LAMBTON COLLEGE

A Project on

[Data Visualization in financial Institution]

121 Brunel Rd, Mississauga

ON L4Z 3E9

A project on Data Visualization of loan and credit data in financial institution.

Big Data Analytics DSMM

Under the supervision
of
Professor Salim Sattar

Submitted BY:
Aadarsha Chapagain (C0825975)
Rajasekhar katta (C0833766)
Rishi Phaneendra Varma (C0825285)

Venkata Sai Manikanta (C0833772)

Submitted To:
Lambton College
Professor Salim Sattar

Submission Date:
17th December 2022

Table of Contents

1. Project Title and Description

In recent years it is seen that a lot of banks and other financial institutions are struggling to get insights from their data. For most companies, data is stored in excel files daily and they store huge amounts of data. These data sit either sit on the company's on-premises servers or the cloud and they has just been accumulated over time. Since the data is not organized and managed properly, querying these data and getting any useful information out of them has become a close to impossible task. Employees have been using excel for some basic visualizations but as the amount of data is growing daily excel is not able to cope up with the volume of the data. This problem demands a solution where these daily growing data can be handled and visualized properly so that insight can be driven from them. When visualized the pattern or the trend of the data can be analyzed.

Dataverse is a data visualization project designed to help financial institutions better understand and optimize their Credit and Sales departments. To achieve this, Dataverse collects and transforms data from various sources within these departments, using ETL (Extract, Transform, Load) tools to ensure that the data is clean and consistent. This transformed data is then stored in a data warehouse, where it can be easily accessed and queried by visualization tools.

One such visualization tool is Power BI, a popular business analytics platform that allows users to create interactive dashboards and reports. Using Power BI, Dataverse can generate a wide range of graphs, plots, and charts to visualize the trends and patterns within the Credit and Sales data. This includes key metrics such as revenue, loan volume, customer satisfaction, and more.

Dataverse is designed to be a flexible and scalable solution that can adapt to the specific needs of each financial institution. It can be configured to collect and visualize data from a wide range of sources, including transactional systems, CRM systems, and other data sources relevant to the Credit and Sales departments. By providing a clear and concise overview of the data, Dataverse helps financial institutions make informed decisions and optimize their operations.

2. Project Manager Assigned and Authority Level

Rajasekhar Katta will be the project manager for this project and will have authority to select team members and determine the final project budget and schedule.

3. Business case

We now live in a world where we can have activities that are out of our price range even before we are able to purchase them. This is made possible by credit cards and banks. Banks, who offer the financial liquidity that enables us to take advantage of the experience, as well as credit card firms have earned enormous profits from issuing credit cards.

A credit card processor's data can be used to pinpoint individual consumer and commercial spending patterns. Customer accounts can be modified in light of this information. Developing marketing campaigns to directly address behaviors grows revenue. These considerations should result in greater sales.

This project targets banking customers when and how they are likely to be most amenable to making a purchase. This provides information to help optimize banking marketing efforts. This knowledge can involve anything. Client may design a new product or service and know when and where to market them. Additionally, if a client experiencing a slowdown in sales, the data will provide insights into what is wrong. This can also relate to the customer experience better.

This project primarily focuses on two reporting abilities.

1. **Individual Cardholder History Review**
2. **Increase consumer spending**

Individual Cardholder History Review

Client can go down more deeply into each customer's history while simultaneously reviewing and analyzing all data on the dashboard. This type of analysis is crucial when deciding whether to approve the customer's credit application because it offers information about the person's spending patterns, payment history, and income. You and your business can see who is meeting payment deadlines and who is most likely to default.

The Dataverse dashboard provides a wealth of information about each customer's credit history and financial behavior. This can be extremely valuable when evaluating a credit application, as it helps you understand the customer's patterns of spending, income, and payment history. By drilling down into each customer's data, you can gain a more detailed and nuanced understanding of their financial situation.

In addition to providing a detailed analysis of individual customers, the Dataverse dashboard also allows you to review and analyze data from all customers in a single view. This makes it easy to identify trends and patterns across your entire customer base, and to compare the financial behavior of different groups of customers.

By combining these two types of analysis, you can make more informed and accurate decisions about credit approval. You can see which customers are meeting payment deadlines and which are at risk of default, and use this information to assess the risk of approving a particular credit application. Overall, the Dataverse dashboard provides a powerful tool for managing credit risk and optimizing your business operations.

Increase consumer spending

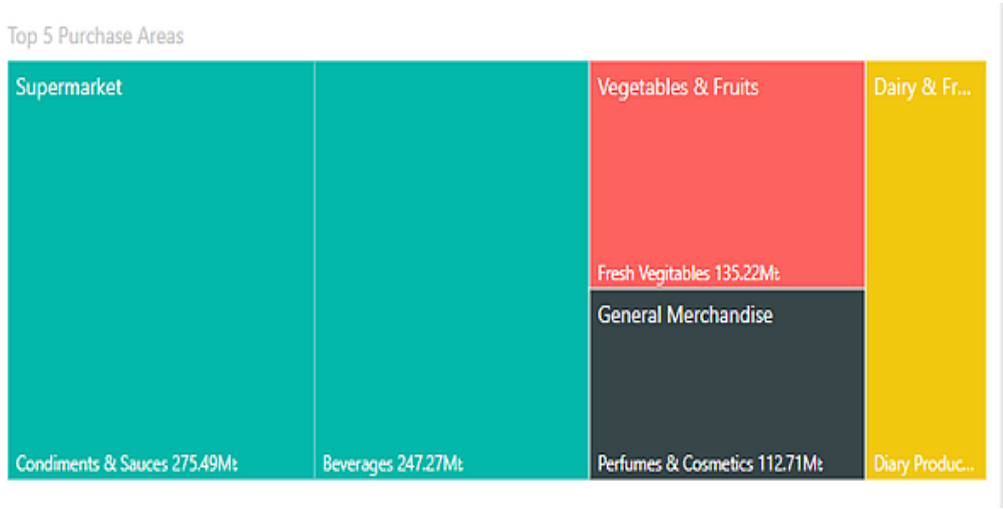
Client or Banking credit card business will be able to determine which categories the cardholder is spending money in by looking at each individual transaction. Do they buy more booze now? even food? or journeys? Insights from this project will be able to target them more effectively with adverts encouraging spending in these areas thanks to that. Client business will also be able to partner with these suppliers to promote even more spending because it will be able to observe which merchants your customers frequently use.

The Dataverse project can provide valuable insights into the spending habits of credit cardholders, helping the client or banking business to better understand their customers' needs and preferences. By analyzing each individual transaction, the business can determine which categories the cardholder is spending money in, such as food, travel, or leisure. This information can be used to target customers with more relevant and effective advertising, encouraging them to spend more in these areas.

In addition to targeting customers with ads, the Dataverse project can also help the client business to partner with merchants that are frequently used by their customers. By observing which merchants are most popular with cardholders, the business can identify opportunities to promote more spending through strategic partnerships or promotions. This can help to drive revenue and boost customer loyalty.

Overall, the Dataverse project provides a powerful tool for understanding and optimizing customer spending patterns, helping the client business to make more informed and strategic decisions about how to engage with their customers and drive business growth.

Visualization at a glance



4. Resources Preassigned

Rajasekhar katta will be the Project Manager and data modeler.

This project leverages Microsoft Azure Public Cloud services as Pay-as-you-Go subscription to avoid capital investments and operational maintenance.

Rishi Varma will be appointed as Data Engineer role to integrate data from various sources, data cleansing, data governance and develop robust automated data pipelines.

Aadarsha Chapagain will be taking care of code, continuous integration and continuous development (CI/CD) pipelines along with the maintaining of Agile frame work and Azure Devops.

Venkata Sai Manikanta will be responsible for the Power BI dashboard development and deployment of Power BI services to the could environment.

5. Key Stakeholder List

Customers

Project Manager - Rajasekhar katta

Project Team Members

- Aadarsha Chapagain
- Rajasekhar katta
- Rishi Varma
- Venkata Sai Manikanta

Company Executives

- Professor Salim Sattar

Project Sponsor

- Professor Salim Sattar

Steering Committee

- Professor Salim Sattar
- Aadarsha Chapagain
- Rajasekhar katta
- Rishi Varma
- Venkata Sai Manikanta

6. Stakeholder Requirements as known

Business unit: **Enterprise Business Intelligence (EBI)**

This unit will be completely focusing on analyzing demographics of Credit card consumers. Consumer demographics are observable and measurable characteristics that include credit card usage, Timely payments, Employment status, Income, Expenses, Geographical location and Age.

The focus of this unit is to thoroughly analyze the demographics of credit card consumers in order to gain a deeper understanding of their characteristics and behaviors. Demographic data can provide valuable insights into how consumers use credit cards, including how often they use them, how they make timely payments, and what types of expenses they incur.

To achieve this, the unit will collect and analyze data on a wide range of demographic characteristics, including employment status, income, expenses, geographical location, and age. By analyzing this data, the unit will be able to identify patterns and trends in credit card usage, and to identify opportunities to optimize the credit card offering for different segments of consumers.

For example, the unit may discover that certain groups of consumers are more likely to use their credit cards for travel expenses, while others tend to use them for everyday purchases. This information could be used to tailor marketing campaigns or to offer targeted promotions that align with the specific needs and preferences of different groups of consumers.

Overall, the goal of this unit is to provide a comprehensive and detailed analysis of consumer demographics, helping the organization to better understand and engage with credit card customers.

Business Operations

Core Business processes include Information technology, Service delivery, Marketing Operation, Sales Operation and Customer Service. Operations techniques and considerations are,

IT Infrastructure operations – Service desk, Asset Management, Physical Security, DevOps

User Interfaces – Information Visualization Using PowerBI

Maintainability – Service Management

Optimization – Business Process Improvement, Quality Control, Customer Satisfaction

Production – Visualize useful insights on PowerBI Dashboards

Production Management – Scheduling, Budget, Product Development, Compliance, Quality Assurance

Customers

Lead Users Aadarsha Chapagain, Rajasekhar katta, Rishi Varma, Venkata Sai Manikanta may contribute user stories or ideas for usability and quality

Subject Matter Experts

Aadarsha Chapagain, Rajasekhar katta, Rishi Varma, Venkata Sai Manikanta will give inputs in their respective areas to improve product quality and efficiency throughout the project.

7. High Level Product Description/Key Deliverables

This project aims to help the clients by reaching the banking customers and shows when they are most likely to be open to making a purchase. By creating the PowerBI dashboard with all the required data the

client can access lot of information about their customers spending patterns in a simple way. This will provide the information about the customers purchase history, their spending limits and many more.

The Dataverse project aims to provide valuable insights into the spending patterns and purchase behavior of banking customers, helping the client to understand when they are most likely to be open to making a purchase. By creating a comprehensive PowerBI dashboard with all the relevant data, the client can access a wealth of information about their customers in a simple and convenient way.

The PowerBI dashboard can provide information about customers' purchase history, including the types of products and services they tend to buy, the frequency of their purchases, and their average spend. It can also provide data on customers' spending limits, allowing the client to better understand the financial resources and constraints of each customer.

In addition to providing this detailed analysis of individual customers, the PowerBI dashboard can also enable the client to analyze trends and patterns across the entire customer base. This can help to identify opportunities to target specific groups of customers with specific marketing campaigns or promotions, or to optimize the pricing and positioning of different products and services.

Overall, the Dataverse project can help the client to better understand and engage with their banking customers, driving business growth and customer loyalty in the process.

The Key deliverables for this project are:

- Customers spending patterns
- Big purchases in last 6 months
- How many times they exceed the credit limit in past 2 years
- Repaying on time or not
- Their top 5 spending product categories
- All these insights will be presented on PowerBI Dashboard

The dashboard could include the following elements:

1. A pie chart that shows customers' spending patterns, including the proportions of their spending on different types of products or services.
2. A bar chart that shows the number of big purchases (i.e., purchases over a certain dollar amount) made by customers in the last 6 months.
3. A line chart that shows the number of times customers exceeded their credit limit in the past 2 years.
4. A stacked bar chart that shows the percentage of customers who made timely repayments versus those who did not.
5. A donut chart that shows the top 5 product categories in which customers spend the most money.

By including a range of different charts and graphs, the PowerBI dashboard can provide a comprehensive and visually appealing overview of the data, making it easy for users to understand and analyze the information.

8. High Level Product Description/Key Deliverables

1. Comprehensive analysis of the given dataset, including identification of any trends, patterns, and relationships within the data.
2. Clear and concise visualizations of the data, including charts, graphs, and maps, to aid in the understanding and interpretation of the findings.
3. Detailed report outlining the key findings and recommendations based on the analysis.
4. Presentation of the results to relevant stakeholders, including a discussion of the implications and potential next steps.
5. Support for the implementation of any recommendations, including assistance with the development of any necessary tools or processes.

9. High Level Assumptions

1. The data sources being used are reliable and accurate.
2. The data sources are structured in a way that allows for easy extraction and transformation.
3. The data sources are available and accessible at the time of the ETL process.
4. The target database or data warehouse is properly configured and able to receive the loaded data.
5. The necessary permissions and access are in place to extract and load the data.
6. The data transformation rules and logic have been properly defined and tested.

It's important to clearly define and document any assumptions made during the data analysis process, as they can impact the accuracy and reliability of the findings.

10. High Level Constraints

1. Time constraints: There may be a limited amount of time available for the data analysis, which could impact the scope and depth of the analysis.
2. Data availability: The data sources and volume of data available for analysis may be limited, which could impact the conclusions that can be drawn.
3. Data quality: The quality and reliability of the data being analyzed can impact the accuracy of the findings.
4. Technical skills and resources: The technical skills and resources available to the data analysis team may impact the tools and techniques that can be used.
5. Budget constraints: There may be limitations on the budget available for the data analysis project, which could impact the resources and tools that can be used.
6. Ethical considerations: There may be ethical considerations to take into account when working with certain types of data or in certain industries.

It's important to be aware of these constraints and to carefully consider their impact on the data analysis

11.Measurable Project Objectives

It's important to set clear and measurable objectives for a data analysis project in order to ensure that the project is focused and valuable to the organization.

1. Identify trends and patterns in the data that can inform business decisions.
2. Determine the relationships between different variables in the data.
3. Improve the efficiency and effectiveness of business processes by identifying bottlenecks and inefficiencies through data analysis.
4. Provide insights into customer behavior and preferences through analysis of customer data.
5. Identify opportunities for cost savings or revenue generation through data analysis.
6. Monitor and track key performance indicators (KPIs) to understand the performance of the business.
7. Develop predictive models to forecast future outcomes or trends.

12.Project Approval Requirements

It's important to ensure that all necessary approvals are obtained before starting a data analysis project, as this will help to ensure that the project is properly supported and has the necessary resources to be successful.

1. A detailed project proposal outlining the goals, objectives, deliverables, and resources required for the project.
2. A clear and concise project plan, including a timeline and budget.
3. Approval from the relevant stakeholders, including any necessary approvals from upper management or the IT department.
4. A signed data processing agreement or other legal documents as required.
5. Any necessary approvals from regulatory bodies or industry associations.

13.Overall Project Risk

Followings are the risk involved for this project.

Data quality issues: The quality of the data being analyzed can impact the accuracy and reliability of the findings.

Data security risks: There may be risks associated with handling sensitive or personal data, including the risk of data breaches or unauthorized access.

Lack of stakeholder buy-in: If the relevant stakeholders do not support the project or do not understand its value, it may be difficult to achieve the desired outcomes.

Limited resources: There may be constraints on the budget or personnel available for the project, which could impact the scope and depth of the analysis.

Technical issues: There may be technical challenges associated with working with large or complex datasets, or with using certain tools or techniques.

Time constraints: There may be a limited amount of time available for the project, which could impact the scope and depth of the analysis.

14. Project Exit Criteria

There need to be clear exit criteria in place for a data analysis project in order to ensure that the project has been successfully completed and that the necessary actions have been taken based on the findings

1. The project deliverables, including any reports or presentations, have been completed and approved.
2. The project goals and objectives have been achieved.
3. All necessary approvals have been obtained.
4. Any necessary documentation, including a final report or presentation, has been completed and shared with relevant stakeholders.
5. The project team has received necessary training and support to continue working with the data and tools used in the project.
6. Any recommendations or action items resulting from the data analysis have been implemented or are in the process of being implemented.
7. The project team has conducted a thorough review of the project and identified any lessons learned.

15.References

https://learn.microsoft.com/en-us/power-bi/fundamentals/service-get-started

https://learn.microsoft.com/en-us/power-bi/create-reports/service-dashboard-create

https://truelayer.com/blog/data-chain-retrieving/

https://docsumo.com/blog/best-bank-data-extraction-software

https://www.inetco.com/use-cases/credit-and-debit-card-analytics/