Advanced Analytics

The **Advanced Analytics** phase focuses on deriving actionable insights. The aim is to perform risk scoring and segment the borrower population to identify high-risk groups.

1. Analysis of borrower solvency
a. Acceptance Rate Based on Work Experience

The acceptance rate based on work experience is an important metric for understanding how the professional background of applicants influences the likelihood of loan approval. By analysing how experience impacts loan acceptance, we can identify trends and refine lending strategies, targeting applicants with a higher chance of loan approval.

SQL query:

```sql
SELECT
    CASE
        WHEN person_emp_exp < 3 THEN 'Under 3'
        WHEN person_emp_exp >= 3 AND person_emp_exp < 5 THEN '3-5'
        WHEN person_emp_exp >= 5 AND person_emp_exp < 10 THEN '5-10'
        WHEN person_emp_exp >= 10 THEN '10+'
    END AS person_emp_exp_range,
    COUNT(*) AS total_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) AS accepted_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) * 100 / COUNT(*) AS acceptence_rate
FROM loan_dataset
GROUP BY
    CASE
        WHEN person_emp_exp < 3 THEN 'Under 3'
        WHEN person_emp_exp >= 3 AND person_emp_exp < 5 THEN '3-5'
        WHEN person_emp_exp >= 5 AND person_emp_exp < 10 THEN '5-10'
        WHEN person_emp_exp >= 10 THEN '10+'
    END
ORDER BY person_emp_exp_range;
```
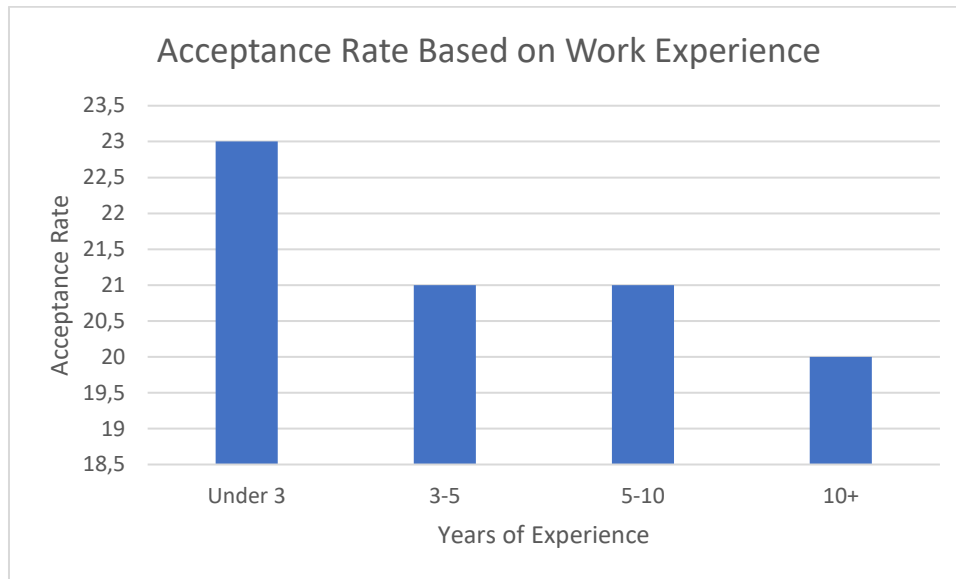
Figure 1 : Acceptance Rate Based on Work Experience

Result:

person_emp_exp_range	total_loans	accepted_loans	acceptence_rate
Under 3	17761	4172	23
3-5	7414	1608	21
5-10	11386	2454	21
10+	8439	1766	20

Figure 2 : Acceptance Rate Based on Work Experience Result

Illustration:

From the chart "Acceptance Rate Based on Work Experience":

- Under 3: 23
- 3-5: 21
- 5-10: 21
- 10+: 20

X-axis: Years of Experience
Y-axis: Acceptance Rate

Interpretation:

From the results, we can observe that applicants with **more work experience** tend to have a higher acceptance rate.

This suggests that HSBC may favour more experienced borrowers, likely due to their higher job stability and creditworthiness. This trend could inform future loan approval criteria and help the bank optimize its risk assessment models.

b. Home Ownership and Loan Acceptance Rate

The relationship between home ownership and loan acceptance is essential for understanding how property ownership influences the likelihood of loan approval. Homeowners are often perceived as more financially stable and less risky by lenders, which could result in a higher rate of loan acceptance. By analysing this relationship, HSBC can adjust lending criteria to better target high-potential applicants.

SQL query:

```sql
SELECT
    person_home_ownership,
    COUNT(*) AS total_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) AS accepted_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) * 100.0 / COUNT(*) AS acceptance_rate
FROM loan_dataset
GROUP BY person_home_ownership
ORDER BY acceptance_rate DESC;
```
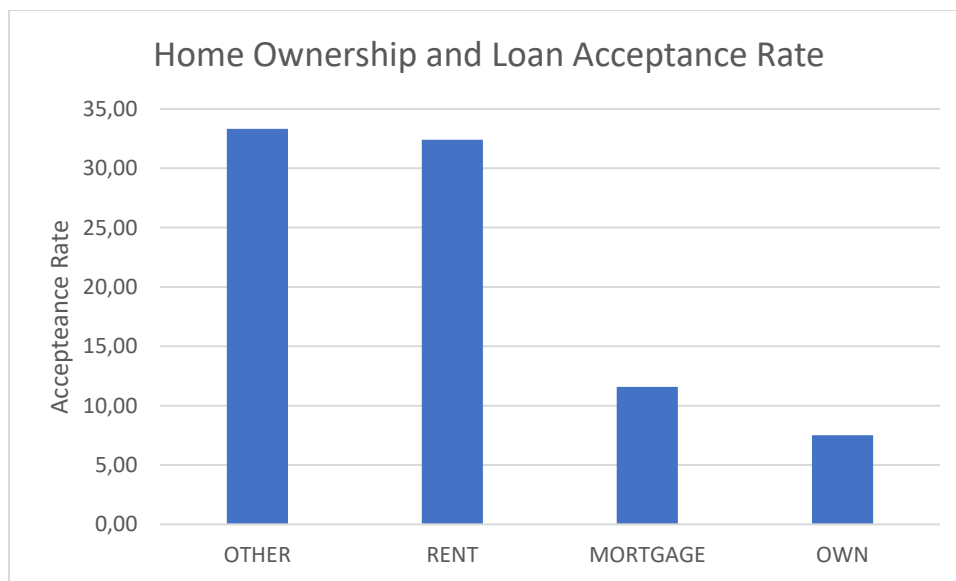
Figure 3 : Home Ownership and Loan Acceptance Rate

Result:

person_home_ownership	total_loans	accepted_loans	acceptance_rate
OTHER	117	39	33.333333333333
RENT	23443	7595	32.397730665870
MORTGAGE	18489	2144	11.596084158148
OWN	2951	222	7.522873602168

Figure 4 : Home Ownership and Loan Acceptance Rate Result

Illustration:

Interpretation:

From the results, we observe the following trends:

- **Homeowners** (those with a mortgage or owning outright) have lower acceptance rates than renters.
- **Others** (people in other living arrangements, like co-ownership or living with family) have an acceptance rate of **33.33%**, which is the highest rates.

The analysis demonstrates that **Others living is positively correlated with loan acceptance**. Own holders are the most likely to don't have their loan applications approved, while renters tend to have a higher acceptance rate.

c. Impact of Previous Loan Defaults on New Loan Approval

The relationship between previous loan defaults and loan acceptance is crucial for understanding the risk associated with applicants who have a history of defaults. Borrowers who have previously defaulted on loans may be perceived as higher risk by lenders, leading to a lower acceptance rate for new loans. Analysing this relationship allows HSBC UK to better assess applicants with a history of loan defaults and refine lending criteria.

SQL query:

```sql
SELECT
    previous_loan_defaults_on_file,
    COUNT(*) AS total_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) AS accepted_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) * 100.0 / COUNT(*) AS acceptance_rate
FROM loan_dataset
GROUP BY previous_loan_defaults_on_file
ORDER BY acceptance_rate DESC;
```
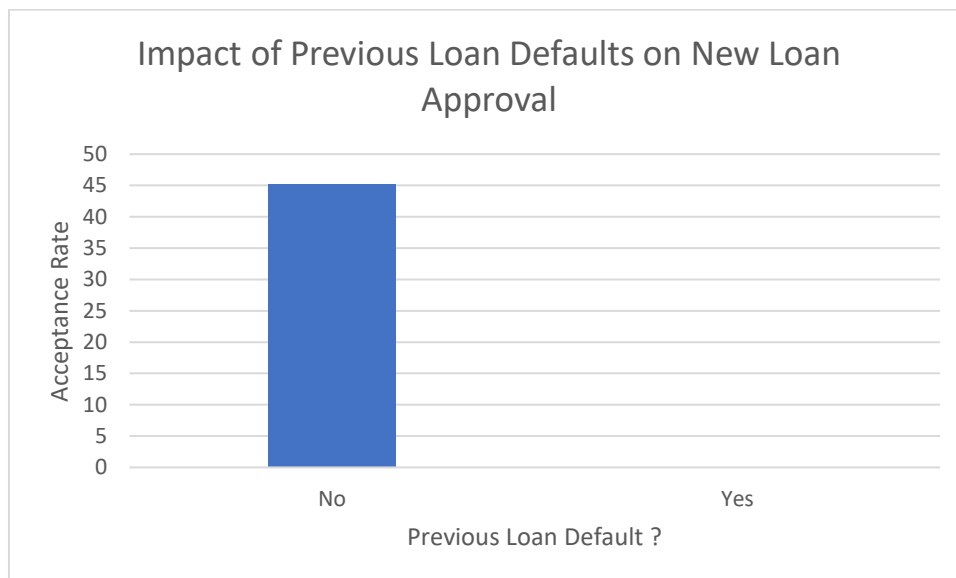
Figure 5 : Impact of Previous Loan Defaults on New Loan Approval

Result:

previous_loan_defaults_on_file	total_loans	accepted_loans	acceptance_rate
No	22142	10000	45.163038569234
Yes	22858	0	0.000000000000

Figure 6 : Impact of Previous Loan Defaults on New Loan Approval Result

Illustration:

Interpretation:

From the results, we can make the following observations:

- **Applicants without previous loan defaults** (indicated by previous_loan_defaults_on_file = No) have a much higher acceptance rate at **45.16%**.
- **Applicants with previous loan defaults** (indicated by previous_loan_defaults_on_file = Yes) have a significantly lower acceptance rate at **0%**.

The analysis shows a clear negative correlation between previous loan defaults and loan acceptance. Borrowers with a history of loan defaults are **not likely** to have their loan applications approved compared to borrowers without such a history. This aligns with the general principle that lenders tend to view applicants with prior defaults as higher risk.

d. Impact of Credit Score on Loan Acceptance Rate

The credit score is a critical factor in determining the likelihood of a borrower being approved for a loan. Typically, the higher the credit score, the more likely the borrower is to be accepted for a loan, as they are considered lower risk. By analysing the relationship between credit score and loan acceptance, HSBC UK can identify the threshold at which the acceptance rate starts to decrease significantly.

SQL query:

```sql
SELECT
    CASE
        WHEN credit_score >= 300 AND credit_score < 500 THEN '300-499'
        WHEN credit_score >= 500 AND credit_score < 600 THEN '500-599'
        WHEN credit_score >= 600 AND credit_score < 700 THEN '600-699'
        WHEN credit_score >= 700 AND credit_score < 800 THEN '700-799'
        WHEN credit_score >= 800 THEN '800+'
    END AS credit_score_range,
    COUNT(*) AS total_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) AS accepted_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 END) * 100.0 / COUNT(*) AS acceptance_rate
FROM loan_dataset
GROUP BY CASE
        WHEN credit_score >= 300 AND credit_score < 500 THEN '300-499'
        WHEN credit_score >= 500 AND credit_score < 600 THEN '500-599'
        WHEN credit_score >= 600 AND credit_score < 700 THEN '600-699'
        WHEN credit_score >= 700 AND credit_score < 800 THEN '700-799'
        WHEN credit_score >= 800 THEN '800+'
    END
ORDER BY credit_score_range;
```
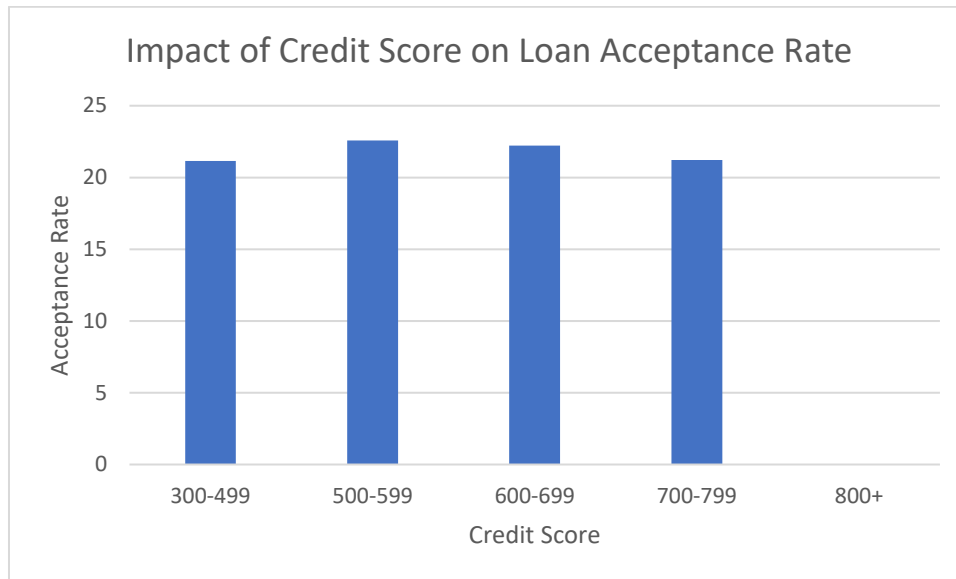
Figure 7 : Impact of Credit Score on Loan Acceptance Rate

Result:

credit_score_range	total_loans	accepted_loans	acceptance_rate
300-499	506	107	21.146245059288
500-599	10322	2330	22.573144739391
600-699	31298	6954	22.218672119624
700-799	2871	609	21.212121212121
800+	3	0	0.000000000000

Figure 8 : Impact of Credit Score on Loan Acceptance Rate Result

Illustration:

Interpretation:

From the results, we can see that the acceptance rate is basically the same for the different ranges of credit score.

2. Behavioural analysis and time trends
a. Speed of Repay Based on Borrowers Loans Amounts

Understanding the relationship between loan size and repayment speed is crucial for assessing the effectiveness of loan terms and repayment behaviour. If smaller loans tend to be repaid more quickly, this could indicate a higher level of borrower financial stability or a less risky profile. Conversely, if larger loans tend to take longer to repay, this may suggest a greater financial burden on borrowers, leading to slower repayments.

SQL query:

```
SELECT
    CASE
        WHEN loan_amnt < 5000 THEN 'Under 5k'
        WHEN loan_amnt >= 5000 AND loan_amnt < 10000 THEN '5k-10k'
        WHEN loan_amnt >= 10000 AND loan_amnt < 20000 THEN '10k-20k'
        WHEN loan_amnt >= 20000 THEN '20k and above'
    END AS loan_amount_range,
    AVG(cb_person_cred_hist_length) AS length_period
FROM loan_dataset
GROUP BY
    CASE
        WHEN loan_amnt < 5000 THEN 'Under 5k'
        WHEN loan_amnt >= 5000 AND loan_amnt < 10000 THEN '5k-10k'
        WHEN loan_amnt >= 10000 AND loan_amnt < 20000 THEN '10k-20k'
        WHEN loan_amnt >= 20000 THEN '20k and above'
    END
ORDER BY loan_amount_range;
```
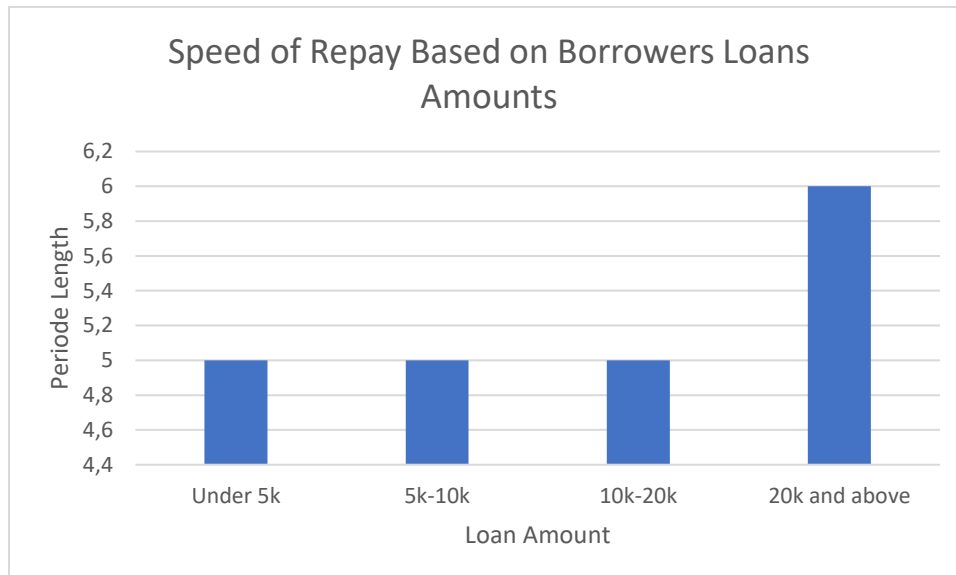
Result:

loan_amount_range	length_period
10k-20k	5
20k and above	6
5k-10k	5
Under 5k	5

Figure 10 : Speed of Repay Based on Borrowers Loans Amounts Result

Illustration:

Interpretation:

From the results, we can say that the loan amount doesn't really impact the duration of the payment as the durations means are practically the same for the different ranges.

b. Distribution of Loans by Purpose (Loan Intent) and Acceptance Rate

Understanding the distribution of loans based on the purpose of the loan and their corresponding acceptance rates is important for assessing how different loan purposes impact the likelihood of loan approval. Some loan purposes may have higher acceptance rates due to their perceived lower risk, while others, like personal loans, might be associated with higher risk and thus have lower approval rates.

SQL query:

```sql
SELECT
    loan_intent,
    COUNT(*) AS total_loans,
    COUNT(CASE WHEN loan_status = 1 THEN 1 ELSE 0 END) AS accepted_loans,
    (SUM(CASE WHEN loan_status = 1 THEN 1 ELSE 0 END) * 100.0) / COUNT(*) AS acceptance_rate
FROM loan_dataset
GROUP BY loan_intent
ORDER BY acceptance_rate DESC;
```

Figure 11 : Distribution of Loans by Purpose (Loan Intent) and Acceptance Rate

Result:

loan_intent	total_loans	accepted_loans	acceptance_rate
DEBTCONSOLIDATION	7145	7145	30.272918124562
MEDICAL	8548	8548	27.819372952737
HOMEIMPROVEMENT	4783	4783	26.301484424001
PERSONAL	7552	7552	20.140360169491
EDUCATION	9153	9153	16.956189227575
VENTURE	7819	7819	14.426397237498

Figure 12 : Distribution of Loans by Purpose (Loan Intent) and Acceptance Rate Result

Illustration:
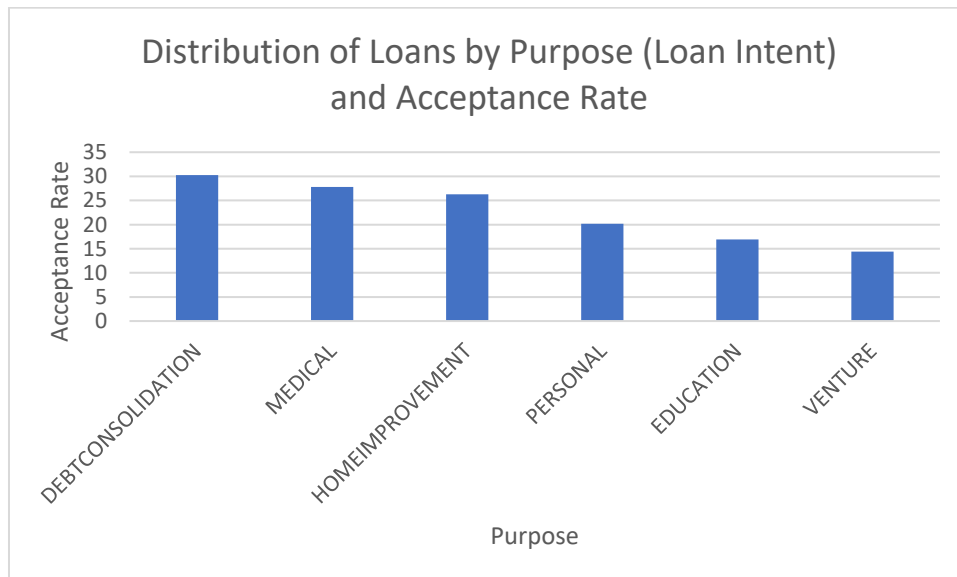
Interpretation:

From the analysis, we can conclude that loans for purposes such as **Debt consolidation** and **Medical** have the highest acceptance rates, while loans for more discretionary or risky purposes like **venture** tend to have lower acceptance rates.

CONCLUSION

Analysis of HSBC's loan acceptance criteria reveals key trends that influence lending decisions and help optimize risk management. The results suggest that HSBC could refine its risk assessment models by incorporating more nuanced criteria, particularly to better consider homeowners and borrowers with past defaults. A more balanced policy, combining preventative approaches and rehabilitation programs, would expand access to credit while maintaining a controlled level of risk.

Finally, adapting eligibility criteria based on the purpose of the loan and optimizing borrower scoring would give HSBC a better ability to target high-potential profiles, while strengthening the profitability and stability of its loan portfolio.