Ministry of Higher Education and Scientific Research
Djilali BOUNAAMA University - Khemis Miliana(UDBKM)
Faculty of Science and Technology
Department of Mathematics and Computer Science

Chapter 5

Data Communication & Visualization

AIBD-M1-UEM112 : Introduction to Data Science

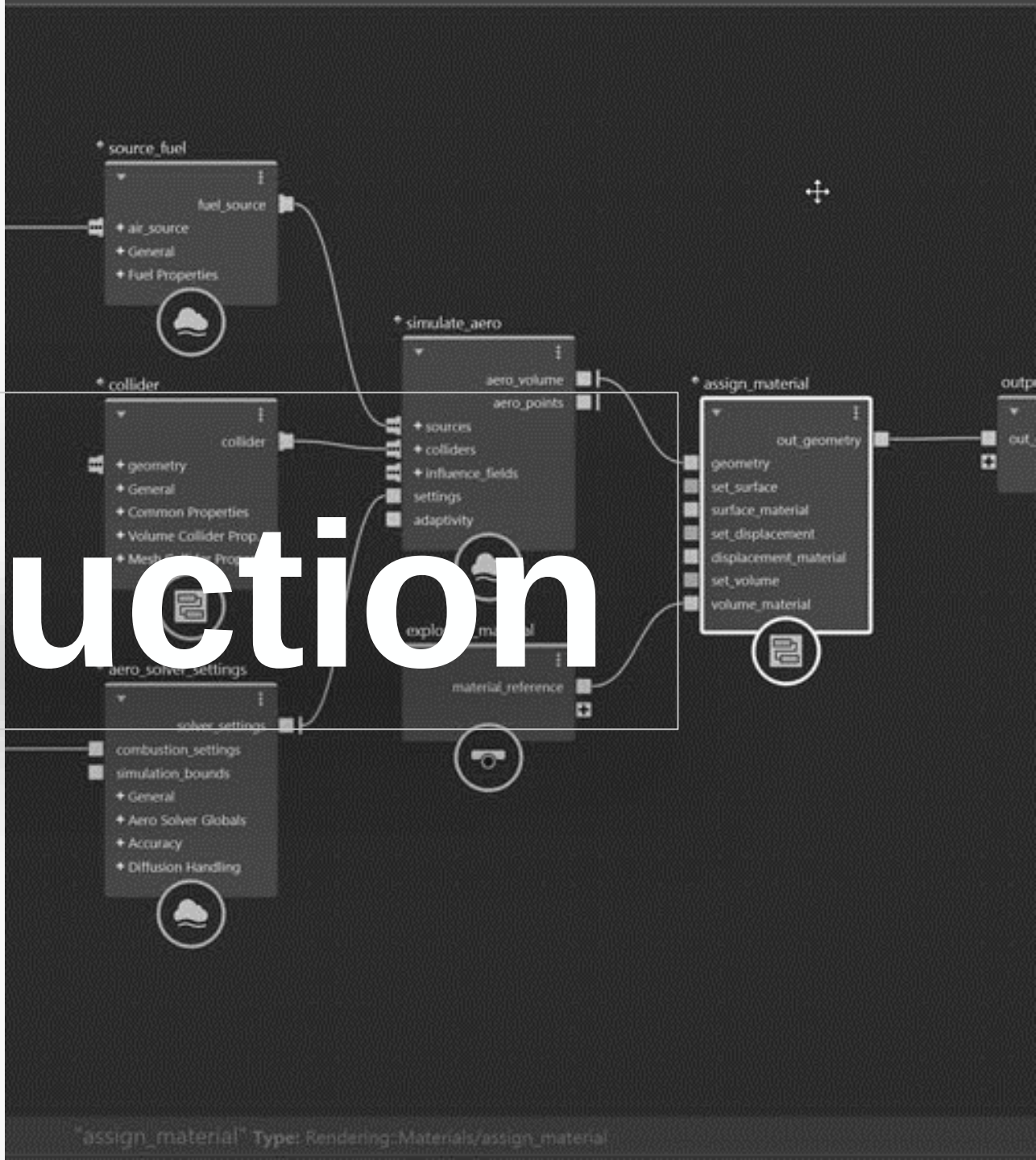
Noureddine AZZOUZA

n.azzouza@univ-dbkm.dz

Course Topics

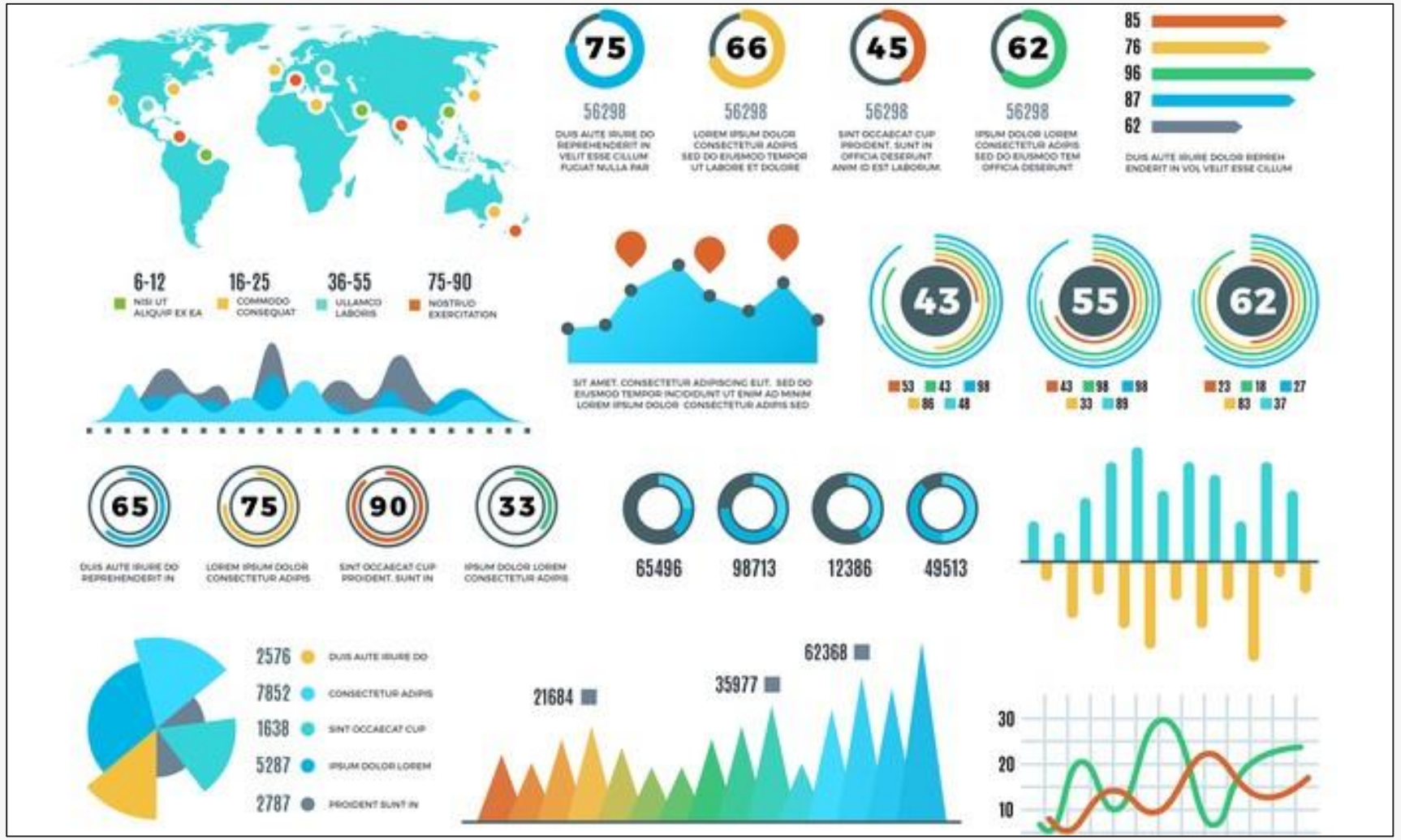
1. Introduction
2. Data visualization
3. Visualization types
4. Software & libraries
3. References

Introduction



Introduction

Introduction



Data visualization

Definition

- ✓ Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from.
- ✓ The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets.
- ✓ The term is often used interchangeably with others, including information graphics, information visualization and statistical graphics

Context

- ✓ Data visualization is one of the steps of the data science process, which states that after data has been collected, processed and modeled, it must be visualized for conclusions to be made.
- ✓ Data visualization is also an element of the broader data presentation architecture (DPA) discipline, which aims to identify, locate, manipulate, format and deliver data in the most efficient way possible.

Use

- ✓ Data visualization is important for almost every career.
- ✓ It can be used by teachers to display student test results, by computer scientists exploring advancements in artificial intelligence (AI) or by executives looking to share information with stakeholders.
- ✓ It also plays an important role in big data projects. As businesses accumulated massive collections of data during the early years of the big data trend, they needed a way to get an overview of their data quickly and easily.
- ✓ Visualization tools were a natural fit.

Benefits

- ✓ The ability to absorb information quickly, improve insights and make faster decisions;
- ✓ An increased understanding of the next steps that must be taken to improve the organization;
- ✓ An improved ability to maintain the audience's interest with information they can understand;
- ✓ An easy distribution of information that increases the opportunity to share insights with everyone involved;
- ✓ Eliminate the need for data scientists since data is more accessible and understandable; and
- ✓ An increased ability to act on findings quickly and, therefore, achieve success with greater speed and less mistakes.

Roles

- Showing change over time
- Showing a part-to-whole composition
- Depicting flows and processes
- Looking at how data is distributed
- Comparing values between groups
- Observing relationships between variables
- Looking at geographical data

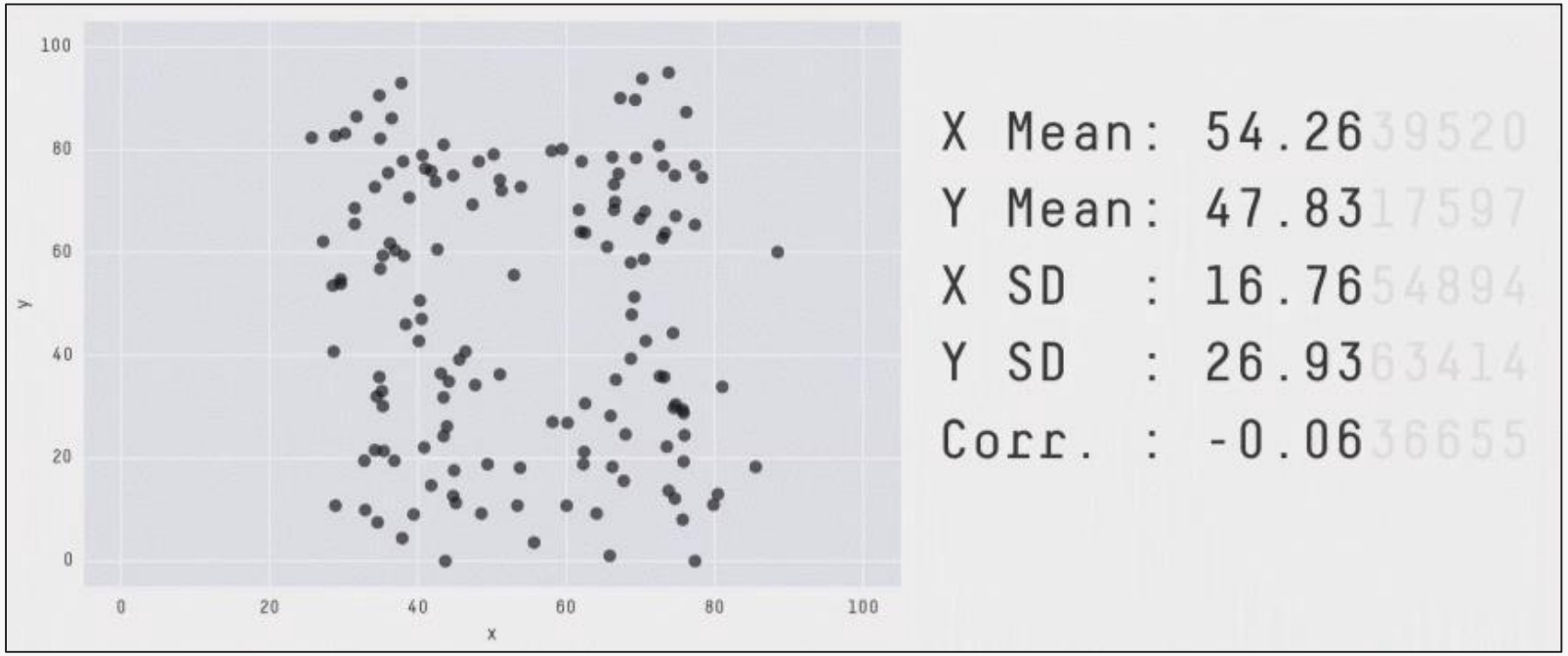
Categories

- ✓ Two types of visualization
  1. Data exploration visualization: figuring out what is true
  2. Data presentation visualization: convincing other people it is true

- ✓ “Data exploration” is much broader than just visualization

Visualization vs. Statistics

✓ Visualization almost always presents a more informative (though less quantitative) view of your data than statistics (the noun, not the field)



Data visualization

Data types

- ✓ **Nominal**: categorical data, no ordering
  - Example – Pet: {dog, cat, rabbit, …}
  - Operations: =, ≠
- ✓ **Ordinal**: categorical data, with ordering
  - Example – Rating: {1,2,3,4,5}
  - Operations: =, ≠, ≥, ≤, >, <
- ✓ **Interval**: numerical data, zero doesn't mean zero "quantity"
  - Example – Temperature Fahrenheit, IQ scores
  - Operations: =, ≠, ≥, ≤, >, <, +, −
- ✓ **Ratio**: numerical data, zero has meaning related to zero "quantity"
  - Example – Temperature Kelvin, Age, Height
  - Operations: =, ≠, ≥, ≤, >, <, +, −, ÷

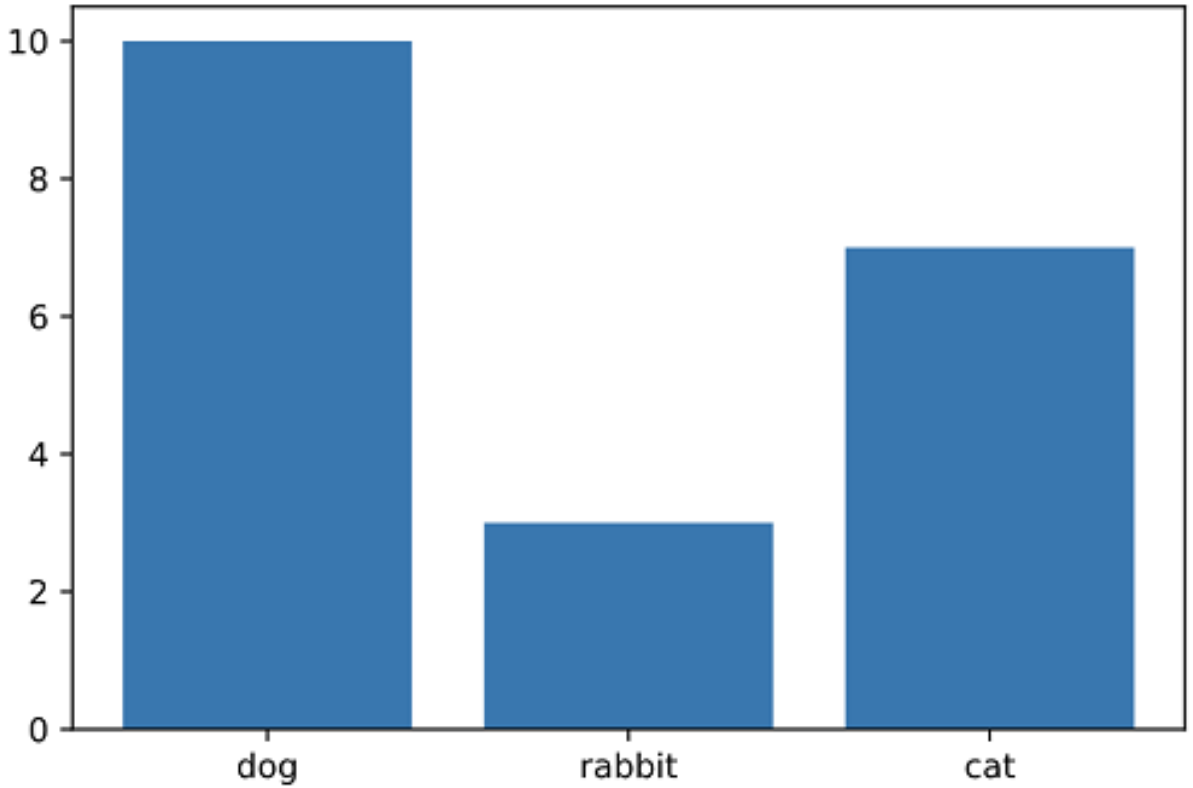
Visualization Types

- ✓ Most discussion of visualization types emphasizes what elements the chart is trying to convey
- ✓ Instead, we are going to focus on the type and dimensionality of the underlying data
- ✓ Visualization types (not an exhaustive list):
  - ➢ 1D: bar chart, pie chart, histogram
  - ➢ 2D: scatter plot, line plot, box and whisker plot, heatmap
  - ➢ 3D+: scatter matrix, bubble chart

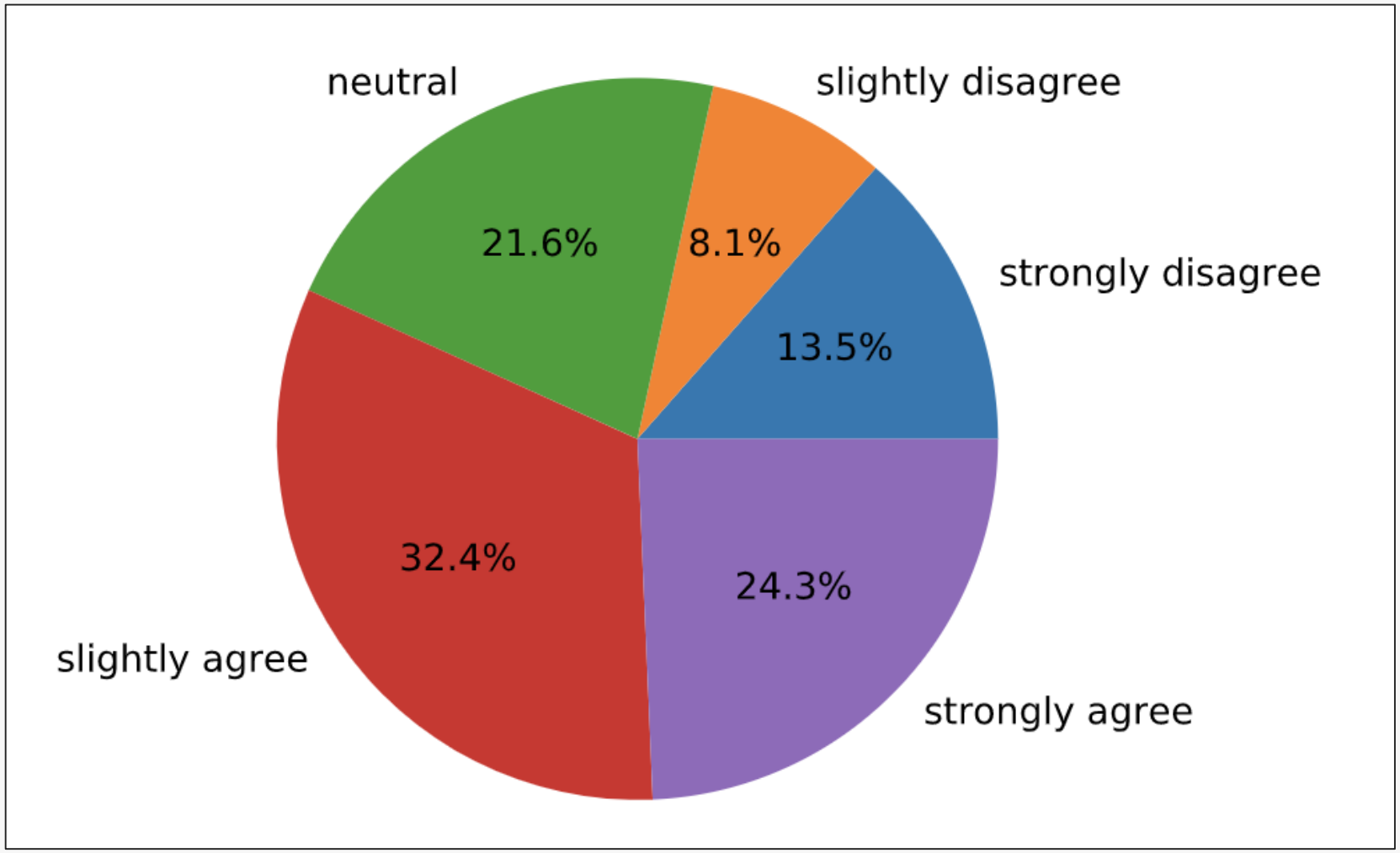
1D Data : Bar chart

	Data
Nominal	✔
Ordinal	✔
Interval	✗
Ratio	✗

Suggestions, not rules

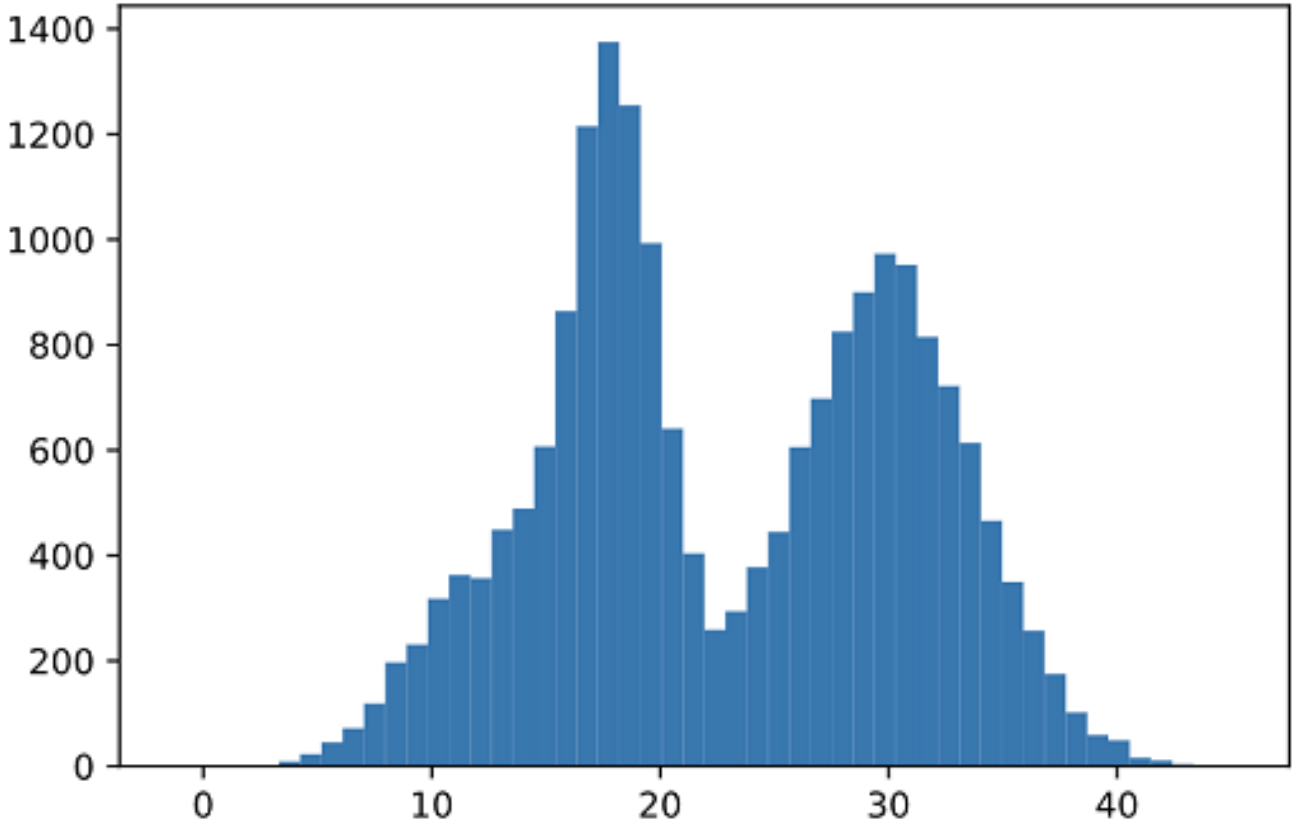


1D Data : Pie chart



1D Data : Histogram

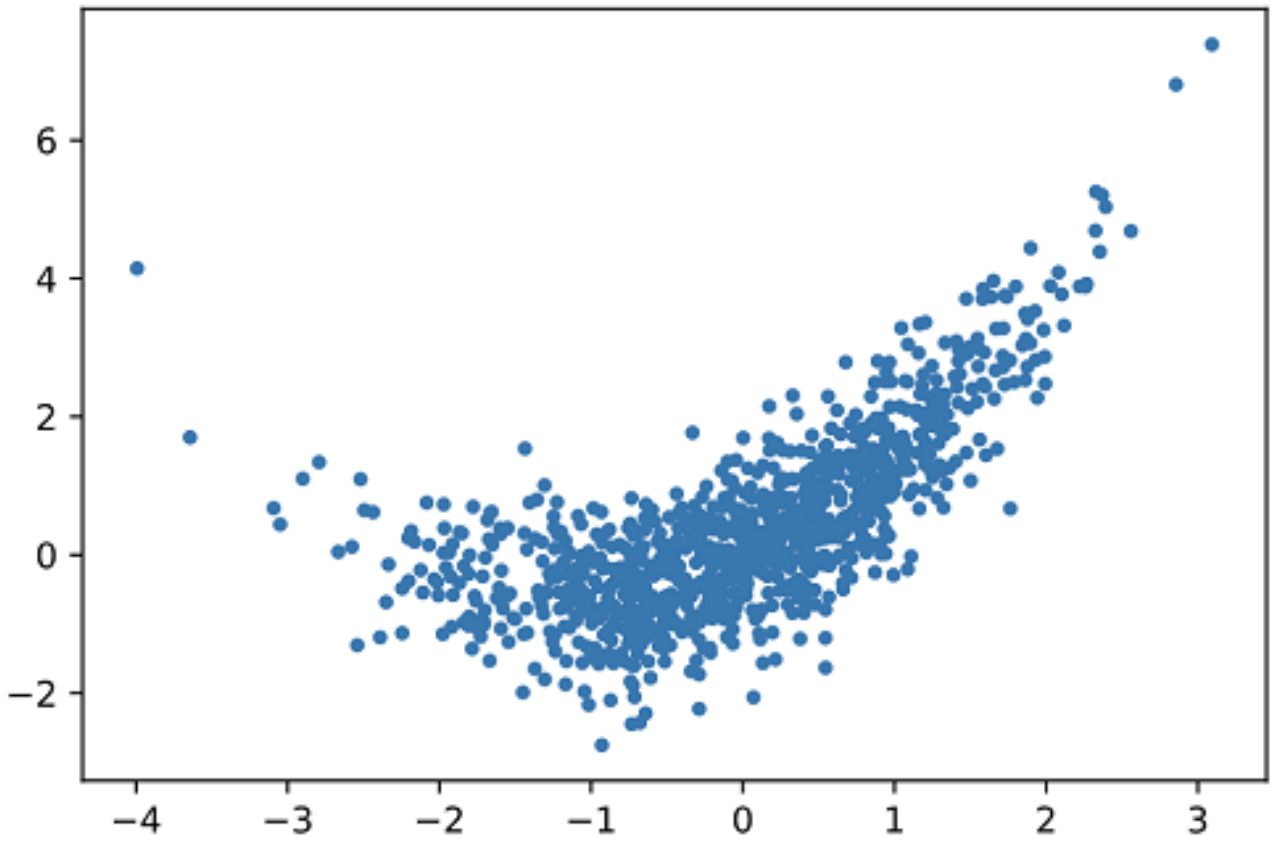
	Data
Nominal	✗
Ordinal	✗
Interval	✔
Ratio	✔



2D Data : Scatter plot

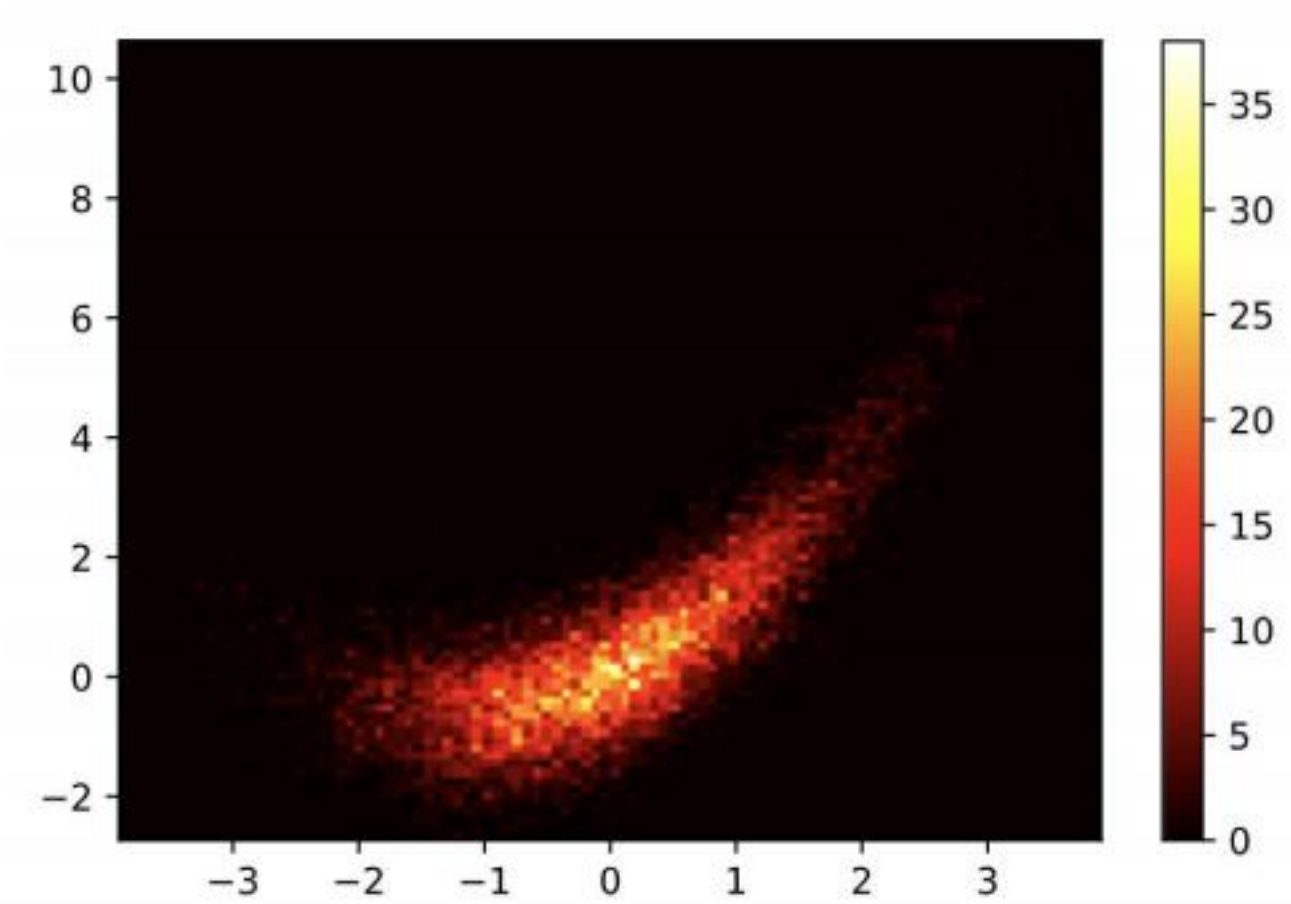
	Dim 1	Dim 2
Nominal	✗	✗
Ordinal	✗	✗
Interval	✓	✓
Ratio	✓	✓

Why not ordinal data in first dimension?



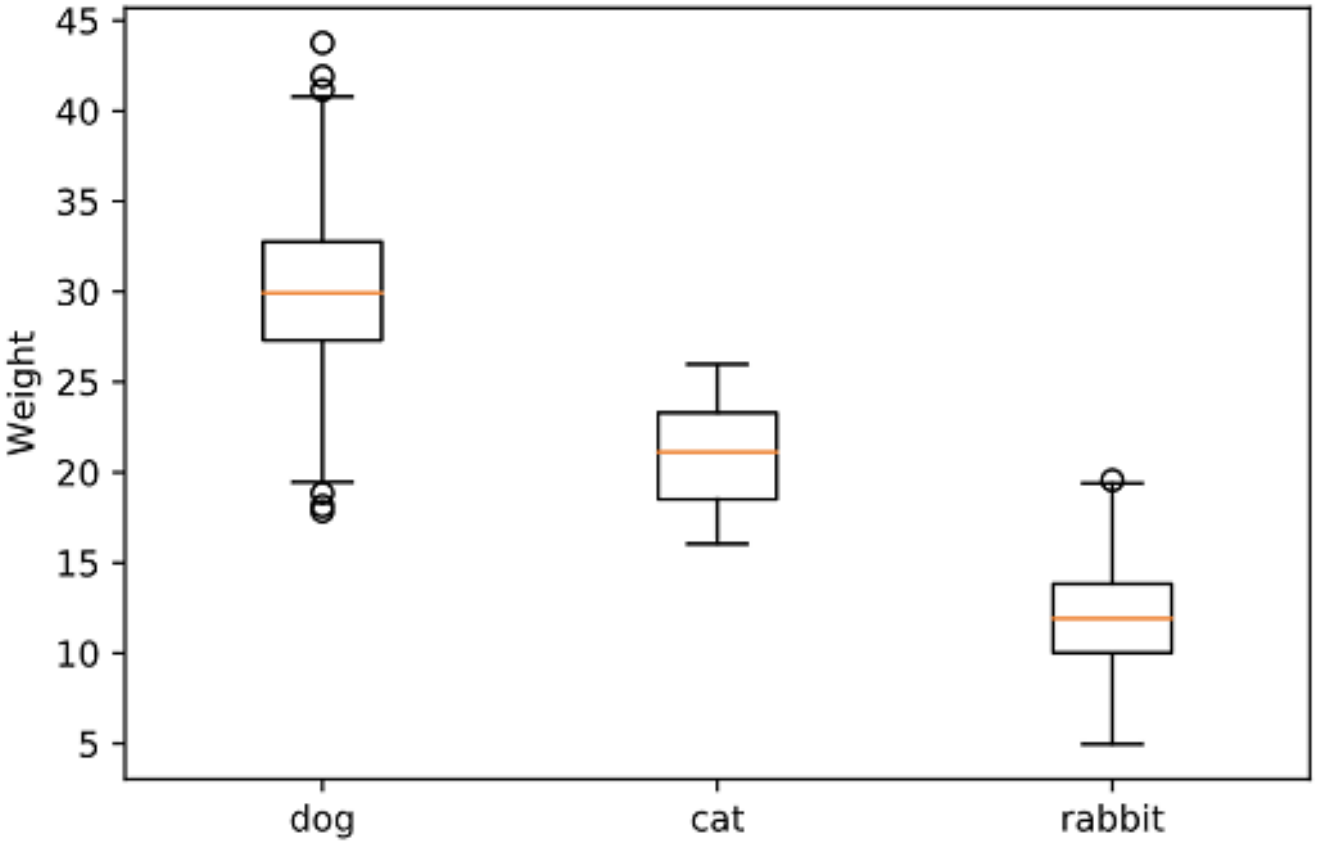
2D Data : Heatmap

	Dim 1	Dim 2
Nominal	✗	✗
Ordinal	✗	✗
Interval	✓	✓
Ratio	✓	✓



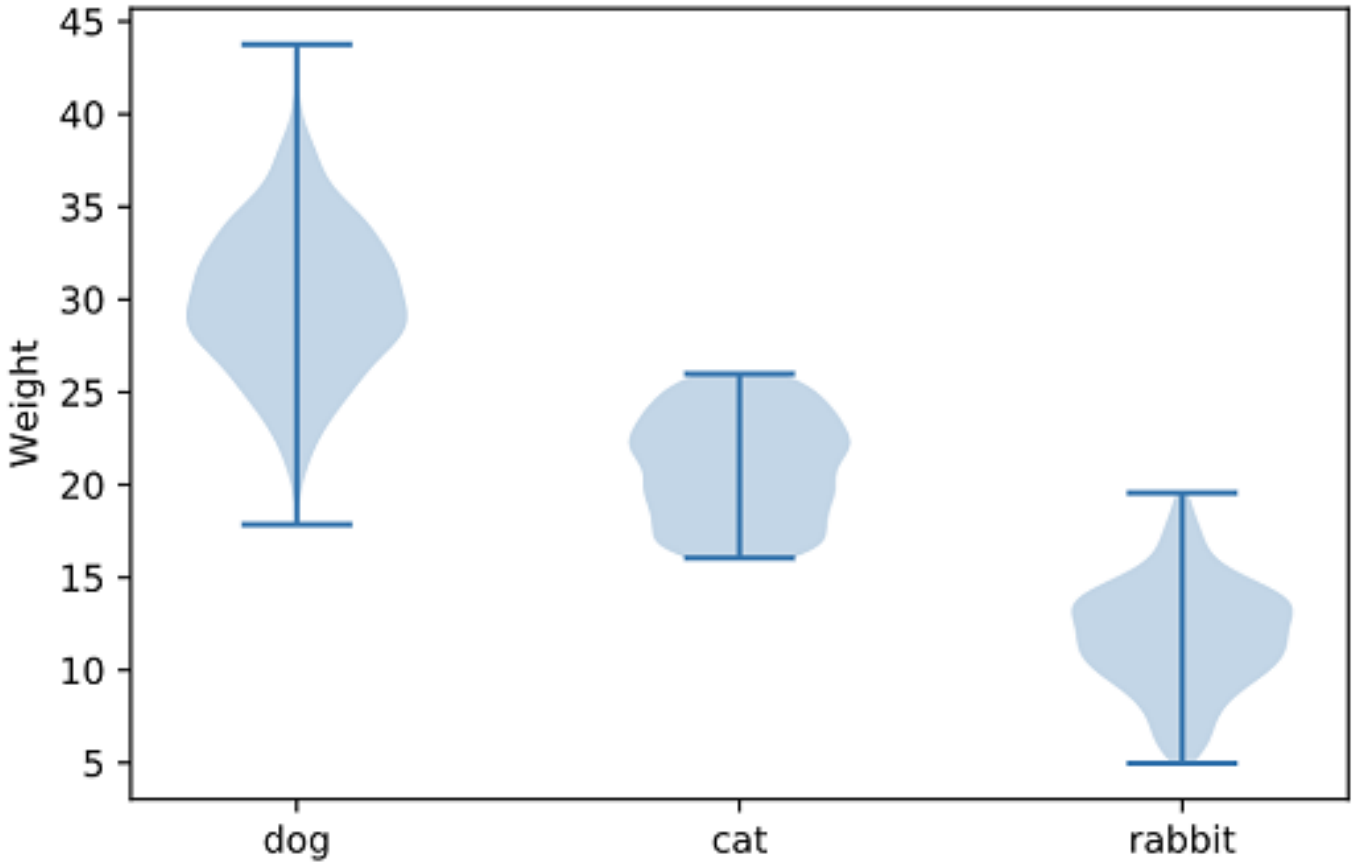
2D Data : Box and whiskers

	Dim 1	Dim 2
Nominal	✔	✗
Ordinal	✔	✗
Interval	✗	✔
Ratio	✗	✔



2D Data : Violin plot

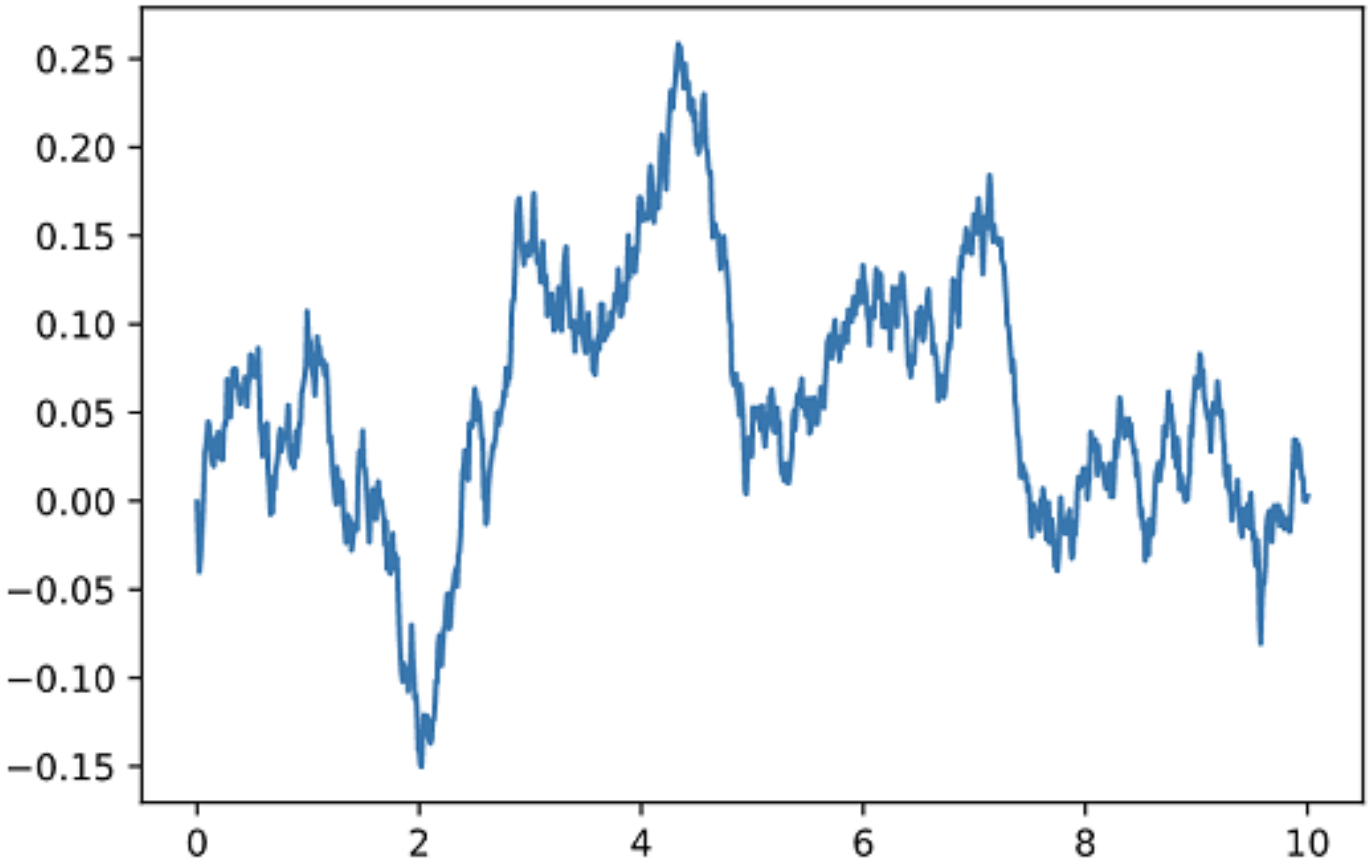
	Dim 1	Dim 2
Nominal	✔	✗
Ordinal	✔	✗
Interval	✗	✔
Ratio	✗	✔



2D Data : Line plot

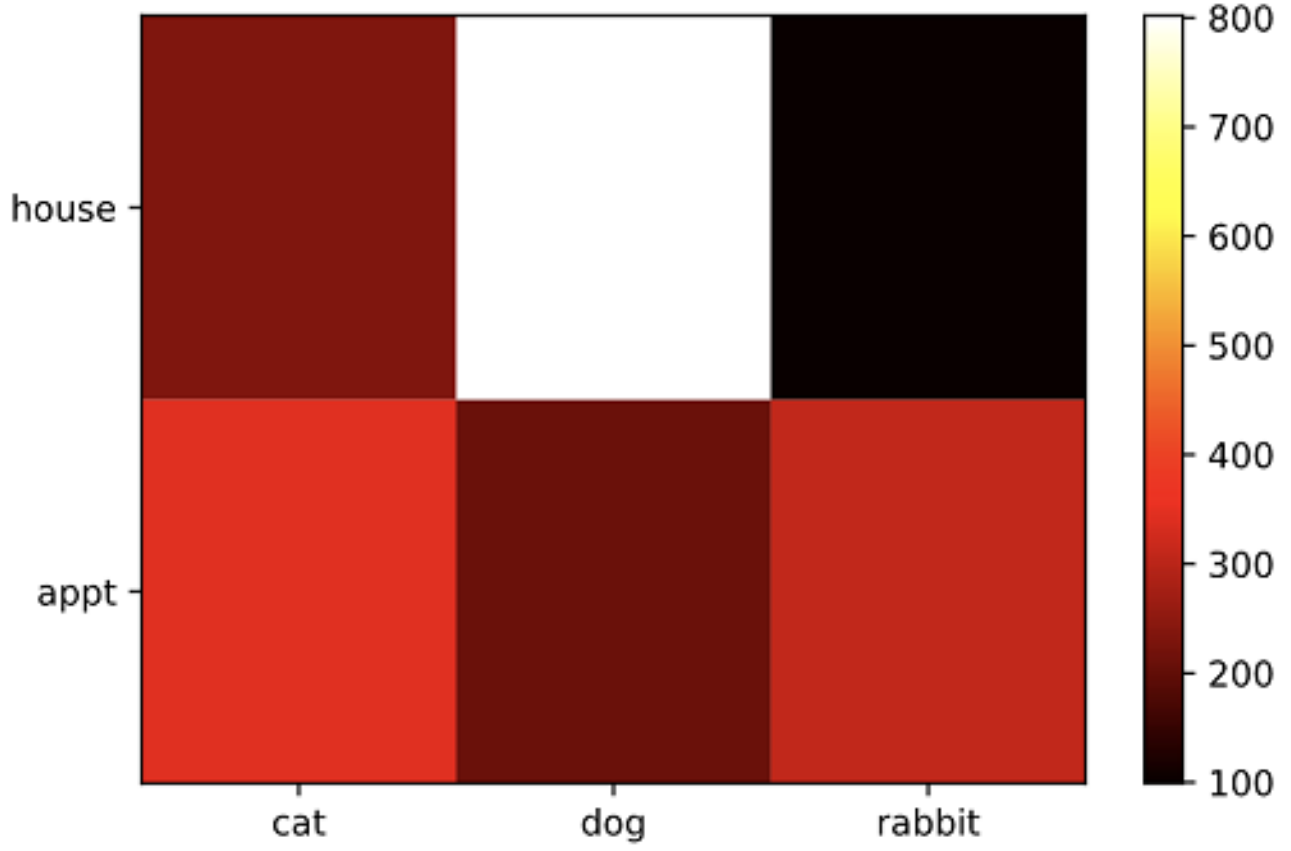
	Dim 1	Dim 2
Nominal	✗	✗
Ordinal	✗	✗
Interval	✓	✓
Ratio	✓	✓

Why not ordinal data in first dimension?



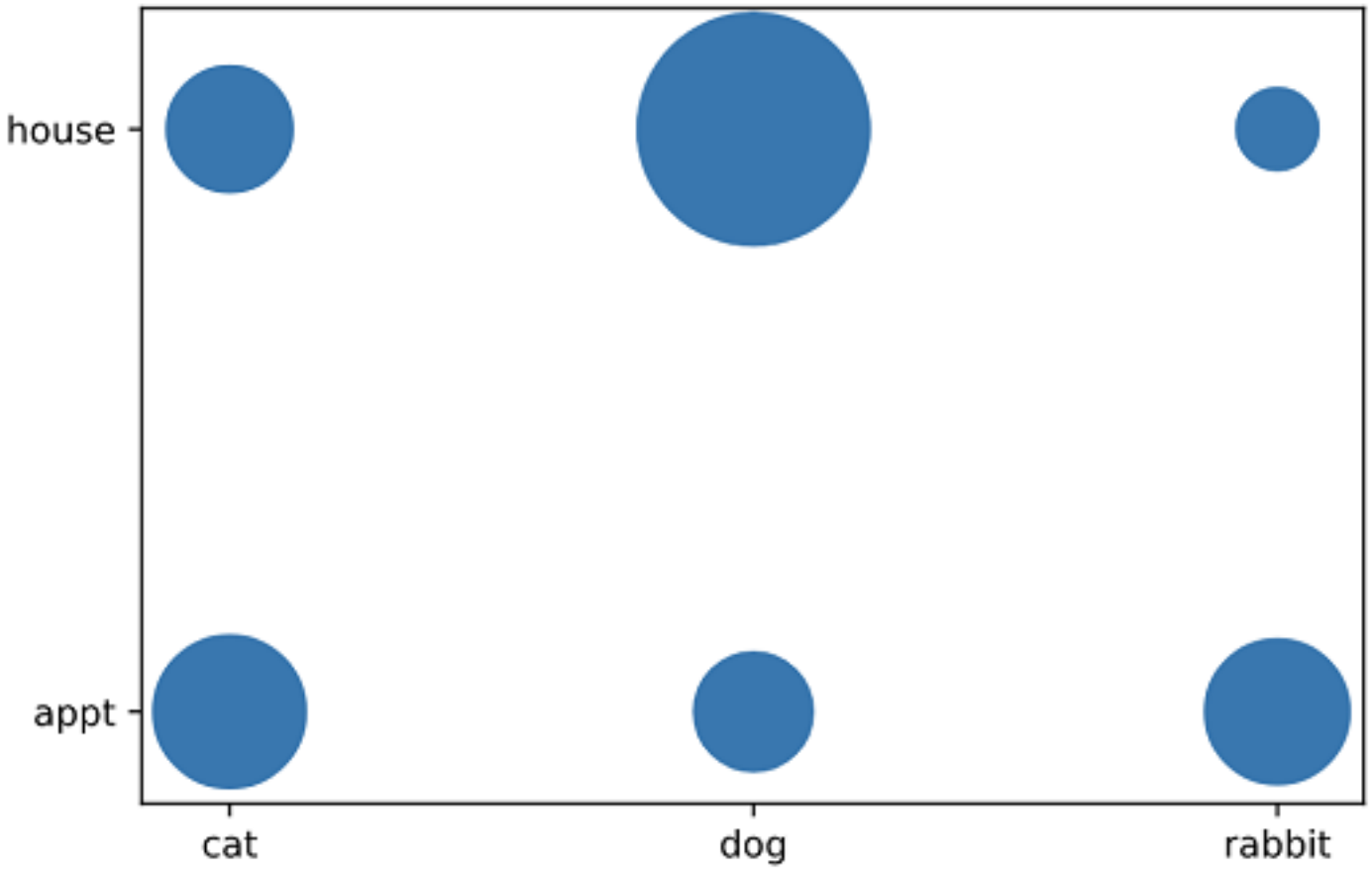
2D Data : Heatmap (matrix)

	Dim 1	Dim 2
Nominal	✔	✔
Ordinal	✔	✔
Interval	✗	✗
Ratio	✗	✗



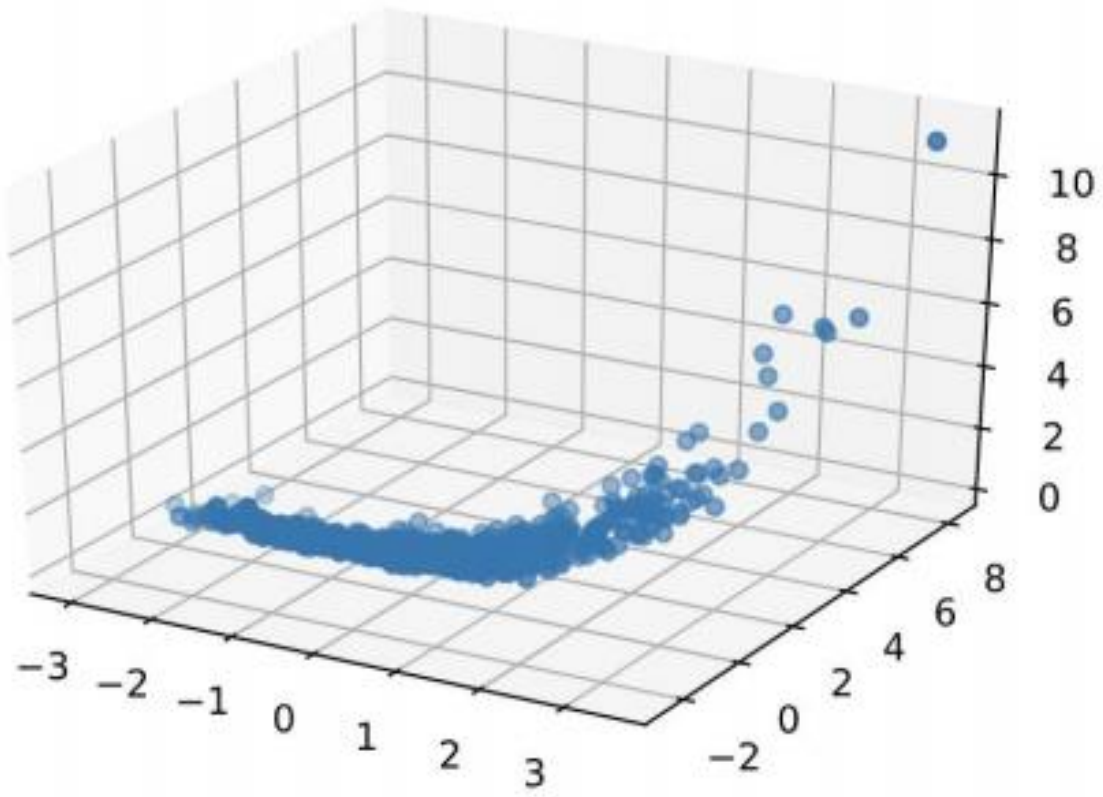
2D Data : Bubble plot

	Dim 1	Dim 2
Nominal	✔	✔
Ordinal	✔	✔
Interval	✗	✗
Ratio	✗	✗



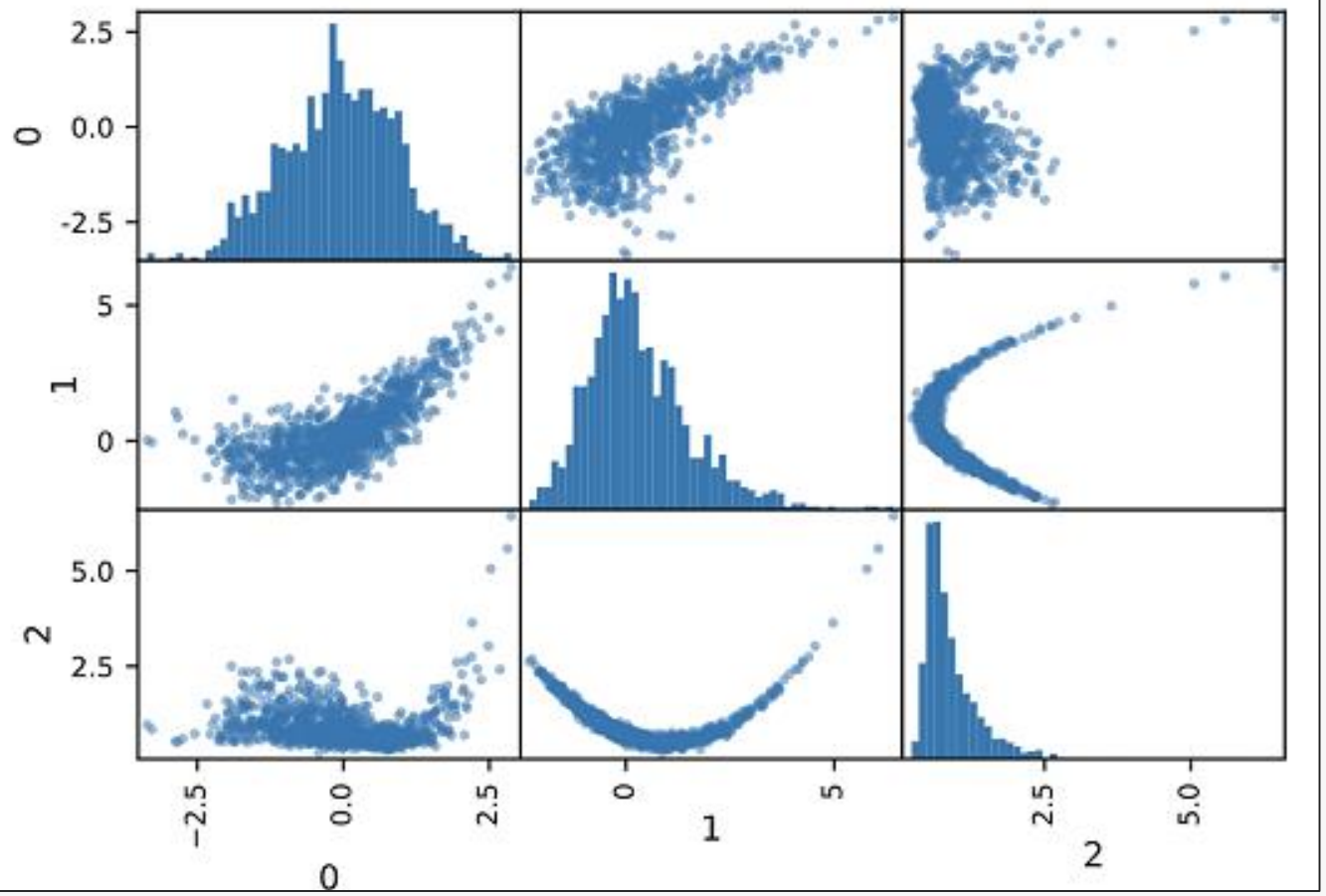
3D Data : 3D scatter plot

	Dim 1	Dim 2	Dim 3
Nominal	✗	✗	✗
Ordinal	✗	✗	✗
Interval	✗	✗	✗
Ratio	✗	✗	✗



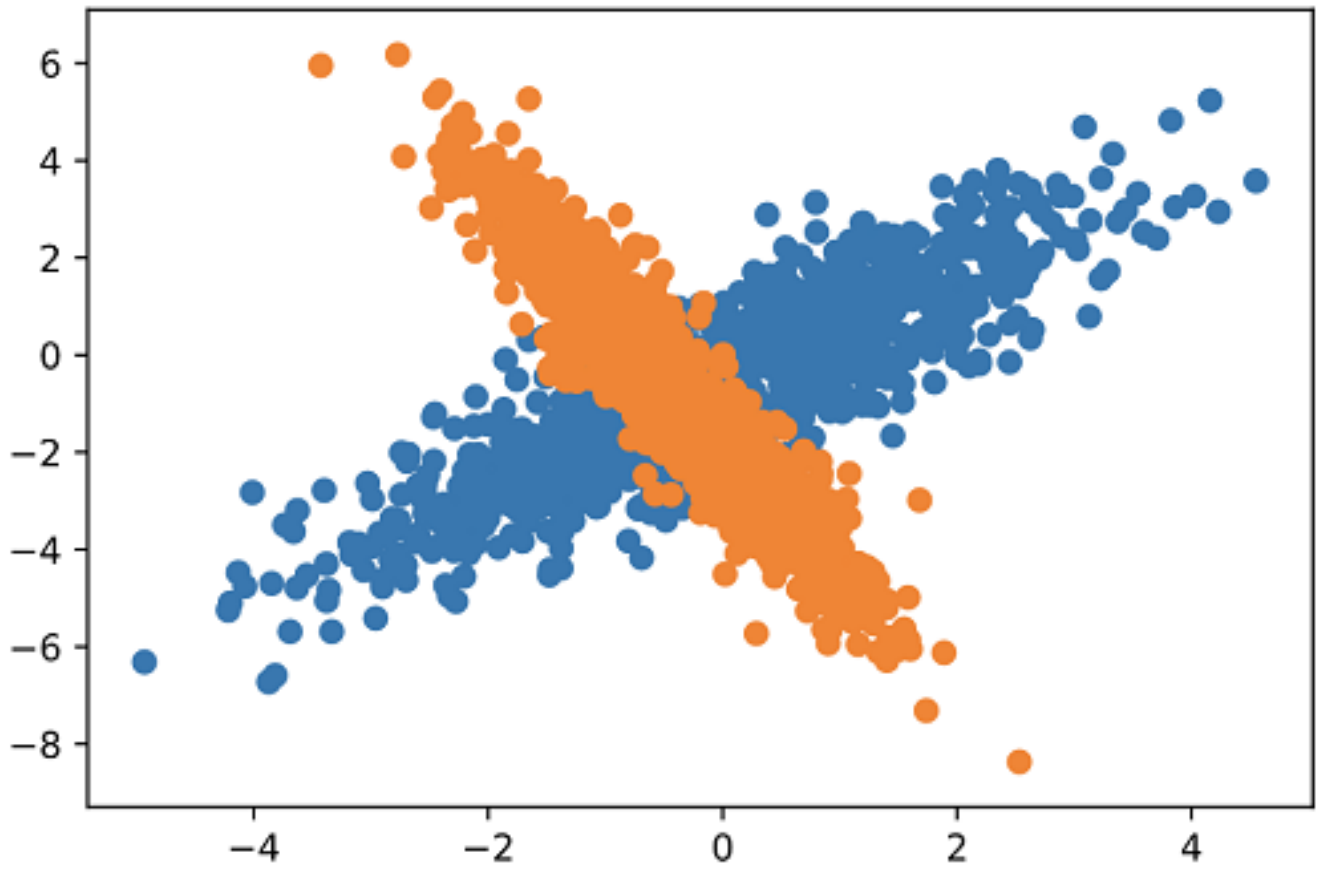
3D Data : Scatter plot matrix

	Dim 1	Dim 2	Dim 3
Nominal	✗	✗	✗
Ordinal	✗	✗	✗
Interval	✓	✓	✓
Ratio	✓	✓	✓



3D Data : Color scatter plot

	Dim 1	Dim 2	Dim 3
Nominal	✗	✗	✓
Ordinal	✗	✗	✓
Interval	✓	✓	✗
Ratio	✓	✓	✗



Visualization Tools

Data Visualization Tools

- ✓ Tableau
- ✓ Infogram
- ✓ ChartBlocks
- ✓ D3.js
- ✓ Google Charts
- ✓ Fusion Charts
- ✓ Chart.js

Visualization using Programming

- ✓ Python
  - matplotlib
  - seaborn
  - plotly
  - pylab
- ✓ R
  - graphics
  - ggplot2

Ministry of Higher Education and Scientific Research
Djilali BOUNAAMA University - Khemis Miliana(UDBKM)
Faculty of Science and Technology
Department of Mathematics and Computer Science

Chapter 5

Data Communication & Visualization

AIBD-M1-UEM112 : Introduction to Data Science

Noureddine AZZOUZA

n.azzouza@univ-dbkm.dz