

HD INSTITUTO HUMAN DELTA^{MR}

A blurred background image of a person's hands typing on a silver laptop keyboard. A credit card is resting on the laptop's trackpad area. To the left, a glass of water is partially visible. The overall color palette is dark and moody.

El desarrollo del espíritu a través del conocimiento al servicio de mis hermanos.

Temario



¿Qué es la analítica?



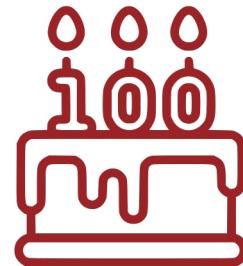
EDA



Probs & Stats
101

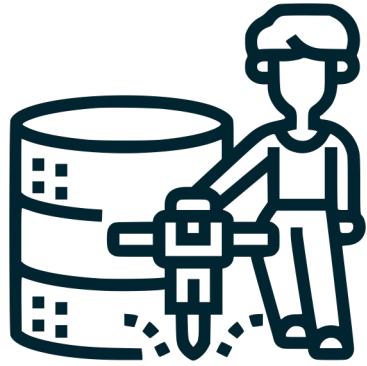


Storytelling



Caso Práctico

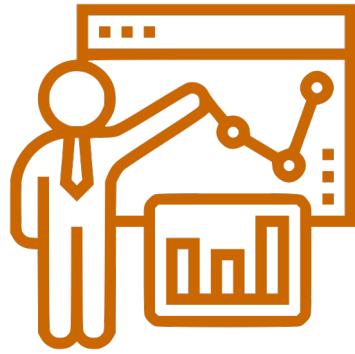
¿Qué es la analítica?



Ingeniero
de datos



Analista de
datos



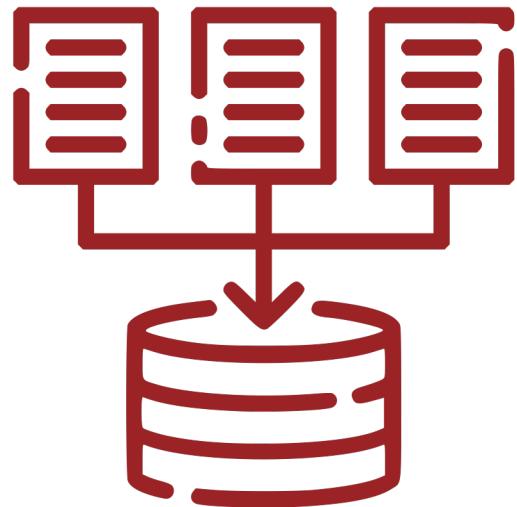
Especialista en
visualización de
datos



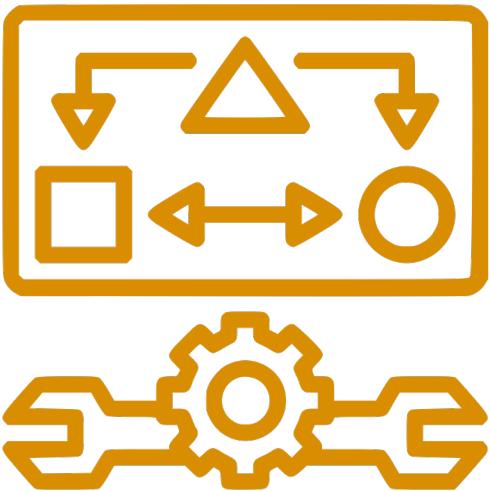
Business
Leader

¿Analítica vs. Estadística?

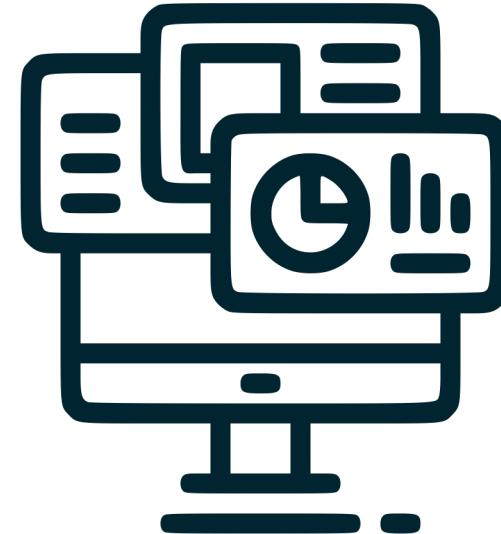
La estadística es una rama de las matemáticas que trabaja con:



Colección



Análisis e
interpretación



Presentación

¿Analítica vs. Estadística?

La estadística tiene aplicaciones en:



Búsqueda



Inventarios



Prediccion
del clima



Prediccion
del elección



Customer
intelligence



Análisis de
mercado



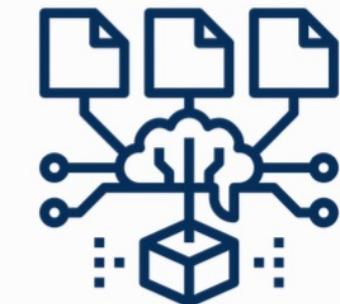
Planeación de
presupuesto



Exploración de
escenario



Manejo de
riesgo



Machine
Learning

Estadística

Descriptiva

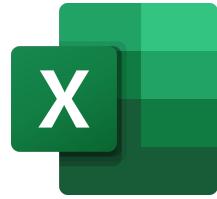
- Describe o resume datos
- Media / Varianza

Inferencial

- Infiere características de una población
- Probabilidad / medidas de confianza

Herramientas clave

Tradicionales



Transformación de datos



Análisis



Otros



¿Qué es EDA (Exploratory Data Analysis)?



- Muestreo
- Observaciones
- Gráficas
- Valores únicos
- Valores nulos
- Matrices
- ...

Variables

Variables

Clasificadas en dos grandes grupos

Categórico

También llamada variables cualitativa, se refiere a características que no pueden cuantificarse

Numérico

También llamada variables cuantitativas, se refiere a características cuantificables, cuyos valores son números

Nominal

Una variable nominal describe un nombre, etiqueta o categoría **sin un orden natural**

Ordinal

Una variable ordinal, es aquella con valores que están **definidos por una relación** de orden entre las diferentes categorías.

Continuo

Una variable es continua si puede **asumir un número infinito de valores reales** dado un intervalo

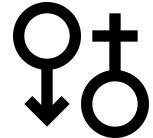
Discreto

Una variable discreta puede asumir **solo valores finitos** de valores reales dado un intervalo

Variables

Variables

Categórico



Numérico



Nominal

- Hombre / Mujer
- Green, Red, Blue

Ordinal

- Small / Medium / Large
- A, B, C (Calificaciones)

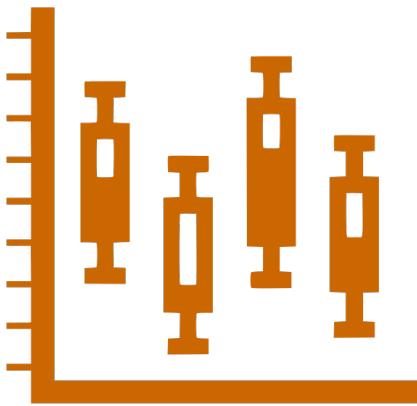
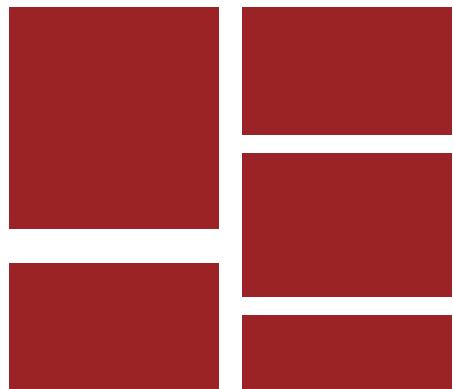
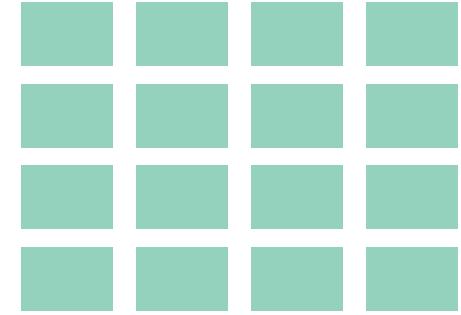
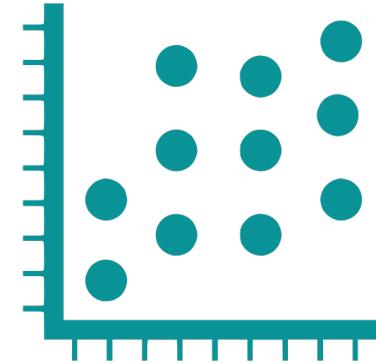
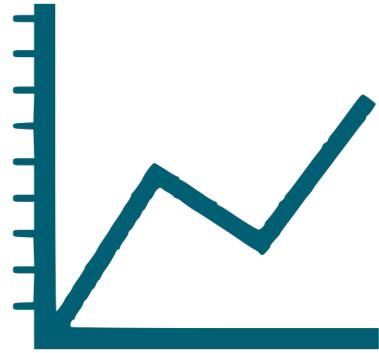
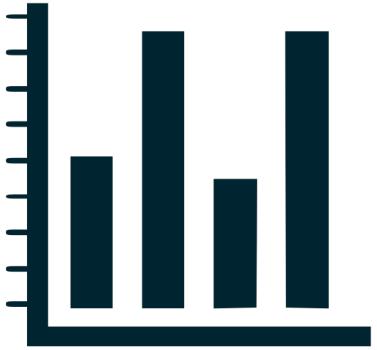
Continuo

- 1, 2, 3, 4, 5, 6, ...
- 7 estudiantes

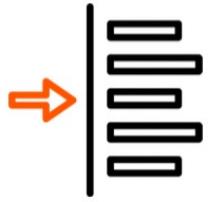
Discreto

- Edad
- Peso

Gráficas

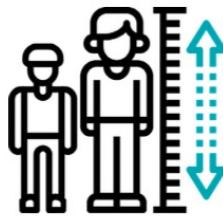


Conceptos estadísticos



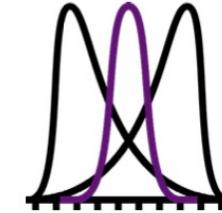
Medidas de tendencia central

- Media
- Mediana
- Moda



Medidas de dispersión

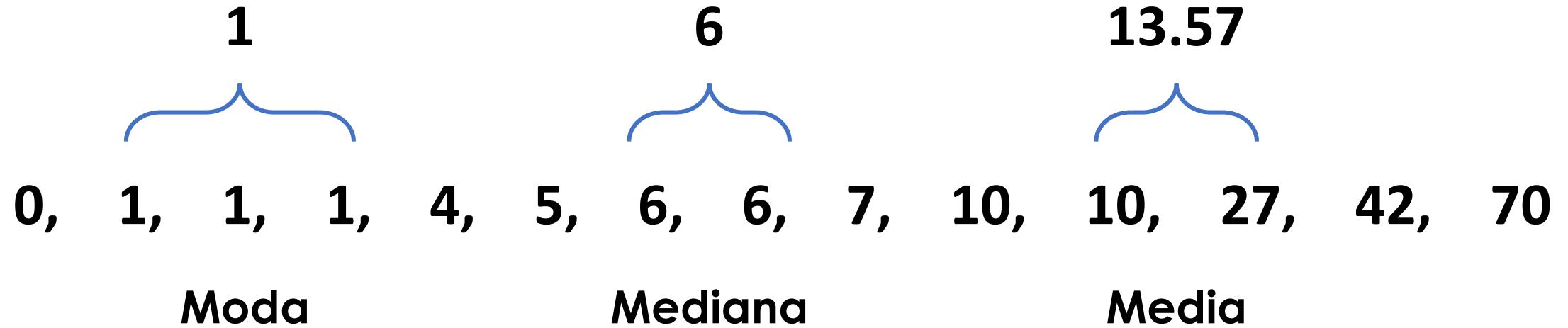
- Rango
- Varianza
- Desv. Estandar



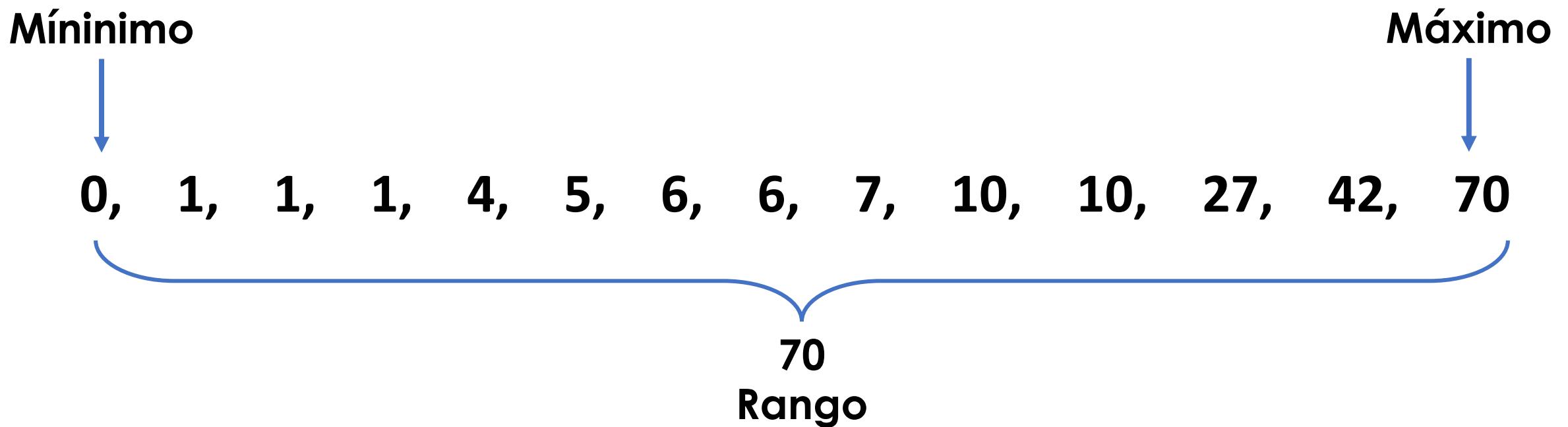
Forma de distribución

- Rango
- Varianza

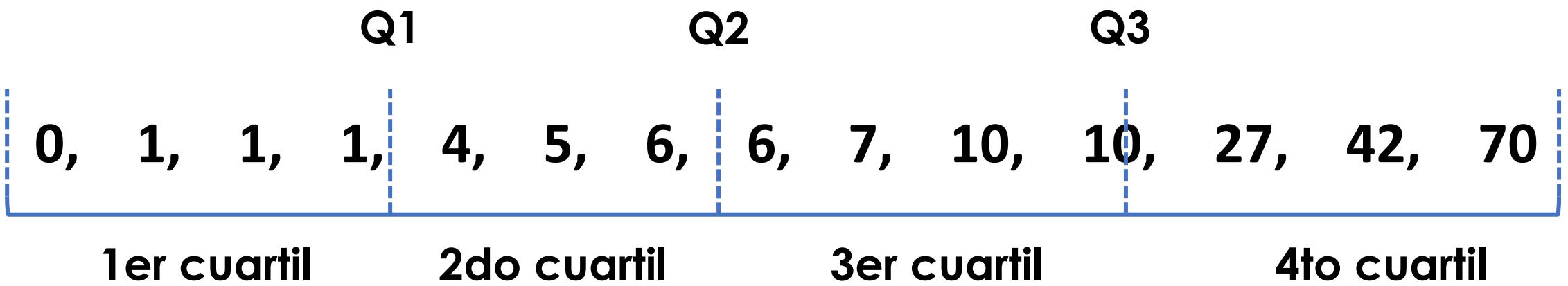
Medias de Tendencia Central



Mínimo, Máximo y Rango



Cuartil

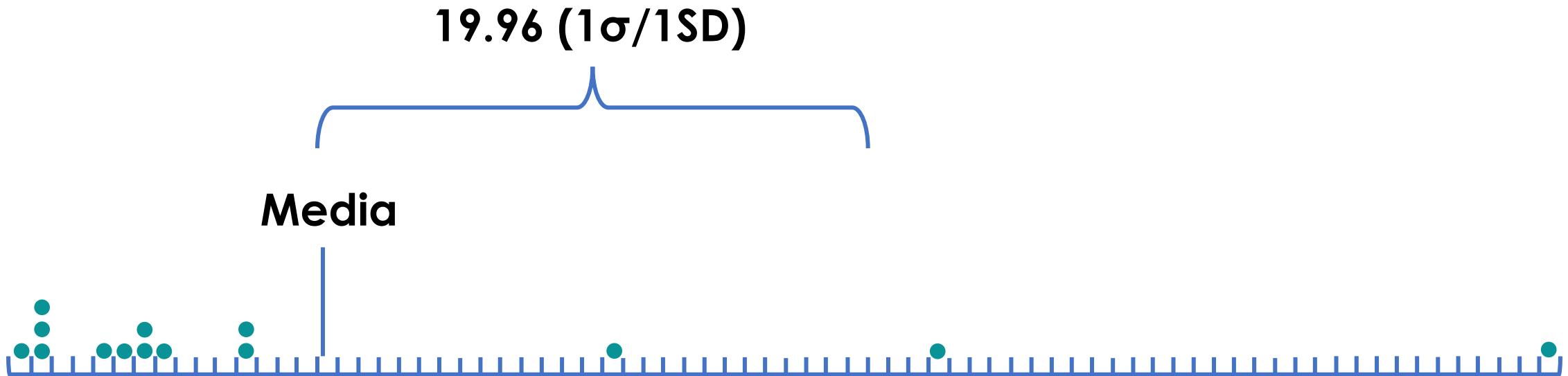


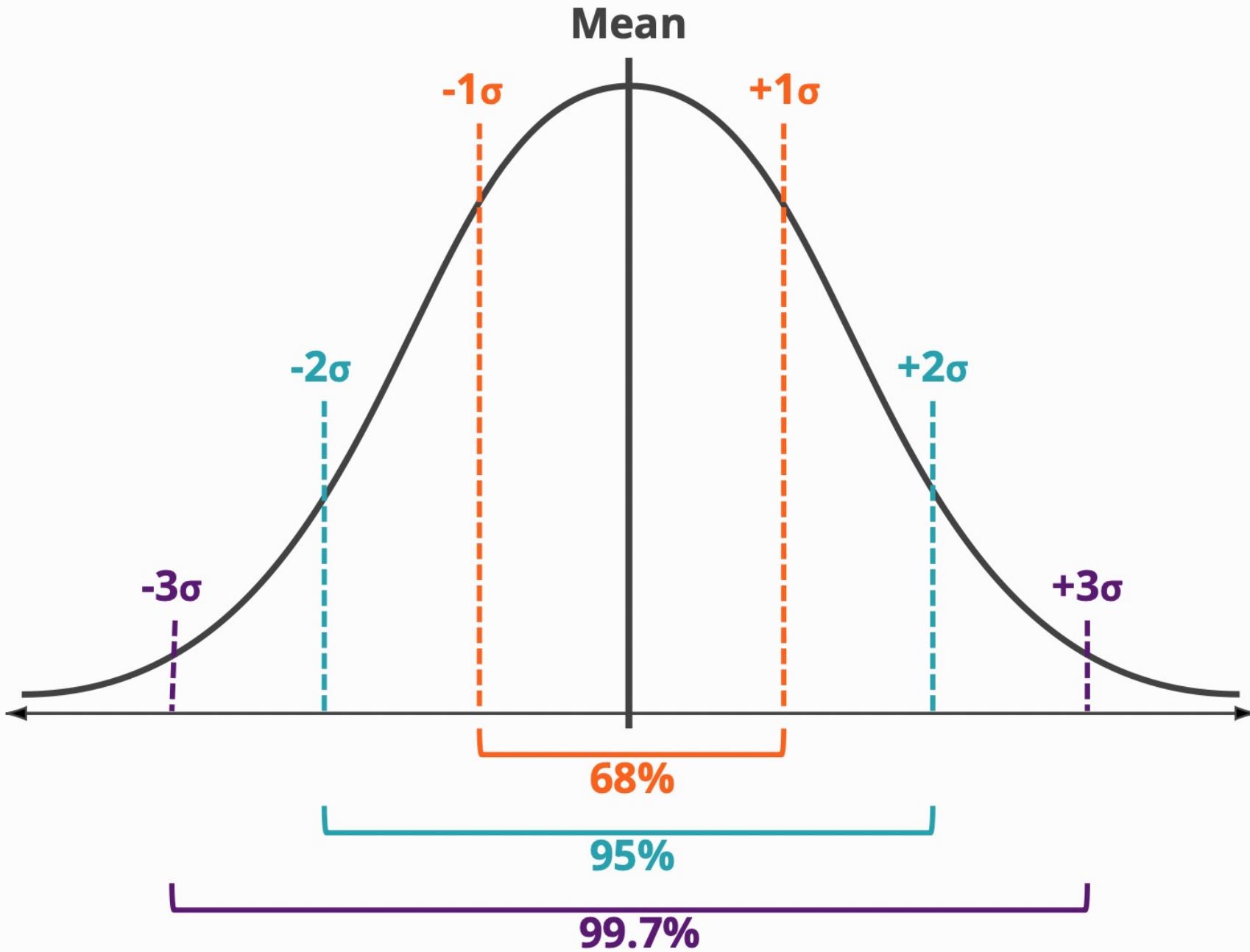
Varianza

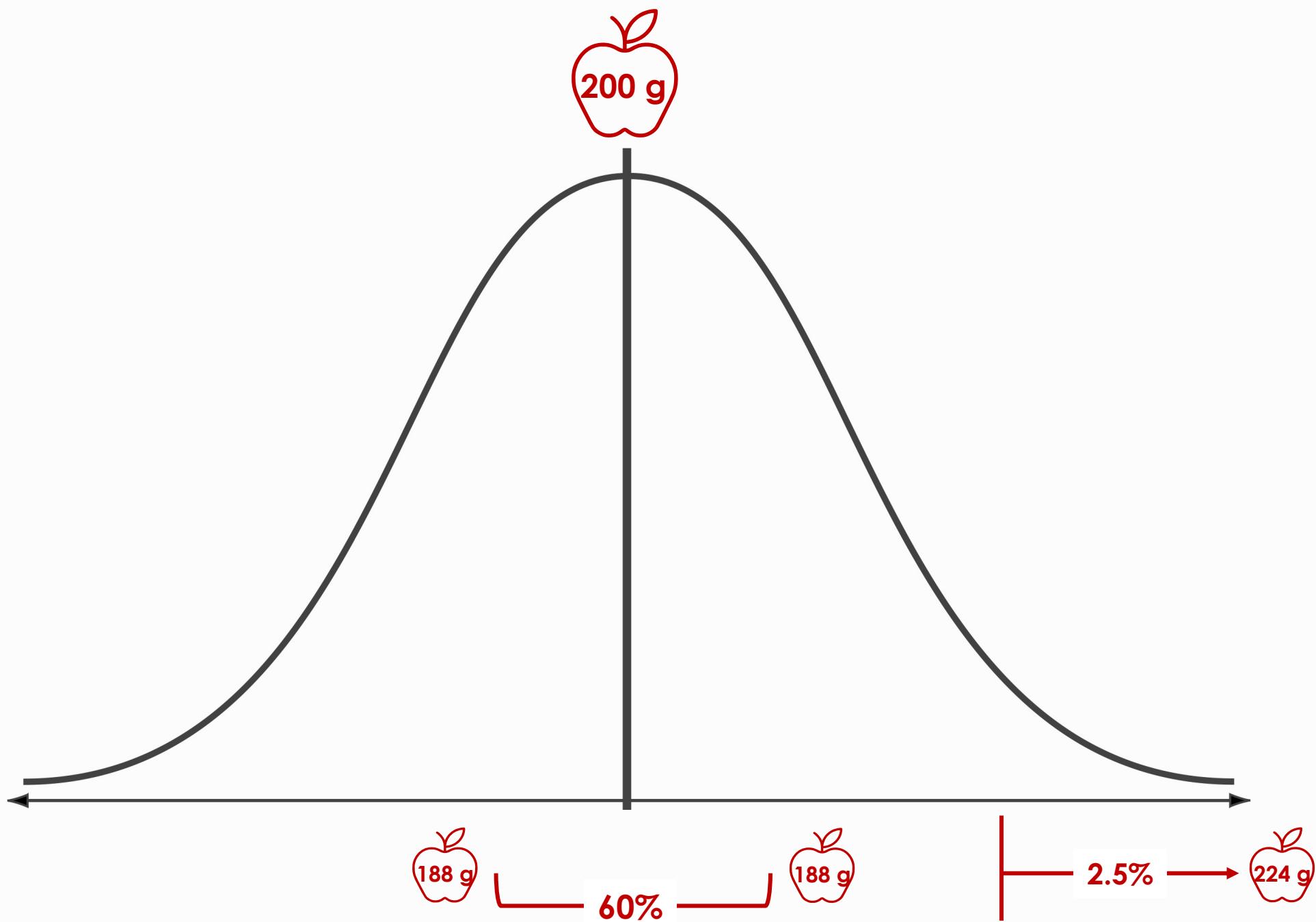
0, 1, 1, 1, 4, 5, 6, 6, 7, 10, 10, 27, 42, 70



Desviación Estándar







Probabilidad y Aleatoriedad



Muestreo



Metodologías para Data Analysis

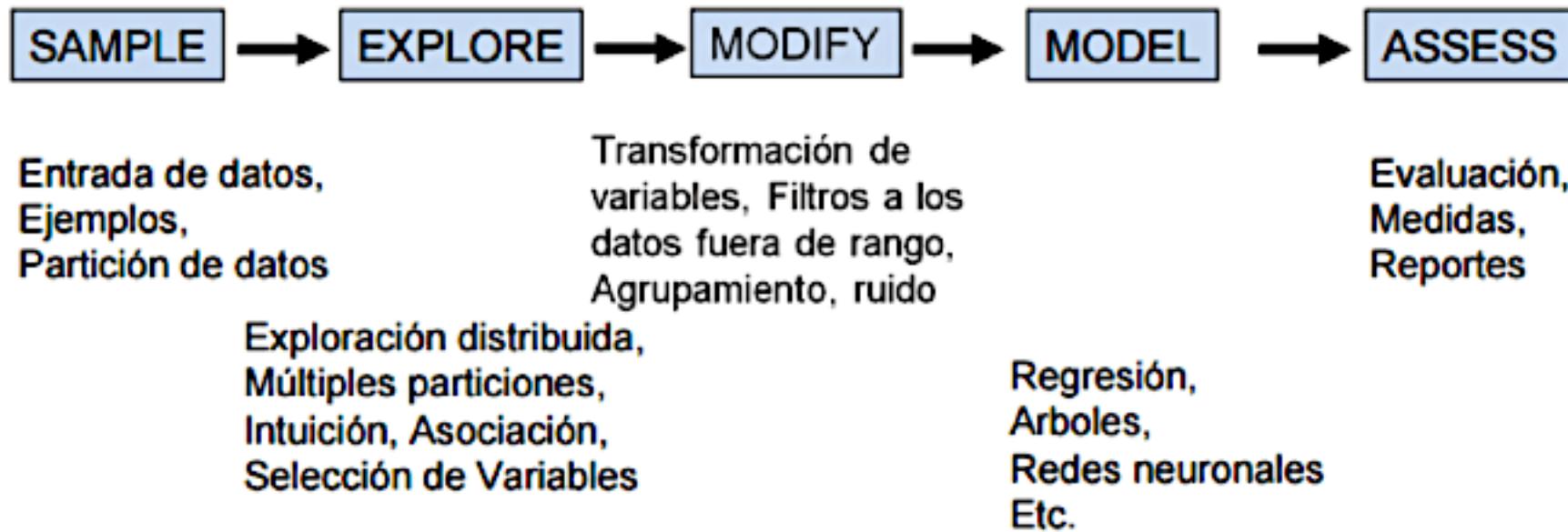


Metodologías SEMMA

- ❑ Fue propuesta por SAS Institute.
- ❑ Se define como el proceso de selección, exploración y modelado aplicado a cantidades significativas de datos.
- ❑ Permitan el descubrimiento de patrones como herramientas de apoyo para el negocio.
- ❑ Es un conjunto de herramientas funcionales enfocadas hacia los aspectos de desarrollo propio de un modelo de minería/analítico.

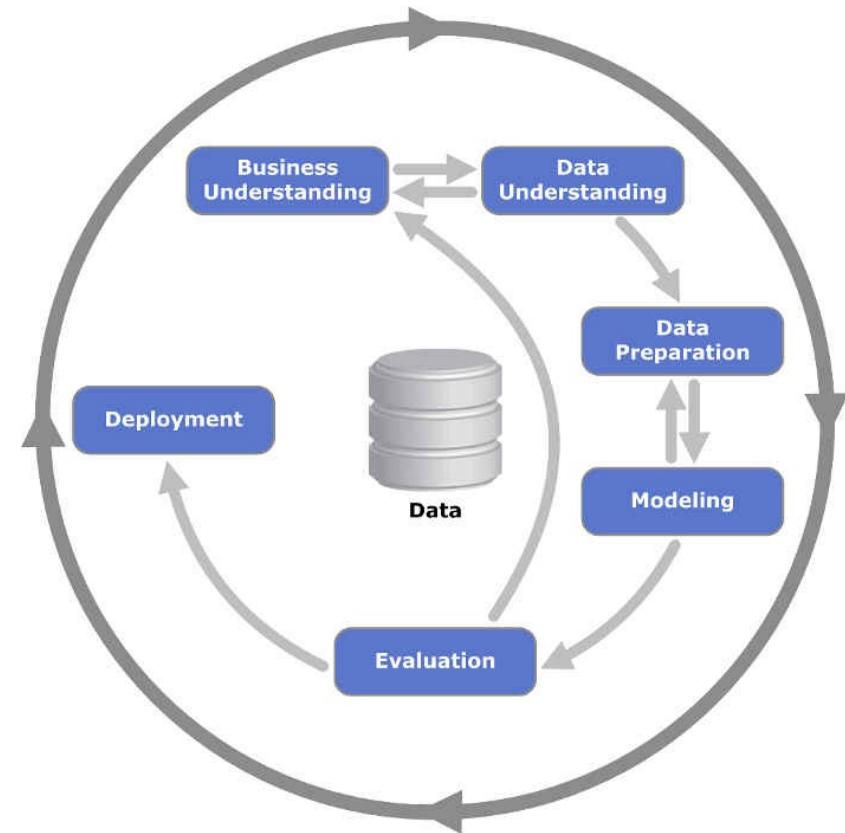


Metodologías SEMMA



Metodologías CRISP-DM

- ❑ CRISP-DM es el acrónimo de **Cross-Industry Standard Process for Data Mining**.
- ❑ Fue diseñada en 1990 por un consorcio de empresas europeas (NCR (Dinamarca), AG (Alemania), SPSS (Inglaterra) y OHRA (Holanda)).
- ❑ Se compone de **seis fases** que permiten el desarrollo óptimo del proyecto.



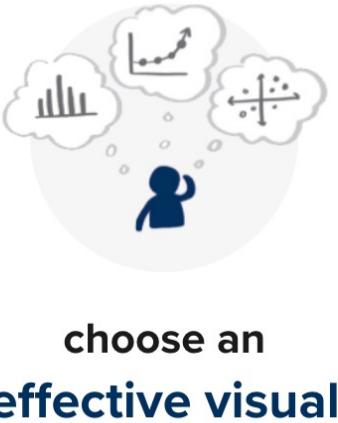
Metodologías CRISP-DM

SEMMA	CRISP-DM
Ya se tiene idea de la problemática.	No está clara la problemática.
Se cuenta con un ABT.	Se crea el ABT.
Se busca solo la Modelación.	Es iterativo en la modelación.
Se evalúa si se resuelve la problemática.	Se evalúa si se resuelve la problemática. (iterativo)

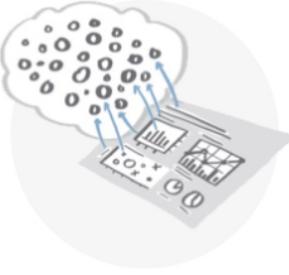
Storytelling



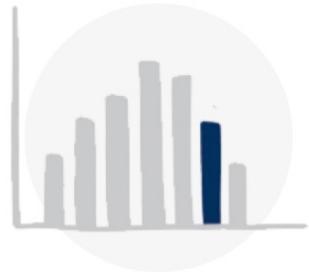
understand the
context



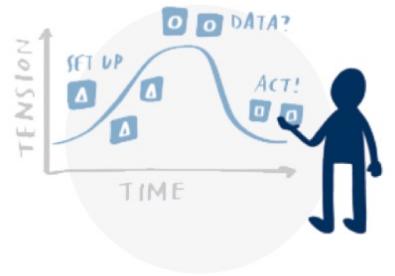
**choose an
effective visual**



**eliminate
clutter**



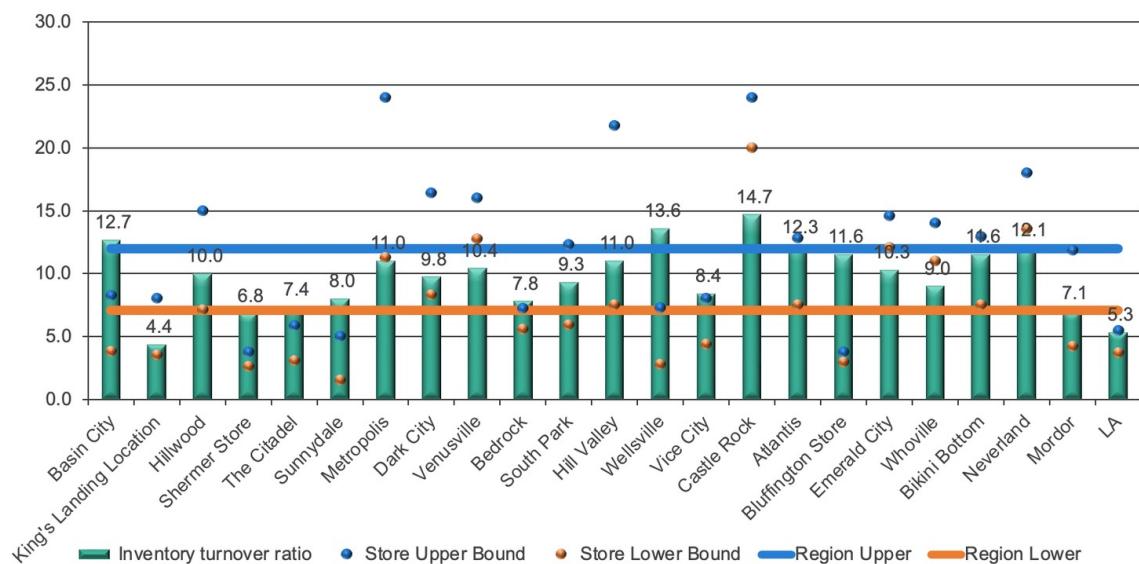
**focus
attention**



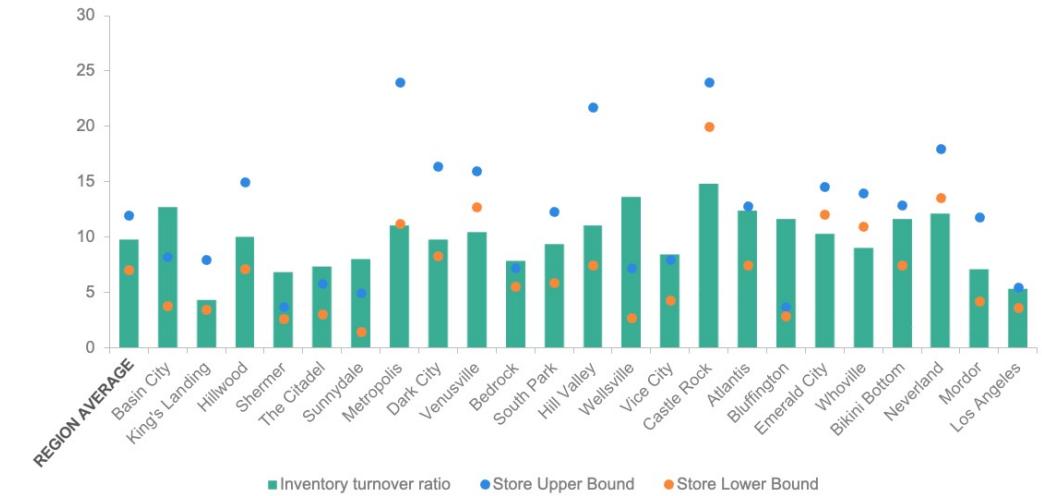
**tell a
story**

Storytelling

Inventory Turnover Ratio



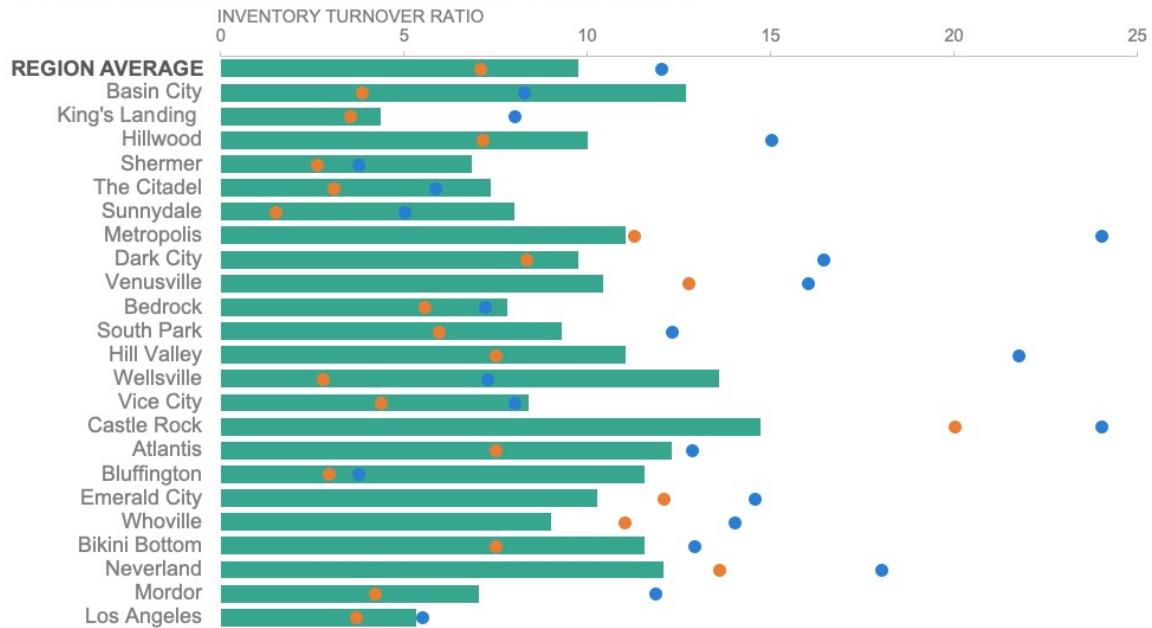
Inventory Turnover Ratio



Storytelling

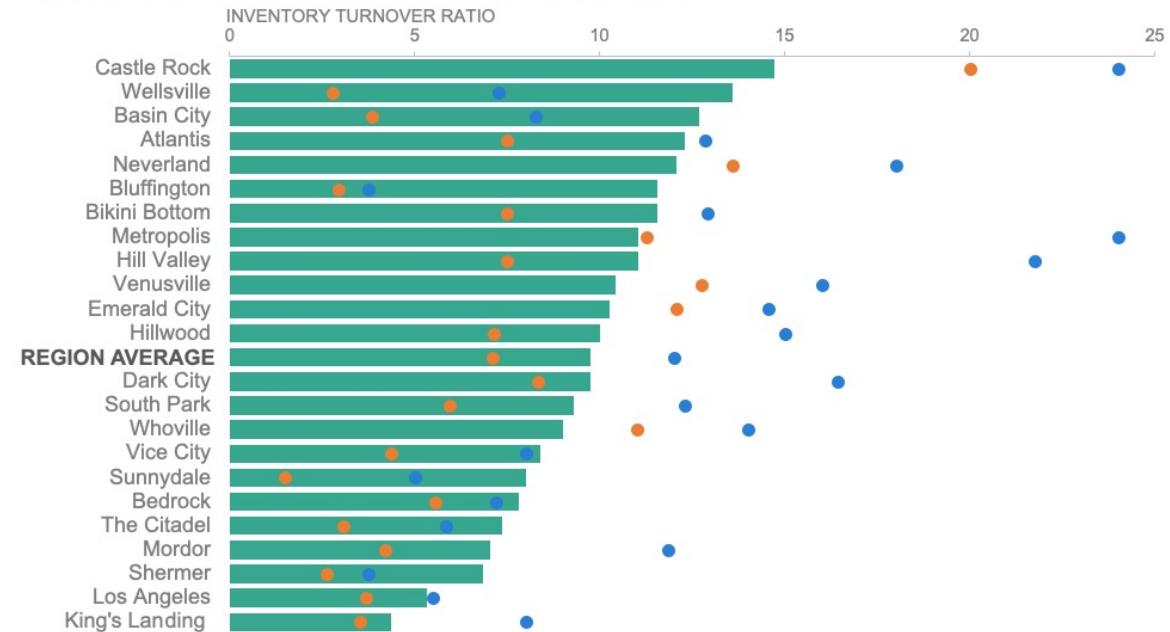
Inventory turnover ratio by store

INVENTORY TURNOVER RATIO | LOWER BOUND | UPPER BOUND



Inventory turnover ratio by store

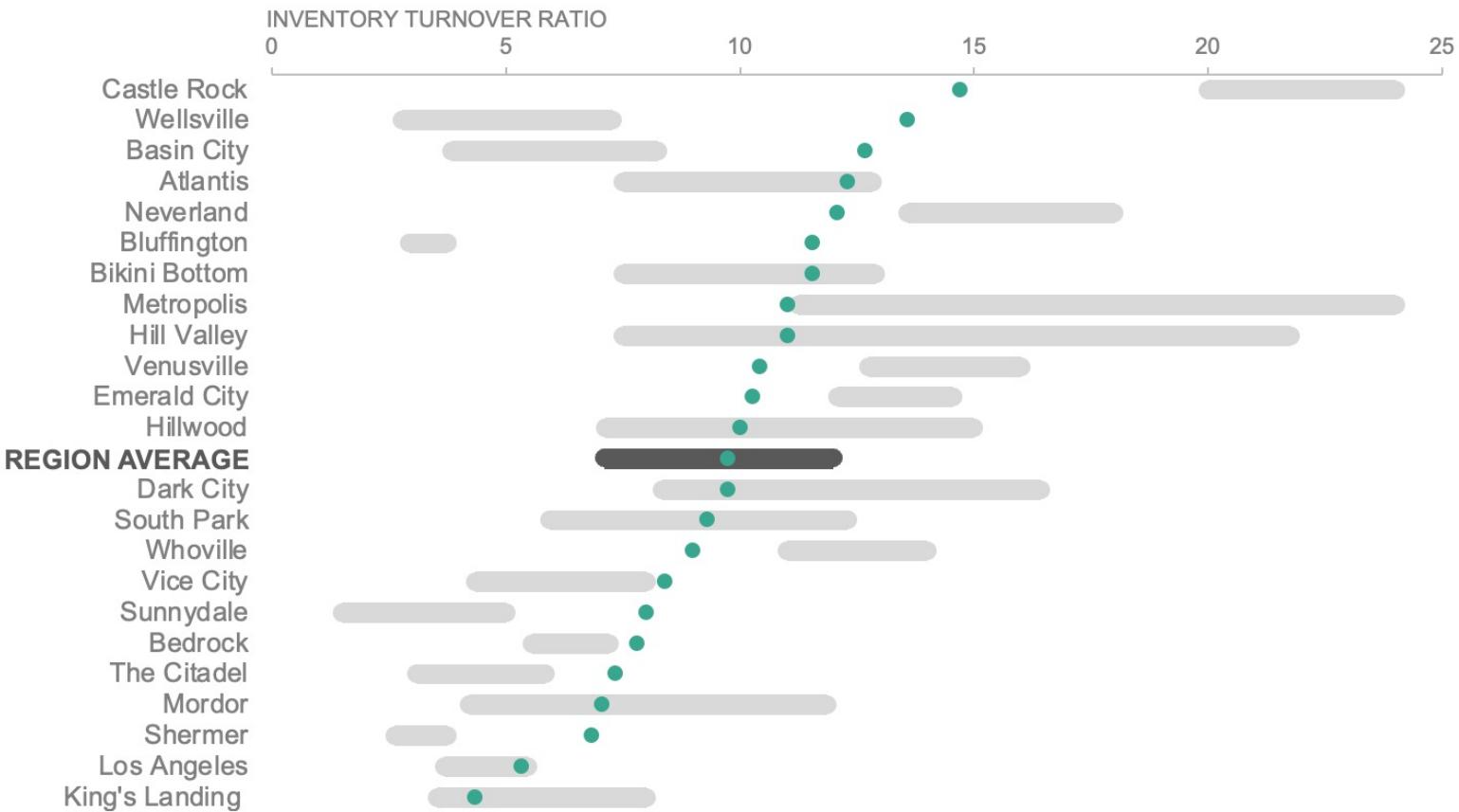
INVENTORY TURNOVER RATIO | LOWER BOUND | UPPER BOUND



Storytelling

Inventory turnover ratio by store

INVENTORY TURNOVER RATIO | FORECASTED RANGE



The Matthew Effect