MBA USP ESALQ

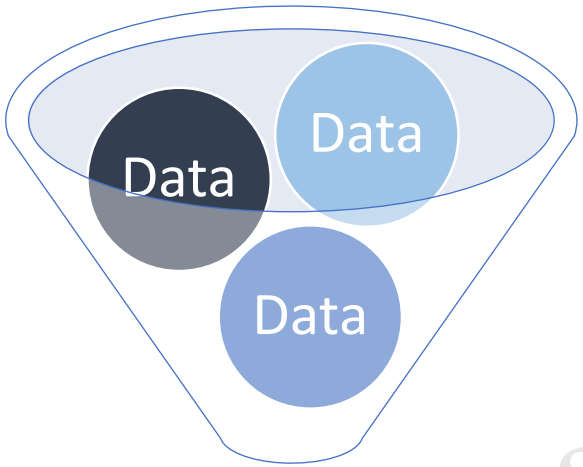
BUSINESS INTELLIGENCE AND DATA VISUALIZATION

Prof. Helder Prado Santos

*The responsibility for trustworthiness, originality, and legality of the didactic content presented is responsibility of the professor.

What is Business Intelligence?

Business



Information

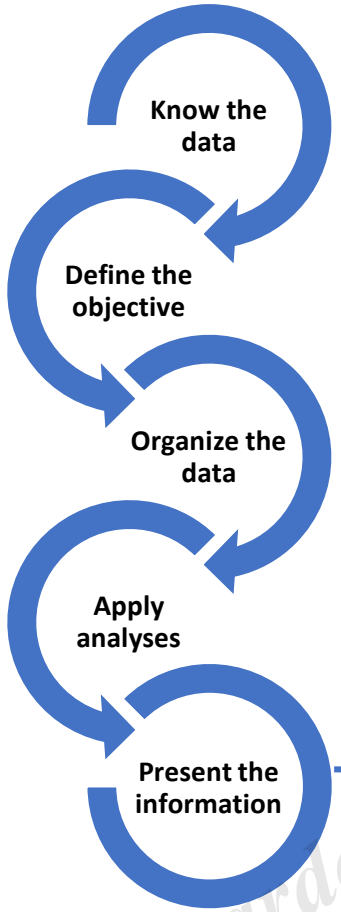
Make intelligent decisions

Business Intelligence (BI) is **not** a software or a tool!!

Has a directly relationship with the business needs

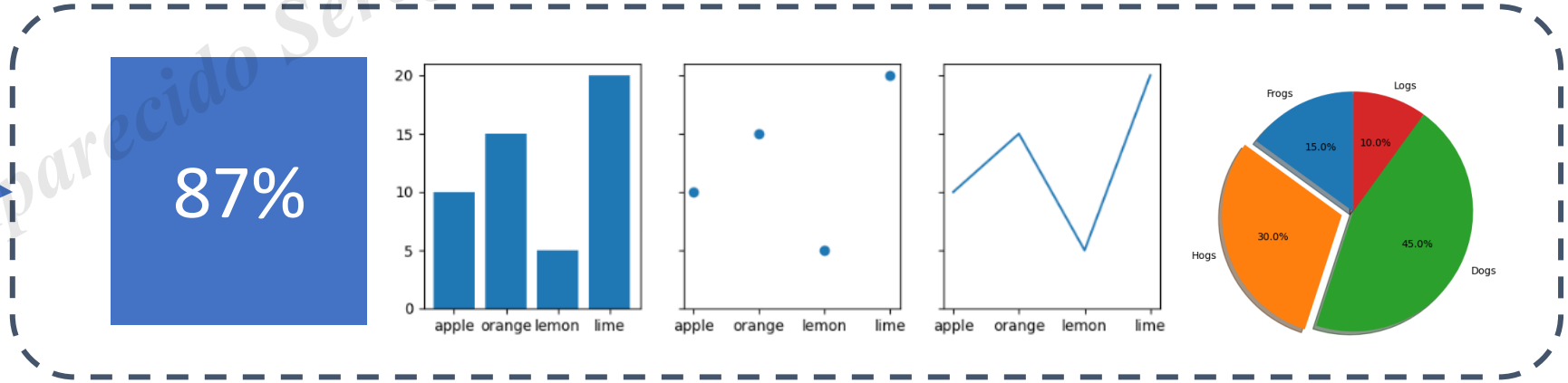
Business objective

How to present information?



CONCEPT:

It refers to the process of collection, organization, analysis, sharing, and information monitoring that offer support to business management.



Information vs storytelling

Tickets in December:

- 177 tickets received
- 140 tickets processed

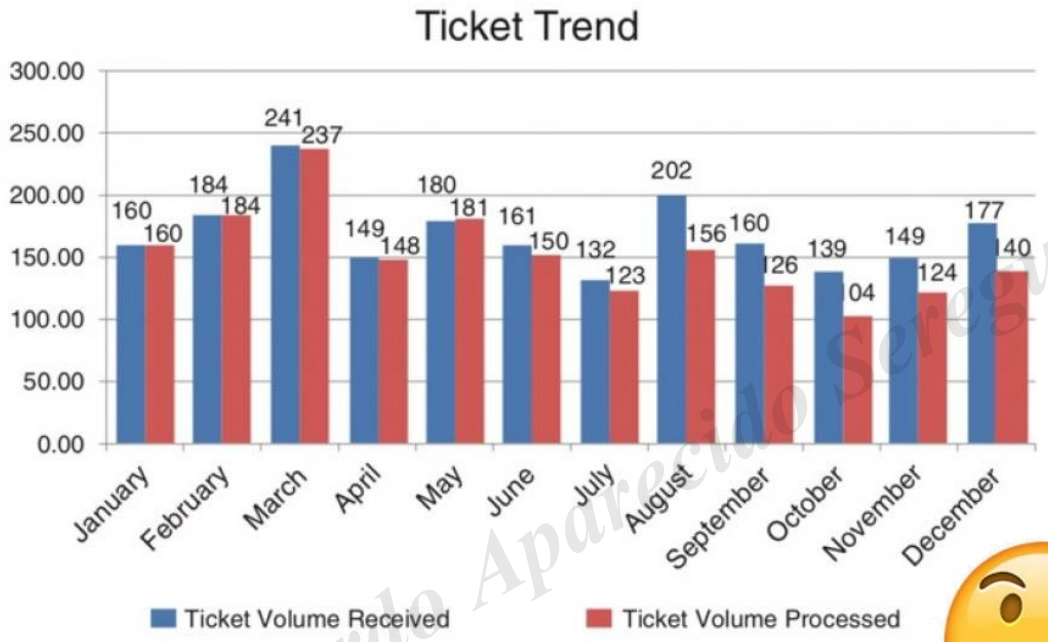


FIGURE 0.2 Example 1 (before): showing data

vs

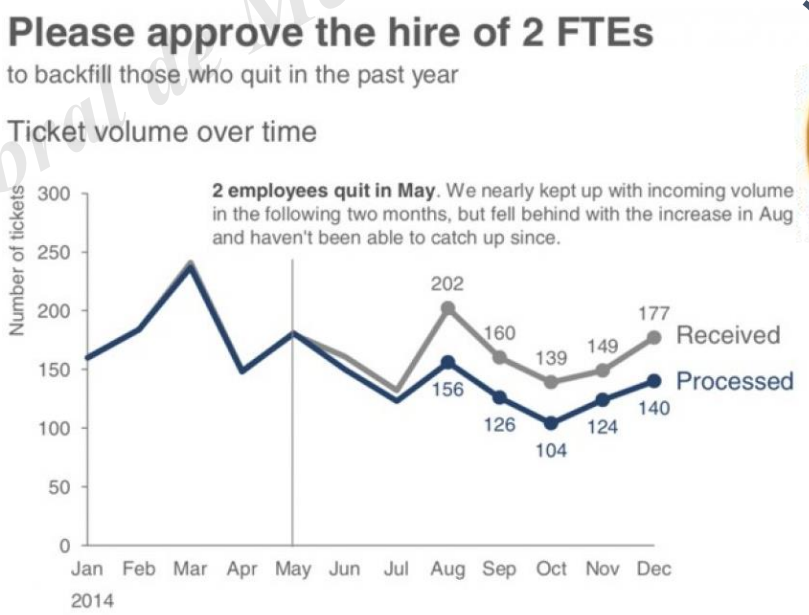
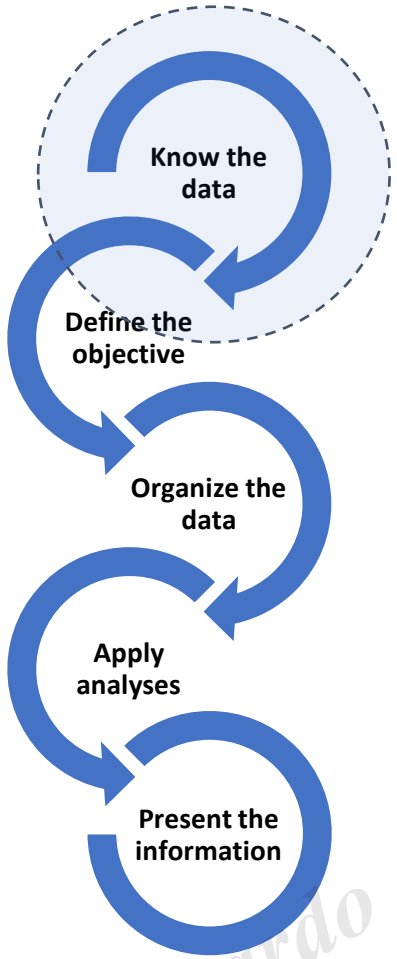


FIGURE 0.3 Example 1 (after): storytelling with data

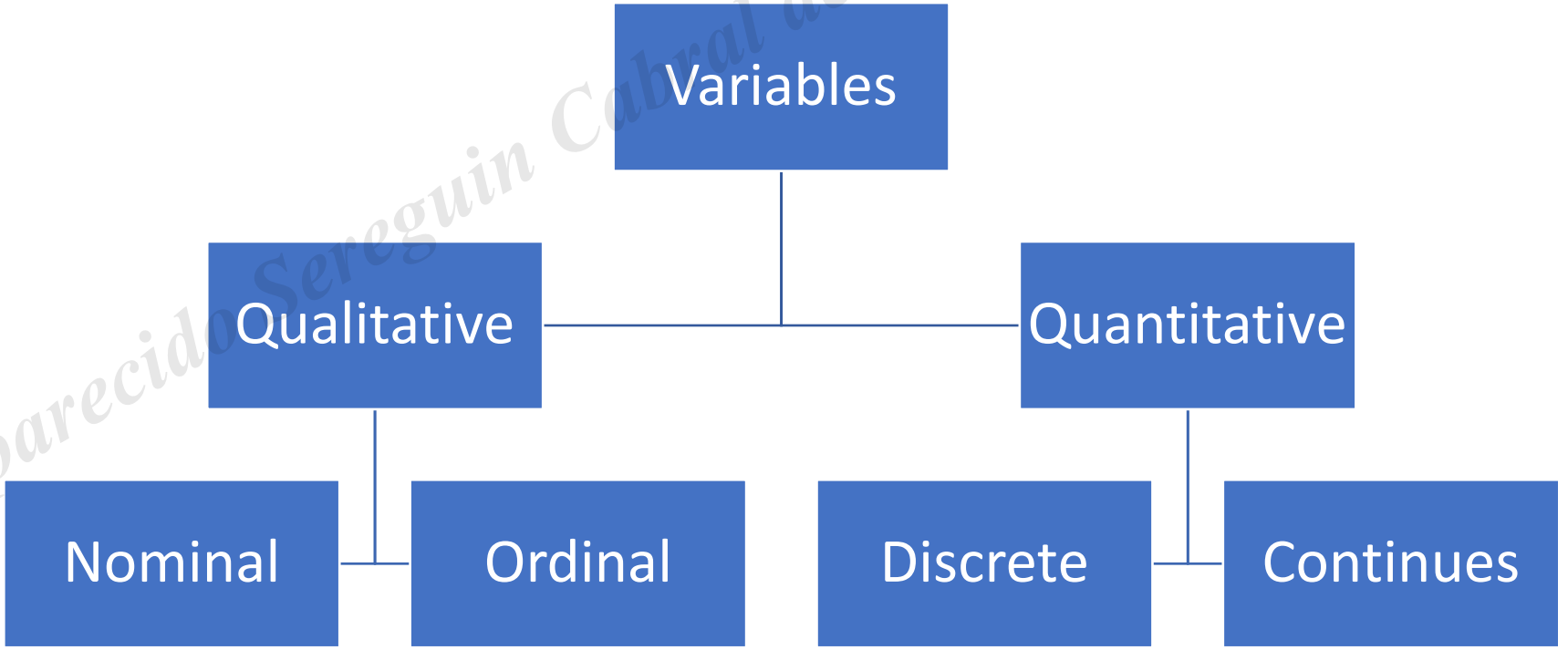
Does it convince?

*Source: Cole Nussbaumer Knaflic, **Storytelling with data**, 2016.*

How to choose the presentation?



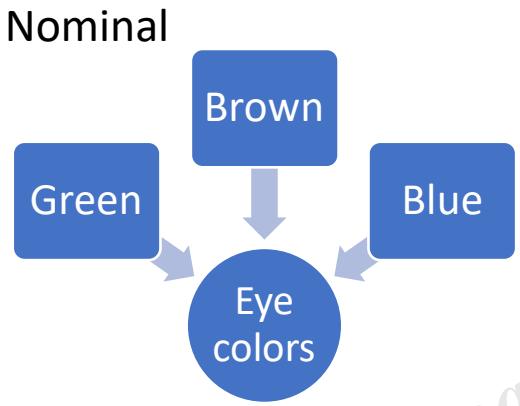
First: Know your variable (data)

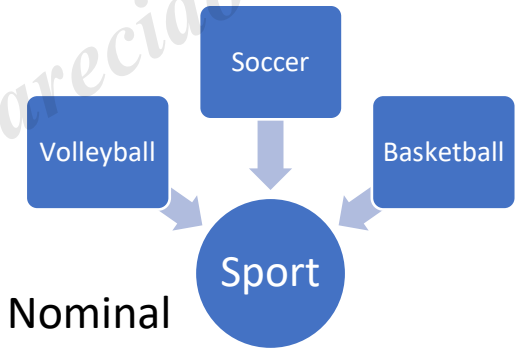


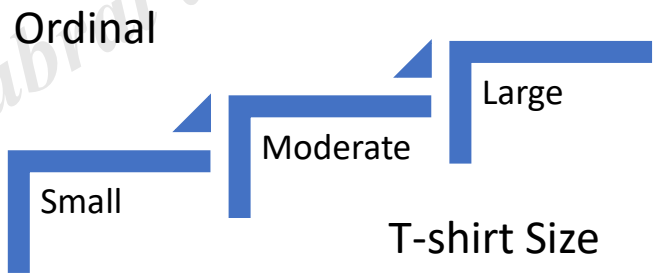
Variables and their types

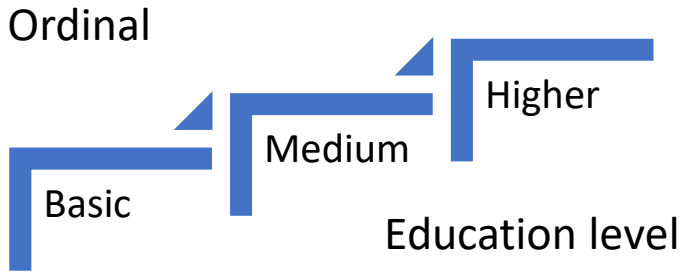
Qualitative Variables: Represent attributes that can be divided into two groups:

- **Ordinal**: when there is an implied order
- **Nominal**: when **there** is not an implied order









Variables and their types

Quantitative Variables: Represent attributes that can be divided into two groups:

- **Discrete**: belong to the interval of finite values and enumerable, and it can be counted. (Counting)

- **Continuous**: belong to the interval of real numbers.

Discrete:

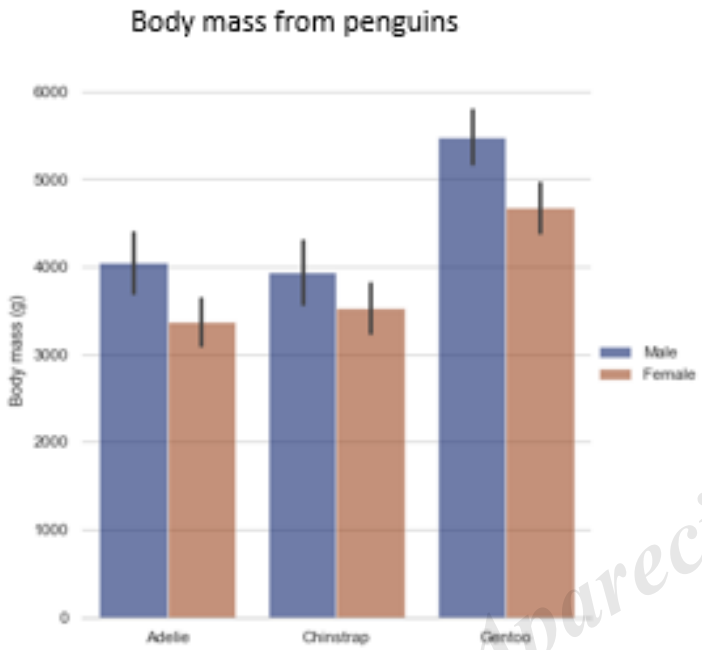
- Number of children (1 child)
- Number of accidents (2 accidents)
- Number of open tickets (20 open tickets)
- Number of parts sold (14 parts sold)

Continuous:

- Weight (88.6 kg)
- Invoicing (R$ 12,599.90)
- Expense (R$ 304.40)
- Height (1.82 cm)

Charts suitable for each type of variable

Grouped column chart



Line chart

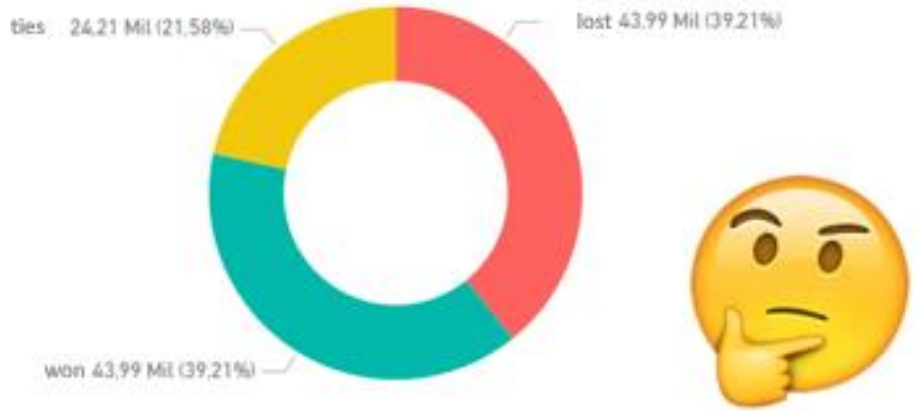
Source: https://seaborn.pydata.org/examples/grouped_barplot.html

Charts suitable for each type of variable

Column and line chart with secondary axis.

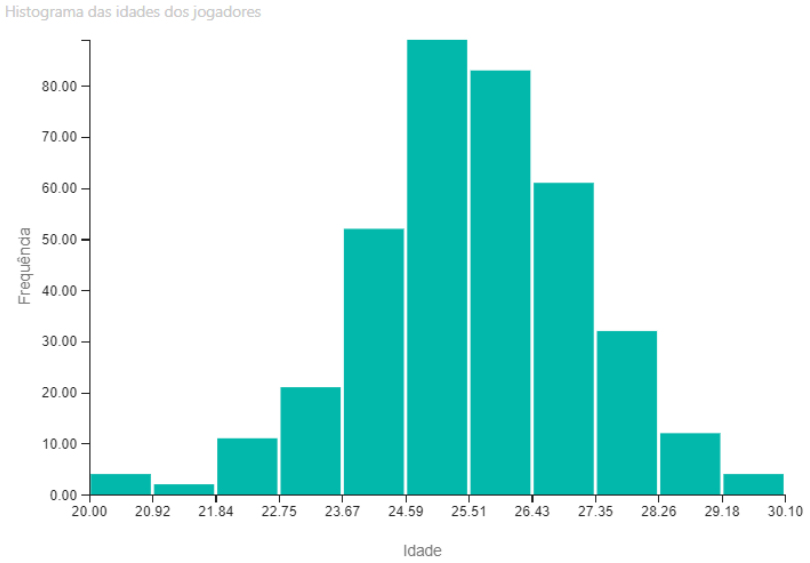
Donut chart (derivation from pie chart)

Attention point!



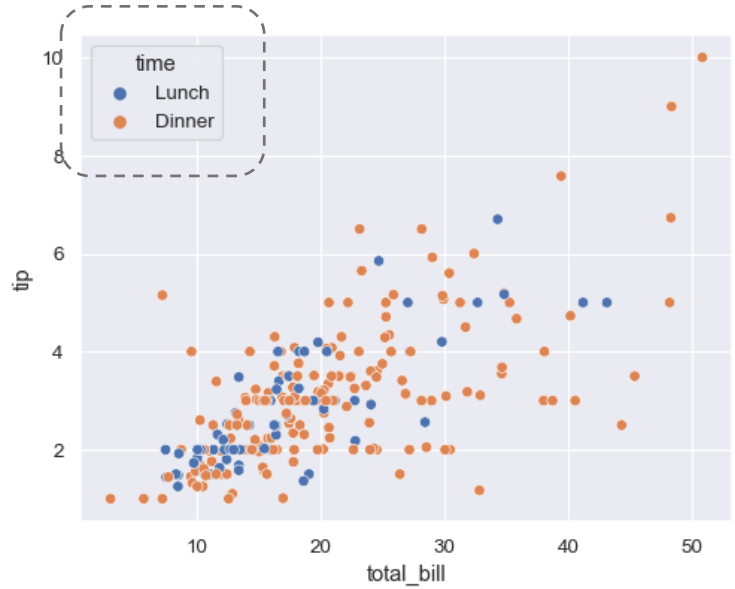
Charts suitable for each type of variable

Histogram



Dispersion chart

Category visualization

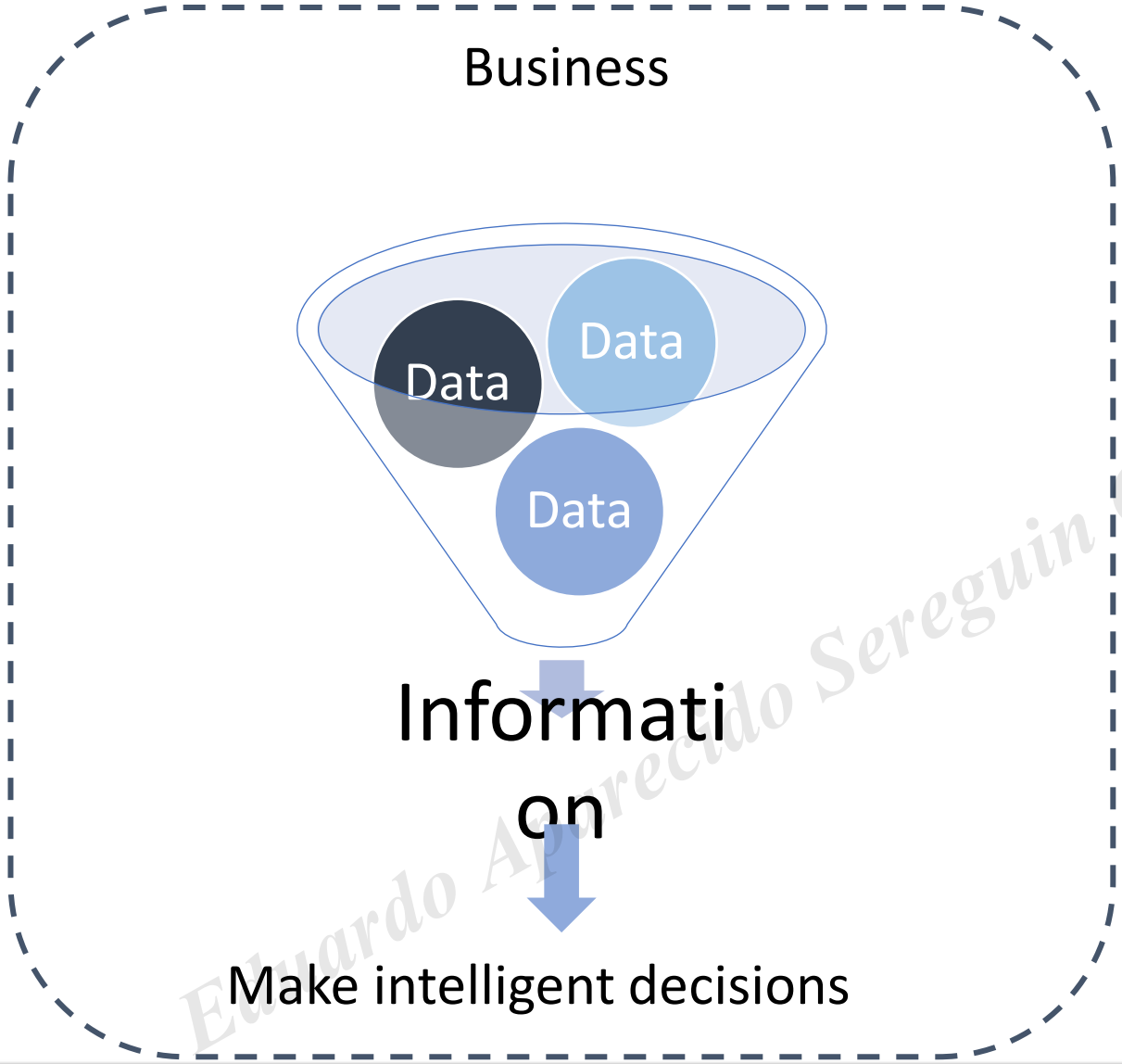


Source: https://seaborn.pydata.org/generated/seaborn.scatterplot.html?highlight=scatter#seaborn.scatterplot

Business intelligence softwares in the market

Many others..

Business intelligence project (BI)



Business: football analytics company

Audience: Managers, technicians, and audience in general

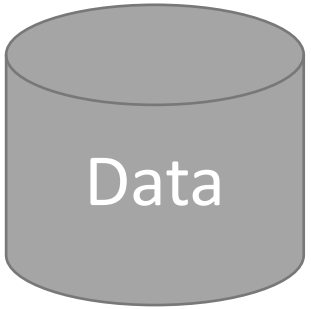
Objectives:

Develop dashboards in which the user can:

- Verify the statistics of a player by:
  - Season
  - Competition
  - If they play in their home ground or not
- Verify the statistics of a club by:
  - Season
  - Competition
  - If they play in their home ground or not

Business intelligence project (BI)

Brand of the fictional business



- 42 championships from Europe
- 55,000+ games of several seasons
- 400+ clubs of these competitions
- 20,000+ players of these clubs
- 300,000+ market values of players over time
- 1,000,000+ apparitions of these players in the championships games.

Source: https://data.world/dcereijo/player-scores (Treatment) https://www.transfermarkt.co.uk/ (Origin)

Business intelligence ´project (BI)

Color palette of the company:

#f1c40f #666666 #ecf0f1 #f00000

It improves the branding in several aspects, becoming a strategy linked to design and marketing.

Help the customer memorizes your brand easily.

It facilitates the creation of the report by limiting options

Which decisions can I make with this information?

Analysis: I am the technician of a team, and I see that the goal average of my opponent is **2.8**.

Decision-making:
Change the position of players to have a more defensive team.

Analysis: The number of yellow cards of a striker of my team increased from one season to the other, it was from **12** to **21**.

Decision-making:
Send the striker to the club psychologist.

Analysis: My team has a goal average very low **(1.4 goals**) at the end of the season.

Decision-making:
Dispense the players that score few goals and give little assistance, and hire players who have a greater average of goals and assistance.

Real example:

Problem:
Liverpool, from England, over many years spent money with strikers and lost money because the team was never the champion.

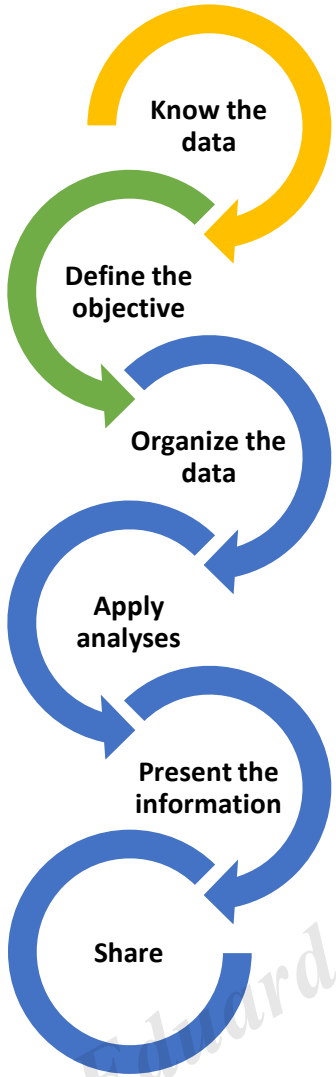
Analysis:
The football analytics department of the club suggested to hire a fullback called Virgin Van Dijk, from a small team, Southampton. This suggestion was based on the historic of conceded goals from Liverpool in the last years, associated to the historic of few conceded goals from the fullback in the small team he played.

Decision-making:
It hired the player for 75 million pounds. (the most expensive fullback in the history at that time)

Result:
The team did not lose in any game with him playing in the season, and they won the championship after more than 20 years.

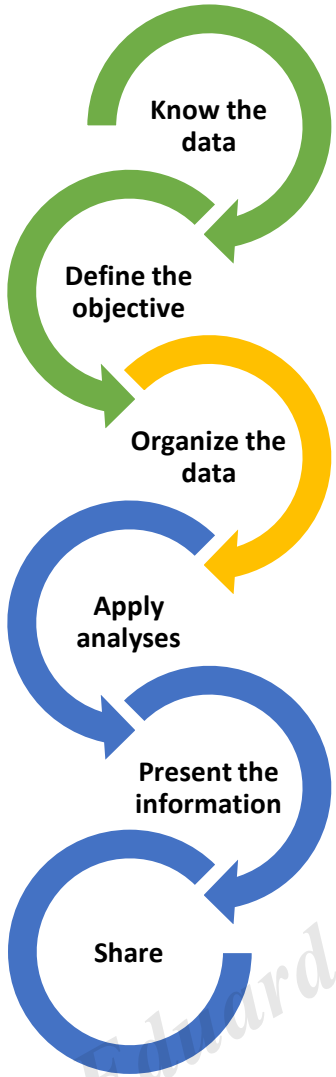
Source: https://oglobo.globo.com/esportes/torcida-do-liverpool-deve-esquecer-preco-de-van-dijk-afirma-klopp-22235031

Know the data



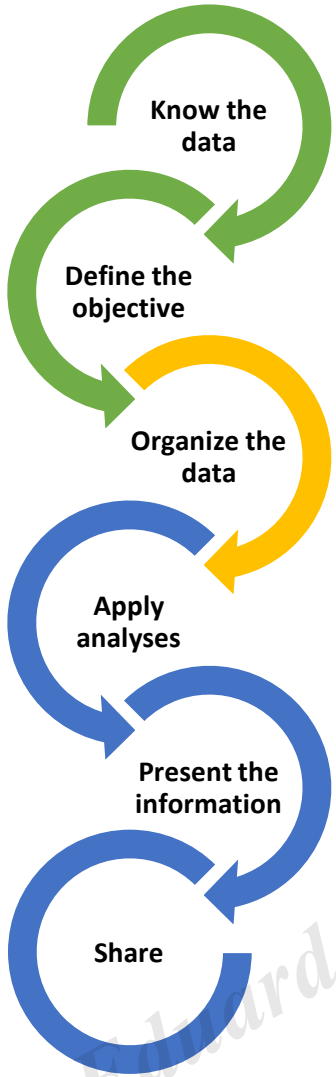
Championships	Clubs
Games	Players
Players apparitions	Player values

Organize the data



Connect data sources → Adjust and standardize data types → Establish relationships between tables

Organize the data



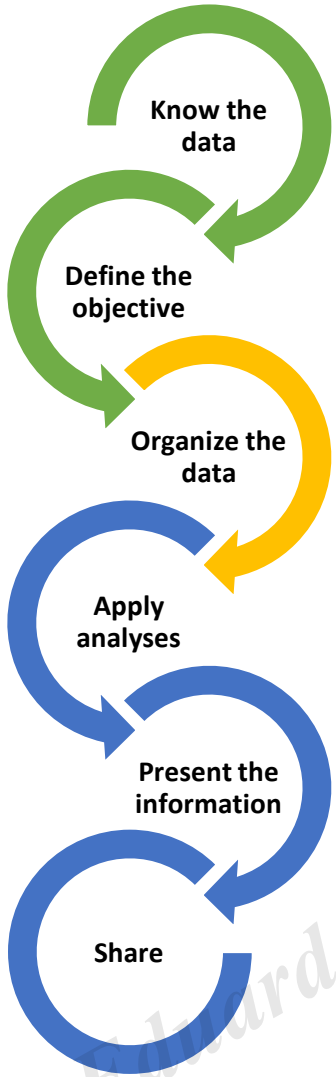
Power BI

Connect data sources → Adjust and standardize data types → Establish relationships between tables

Data sources:
- .csv
- .txt
- .json
- .xlsx
- SQL database
- Among many others

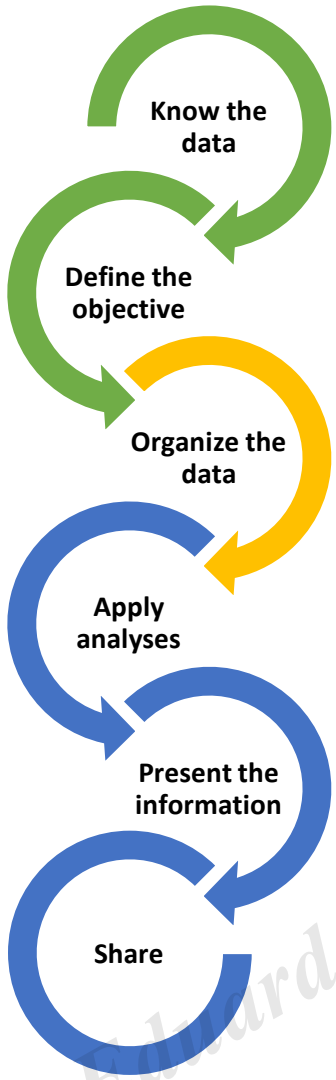
Connection type:
- Import
- DirectQuery

Organize the data



Connect data sources → Adjust and standardize data types → Establish relationships between tables

Organize the data



Power BI

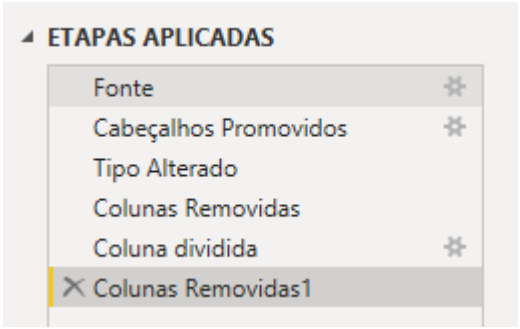
Connect data sources → Adjust and standardize data types → Establish relationships between tables

Power Query Editor within Power BI

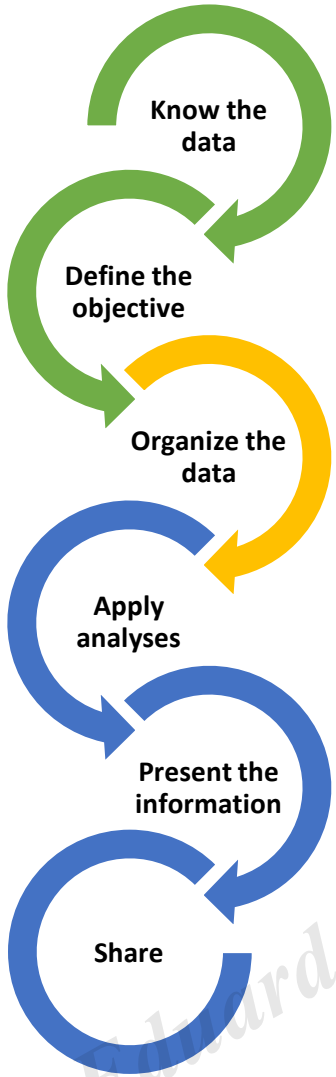
Power BI **ETL** Interface:

- **E**xtract the data
- **T**ransform the data
- **Lo**ad the data

Series of stages



Organize the data



Power BI

Connect data sources → Adjust and standardize data types → Establish relationships between tables

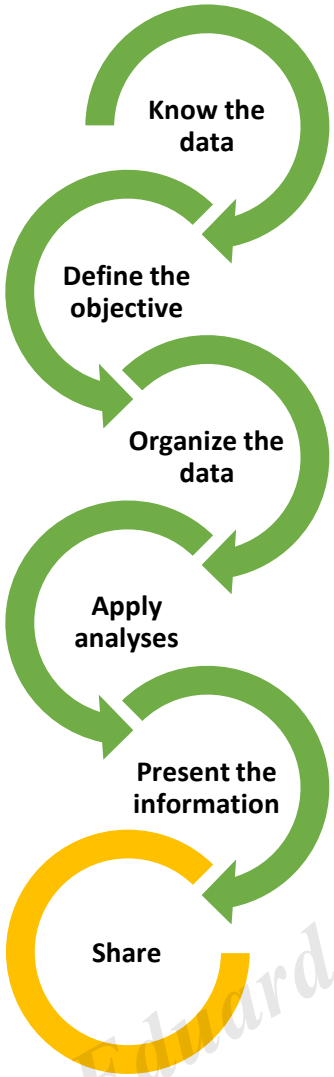
Type of relationship between tables: (Cardinality)

- One to many (1,*)
- Many to one (*, 1)
- One to one (1, 1)
- Many to many (*, *)

→ Both

→ Unique

Organize the data



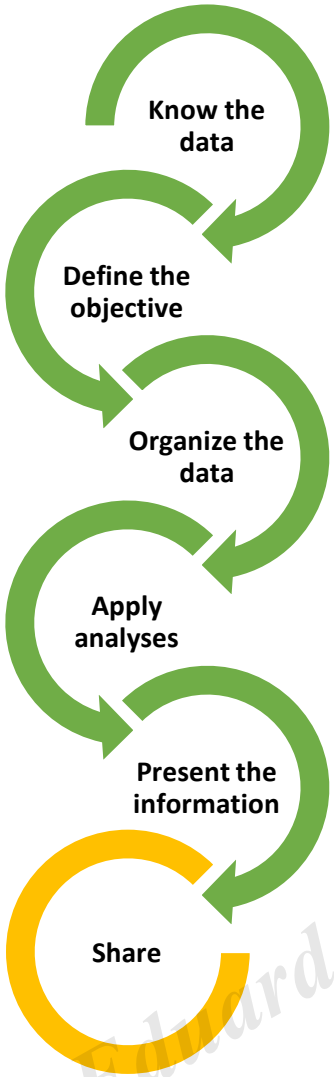
Power BI

Publish → Share

Type of publication:

- Website or web portal
- Web (public)
- Sharepoint Online
- Other options for the developer

Organize the data



Publish → Share

Type of sharing:

- Audience
- Constrained

MBA
USP
ESALQ

Prof. Helder Prado Santos

https://www.linkedin.com/in/helderprado/