**1. Business Understanding**

**Business Overview**

MTN Group Limited, formerly M-Cell, is a South African multinational mobile telecommunications company, operating in many African, European and Asian countries. Its head office is in Johannesburg. As of 30 June 2016, MTN recorded 232.6 million subscribers, making it the eighth largest mobile network operator in the world, and the largest in Africa. Active in over 20 countries, one-third of company revenue comes from Nigeria, where it holds about 35% market share. MTN Cote D' Ivoire is a branch of the MTN group that offers telecom services In Côte D' Ivoire.

**Business Objective**

The main objective of this report is to identify the cities that should be prioritized by MTN when they carry out their system upgrade plan.

**Business Success Criteria**

To compile a list of cities that will increase the return on investment of the

MTN Cote D’ Ivoire System upgrade we will conduct the following;

**Assessing the Situation**

1. **Resource Inventory**
2. Datasets:

i. CDR\_description[“CDR\_description.xlsx”]

Ii. cells\_geo[“cells\_geo.csv”]

Iii. cells\_geo\_description[“cells\_geo\_description.xlsx”]

iv. Telcom\_dataset[“Telcom\_dataset.csv”]

v. Telcom\_dataset2[“Telcom\_dataset2.csv”]

vi. Telcom\_dataset3[“Telcom\_dataset3.csv”]

1. Software( Github, Google Collaboratory)

2. **Assumptions**

1. The data provided is correct and up to date

3. **Constraints**

1. There are no constraints

**Data Mining Goals**

Our data mining goals for this project are as follows:

1. Calculate the most used city for the three days?
2. Which region has the largest number of users.
3. Which service is used most in a particular city.
4. Which city has a greater billing price.

**Data Mining Success Criteria**

Our success criteria will be measured by the following criteria;

* We target the cities that have the most users and the cities that have most users for both during business and home hours.
* We target the type of service that is mostly used in the city with the highest number of users.

**2. Data Understanding**

**Data Understanding Overview**

For this project, we are using the availed dataset by the company. These datasets are

i. CDR\_description

Ii. cells\_geo

Iii. cells\_geo\_description

iv. Telcom\_dataset

v. Telcom\_dataset2

vi. Telcom\_dataset3

**Data Description**

We have two datasets available for this project. A detailed description of the datasets is

provided as follows:

**Cells\_geo\_description -** This dataset contains an explanation of the column names used in the Cells geo dataset. It consists of three columns;

* Column name
* Description
* Format

**Cells\_geo**- This dataset, on the other hand, focuses on the data of MTN telecom services. It contains eleven columns;

* VILLES
* STATUS
* LOCALIZATION
* DECOUPZONE
* ZONENAME
* LONGITUDE
* LATITUDE
* REGION
* AREA
* CELL\_ID
* SITE\_CODE

**CDR\_description -** This dataset gives a description of the column names of datasets ; Telcom\_dataset, Telcom\_dataset1, Telcom\_dataset2.

It has three columns:

* Column name
* Description
* Format

**Telcom\_dataset, Telcom\_dataset1, Telcom\_dataset2 -** Thesethree datasets have similar columns. They have ten columns;

PRODUCT

VALUE

DATETIME

CELL\_ON\_SITE

DW\_A\_NUMBER\_INT

DW\_B\_NUMBER\_INT

COUNTRY\_A

COUNTRY\_B

CELL\_ID

SITE\_ID

**Verifying Data Quality**

The **Cells\_geo** dataset consists of a column “AREA” which contains Null values. The rest of the datasets do not have null values.

**3. Data Preparation**

These are the steps followed in preparing the data

1. Loading Data

Loaded the datasets from the CSV and then created an dataframes from

them.

1. Cleaning Data

Created our four data frames from the datasets provided.

Got rid of all the null value rows and worked with columns with values.

1. Merging of the Datasets

Merged our four datasets into one large dataset so that it would be easier to get good insight from the data.

1. Deriving New Attributes

Observed several key insights from the data;

* 1. Most used city in 3 days.
  2. Which region has the largest number of users.
  3. Which service is used most in a particular city.
  4. Which city has a greater billing price.

**4. Analysis**

During our analysis, we were able to single out the following cities;

1. Most used cities in three days(Descending order).

Cocody : 241

Abobo : 211

Yopougon : 184

Marcory : 96

Adjame : 81

1. Most used regions in three days(Descending order).

Lagune : 1290

Bas Sassandra : 155

Vallee Du Bandama : 101

Lacs : 96

1. Compare the number of users in Abidjan and those outside Abidjan.

Interieur - Outside Abidjan - 1431

Abidjan - 1173

1. City with greater billing price.

Cocody - Voice : 121

- data : 9

- sms : 111

1.The above analysis was done using Pandas. The full analysis can be found in the following notebook. [”https://drive.google.com/open?id=13RPmbkse7Kishj66MPKeb7c-xK3nGyqW”]

1. **Recommendations**

From our analysis, we would recommend that MTN should prioritize the above listed cities. Our main reason behind this recommendation would be that the cities listed above have the highest descending ratio.

Therefore, prioritizing these cities would not only increase the chances of Increased revenue but also would increase the return on investment in the entire MTN systems upgrade technology infrastructure.