BASIC CONCEPTS

Summary

- Why learn statistics?.
- Exploratory Data Analysis.
- About data concept.
- About variable concept.
- Work with some data types.

Why to Learn Statistics?

- Because data are everywhere.
- Help us make better decisions in our life.
- Decision making process needs statistical data analysis.

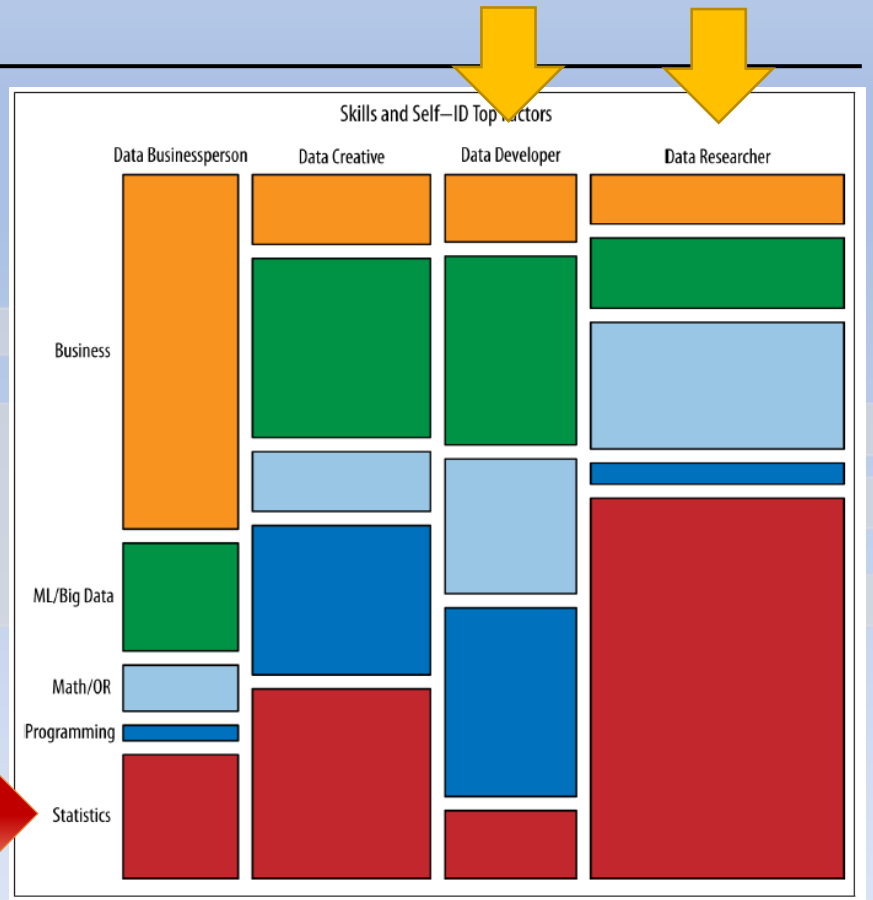


Figure 1-4. Harlan Harris's clustering and visualization of subfields of data science from Analyzing the Analyzers (O'Reilly) by Harlan Harris, Sean Murphy, and Marck Vaisman based on a survey of several hundred data science practitioners in mid-2012

Source: (Book) Doing Data Science

Exploratory Data Analysis

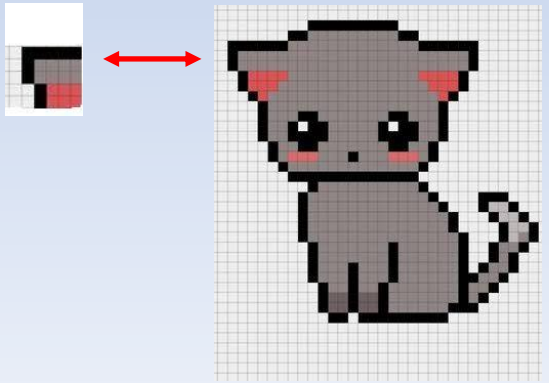
- Digging into the data collected.
- Know its basic properties.
- Make less effort in future analyzes.
  - plots, graphics, statistical summary

What is data?

- Data are characteristics or information usually numerical, that are collected through observation. Data are a set of qualitative or quantitative variables.

- Certain data, viewed in an unrelated way, may not contain interesting information, unless they are analyzed as a whole, under a certain approach or hypothesis.

Source: https://en.wikipedia.org/wiki/Data



What is a variable?

- A variable is a data set feature.

This is a variable

This is a sample

Employees List

No	Employee ID	First Name	Last Name	Age	Worked years	Salary	Status	Grade
1	1000001	John	Denver	23	1	$500	Single	Elementary
2	1000002	Peter	Hank	30	3	$900	Married	High School
3	1000003	Jack	Sullivan	27	2	$900	Married	High School
4	1000004	Marco	Aurelio	40	8	$1,500	Married	Master Degree
5	1000005	Claudia	Perez	35	5	$1,300	Single	Master Degree

This is a data

Data types

- Quantitative
  - Weight, height, age, rate heart, body mass index (BMI).
  - household income, purchases.
- Qualitative
  - Man, Woman.
  - Low income, Average income, High income.

M W Low Average High

1 2 1 2 3