# Analytics Startup Plan

**Synopsis: *This document provides a high-level walkthrough of the activities required to guide the completion of the analysis.***

|  |  |
| --- | --- |
| **Project** | *Credit Card Default Prediction* |
| **Requestor** | *Centennial College* |
| **Date of Request** | *July 4th 2023* |
| **Target Quarter for Delivery** | *Quarter 3 2023* |
| **Epic Link(s)** | *Not applicable. There is no Agile group.* |
| **Business Impact** | *Safeguard Company’s Financial Assets.* |

**PREPARED BY:**

**SAMSON OGWUCHE: 301196569**

**CHARLES NTAMACK: 301209795**

## 1.0 Business Opportunity Brief

|  |  |
| --- | --- |
|  | Clearly articulated business statement of the Ask, opportunity, or problem you are trying to solve for. An important step is to understand the nature of the business, system or process and the desired problems to be addressed. This will be communicated back to All stakeholders for alignment. |

Credit card default risk has become a serious concern to lenders and the financial system at large. We intend to build a model that can help to predict credit card default by customers on their payments thereby posing credit risk to the financial system. The model will help banks and other lenders to identify risk factors in credit card customers and how to manage the risk.

**Business Case**

The total household debt in debt as of Q1 2023 stood at $2.32 trillion and the total number of Canadians with access to credit of about 30.6 million. The number of subprime consumers who are in the risk tiers increased by 8.3% and this constituted the largest growth. Total outstanding credit balances increased by 5.6% in Q1 2023. Credit card consumer balances grew by 11.36% in Q1 2023 representing the highest growth in a long time. However, prime consumers still constitute about three-quarters of the consumers indicating a relatively healthy distribution of credit consumers. Despite this, the delinquency rate on credit balances increased by 5.99% YoY in Q1 2023 and this therefore becomes imperative to make predictions to proactively identify consumers who are potential defaulters and manage the credit risk associated with such consumers.

**Opportunity**

The business case for this project is to prevent credit losses associated with default on credit card consumers by proactively identifying the risk and managing them. This project will provide lending institutions with the opportunity to protect their financial asset and have more liquidity for investment in better business opportunities that meet their business strategic objectives.

**Target Client**

Our target clients are financial institutions and other lenders such as auto leasing companies, mortgage lenders, grocery stores, etc.

Retail banks are credit institutions that finance the economy and this is extremely linked to its proper functioning. A retail bank is therefore an economic player whose main activities are:

* The provision of means of payment to customers;
* Collection and administration or investment (savings) of funds deposited by clients;
* The granting of loans.

As a result of these activities, Retail banking is subject to several types of risks (credit risks, operational risks, and market risks). For the purpose of this capstone project, we are going to focus on credit risks. Credit risk can be defined as the risk that the customer will no longer be able to meet the commitments he has made to the bank. This problem poses significant challenges for both cardholders and financial institutions, leading to adverse effects on personal finances, credit ratings, profitability, and economic stability. The primary issue contributing to credit card defaults is inadequate credit risk management by lenders. Addressing this problem requires implementing effective measures to improve risk assessment and mitigation strategies.

**The specific task:**

*Clearly articulate the specific task you will be conducting to help achieve the opportunity*

To achieve this opportunity of safeguarding the asset of financial institutions, we will

1. Collate data on credit card customers
2. Exploration of Data and Analysis (EDA)
3. We will do feat engineering on dataset
4. We will split the data into dependent and independent variables
5. We will train and fit the data
6. We will use different model algorithms to predict which set of customers is likely to default
7. Finally, we will interpret the result of our model and advise the lenders on the risk factors to look out for using the variables in our dataset.

## 1.1 Supporting Insights

|  |  |
| --- | --- |
|  | Define any supporting insights, trends and research findings. Where relevant, list key competitors in the market. What are their key messages, products & services? What is their share of market, nationally and regionally? |

**Canada’s Consumer Credit Balances**

As Canada’s economy continues to face difficult times, the number of individuals and corporations who apply to obtain and use credit cards across the country continues to rise. According to Transunion.ca, as of the end of Q1 2023, Canada’s total consumer debt climbed to $ 2.32 trillion as consumers lean on credit cards. The number of Canadians with access to credit stood at 30.6 million as of the end of Q1. In the same period, the number of credit consumers with balances in Q1 increased by 3.1% while outstanding credit balances increased by 5.6% year-on-year. The rise in total credit card debt comes with the problem of customers’ default on payment for the amount used thereby posing a credit risk to the financial industry and the economy at large.

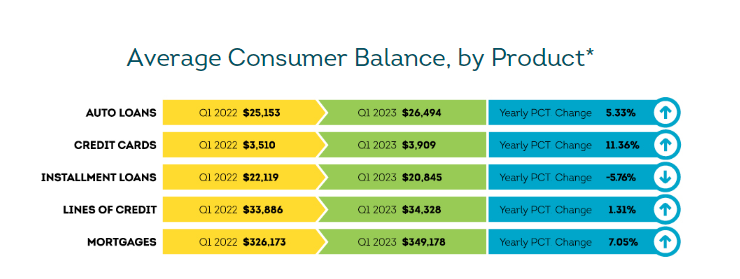
[](https://www.transunion.ca/lp/iir)

Figure 1 An Extract from Transunion.ca

**Growing Concerns**

The Bank of Canada and the financial industry are concerned about the rising household debt loads and the ability of households to meet their financial obligations. One of the key takeaways from the assessment of the annual Financial System Review 2023 is the risk of stability of Canada’s financial system and one of the factors that will affect the stability of the financial system is credit risk. Apart from cybersecurity attacks, Credit card default is one of the major risks to the financial system.

Several factors contribute to the recent rise in the volume of credit issuance and the rise in debt. One of the factors is the Bank of Canada’s (BC) hike in interest rates. The aggressive hike in interest rates by BC has continued to put many households under severe pressure making it harder for them to meet their mortgage and other debt payment obligations. If a sustained period of financial stress occurs, it would put further dwindling liquidity in the fixed-income markets. It could also add to the bank funding pressures and the cost of credit, further adding more difficulty to the ability of households to service their debt.

**Credit Card Debt in Canada**

Canadians added over $800 million to their credit card debt in December of 2022, thereby bringing the outstanding debt to a record high of $91.5 billion. Total credit card debt grew by about 22% YoY as of Q4 2022 compared to the same period in 2021. According to Equifax.ca, over 1.5 million new credit cards were issued in 2022 representing about a 22.5% increase compared to the previous quarter. Credit card spending was at an all-time high in Q3 2022 compared to the previous periods. Average monthly credit card spending of $2,447 in Q3 2022.

A survey by Statistics Canada in the Fall of 2022 revealed that 35% of Canadian household is struggling to meet their financial needs and this put pressure on their use of credit cards. At the same time, new credit consumers increased by 37.4% in 2022 compared to 2021. Given this backdrop, we have identified that higher interest rates have exposed weaknesses in the current business models of most financial institutions. Therefore, our project is aimed to predict credit card default by customers using other risk factors than the traditional credit score rating typically used by financial institutions to appraise credit card customers. This project will help banks and the financial services industry to be able to identify the risk factors in customers who apply for credit cards by being proactive in deciding which customers to approve for credit cards and whom to reject to avoid the credit risk of default and to manage the credit risk of their existing credit cardholders.

## 1.2 Project Gains

|  |  |
| --- | --- |
|  | *Describe any revenue gains, quality improvements, cost and time savings (as applicable). What will you do differently and why would our customers care. What are the implications if we do nothing? This section is particularly key for prioritization against company goals and KPI’s.* |

The gains of this project are to

1. Ensure a sound financial system
2. Guarantee stability in the financial system
3. Protect depositors’ funds
4. Guarantee investors’ confidence in the financial system
5. Save costs that are associated with debt recovery and provisioning for bad debt by financial institutions
6. Protection of cash flow.
7. Identifying opportunities and freeing up the company’s working capital for critical investment.

The following are revenue gains that effective management of credit card default brings to financial institutions

1. By effectively managing credit card defaults, financial institutions can reduce the number of delinquent accounts and increase the likelihood of timely repayments. This helps to maintain a steady stream of interest income from credit card balances, improving overall profitability.
2. Credit card defaults result in financial institutions needing to set aside provisions or reserves to cover potential losses. By addressing credit card defaults and minimizing default rates, financial institutions can reduce the need for significant loss provisions, thereby freeing up capital for other revenue-generating activities.
3. When individuals have positive credit card experiences with a low incidence of defaults, they are more likely to continue using their credit cards for purchases and transactions. This increased usage leads to higher transaction volumes, merchant fees, and interchange income for financial institutions.
4. Effectively addressing credit card defaults can help enhance customer satisfaction, trust, and loyalty. By providing proactive assistance, tailored repayment options, and effective default management solutions, financial institutions can retain their existing customers and attract new ones. Loyal customers are more likely to use additional banking services, such as loans, mortgages, or investments, generating additional revenue streams.
5. Proactively addressing credit card defaults can minimize the need for extensive collection efforts and associated costs. By implementing effective default prevention strategies and early intervention measures, financial institutions can reduce the expenses incurred in pursuing delinquent accounts through collections agencies or legal action.
6. Financial institutions that demonstrate a commitment to responsible lending and effective credit risk management earn a positive reputation among customers, regulators, and investors. A strong reputation can attract new customers, improve market share, and potentially lead to partnerships and collaborations that generate additional revenue opportunities.

Here are some potential quality improvements associated with addressing credit card defaults:

1. By implementing proactive measures, such as analytics in credit risk assessment, borrower education, and early intervention programs, financial institutions can reduce the incidence of credit card defaults. This leads to lower default rates, indicating improved credit quality and reduced financial distress for cardholders.
2. Addressing credit card defaults involves providing support and assistance to cardholders facing financial challenges. By offering financial education, debt counseling, and tailored repayment options, financial institutions can help individuals regain control of their finances and improve their overall financial well-being.
3. Addressing credit card defaults requires financial institutions to review and enhance their lending practices. This includes implementing robust credit risk assessment processes, ensuring appropriate credit limits, and promoting responsible borrowing behavior. These practices contribute to a healthier credit ecosystem and reduce the risk of customers obtaining credit beyond their means.
4. Effective default management strategies, such as responsive customer service, clear communication, and fair treatment of cardholders in financial distress, enhance the overall customer experience. By addressing credit card defaults promptly and providing supportive services, financial institutions can build stronger relationships with their customers.
5. Effective default management strategies, such as responsive customer service, clear communication, and fair treatment of cardholders in financial distress, enhance the overall customer experience. By addressing credit card defaults promptly and providing supportive services, financial institutions can build stronger relationships with their customers.
6. Addressing credit card defaults necessitates robust risk management practices. Financial institutions that implement comprehensive credit risk management frameworks, including effective monitoring systems, risk assessment models, and stress testing, strengthen their ability to identify, measure, and mitigate credit risk. This contributes to the overall quality of their risk management practices.
7. By demonstrating a commitment to addressing credit card defaults and assisting cardholders in financial distress, financial institutions can build trust and enhance their reputation. Trustworthy and reputable institutions are more likely to attract new customers, retain existing ones, and establish long-term relationships with stakeholders.

## *Note: Completion of the following sections is possible only after a careful assessment and triage of the Ask. This is required to determine scope, resource, time, priority and data availability.*

## 2.0 Analytics Objective

|  |  |
| --- | --- |
|  | List the key questions, assumptions and define the hypotheses. Often the deliverable may not just be an analysis output, however a recommended operating model or blueprint for a pilot etc.  Note: Asking the right questions and truly understanding the problem will lead to the right data, right mathematics, and right techniques to be employed. |

This project's main objective is to build a model to predict customers with the likelihood of default thereby helping lenders to minimize financial losses resulting from credit card defaults. Financial institutions aim to reduce the amount of unpaid credit card debt and associated interest charges, which can have a significant impact on their profitability and financial stability.

The second objective of this project is to preserve the quality of the credit card portfolio. By proactively managing defaults, financial institutions can mitigate the risk of deteriorating asset quality and maintain a healthier balance sheet.

The third objective of our project is to enhance customer relationships through a quality selection of customers for credit cards and manage them effectively. Our model will help financial institutions to work closely with cardholders, providing support, education, and tailored repayment options. By demonstrating empathy and flexibility, financial institutions can enhance customer loyalty and retention. This objective aligns with broader financial inclusion and responsible lending initiatives.

By implementing sound credit risk management practices and managing defaults effectively, institutions contribute to the stability and resilience of the overall financial ecosystem.

Key Questions:

1- How can we improve our credit risk through the use of analytics to identify potential defaulters?

2- Are there opportunities to leverage technology and data analytics to improve the early detection of potential default risks and implement proactive measures to minimize credit losses?

## 2.1 Other related questions and Assumptions:

|  |  |
| --- | --- |
|  | *List any assumptions that may affect the analysis* |

It is assumed that the necessary data related to credit card defaults, customer behavior, and financial performance is available for analysis and decision-making. Sufficient and accurate data is crucial for effective risk assessment, identification of default risks, and evaluation of the impact of strategies used.

It is assumed that the financial institution has the necessary resources, including financial, technological, and human resources, to effectively address credit card defaults. A sufficient budget, staffing, and infrastructure are essential for implementing strategies and initiatives to achieve the objectives.

It is assumed that financial institutions have access to appropriate technological tools and systems to support credit risk assessment, data analysis, reporting, and customer communication. Adequate technological capabilities enable efficient and accurate implementation of strategies and initiatives.

It is assumed that the financial institution is committed to continuous learning, adaptation, and improvement. This includes actively monitoring industry trends, best practices, and regulatory updates to refine strategies and ensure they remain effective over time.

## 2.2 Success measures/metrics

|  |  |
| --- | --- |
|  | *What does success look like? Define the key performance indicators (success definition/indicators, drivers and key metrics) against which the objectives will be analyzed. These should be drawn from the interlock meeting with key stakeholders and will inform the approach and methodology for the analysis.* |
|  |  |
|  |  |

The metrics that will be used to define the success of this analytics project will be geared toward achieving the objective of minimizing the financial losses arising from credit card default and improving collection rates. The following will be the primary metrics for measuring the success of the project upon implementation:

* Credit card default (delinquency) rate
* Credit card utilization and collection rate

## 2.3 Methodology and Approach

|  |  |
| --- | --- |
|  | *Now that you have a good understanding of the Ask and deliverable, detail the recommended approach/methodology.* |

**Type of Analysis:**

Based on our objective, the approach we will adopt in this project is to use predictive models such as Random Forest, Decision Tree, Logistic Regression, Neural Networks, XG Boosting, and Machine models such as Variational Auto-Encoder (VAE) to predict credit card default on the available dataset.

**Methodology:**

The methodologies we will on this project are as outlined below:

* Identification of project/analytics approach
* Identification of our most important variables from the dataset
* Prepare the foundation for subsequent steps in data modeling and analysis of the model
* Delivering Sprint 1 by outlining the projective, project, and project vision.
* Evaluation of the model performance
* Peer quality review
* Model update/improvement after peer review
* Delivery of Sprint 2
* Project implementation
* Project approval.

**Output: *We will include the output once we start the EDA and modeling on the dataset.***

## 3.0 Population, Variable Selection, considerations

|  |  |
| --- | --- |
|  | Capture learning about the data available today location, structure, and reliability; this would include data in operational systems including dealer sourced, data warehouse and any CRM or email marketing systems available today. |

**Audience/population selection: Banks and other lenders**

**Observation window: Historical data on credit card default**

**Inclusions: At the moment, we are yet to agree on inclusions. We are using all the variables in the dataset for now.**

**Exclusions: At the moment, we are yet to agree on exclusions. We are using all the variables in the dataset for now but will decide later on exclusions.**

**Data Sources:** Kaggle

**Audience Level:** Professional

**Variable Selection: Credit card default prediction.csv**

**Variables: Limit\_Bal, Sex, Education, Age, Pay\_0, Pay\_2, Pay\_3, Pay\_4, Pay\_5, Pay\_6, Bill\_Amt1, Bill\_Amt2, Bill\_Amt3, Bill\_Amt4, Bill\_Amt5, Bill\_Amt6, Pay\_Amt1, Pay\_Amt2, Pay\_Amt3, Pay\_Amt4, Pay\_Amt5, Pay\_Amt6, Default.**

**Derived Variables: Charts are derived from the dataset and displayed on the model.**

**Assumptions and data limitations:**

**Assumptions:** Our assumptions have been listed under section 2.1.

**Limitations:**No real-time data, we only have historical data.

## 4.0 Dependencies and Risks

|  |  |
| --- | --- |
|  | Identification of key factors that may influence the outcome of the project and the likelihood of it happening: |

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Likelihood (based on historical data)** | **Delay (based on historical data)** | **Impact** |
| *Data collection* | *Low* |  | *We are relying upon historical/third-party data not collated by ourselves.* |
| *Unforeseen events* | *High* |  | *Events such as loss of job, sickness, and other social issues such as marital challenges could influence customers’ ability to pay credit card debt. These are not possible to predict.* |
| *Geographical relocation* | *Medium* |  | *The geographical location of credit cardholders may influence default rates due to the country’s foreign transfer regulations, and economic conditions.* |
| *Credit Utilization ratio* | *High* |  | *High utilization may lead to higher defaults.* |
| *Payment history* | *Medium* |  | *We have only 6 months of payment history of cardholders in the dataset.* |
| *Credit history/credit score* | *High* |  | *No credit history/score of the cardholders.* |
| *Level of financial literacy of cardholders* | *Low* |  | *The level of financial literacy of cardholders can influence the default rate.* |

## 5.0 Deliverable Timelines

|  |  |
| --- | --- |
|  | List key dates and timelines as a work-back schedule. Activate line items based on complexity and line-of-sight required. Will set the stakeholder expectations for the process. |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Major Events / Milestones** | **Description** | **Scope** | **Days** | **Date** |
| 1. | Kick-off / Formal Request | *Project Proposal and agreement on the Capstone topic.* | *Sourcing for the dataset.* | *1* | *4th July 2023.* |
| 2. | Assessment / Triage | *Identification of project scope, assessment of the issues, requirements, and the quality of the dataset shall be conducted.* | Identifying project approach | 3 | *7th July 2023.* |
| 3. | Prioritization | Identification of the target and the predictors (dependent and independent variables). | Identifying the important variables in the model | 2 | *9th July 2023* |
| 4. | Data Exploration & Analysis   * Issues with duplicates * Issues with Spend data | Importation of libraries and Preprocessing of the dataset (to check for missing values, imputation of missing values, variable identification, duplication in the data, outliers, skewness, and special characters, selecting important variables, dropping correlated variables, standardizing and normalization of the data) | Provide a foundation for subsequent modeling steps and data analysis. | 15 | *24th July 2023* |
| 5. | Story Board 1 | To describe the business opportunity, project gains, and project vision aimed at minimizing financial losses arising from credit card default. | *Delivery of Sprint 1.* | *1* | *25th July 2023* |
| 6. | QA Output | This involves the process of testing the data model to check if matches our project criteria. | *Performance of quality evaluation.* | *13* | *7th August 2023* |
| 7. | Internal team Presentation | This involves peer reviews with other classmates on the modeling processes and feedback for corrections. | *Peer quality review of modeling/presentation processes.* | *5* | *12th August 2023* |
| 8. | Go/No Go | This is the outcome of the peer review with other classmates. | *Making updates to the modeling taking into consideration suggestions and observations by classmates.* | *2* | *14th August 2023* |
| 9. | Story Board 2 | Classified analytics objectives, supporting insights, success metrics, methodologies, and approaches. | *Delivery Sprint 2.* | 1 | *15th August 2023* |
| 10. | Pilot | Testing feasibility and risk of the full-scale project. | *Project Implementation* | *2* | *17th August 2023* |
| 11. | Delivery & sign-off | Presentation of the project to the Advisor and his approval. | *Final project approval* | *1* | *18th August 2023.* |