Data Visualization

Date: 29-01-2021

By: Shraddha P Jain

Roll No: 20-PBD-002

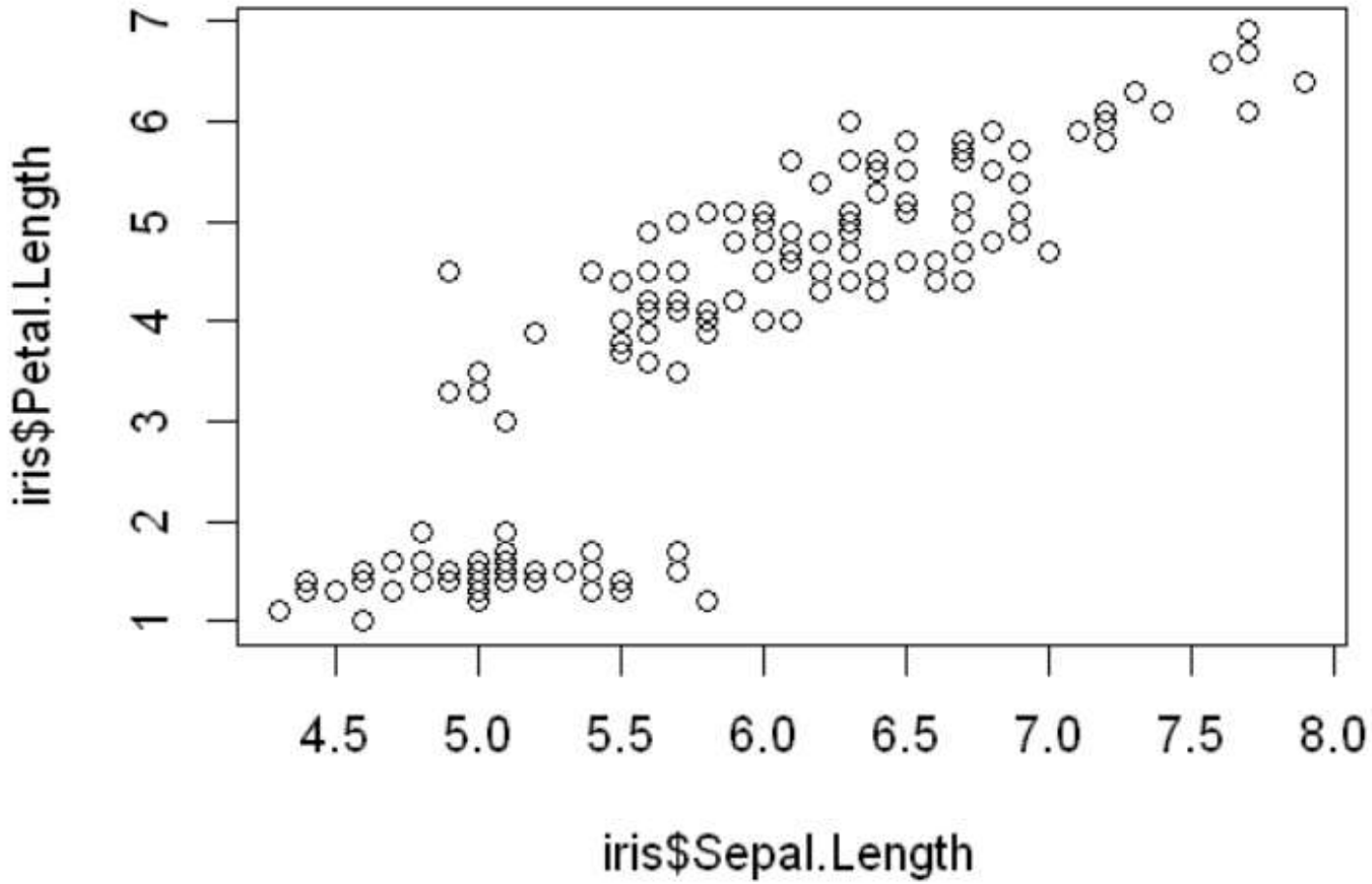
This is my presentation for the subject 1801:Statistical Methods
For CIA-2

What is data visualization?

- Data Visualization is Representation of data graphically.
- It is turning raw data into insights that can be easily interpreted by the readers.
- It is an art through which information, numbers, and measurements can be made more understandable.

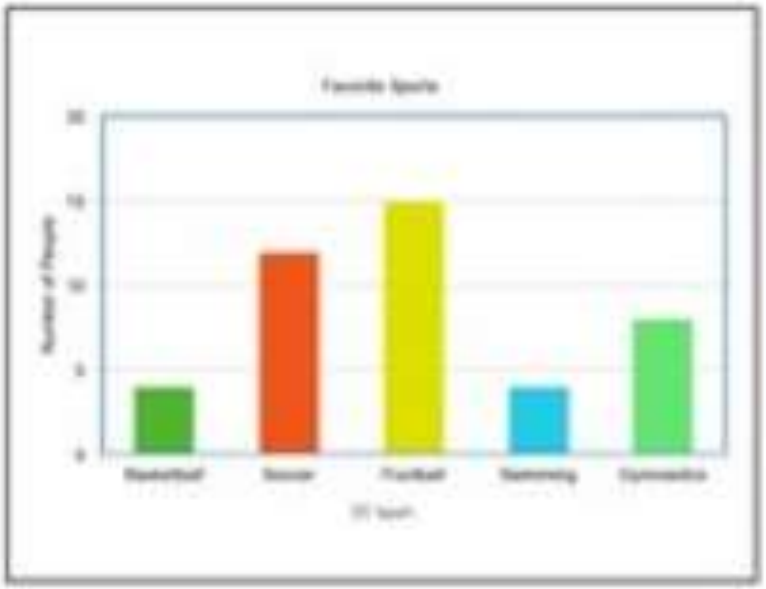
Goal of Data Visualization

- Data Visualization is used for throughout data analysis, mainly in exploratory data analysis and to communicate our results.
- The purpose of data visualization at different stages of data analysis are different
  - In exploratory data analysis, graphs are used to look for insights of data, like the distribution of the data, relationship among variables, etc., which guide us on how to proceed further with our analysis


iris$Petal.Length
1
2
3
4
5
6
7
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
iris$Sepal.Length

Goal of Data Visualization (Contd.)

- Often times, we need to convey our inferences from the data to people who have little or no knowledge about all the statistical terms and techniques we need for our analysis. Visual techniques help us communicate our findings in layman's terms, that can be understood by everyone.



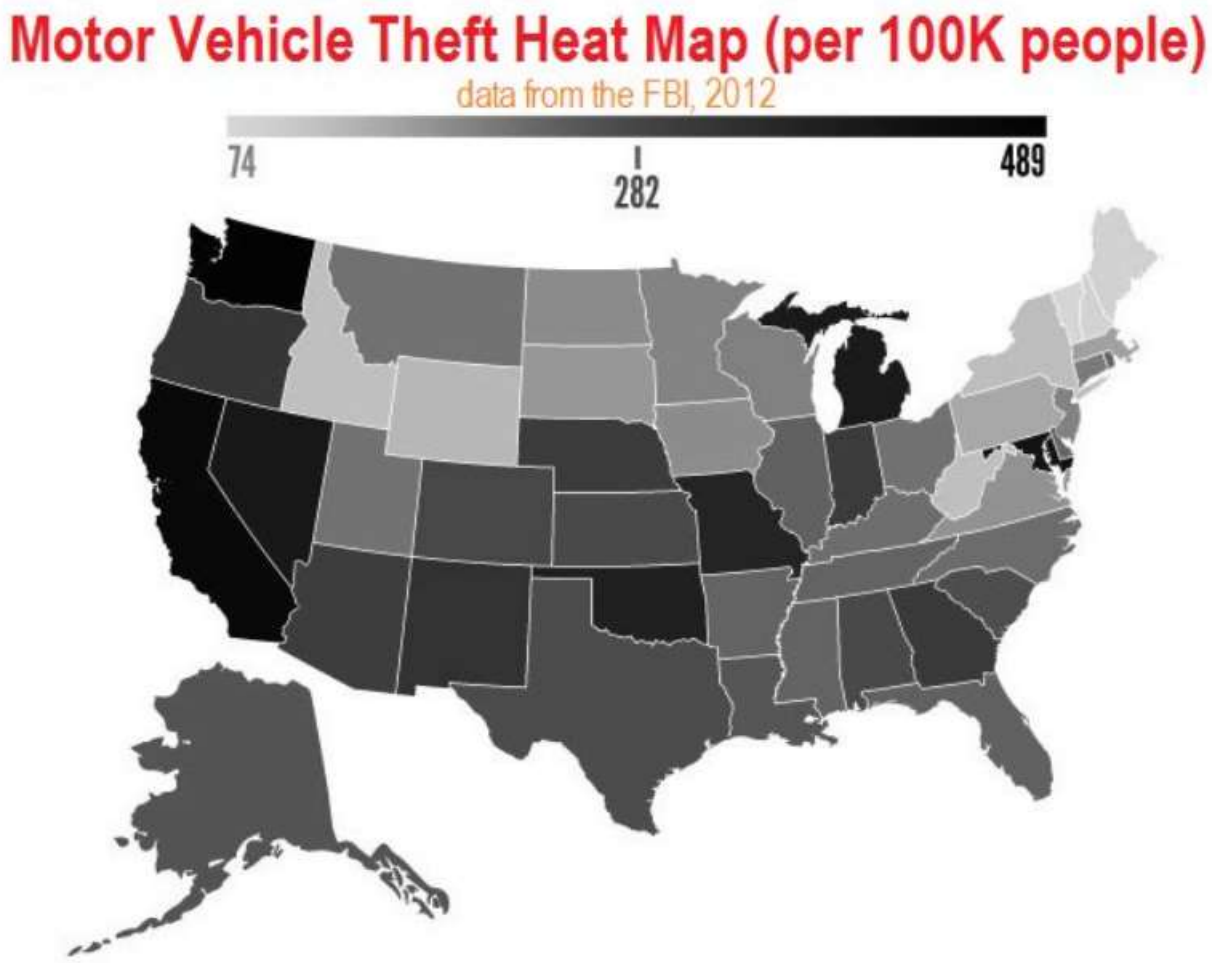
Favorite Sports	
Swimming	4
Gymnastics	8
Football	15
Basketball	4
Soccer	12

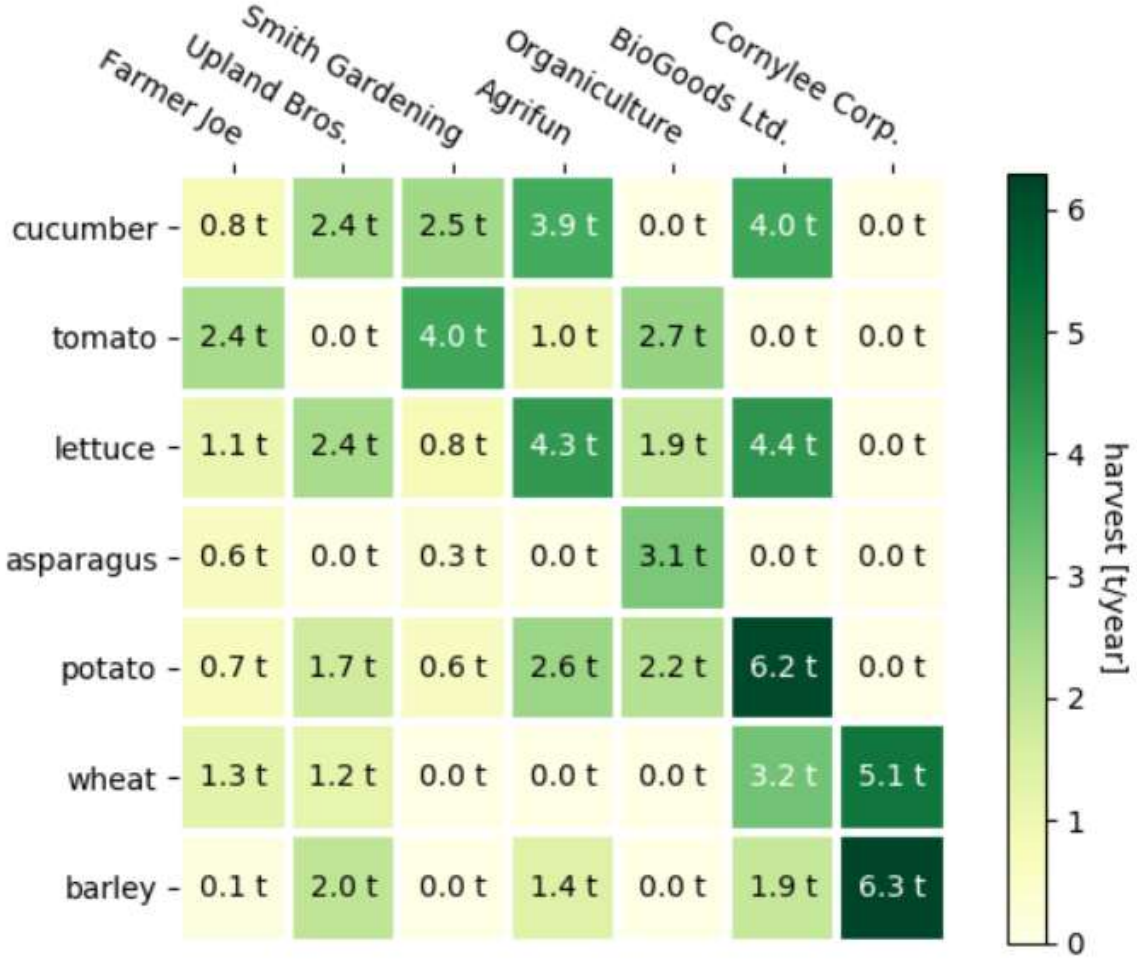
Types of visualizations

- We have already studied many of the visualizations in the class, which include
  - Box plots
  - Pie plots
  - Scatter plots
  - Line charts
  - Ogives
  - Histograms
  - Pareto graphs

- We will not go in details about this, but let us see some other types of interesting and useful types of graphs.

Heat map


Motor Vehicle Theft Heat Map (per 100K people)
data from the FBI, 2012
74
282
489


Farmer Joe
Upland Bros.
Smith Gardening
Agrifun
Organiculture
BioGoods Ltd.
Cornylee Corp.
cucumber
tomato
lettuce
asparagus
potato
wheat
barley
0.8 t
2.4 t
2.5 t
3.9 t
0.0 t
4.0 t
0.0 t
2.4 t
0.0 t
4.0 t
1.0 t
2.7 t
0.0 t
0.0 t
1.1 t
2.4 t
0.8 t
4.3 t
1.9 t
4.4 t
0.0 t
0.6 t
0.0 t
0.3 t
0.0 t
3.1 t
0.0 t
0.0 t
0.7 t
1.7 t
0.6 t
2.6 t
2.2 t
6.2 t
0.0 t
1.3 t
1.2 t
0.0 t
0.0 t
0.0 t
3.2 t
5.1 t
0.1 t
2.0 t
0.0 t
1.4 t
0.0 t
1.9 t
6.3 t
harvest [t/year]
0
1
2
3
4
5
6

Word Cloud

State of the Union Address, 2002 vs. 2011

President Bush, January 29, 2002

President Obama, January 25, 2011

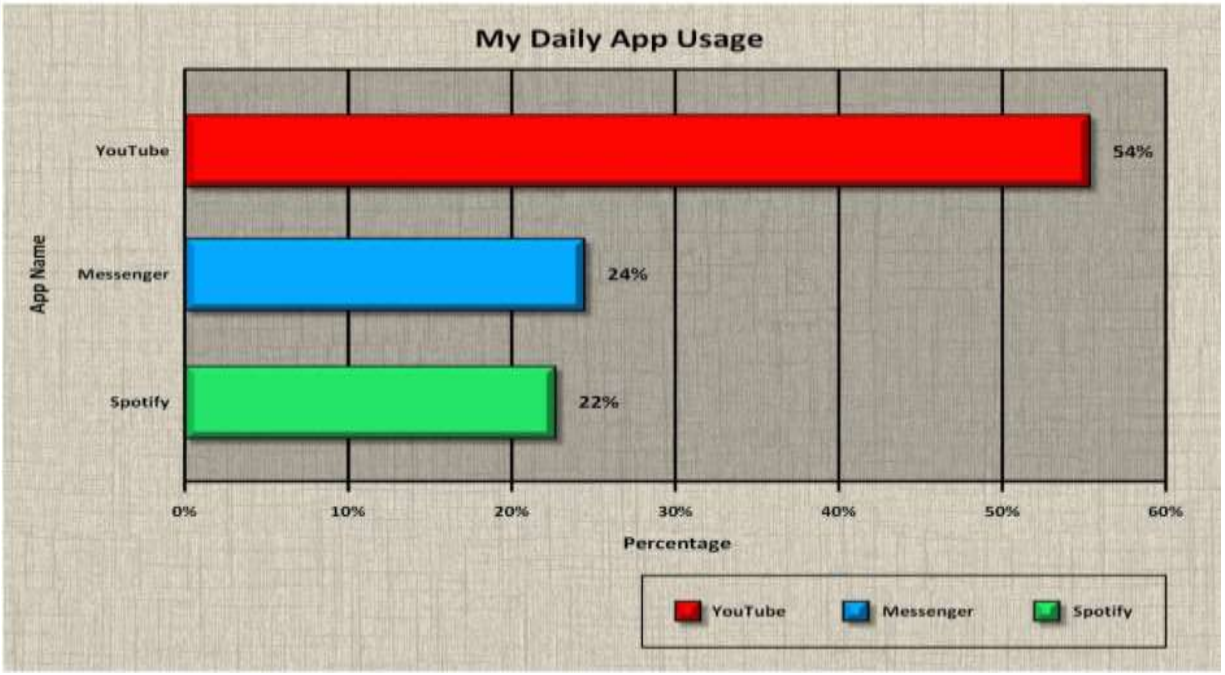
Elements of visualization

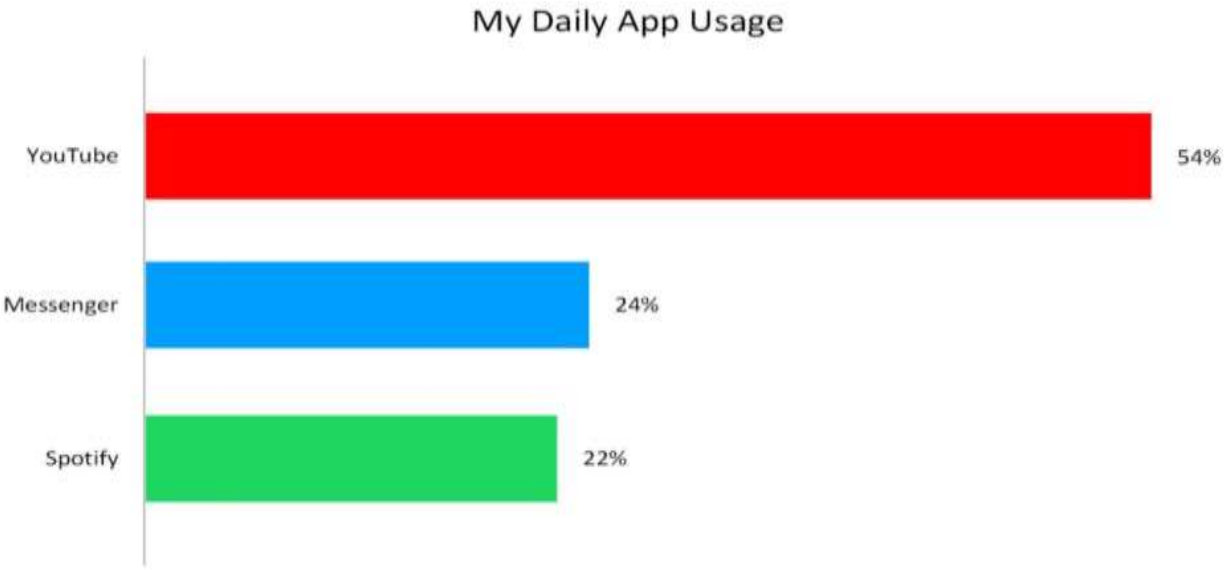
- In his book ‘The functional Art’, Albert Cairo gives six pairs of elements of visualization that are opposite to one another, with the first characteristic falling in the complex and deeper region and the second characteristic falling in the more intelligible and shallower region:

- **Abstraction – Figuration**
  - *Boxes and charts (abstraction) or real-world physical objects (figuration)*
- **Functionality – Decoration**
  - *No embellishments (functionality) or artistic embellishments (decoration)*
- **Density – Lightness**
  - *Must be studied in depth (density) or understandable at a glance (lightness)*
- **Multidimensional – Unidimensional**
  - *Different aspects of phenomena (multidimensional) or single or few items of phenomenal (unidimensional)*
- **Originality – Familiarity**
  - *Novel methods of visualization (originality) or established and well understood methods of visualization (familiarity)*
- **Novelty – Redundancy**
  - *Explaining each item once (novelty) or encoding multiple explanations of the same phenomena (redundancy)*

Graphical Heuristics

- In his book, Visual Display Of Quantitative Information, Edward Tufte introduces two heuritics:
  - **Data Ink Ratio** : The **data-ink ratio** is the proportion of Ink that is used to present actual data compared to the total amount of ink (or pixels) used in the entire display.
  - **Chart junk**: he suggests that artistic decorations on statistical graphs are like weeds in our data graphics.


My Daily App Usage
YouTube
Messenger
Spotify
54%
24%
22%
App Name
0%
10%
20%
30%
40%
50%
60%
Percentage
YouTube
Messenger
Spotify


My Daily App Usage
YouTube
Messenger
Spotify
54%
24%
22%


$41 BILLION IN SALES
BILLIONS OF DOLLARS
GDP OF AFGHANISTAN
$21 BILLION
$4.1 BILLION IN SALES
$4.3 BILLION IN SALES
$8 BILLION IN SALES
$8.2 BILLION IN SALES
$9.4 BILLION IN SALES
$11.3 BILLION IN SALES
40
35
30
25
20
15
10
5
Pizza Hut
KFC
Wendy's
BURGER KING

Is chart junk really bad?

- There is research going on whether chart junk helps in recall of information,
- an experiment that compared embellished charts with plain ones was performed, and measured both interpretation accuracy and long-term recall. It was found that people's accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better.
- The study was :

Bateman, S., Mandryk, R. L., Gutwin, C., Genest, A., McDine, D., & Brooks, C. (2010, April). Useful Junk?: The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2573-2582). ACM.

Thank You