<u>**DATA REPORT**</u>

The following is a data report on how MTN Cote d'Ivoire would go about upgrading its technology.

Problem Description

Currently MTN Cote d'Ivoire would like to upgrade its technology infrastructure for its mobile users in Ivory Coast. Studying the given dataset, how does MTN Cote d'Ivoire go about the upgrade of its infrastructure strategy within the given cities?

PHASE 1: BUSINESS UNDERSTANDING

- Identify goals
  The goal of this project is to upgrade technology infrastructure for MTN users

- Assess situation
  Six datasets are available. We have data on calls,sms voice usage for three days
  The data is stored in csv format and no security issues exist preventing access to required data.
  As for personnel, a data analyst is on board.
- Define data mining goals.
  Our data mining goals will be based on what products i.e voice or calls to upgrade network infrastructure
- Produce a project plan.
  The project plan will take about 10 weeks ; with phase one to take one week, phase two 2 weeks, phase three, 3 weeks, phase four 2 weeks, phase five and six 3 weeks.

PHASE 2: DATA UNDERSTANDING

- Gather data.
  The data available is from existing sources and hence no needd to purchase external data.
  The attributes of the data are ten columns in the three day datasets.
  We shall merge the three day data sets to our discretion
  The data sets that contain missing values have been filled with the value 0
  The datasets containing duplicate values have been dropped as well

- Describe data.

  The data types of the various datasets are object,integers and floats

- Explore data.

  We shall explore data via pandas library

- Verifying data quality

  From the datasets used we noted some missing values and duplicate values.

PHASE THREE: DATA PREPARATION
- Select

  The data was downloaded from the links provided and then loaded on to the Google Collab via csv importation.

- Clean.

  Cleaning was done in several stages:
  - We dropped any duplicate values
  - We filled missing values with the value 0
  - We converted data types

- Construct

  The new column values ranked was created for day 1
- Integrate

  We merged datasets of day 1,2 and 3.
- Format

  All data is accessible from within the SPSS modeller.

PHASE FOUR: MODELING
- Select modeling technique

  The SPSS technique will be suitable
- Design test

  We will test the success of supervised and unsupervised data types.
- Build model

  Parameter settings will be adjusted in this stage, We shall run the model and draw meaningful conclusions
- Assess

  The results are as dramatic as expected and by knowing states with the most local votes, we are able to win the elections for our candidate.

PHASE FIVE: EVALUATION
- Evaluate results

The rank of the model was done through statistical analysis and ease of interpretation. The results are stated clearly and in a representable form.

The unique findings are that usage of voice calls increased at certain hours of the day and that sms usage was not as high as first expected.

The results answer the questions set in the objectives to a satisfactory degree.

- Review process

A return to the exploration process is always warranted when something unusual appears in
another phase of the CRISP-DM process.

Data preparation, requires patience, since it can take a very long time.

It is vital to stay focused on the business problem at hand, because once the data is ready for analysis, it's all too easy to start constructing models without regard to the bigger picture.

Once the modeling phase is over, business understanding is even more important in deciding
how to implement results and determine what further studies are warranted

- Determine next steps

The next step is to continue to the deployment phase.

PHASE SIX: DEPLOYMENT
- Plan, monitor and maintain

The column variables of product,value are fluctuational and need to be monitored

The accuracy of the model is when m mobile users in the MTN network increase as this will be a sign of customer satisfaction

The model is subject to changes in the market.

The model is specific to the Cote d' Ivoire region.

- Report final results

No deviations exist from the original project plan.

The proposed plan for deployment is right after analysis has been done and all parties involved have met on a decision to deploy the model

- Review final results

The overall impressions of the project are that it performed well subject to constraints.

The observations were that more advanced modeling techniques were needed.

The difficulties arose in the data preparation and modeling stages