

DATA SCIENCE INTERNSHIP

- at Unified Mentor

BHUVANESH M

B.E. Computer Science Engineering

bhuvaneshm.in | linkedin.com/in/bhuvaneshm-developer

bhuvaneshm.developer@gmail.com | github.com/bhuvanesh-m-dev

Internship Duration : June–July 2025

Mode : Remote

This Internship project is hosted on GitHub at github.com/bhuvanesh-m-dev/ds-intern-unified-mentor

My First Step into Data Science Begins Here – Grateful to Unified Mentor for the Opportunity.

Table of Contents

- Introduction
- Technologies & Core Skills
- Netflix Data Analysis
- Customer Satisfaction Prediction
- Boxplot with IQR
- Standard Deviation
- Folium Map
- Social Media Pie Chart
- Wine Data Analysis
- Conclusion

Introduction

Internship Highlights:

- Real-world projects solving data challenges
- Hands-on use of tools: Python, Pandas, NumPy, Matplotlib, etc.
- End-to-end ownership: data cleaning, analysis, visualization.
- Exposure to Git, GitHub version control and report writing.
- Learned professional documentation.

Technologies & Core Skills

Category	Tools & Skills
Programming	Python
Data Handling	Pandas, Numpy
Visualization	Matplotlib, Seaborn
IDE/Environment	Jupyter, Google Colab, VS Code

Additional Skills:

- Exploratory Data Analysis (EDA)
- Data Cleaning & Preprocessing
- Version Control with Git
- Report Writing & Visualization Storytelling

Netflix Data Analysis

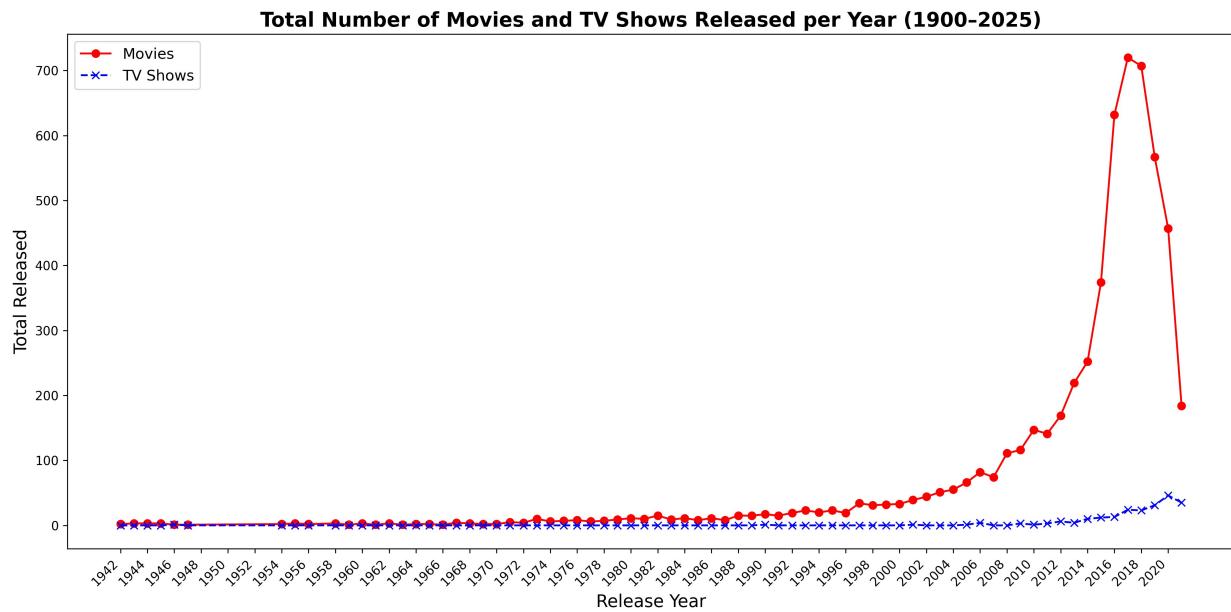
Objective : Analyze content trends in Netflix OTT library.

Steps :

- Cleaned netflix.csv : removed NaNs, standardized data
- Visualized content growth by year (movies & TV shows)
- Merged plots for comparative analysis

Output : Trend lines showing content growth.

Repository URL : [Netflix](#)



Customer Satisfaction Prediction

Objective : Predict satisfaction using support ticket data

Dataset : customer_support_tickets.csv → Cleaned to data.csv

Steps:

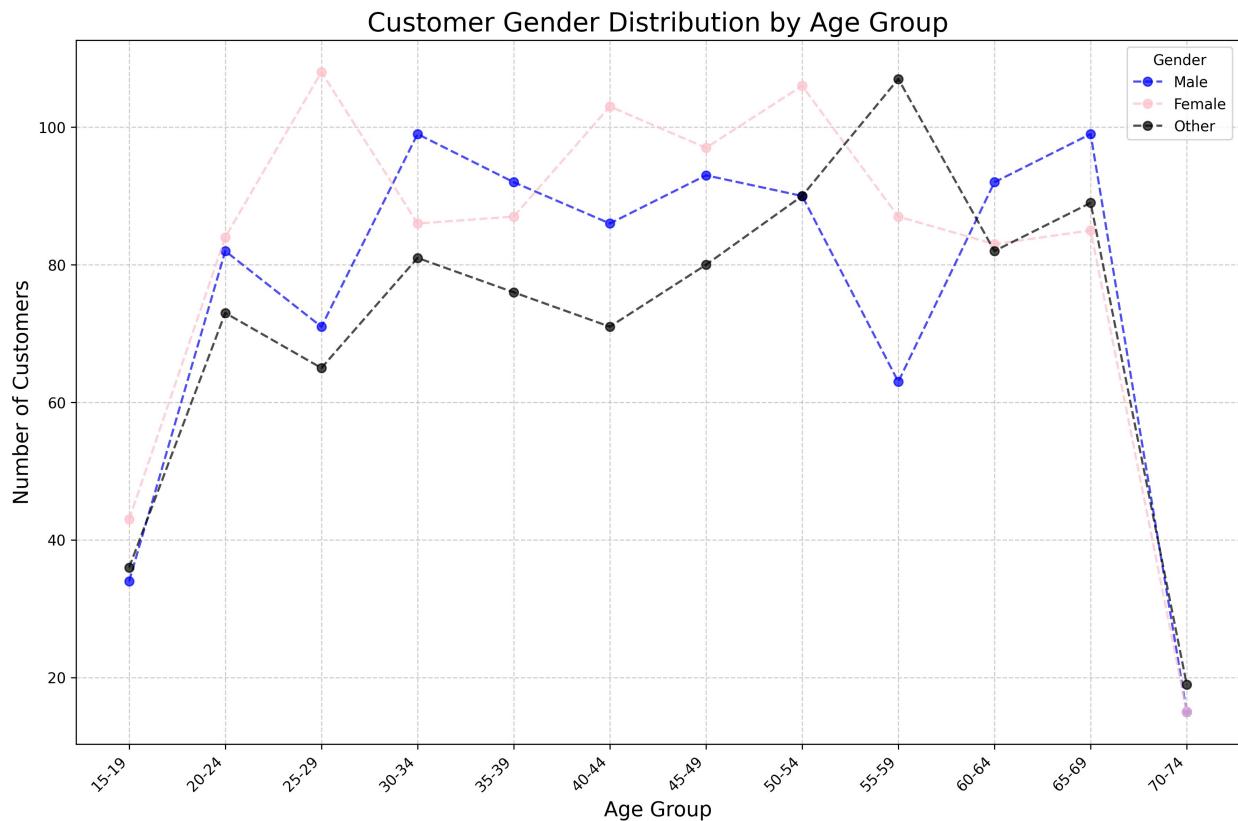
- Cleaned rows with missing values (Resolution, First Response)
- Removed duplicates and saved to new file
- Prepared data for future

Script : data_cleaning.py

Result : Cleaned data → 2,769 rows from 8,469

Learning Outcome : Importance of data hygiene in Data Science

Repository URL : [Customer](#)



Boxplot with IQR

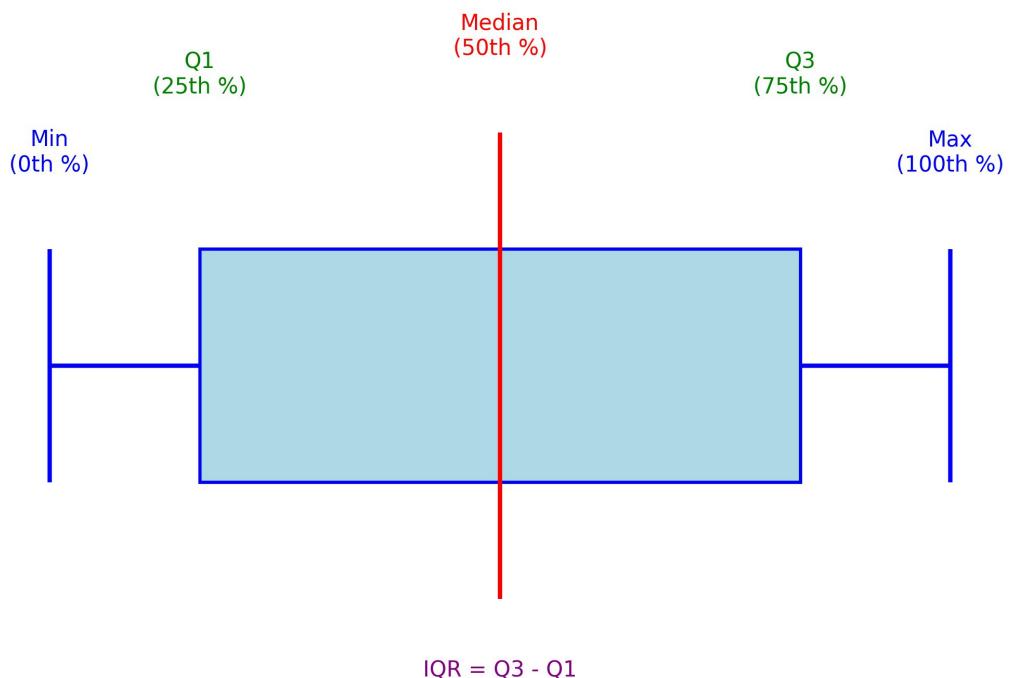
Objective : Visualize data distribution and detect outliers

Approach :

- Used NumPy for stats (Q1, Q2, Q3, IQR)
- Created annotated boxplot using matplotlib
- Explained IQR and its role in outlier detection

Output : Labeled boxplot visual

Boxplot Structure: IQR, Quartiles, Min, Max, Median



Key Takeaway : Boxplot is ideal for understanding data spread and central values

Repository URL : [IQR](#)

Standard Deviation Comparison

Objective : Differentiate between sample and population std. deviation

Workflow :

- Generated 500 values, took random sample of 50
- Stored values using pickle in data.dat
- Calculated population vs sample standard deviation

Scripts : create_data.py, run.py

Concepts Covered : Pickle I/O, variance formulas, real data comparison

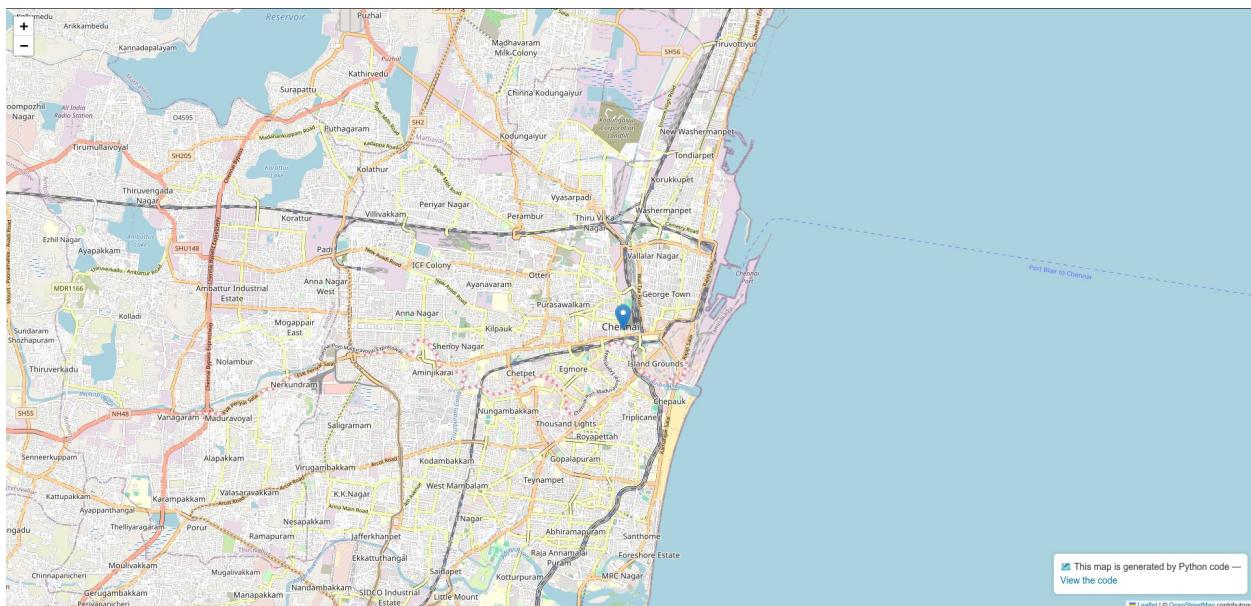
Repository URL : [Deviation](#)

Folium Map Visualization

Objective : Create an interactive map centered on Chennai

Steps:

- Used folium.Map to set coordinates
- Added markers with popups & tooltips
- Saved map as HTML



Output : Web-ready interactive map

Learning : Practical intro to geospatial visualization using Python

Repository URL : [Folium](#)

Social Media Usage Pie Chart

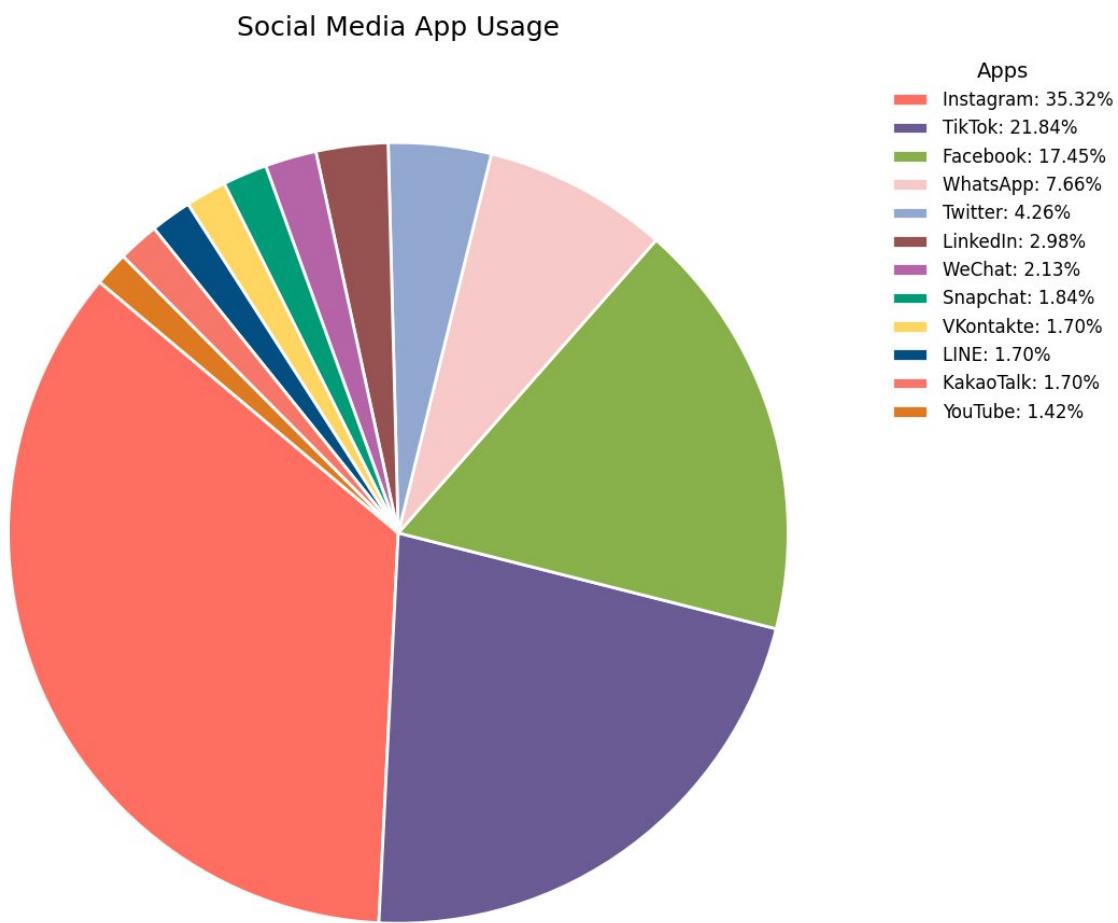
Objective : Visualize social media usage data from a survey

Tools : pandas, matplotlib

Steps :

- Read dataset and calculated platform percentages
- Created colorful pie chart with legend and labels

Script : pie_chart.py



Visualization Outcome : Clear insight into user behavior patterns

Insight : Instagram & TikTok dominate among surveyed users

Repository URL : [Pie Chart](#)

Wine Data Analysis

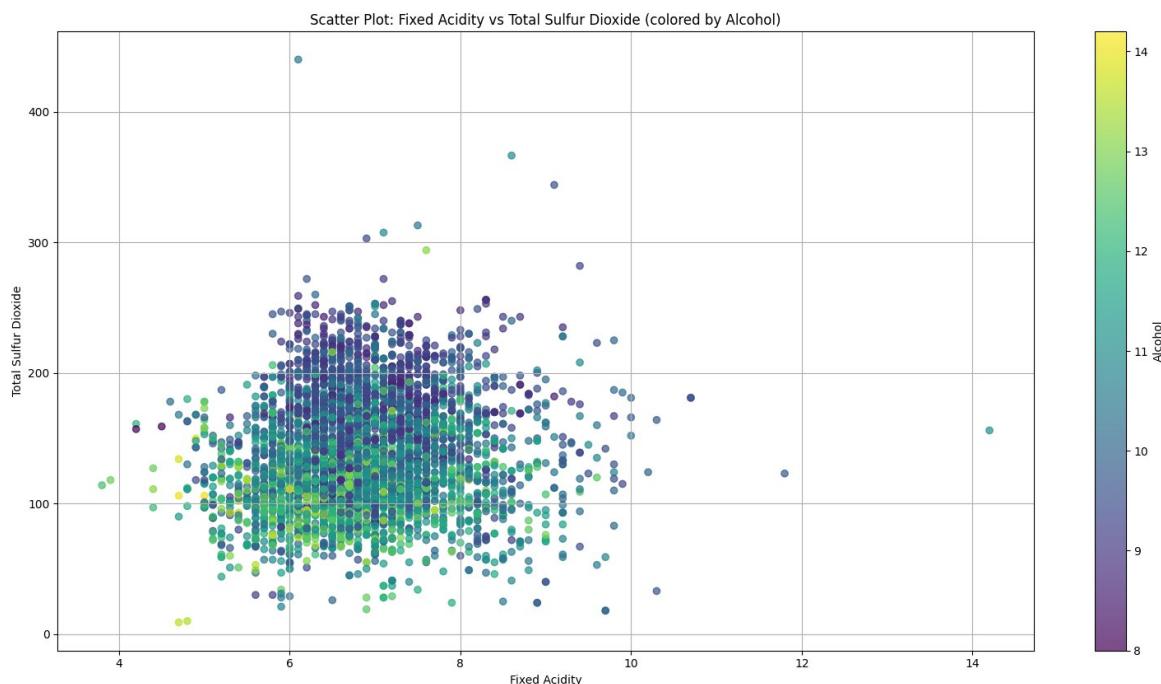
Objective : Explore properties of wine dataset

Key Tasks :

- Donut chart comparing mean values
- Scatter plot visualizing correlations
- Bar plots for statistical metrics

Files : donut.py, scatterplot.py, mean.py, median.py, mode.py

Dataset : Wine Quality Dataset.csv



Learning : Visual analytics aid in understanding feature behavior

Repository URL : [Wine Data Analysis](#)

Conclusion

This internship has provided a solid foundation in Data Science. I've gained practical experience in data wrangling, visualization, statistical analysis, and project documentation. I sincerely thank Unified Mentor for this incredible learning opportunity and look forward to applying these skills in future roles.