Independent Project Week 3: Data Report

Peter Kiragu

Moringa School

Jedidah Achieng

**Business Understanding**

A telecommunication company like MTN needs to have a clear strategy for expanding its infrastructure. This is because it needs to ensure that its infrastructure is distributed properly across the different cities to optimize the number of users who have access to their services.

MNT needs to rank the cities based on revenue. This is because the goal is to understand the cities that bring in more revenue and expand infrastructure in these areas. This is because cities with most billing prices are going to make a return on investment faster. Ranking cities by billing price also helps the business understand locations where its network is being used a lot. These are the locations where the network is likely to be jammed because of many users. As a business, MTN needs to prioritize, and analyzing the data on its users in various cities is very important.

**Data Understanding**

The data available is in four datasets. One dataset has the geo-location of various cell towers. The information in this data set has the name of the cities, the cell tower information, site information, and the exact location in terms of latitude and longitude. The data set also has information on the status of the cell tower, denoting whether it is in service or not.

The second set of datasets has the product information and billing information, and time spent as well as information on the cell site being used. These data sets are for three days, and they have valuable information that can be used to evaluate where MTN can start when upgrading their infrastructure.

The data available is in the form of strings, floats, and integers. The data also has a lot of missing values and misspelled names. It is essential to clean the data to ensure that it is ready for analysis.

**Data Preparation**

This stage of the data analysis includes selecting the data, cleaning data, and integrating the data for analysis.

The selected data will be uploaded and converted into four data frames. Data selection will ensure that columns and rows required for effective analysis are retained. The rationale for inclusion or exclusion of data shall be based on the business understanding. The data available seems to be sufficient, but further filtering will be necessary to ensure that we can get insightful information from the data.

The data cleaning process will include removing the null values, removing the duplicate values, and changing column names so that the data frames are easier to integrate. The data cleaning process is the most time consuming part of this project, and more resources will be dedicated to this process.

Integrating the data will take place in two stages. The first stage is integrating the product information data that has product information and cell tower location. The next stage will involve merging the data containing the geolocation of the cell towers in each city with the merged data frame containing product information.

**Analysis**

The analysis will mainly include grouping the final data frame by cities and selecting the billing information. The billing information is represented by the value figures. The Following questions are answered based on the analysis conducted.

1. Which ones are the most used city for the three days?

2. Which cities were most used during business hours?

3. The most used city for the three days?

|  |  |
| --- | --- |
| **City** | **Value** |
| Yopougon | 212,584 |
| Cocody | 190,748 |
| Abobo | 120,835 |
| Koumassi | 71,855 |
| Treichville | 66,381 |
| Adjame | 62,641 |
| Marcory | 61,992 |
| Yamoussoukro | 45,798 |
| San-Pedro | 44,937 |
| Port-Bouet | 39,081 |

**Recommendations**

MTN should focus on the first cities in the list above because these are the cities that offer the most value. The cities at the top of the table above have the highest overall revenue for the three days, which means that they have the most active users. Following the list above will ensure that the infrastructure upgrade is done on a place where returns will be higher so that the company can use the money to continue with the infrastructure upgrade.

**Evaluation**

The model built in this project meets the basic business objectives. The goal of the business is to identify the most important cities that can be targeted for an infrastructure upgrade. MTN wants to ensure that it is efficiently using its resource. The list provided is based on the billing price paid in the different cities for the three days.