

Inteligencia de Negocios

Giovany Babativa, PhD

Sobre Mi

PhD en Estadística, MSc en Analytics & Big Data, MSc en Estadística. Con 20 años de experiencia, actual director de analítica en el CNC, miembro del comité de expertos en pobreza en el DANE y consultor de la División de Estadística de la CEPAL. Ex-decano de la Facultad de Estadística USTA, ex-director de operaciones en el ICFES, PM CEV ...

Puedes encontrarme en:

-  [Google scholar](#)
-  [GitHub. https://github.com/jgbabativam](#)
-  [linkedin](#)
-  j.babativamarquez@uniandes.edu.co

Reglas del juego

- Viernes 17 de enero, de 5:30 p.m. a 9:30 p.m.
- Sábado 18 de enero, de 8:00 a.m. a 12:00 p.m.
- Martes 21 de enero, de 5:30 p.m. a 9:30 p.m.
- Viernes 24 de enero, de 5:30 p.m. a 9:30 p.m.
- Sábado 25 de enero, de 8:00 a.m. a 12:00 p.m.
- Martes 28 de enero, de 5:30 p.m. a 9:30 p.m.



3

- Mecanismo de evaluación:
 - **Proyecto precios:** 35%, se permite la entrega por parejas
 - **Proyecto satisfacción:** 35%, se permite la entrega por parejas
 - **Proyecto de tablero de control/visualización:** 30%, trabajo individual



La cadena de valor de los datos



Los datos como recurso estratégico: Nuevas oportunidades, segmentos del mercado, entendimiento del consumidor

- Entender mejor el consumidor
- Necesidades, expectativas
- Quién es el consumidor
- Cuáles son sus hábitos de consumo
- Cuántos segmentos existen en el mercado
- Factores relevantes en momentos de compra
- Cuáles son los momentos del servicio que más afectan la satisfacción



Decisiones informadas



El término fue acuñado por Gartner a mediados de los años 90, pero el concepto se origina en los años 70 con el comienzo de los sistemas de información gerenciales (MIS, Management Informations System). Hoy en día se implementan sistemas conocidos como:

- **CRM** (Customer Relationship Management), que es un sistema diseñado para gestionar las relaciones con los clientes y mejorar las interacciones con ellos, buscando mejorar la experiencia del cliente, aumentar la lealtad y la retención, así como la optimización de procesos de ventas y marketing.

El término fue acuñado por Gartner a mediados de los años 90, pero el concepto se origina en los años 70 con el comienzo de los sistemas de información gerenciales (MIS, Management Informations System). Hoy en día se implementan sistemas conocidos como:

- **ERP** (Enterprise Resource Planning), es un sistema integrado que gestiona las operaciones internas de una empresa, desde finanzas hasta inventarios, buscando centralizar la información, mejorar la eficiencia operativa, reducir errores y duplicación de datos.

Gobierno de datos

El término fue acuñado por Gartner a mediados de los años 90, pero el concepto se origina en los años 70 con el comienzo de los sistemas de información gerenciales (MIS, Management Informations System). Hoy en día se implementan sistemas conocidos como:

- **SCM** (Supply Chain Management), ayuda a gestionar y optimizar la cadena de suministro, desde el aprovisionamiento de materias primas hasta la entrega al cliente final, busca reducir costos de operación, mejorar la eficiencia logística, asegurar la calidad y disponibilidad de productos.

Inteligencia de negocios

Un sistema de inteligencia de negocios incluye **herramientas** y **técnicas** que proporcionan grandes capacidades para la transformación de los datos en conocimiento que contribuyan en la toma estratégica de decisiones que se conviertan en acciones oportunas.

INFORMACIÓN + CONVERSACIÓN + ACCIÓN = TRANSFORMACIÓN

- ¿Qué se imagina cuando le piden una estrategia de inteligencia de negocios?
- ¿Cuáles son los aspectos que se deberían considerar?

05:00

12

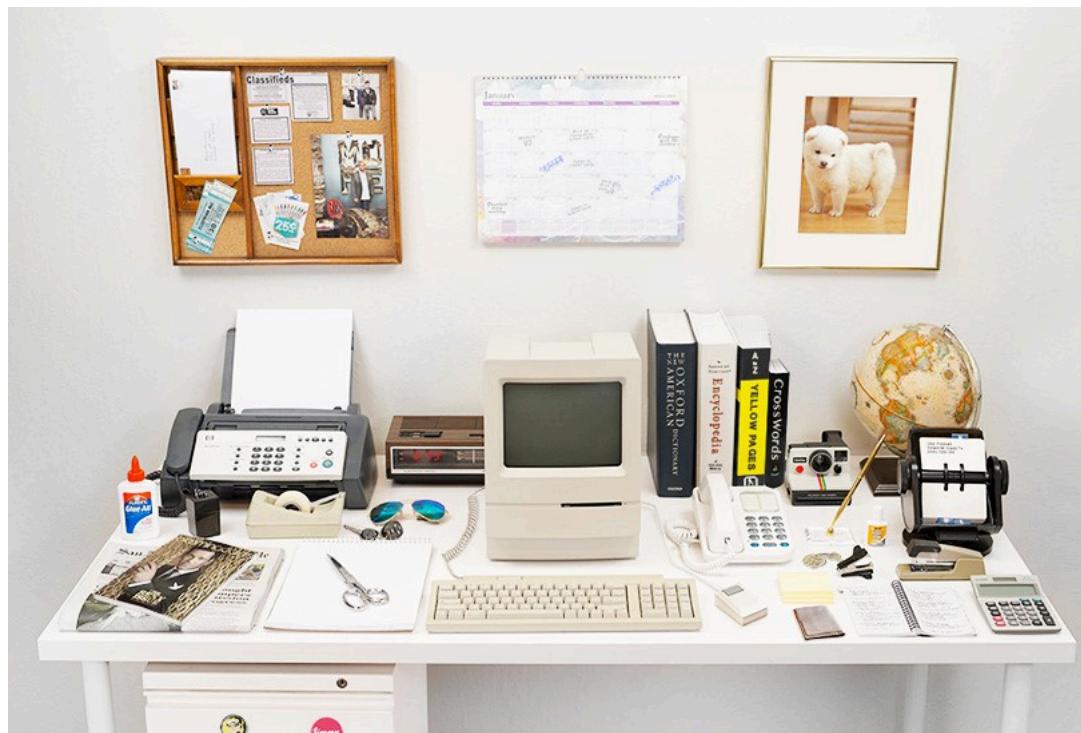
Estaremos hablando de conceptos, técnicas y herramientas que constituyen el soporte de la inteligencia de negocio y la analítica de negocios como:

- Bases de datos
- Metadatos
- *Data Warehouse y Data Marts*
- *Data Lakes*
- Integración de datos
- Datos estructurados y no estructurados
- Herramientas de visualización
- Métodos de visualización
- Analítica de datos
- Modelos de aprendizaje automático
- Ciencia de datos
- Big Data

HACÍA EL RETO BIG DATA

14

Transformación digital



Haz clic para ver el video

Transformación digital



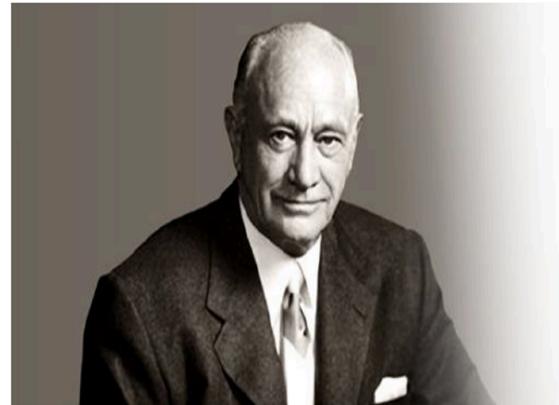
Transformación digital

fundador Airbnb Brian Chesky



Visión digital... repensar el negocio

fundador hoteles Hilton:
Conrad Hilton



Transformación digital

1

CONSUMO
sharing economy



2

DISEÑO Y FABRICACIÓN
open design & manufact.



WIKIPEDIA
The Free Encyclopedia



creative commons



3

CONOCIMIENTO
open knowledge

4

FINANZAS
crowdfunding, lending



5

GOBIERNO
opengov & flat orgs



Holacracy®

6

SISTEMAS DE INTERCAMBIO
currencies, timebanks



7

TRANSPORTE



8

SOFTWARE



OUISHARE

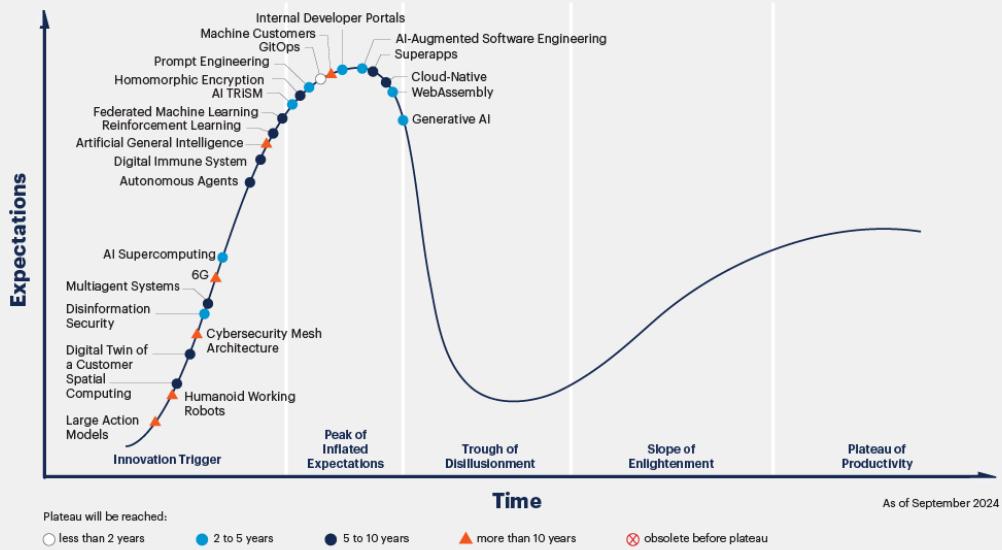
Abundancia de datos



19

Hype Cycle for Emerging Technologies

Hype Cycle for Emerging Technologies, 2024



Source: Gartner
Commercial reuse requires approval from Gartner and must comply with the Gartner Content Compliance Policy on gartner.com.
© 2024 Gartner, Inc. and/or its affiliates. All rights reserved. 3205434

Gartner®

Aspectos clave: Adopción, madurez y aplicación

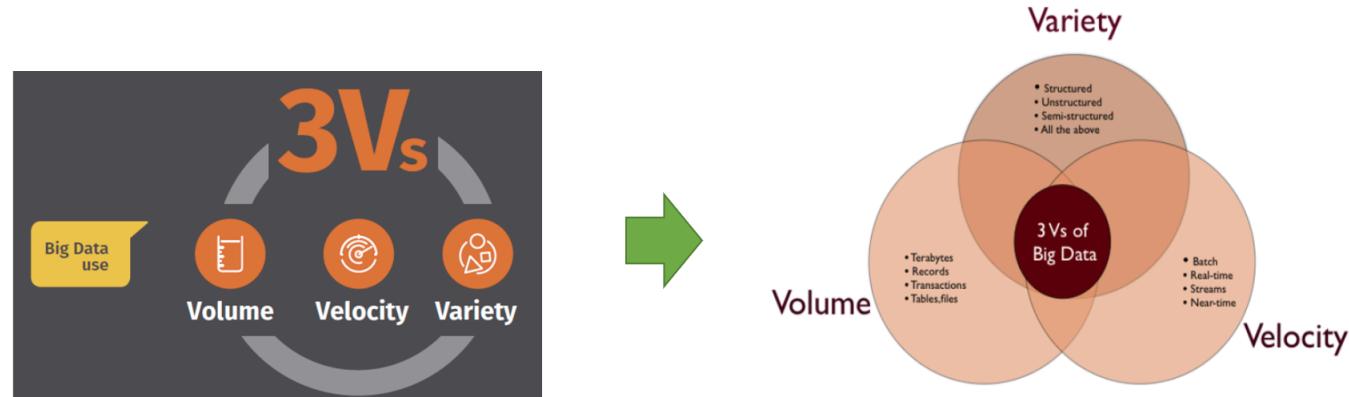
- Eje X (Tiempo): Representa las etapas por las que pasa una tecnología desde su descubrimiento hasta su adopción generalizada o abandono.
- Eje Y (Expectativas): Refleja el nivel actual de expectativas o entusiasmo en torno a esa tecnología.

Fases

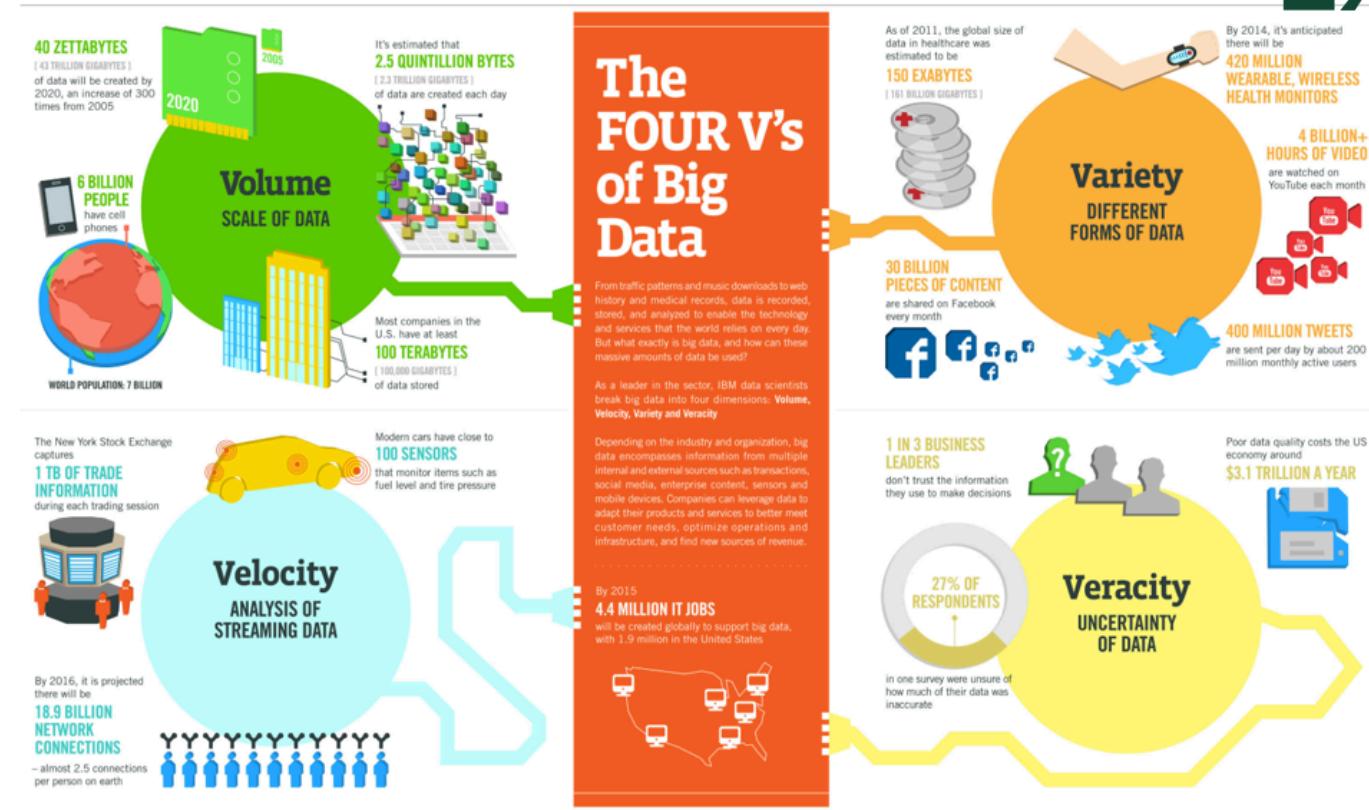
- **Innovation Trigger (Desencadenante de Innovación)**: Marca el inicio de una tecnología emergente donde las expectativas empiezan a crecer debido a anuncios y primeros prototipos.
- **Peak of Inflated Expectations (Pico de Expectativas Infladas)**: La tecnología recibe una gran atención mediática, generando entusiasmo exagerado.
- **Trough of Disillusionment (Valle de Desilusión)**: La tecnología no cumple con las expectativas iniciales, y muchas empresas abandonan sus esfuerzos. Solo los jugadores más comprometidos continúan refinando la tecnología.
- **Slope of Enlightenment (Pendiente de Iluminación)**: La comprensión sobre la utilidad real de la tecnología comienza a aumentar.
- **Plateau of Productivity (Meseta de Productividad)**: La tecnología alcanza su madurez, demostrando beneficios concretos.

Convenciones: Gartner suele incluir una estimación del tiempo que falta para que la tecnología alcance la madurez.

Definiciones tradicionales



Definiciones tradicionales



Definiciones tradicionales



Definiciones tradicionales

Big Data se asocia con:

- Grandes volúmenes de datos.
- Análisis de redes sociales.
- Datos en tiempo real.
- Datos complejos de diferentes tipos
- Se generan a gran velocidad.

Se requiere que las técnicas de analítica permitan descubrir información útil en tiempos razonables.

Herramientas de BI

26

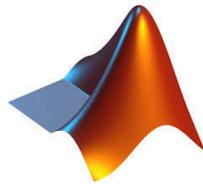
Lenguajes, librerías y herramientas



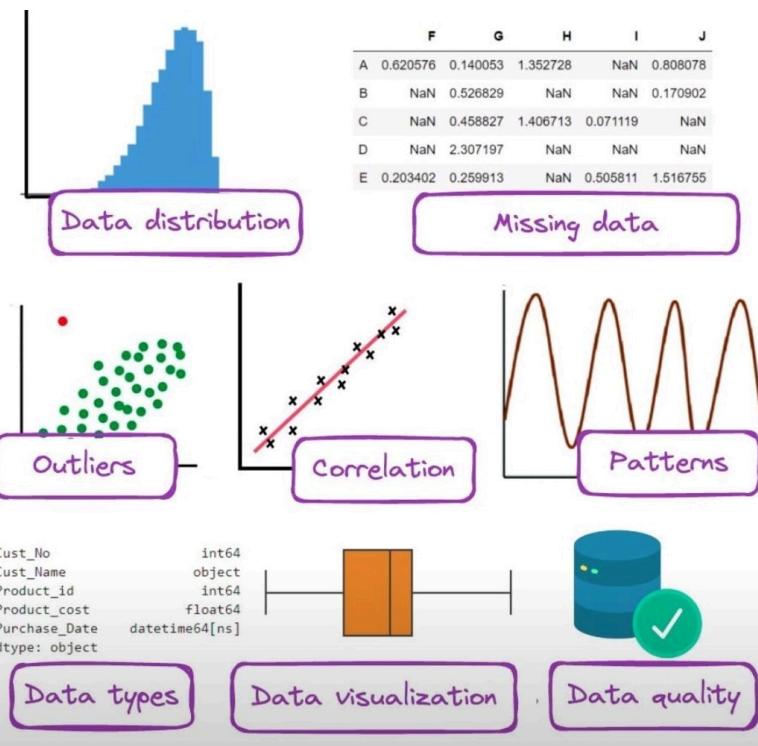
Not
Only
~~Relational~~
NoSQL

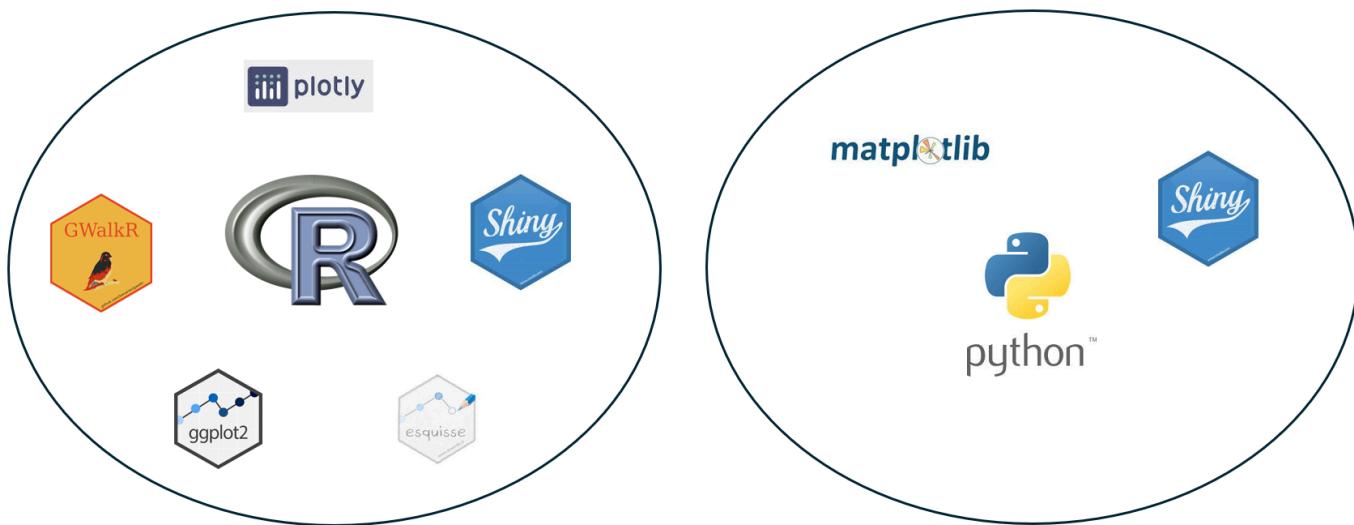


IBM SPSS
MODELER



1. Tipos de variables.
2. Visualizar los datos
3. Identificar relaciones
4. Datos atípicos
5. Datos faltantes



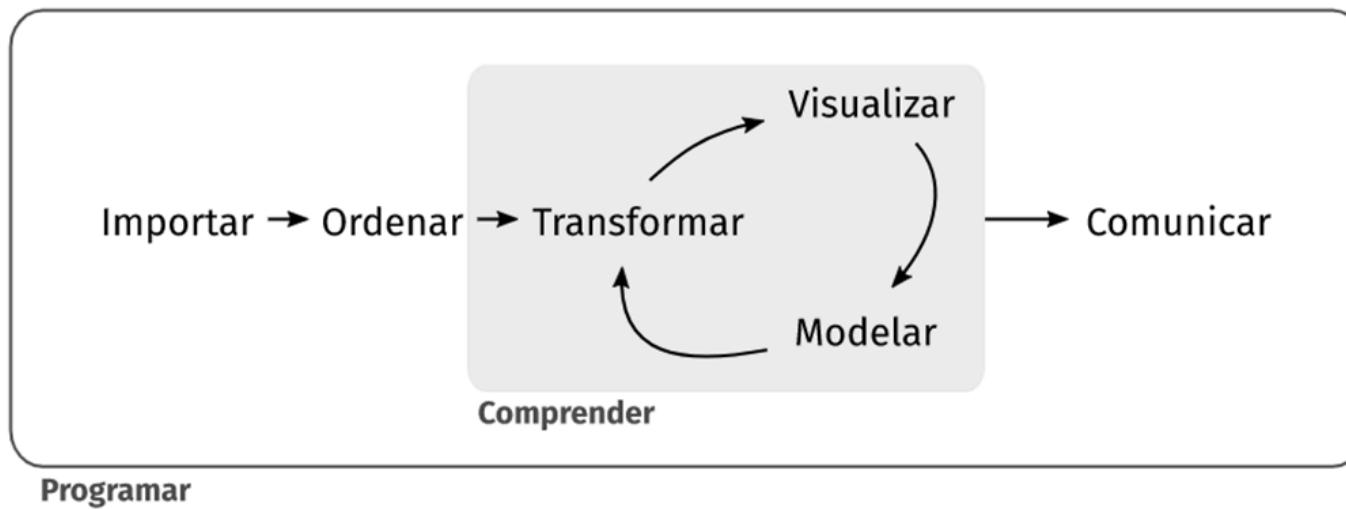


Requisitos

- Instalar R y R-Studio: [Ver video](#)
- Instalar Power BI de escritorio: [Ver video](#)
- Instalar tableau de escritorio
- Abra una cuenta en ChatGPT

Empecemos con R

31



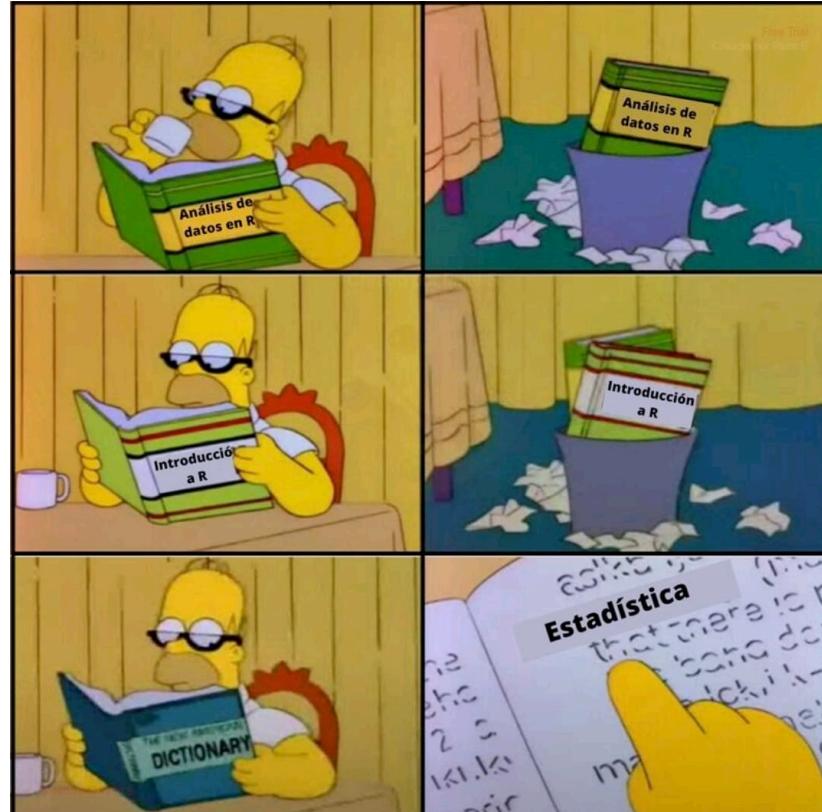
Wickham, H. y otros (2023)

Paquete R

- Paquete de uso libre
- Tiene todos los métodos que necesitamos
- Incluye paquetes específicos para la visualización como `ggplot2`, `Plotly`, `esquisse`, `tidyplots`, etc



Etapas del proceso



Paquete estadístico



Arte de Allison Horst

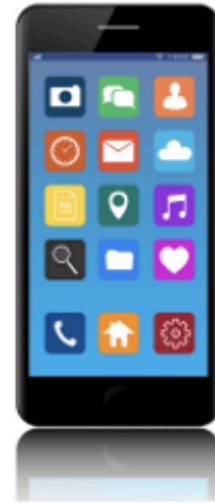
R y R-Studio



Arte de Allison Horst

Trabajando en R

R: Nuevo teléfono



Paquetes: Aplicaciones que se pueden descargar



El entorno tidyverse

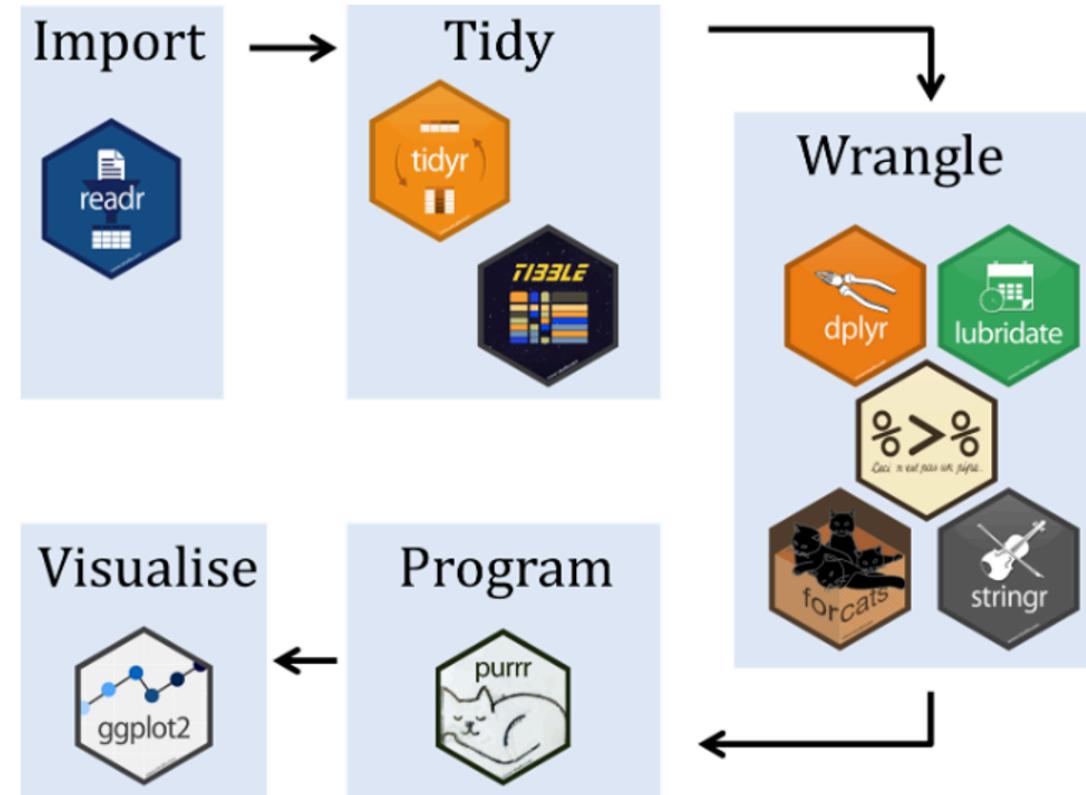


```
library(tidyverse)
```



```
library(readr)  
library(dplyr)  
library(tidyr)  
library(ggplot2)  
library(purrr)  
library(tibble)  
library(stringr)  
library(forcats)
```

Flujo de trabajo



```
1 library(pacman)
2 p_load(tidyverse, ganimate, gapminder, gifski)
3
4 datos <- gapminder |>
5     filter(year >= 1952) |>
6     mutate(poblacion = pop / 1e6)
7
8 paises_interes <- c("Colombia", "United States", "China", "India", "Nigeria")
```

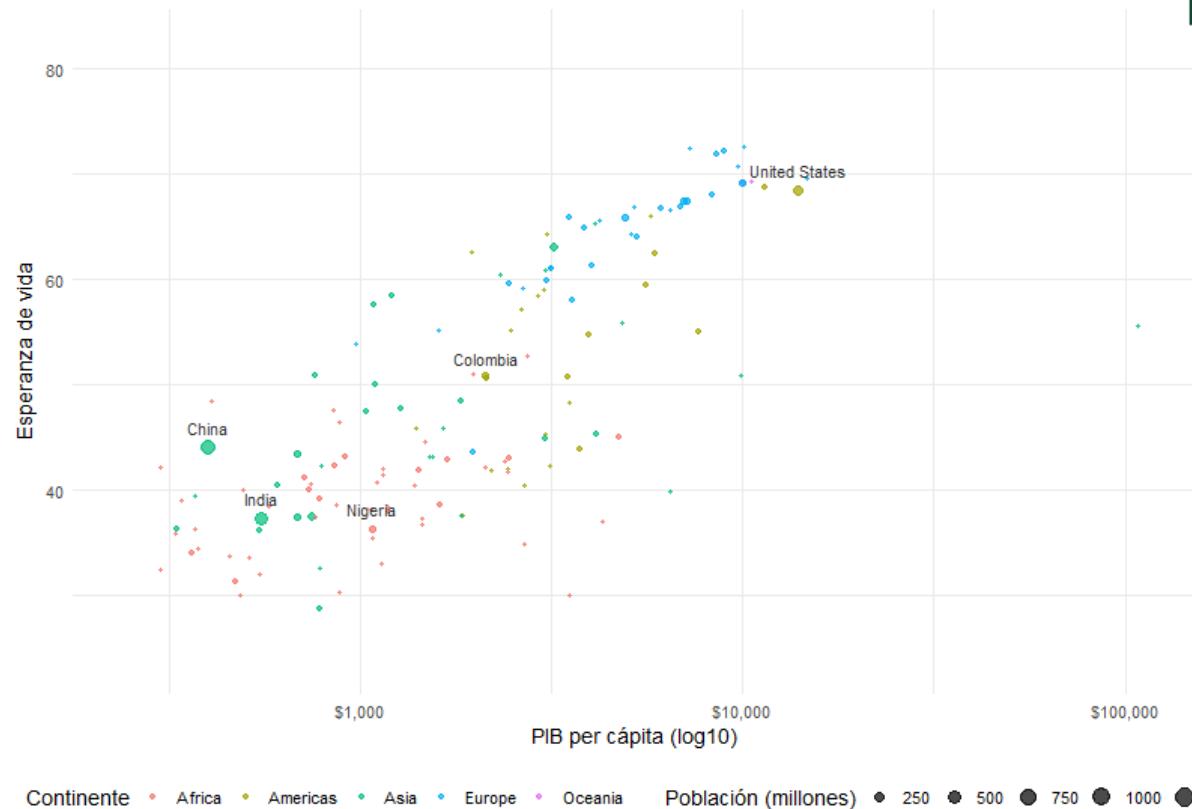
Explore el conjunto de datos

Ejemplo DataViz

```
1 plot <- ggplot(datos, aes(x = gdpPercap, y = lifeExp, size = poblacion, color = continente)) +
2   geom_point(alpha = 0.7, show.legend = TRUE) +
3   geom_text(data = datos |> filter(country %in% paises_interes),
4             aes(label = country), size = 4, vjust = -1, hjust = 0.5, color = "black", alpha = 0.8)
5   scale_x_log10(labels = scales::dollar_format(prefix = "$")) +
6   scale_y_continuous(labels = scales::comma) +
7   labs(title = "Relación entre PIB per cápita y la esperanza de vida",
8        subtitle = "Año: {frame_time}",
9        x = "PIB per cápita (log10)",
10       y = "Esperanza de vida",
11       size = "Población (millones)",
12       color = "Continente") +
13   theme_minimal(base_size = 14) +
14   theme(legend.position = "bottom") +
15   transition_time(year) +
16   ease_aes('linear')
17
18 #anim <- animate(plot, fps = 20, duration = 10, width = 800, height = 600, renderer = gifski_renderer())
19 #anim_save("gapminder_animation.gif", anim)
```

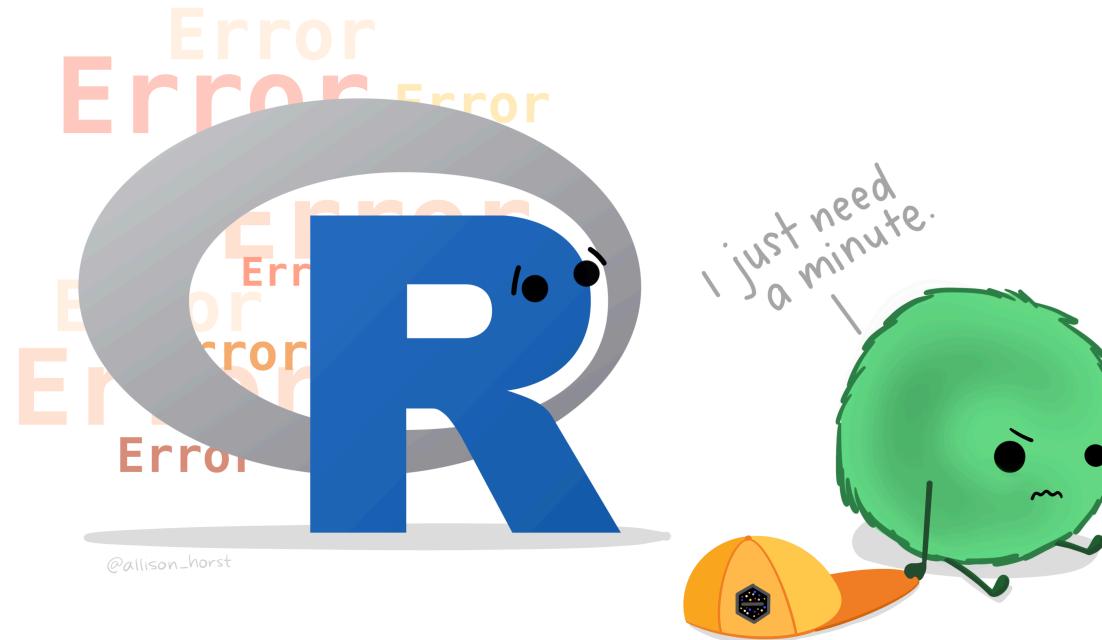
Relación entre PIB per cápita y esperanza de vida

Año: 1952



IA y Programación

La etapa de la frustración



Arte de Allison Horst

La IA como herramienta

Estamos en un mundo de constante evolución, ¿la IA nos va a reemplazar?

- Enviar una carta en papel por correo
- Pedir un domicilio por teléfono
- Solicitar un taxi por teléfono
- Orientarse en una ruta con un mapa de papel

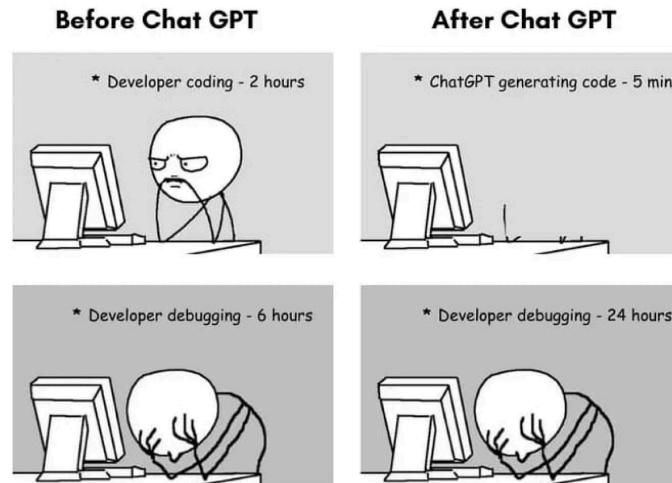


Imagen de Caracol Radio

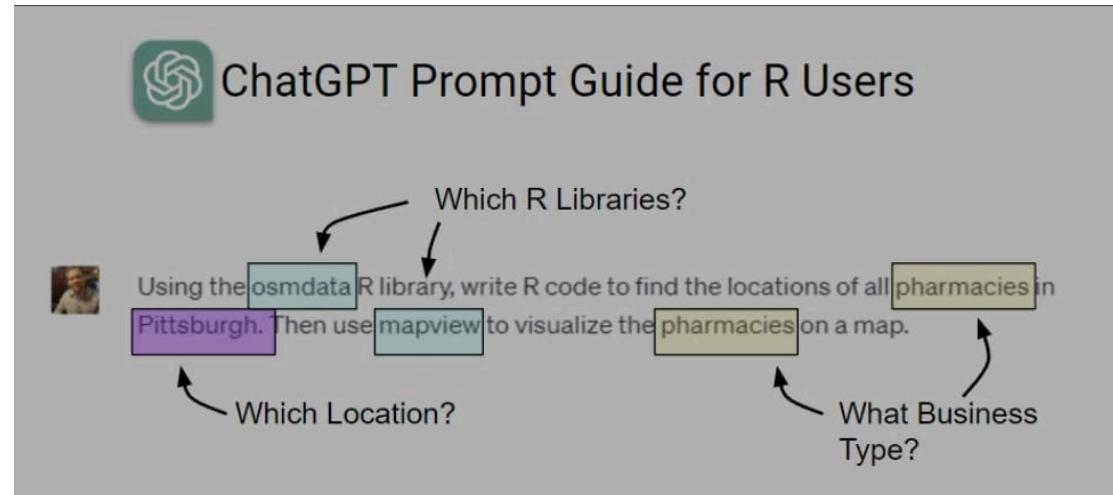
@tiangolo

GPT y Modelos LLMs

El GPT (Generative Pre-trained Transformer) es un modelo de LLM (Large Language Model). Mientras que GPT-3 usaba 175 MM de parámetros usando modelos soportados en texto, GPT-4 usa 100 BN de parámetros usando modelos soportados en texto e imágenes.



No pretenda que todo ocurra en un solo paso, a veces se obtienen mejores resultados precisando un *prompt* en cada paso.



@mdancho84

Herramientas

- <https://rtutor.ai/>
- <https://www.codeconvert.ai/r-to-python-converter>
- ChatGPT
- Copilot
- Gemini
- Claude
- GitHub Copilot
- ...

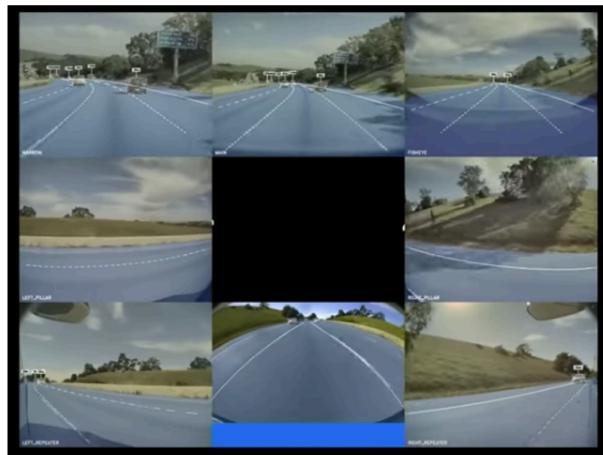
Ejemplo

Escriba el siguiente *prompt* en ChatGPT y transcriba el código a R:

“use el paquete osm de R para crear un mapa y la base de datos de droguerias en bogotá, luego use mapview para visualizar las droguerias”

Necesidades de información

Tesla Vision



Red neuronal de reconocimiento: ¿Cuál es la iguana?



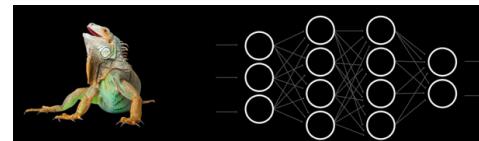
Compuesta por millones de pixeles e intensidades que permiten determinar que es una iguana.

Necesidades de información

Pero...



Iguanas hay de diferentes tipos de
diferentes condiciones...



Agilidad de entrenamiento, aprendizaje profundo, algoritmos eficientes.

Método de descenso de gradiente estocástico

Velocidad, Veracidad, Variedad, ...

Millones de imágenes...



Diferentes posiciones
o ángulos... solo para
entrenar al algoritmo



Agilidad de entrenamiento, aprendizaje profundo, algoritmos eficientes.

Método de descenso de gradiente estocástico

Variedad en el conjunto de datos... y situaciones poco comunes



Retransmite y entrena ...

Modelo TESLA

Variedad en el conjunto de datos... y situaciones poco comunes



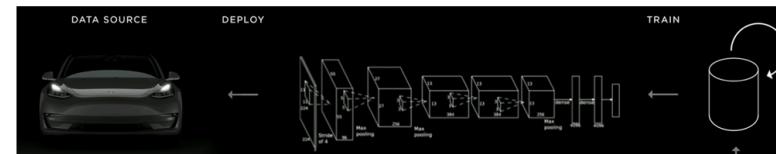
Ahora el vehículo y la bici se deben procesar como un solo objeto

Modelo TESLA

Más Variedad



El tiempo para procesar es ...



AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking

by Michael Gerlich  

Center for Strategic Corporate Foresight and Sustainability, SBS Swiss Business School, 8302 Kloten-Zurich, Switzerland

Societies 2025, 15(1), 6; <https://doi.org/10.3390/soc15010006>

Submission received: 14 October 2024 / Revised: 18 December 2024 / Accepted: 29 December 2024 / Published: 3 January 2025

Download ▾

Browse Figures

Versions Notes

El uso frecuente de herramientas de IA está determinando nuestras habilidades cognitivas, fenómeno llamado “descarga cognitiva”, este fenómeno afecta negativamente nuestra capacidad de pensamiento crítico. La investigación revela que los jóvenes de 17 a 25 años son los más afectados, en comparación con los mayores de 46 años.

- Procesos de codificación
- Encuestas telefónicas
- Observatorio nutricional
- Rating - RPD



Material recomendado

Modelos de IA generativa: <https://www.futureofai.mit.edu/>

- ChatGPT
- Stable-Diffusion & Dall-E
- Neural Networks
- Supervised Learning
- Representation & Unsupervised Learning
- Reinforcement Learning
- Generative AI ...

Material recomendado (español)

- Introduction to Generative AI (45 minutos)
- Encoder-Decoder Architecture (2.5 horas)
- Transformer Models and BERT Model (2 horas)
- Create Image Captioning Models (2 horas)
- Introduction to Vertex AI Studio (2 horas)

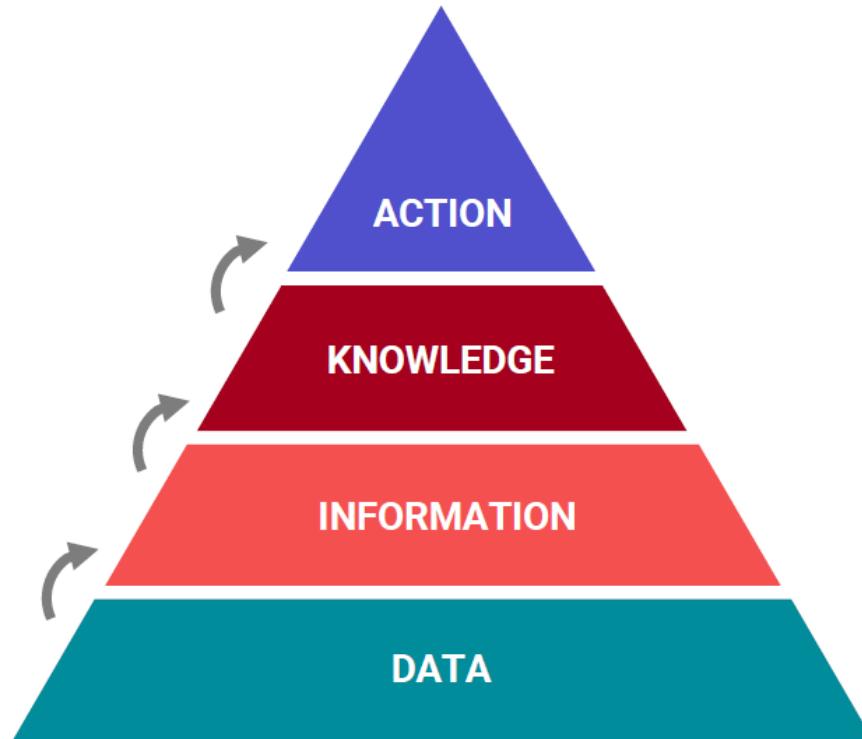
FUNDAMENTOS DEL DATA VIZ

60

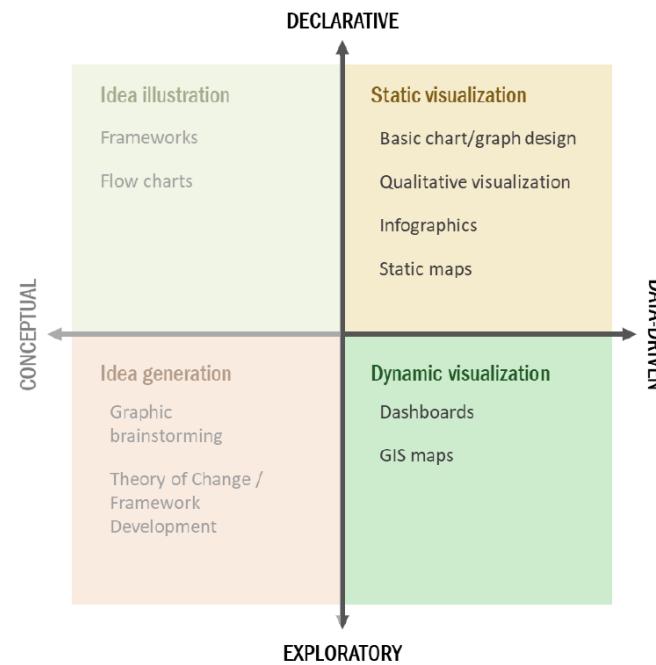
- **Tipografía:** Seleccione fuentes apropiadas, limitándose a 2-3 y siendo consistente en su uso.
- **Contraste:** Utilizar contraste (claro/oscuro, grande/pequeño, grueso/delgado) para resaltar y enfatizar.
- **Color:** Usar el color para enfatizar o reforzar valores, evitando el uso excesivo de colores que pueda distraer.
- **Balance:** Asegurar que el diseño sea visualmente equilibrado para que sea agradable y destaque la información importante.
- **Jerarquía:** Proporcione pistas visuales sobre la parte más importante del mensaje.

- **Despejar el gráfico:** Elimine líneas, bordes y marcas de verificación innecesarias para evitar distracciones.
- **Títulos y etiquetas con propósito:** Ser claro y conciso en los títulos y usar etiquetas de datos con moderación.
- **Usar el eje completo comenzando en cero** para evitar malinterpretaciones de los datos.
- **Evite gráficos de pastel** especialmente para comparaciones, ya que son difíciles de interpretar visualmente.

Lectura recomendada: [Data Visualization Principles](#)

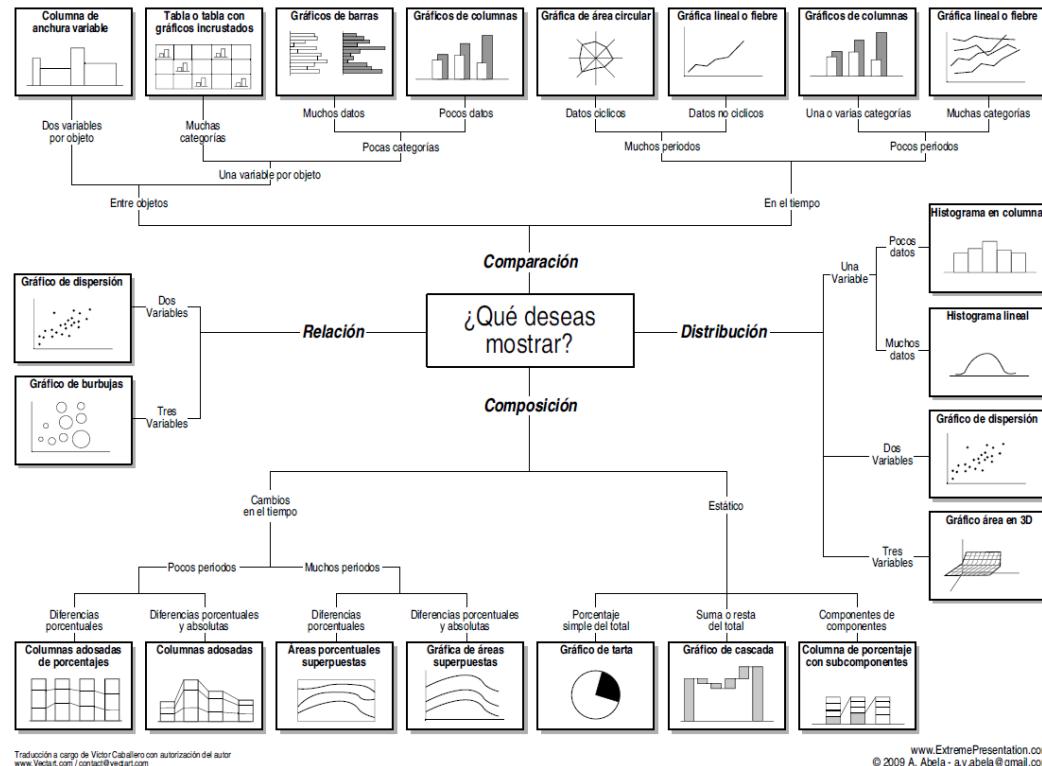


Tipos de comunicación visual



- **Eje horizontal** representa el tipo de información que se está comunicando, va desde conceptos o ideas (izq) hasta datos concretos (der).
- **Eje vertical** describe la interacción que los usuarios tienen con la visualización. Las visualizaciones declarativas cuentan una historia clara, mientras que la comunicación exploratoria permite a los usuarios interactuar con los datos para llegar a sus propias interpretaciones.

Tipos de gráficos



Galerías de gráficos en R: <https://r-graph-gallery.com/>

- Conceptos gráficos
- Dataviz Inspiration
- Galerías ggplot2
- Galerías R

Algunos ejemplos

Con el conjunto de datos artificial *study* realice un gráfico de cajas que presente el promedio de la variable *score* por cada tratamiento (asigne un color diferente a las barras según el tratamiento)

```
1 library(pacman)
2 p_load(tidyverse, tidyplots)
3
4 g1 <- study |>
5     tidyplot(x = treatment, y = score, color = treatment, width = 100, height = 100) |>
6     add_boxplot()
```

Algunos ejemplos

Ahora presente las diferencias estadísticas:

```
1 library(pacman)
2 p_load(tidyverse, tidyplots)
3
4 g1 <- study |>
5   tidyplot(x = treatment, y = score, color = treatment, width = 100, height = 100) |>
6   add_boxplot() |>
7   add_test_pvalue(ref.group = 1)
```

Algunos ejemplos

Haga un gráfico que presente el promedio del *score* por grupo y tipo de dosis

```
1 g2 <- study |>
2     tidyplot(x = group, y = score, color = dose, width = 100, height = 100) |>
3     add_mean_bar(alpha = 0.4) |>
4     add_mean_dash() |>
5     add_mean_value()
```

Business Analytics

70

Modelos para generar valor

Business Analytics (BA) es el uso de datos, análisis estadístico y modelos para apoyar la toma de decisiones estratégicas y operativas en las organizaciones. Su utilidad radica en transformar datos en insights accionables que permitan optimizar procesos, identificar oportunidades y mejorar la competitividad empresarial.



1. **Predicción:** Anticipar tendencias y comportamientos futuros (metas, mercado potencial, proyecciones).
2. **Diagnóstico:** Identificar causas subyacentes de problemas o patrones (análisis de la caída en ventas o de la satisfacción del cliente). 3 **Prescripción:** Proponer acciones óptimas para alcanzar objetivos (estrategias de precios, asignación de recursos).
3. **Descriptivo:** Resumir datos para entender comportamientos (percepción, hábitos).

Modelos de Predicción de Churn (Deserción):

- Regresión Logística: Predecir la probabilidad de que un cliente abandone.
- Análisis de Supervivencia: Identifica el tiempo hasta que un cliente abandona.
- Árboles de Decisión: Identificar reglas simples que expliquen la deserción.
- Modelos de Machine Learning: Redes neuronales, Random Forest, SVM, modelos basados en XGBoost o LightGBM.
- Análisis RFM (Recency, Frequency, Monetary): Segmentación de clientes según comportamiento histórico.

Optimización de Campañas de Marketing:

- Clustering: Segmentación de clientes para personalizar mensajes o definir estrategias.
- Modelos de Propensión: Para identificar la probabilidad de que un cliente responda a una campaña.
- A/B Testing: Evaluar la efectividad de diferentes campañas.
- Sistemas de Recomendación: Filtrado colaborativo y basado en contenido para ofertas personalizadas (busqueda en redes sociales).

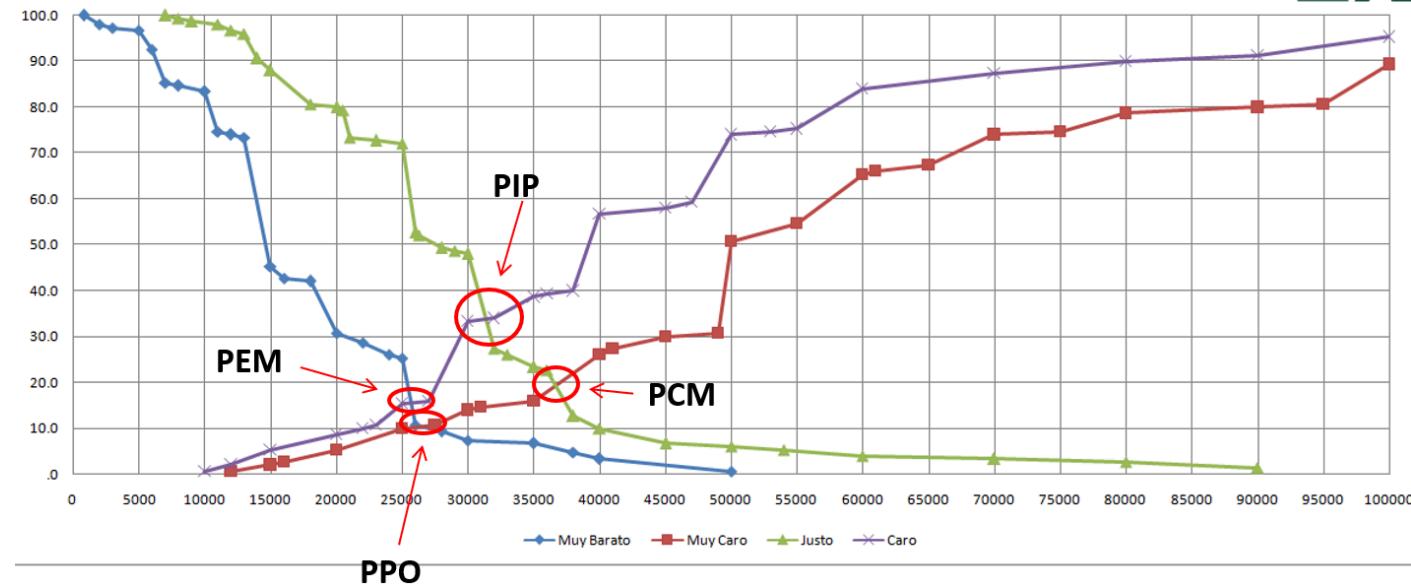
Satisfacción y Experiencia del Cliente:

- Análisis de Sentimientos (NLP): Identificar opiniones positivas o negativas en encuestas o redes sociales.
- Análisis de redes semánticas: Descubrir temas recurrentes en comentarios de clientes.
- Modelos de asociación: Relacionar la satisfacción con factores específicos.

Gestión de Precios y Estrategias de Revenue Management

- Elasticidad de Precios: Modelos Regresión Logística o Lineal para entender el impacto de cambios de precios.
- Optimización de Precios Dinámicos: Programación matemática o Machine Learning (pasajes áereos, tarifa dinámica de uber).
- Modelos de Simulación: Para predecir ingresos en diferentes escenarios (logit anidado).
- Utilidad: Maximizar ingresos mediante precios optimizados según la demanda y competencia (CBC).

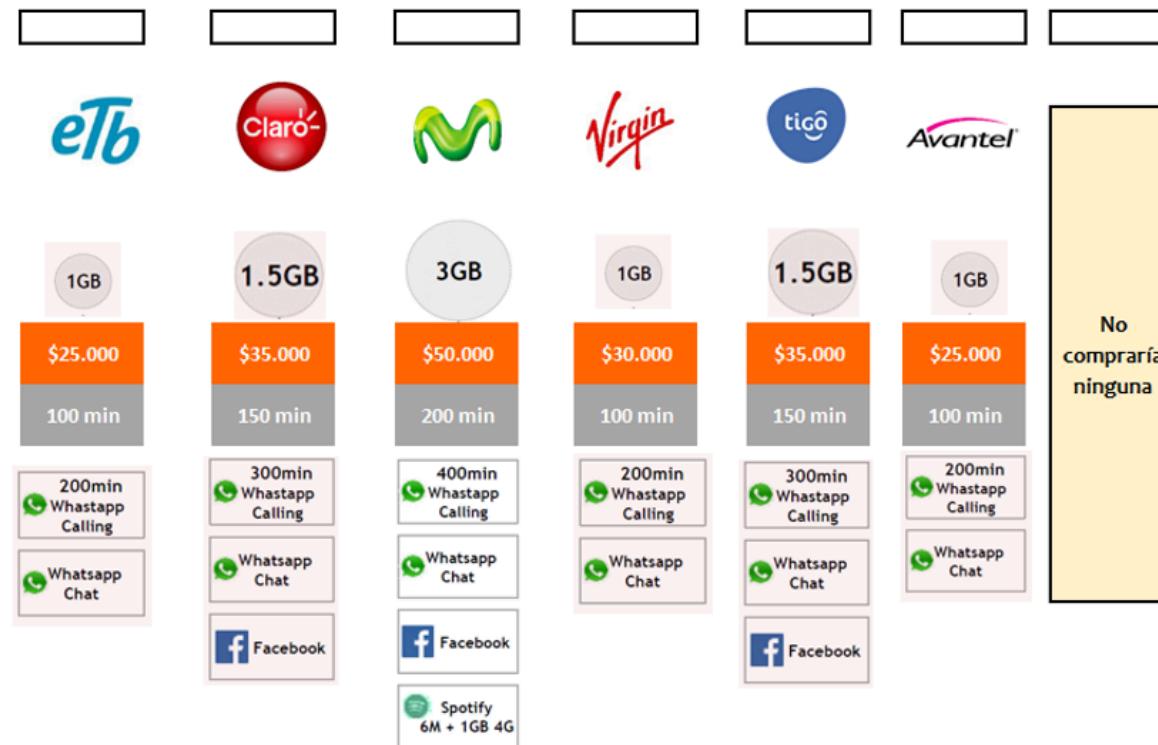
Decisiones sobre precios



Van Westendorp, Gabor-Granger, Newton-Miller-Smith, Conjoint Analysis.

Lectura recomendada: How to price products for maximum profitability

Ejemplo: Dashboard modelo de CBC



Mostrar simulador de shiny.

El medidor de sensibilidad al precio (PSM) fue propuesto por Peter Van Westendorp (1976). La técnica permite determinar el rango de precios aceptable y el precio óptimo de los productos.

Newton, Miller y Smith (1993) ampliaron el PSM con dos preguntas de seguimiento que se pueden utilizar para generar curvas de demanda de productos y curvas de ingresos.



Bombillo 100W
DURAMAX
\$?????



Bombillo 100W
SYLVANIA
\$ 11.400



Bombillo 100W
PHILIPS
\$ 13.800



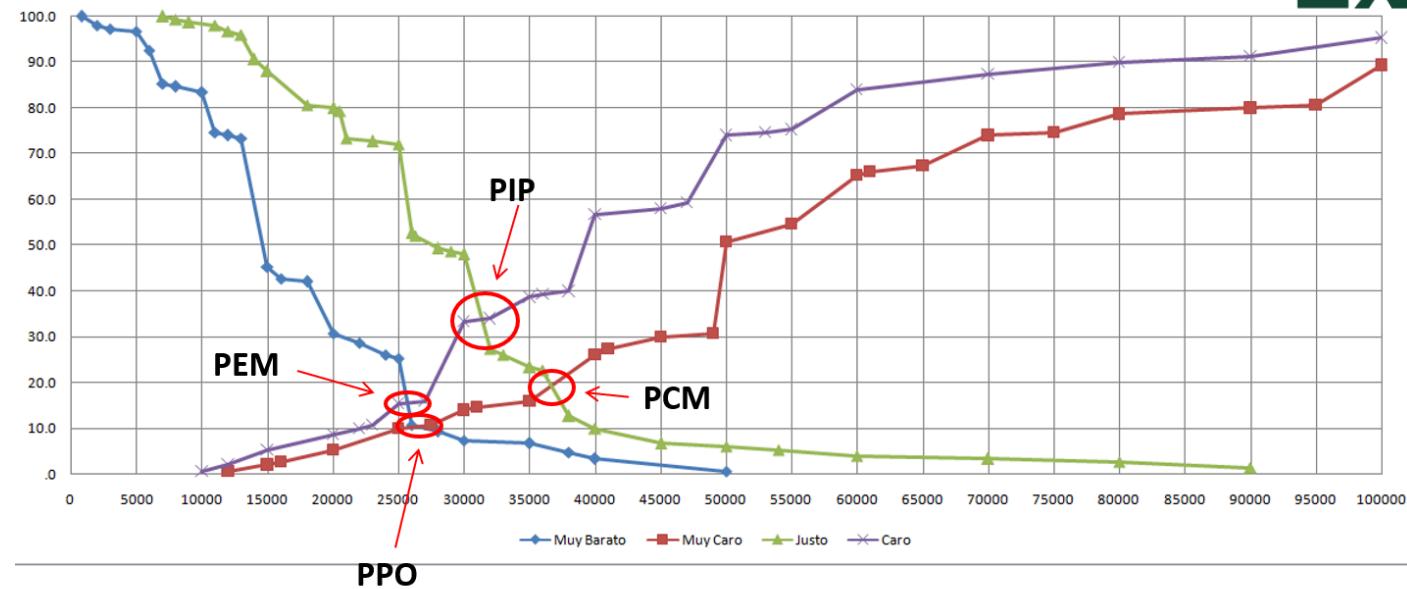
Bombillo 100W
ILUMAX
\$ 12.500

La competencia debe ser evaluada con las características más similares (W, Color de luz,...)

Aplicación



Encuesta



Punto de indiferencia al precio (PIP): Es el precio considerado como base o justo del mercado.

Punto de precio óptimo (PPO): Es el precio viable para el mercado.

Punto Económico Marginal (PEM): Es el precio más bajo que el usuario esperaría*.

Punto Caro Marginal (PCM): Es el precio máximo tolerable por el consumidor*.

PROYECTO 1: SENSIBILIDAD DE PRECIOS

- **VALOR:** 35%
- **FECHA DE ENTREGA:** Próximo 25 de enero, hasta 11:59 pm
- **FORMA DE ENTREGA:** j.babativamarquez@uniandes.edu.co
- **ASUNTO:** Proyecto Precios

BRIEF

- **Objetivo:** Determinar el rango de precio óptimo para un NUEVO producto en el mercado.
- **Target:** Consumidores de la categoría
- **Cobertura:** La debe especificar el estudiante.

PROYECTO 1: SENSIBILIDAD DE PRECIOS

- **Muestra:** Al menos 20 respuestas (Nota máx. de 4.0). Por cada encuesta adicional se sumará 0.1.
- **Preguntas (Análisis):** Mínimo agrado, intención de compra y las 4 de precios.
- **Resultados esperados:** Presentación comercial (Power Point, canvas, Beamer, etc.) con ficha técnica de la investigación (Investigar), contextualización del producto, gráfico VW, rango de precios y conclusiones con recomendaciones. Debe entregar la Base de datos y el enlace con el que se realizó la encuesta.

*La estética y forma de presentación tienen una ponderación en la nota



ANÁLISIS DE SENSIBILIDAD DE PRECIO



- Este producto tiene como finalidad satisfacer las necesidades de las personas que pasan más tiempo fuera de sus casas, pues le ofrece mayor comodidad y facilidad a la hora de la limpieza oral.

PRODUCTO

Es un Cepillo de dientes que le brinda 2 funciones en un solo producto. (incluye la crema y la seda dental).

- ✓ Tiene un diseño atractivo y práctico
- ✓ Es multiusos, fácil de adquirir y de llevar.
- ✓ Limpieza rápida y segura.



87

Estructuras de datos

88

Proceso de ETL: Extracción, Transformación y Carga



Bases de datos relacionales

Estructura de datos que cumple el modelo entidad-relación (objeto-atributos).

Es un conjunto de tablas estructuradas en registros (filas) y campos (variables, columnas), que se vinculan entre sí por un campo en común (índice).

El software que gestiona la BD se denomina SGDBR-Sistema Gestor de la Base de Datos Relacional(RDBMS -Relational Data Base Management System). Ej.: PostgreSQL, Oracle, MySQL, DB2, SQL Server, Access...

Bases de datos relacionales

a	m	m
b	m	m
c	m	m

+

a	1
b	2
c	3

→

a	m	m	a	1
b	m	m	b	2
c	m	m	c	3

Relación uno a uno. Cada registro de la tabla 1 es pegado con un solo registro de la tabla 2.

a	m	m
b	m	m
c	m	m

+

a	1
b	2
c	3
c	4

→

a	m	m	a	1
b	m	m	b	2
c	m	m	c	3
c	m	m	c	4

Relación uno a muchos. Cada registro de la tabla 1 es pegado con varias filas de la tabla 2.

a	m	m
b	m	m
c	m	m
c	m	m

+

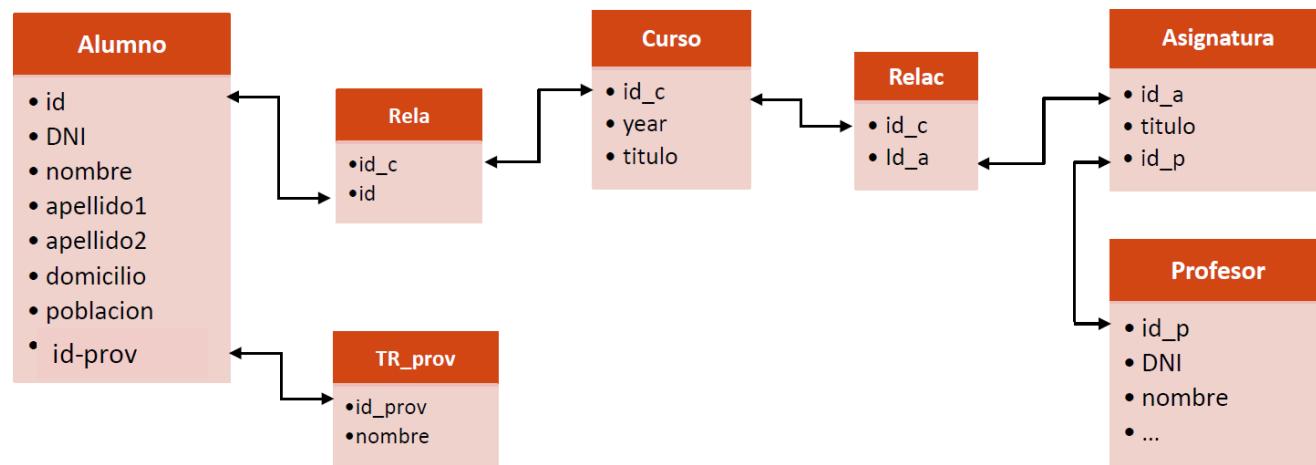
a	1
b	2
c	3
c	4

→

a	m	m	a	1
b	m	m	b	2
c	m	m	c	3
c	m	m	c	4
c	m	m	c	3
c	m	m	c	4

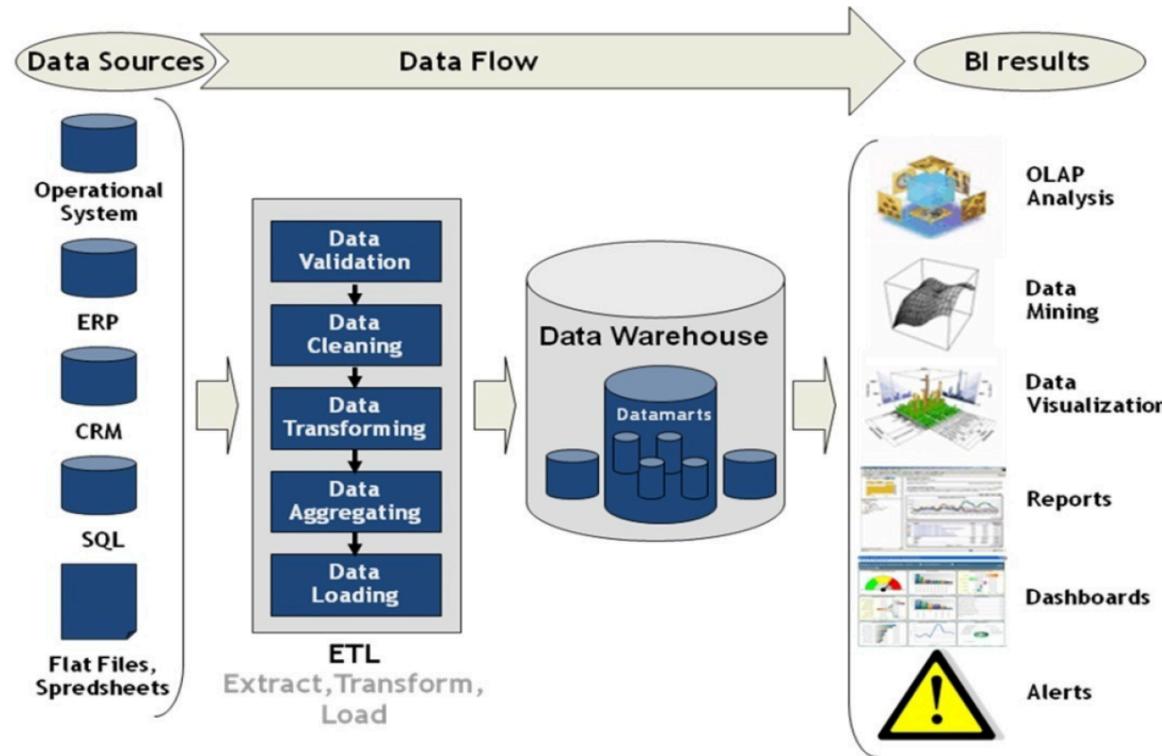
Relación muchos a muchos. Varios registros de la tabla 1 es pegado con varias filas de la tabla 2. Cada fila de la tabla 1 es pegada con todos los coincidentes en la tabla 2.

Bases de datos relacionales



```
SELECT Alumno.DNI, Alumno.nombre, Alumno.apellido1, Curso.titulo
FROM Alumno
INNER JOIN Rela ON Alumno.id = Rela.id
INNER JOIN Curso ON Rela.id_c = Curso.id_c
```

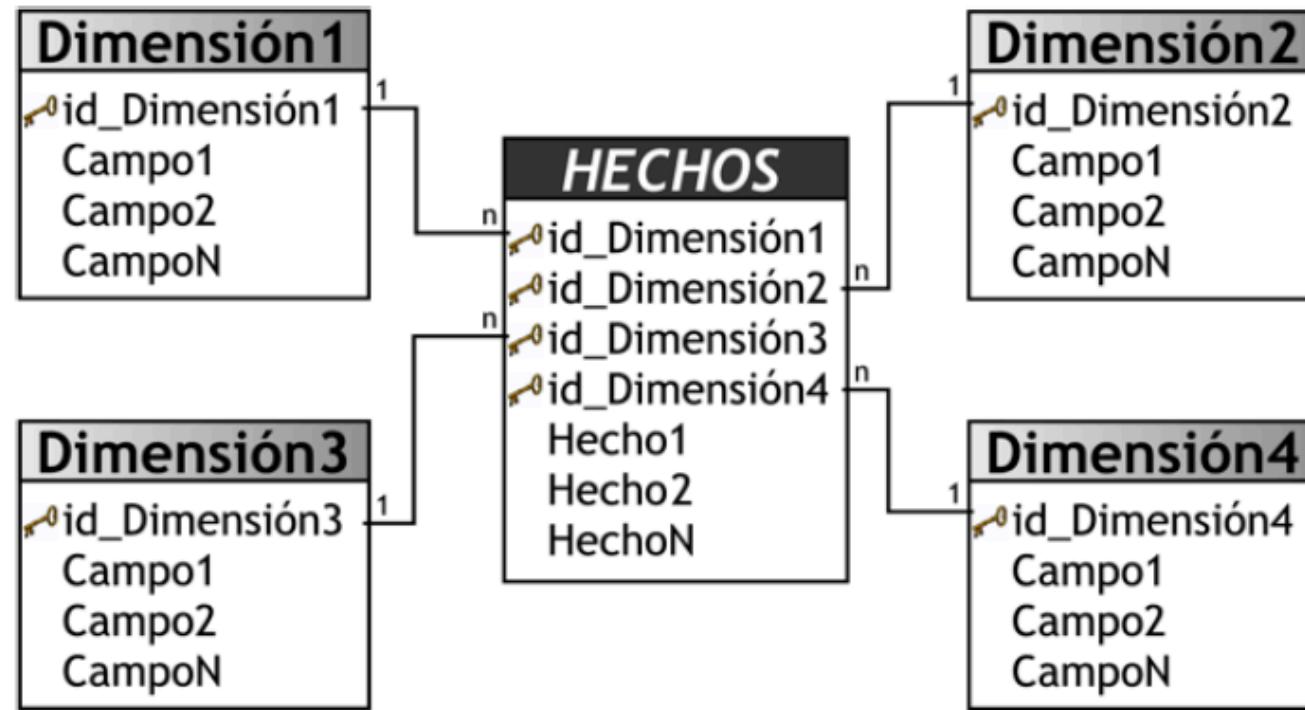
- **Tabla de datos:** Estructura de datos de dos dimensiones.
- **Base de datos:** Conjunto de tablas de datos que generalmente pertenecen al mismo contexto
- **Cubo OLAP:** Base de datos multidimensional para análisis interactivo.
- **Bases de datos NoSQL:** Sistemas de almacenamiento que no cumplen el modelo entidad-relación.



[Ver artículo](#)

Construcción de un DW

Metadatos: recoge la información de toda la estructura y datos



Dato vs Información

El DW es el motor del **Business Intelligence (BI)** y está diseñado para favorecer el análisis y la divulgación eficiente de datos.

El BI usa los datos de una empresa para facilitar la toma de decisiones.

DATOS -> INFORMACIÓN -> CONOCIMIENTO

Características de un BI:

- Accesibilidad a la información
- Apoyo a la toma de decisiones
- Orientación al usuario final

Aplicación: Marketing Directo

¿Qué significa?

Es la comunicación a través de algunos medios con el fin de generar una reacción.



Aplicación: Marketing Directo

Objetivos

Ganar clientes y fomentar la fidelidad de los mismos. El objetivo de la fidelidad de los clientes puede estar dirigido a que repitan la compra o a mantener la adquisición permanente de un producto.

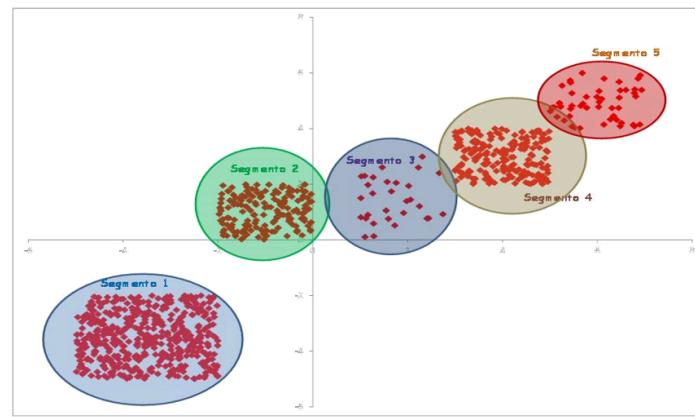
Metodología

Hacer un continuo seguimiento de los clientes con el fin de detectar y responder: Cómo están agrupados?. Quiénes son?. Donde están? Cuáles son los mejores? Alarmas sobre quiénes están en riesgo, etc..

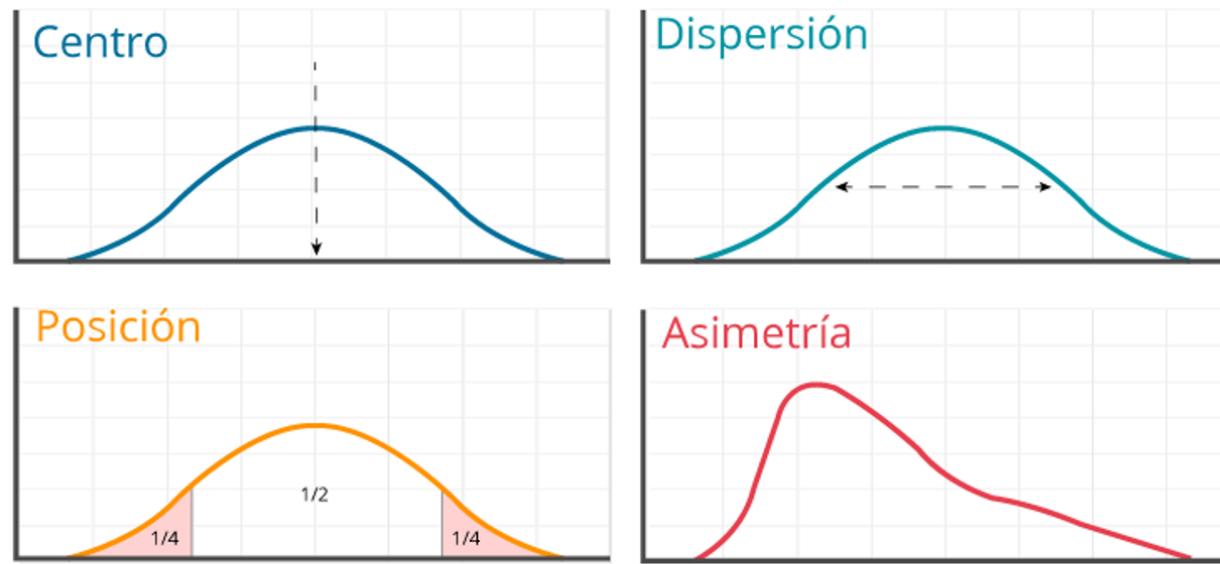
Segmentación

Es el proceso de dividir una población en grupos más pequeños que tengan características y necesidades semejantes dentro de ellos y distintas entre ellos.

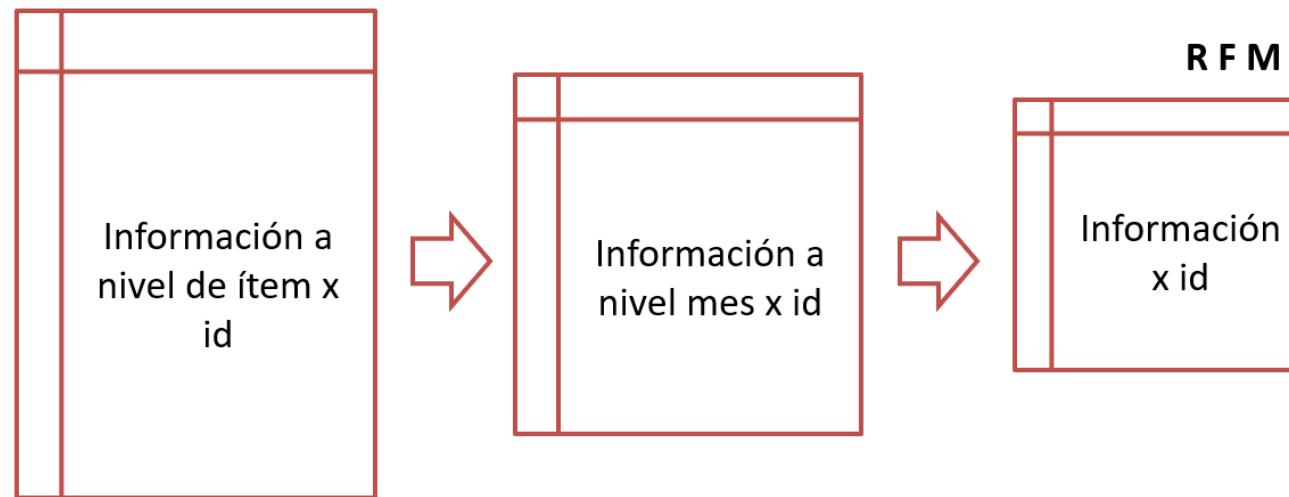
Se identifican grupos con alto nivel de relación para definir un plan de contactos adecuado, con el mensaje correcto, a través de canales precisos, a los clientes adecuados.



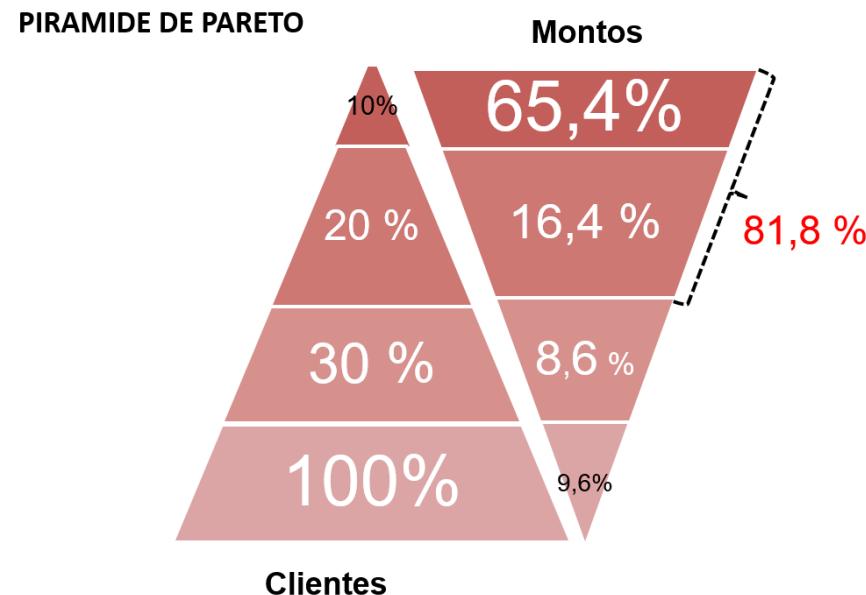
Resumen de variables cuantitativas



Nuestros datos



Segmentación - Marketing Directo



Identifique el nivel de riesgo de su negocio ¿de cuántos depende hoy?. Implemente estrategias para cuidar esos clientes importantes. Identifique, ¿Quiénes son?, ¿Dónde están?.

Segmentación RFM

(i)	Semanas procesadas	24
(j)	Corte del periodo	17-jun

Id	Nombre	Ultima Transacción	Total de visitas en 24 semanas	Total de compras en 24 semanas (Miles de \$)	Venta Reciente (Días) = j-b	Prom. Visitas por semana = c/i	Prom. Montos por visita = d/c	Prom. Montos por semana = d/i	Perc.Rec	Perc.Fr	Per.Mon	R	F	M
									(A)	(D)	(D)	R	F	M
1	Karen	17-jun	78	295,2	0	3,25	3,78	12,3	1	1	8	1	1	2
2	Felipe	02-jun	38	1211,52	15	1,58	31,88	50,48	5	5	4	2	2	2
3	Andrés	30-may	60	1256,16	18	2,50	20,94	52,34	6	2	3	2	1	1
4	Juanita	16-jun	15	1488,24	1	0,63	99,22	62,01	2	8	2	1	2	1
5	Daniel	15-jun	19	674,64	2	0,79	35,51	28,11	3	7	6	1	2	2
6	Alex	02-feb	2	289,92	135	0,08	144,96	12,08	10	10	9	2	2	2
7	Patricia	29-mar	15	549,84	80	0,63	36,66	22,91	9	9	7	2	2	2
8	Daniela	18-may	58	1090,56	30	2,42	18,80	45,44	7	3	5	2	1	2
9	Jeisson	14-jun	47	1501,92	3	1,96	31,96	62,58	4	4	1	2	2	1
10	William	11-abr	26	240	67	1,08	9,23	10	8	6	10	2	2	2

Adaptado de Woolf, B (2002). Loyalty Marketing: The Second Act

Segmentación RFM

(i)	Semanas procesadas	24
(j)	Corte del periodo	17-jun

30-70

Id	Nombre	Última Transacción	Total de visitas en 24 semanas	Total de compras en 24 semanas (Miles de \$)	Venta Reciente (Días) = j-b	Prom. Visitas por semana = c/i	Prom. Montos por visita = d/c	Prom. Montos por semana = d/i	Perc.Rec	Perc.Fr	Per.Mon	R	F	M
									(A)	(D)	(D)	R	F	M
1	Karen	17-jun	78	295,2	0	3,25	3,78	12,3	1	1	8	1	1	2
2	Felipe	02-jun	38	1211,52	15	1,58	31,88	50,48	5	5	4	2	2	2
3	Andrés	30-may	60	1256,16	18	2,50	20,94	52,34	6	2	3	2	1	1
4	Juanita	16-jun	15	1488,24	1	0,63	99,22	62,01	2	8	2	1	2	1
5	Daniel	15-jun	19	674,64	2	0,79	35,51	28,11	3	7	6	1	2	2
6	Alex	02-feb	2	289,92	135	0,08	144,96	12,08	10	10	9	2	2	2
7	Patricia	29-mar	15	549,84	80	0,63	36,66	22,91	9	9	7	2	2	2
8	Daniela	18-may	58	1090,56	30	2,42	18,80	45,44	7	3	5	2	1	2
9	Jeisson	14-jun	47	1501,92	3	1,96	31,96	62,58	4	4	1	2	2	1
10	William	11-abr	26	240	67	1,08	9,23	10	8	6	10	2	2	2

Adaptado de Woolf, B (2002). Loyalty Marketing: The Second Act

Segmentación RFM

(i)	Semanas procesadas	24
(j)	Corte del periodo	17-jun

30-70

Id	Nombre	Ultima Transacción	Total de visitas en 24 semanas	Total de compras en 24 semanas (Miles de \$)	Venta Reciente (Días) = j-b	Prom. Visitas por semana = c/i	Prom. Montos por visita = d/c	Prom. Montos por semana = d/i	Perc.Rec	Perc.Fr	Per.Mon	R	F	M
									(A)	(D)	(D)	R	F	M
1	Karen	17-jun	78	295,2	0	3,25	3,78	12,3	1	1	8	1	1	2
2	Felipe	02-jun	38	1211,52	15	1,58	31,88	50,48	5	5	4	2	2	2
3	Andrés	30-may	60	1256,16	18	2,50	20,94	52,34	6	2	3	2	1	1
4	Juanita	16-jun	15	1488,24	1	0,63	99,22	62,01	2	8	2	1	2	1
5	Daniel	15-jun	19	674,64	2	0,79	35,51	28,11	3	7	6	1	2	2
6	Alex	02-feb	2	289,92	135	0,08	144,96	12,08	10	10	9	2	2	2
7	Patricia	29-mar	15	549,84	80	0,63	36,66	22,91	9	9	7	2	2	2
8	Daniela	18-may	58	1090,56	30	2,42	18,80	45,44	7	3	5	2	1	2
9	Jeisson	14-jun	47	1501,92	3	1,96	31,96	62,58	4	4	1	2	2	1
10	William	11-abr	26	240	67	1,08	9,23	10	8	6	10	2	2	2

Adaptado de Woolf, B (2002). Loyalty Marketing: The Second Act

Interpretación

- ¿Qué significa que el resultado sea 111?
- ¿Qué significa que el resultado sea 211?
- ¿Cuál debería ser la estrategia de retención en cada segmento?

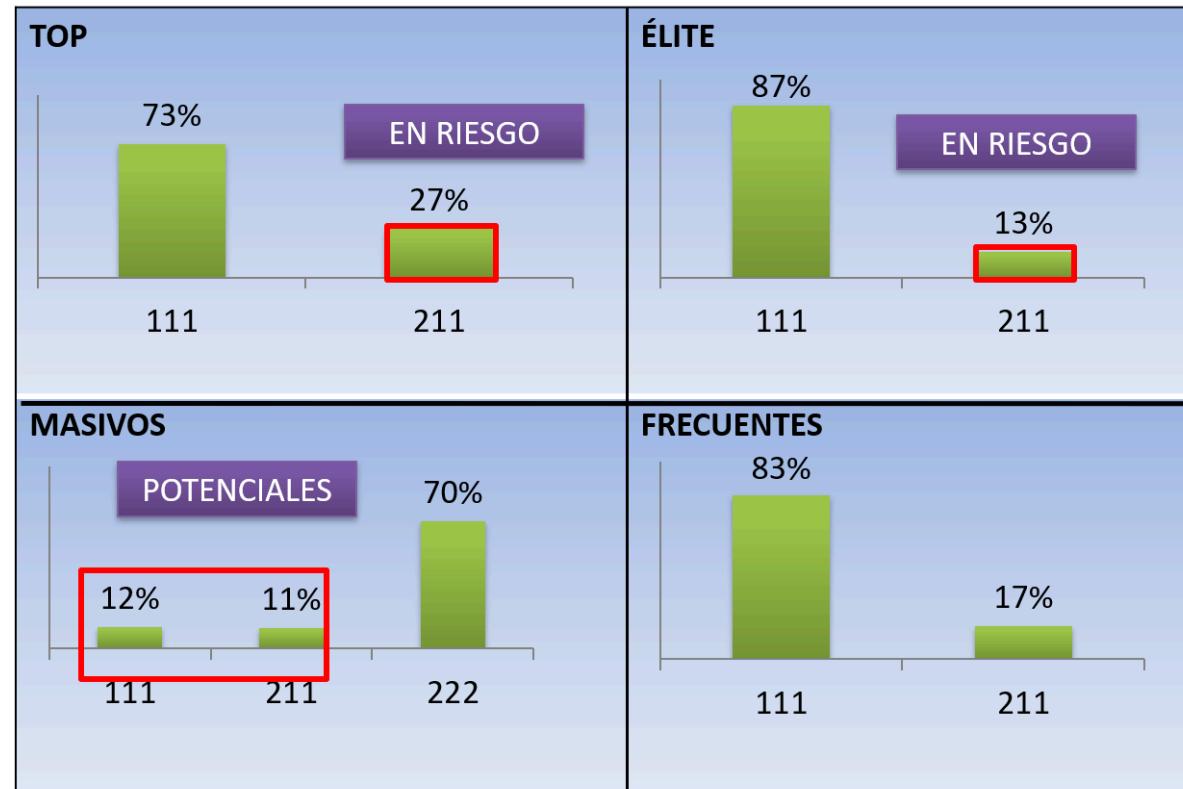
03:00

106

Visualización: Mapa de segmentación



Visualización: Mapa de acción



108

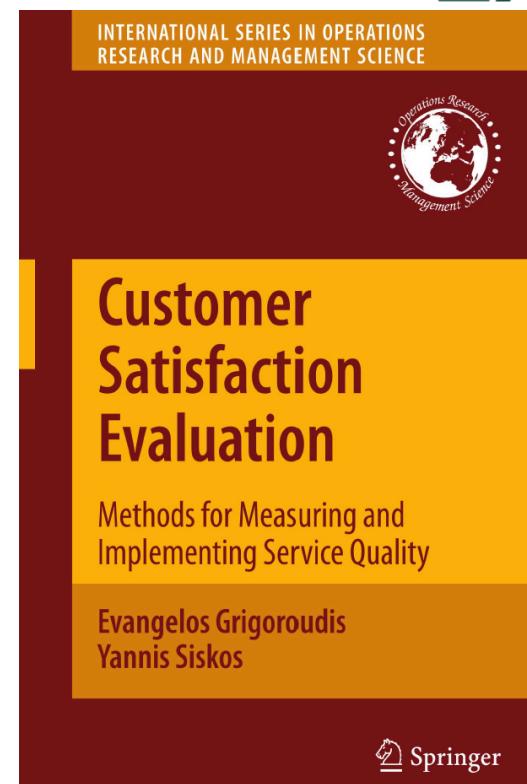
Tu turno: Use los conjuntos de datos “rfm_transacciones.xlsx” y “rfm_ejemplo.xlsx” para hacer una segmentación RFM. Discuta sobre las estrategias que aplicaría.

15:00

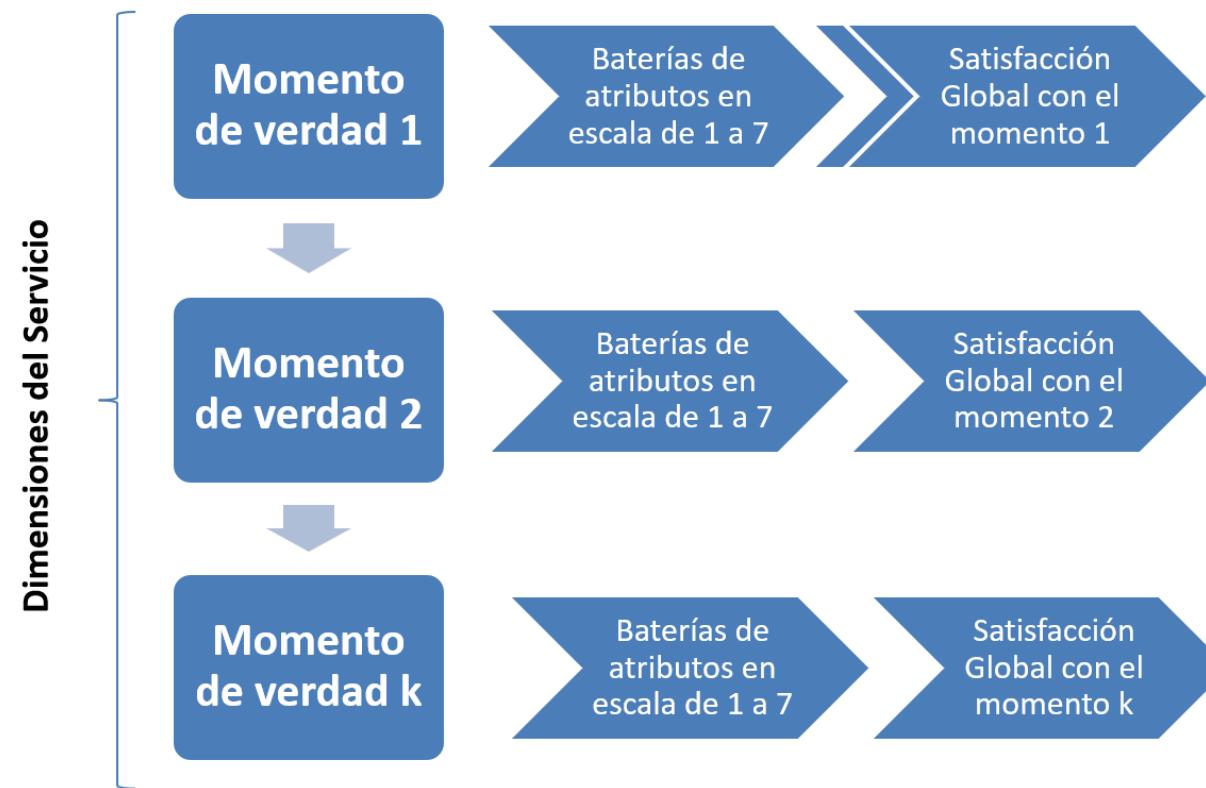
109

Satisfacción y Experiencia del Cliente

Los modelos de satisfacción y lealtad son herramientas analíticas utilizadas para comprender y predecir el comportamiento de los clientes, evaluando cómo diferentes factores influyen en su satisfacción con un producto o servicio y en su disposición a permanecer fieles a una marca o empresa.



Esquema de medición



111

Matriz de acción

DESEMPEÑO	ALTO	MANTENER	ATRIBUTOS FUERTES - APROVECHAR OPORTUNIDAD
		(Oportunidad de transferir recursos sin perder rendimiento)	(Reforzar y Mantener los esfuerzos realizados)
BAJO	MEJORAR AL LARGO PLAZO	ACCIONAR DE FORMA INMEDIATA	(Atributos que requieren de atención prioritaria)
	(No se requiere una acción inmediata, son de mejora secundaria)	BAJA	ALTA

- 1. Mejorar al largo plazo:** Son atributos con un desempeño bajo y percibidos como de baja importancia a la hora de evaluar la satisfacción general.
- 2. Mantener:** Son atributos que están con desempeño alto y que no fueron percibidos de forma importante al evaluar la satisfacción general.
- 3. Atributos Fuertes:** Son aquellos que fueron evaluados con un desempeño alto y son percibidos importantes al evaluar la satisfacción general, pueden ser usados como ventaja competitiva.
- 4. Accionar de forma Inmediata:** Son los que necesitan la mayor atención, fueron percibidos importantes al evaluar la satisfacción general pero la Compañía tiene un desempeño bajo.

Formulación del modelo

Sea $\mathcal{D} = \{(y_i, \mathbf{x}_i) : i = 1, \dots, n\}$, con y_i la i -ésima respuesta medida en una escala continua; $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^t \in \mathbb{R}^p$ es el vector de variables predictoras; y n ($\gg p$) es el tamaño de la muestra. El modelo lineal se especifica así:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i \text{ con } \varepsilon_i \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$$

Tu turno: Usando la base Satisfaction

1. Identifique los momentos del servicio
2. Identifique los atributos
3. Haga un modelo en cascada
 - Modelo general global con las generales de las dimensiones
 - En cada dimensión su general con los respectivos atributos.
4. Realice las matrices de acción.

INTERPRETE

La lectura del conjunto de datos se puede realizar directamente desde el repositorio de [GitHub](#):

```
1 library(pacman)
2 p_load(tidyverse, haven, broom, ggrepel, skimr)
3
4 url <- "https://github.com/jgbabativam/BI/raw/main/datos/Satisfaction.sav"
5
6 datos <- read_sav(url)
```

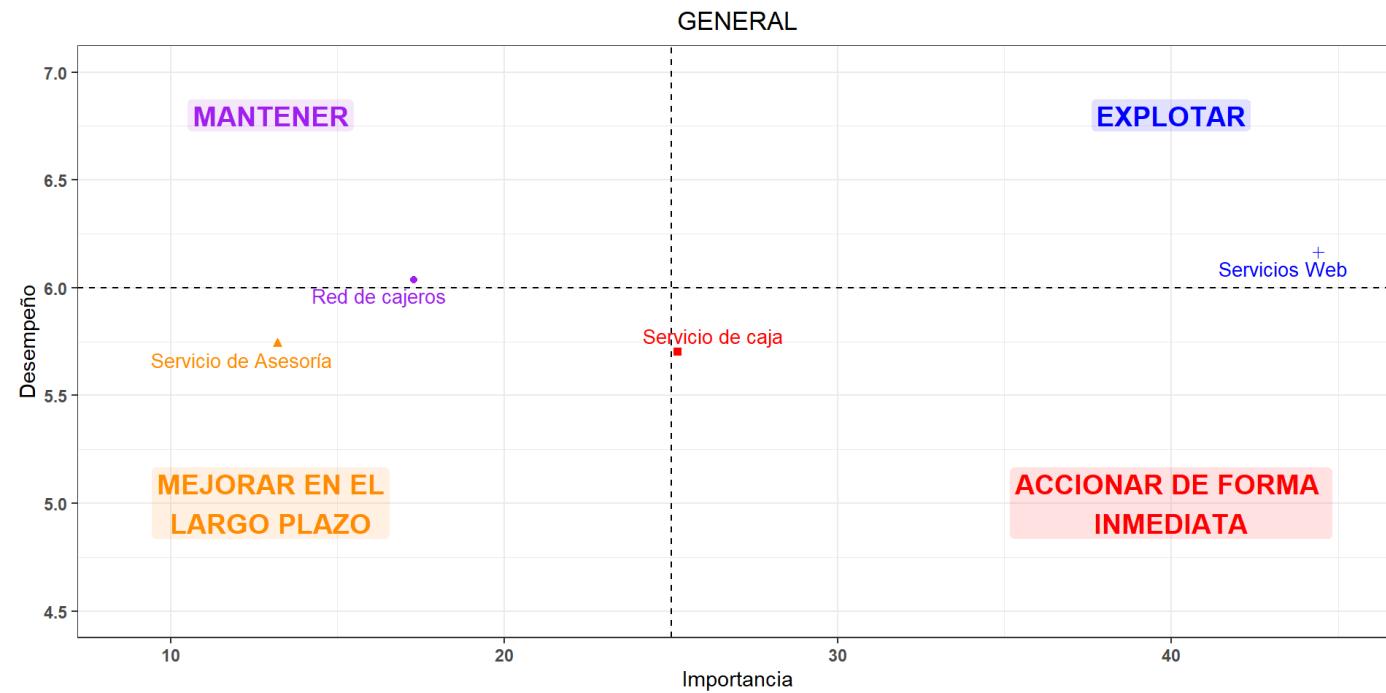
Use los comandos `skim()` y `glimpse()` para explorar el conjunto de datos.

```
1 desempe <- datos |>
2     select(Pp1f, Pp2e, Pp3h, Pp4e) |>
3     summarise(across(everything(), ~mean(.))) |>
4     pivot_longer(everything(), names_to = "term", values_to = "desempeño")
5
6 general <- lm(Pp5 ~ Pp1f + Pp2e + Pp3h + Pp4e - 1,
7                 data = datos) |>
8     tidy() |>
9     mutate(score = estimate * (1 - p.value)) |>
10    mutate(importancia = round(100 * score / sum(score), 1)) |>
11    left_join(desempe, by = "term") |>
12    select(term, importancia, desempeño)
13
14 res <- general |>
15     mutate(impacto = factor(case_when(
16         importancia < mean(importancia) & desempeño > 6 ~ 1,
17         importancia < mean(importancia) & desempeño <= 6 ~ 2,
18         importancia >= mean(importancia) & desempeño <= 6 ~ 3,
19         importancia >= mean(importancia) & desempeño > 6 ~ 4),
20         levels = 1:4, labels = c("Mantener", "Mejorar en el largo plazo",
21                               "Acción Inmediata", "Explotar")))) |>
```

Matriz de acción

```
1 paletaAEML <- c("Acción Inmediata"= "red",
2                     "Explotar"= "blue",
3                     "Mantener"= "purple",
4                     "Mejorar en el largo plazo"="darkorange")
5
6
7 res |>
8 ggplot(aes(x= importancia, y= desempeño)) +
9 geom_point(aes(shape=impacto, color=impacto)) +
10 labs(x="Importancia", y="Desempeño", title = "GENERAL") + ylim(4.5,7) + xlim(9, 45) +
11 geom_vline(xintercept = mean(res$importancia), color = "black", linetype = 2) +
12 geom_hline(yintercept = 6.0, color = "black", linetype = 2) +
13 geom_text_repel(aes(label = str_wrap(names, 30), colour = impacto), size=3.5, segment.color = "grey50")
14 scale_colour_manual(values = paletaAEML) +
15 theme_bw() +
16 theme(axis.text = element_text(face = "bold"), legend.position = "none") +
17 theme(plot.title = element_text(hjust = 0.5), plot.subtitle = element_text(hjust = 0.5)) +
18 annotate("label", x= 40, y = 5.0, label = "ACCIONAR DE FORMA \nINMEDIATA", size = 5, color = "red", label.size = 0, label.padding = 0)
19 annotate("label", x= 13, y = 5.0, label = "MEJORAR EN EL\nLARGO PLAZO", size = 5, color = "darkorange", label.size = 0, label.padding = 0)
20 annotate("label", x= 13, y= 6.8, label = "MANTENER", size = 5, color = "purple", label.size = 0, label.padding = 0)
21 annotate("label", x= 40, y= 6.8, label = "EXPLOTAR", size = 5, color = "blue", label.size = 0, label.padding = 0)
```

Matriz de acción



PROYECTO 2: SATISFACCIÓN

- **VALOR:** 35%
- **FECHA DE ENTREGA:** 29 de enero.
- **FORMA DE ENTREGA:** Correo electrónico
- **ASUNTO:** Trabajo 2 BI. Satisfacción

BRIEF

- **Objetivo:** Determinar la importancia de los momentos de verdad y atributos dentro de un target de clientes de un servicio.
- **Target:** Clientes del servicio
- **Cobertura:** La debe especificar el estudiante.

PROYECTO 2: SATISFACCIÓN

- **Muestra:** Al menos 30 respuestas (Calificación de 4.0). Por cada respuesta adicional 0.2 hasta completar una nota de máximo 10.
- **Momentos:** Medir al menos 5 aspectos generales o del mismo momento de verdad.
- **Resultados esperados:** Presentación comercial con: ficha técnica, link del instrumento, Explicación del ciclo del servicio, Importancia de los momentos/atributos, Desempeño del modelo, Matriz de acción. Base de datos. Conclusiones y recomendaciones.

*La estética y forma de presentación tienen una ponderación en la nota

Videos clase: Modelo de satisfacción

- Parte 1
- Parte 2
- Parte 3

GRACIAS!

122

Referencias

- Howson, C. (2013). Successful Business Intelligence, Second Edition: Unlock the Value of BI & Big Data. 2d Edition. McGraw Hill.
- Sherman, R. (2014). Business Intelligence Guidebook: From Data Integration to Analytics. Morgan Kaufmann
- Sinha, C. (2024). Mastering Power BI: Build business intelligence applications powered with DAX calculations, insightful visualizations, advanced BI techniques, and loads of data sources. BPB Publications.
- Woolf, B. P. (2002). Loyalty marketing: The second act. Teal Books.

Citación y derechos de autor

Este material ha sido creado por [Giovany Babativa-Márquez](#) y es de libre distribución bajo la licencia [Creative Commons Attribution-ShareAlike 4.0](#).

Si se copia parcial o totalmente, debe citar la fuente como:

Babativa-Márquez, J.G. *Diapositivas del curso de inteligencia de negocios*. URL: <https://jgbabativam.github.io/Muestreo-I/Semana1.html>. 2024.