The background of the slide is a blurred image of a laptop screen. On the screen, there is a line chart with a blue line showing an upward trend, and a pie chart with a large blue section and a smaller green section. The text is overlaid on this background.

Data Analytics and Visualization

with Applications in Financial Domain

December 14, 2021

FDP - XIM University, Bhubaneswar

Sohom Ghosh

About Me

“

I am a Data Scientist at the Artificial Intelligence, Centre of Excellence of Fidelity Investments. My research interests include Applications of Natural Language Processing (NLP) in FinTech.

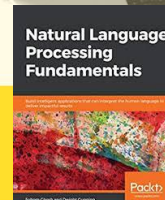
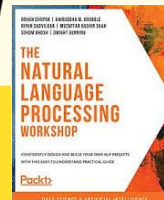
Over the last few years, I have been working on improving digital experience & financial well-being of millions of users across different industries like Financial Services, Internet and so on. In addition to being co-author of the books NLP Fundamentals and The NLP Workshop, I have several publications in proceedings of international conferences and refereed journals.

I hold a Master's Degree in Software Systems (with specialization in Data Analytics) from BITS Pilani, India and a Bachelor's Degree in Computer Science & Engineering.

Outside work, I like to play harmonica & cajon. Being an adventure lover and a fitness buff, I believe that “Health is Wealth”.

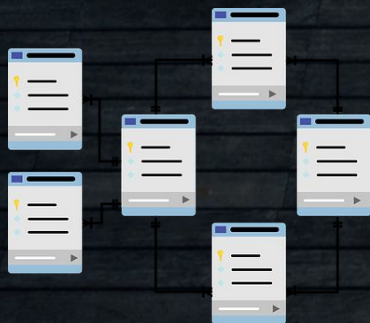
SOHOM GHOSH

sohomghosh.github.io



**SOHOM
GHOSH**

Data→Information→Insights



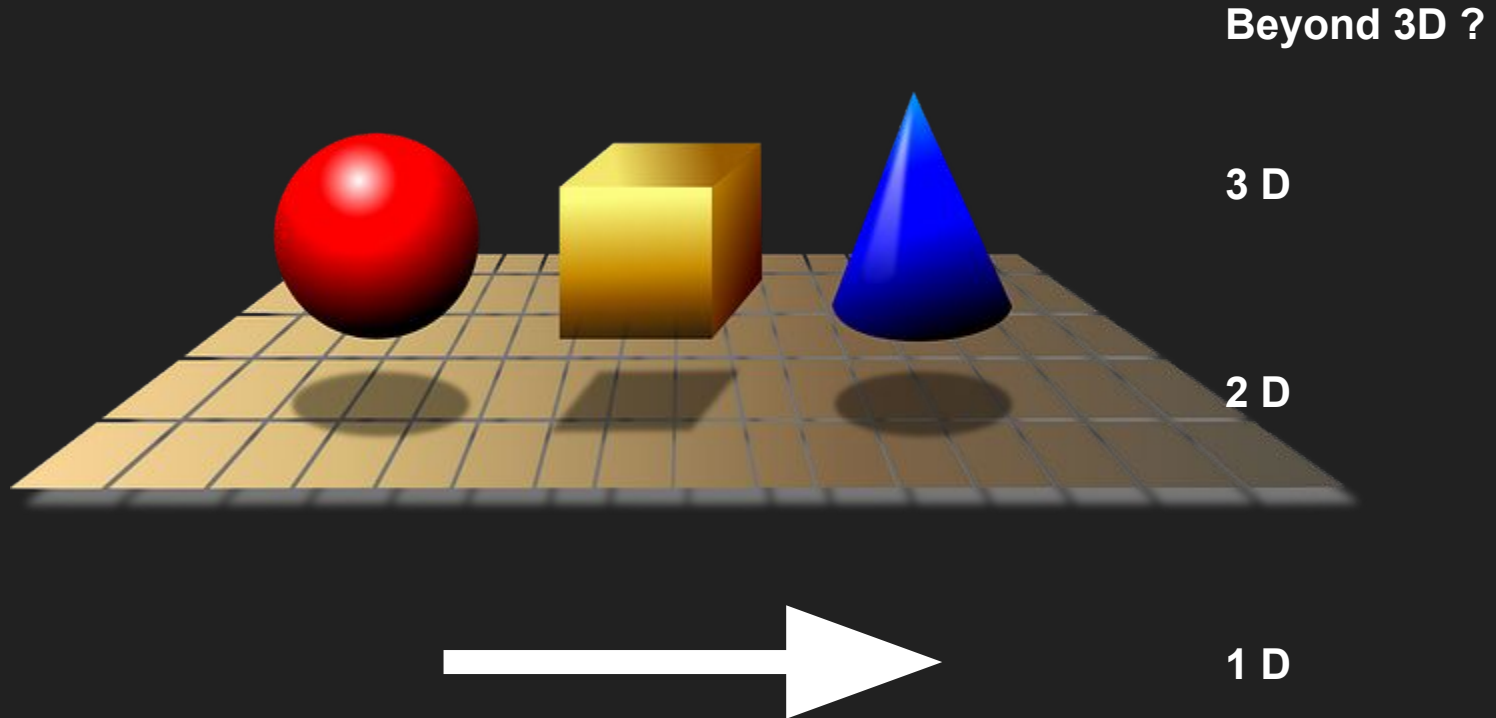
Structured (e.g. csv, excel), Semi-structured (e.g HTML, XML), Unstructured

Why should we visualize data?

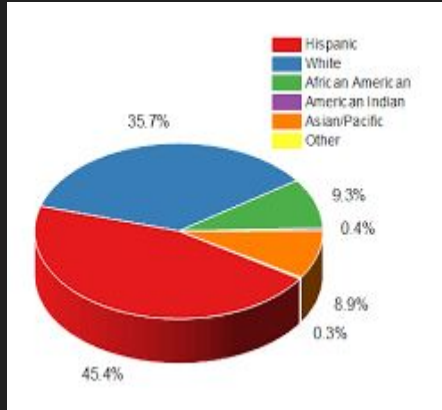


"A picture is worth a thousand words"

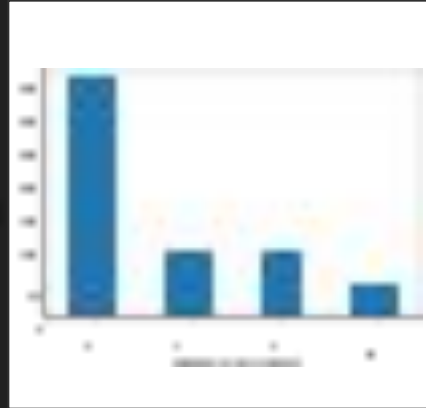
What we can/(not) visualize?



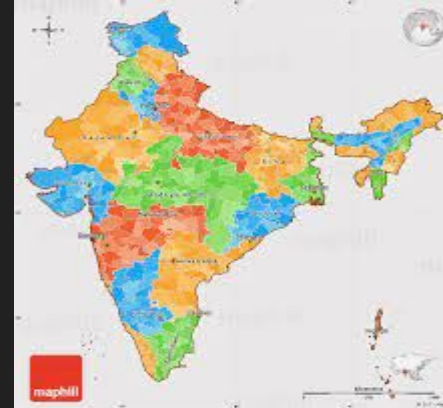
Common mistakes in Data visualization



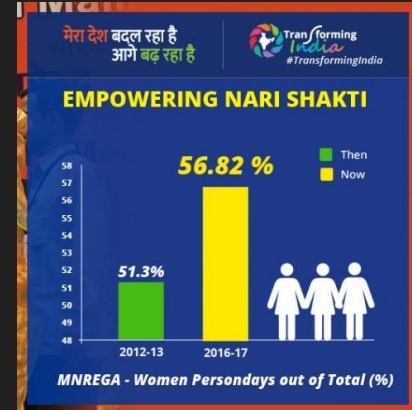
1



2



3



4

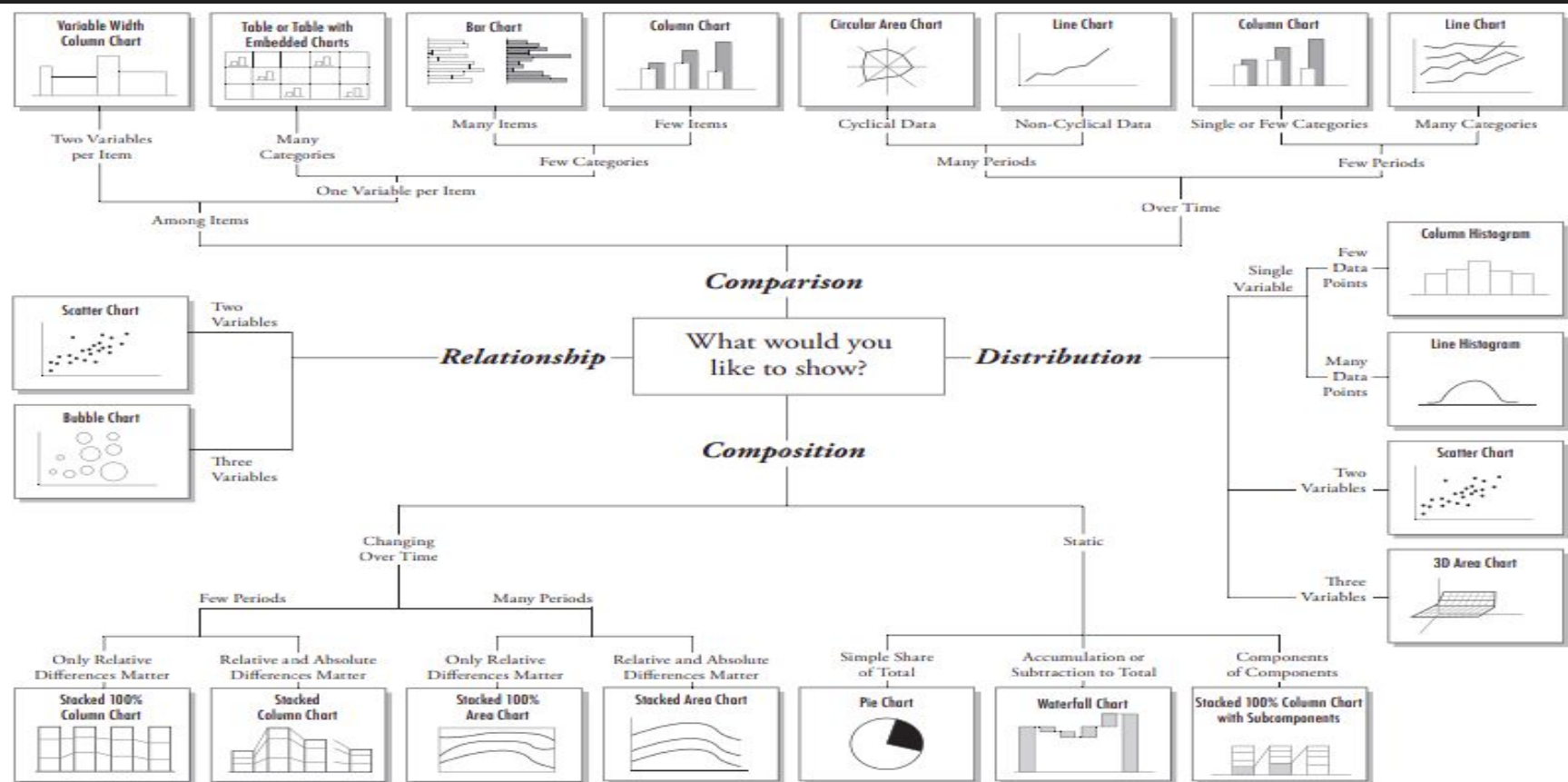
Other points to remember:

- Don't forget to label the axes
- Avoid putting too much information in one go
- For publishing in academic journals try to use shape(s) instead of colour(s).

Different kinds of visualization



How to decide which visualization would be appropriate?



Case study 1: Visualizing Loan Application Data



Case study 2: Visualizing Financial Tweets

- Word Cloud
- Clusters using TSNE
- Named Entities
- Dependency trees



Thank you

Any questions?

All the training materials are available at:

https://github.com/sohomghosh/Data_Visualization_FDP

Reach me at:

Email: sohom1ghosh@gmail.com

Website: <https://sohomghosh.github.io/>

LinkedIn: <https://www.linkedin.com/in/sohomghosh>



Sohom Ghosh

Data Scientist at AI, CoE - Fidelity

