

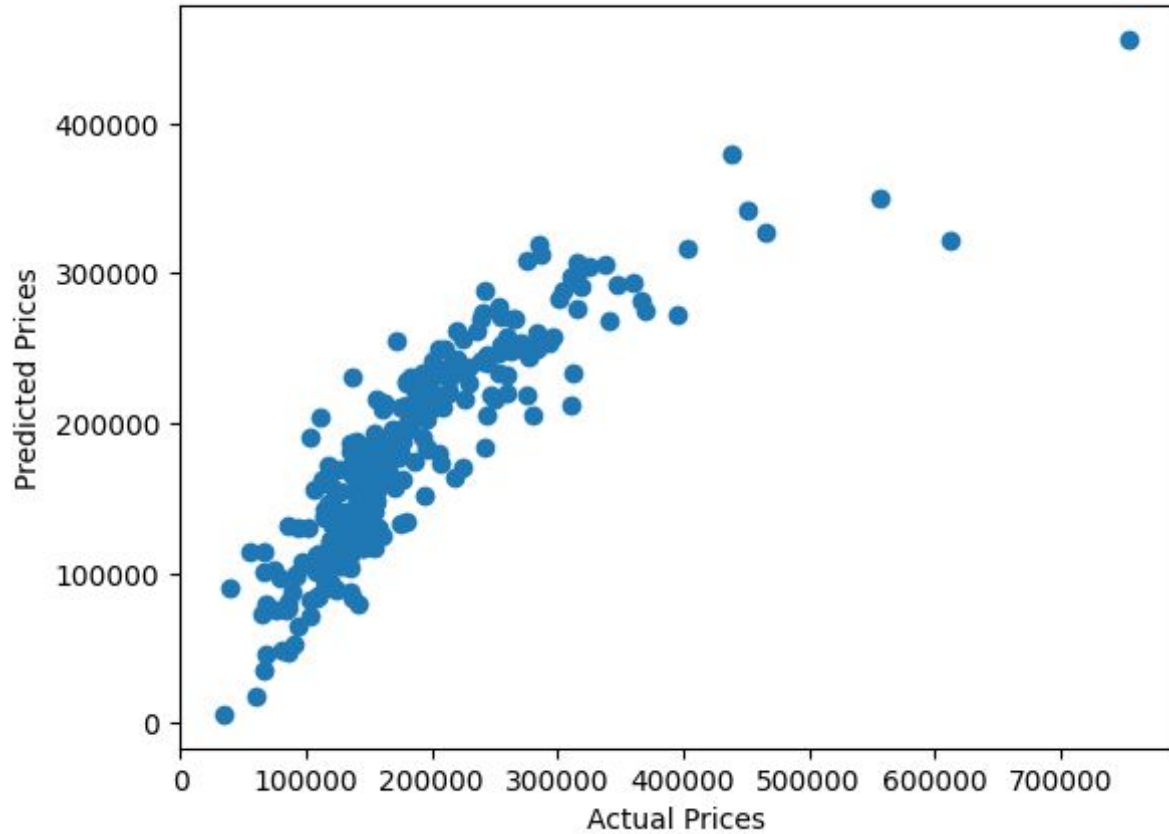
My Data Analysis Project Portfolio

A journey through 7 hands-on data analysis & machine learning projects

Project 1: House Price Prediction

- **Brief Description:** Built a regression model using Linear Regression & Random Forest to predict home prices based on area, year built, and quality.
- **Unique Learning:** Handling missing values, understanding feature importance, and encoding categorical variables.
- **Key Skills:** Regression, Pandas, Matplotlib, Feature Engineering

Actual vs Predicted Prices



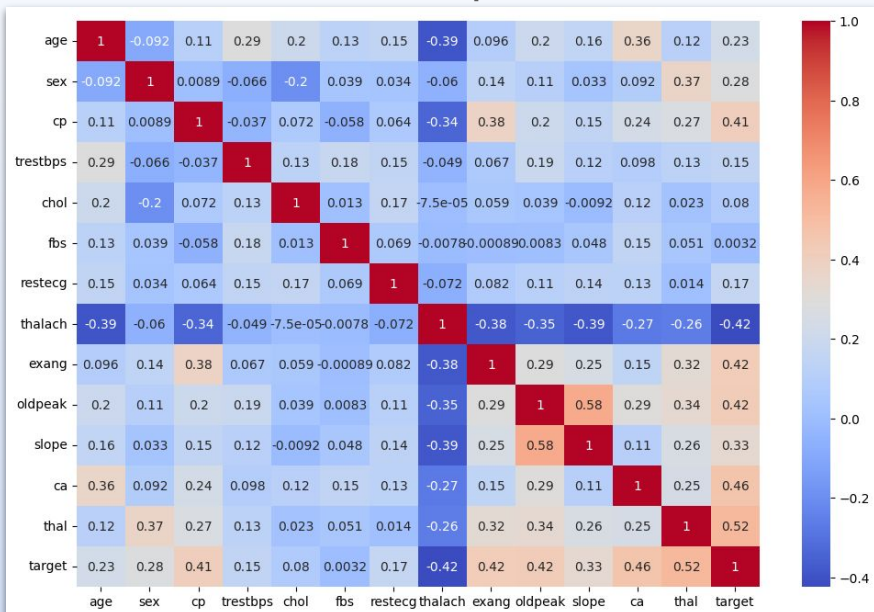
Trained model
with 85%
accuracy.
Visualizing
prediction
patterns.

House Price Predictions

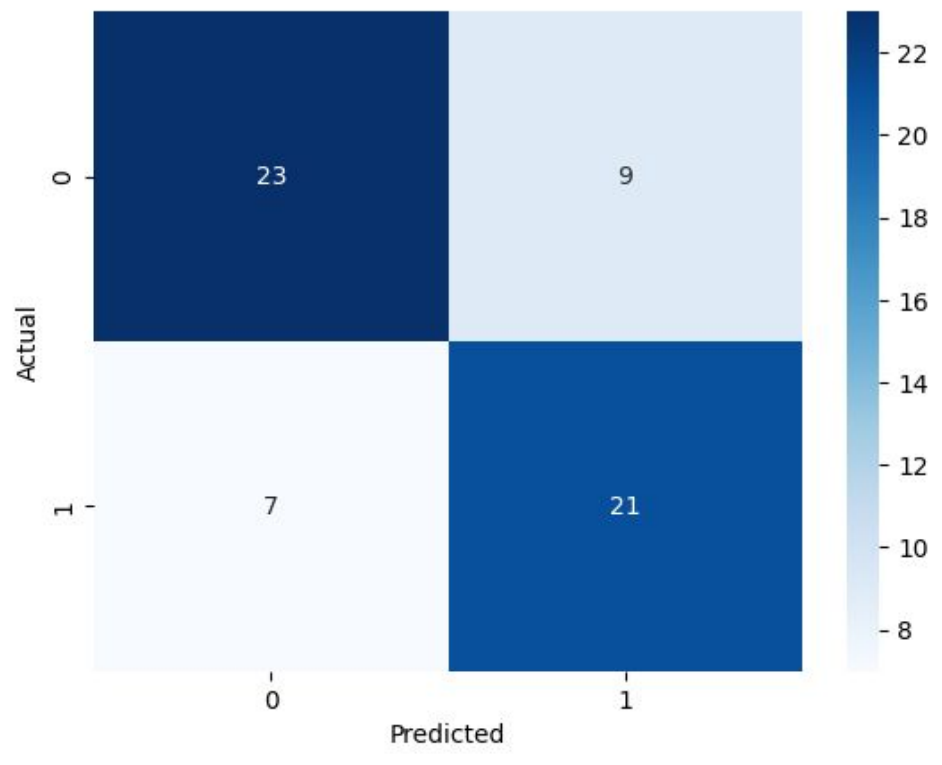
Project 2: Heart Disease Prediction

- **Description:** Classification model predicting likelihood of heart disease using logistic regression and KNN.
- **Unique Learning:** Dealing with imbalanced data and importance of evaluation metrics.
- **Skills:** Logistic Regression, KNN, Confusion Matrix, Scikit-learn

HeatMap



Confusion Matrix



Project 3: Customer Segmentation

- **Description:** Grouped customers by spending habits using unsupervised learning.
- **Unique Learning:** Choosing K using Elbow Method and visualizing clusters.
- **Skills:** Clustering, KMeans, PCA, Elbow Method

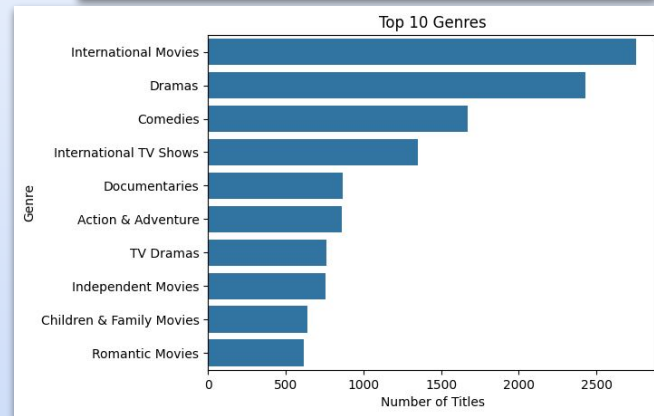
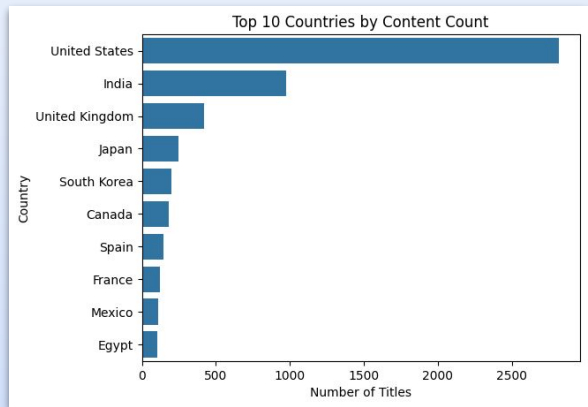
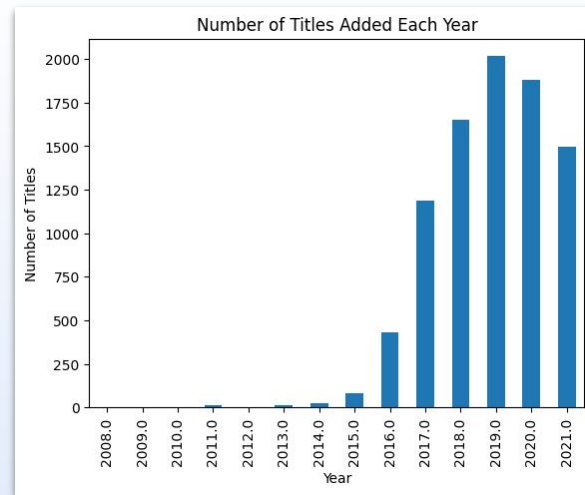
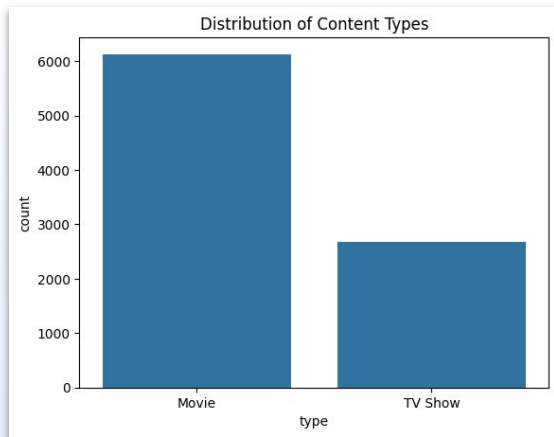


- Visualizing Customer Segments Based on Annual Income and Spending Score using K-Means Clustering.
- This clearly shows the separation between customer groups and highlights behavioral patterns useful for targeted marketing strategies.

Customer Segmentation (K-Means)

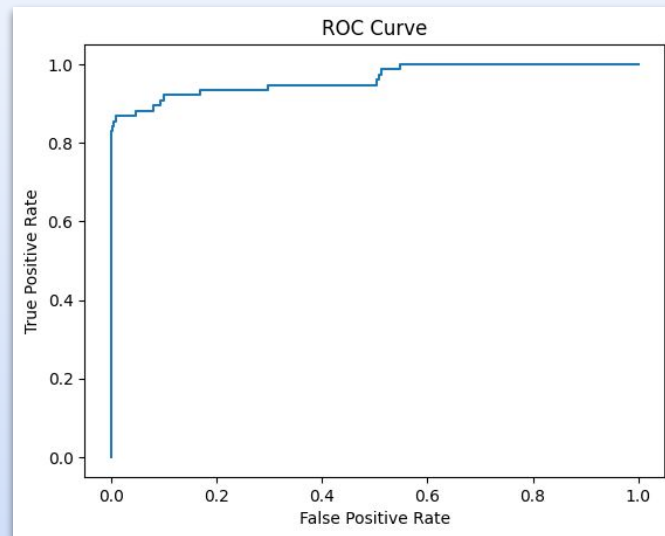
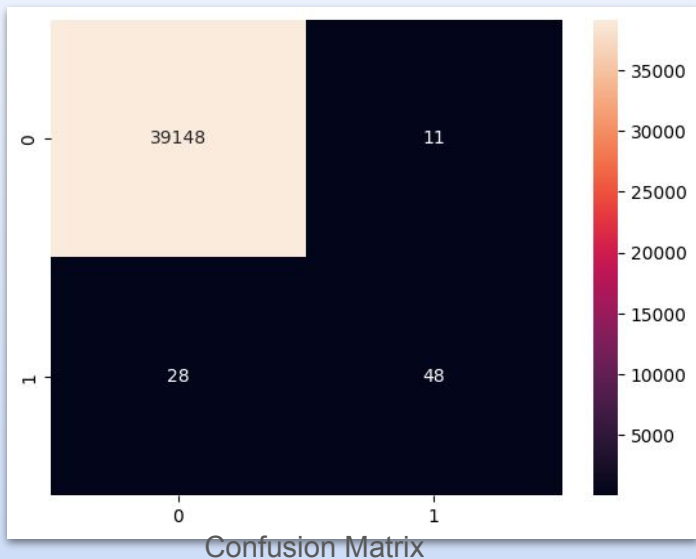
Project 4: Netflix EDA

- **Description:** Explored trends in genres, release years, and countries on Netflix.
- **Unique Learning:** Parsing and converting non-standard date formats.
- **Skills:** EDA, Pandas, Data Cleaning, Time-Series



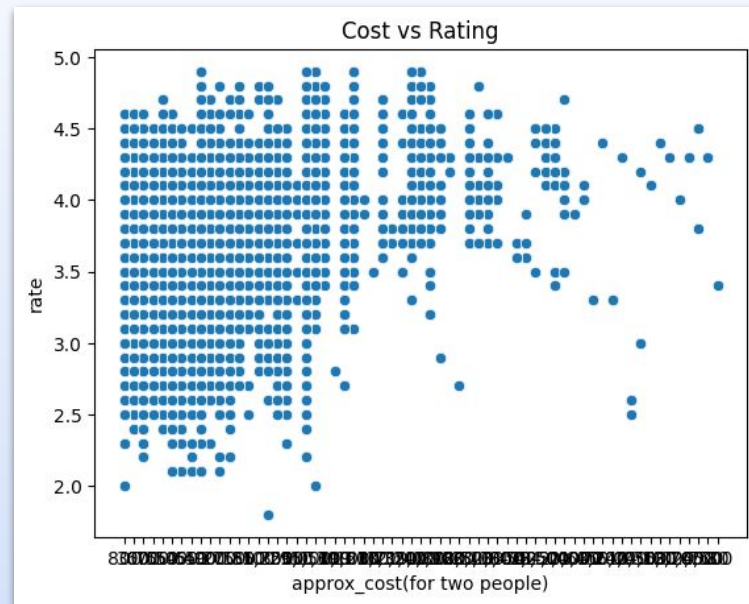
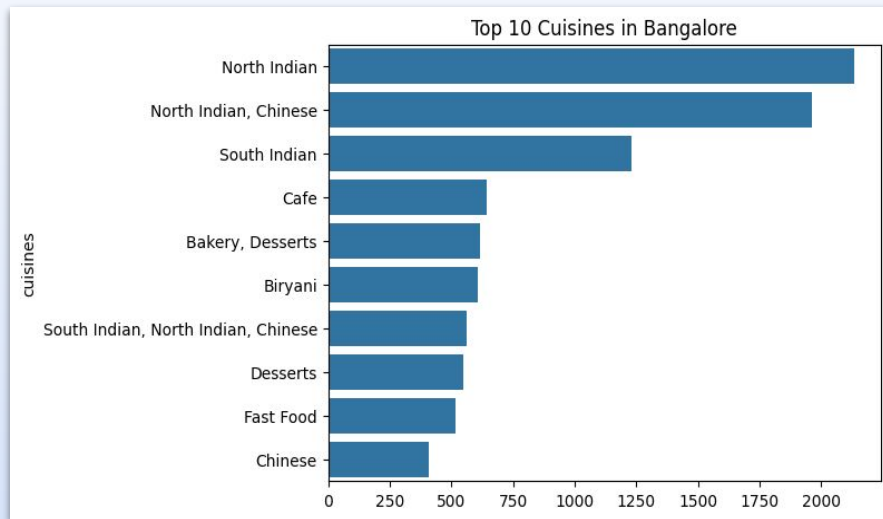
Project 5: Credit Card Fraud Detection

- **Description:** Built a model to detect fraud in imbalanced datasets.
- **Unique Learning:** Using ROC, AUC, and dealing with class imbalance.
- **Skills:** Classification, Imbalanced Data, Model Evaluation



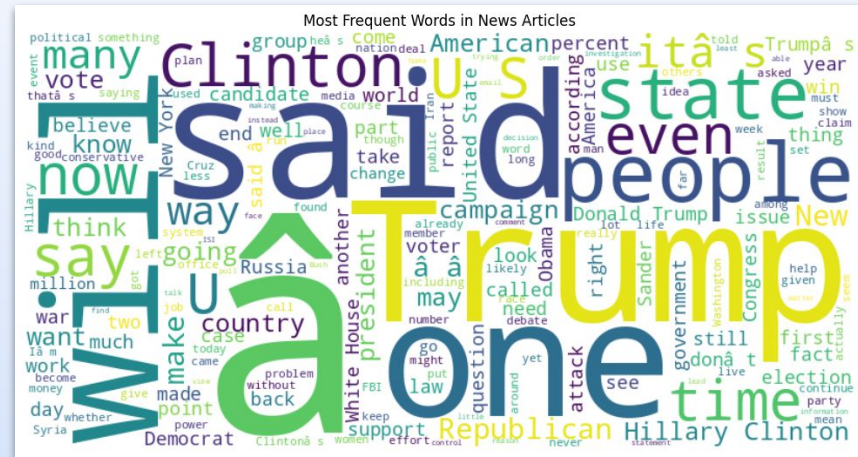
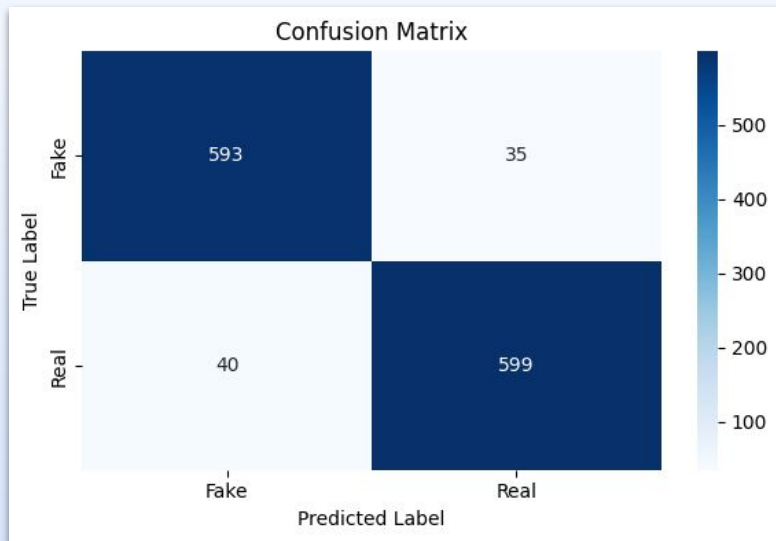
Project 6: Zomato Restaurant EDA

- **Description:** Explored food cost, cuisines, and city-level trends.
- **Unique Learning:** Encoding issues solved using ISO formatting and data wrangling.
- **Skills:** EDA, Python, Text Cleaning, Visualization



Project 7: Fake News Detection

- **Description:** Text classification using TF-IDF and Passive-Aggressive Classifier.
- **Unique Learning:** Complete NLP pipeline from cleaning to evaluation.
- **Skills:** NLP, TF-IDF, Text Classification, Scikit-learn



Word cloud showing most frequent words

Thanks for scrolling! 🙌

I'm always open to collaboration, feedback, and opportunities in Data Science, Analytics & Web Development.

LinkedIn- <https://www.linkedin.com/in/aryan-pange-a11376247/>

GitHub- <https://github.com/aryanpange>

Email- aryanpange010@gmail.com