**EDA Summary**

* Outcome of your EDA
* What do you feel was missed during the analysis?
* Were there any variables you felt could have helped in the analysis?
* Were there any assumptions made you felt were incorrect?
* What challenges did you face, what did you not fully understand?

My final project was based on Telecom customer churn” analysis based on Kaggle data.

**Outcome of your EDA:**

* We have 21 different variables in our dataset but below variables are primary used to find the impact on churn. Primary variable component that will be utilize for correlation is Churn with Monthly Charges and Total Charges. We also made a bin/group based on Tenure to further find the data impacts
  + Churn
  + Monthly Charges
  + Total Charges
  + Tenure
  + Gender
  + Senior Citizen
* *Churn is high when Monthly Charges are high, Telecom companies needs to respond and review with customer whenever they reach out to customer services for any charge related issues*
* *Tenure also has significant impact on churn*
* *However, if we combine the insights of 3 parameters Tenure, Monthly Charges & Total Charges then the outcomes are clear: - Higher Monthly Charge at lower tenure results into lower Total Charge. Hence, all these 3 factors viz Higher Monthly Charge, Lower tenure and Lower Total Charge are linked to High Churn*

**What do you feel was missed during the analysis?**

* The Senior Citizen field has some relationship with churn which needs to be explored more
* I did not get time to explore and test this dataset with other models to compare the performance
* Real world data will have more challenges and issues than these test data what we use so we should try more with other data as well

**Were there any variables you felt could have helped in the analysis?**

* The customer Service information about each customer / account was not available which would have been another
* Geo data about of city and exact location will give more insights on these challenges

**Were there any assumptions made you felt were incorrect?**

* Category like postpaid(contract) and prepaid(No contract) were not defined in dataset
* Location information’s are not that detailed or correct
* Relationship between Total charges and Monthly charges does not seems to be correct

**What challenges did you face, what did you not fully understand?**

* As I mentioned earlier, we need more time to explore the other models and performance comparisons
* The same churn can be explored more from other variable point of you so we should try out those as well!
* Also removed some missing value data, specially based on Total charges, yet to explore the impacts