

Data Analysis and Visualization

CentraleDigitalLab@Nice

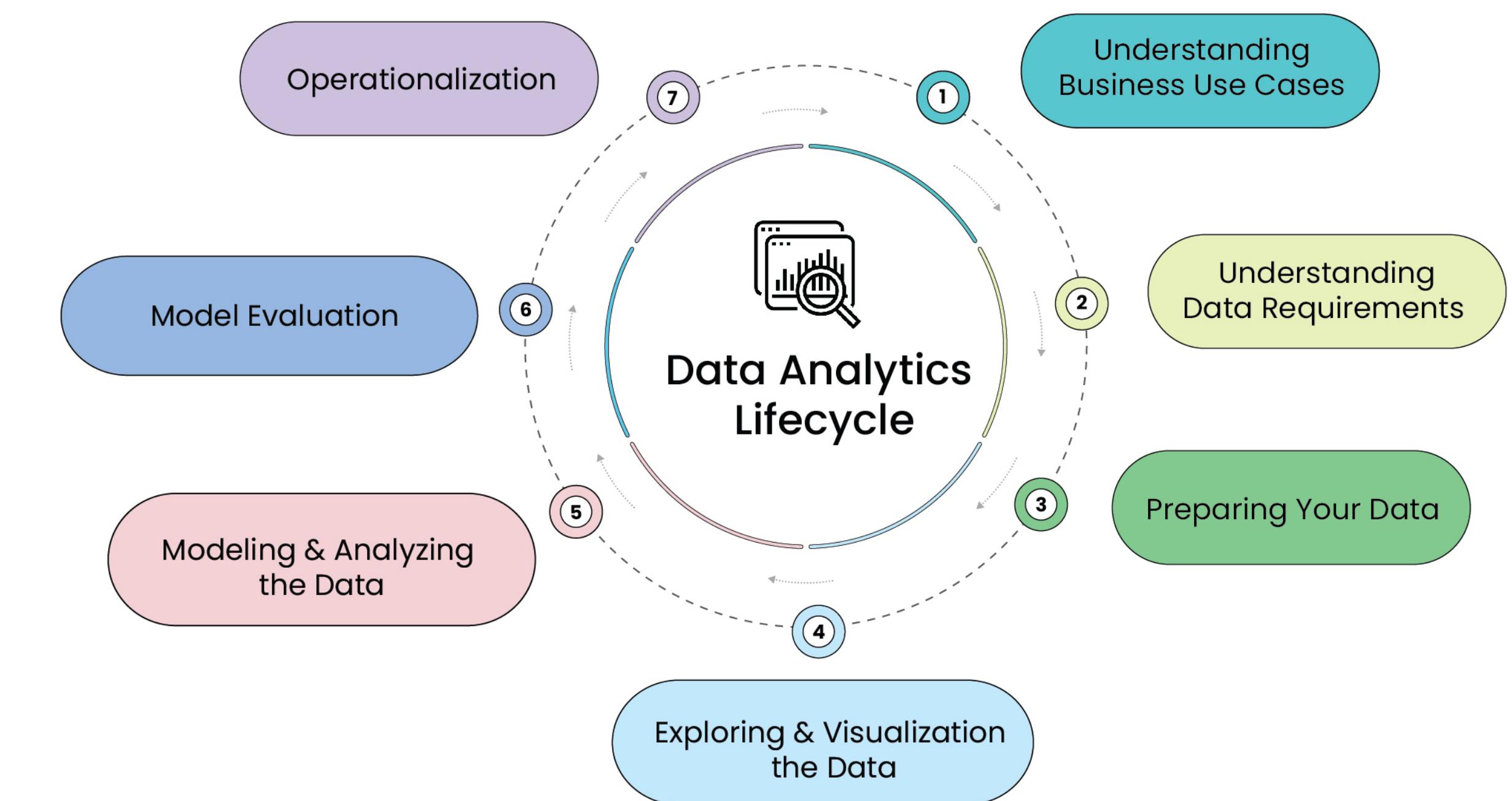
Deborah Dore - PhD - ddore@i3s.unice.fr

MARIANNE, Université Côte d'Azur, CNRS, INRIA, I3S

Introduction to Data Analysis

Understanding Data Analysis

Data analysis is the process of examining, cleaning, transforming, and modelling data to extract meaningful insights



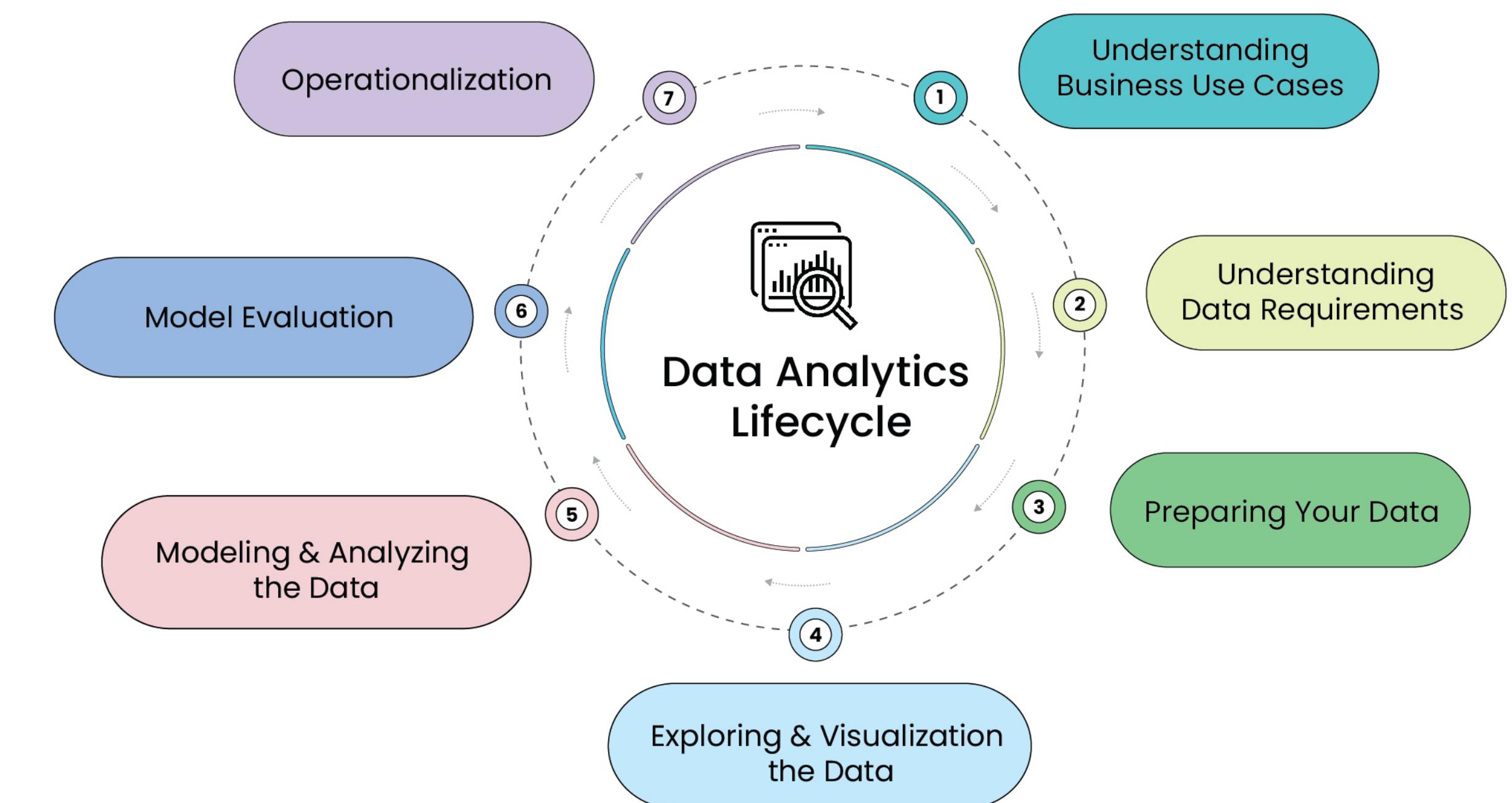
Introduction to Data Analysis

Understanding Data Analysis

Data analysis is the process of examining, cleaning, transforming, and modeling data to extract meaningful insights

Key goals:

1. Identify *patterns and trends*
2. Support the *decision-making process*
3. Predict *future outcomes*



Introduction to Data Analysis

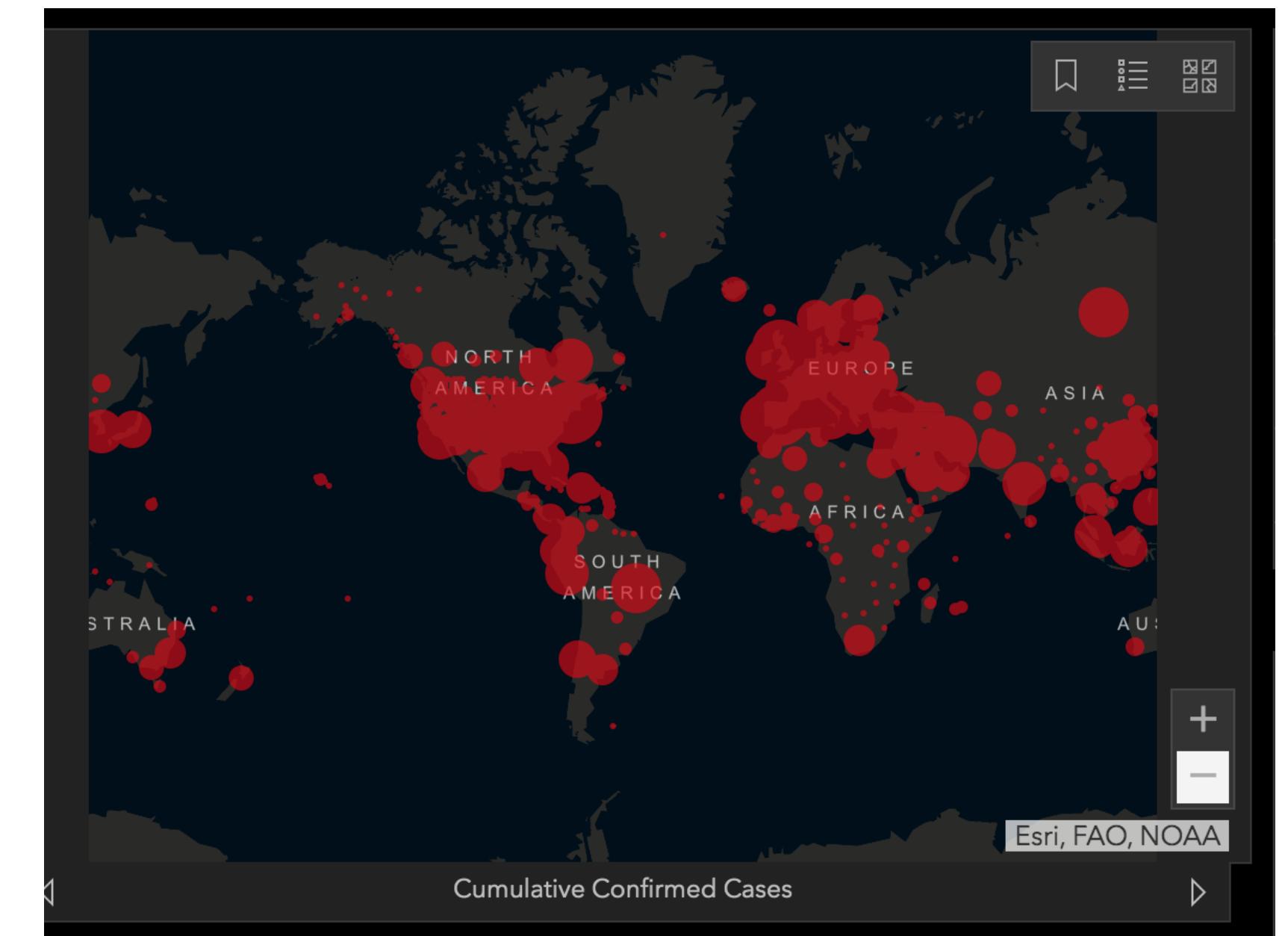
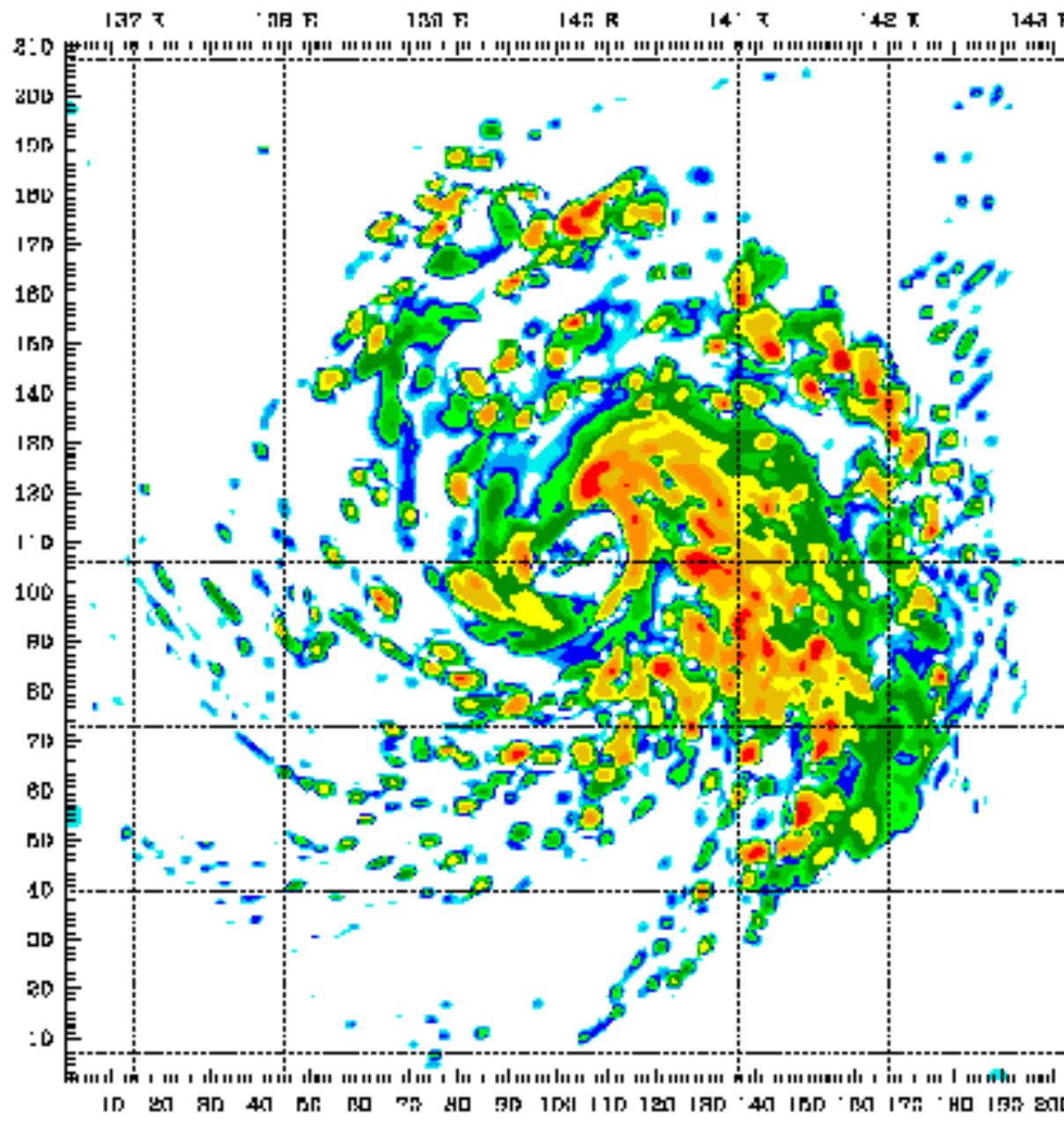
The Role of Data Analysis

- Optimizing processes (e.g., logistics, manufacturing)
- Enhancing customer experiences (e.g., personalization)
- Driving innovation (e.g., AI, machine learning)

Introduction to Data Analysis

The Role of Data Analysis

Real world applications: predicting weather, tracking disease outbreaks, or creating recommendation systems



Introduction to Data Analysis

What is data visualization?

Data visualization is the graphical representation of information and data using visual elements like charts, graphs, and maps

Introduction to Data Analysis

Why data visualization is important?

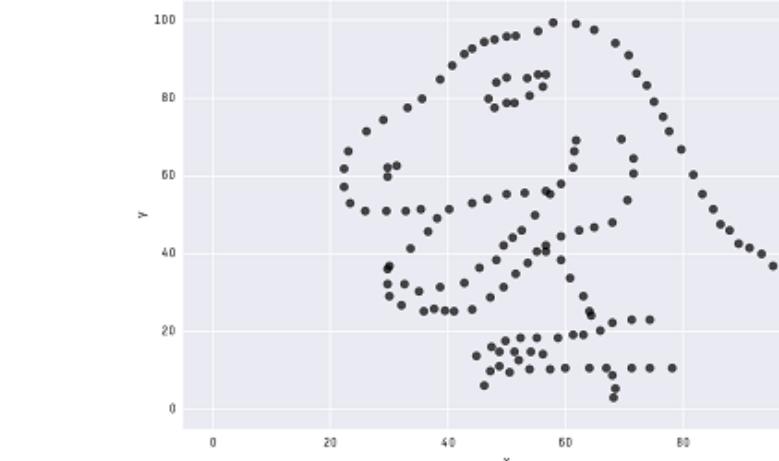
Data visualization helps to:

- Identify relevant information in dataset
- Detect patterns and correlations
- Experiment and test hypotheses
- Communicate results effectively
- Identify relevant features for machine learning models

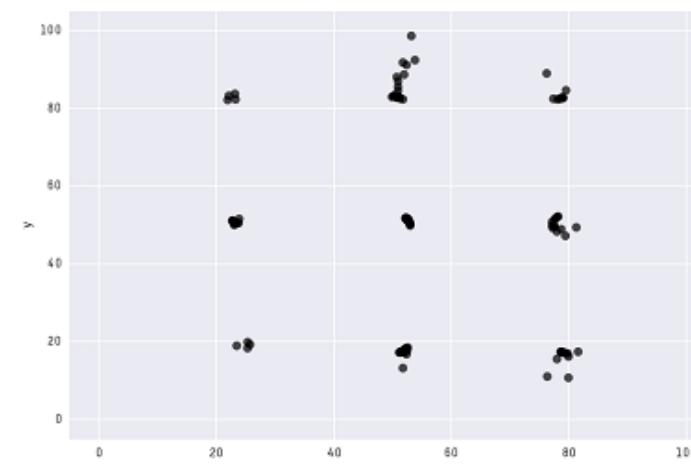
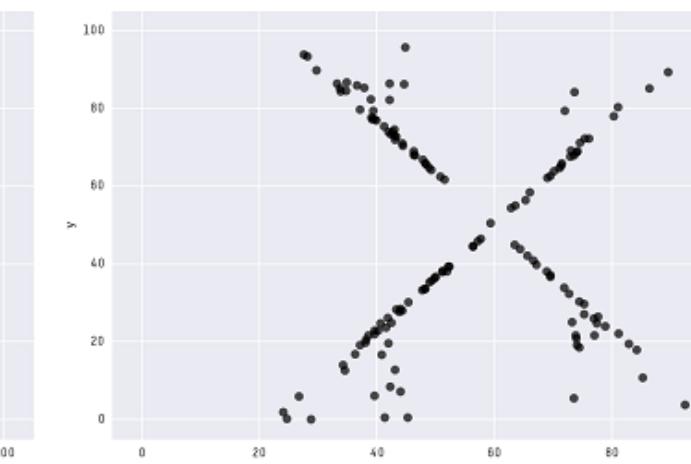
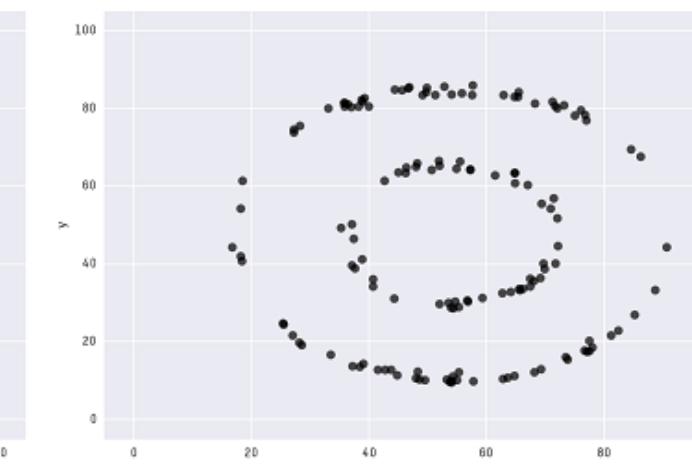
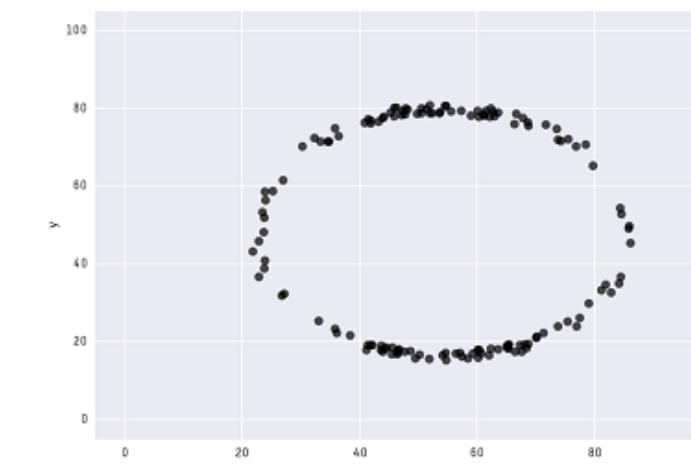
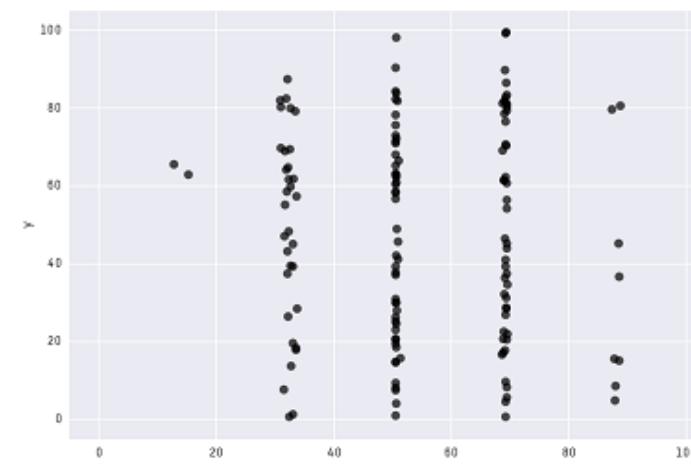
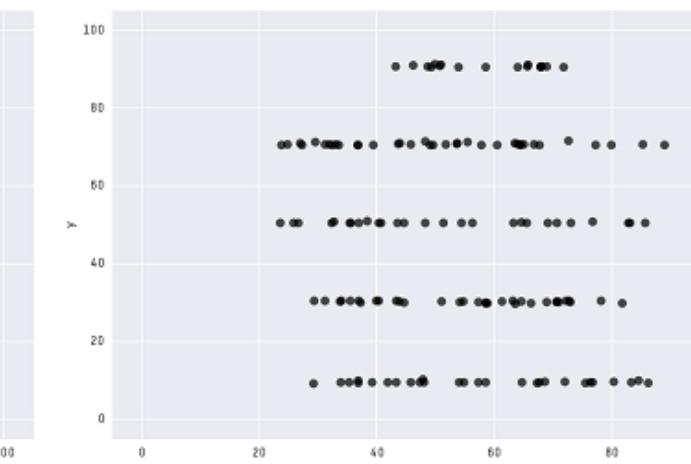
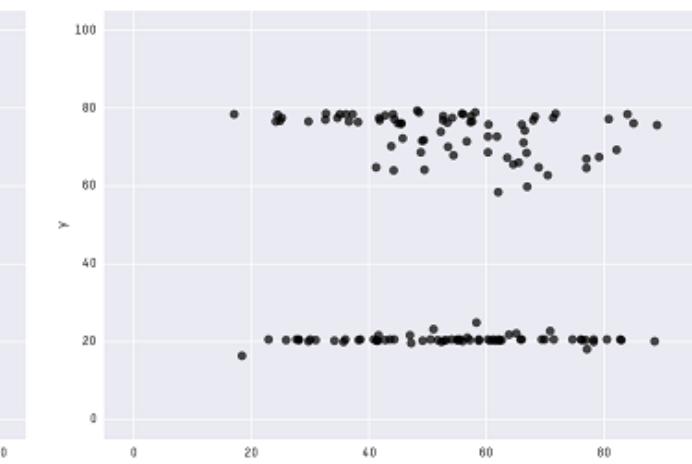
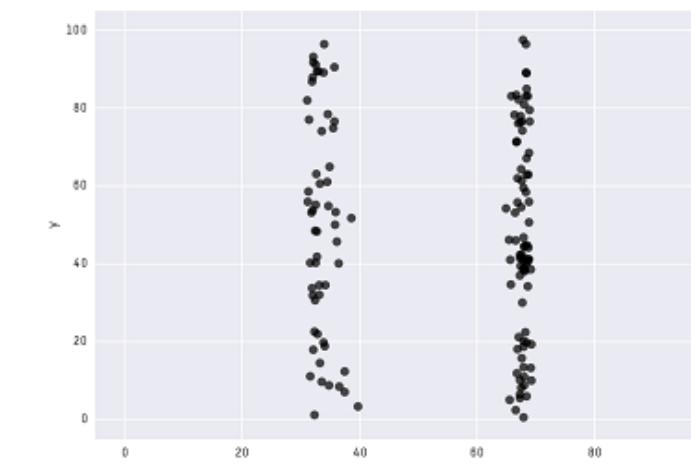
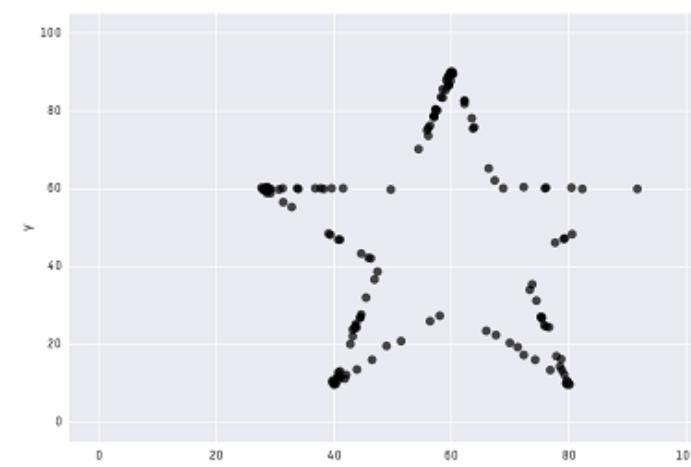
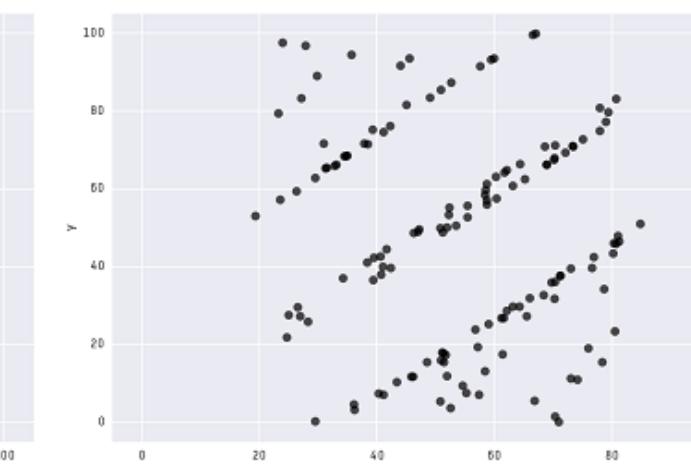
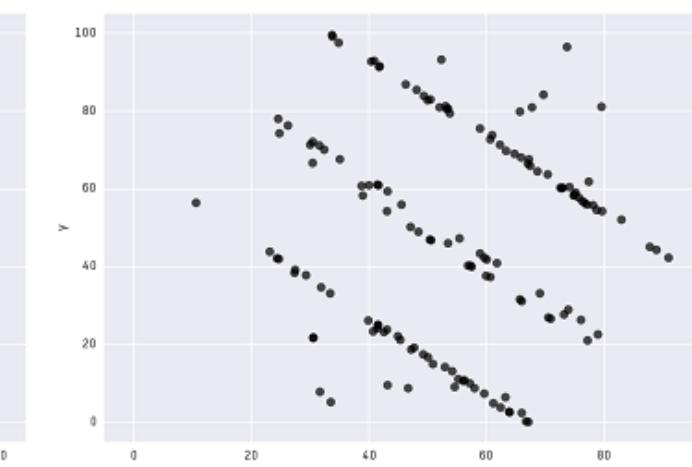
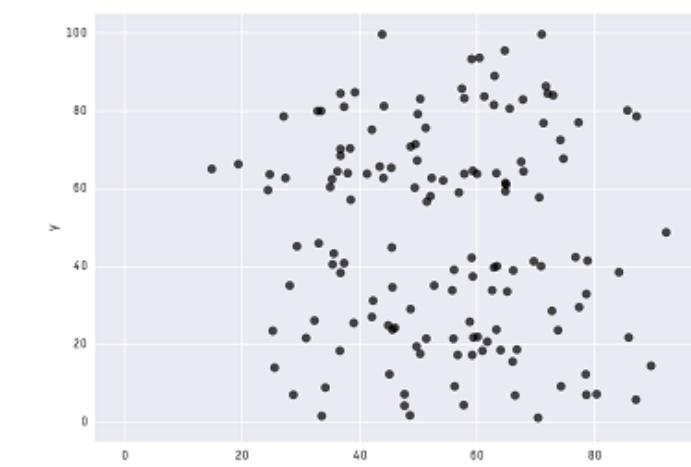
Introduction to Data Analysis

Why data visualization is important?

“We don’t know really understand the data until we visualise them”



X Mean: 54.26
Y Mean: 47.83
X SD : 16.76
Y SD : 26.93
Corr. : -0.06



Introduction to Data Analysis

Why visualization work

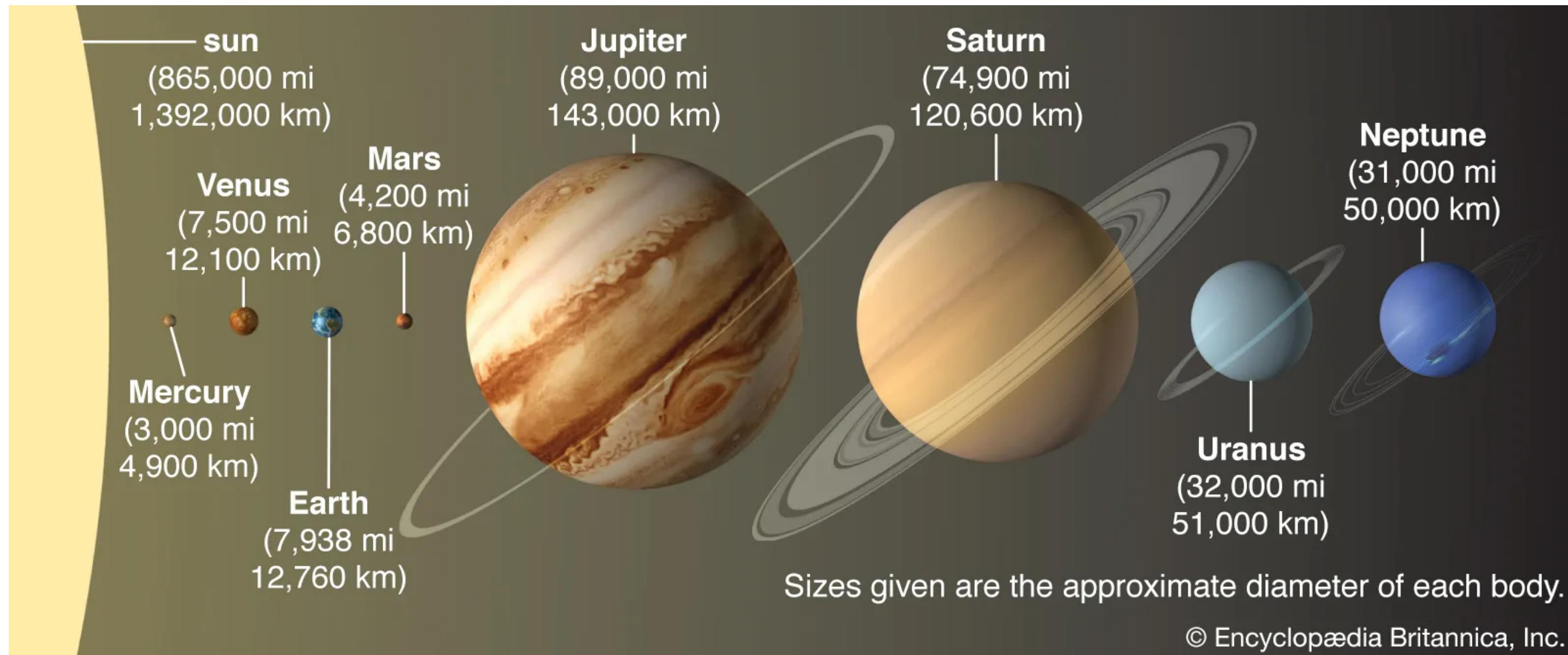
- Humans process visuals 60,000 times faster than text



Introduction to Data Analysis

Why visualization work

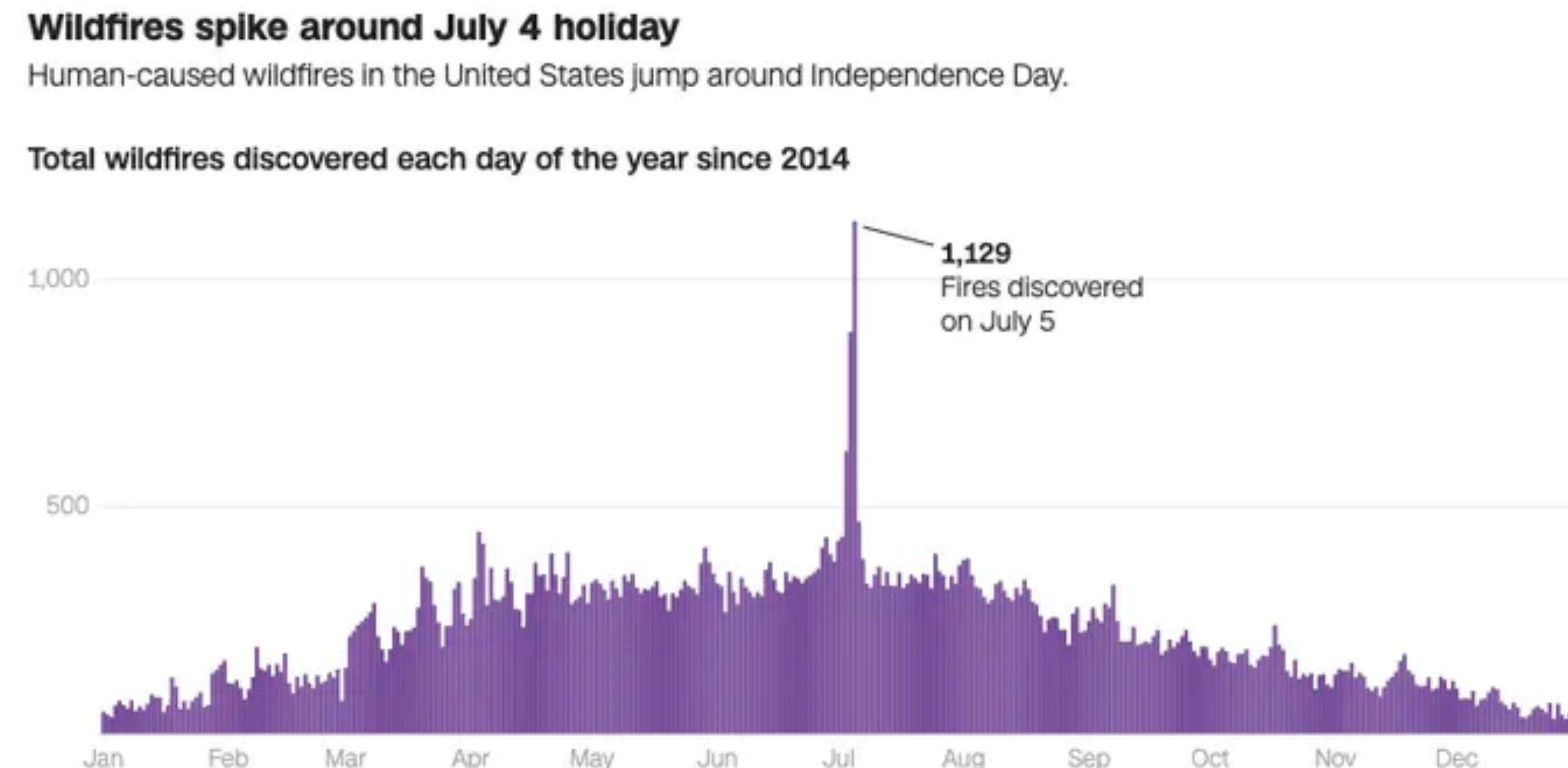
- We remember 80% of what we see but only 20% of what we read



Introduction to Data Analysis

Why visualization work

- Patterns and outliers are easier to spot visually



Human-caused fires, excluding prescribed fires. 2022 fires included through June 30. All incident times Eastern.

Sources: CNN analysis of data from the National Interagency Fire Center
Graphic: John Keefe, CNN

CNN

Introduction to Data Analysis

Example



Introduction to Data Analysis

Example



<https://www.maptive.com/data-visualization-examples/>

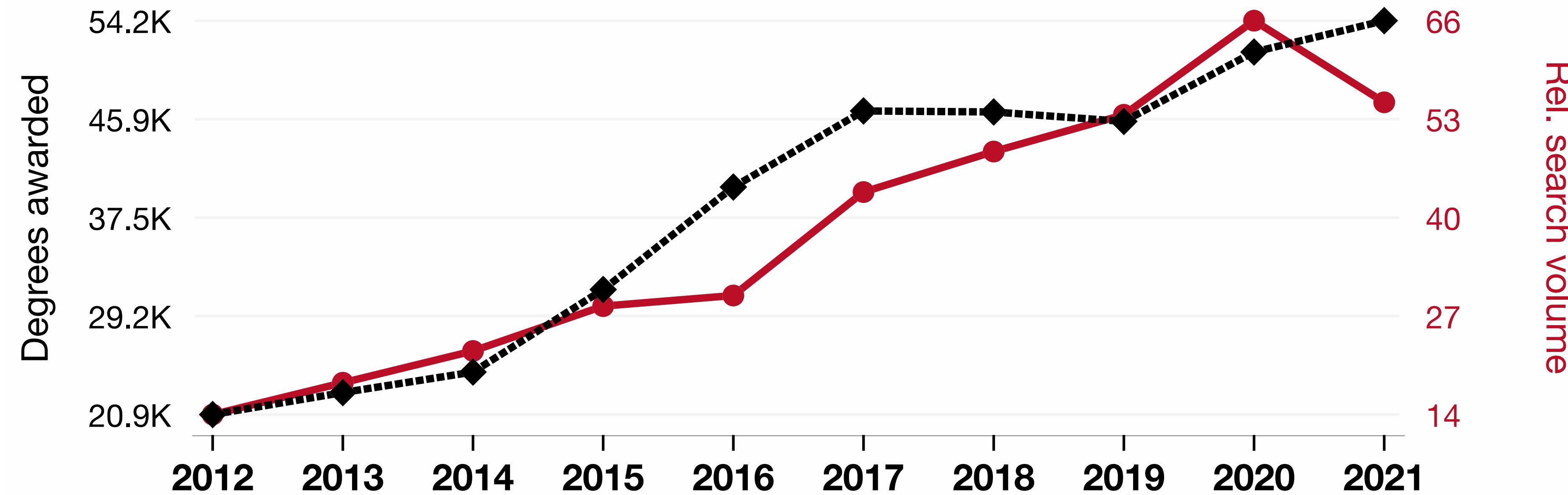
Introduction to Data Analysis

Example

**Master's degrees awarded
in information sciences**

correlates with

Google searches for 'how to learn python'



https://www.tylervigen.com/spurious/correlation/920/_masters-degrees-awarded-in-information-sciences_correlates-with_google-searches-for-how-to-learn-python

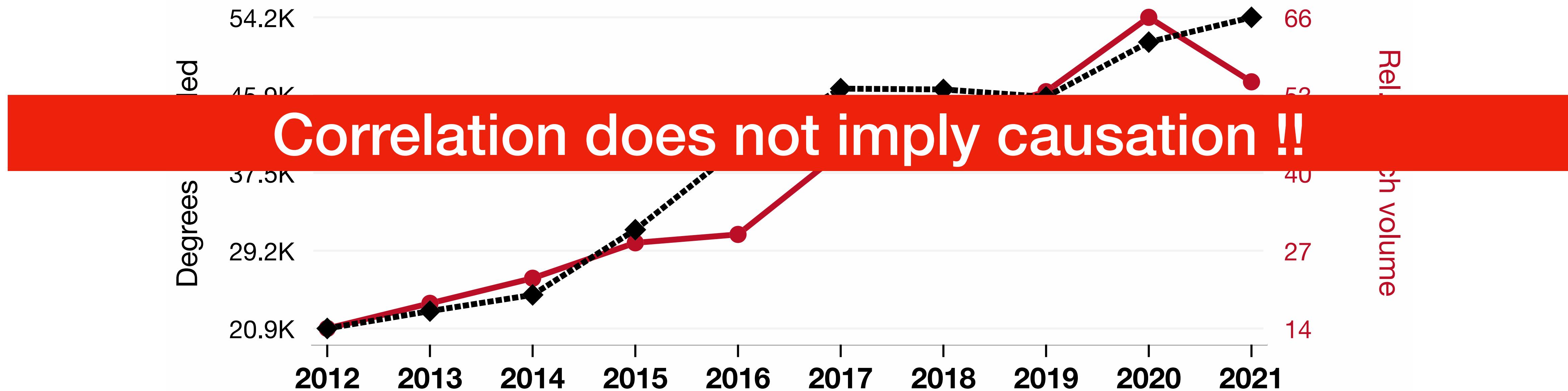
Introduction to Data Analysis

Example

Master's degrees awarded
in information sciences

correlates with

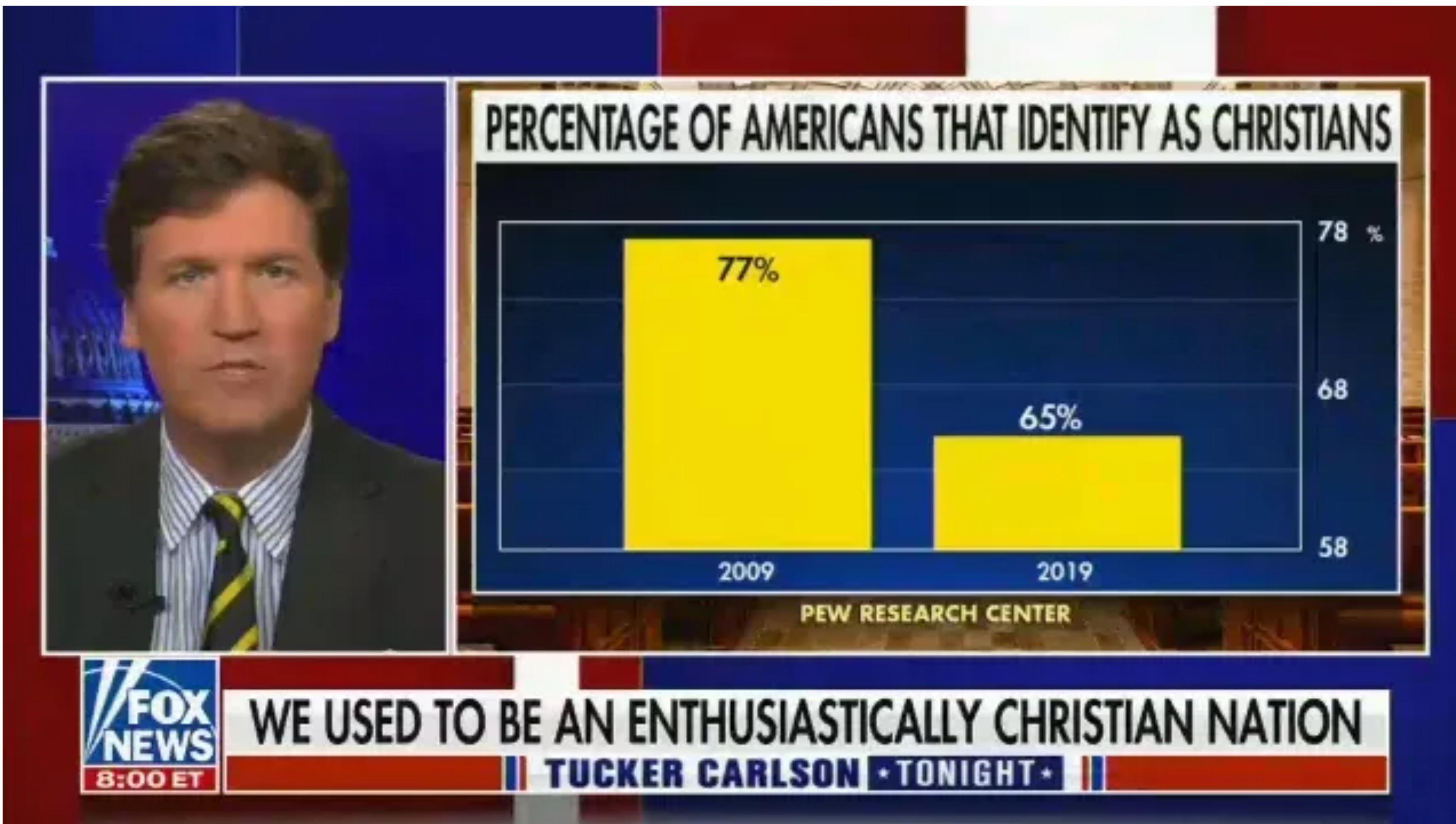
Google searches for 'how to learn python'



https://www.tylervigen.com/spurious/correlation/9207_masters-degrees-awarded-in-information-sciences_correlates-with_google-searches-for-how-to-learn-python

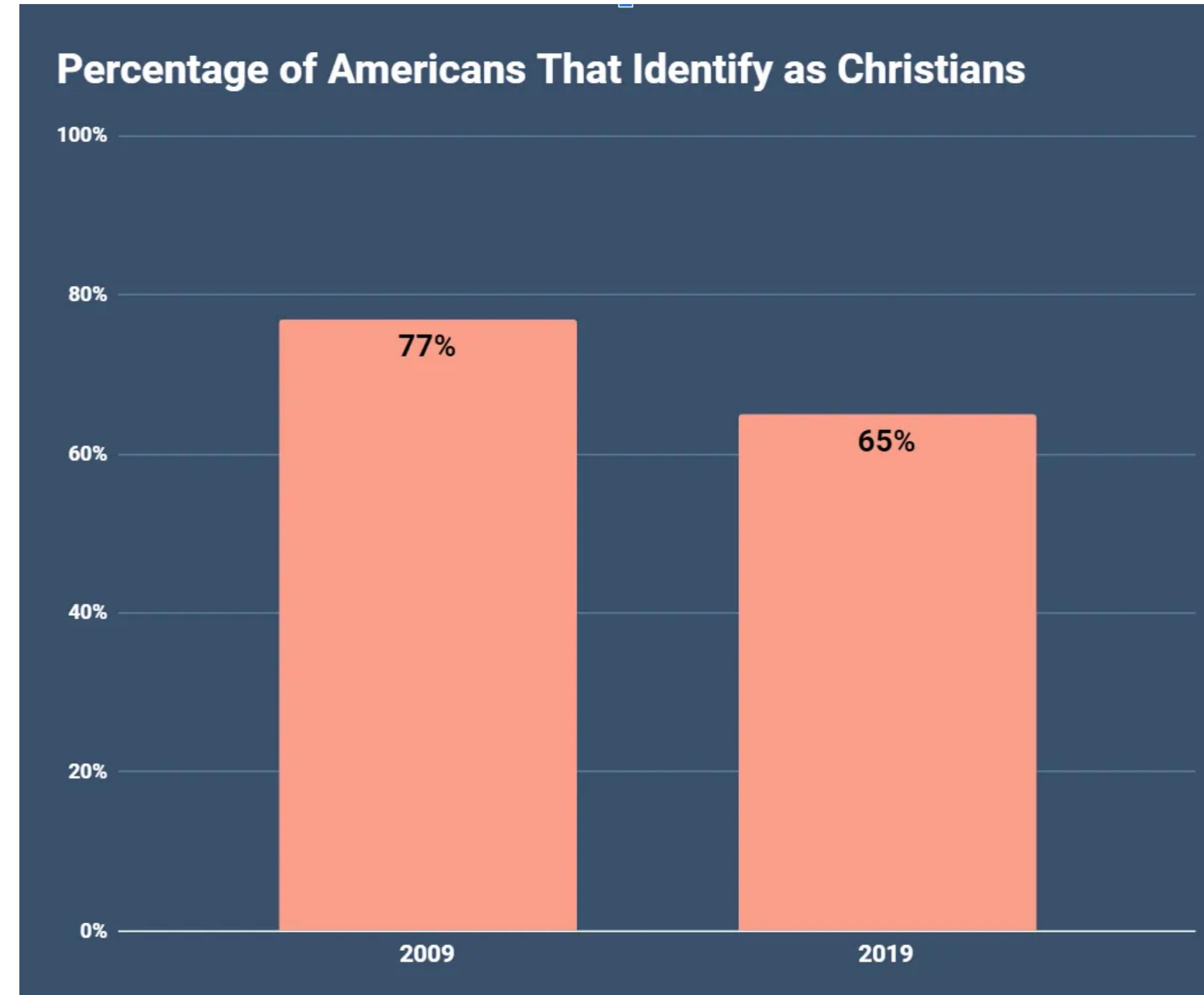
Introduction to Data Analysis

Tricky Visualisation



Introduction to Data Analysis

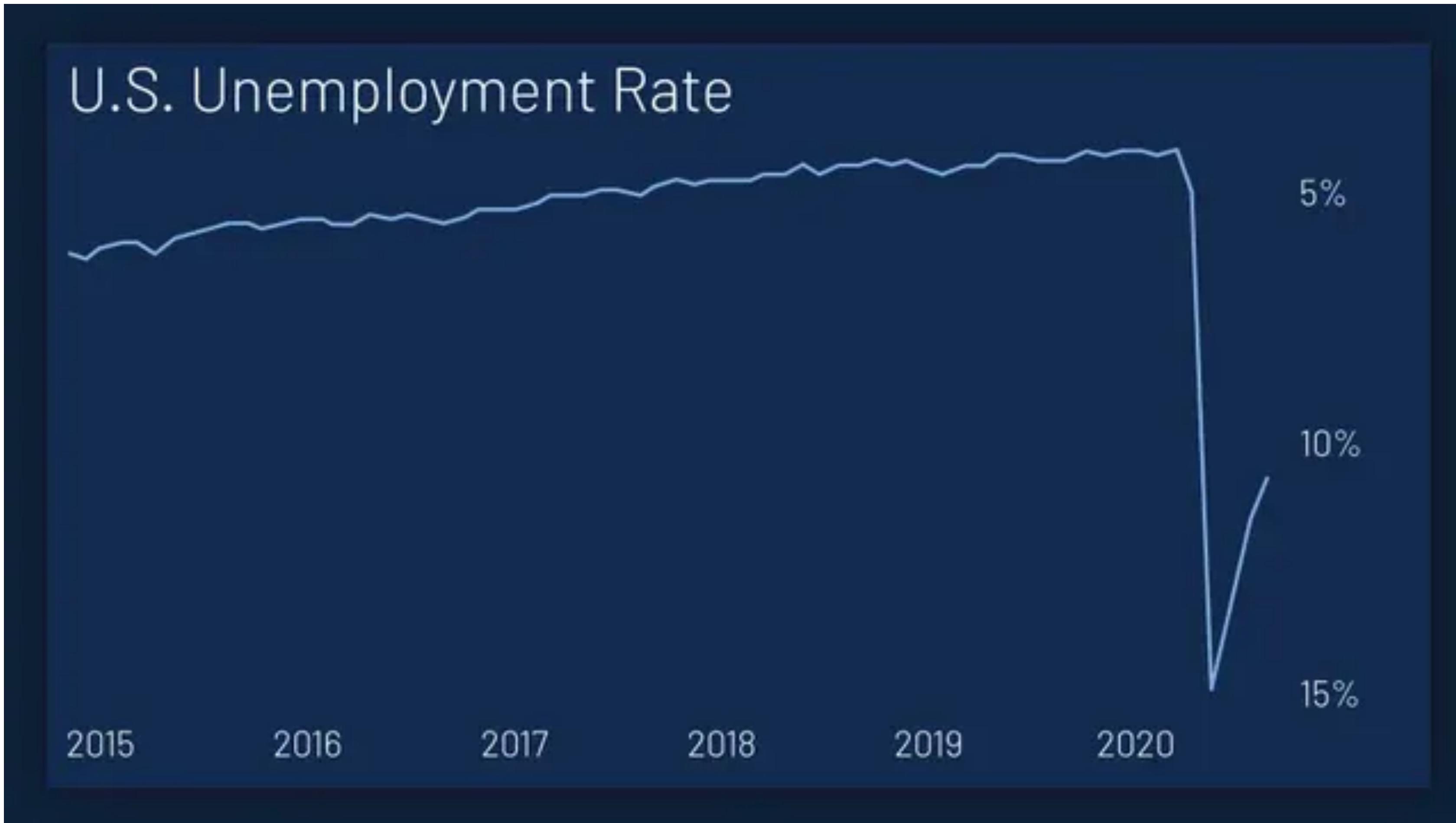
Tricky Visualisation



<https://medium.com/@thomas.ellyatt/bad-data-visualisation-real-life-examples-out-there-in-the-wild-eb5032329aeb>

Introduction to Data Analysis

Tricky Visualisation



<https://medium.com/@thomas.ellyatt/bad-data-visualisation-real-life-examples-out-there-in-the-wild-eb5032329aeb>

Introduction to Data Analysis

Summary

Data **Analysis**:

Examines and interprets raw data to extract insights

Uses techniques like statistical analysis and modeling

Data **Visualization**:

Transforms data into graphical representations

Highlights trends, patterns and relationships for better understanding

Any questions?