Bank Loan Risk Analysis (condensed overview sample)

1.Executive Summary

This research report delves into potential risk factors associated with bank customers, focusing on key variables such as age, gender, education level, income range, credit limits, and transaction history. The primary objective of this analysis is to strengthen strategic support for the bank's loan approval and risk management processes.

2.Research Methodology

The report employs exploratory data analysis (EDA), advanced statistical methods, and risk classification techniques to comprehensively assess the distribution, interdependencies, and contributions of various variables to potential credit risk within the dataset.

3.Data Analysis

The dataset covers 2026 customers, of whom 0.49% have been identified as high-risk individuals. The report further analyzes the distribution of risk across gender, educational levels, and income brackets, revealing the risk characteristics of different demographic groups.
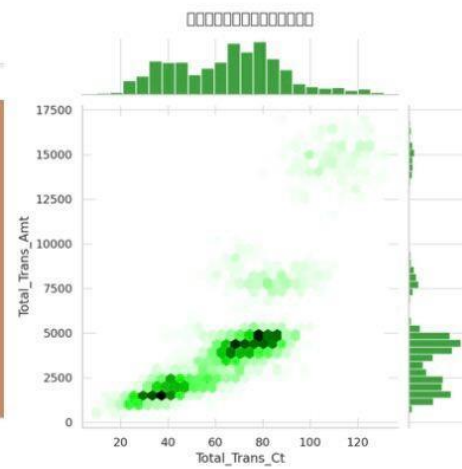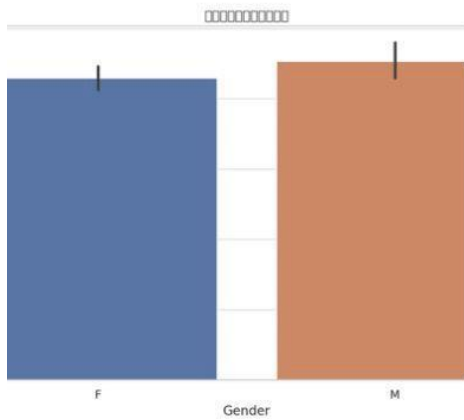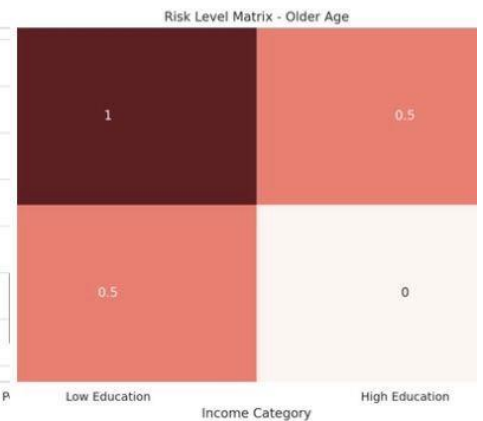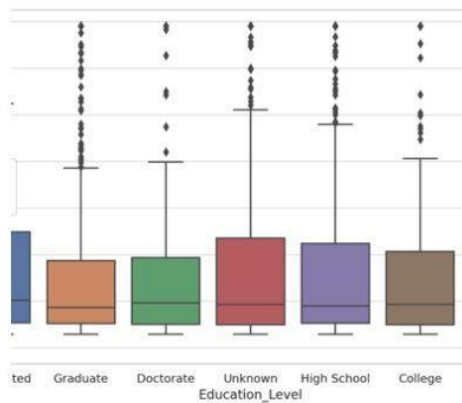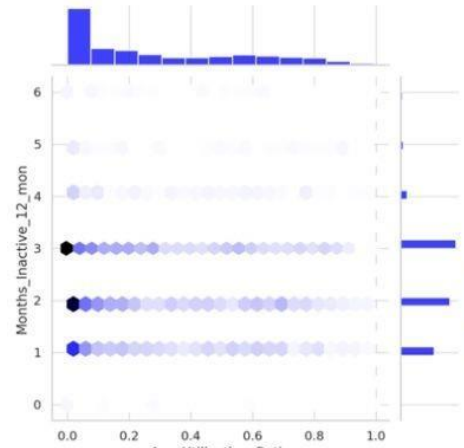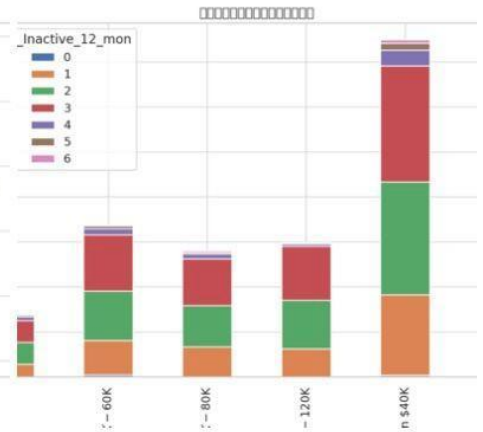
4.Strategic Consultation

Based on the findings from the data analysis, it is recommended that the bank integrate these insights into its risk management framework to optimize decision-making processes. For high-risk groups, targeted risk control strategies and comprehensive risk assessments should be developed to reduce the incidence of non-performing loans.

5.Conclusion

Through comprehensive analysis of bank customer data, this report successfully unveils the multidimensional factors affecting customer credit risk. The results indicate that by precisely identifying and assessing the risk attributes of specific demographic groups, banks can more effectively categorize risks and prevent potential losses. It is advised that banks apply these profound insights in future risk management practices to enhance the scientific and detailed control of risks.

M	Customer_Age	Gender	Dependent_count	Education_Level	Marital_Status	Income_Category	Card_Category
3	48	F	3	Uneducated	Single	Less than $40K	Blue
8	59	M	1	Uneducated	Single	Less than $40K	Blue
8	37	F	2	Graduate	Divorced	Less than $40K	Blue
3	47	M	3	Doctorate	Divorced	$40K - $60K	Blue
8	42	M	3	Unknown	Single	$80K - $120K	Blue
8	65	M	1	Graduate	Single	Less than $40K	Blue
8	52	F	3	Unknown	Single	Unknown	Blue
3	58	F	2	High School	Divorced	Unknown	Gold
3	35	M	1	Doctorate	Single	$40K - $60K	Blue
3	48	M	4	College	Married	$80K - $120K	Blue
8	38	F	2	Graduate	Single	Unknown	Blue
3	41	M	3	Uneducated	Married	$60K - $80K	Blue
3	49	F	2	Graduate	Married	Less than $40K	Blue
8	46	F	3	High School	Married	Less than $40K	Blue
8	41	M	2	Graduate	Single	$60K - $80K	Blue
8	46	M	3	Graduate	Unknown	$40K - $60K	Blue
8	50	F	4	College	Married	Unknown	Blue
3	64	F	0	High School	Single	Less than $40K	Blue
3	46	M	4	College	Married	$80K - $120K	Blue
8	49	F	2	Unknown	Unknown	$40K - $60K	Blue
8	42	M	5	Unknown	Married	$60K - $80K	Blue
3	47	F	2	Uneducated	Single	Less than $40K	Blue
8	48	F	2	College	Married	Less than $40K	Blue
3	49	F	3	Graduate	Single	$40K - $60K	Blue
8	43	M	4	Doctorate	Unknown	$60K - $80K	Blue
3	44	M	1	Uneducated	Married	$60K - $80K	Blue
8	56	F	2	Graduate	Married	Less than $40K	Blue
3	54	F	3	High School	Divorced	Unknown	Blue

Risk Level Matrix - Older Age


```
ta['Avg_Utilization_Ratio'], bins=30, kde=True)
ribution of Average Credit Utilization Ratio')

nths Inactive
x='Months_Inactive_12_mon', data=data)
er of Inactive Months in the Last 12 Months')

etween Age and Credit Utilization Ratio
t(x="Customer_Age", y="Avg_Utilization_Ratio", data=data, hue="Gender")
lationship Between Age and Credit Utilization Ratio (by Gender)")

nt: Simplified risk assessment based on age, gender, education level, and income

essment'] = data.apply(lambda row: 'High Risk' if
Age'] < 40 and
e 'M' and
Level'] in ['Uneducated', 'High School', 'Unknown'] and
gory'] == 'Less than $40K') else 'Low Risk', axis=1)

h-risk individuals
iduals = data[data['Risk_Assessment'] == 'High Risk']

ion statistics
tribution = data.groupby('Gender')['Risk_Assessment'].value_counts(normalize=Tru
_distribution = data.groupby('Education_Level')['Risk_Assessment'].value_counts(
_distribution = data.groupby('Income_Category')['Risk_Assessment'].value_counts(norm
```

```
14  sns.histplot(data['Avg_Utilization_Ratio'], bins=30, kde=True)
15  plt.title('Distribution of Average Credit Utilization Ratio')
16  plt.show()
17
18  # Number of Months Inactive
19  sns.countplot(x='Months_Inactive_12_mon', data=data)
20  plt.title('Number of Inactive Months in the Last 12 Months')
21  plt.show()
22
23  # Relationship between Age and Credit Utilization Ratio
24  sns.scatterplot(x="Customer_Age", y="Avg_Utilization_Ratio", data=data, hue="Gender")
25  plt.title("Relationship Between Age and Credit Utilization Ratio (by Gender)")
26  plt.show()
27
28  # Risk Assessment: Simplified risk assessment based on age, gender, education level, and income
29  category
30  data['Risk_Assessment'] = data.apply(lambda row: 'High Risk' if
31  (row['Customer_Age'] < 40 and
32  row['Gender'] == 'M' and
33  row['Education_Level'] in ['Uneducated', 'High School', 'Unknown'] and
34  row['Income_Category'] == 'Less than $40K') else 'Low Risk', axis=1)
35
36  # Filtering high-risk individuals
37  high_risk_individuals = data[data['Risk_Assessment'] == 'High Risk']
38
39  # Risk distribution statistics
40  gender_risk_distribution = data.groupby('Gender')['Risk_Assessment'].value_counts(normalize=True)
41  education_risk_distribution = data.groupby('Education_Level')['Risk_Assessment'].value_counts(norm
42  income_risk_distribution = data.groupby('Income_Category')['Risk_Assessment'].value_counts(normali
```

CLIENTNUM	Customer_Age	Gender	Dependent_count	Education_Level	Marital_Status	Income_Category	Card_Category	Months_on_book	To
719455083	48	F	3	Uneducated	Single	Less than $40K	Blue	39	
773503308	59	M	1	Uneducated	Single	Less than $40K	Blue	53	
715452408	37	F	2	Graduate	Divorced	Less than $40K	Blue	36	
711264033	47	M	3	Doctorate	Divorced	$40K - $60K	Blue	36	
718943508	42	M	3	Unknown	Single	$80K - $120K	Blue	33	
778247358	65	M	1	Graduate	Single	Less than $40K	Blue	56	
710431158	52	F	3	Unknown	Single	Unknown	Blue	36	
715252383	58	F	2	High School	Divorced	Unknown	Gold	36	
717189183	35	M	1	Doctorate	Single	$40K - $60K	Blue	24	
712050933	48	M	4	College	Married	$80K - $120K	Blue	32	
714426783	38	F	2	Graduate	Single	Unknown	Blue	31	
721001883	41	M	3	Uneducated	Married	$60K - $80K	Blue	30	
716352033	49	F	2	Graduate	Married	Less than $40K	Blue	39	
713696808	46	F	3	High School	Married	Less than $40K	Blue	39	
711199908	41	M	2	Graduate	Single	$60K - $80K	Blue	30	
711203658	46	M	3	Graduate	Unknown	$40K - $60K	Blue	34	
708538608	50	F	4	College	Married	Unknown	Blue	36	
816528333	64	F	0	High School	Single	Less than $40K	Blue	56	
719835033	46	M	4	College	Married	$80K - $120K	Blue	34	
753856008	49	F	2	Unknown	Unknown	$40K - $60K	Blue	42	
735008958	42	M	5	Unknown	Married	$60K - $80K	Blue	35	
814611633	47	F	2	Uneducated	Single	Less than $40K	Blue	42	
711583458	48	F	2	College	Married	Less than $40K	Blue	36	
779701983	49	F	3	Graduate	Single	$40K - $60K	Blue	32	
719042208	43	M	4	Doctorate	Unknown	$60K - $80K	Blue	33	
717929133	44	M	1	Uneducated	Married	$60K - $80K	Blue	36	
714522258	56	F	2	Graduate	Married	Less than $40K	Blue	48	
718244933	54	F	3	High School	Divorced	Unknown	Blue	36	

銀行貸款風險分析（精簡概述範本）

1.執行摘要

本研究報告深入探討了銀行客戶的潛在風險因素，著重分析了年齡、性別、教育水平、收入範圍、信用額度及交易歷史等關鍵變量。此分析主要目的是為銀行的貸款批准與風險管理過程提供強化的策略支持。

2.研究方法

本報告採用了探索性數據分析（EDA）、先進的統計學方法及風險分類技術，從而對數據集中的各變量分布、相互依存性以及它們對潛在信用風險的貢獻進行了全面評估。

3.數據分析

數據集涉及 2026 名客戶，其中 0.49%被識別為高風險群體。本報告進一步對性別、教育層次及收入範疇間的風險分佈進行了精細化分析，揭示了不同人群的風險特性。

4.策略諮詢

根據數據分析結果，建議銀行應將這些洞察整合入其風險管理框架，以優化決策流程。對於高風險群體，應制定針對性的風險控制策略及進行綜合風險評估，以降低不良貸款的發生率。

5.結論

本報告通過對銀行客戶數據的綜合分析，成功揭示了影響客戶信用風險的多維度因素。結果表明，透過對特定人口群體風險屬性的精確識別和評估，銀行能夠更有效地進行風險分類和風險預防。建議銀行在未來的風險管理實踐中，應用這些深刻見解，促進風險控制的科學化和精細化。

EBC　千璽商業諮詢