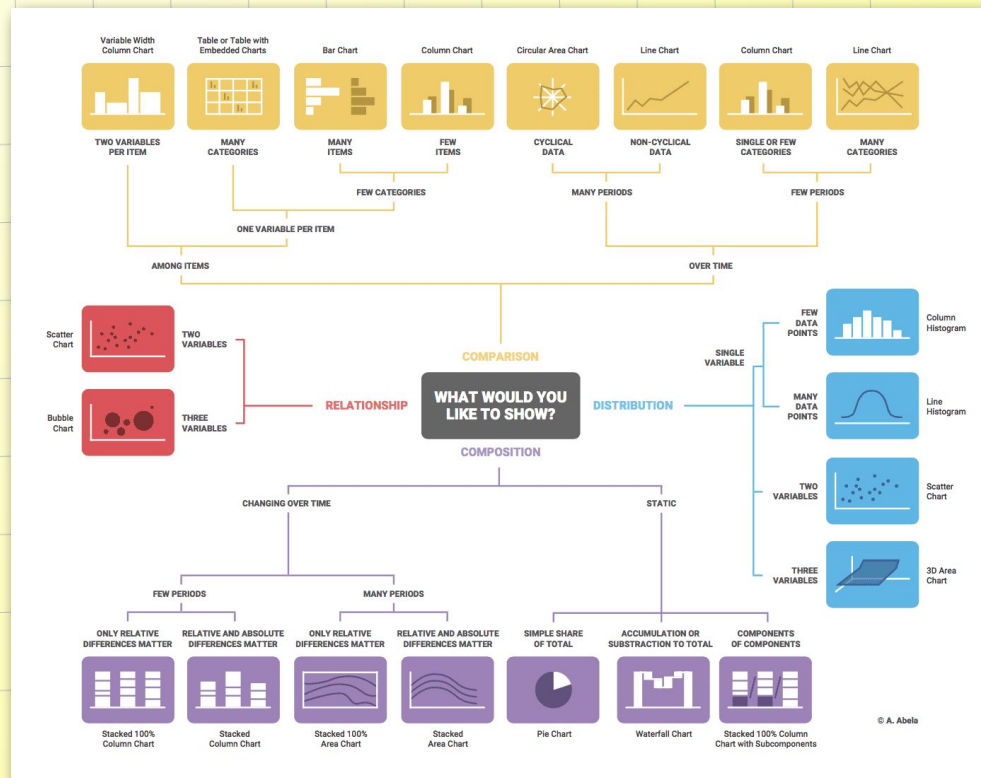
# Elección de gráficos

Todo comienza preguntándose qué se desea representar con los datos, y según la respuesta los gráficos se pueden agrupar en cuatro tipos principales:

- **Comparación**: cuando queremos comparar entre sí los diferentes valores o atributos dentro de los datos.
- **Composición**: cuando queremos saber cómo están compuestos los datos, es decir, qué características generales están presentes en el conjunto de datos.
- **Distribución**: cuando queremos comprender cómo se distribuyen los puntos de datos individuales dentro de todo el juego de datos (gráficos de barras, líneas o gráficos de dispersión).
- **Relación**: estamos interesados en saber cómo los valores y atributos se relacionan entre sí (gráficos de dispersión, gráficos de burbujas).

# Gráficos

La mayoría de los casos se pueden solucionar con unos pocos tipos de gráficos básicos, aunque pueden aplicarse múltiples variantes para cada caso concreto.



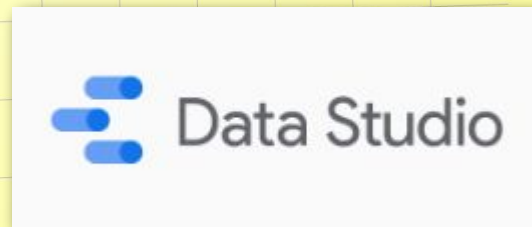


# Herramientas





# Herramientas





# Power BI

- Se posicionó como herramienta líder del mercado debido a su curva de aprendizaje, actualizaciones mensuales, comunidad y bajo costo de adopción.
- 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.
- Permite conectarse con facilidad a los orígenes de datos y aplicar procesos de transformación sencillos sin recurrir a ETL complejos.



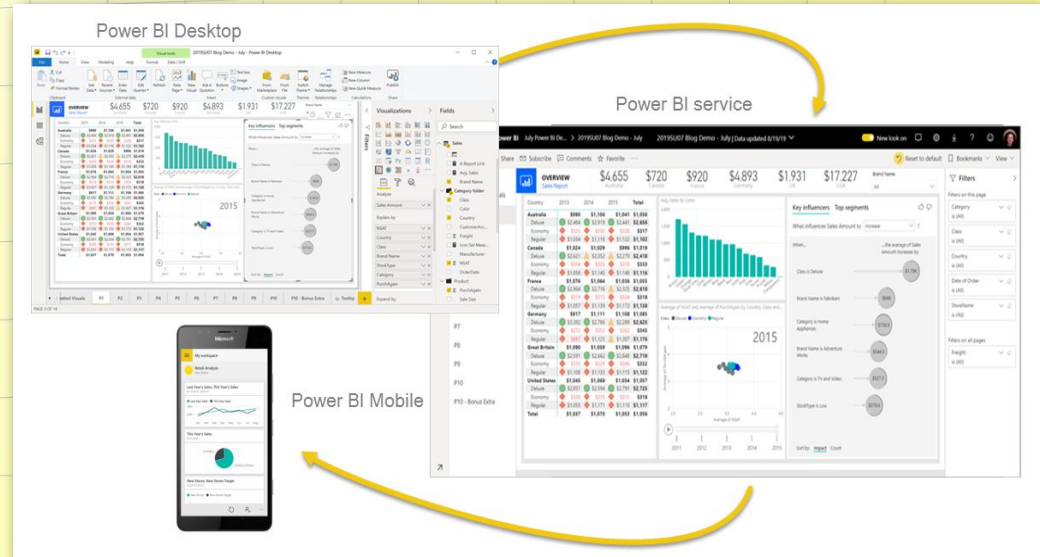




# Power BI

Power BI consta de varios elementos que funcionan de manera conjunta, empezando por estos tres conceptos básicos:

- **Power BI Desktop:**  
Aplicación de escritorio de Windows.
- **Servicio de Power BI:**  
Servicio SaaS (software como servicio) en línea.
- **Aplicaciones móviles** de Power BI para dispositivos Windows, iOS y Android.





# ¿Alguien dijo Homework?





# HENRY

