Applicant Data Dashboard

Upload your file (csv/txt/xlsx/xls)

Drag and drop file here
Limit 200MB per file • CSV, TXT, XLSX, XLS

Browse files

No file uploaded. Using default file from: C:\Users\aashi\Downloads\archive\Applicant_details.csv

Unique Values in Categorical Columns

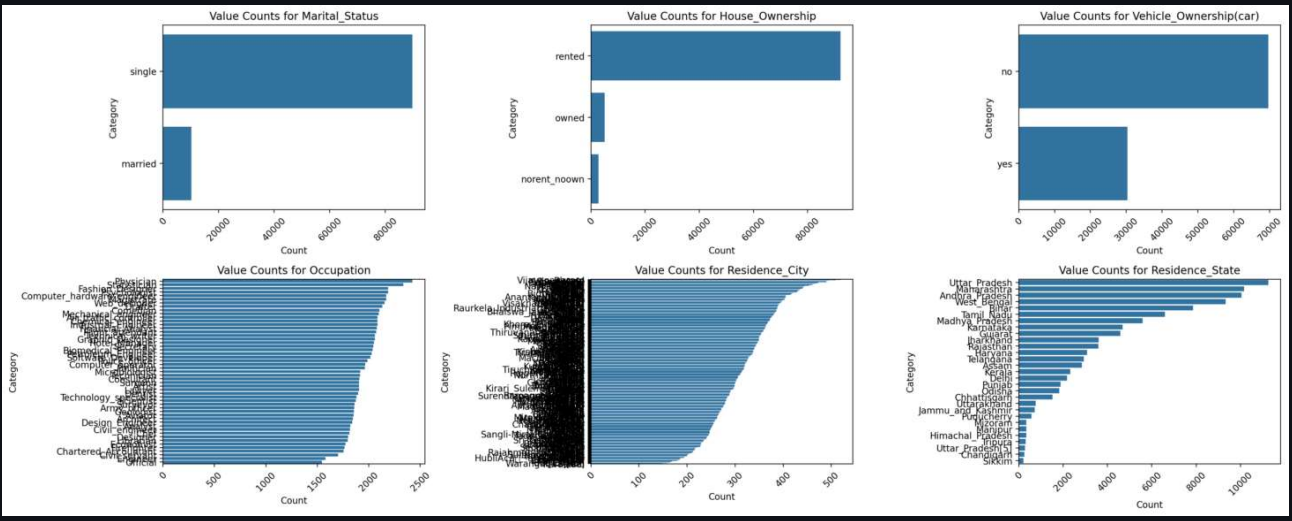
Select Column for Unique Value Count

Marital_Status

Marital_Status: 2 unique values

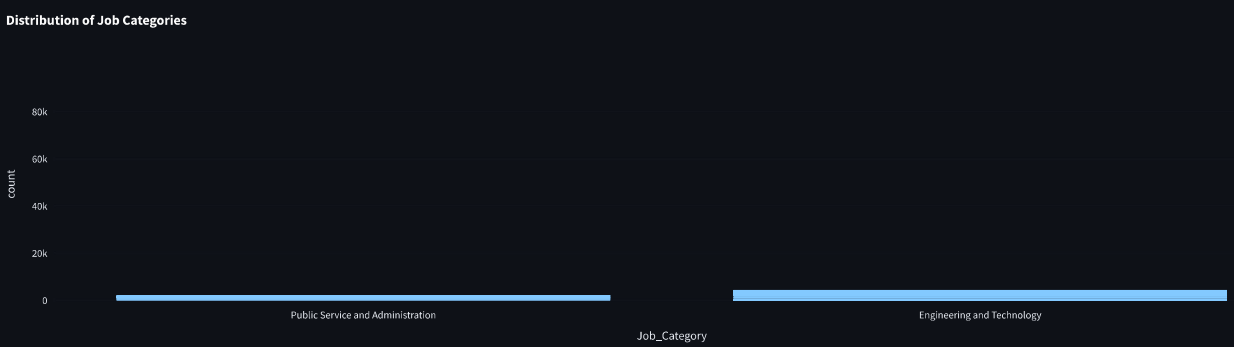
value
single
married

General Value Counts for Categorical Columns



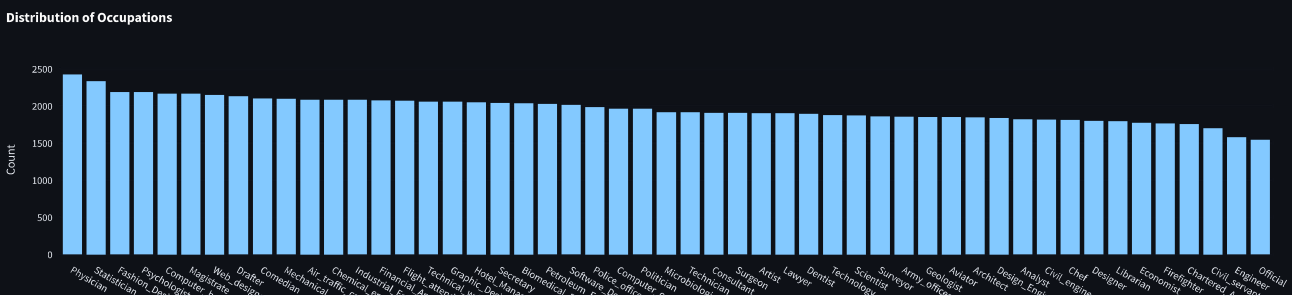
Interpretation: The above visuals show the distribution of categorical variables, providing insights into the diversity of applicant profiles.

Distribution of Job Categories



Interpretation: This chart shows the distribution of applicants across different job categories.

Occupation Distribution

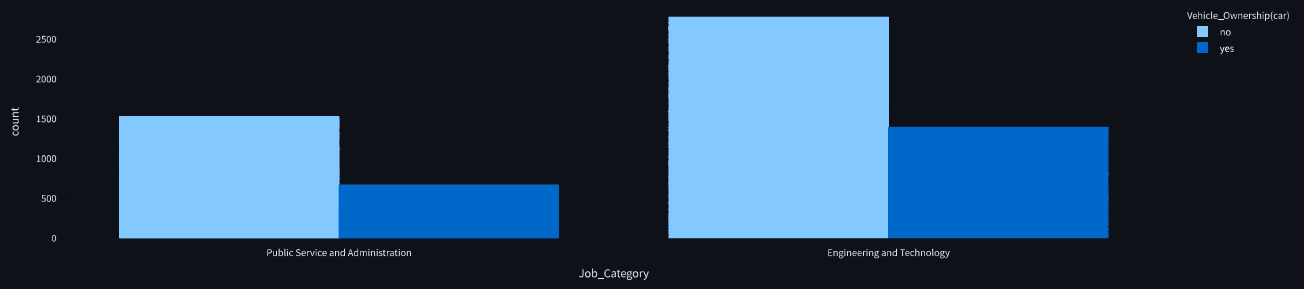


Interpretation: The above chart displays the number of applicants for each unique occupation.

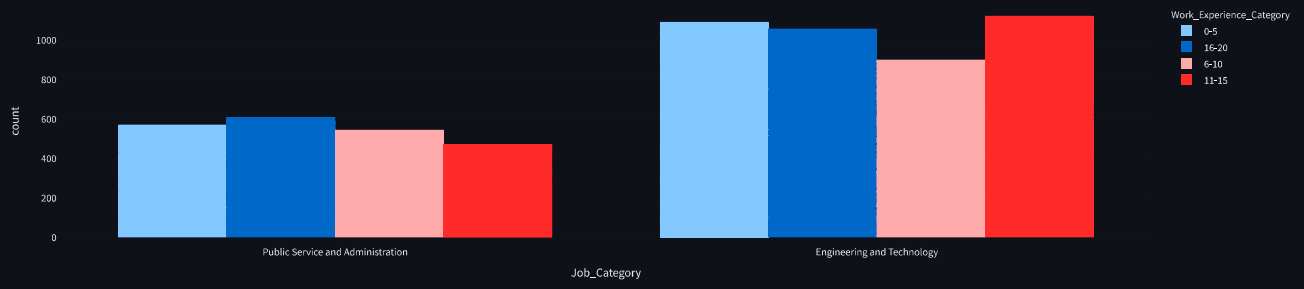
Value Counts by Job Category

Value Counts for House_Ownership by Job Category

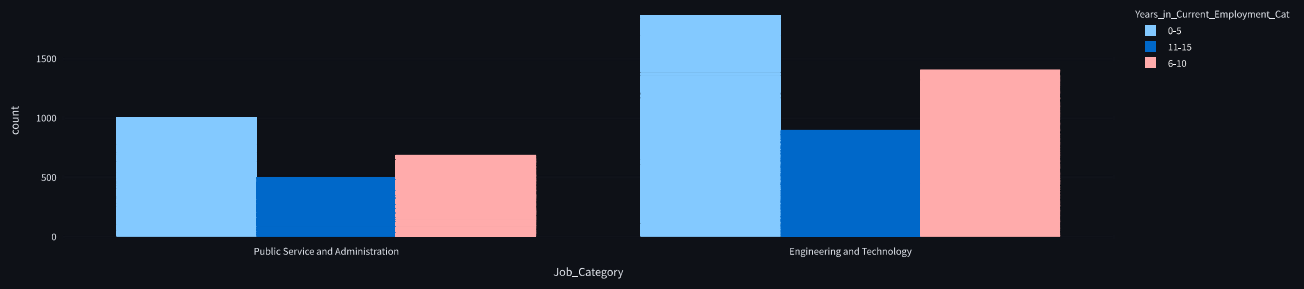
Value Counts for Vehicle_Ownership(car) by Job Category



Value Counts for Work_Experience_Category by Job Category

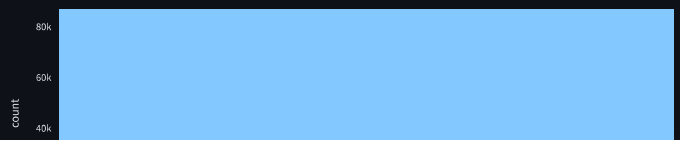


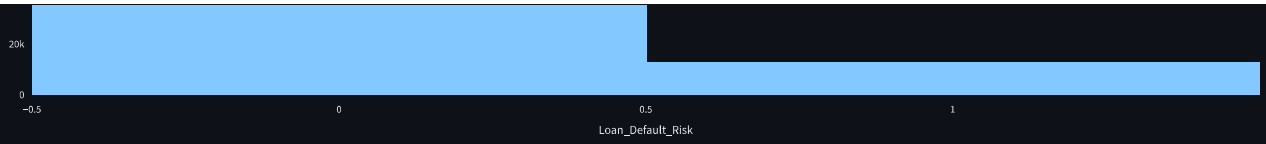
Value Counts for Years_in_Current_Employment_Cat by Job Category



Target Variable Distribution (Loan Default Risk)

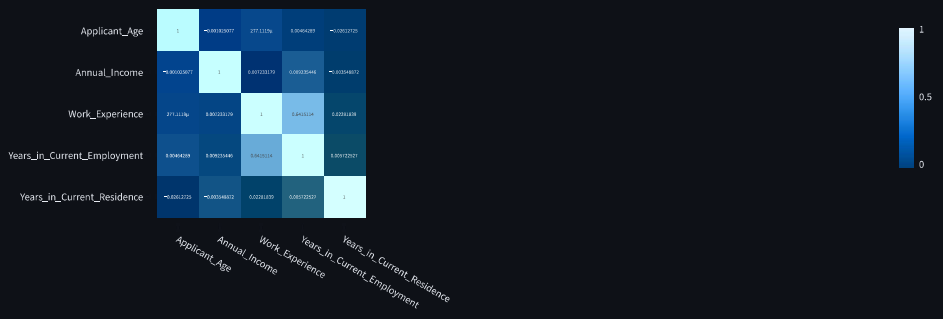
Distribution of Loan Default Risk





Correlation Heatmap

Correlation Matrix



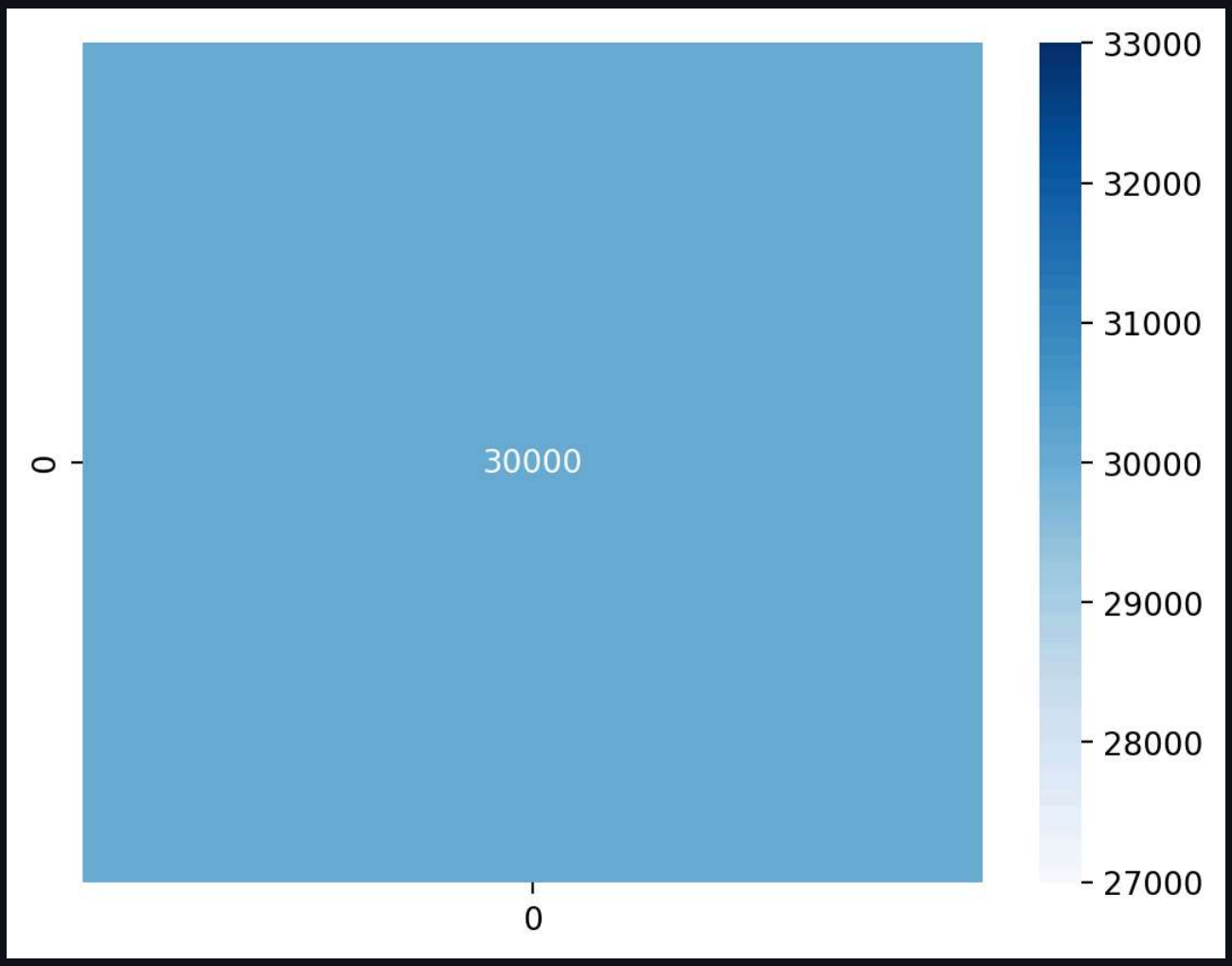
Loan Default Prediction with Decision Tree

Decision Tree Accuracy: 1.00

Decision Tree Classification Report:

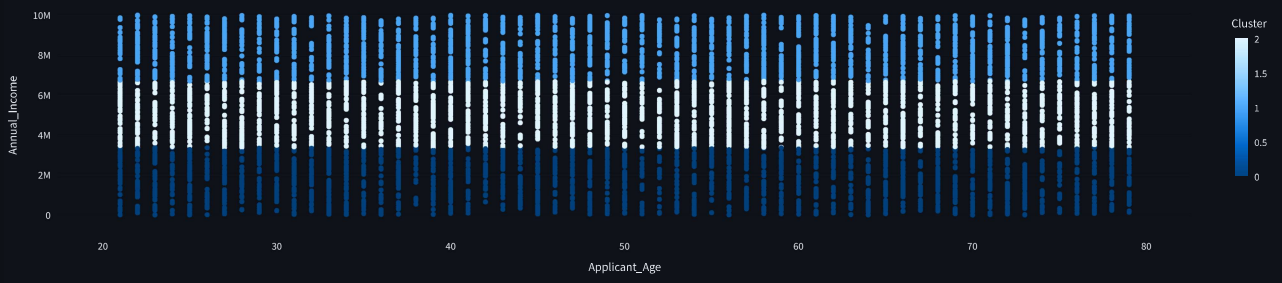
	precision	recall	f1-score	support
0	1.00	1.00	1.00	30000
accuracy			1.00	30000
macro avg	1.00	1.00	1.00	30000
weighted avg	1.00	1.00	1.00	30000

Decision Tree Confusion Matrix



KMeans Clustering for Applicant Segmentation

KMeans Clustering: Applicant Segments



Interpretation: KMeans clustering segments applicants into distinct groups based on age and income.

Final Insights

- The decision tree classifier performed well in predicting loan default risk.
- KMeans clustering helps identify distinct applicant segments based on key numerical features.
- Outlier detection and correlation analysis provide further insights into the dataset.