

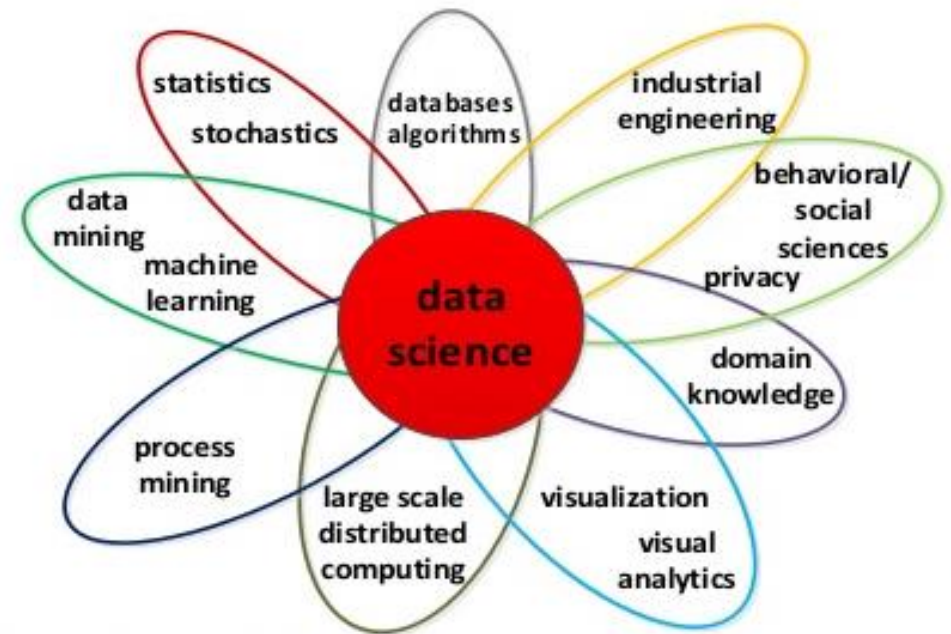
DATA SCIENCE 101



Intro to Data Science

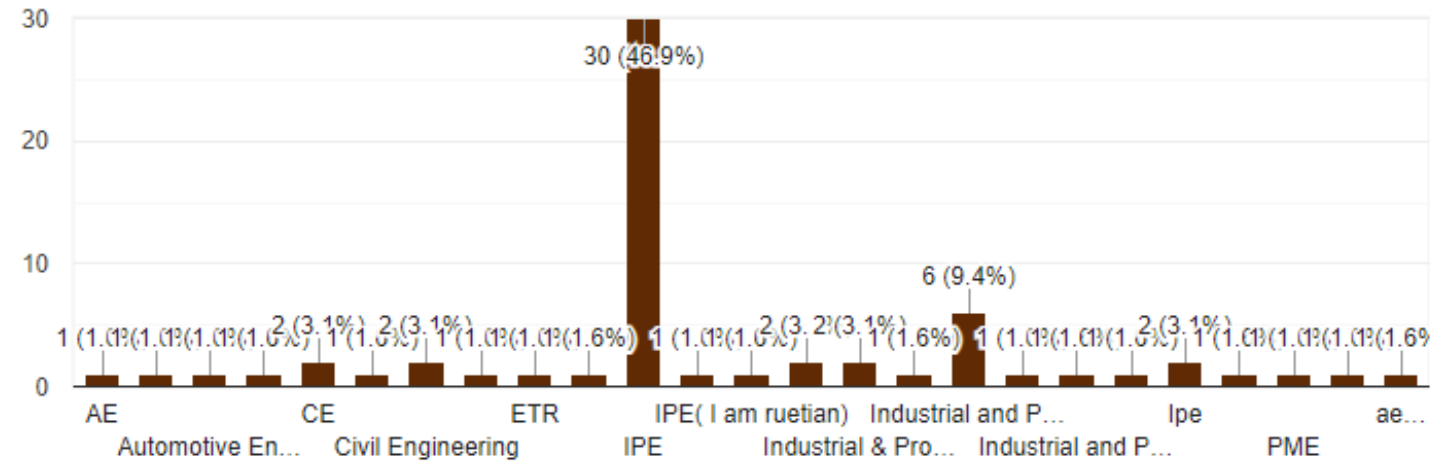
The ability to take data - to be able to understand it, to process it, to extract value from it, to visualize it, to **communicate** it.

- Hal Varian, chief economist at Google



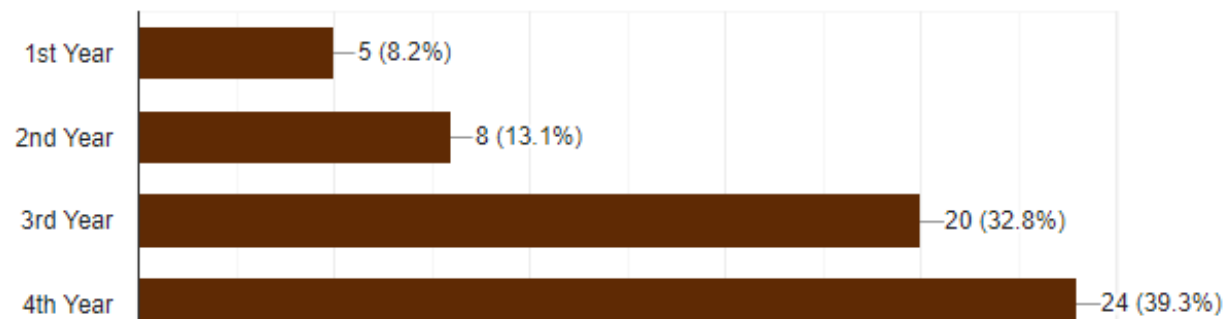
Department

64 responses



Year

61 responses





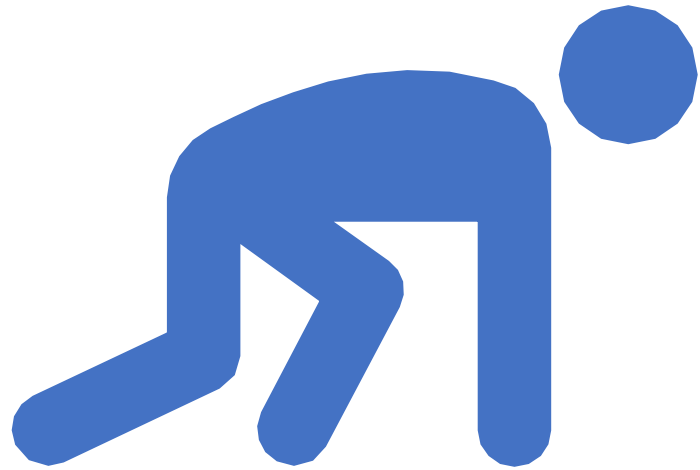
Trainer's bio:

Tanmoy Das

- Ph.D. Researcher in Training, Industrial Engineering, Dalhousie University, Canada.
- Masters in Industrial Engineering, Florida State University, US
- Former Asst professor, MIST
- Experience in Industrial Engineering with a focus in Data Science.
- www.linkedin.com/in/tanmoyie



Showing some issues (e.g.
certification) from LinkedIn



*Expert in any
field was once
a beginner*

Data





Q&A

Question from audience

- Data Science in Industrial Engineering undergrad program?
 - Show from syllabus of MIST

- https://www.linkedin.com/posts/tanmoyie_department-of-industrial-and-production-engineering-activity-6709086444037689344-nvcN

Sources to start learn **Data Science/ Python**

- [Data Science A-Z from UDEMY](#)
- [Youtube playlist of Simplilearn](#)
- [Learn Python with Socratica](#)
- [Active learning](#)
- Which is the best IDE
(Anaconda – spyder, Pycharm)
for Python.
- Upwork, people per hour

Q&A

- How to improve data science skills
- Scope For abroad masters and PhD
 - Don't bother about the department offering related course.
- Want to learn the fundamentals of data science
 - <https://www.kaggle.com/tanmoyie/how-to-learn-data-science-use-kaggle>
 - <https://www.linkedin.com/pulse/how-learn-data-science-non-cs-major-tanmoy-das/>

RAIN



0, 1, 5, 1, 0,
0, 0, 0, 2, 0,
0, 3, 0, 1

Rainfall

- What'd be the amount of rainfall tomorrow?
- What's the mean rainfall?
- Will it rain tomorrow?



**Shouldn't we
just delete
outliers?**



[www.linkedin.com/pulse/
how-learn-data-science-
non-cs-major-tanmoy-das](https://www.linkedin.com/pulse/how-learn-data-science-non-cs-major-tanmoy-das)

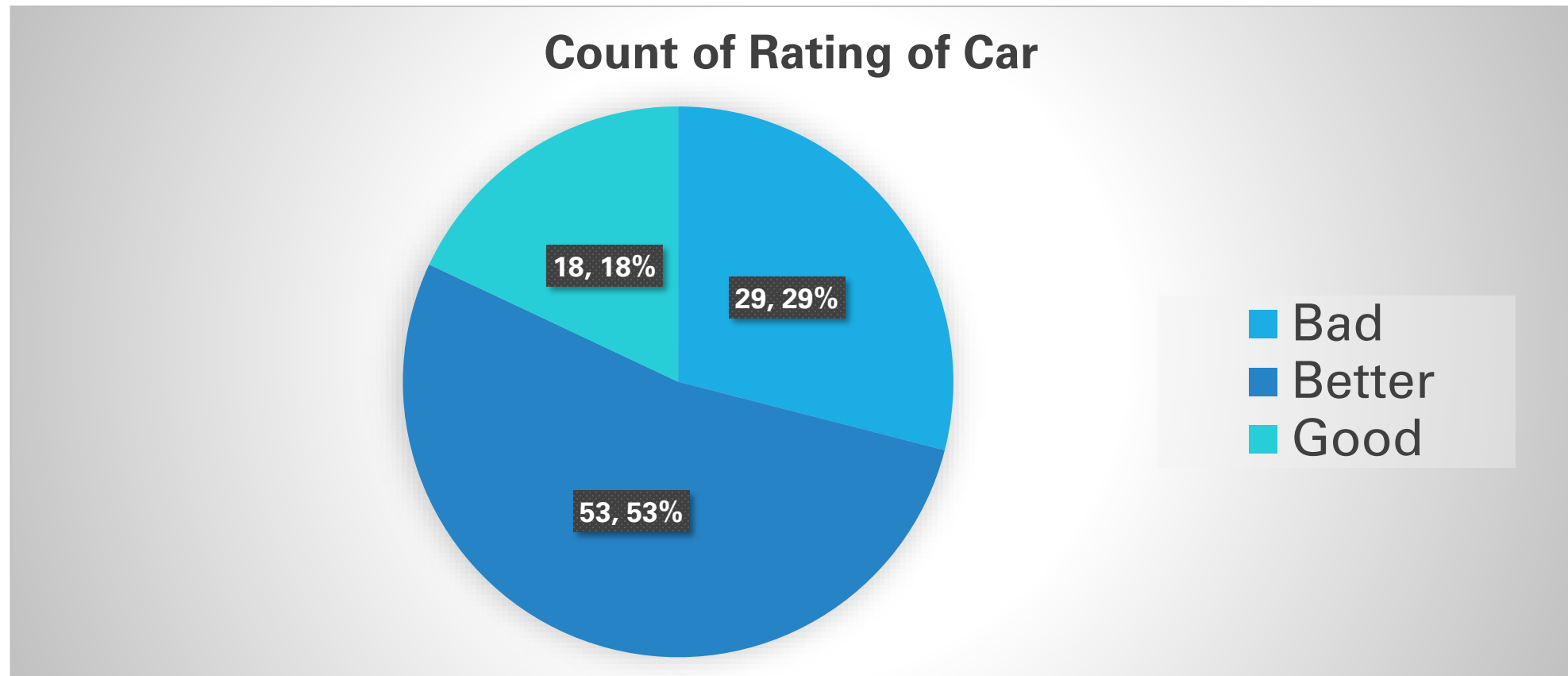
PRICE OF CAR



Dataset

Record no.	Color of Car	Rating of Car	Price
1	Red	Good	17000
2	Red	Better	24000
3	Red	Bad	10000
4	Red	Good	20000
5	Red	Better	28000
6	Red	Bad	6000
7	Red	Good	19000
8	Red	Better	25000
9	Red	Bad	11000
10	Red	Good	19000
...
100	Blue	Good	18000

Pie Chart



HomeTask

Why **PIE CHART** is suggested from **RECOMMENDED CHART**
But

PIE CHART is NOT suggested when I click **PIE CHART**
manually
in EXCEL
?

I am showing you what I mean!

When we compare quantitative
with qualitative

Box plot





www.facebook.com/DataIE

Q&A

- Learn about how to improve data science skills
 - Start with small projects
 - Finish some online courses
 - Start random projects even in Upwork

Q&A

- the process to avoid the overfitting a model.
- What are the mistake to avoid?
 - Certification?
- Some good project to start with
- data science ki aitai sobai jante chaise sir?



Section 1: Data Exploration

Scatterplot, Heatmap, storytelling,
dashboard

Data Visualization

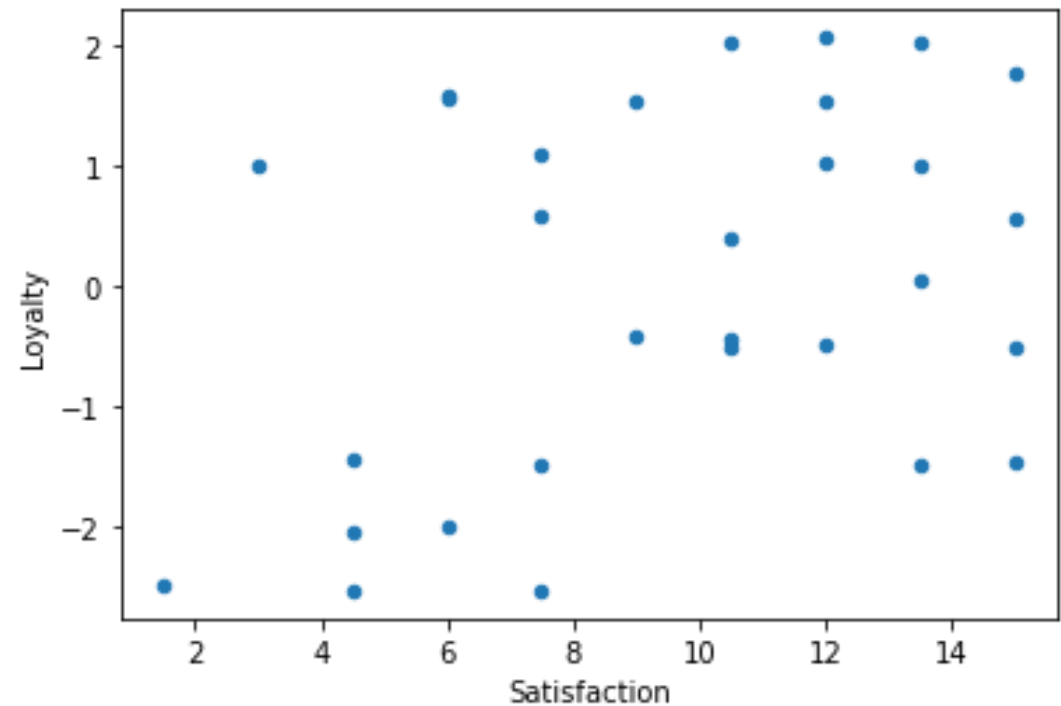
- Graphical representation of information and **data**.
- By using visual elements like charts, graphs, and maps, **data visualization** tools provide an accessible way to see and understand trends, outliers, and patterns in **data**



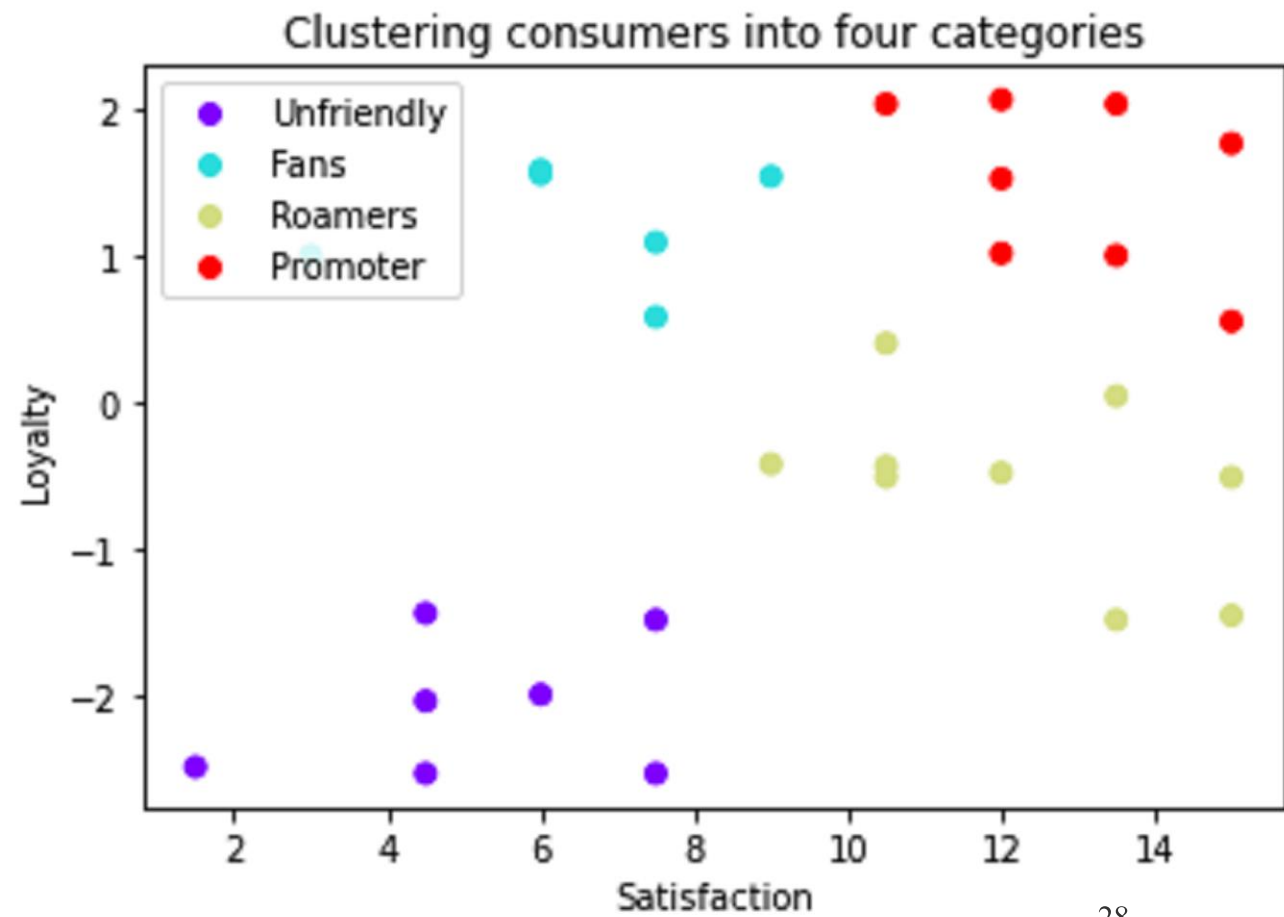
Scatterplot: Satisfaction vs loyalty

- Audience will write the Python program from scratch
- Let's draw this Scatter plot

Satisfaction	Loyalty
6	-1.995
9	-0.42
7.5	-1.485
10.5	-0.435
6	1.59
...	...



www.kaggle.com/tanmoyie/consumer-cluster



SOME SUGGESTIONS





**DO I NEED
PROGRAMMING?**



Python or R?

Online presence of Tanmoy Das

www.tanmoyie.com

www.youtube.com/c/Tanmoyle

<https://github.com/tanmoyie>

www.linkedin.com/in/tanmoyie

www.kaggle.com/tanmoyie

www.facebook.com/tanmoyie

www.facebook.com/dataie

