**ICA 8 – Group Work Assignment Preliminary Project Presentations**

**Arman –**

Goal: create an app that summarizes F1 race details from multiple years for use by engineering team or drivers to find insights to achieve quicker times.

Data: Pulled from an api (didn’t catch name). It was pretty clean and did not have many missing values or need for imputation. For some races lap times were not recorded for drivers because they did not finish in time so he tried to impute lap times in this case using the mean from that race.

App: The app utilizes multiple pages, dropdown boxes, and sliders to allow the user many options for customizing their visuals.

**Brandon (myself) –**

Feedback received: I received some feedback on how to make my app cleaner. One was to sort my histograms by value first before plotting. Another was to add some specific captions ot my figures for understanding, add version requirements to ‘requirements.txt’ so the app does not break in the future with an update.

**Sanjeev –**

Goal: Create an app that will document their entry into a machine learning prediction contest. The contest is to create the best model based on RMSE to predict the duration of a NYC taxi cab drive.

Data: The data is hosted on google cloud services. To access the data he had to go through the Google BigQuery data warehouse.

App: He has not created any streamlit portions yet.

**Hussain –**

Goal: Create an app that matches data science jobs with job seekers based on their desires and skills.

Data: He scraped data from Indeed.com using a custom scraper. Making the scraper involved many subjective decisions about how to chop up the data. There are a lot of missing values for one of the most important factors, salary. He has explored mean and KNN imputation so far but will try MICE.

App: He makes good use of the page feature to break up the app into multiple parts. Some of the data IDA and EDA is included in the app, while me and Arman seemed to keep that separate. He was having trouble putting seaborn figures on streamlit so might switch to plotly.

**Siddish –**

Goal: Combine car insurance and automobile data to help predict the insurance cost of a car for prospective buyers. (at this point our group was running out of time so my notes become more sparse).

Data: I did not catch where the data was from but the automobile data appeared to be like the well known “cars” dataset hosted on seaborn and other places. There was not a lot of missingness in the data he used. He performed a lot of univariate and bivariate statistics and has lots of conclusions about how various aspects of a car might affect insurance.

App: He has not created a streamlit app yet so there is not much to write here. Our group did spend some time discussing how to make interactive plots with the streamlit functions. First write a function that creates a figure. Then have the input of that function be some parameters from an interactive element such as numerical input from a slider or category selection from a drop-down box. I think adding something like this can take my own app to the next level.