# Capstone Project Submission

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

**Team Member's Name, Email, and Contribution:**

**Team Member's Role:-**

- **RAM MANOHAR THAKUR**

  **eMail- rayanka148@gmail.com**
  - Data Understanding
  - Feature Analysis
    - Calls Data, international call, voice mail call, No. of voice mail
  - Feature Engineering
    - Null value check
  - Data Visualization
    - Donut plot, box plot
  - Multivariate Analysis
    - Correlation matrix
  - Research Analytics
    - Technical documentation

- **Satyasuryapavan Chanduri**

  **eMail- satyasuryapavanch@gmail.com**
  - Data Understanding
  - Feature Analysis
    - Area code, account length, state
  - Feature Engineering
    - Missing value
  - Data Visualization
    - Count plot
  - Multivariate Analysis
    - Heat map
  - Research Analytics
    - Technical documentation

- **Avinash Arora**

  **eMail- balarampanigrahy42@gmail.com**
  - Data Understanding
  - Feature Analysis
    - Customer service call, churn
  - Feature Engineering
    - Duplicate value
  - Data Visualization
    - Scatter plot

- Multivariate Analysis
- Research Analytics
  - Technical documentation

**Please paste the GitHub Repo link.**

Github Link:- https://github.com/rayanka148/capstone-project-python

**Please write a summary of your Capstone project and its components. Describe the problem statement, your approaches, and your conclusions. (200-400 words)**

PROBLEM STATEMENT:
Finding factors and cause those influence customers to churn. Retain churn customers by taking appropriate steps providing offers based on affecting factors. Using the data provided, this paper aims to analyze the data to determine what variables are correlated with customer churn, if any. To identify the people that might churn, will also be analyze.

APPROACH:
Exploratory Data Analysis (EDA) is an approach to analyze data. The first and foremost task is that the data analysis to view the data and tries to make some sense out of it. Later we figure out what question we want to ask and how to use the available data to get the insights and all the answers that we need from the data set. EDA helps us to

- Delve into the dataset
- Examine the relationships among the variables
- Identify the interesting observation
- Develop an initial idea of possible associations among the predictors and the target variable.

Our telecom data set consists of 20 features and 3333 records. out of which 3 objective, integer, and float both have 8 and 1 Boolean data type.
CONCLUSION:
The telecom market is saturated and customer growth rates are low. The key focus of market players therefore is on retention and churn control. This project explores the churn dataset to identify the key drivers of churn and grab key insights from the dataset. The insights we have gained into the churn data set through the use of exploratory data analysis:

- The four charge fields are linear functions of the minute fields.
- The area code field and the state field are anomalous and can be omitted.
- The correlation among the remaining predictor variables are weak, allowing us to retail them all for any data mining model.
- Customer with the international plan tend to churn more frequently.
- Customers with four or more customer service call churn more than four times as often as do the other customers.
- Customers with high day minutes and evening minutes tend to churn at a higher rate than do the other customer.

There is no obvious association of churn with the variable day calls, evening calls, night calls, international calls, night minutes, international minutes, account length or voice mail message.