Description of my work:

* Performed EDA on the dataset to help in finding best ways to answer our problem
* Performed first stage data preprocessing which included the following;
  + Creating new column by extracting data from different column (Year, Time duration…etc.) which assisted in analysis.
  + Imputed missing values and verified the authenticity of the imputation
  + Collapsed subcategories to prepare variables for modeling (weather condition & wind speed)
* Performed second stage data preprocessing which included the following;
  + Performing preprocessing on different data set and combined it with the original data to enhance model performance that was weakened by the imbalanced data.
  + Tried different missing value imputation methods (regression) to impute values for model features with highest missing values (precipitation and Wind chill)
* Assisted with modeling requirements (providing different clean datasets and brainstorm meetings)
* Provided weather conditions EDA Plots for the GUI.
* Performed initial logistic regression model that included weather condition variables only