Analyzing Loan Default Risk Indicators

This document analyzes data to visualize possible indicators that a customer will not be able to pay their loan. It examines numeric and categorical features, comparing "Charged Off" and "At-Risk" customers to identify key risk factors and provide recommendations for managing loan default risk.

by Anikó Mádi

Numeric Features Analysis

1. **Annual Income (annual_inc)**:

- Both **Charged Off** and **At-Risk** customers exhibit very similar distributions of annual income, with the median close to 8 (in log scale). There's no significant distinction in income between the two groups.

- **Insight**: Annual income does not seem to be a strong indicator of loan default risk.

2. **Debt-to-Income Ratio (dti)**:

- **At-Risk** customers tend to have a slightly higher median DTI compared to **Charged Off** customers, though the difference is marginal.

- **Insight**: DTI could be a minor contributing factor, but it is not a dominant indicator of default risk based on this comparison.

3. **Employment Length (employment_length)**:

- The distribution of employment length between the two groups is nearly identical, suggesting no major differences in the length of employment.

- **Insight**: Employment length does not appear to be a key differentiator between **Charged Off** and **At-Risk** customers.

4. **Loan Amount (loan_amount)**:

- The distributions of loan amounts are very similar between the two groups, with slightly lower median loan amounts for **At-Risk** customers.

- **Insight**: Loan amount is not a strong indicator of default risk as both groups share similar characteristics.

5. **Open Accounts (open_accounts)**:

- The number of open accounts shows similar distributions between both groups, though **At-Risk** customers have slightly higher median open accounts.

- **Insight**: This feature doesn't provide a clear distinction between the two groups in predicting loan defaults.

6. **Outstanding Principal (out_prncp)**:

- The **At-Risk** group shows much higher outstanding principal values compared to **Charged Off** customers, whose values are near zero.

- **Insight**: This is a critical indicator of risk since customers with higher outstanding principal amounts are more likely to default. The large outstanding principal in the **At-Risk** group indicates a potential future loss.

7. **Total Accounts (total_accounts)**:

- There's no significant difference in the total number of accounts between **Charged Off** and **At-Risk** customers.

- **Insight**: Total accounts is not a reliable indicator of default risk.

8. **Total Payment (total_payment)**:

- **At-Risk** customers have made higher total payments than **Charged Off** customers, suggesting they may still be active, but with delayed payments.

- **Insight**: This could indicate that while these customers are making some payments, they may still pose a future risk of default if their payments slow or stop.

Annual Income

Similar distributions for both groups. Not a strong indicator of default risk.

Outstanding Principal

Critical indicator. At-Risk group shows much higher values, indicating potential future loss.

Total Payment

At-Risk customers made higher payments, but may still pose future default risk.

Categorical Features Analysis

1. **Loan Grade (grade)**:

- **Charged Off** customers are concentrated more heavily in lower loan grades (D, E, F, G), while **At-Risk** customers are more evenly distributed across higher loan grades (A, B, C).

- **Insight**: Loan grade is a significant indicator, as customers with lower grades are more likely to default.

2. **Home Ownership (home_ownership)**:

- Both **Charged Off** and **At-Risk** customers are mostly renters or mortgage holders, though **Charged Off** customers have a higher proportion of **renters**.

- **Insight**: Home ownership (specifically renting) seems to be associated with a higher risk of loan default.

3. **Loan Purpose (purpose)**:

- **Charged Off** customers primarily took loans for **debt consolidation** and **credit card payments**, while **At-Risk** customers also have a high concentration in these areas, but with a more balanced spread across other purposes.

- **Insight**: The loan purpose, especially **debt consolidation** and **credit card repayment**, is a strong indicator of default risk.

1 Loan Grade

Lower grades (D, E, F, G) are strongly associated with default risk.

2 Home Ownership

Renters show a higher tendency to default than homeowners or those with mortgages.

3 Loan Purpose

Debt consolidation and credit card repayment loans have a higher likelihood of default.

Recoveries Analysis

Recoveries (recoveries):

- Both groups have minimal recoveries, indicating that once a customer defaults, there is little chance of recovering the outstanding amounts.

- **Insight**: Recoveries do not appear to be a meaningful factor for understanding default risk.

Key Indicators of Default Risk:

Based on this analysis, the following features are most likely to contribute to **loan default risk**:

1. **Outstanding Principal (out_prncp)**: Customers with high outstanding principal are at significant risk of default.
2. **Loan Grade**: Lower grades (D, E, F, G) are strongly associated with default.
3. **Loan Purpose**: Debt consolidation and credit card repayment loans tend to have a higher likelihood of default.
4. **Home Ownership**: Renters show a higher tendency to default than homeowners or those with mortgages.

Recommendations

1. **Focus on Monitoring Lower Loan Grades**: Customers with low loan grades (D, E, F, G) should be closely monitored for payment delays or increased risk of default.
2. **Prioritize Debt Consolidation Loans for Early Action**: Debt consolidation and credit card repayment loans should be given priority when assessing risk and offering restructuring options.
3. **Target Renters with Support Options**: Since renters are at higher risk, the company should consider targeted financial support or early intervention for these customers.
4. **Monitor Customers with High Outstanding Principal**: A significant outstanding principal balance should be flagged as an early warning sign for potential default.

Let me know if you'd like further details or specific actions based on this analysis!

1

Monitor Lower Loan Grades

Focus on customers with grades D, E, F, G for potential default risk.

2

Prioritize Debt Consolidation Loans

Assess risk and offer restructuring options for these high-risk loan types.

3

Support Renters

Provide targeted financial support or early intervention for renters.

4

Track High Outstanding Principal

Flag significant outstanding balances as early warning signs.

Additional Categorical Comparison Findings

Home Ownership:

- **At-Risk Customers**: 202 renters, 192 mortgage holders, 46 owners, and no "other."
- **Charged Off Customers**: Higher proportion of renters (694) and mortgage holders (690), with fewer owners (141) and very few "other" (2).
- **Insight**: Customers who rent their homes seem to have a higher likelihood of being charged off compared to those who own or have a mortgage. Renting appears to be a strong indicator of risk.

Loan Purpose:

- The distribution of **loan purposes** (e.g., debt consolidation, credit card repayment) is similar between the two groups. However, there seems to be a higher number of **Charged Off** customers in certain categories.
- **Insight**: Loan purpose, especially for debt consolidation and credit card repayment, continues to be a risk factor for default, as observed in earlier findings.

Loan Grade:

- The distribution of loan grades shows that **Charged Off** customers have more representation in the lower loan grades, which correlates with earlier observations.
- **Insight**: Customers in lower loan grades (D, E, F, G) are at a higher risk of defaulting.

Feature	At-Risk Customers	Charged Off Customers
Renters	202	694
Mortgage Holders	192	690
Owners	46	141

Indicator Correlation Summary and Conclusion

Indicator Correlation Summary:

1. **Loan Amount**: **At-Risk Customers**: Average loan amount is slightly higher (14,408) compared to **Charged Off** customers (14,322). **Insight**: Loan amount does not seem to significantly differ between the two groups, and thus it's not a strong predictor of default risk.
2. **Annual Income**: Both groups have very similar **annual income** levels, with slight variations. **Insight**: Annual income, as in earlier analyses, does not provide a strong distinction between **At-Risk** and **Charged Off** customers.
3. **Debt-to-Income Ratio (DTI)**: **At-Risk Customers** have a higher **DTI** (19.22) compared to **Charged Off** customers (17.87). **Insight**: DTI could be a moderate indicator of risk, with higher ratios suggesting a potential struggle to meet debt obligations.
4. **Employment Length**: Both groups have very similar average **employment lengths**, with **At-Risk** customers having a slightly higher average. **Insight**: Employment length does not appear to be a distinguishing factor between these two groups.
5. **Open and Total Accounts**: **At-Risk Customers** have more **total accounts** (23.88) and slightly more **open accounts** (3.20) compared to **Charged Off** customers (24.73 total accounts and 3.16 open accounts). **Insight**: While not a major differentiator, a higher number of total accounts could indicate potential risk of default due to a heavier debt burden.
6. **Outstanding Principal (out_prncp)**: **At-Risk Customers** have a much higher **outstanding principal** (4.27) compared to **Charged Off** customers (0). **Insight**: This is a crucial risk indicator. Customers with higher outstanding principal amounts pose a significant risk of future default.
7. **Total Payment**: **At-Risk Customers** have made larger total payments (52.28) compared to **Charged Off** customers (43.59). **Insight**: Despite making more payments, **At-Risk** customers are still falling behind, indicating that while they continue making partial payments, they remain at high risk of default.

Conclusion and Recommendations:

1. **Key Risk Indicators**:

- **Home Ownership**: Renters are at higher risk.
- **Loan Grade**: Lower loan grades (D, E, F, G) correlate with a higher chance of default.
- **Debt-to-Income Ratio**: Higher DTI is associated with higher risk.
- **Outstanding Principal**: Customers with higher outstanding principal are at risk of defaulting.
- **Total Payment**: Customers making higher total payments may still be at risk if they are falling behind in installments.

2. **Next Steps**:

- Focus on **outstanding principal** and **debt-to-income ratios** to identify at-risk customers early.
- Monitor customers in lower loan grades (D-G) and renters more closely.
- Prioritize actions for customers with high DTI and outstanding balances to offer restructuring or support before they default.

Home Ownership

Renters at higher risk

Loan Grade

Lower grades correlate with default

DTI & Principal

Higher values indicate risk

Early Action

Monitor and support at-risk customers