

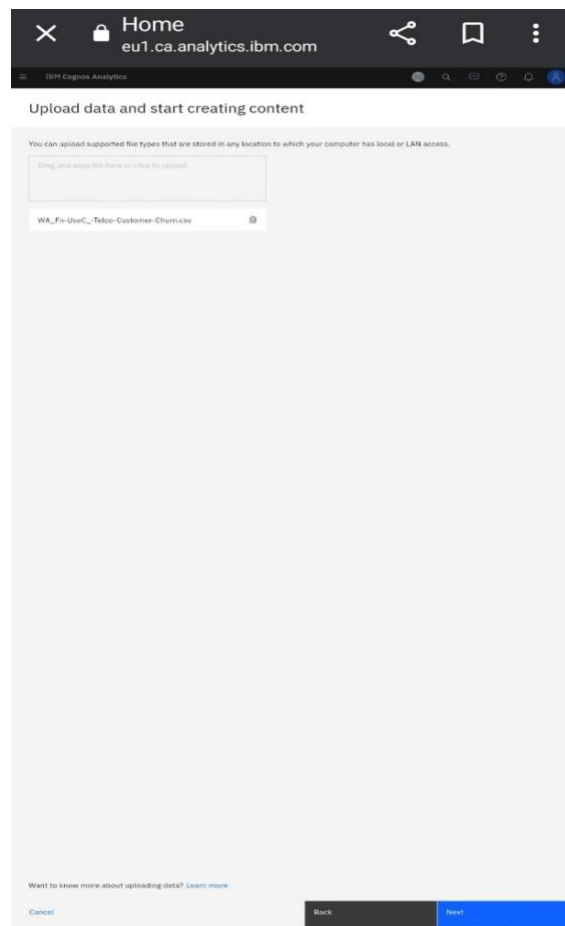
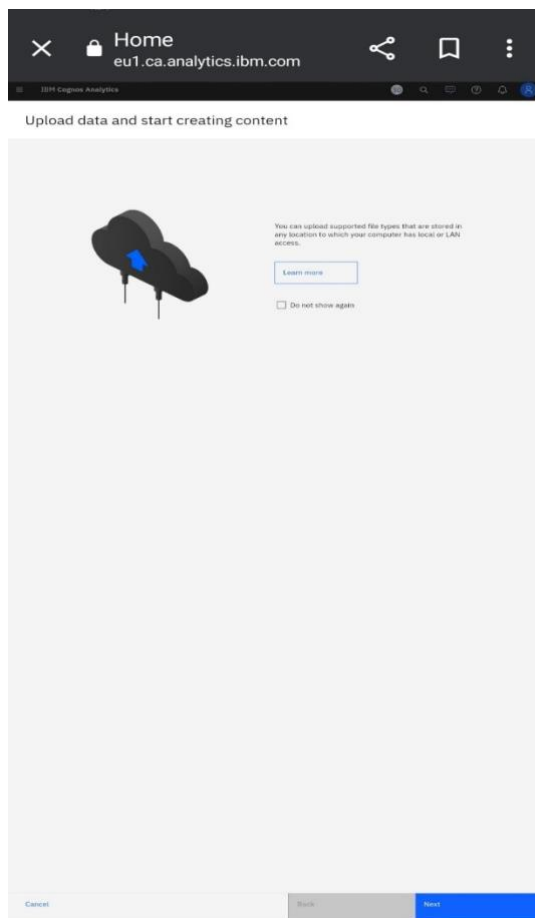
Customer Churn Prediction

Phase 3

Dataset loading and preprocessing

Uploading dataset:

- In this project phase we are going to upload the dataset provided from kaggle in skillup platform
- First, we have to register in the IBM Cloud service and IBM cognos analytics site
- After registering to IBM account we need to create a project and upload our dataset from our device



Preprocessing dataset

- Clean the data by handling missing values, outliers, data quality issues
- Transform and reshape the data as needed . This might include feature engineering , Data aggregation , or other data preparation.

The screenshot shows a dialog box titled "Create a data group (numeric style)". The "Name" field is set to "TotalCharges (Group)". Under "Group names", there is a "Groups" dropdown set to "5". The "Range border values" section shows five ranges on the left and their corresponding numerical borders on the right, connected by lines. The ranges are: "6951.60 and above", "5218.40 to < 6951.60", "3485.20 to < 5218.40", "1752.00 to < 3485.20", and "less than 1752.00". The borders are: "Higher", "6951.6", "5218.4", "3485.2", "1752", and "Lower". At the bottom, there is a checkbox "Group NULL values as" which is checked, with a dropdown set to "Null". There are three buttons at the bottom: "Create a data group (text style)", "Cancel", and "Create".

Group names	Groups	Range border values
6951.60 and above	5	Higher
5218.40 to < 6951.60		6951.6
3485.20 to < 5218.40		5218.4
1752.00 to < 3485.20		3485.2
less than 1752.00		1752
		Lower

☒ Group NULL values as Null

Create a data group (text style) Cancel Create

- Data Preprocess is one of the crucial step in the Data Analytics . Otherwise our analysis can be irrelevant due to unwanted noises in the data.
- This includes null value removal , formatting ,etc.

The screenshot shows a dialog box titled "Create a data group (numeric style)". The "Name" field is set to "MonthlyCharges (Group)". Under "Group names", there is a "Groups" dropdown set to "5". The "Range border values" section shows five ranges on the left and their corresponding numerical borders on the right, connected by lines. The ranges are: "98.650 and above", "78.550 to < 98.650", "58.450 to < 78.550", "38.350 to < 58.450", and "less than 38.350". The borders are: "Higher", "98.65", "78.55", "58.45", "38.35", and "Lower". At the bottom, there is a checkbox "Group NULL values as" which is checked, with a dropdown set to "Null". There are three buttons at the bottom: "Create a data group (text style)", "Cancel", and "Create".

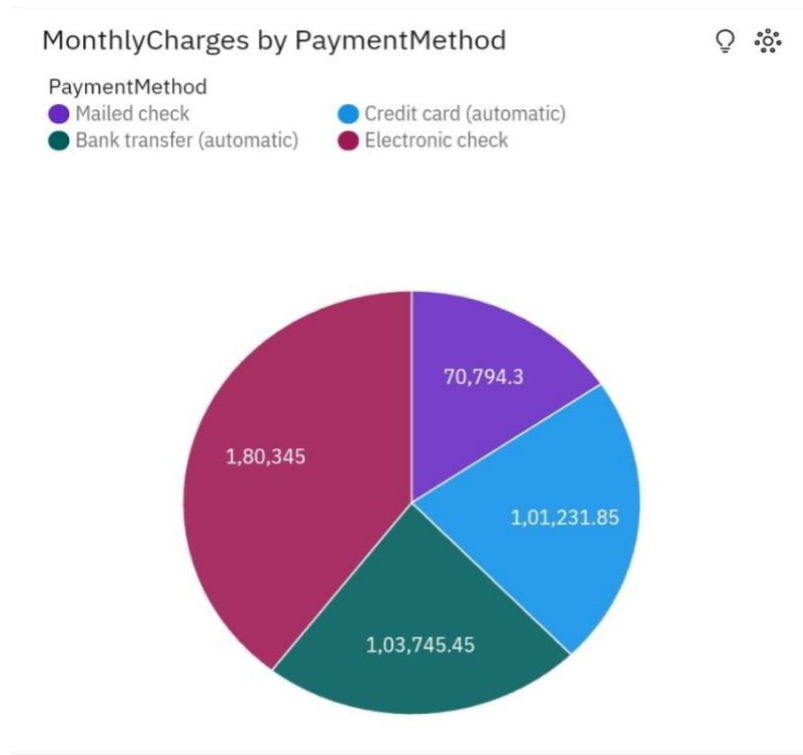
Group names	Groups	Range border values
98.650 and above	5	Higher
78.550 to < 98.650		98.65
58.450 to < 78.550		78.55
38.350 to < 58.450		58.45
less than 38.350		38.35
		Lower

☒ Group NULL values as Null

Create a data group (text style) Cancel Create

Analysis and visualization(Dashboard creation)

Analysis

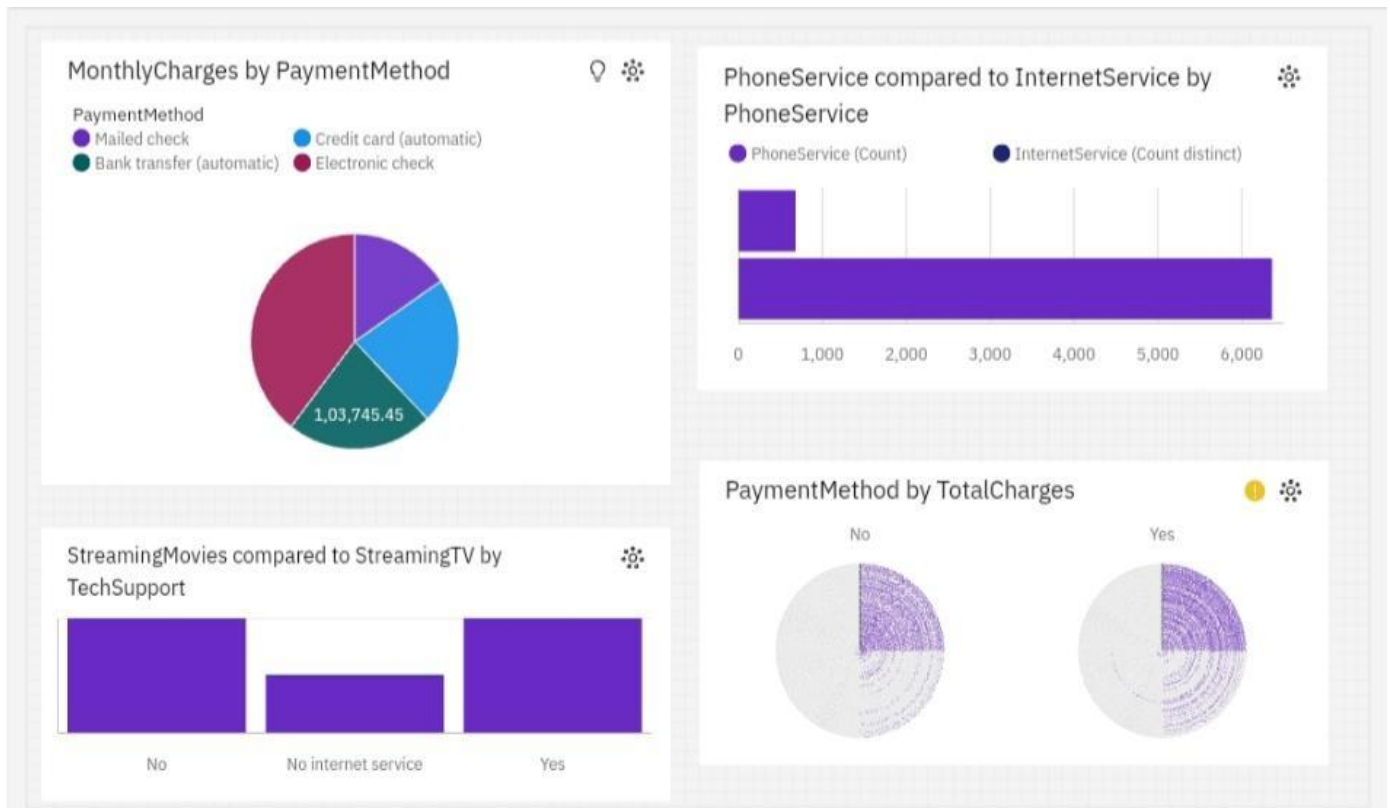


- Use IBM Cognos for descriptive and exploratory data analysis. Conduct statistical analysis to understand the dataset's characteristics.
- Identify patterns, trends, and correlations within the data. Look for potential factors that may contribute to website traffic analysis.

Visualization

- Create visualizations using IBM Cognos to present your findings. This might include:
- Histograms, scatter plots, and box plots to visualize data distributions and relationships.
- Line charts or time series visualizations to explore trends over time.
- Dashboards that summarize key insights in a user-friendly manner.

Dashboard



By,

Abdulla A (au720921243001)

Abhijith J (au720921243002)

Don M S (720921243019)

Mathews M Jaico (au720921243032)

Mohammed Fateen (au720921243035)