Data Management and Data Analytics Capstone Topic Approval Form

Capstone Topic Approval Form

The purpose of this document is to help you clearly explain your capstone topic, project scope, and timeline. Identify each of the following areas so you will have a complete and realistic overview of your project. Your course instructor cannot approve your project topic without this information.

Student Name: [redacted]

Student ID: [redacted]

Capstone Project Name: Analysis of JGA Bank Credit Card Customer Data and Churn Characteristics

Project Topic: This project will assess profile data of both current and former credit card holders to help identify trends that may indicate which customers are at risk of churn.

Research Question: Are customers of a certain age group, income level or educational background more likely to churn?

> **Hypothesis:** Customers of an older age group, higher income and educational level will be more likely to churn.

Context: JGA Bank (fictional), is a financial institution that provides banking and lending solutions to over 25,000 customers in its local community. JGA and its associates strive to provide their customers with a wide variety of products that both meet their customer's needs and provides additional revenue streams to fuel the banks growth. Of the services offered, JGA Bank launched its first credit card five years ago and has since seen much success with its initial card offering. Since its inception, JGA's credit card has been adopted by nearly half of its current customers. Due to this success, the bank is looking to increase its credit card portfolio with a new credit card offering that will allow JGA to expand its customer base and further grow revenue. Since launching its first credit card, the bank has built a database containing characteristics of each customer, purchase patterns and credit card usage. As JGA is looking to develop its new credit card program, the bank seeks to gain insights from its existing database of card holders to help better build their new product. JGA Bank is particularly interested in leveraging its existing customer dataset to identify characteristics of customers that have churned since acquiring its credit card. While overall, JGA has been very successful in retaining its credit card customer base, it is particularly concerned about the potential impact churn may have on its credit card revenue as it seeks to expand its customer base with the new credit card product. The bank believes that utilizing its current data to understand its existing and former customers, will allow it to better build its new product and reduce the risk that customer churn poses to its revenue.

Data: I will need to collect credit card customer data, containing customer attributes and usage measures.

I will be collecting the dataset Credit Card Customers from Kaggle.com.

- Credit Card Customers – Kaggle (https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers)

The dataset is made publicly available on Kaggle for use of data analytics projects. I will not be utilizing private or confidential information for the purposes of this analysis.

Data Gathering: I will download a tabular dataset from Kaggle.com and will clean the data as needed for this analysis.

Data Analytics Tools and Techniques: I will use descriptive analytics to evaluate the attributes contained in the dataset and derive necessary values applicable to the objective of this project. An example of this will be to create age, income and education brackets that will better allow me to generalize how a range of these attributes correlates to churn status. I will then determine what percentage of total customers are in each age, income, and educational background and what percentage each bracket represents for active and churned customers. I will then perform a linear regression to review any possible correlations between age, income, education, and customer status, active or churned. I will use both Python to create custom code to cleanse, manipulate and analyze the dataset.

Justification of Tools/Techniques: Due to the number of customers represented in the data of varying backgrounds, I will need to be able to apply filters to the data to explore and summarize how each attribute being observed is represented within the data. I will be able to accomplish using descriptive analytic techniques. I will be utilizing multivariate analysis to identify if there are any correlations between each customer attribute and churn. I will create multiple scatter plots to visualize potential correlations, as well as linear regression further explore correlation. Python provides an open-source language that will allow me to perform these analyses.

Application Type, if applicable (select one):

☐ Mobile
☐ Web
☒ Stand-alone

Programming/Development Language(s), if applicable: Python

Operating System(s)/Platform(s), if applicable: N/A

Database Management System, if applicable: N/A

Project Outcomes: The outcome of this project will be the production of a clean, summarized dataset based upon the Credit Card Customers data retrieved from Kaggle. The resulting dataset will include a summary of any correlations found between customer age, income level, educational background, and churn trends. The product of this project will include all applicable code created to clean the data and produce the summarized results of the analysis.

Projected Project End Date: 10/31/2022

Sources:

- Churn Rate (https://www.investopedia.com/terms/c/churnrate.asp)
- Why Retaining Customers for Banks Is As Important As Winning New Ones (https://www.forbes.com/sites/forbestechcouncil/2020/05/27/why-retaining-customers-for-banks-is-as-important-as-winning-new-ones/?sh=5e571a203f98)
- Why Churn Analysis is the Key to Improving Customer Retention (https://amplitude.com/blog/why-churn-analysis-is-the-key-to-improving-customer-retention)

Human Subjects or Proprietary Information
Does your project involve the potential use of human subjects? (Y/N): N
Does your project involve the potential use of proprietary company information? (Y/N): N

STUDENT SIGNATURE

[redacted]

By signing and submitting this form, you acknowledge that any cost associated with the development and execution of your data analytics solution will be your (the student) responsibility.

TO BE FILLED BY A COURSE INSTRUCTOR

The capstone topic is approved by a course instructor.

COURSE INSTRUCTOR'S NAME AND SIGNATURE: William L. Dean Jr., Ph.D.

COURSE INSTRUCTOR APPROVAL DATE: 16 October 2022

Project Compliance with IRB (Y/N):Y