**MTN Cote d’Ivoire Infrastructure Upgrade Strategy:**

MTN Côte d'Ivoire would like to upgrade its technology infrastructure for its mobile users in Ivory Coast. We have been provided some datasets for facilitations of this.

Below is the strategy to be used for this

# **Business Understanding**

MTN Côte d'Ivoire is a leading telecommunications company that was founded on1 st July 2005 in Côte d'Ivoire, with the acquisition by the South African group M-Cell, which later became MTN international, of the mobile telephony license of Loteny Telecom (Telecel).

Since July 2005, in application of the provisions of the transfer agreement, the share capital of MTN is 2.865 billion CFA francs , distributed as follows: 64.67% are held by the MTN group and 35.33% by the other shareholders. It provides services for voice, sms & data.

From the provided data, we can see that it has a presence in Abidjan North, Abidjan South, Centre North, East, North West and SouthWest Areas.

In the past few months they have noticed a surge in subscribers as well as network usage. The surge has caused network issues as the current infrastructure is overworked,leading to a lot of complaints from current clients. This also resulted in migration of clients from MTN to Orange Group’s local unit, Orange CI.

MTN Côte d'Ivoire would like to upgrade its technology infrastructure to improve customer experience by providing better quality services to its users, minimise disruptions and maintain its current subscriber base.

**Business Objectives**:

1. To find out the regions to be prioritised during the upgrade by checking the usage.
2. To find out which products have a heavy usage so as to prioritise them.

**Success Criteria:**

1. Descriptive Statistics
2. Clean Data.
3. **Resource Inventory:**

* Datasets (cells\_geo.csv; cells\_geo\_description, CDR\_description, CDR 20120507, CDR 20120508, CDR 20120509)
* Software (Google Colab Notebook, Google Docs, GitHub, Python)

1. **Assumptions:**

* Data provided is correct

1. Constraints:

* Linking of the Telcom data sets with the cells\_geo data set as there were no common values to use while linking.
* Errors in the CELL\_ID & SITE\_ID columns as they contain some characters such as ‘/’, ‘=’ & ‘+’.

**Data Mining:**

Seeing as the data has already been provided, we shall not carry out any data mining.

# 2. Data Understanding:

Below is a table with the datasets provided to guide this exercise:

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Size of data** | **Link to data** | **Description** |
| cells\_geo.csv | 3974 rows × 1 columns  (The dataset has multiple columns in one column) | [Cells\_geo.csv](https://drive.google.com/a/moringaschool.com/file/d/1ABZux280OjL3yWcOn8BDA_f5QsyO0QPU/view?usp=sharing) | Geographical location. Contains latitudes & longitudes, the cell\_IDs, the site codes, the cities and the regions. |
| CDR 20120507 | 5001 rows × 10 columns | [Telecom Dataset1.csv](http://bit.ly/Telcom_dataset1) | Contains values such as billing price, phone numbers for both parties, countries, cell\_ID and site\_ID for 7/05/2012 |
| CDR 20120508 | 5001 rows × 10 columns | [Telecom Dataset2.csv](http://bit.ly/Telcom_dataset2) | Contains values such as billing price, phone numbers for both parties, countries, cell\_ID and site\_ID for 8/05/2012 |
| CDR 20120509 | 5001 rows × 10 columns | [Telecom Dataset3.csv](http://bit.ly/Telcom_dataset3) | Contains values such as billing price, phone numbers for both parties, countries, cell\_ID and site\_ID for 9/05/2012 |

MTN also provided descriptions for each type of data,explaining what is contained in each column.

We shall go through this data to explore it and check the columns provided, as well as their data types for ease of calculation later. We will make sure to understand what exactly each column represents and how this would help us in our calculations. We will verify data quality for missing or erroneous data or any unusable data.

# 3. Data Preparation

We will download the datasets and load them into our database and also load the pandas, numpy & datetime libraries

We will also need to clean the data. We will rename some of the columns on the Telcom datasets as we noticed some errors in spelling such as “SIET\_ID” instead of “SITE\_ID”. We will also need to have uniform names on the 3 datasets as we will combine the datasets into one later. We also checked on the column titles for the geographical location dataset.

We then converted the “DATETIME” column into datetime datatype, split it into dates and time. Then we converted the time column back into datetime data type.

We wanted to create datasets which fit into either business hours or home hours to know when most users use the network. However, we were unable to.

We then did the summary statistics by Area and Product

# 4. Analysis

The data was analysed using python in [this python notebook](https://docs.google.com/document/d/1vLvac9TeRmMTWdSQNpdQJGPgEqfc1sZ43ewfRPGsh1s/edit?usp=sharing).

Below are some of the summary statistics obtained:

|  |  |  |
| --- | --- | --- |
| **Product** | **Count** | **Mean** |
| **Voice** | 5660 | 55.1699 |
| **Data** | 2266 | 7.628861 |
| **SMS** | 7071 | 7.555791 |

|  |  |  |
| --- | --- | --- |
| **Area** | **Count** | **Unique** |
|  | 23 | 23 |
| **Abidjan North** | 1222 | 1196 |
| **Abidjan South** | 655 | 645 |
| **Centre North** | 615 | 614 |
| **East** | 342 | 338 |
| **North West** | 583 | 569 |
| **South West** | 534 | 534 |

# 5. Recommendation:

We would like to recommend SMS be given a higher priority during the upgrade as it has the heaviest usage.

We would like to also recommend that Abidjan North be prioritised as it has the highest usage.