Enhanced Project Report

1. Customer Lifetime Value Prediction Model

Objective: Predict the lifetime value (LTV) of customers based on their purchase behavior to aid in targeted marketing.

Tools Used: Python (Sklearn, XGBoost), Excel

Solved Problems:

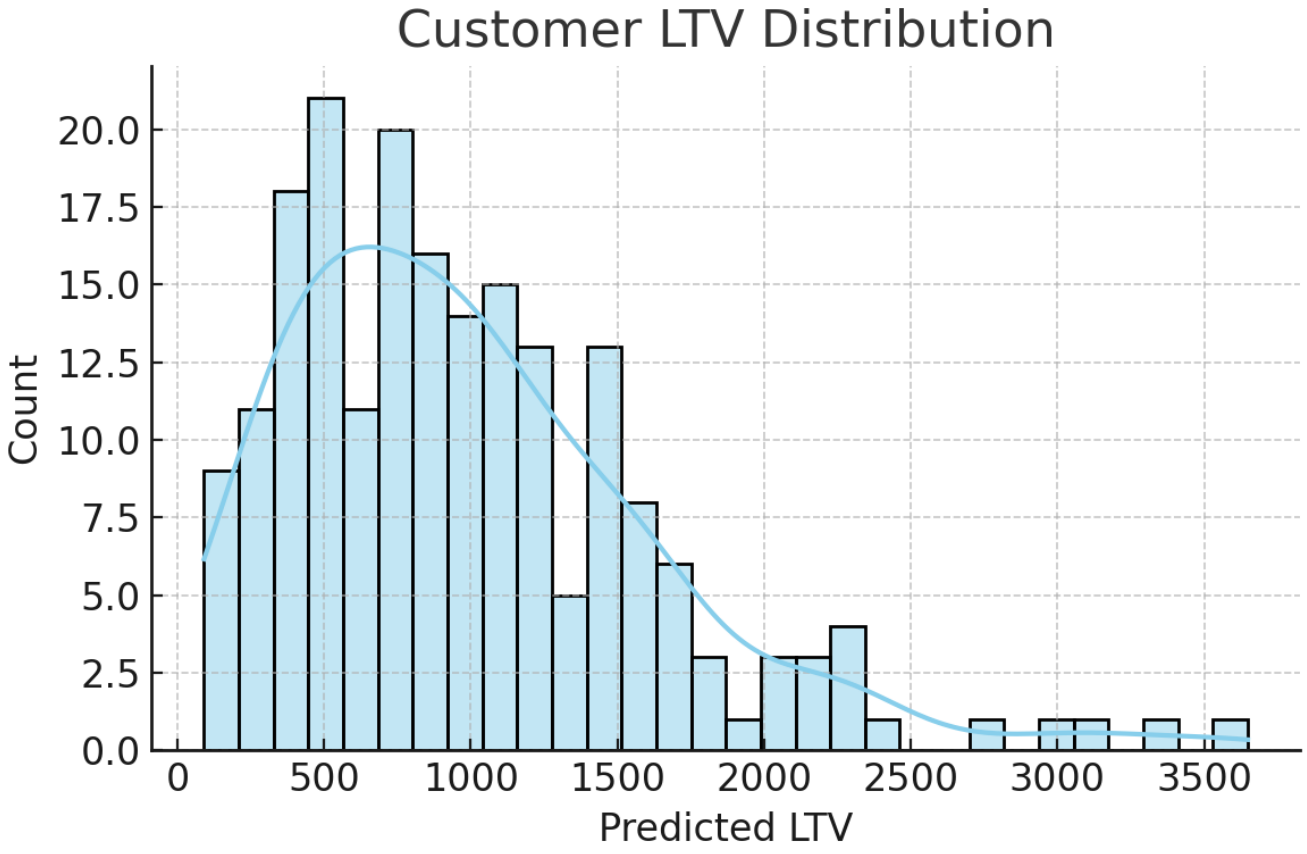
- Identified valuable customers using LTV scores.
- Enabled targeted marketing by segmenting high and low-value customers.
- Provided insights into purchasing behavior through feature engineering.

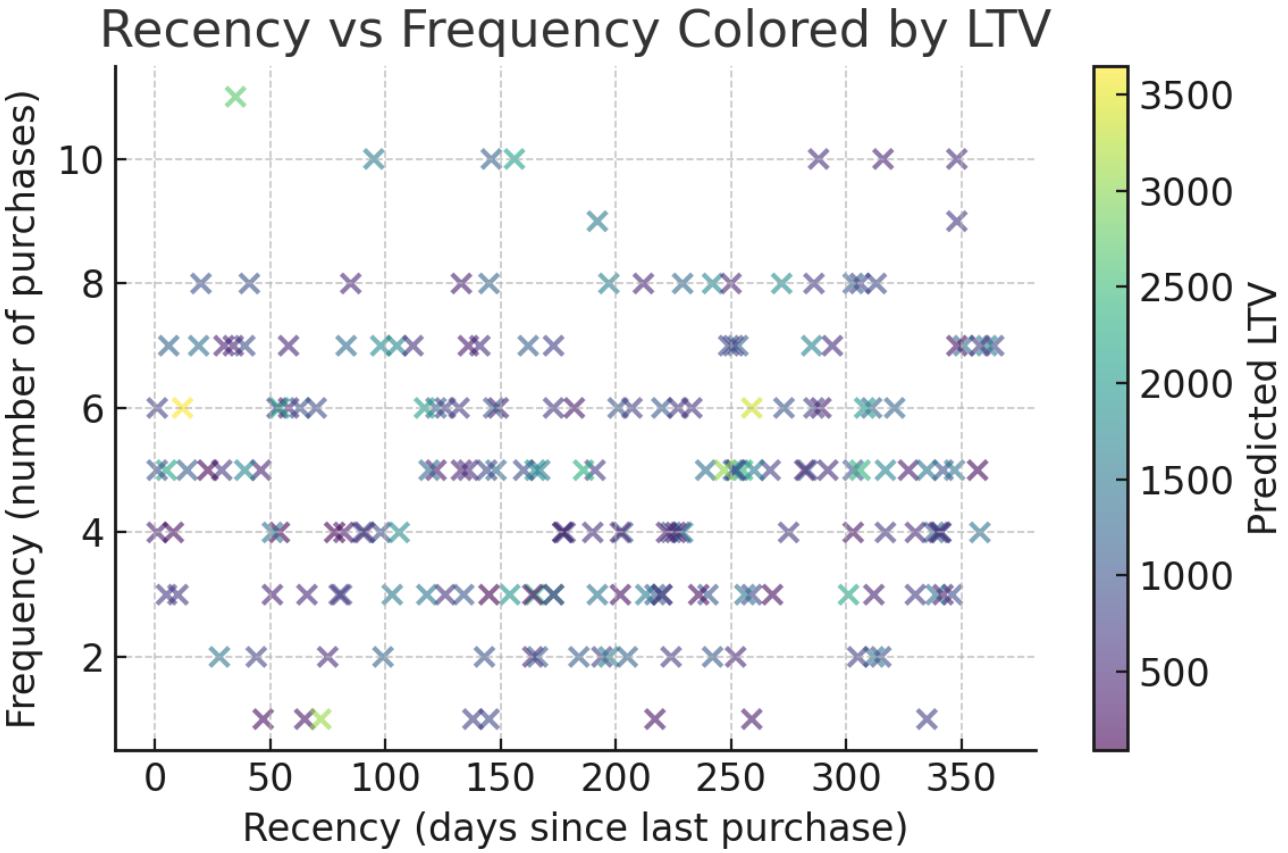
Steps Followed:

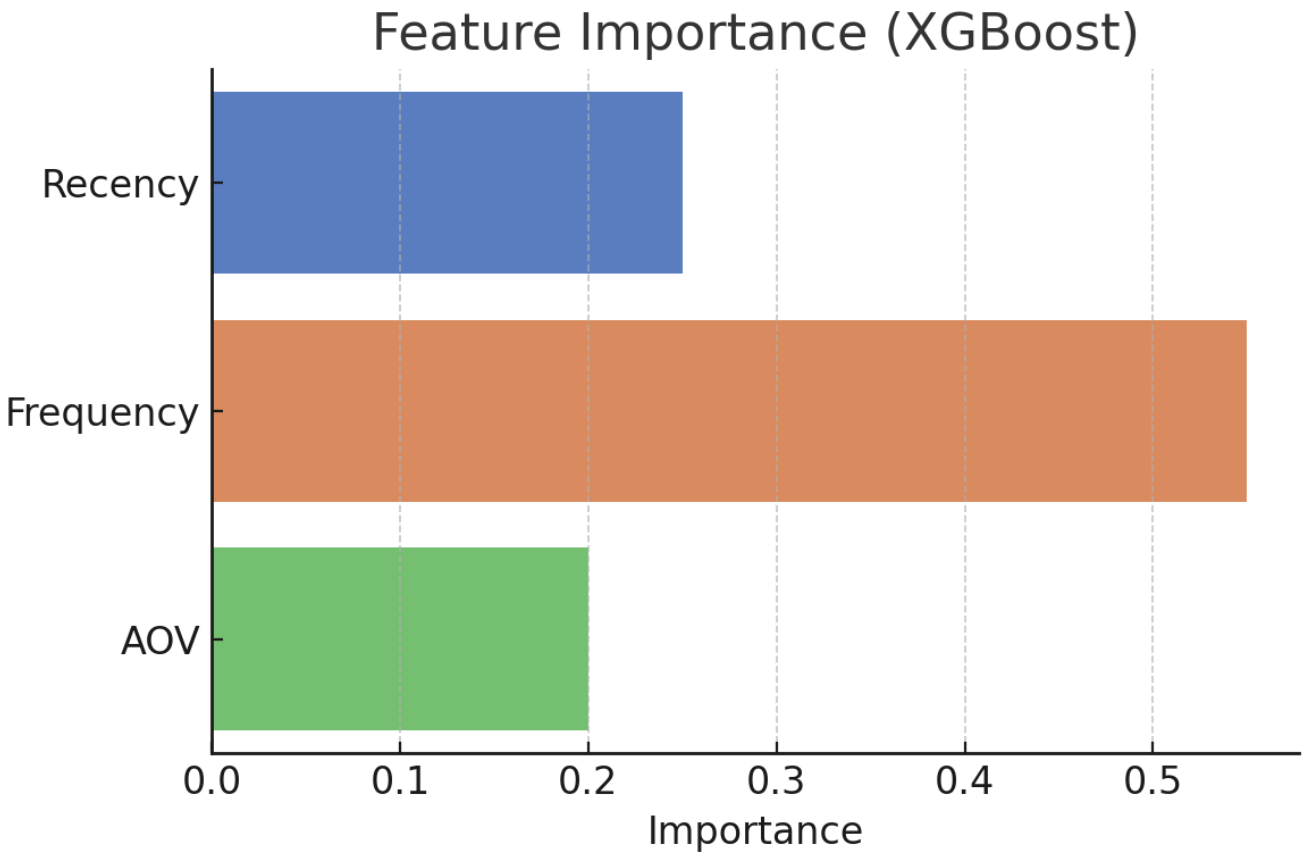
1. Data preprocessing: Merged transaction data with customer profiles.
2. Feature engineering: Calculated Recency, Frequency, and AOV (Avg Order Value).
3. Model training: Used XGBoost Regressor to predict LTV.
4. Model validation: Evaluated using MAE and RMSE.
5. Customer segmentation based on LTV quartiles.

Project Analysis:

- Frequency was the most important predictor of LTV.
- Recency had an inverse relationship with LTV.
- High-frequency, recent buyers were top-value customers.


Customer LTV Distribution
Count
Predicted LTV
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
0
500
1000
1500
2000
2500
3000
3500


Recency vs Frequency Colored by LTV
Frequency (number of purchases)
Recency (days since last purchase)
Predicted LTV
2
4
6
8
10
0
50
100
150
200
250
300
350
500
1000
1500
2000
2500
3000
3500


Feature Importance (XGBoost)
Recency
Frequency
AOV
0.0
0.1
0.2
0.3
0.4
0.5
Importance

2. LinkedIn Job Trend Analysis (Web Scraping)

Objective: Scrape LinkedIn job postings to analyze skill demand trends across cities and roles.

Tools Used: Python (BeautifulSoup, Pandas), Excel

Solved Problems:

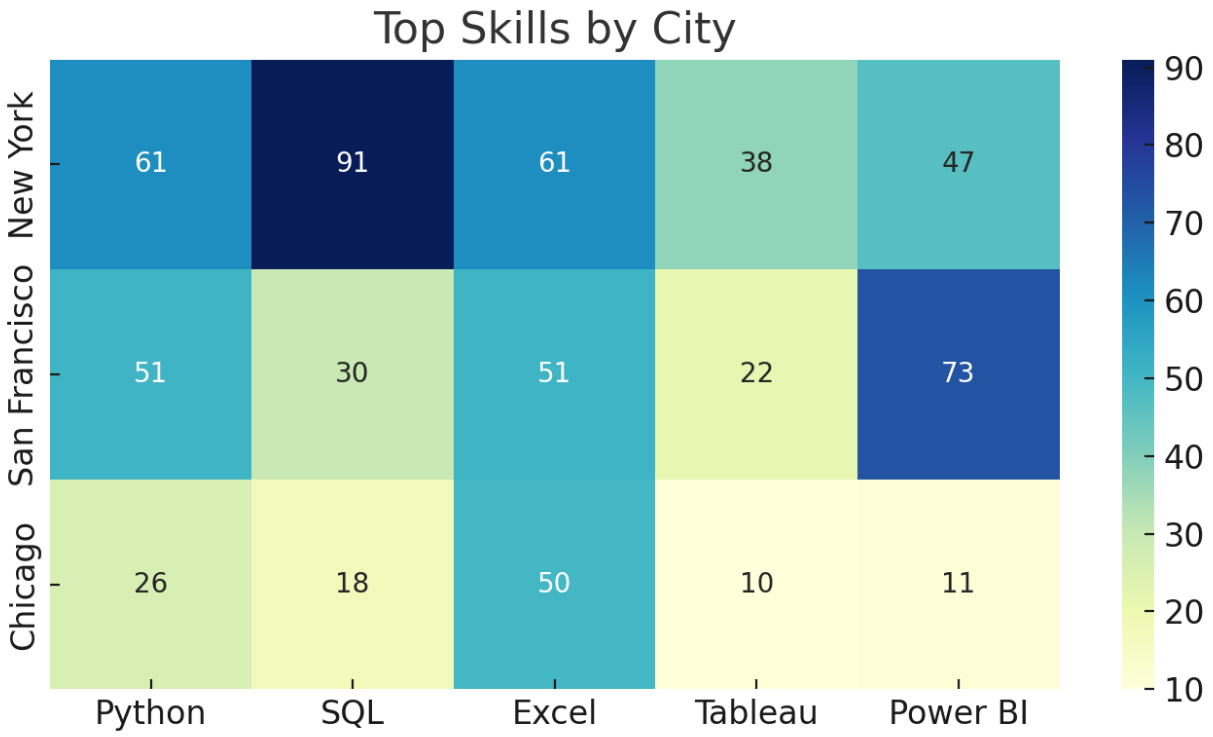
- Identified top in-demand skills by city and role.
- Provided a skill vs role matrix to guide job seekers and recruiters.
- Enabled visual understanding of job market demand.

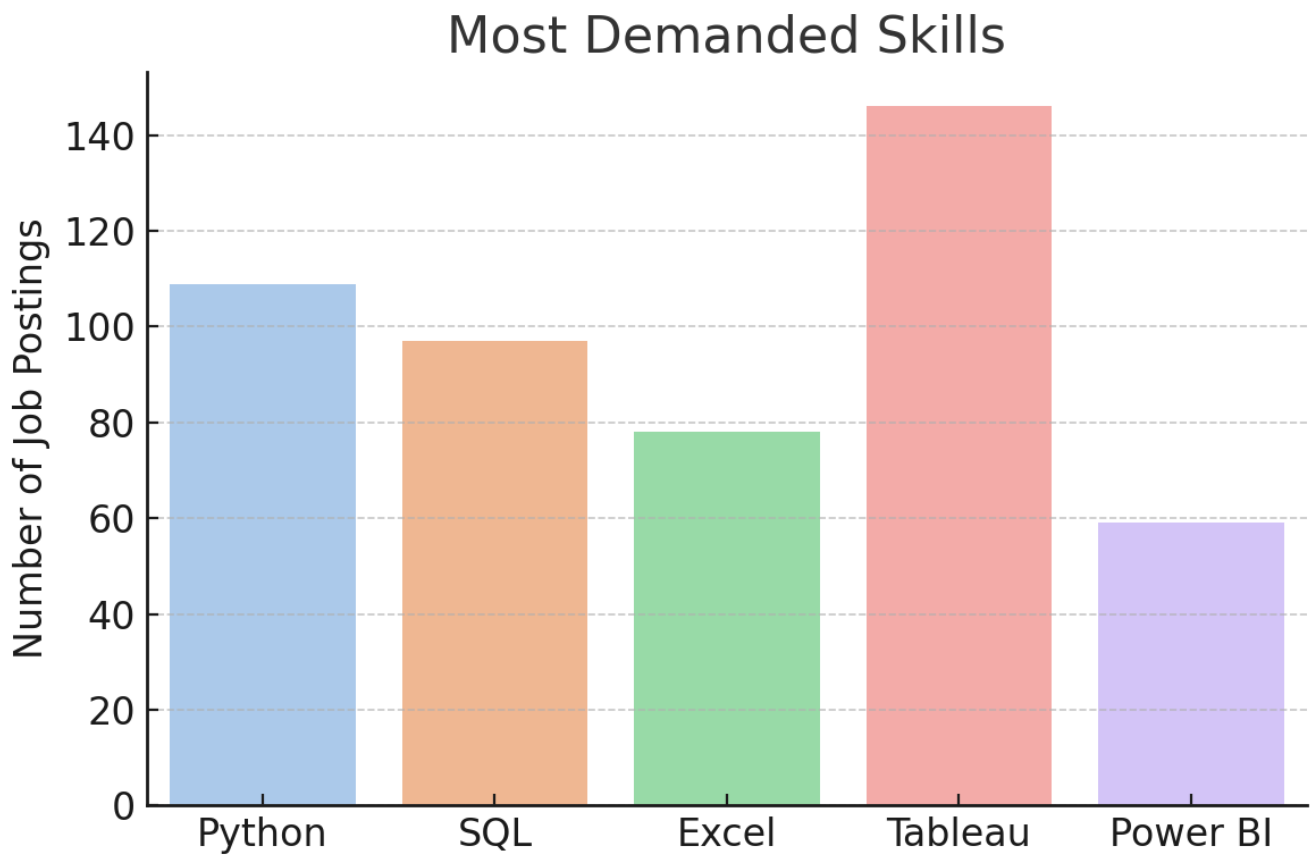
Steps Followed:

1. Scraped job titles, locations, and skills from LinkedIn using BeautifulSoup.
2. Cleaned and parsed the skill tags.
3. Grouped job data by city and role.
4. Generated heatmaps and bar charts to visualize the skill demand.

Project Analysis:

- Python, SQL, and Excel were consistently top-demanded skills.
- Data-centric roles were in higher demand in cities like New York and San Francisco.
- Skill heatmaps revealed specific tool preferences by region.


Top Skills by City
New York
San Francisco
Chicago
61
91
61
38
47
51
30
51
22
73
26
18
50
10
11
Python
SQL
Excel
Tableau
Power BI
90
80
70
60
50
40
30
20
10


Most Demanded Skills
Number of Job Postings
0
20
40
60
80
100
120
140
Python
SQL
Excel
Tableau
Power BI