This is a basic structure for a data science project summary report. You can modify and expand this template to suit your project's specific needs. Remember to clearly and concisely communicate your work, using visual aids where appropriate, and interpret your findings for both technical and non-technical audiences.

Understanding the process before and after the policy change will be crucial in your analysis.

This will be your final deliverable. Keep these requirements in mind as you work through the task.

Familiarize yourself with these data elements. This will aid in understanding the structure of the data and how to handle it in your analysis.

the relationship between the data in the two files. Use this information to guide your data handling and analysis process.

Prior the policy change, filing a claim with FIS was a required step before people could be reimbursed for theft, so we would expect a closer relationship between the volume of claim activity and reimbursement activity in the pre-policy period. This provides a key understanding of how the policy change might have affected the relationship between claims and reimbursements. It will be crucial to understand and account for this when analyzing the effect of the policy change on claims and reimbursements.

Additionally, not all claims to FIS would have a corresponding theft reimbursement. This part explains why not all claims might result in reimbursements, which is another factor to consider in your analysis.This highlights the complexity of the relationship between claims and reimbursements, and should be taken into consideration when interpreting your results.

The goal of this exercise is not only to perform the tasks, but also to demonstrate your thought process, approach, and ability to interpret results. Remember to document your steps, assumptions, and decisions as you work through the task.

\*\*\* Your analysis should focus on the extent to which the volume and value of claims and theft reimbursements change following the policy change on January 27th – looking at the connections between cases/individuals in both data sets is not required but could complement your findings. This part restates the main objective of the analysis and provides a suggestion for a potential supplementary analysis. The focus should be on analyzing how the policy change affected the volume and value of claims and theft reimbursements. However, exploring connections between cases/individuals in both datasets could provide additional insights.

Remember, the focus of the analysis is on the extent to which the volume and value of claims and theft reimbursements changed following the policy change, not necessarily on the connections between cases/individuals in both data sets.

Google Colab

Estimate the change in administrative theft claims and reimbursement payments

Author: Steven Archuleta

Date: 19 May 2023

Table of Contents

Executive Summary

Introduction

Methodology

Results

Data Cleaning and Preprocessing

Exploratory Data Analysis

Machine Learning Model

Discussion

Conclusion and Recommendations

References

1. Executive Summary

This report presents the process and findings of our predictive analysis project, where we aimed to predict the likelihood of a hypothetical event (replace with your actual event) using a binary classification model. Our methodology involved data cleaning and preprocessing, exploratory data analysis, and building a machine learning model. Our key findings indicate that (insert key findings). Based on these results, we recommend (insert recommendations).

2. Introduction

In this project, we sought to predict (insert what you're predicting). This problem is significant because (insert significance of the problem). Our objective was to build a model that accurately predicts the likelihood of (insert event) based on historical data.

3. Methodology

Our methodology involved three main steps:

Data Cleaning and Preprocessing: We cleaned and preprocessed our dataset to prepare it for analysis and modeling.

Exploratory Data Analysis (EDA): We performed EDA to gain insights from our data and inform our modeling approach.

Machine Learning Model Building: We built a binary classification model to predict the likelihood of (insert event).

4. Results

4.1 Data Cleaning and Preprocessing

We identified and handled missing values, outliers, and other inconsistencies in our data. We also preprocessed our data by (insert preprocessing steps, e.g., encoding categorical variables, scaling numerical variables, etc.).

4.2 Exploratory Data Analysis

Our EDA revealed several interesting findings:

Finding 1

Finding 2

Finding 3

(Insert charts, graphs, etc. to illustrate these findings)

4.3 Machine Learning Model

We chose a binary classification model for our analysis. Our model achieved an accuracy of (insert accuracy) and a precision of (insert precision).

(Insert any charts, confusion matrices, ROC curves, etc. that illustrate your model's performance)

5. Discussion

Our findings suggest that (interpret your findings). The model's performance indicates that (interpret your model's performance). These results suggest (interpret the implications of your results).

6. Conclusion and Recommendations

In conclusion, our analysis and model provide valuable insights into (insert problem). Based on our findings, we recommend (insert recommendations). Future work could explore (insert potential future work).

7. References

Reference 1

Reference 2

Reference 3

This is a basic structure for a data science project summary report. You can modify and expand this template to suit your project's specific needs. Remember to clearly and concisely communicate your work, using visual aids where appropriate, and interpret your findings for both technical and non-technical audiences.