**Project 4: Regression Analysis**

**Team Members:**

**Rachel Menezes: 905508013**

**Diplav : 605627748**

**Evyn Chiappe : 605201321**

**QUESTION 9:**

**About the Data:**

**9.1 Report the following statistics for each hashtag:**

**ANS:**

**Average number of tweets per hour:**

Average number of tweets per hour #goHwaks: 292.09326424870466

Average number of tweets per hour #gopatriots: 40.888695652173915

Average number of tweets per hour #nfl: 396.97103918228277

Average number of tweets per hour #patriots: 750.6320272572402

Average number of tweets per hour #sb49: 1275.5557461406518

Average number of tweets per hour #superbowl: 2067.824531516184

**Average number of followers:**

Average number of followers of users posting tweet #goHwaks: 2217.9237355281984

Average number of followers of users posting tweet #gopatriots: 1427.2526051635405

Average number of followers of users posting tweet #nfl: 4662.37544523693

Average number of followers of users posting tweet #patriots: 3280.4635616550277

Average number of followers of users posting tweet #sb49: 10374.160292019487

Average number of followers of users posting tweet #superbowl: 8814.96799424623

**Average number of retweets per tweet:**

Average number of retweets per tweet #goHwaks: 2.01320939913198

Average number of retweets per tweet #gopatriots: 1.408191910169

Average number of retweets per tweet #nfl: 1.534460265554

Average number of retweets per tweet #patriots: 1.78528712884

Average number of retweets per tweet #sb49: 2.5271344411

Average number of retweets per tweet #superbowl: 2.3911895819207

**9.2 Plot “number of tweets in hour” over time for #SuperBowl and #NFL:**

**ANS:**

**Chart

Description automatically generated**

**Chart, histogram

Description automatically generated**

**QUESTION 10:**

**ANS:**

I have performed following task on the provided twitter dataset.

1. Fan Base analysis before, during and after the match.
2. Analyzing the average Sentiment trend during the game.
3. Identifying the Key Moment and Key Player in a Game at different duration
4. Predicting Fanbase from Tweet whether Massachusetts or Washington.

Fan Base analysis before, during and after the match

**Describe your task?**

The aim for this task was to find which team has better fanbase in term of volume by analyzing the volume of tweet in support for both the teams at different time interval before, during and after the game.

**Explore the data and any metadata?**

For this we explored tweets\_#gohawks.txt and tweets\_#gopatriots.txt which contains the relevant positive tweets for patriots and hawk team.

The tweets\_#gohawks.txt contains around 169122 tweets while tweets\_#gopatriots.txt contains around 23511 tweets.

Chart, bar chart

Description automatically generated