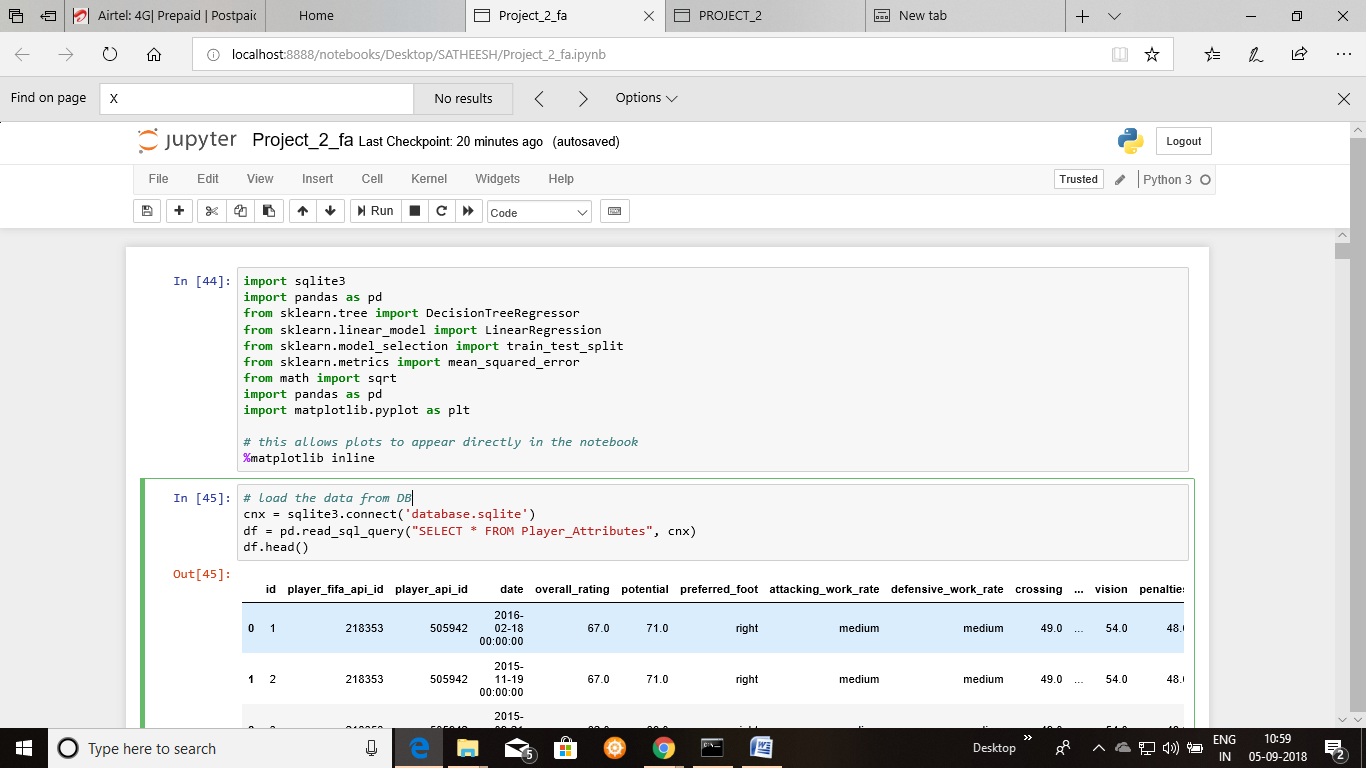
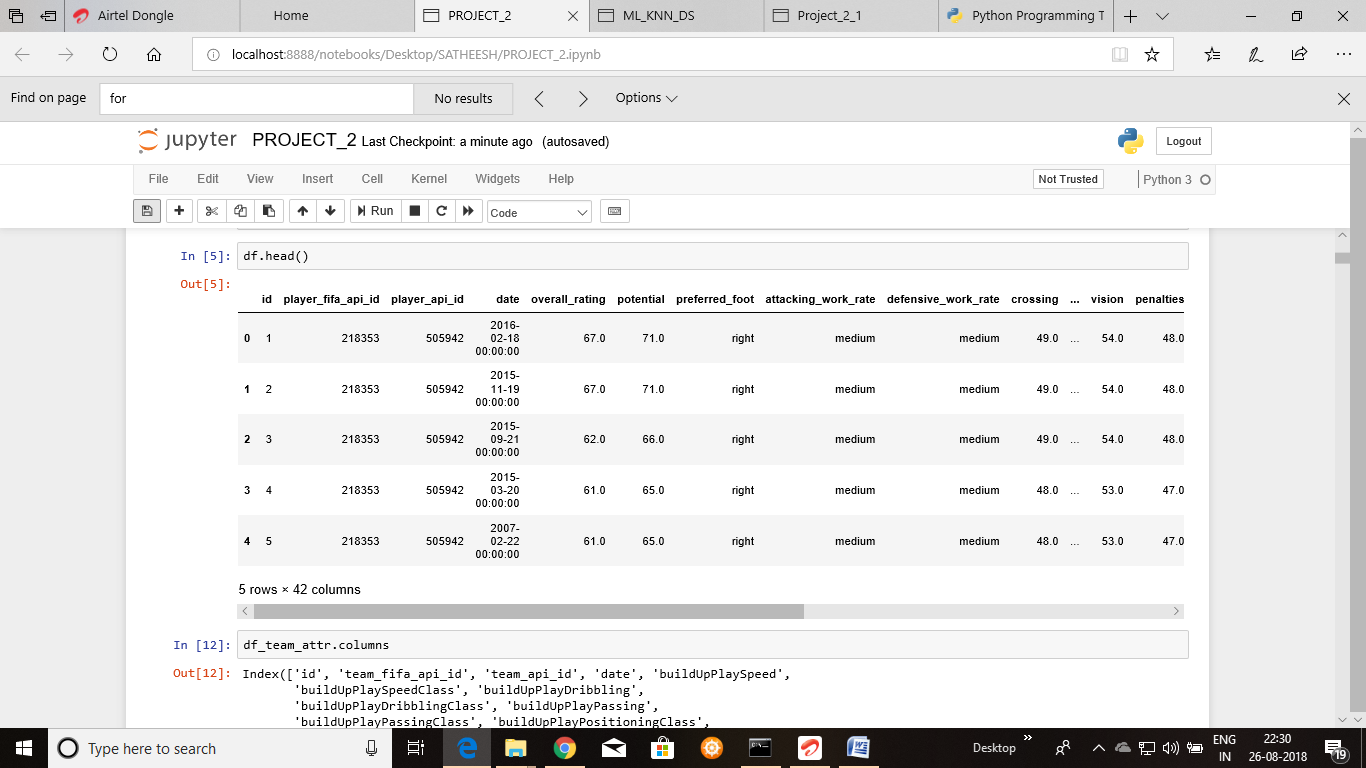
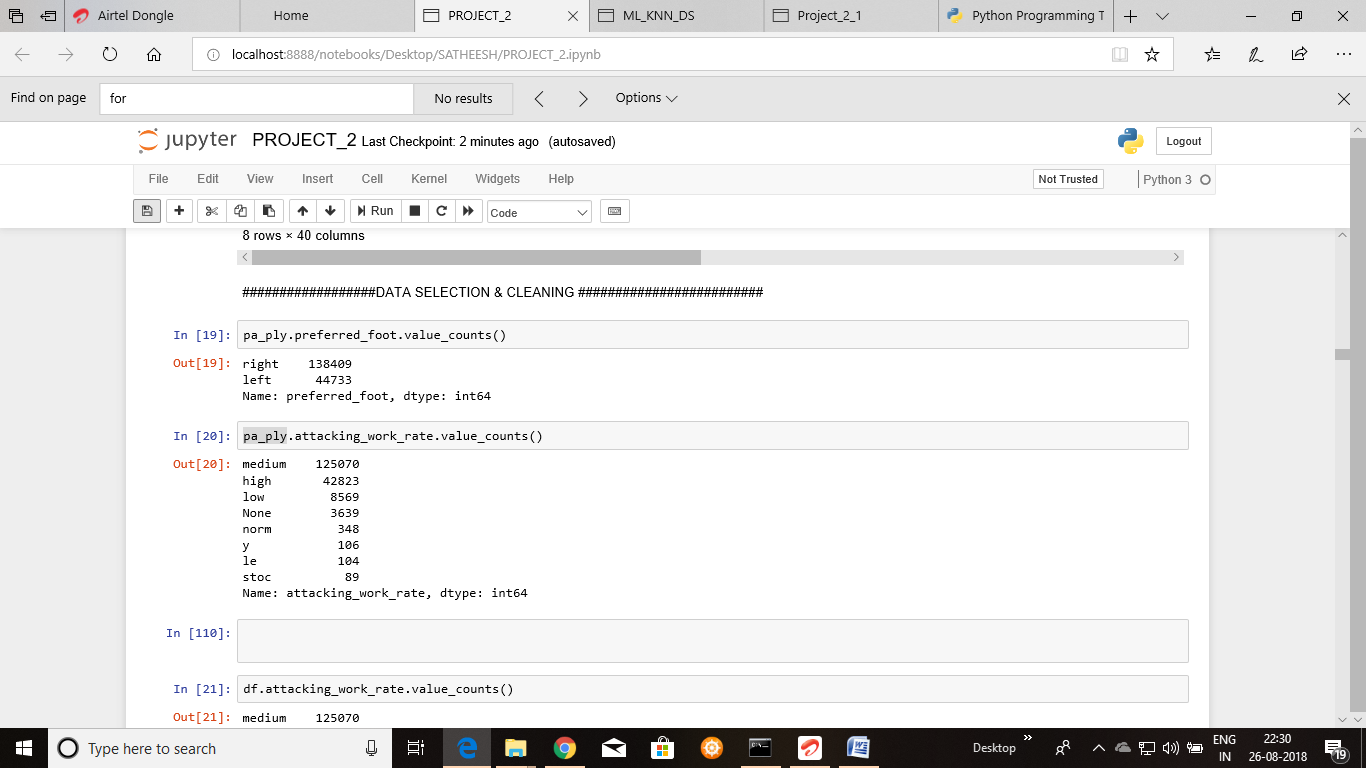
**OUTPUT FOR PROJECT-2**

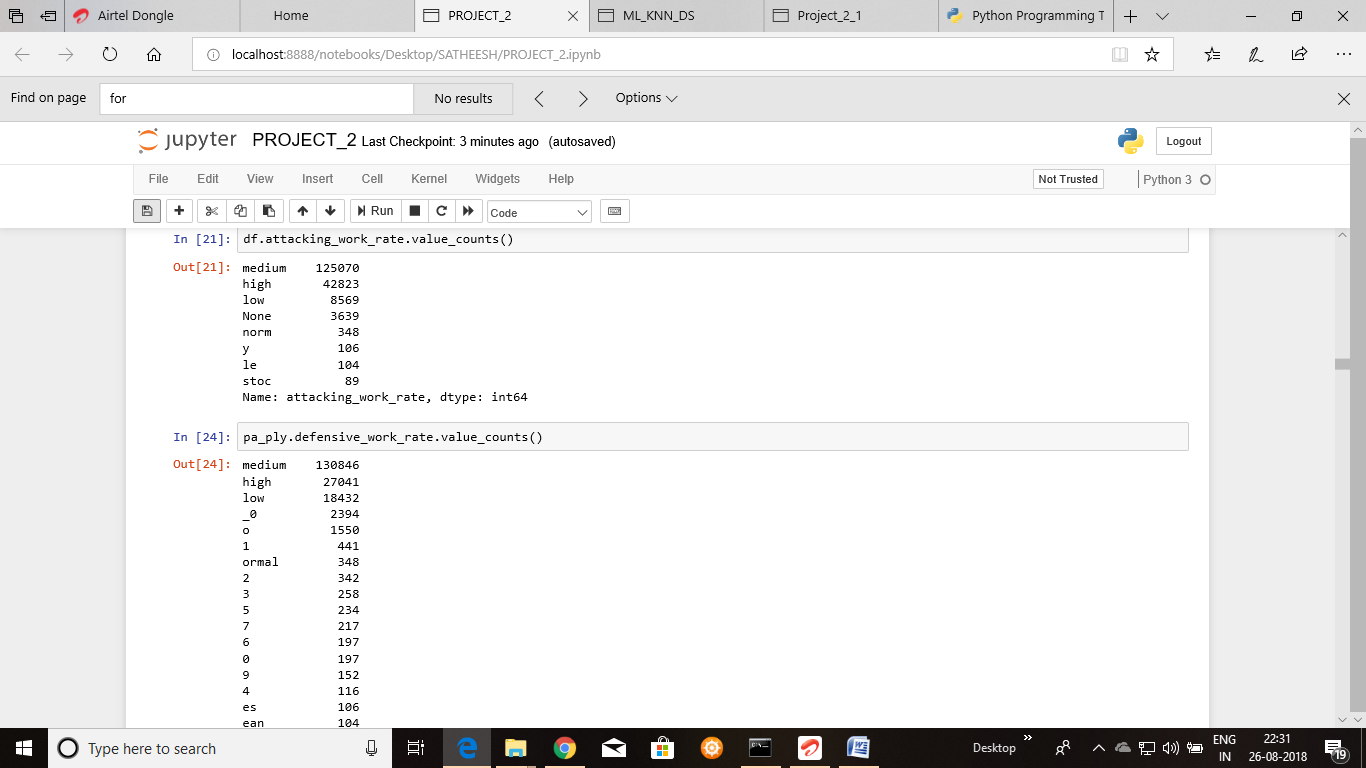
**1.Loading data**

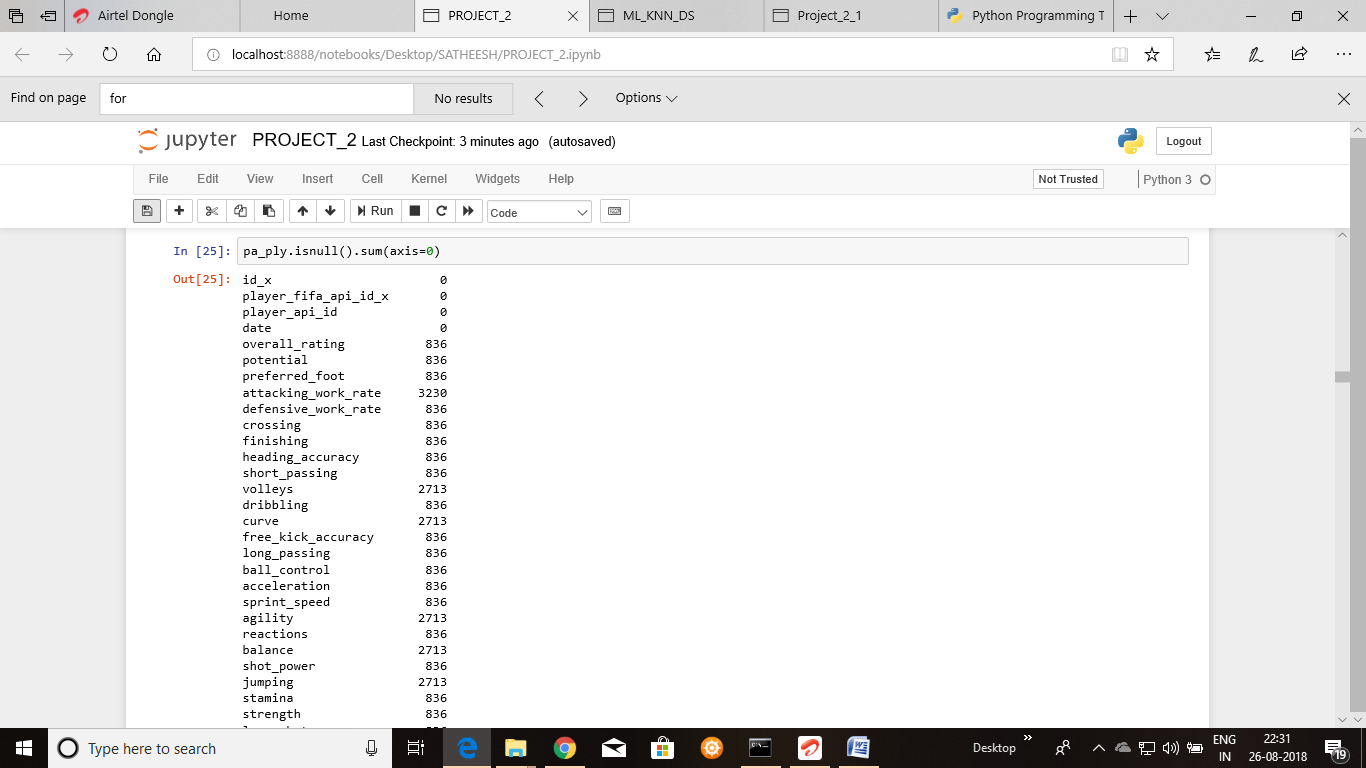
****

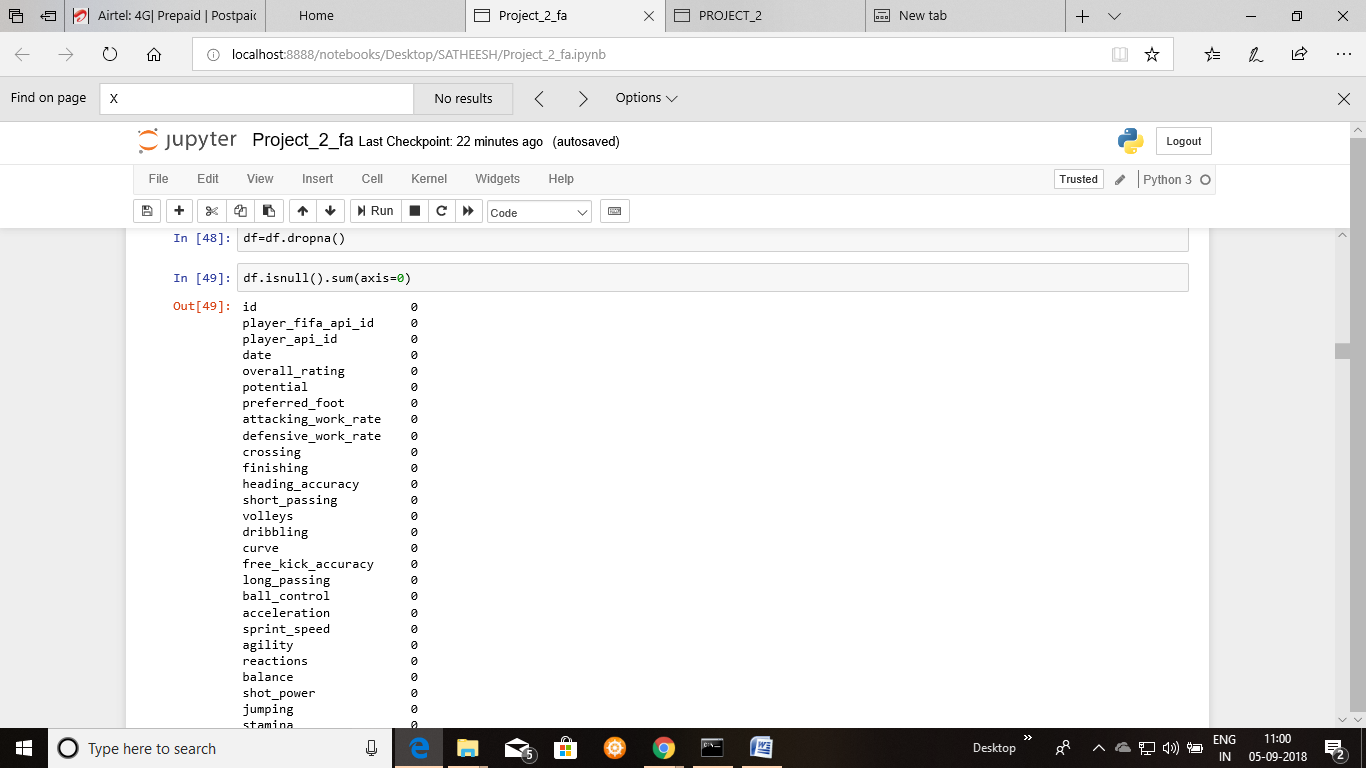
****

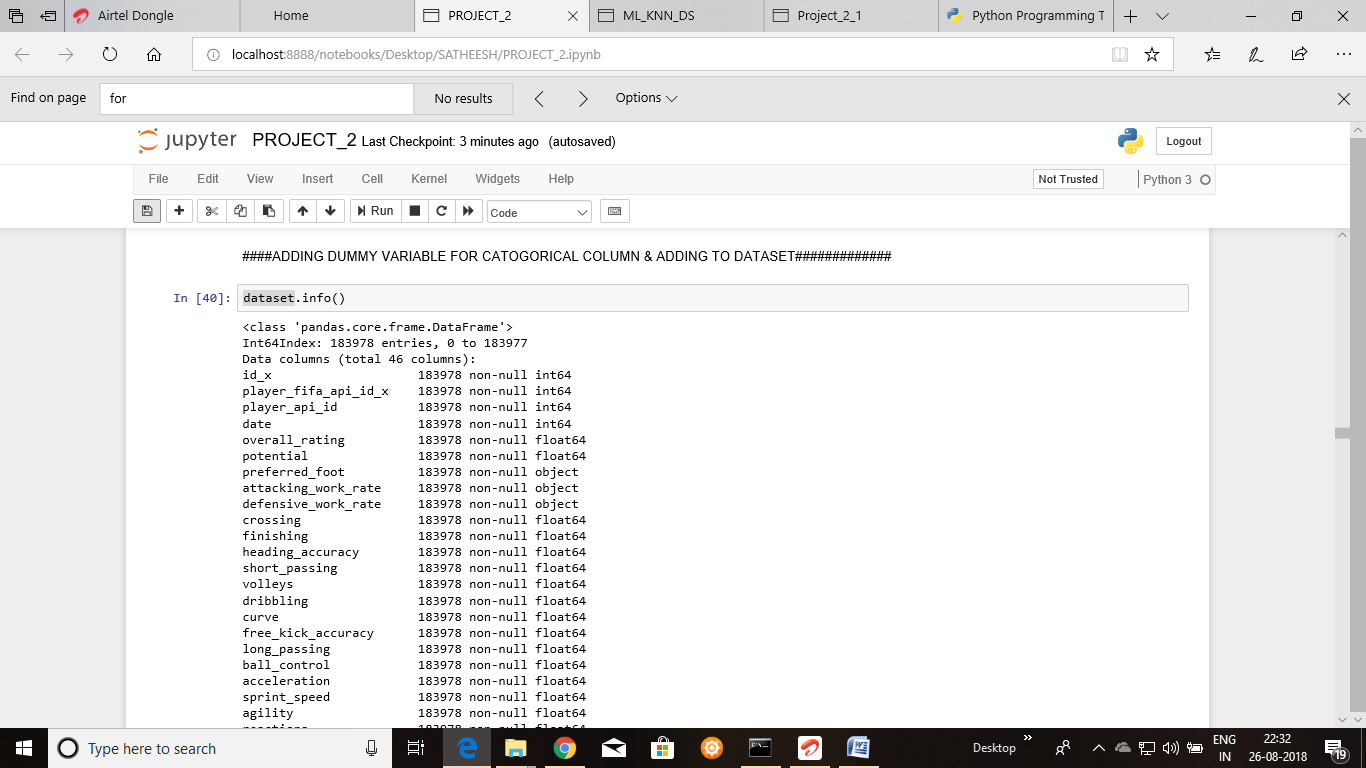
**2) Data Analysis and cleaning**

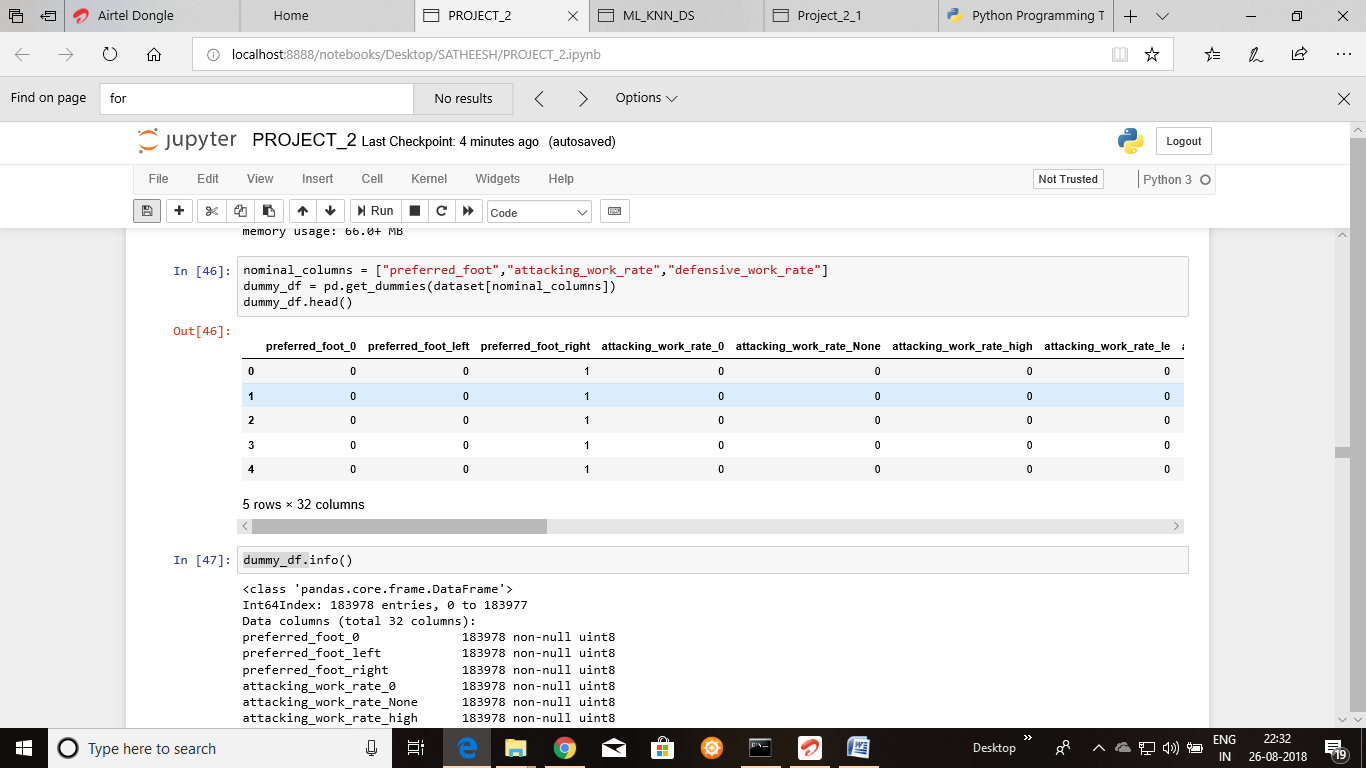
****

****

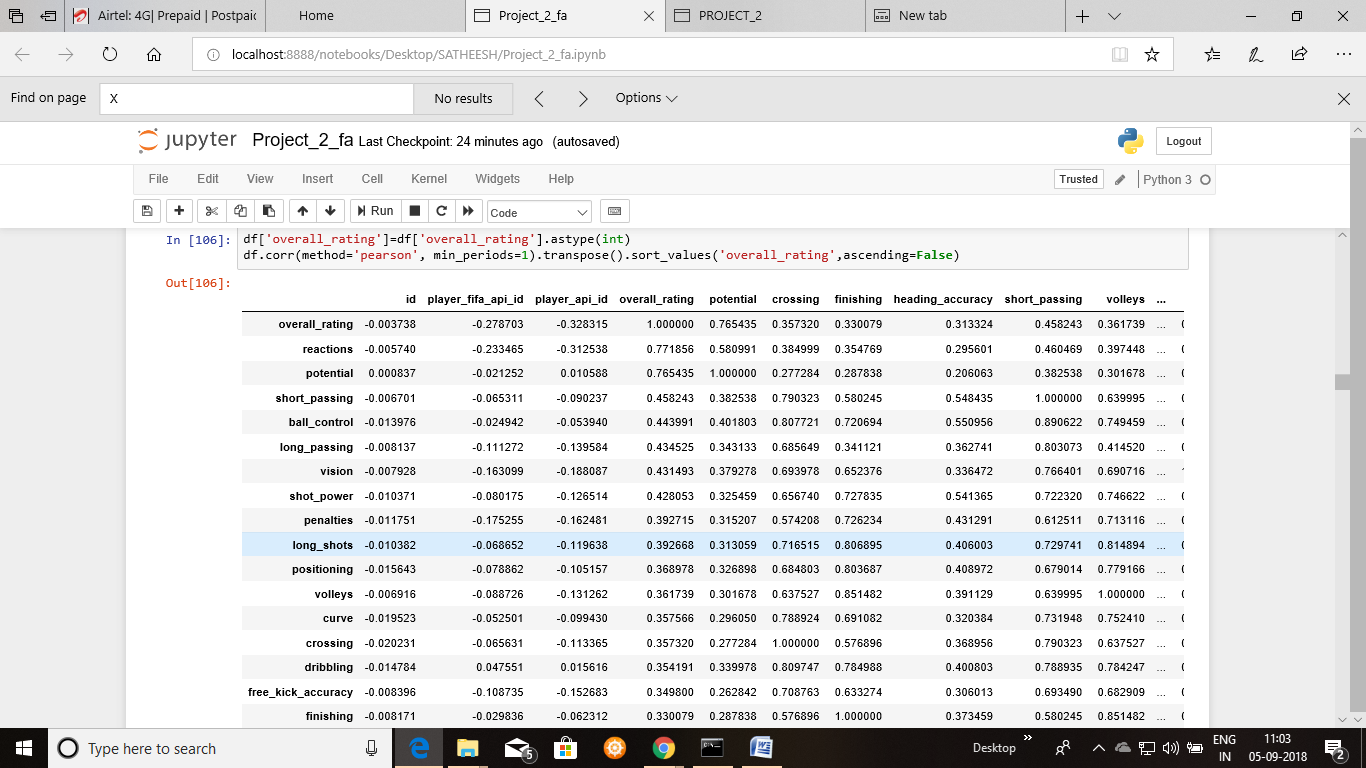
****

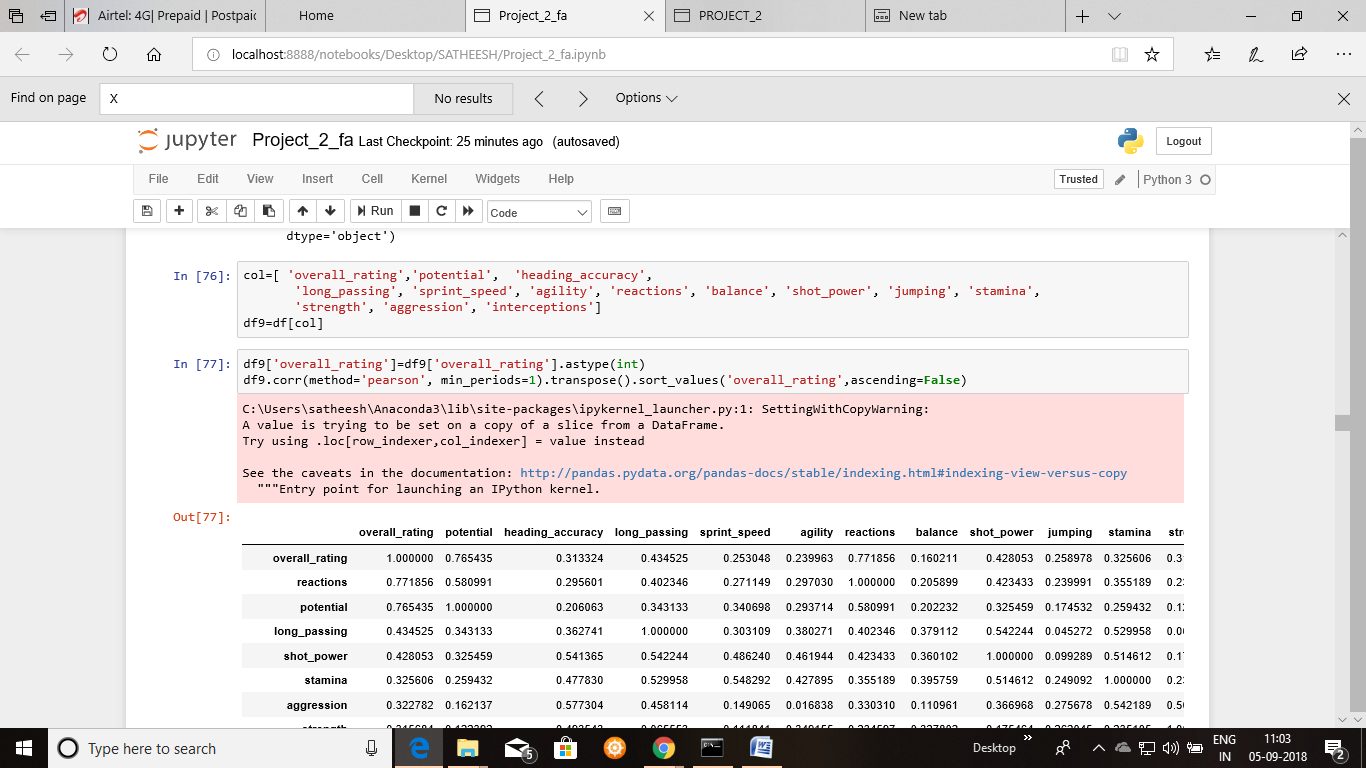
****

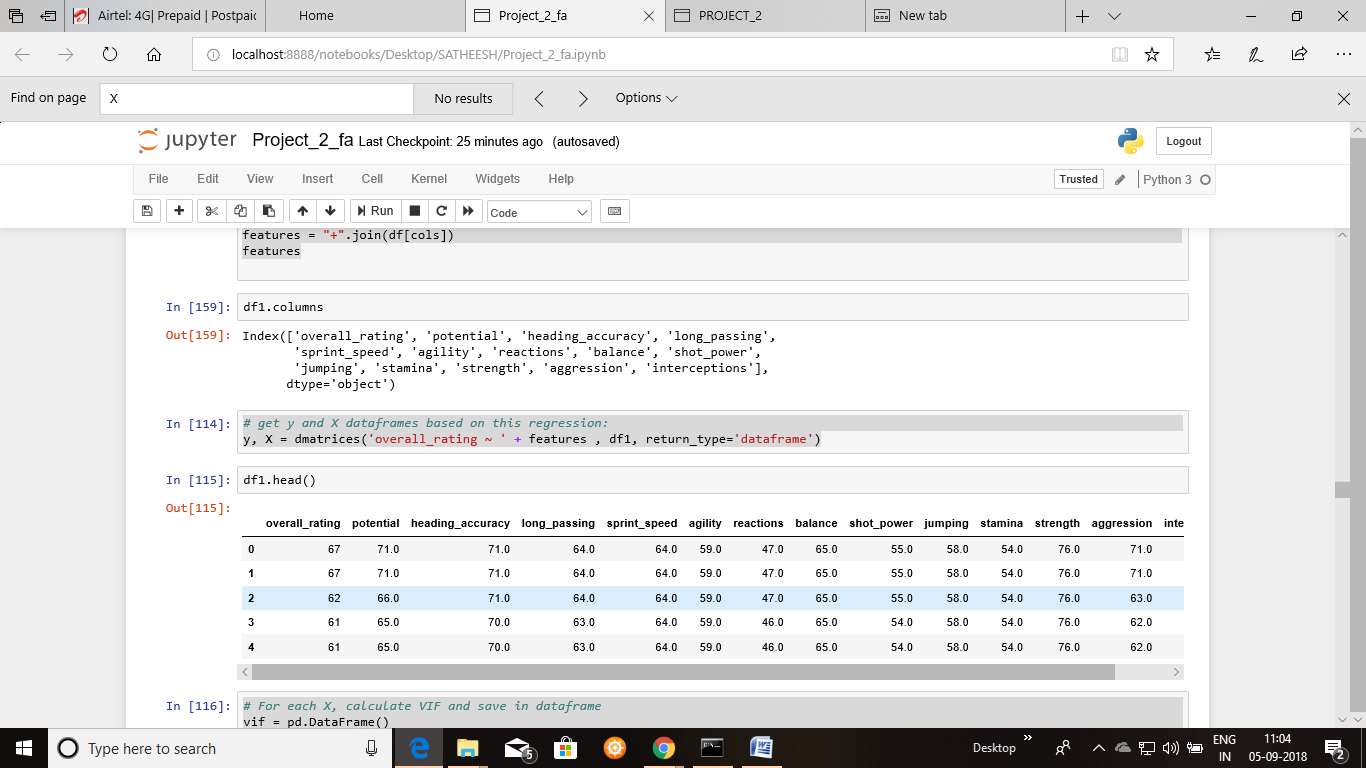
****

****

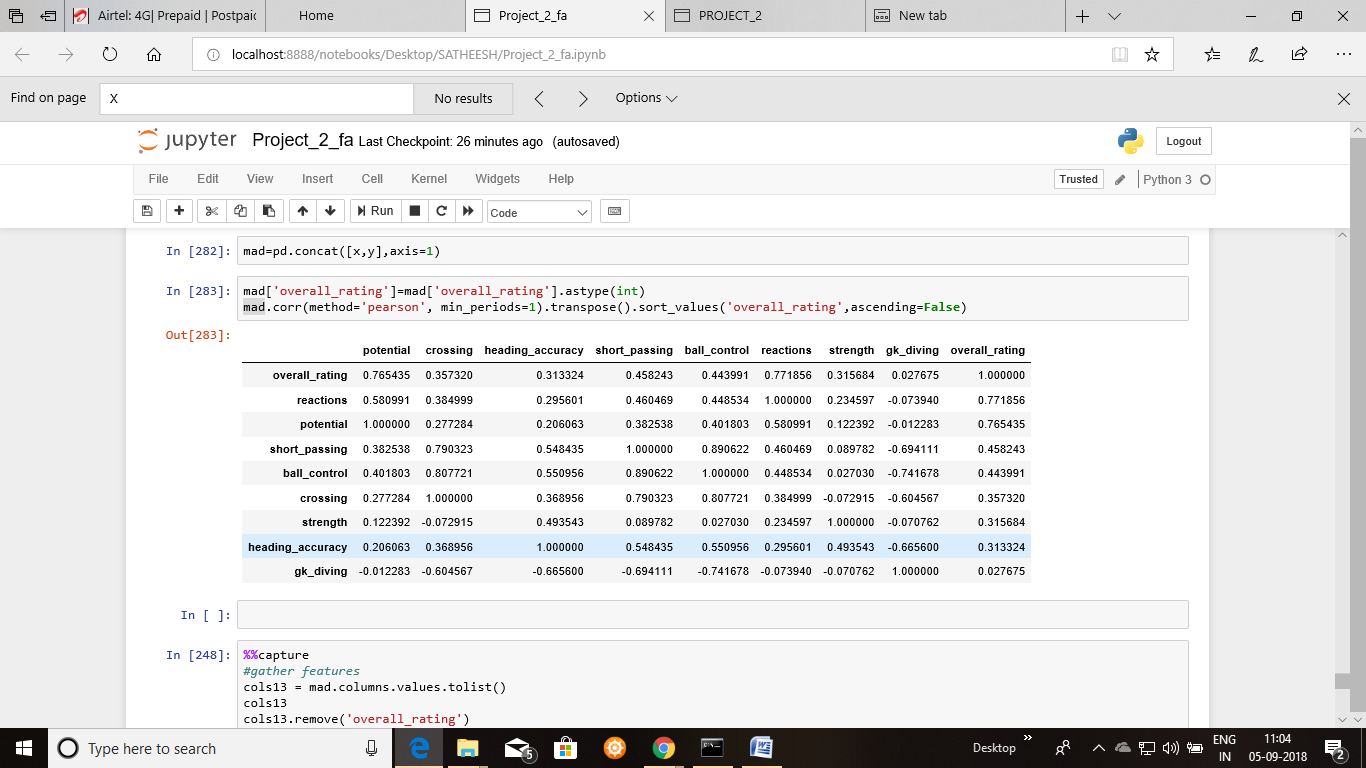
**FEATURE SELECTION WITH PEARSON CORRELATION AND VIF**

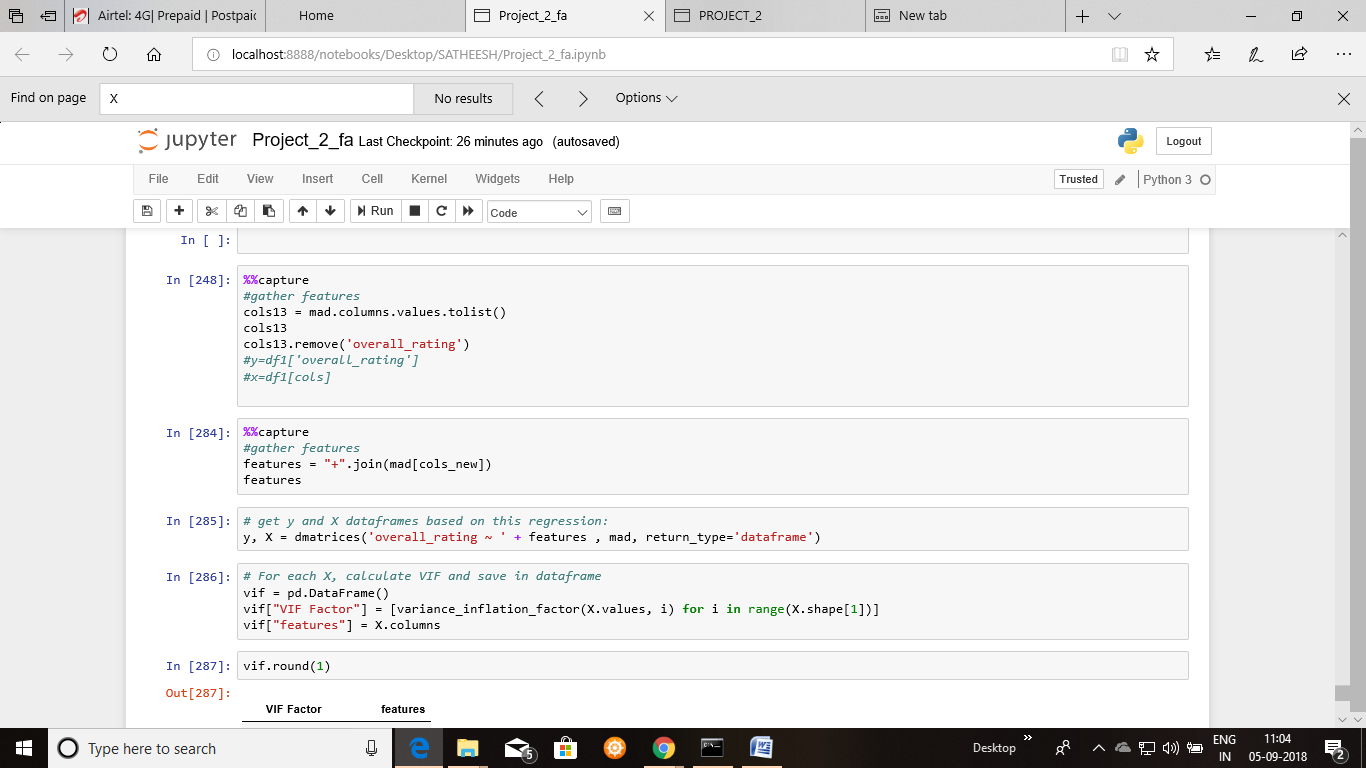
****

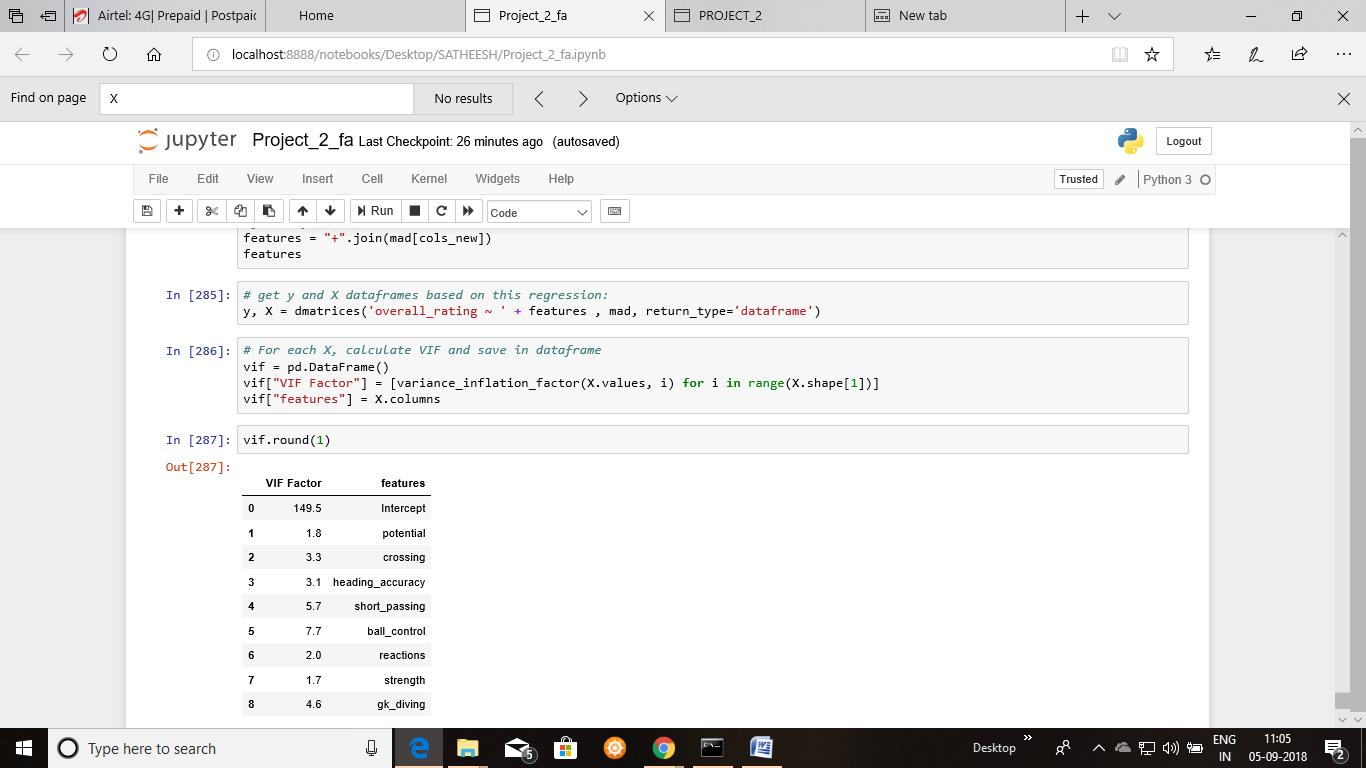
****

****

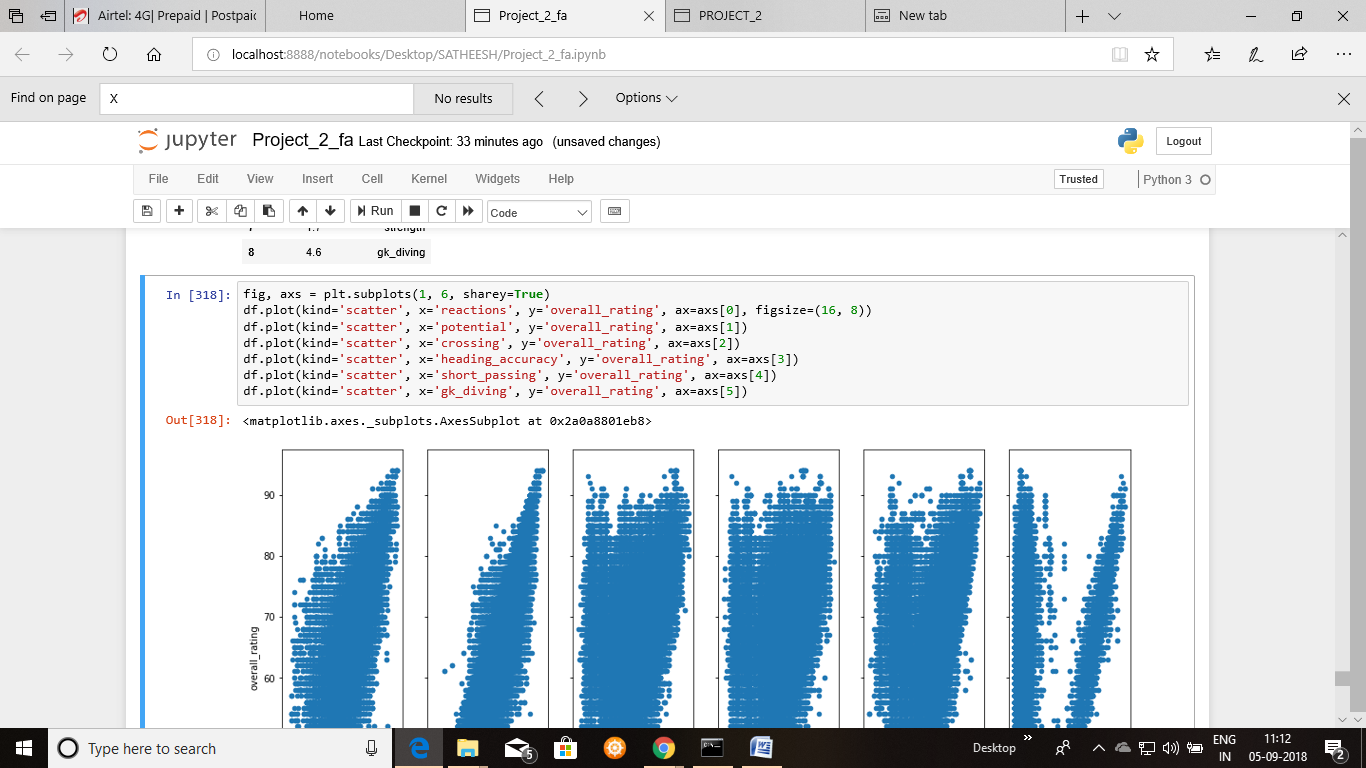
****

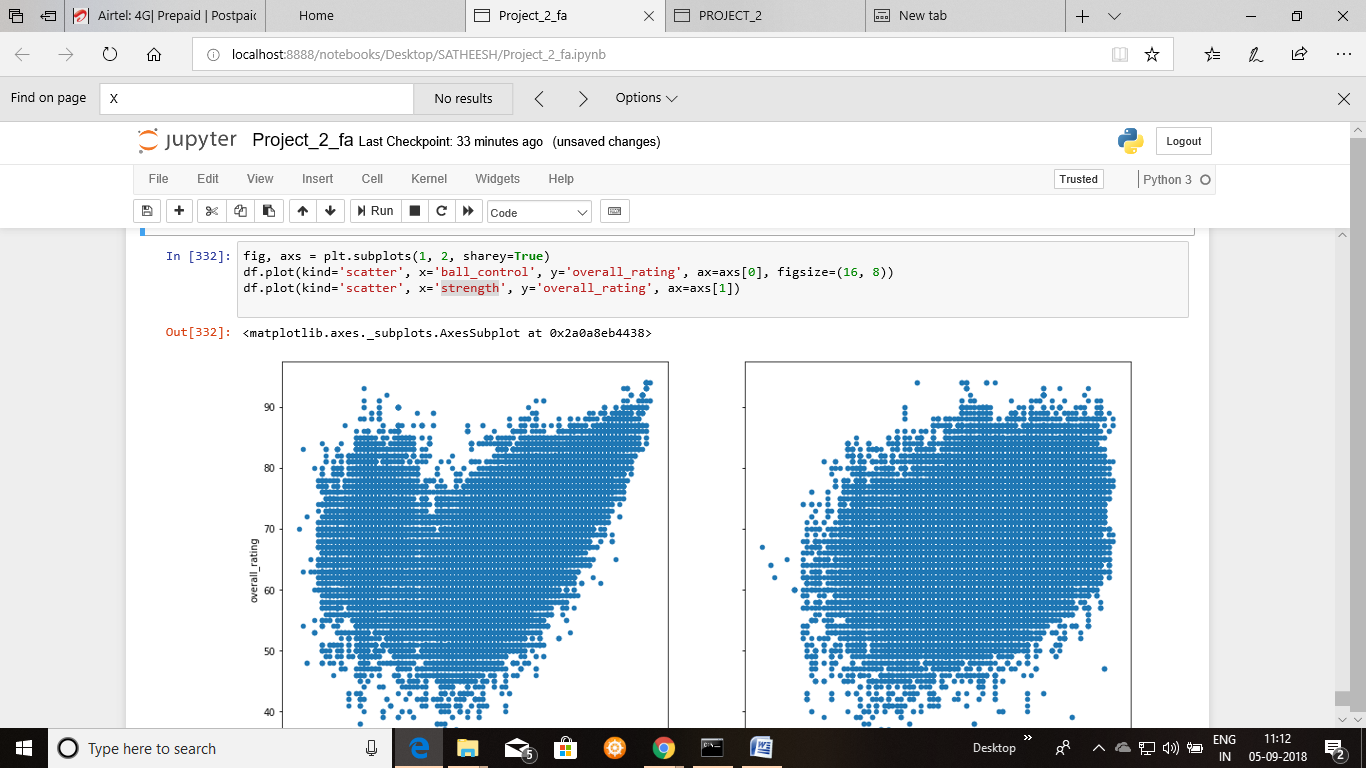
****

****

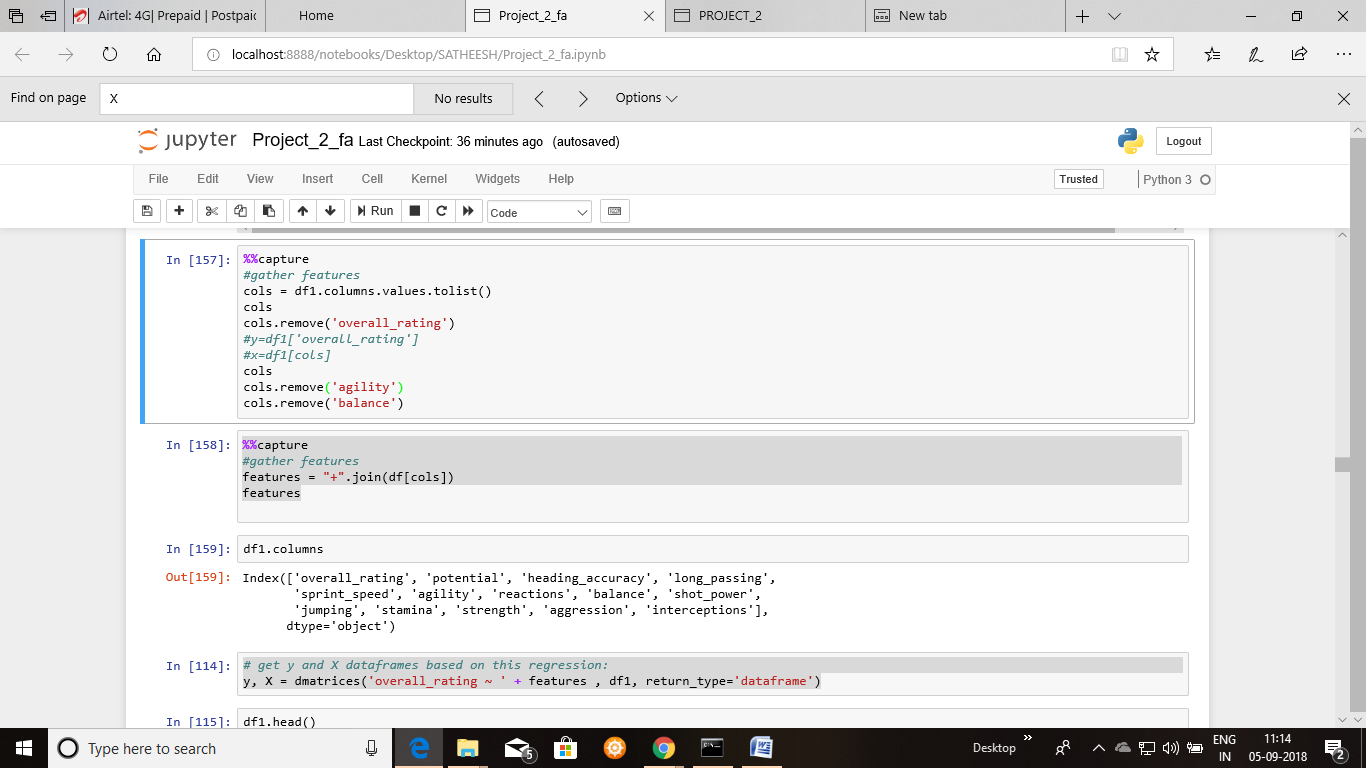
****

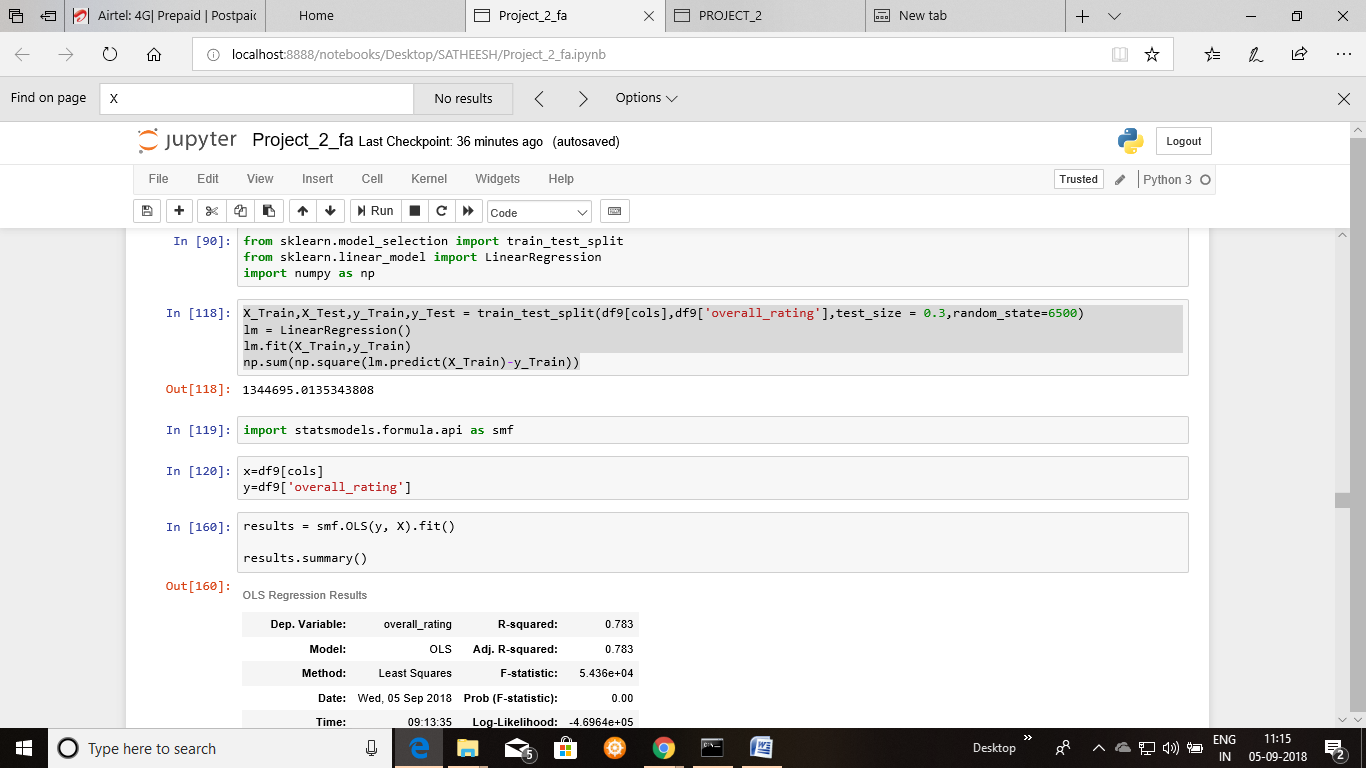
**Visualisations for features:**

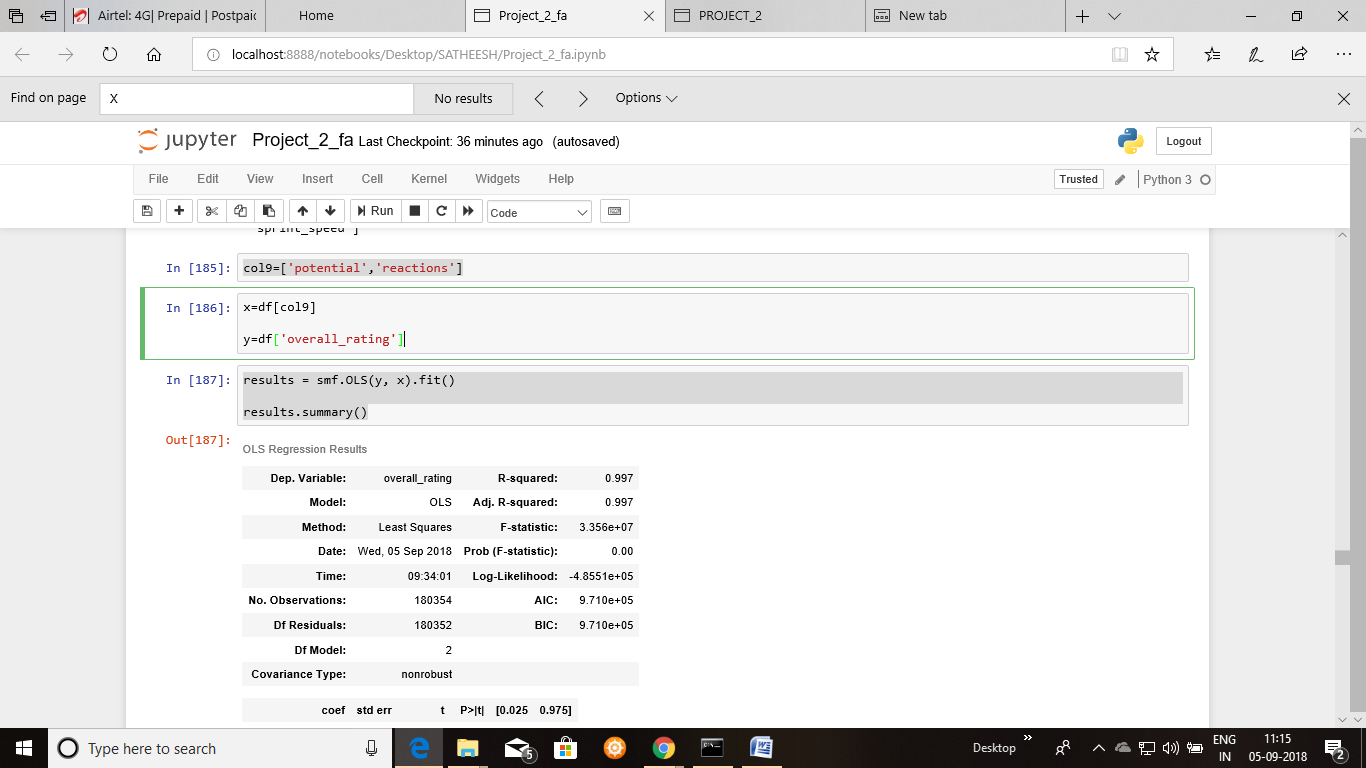
****

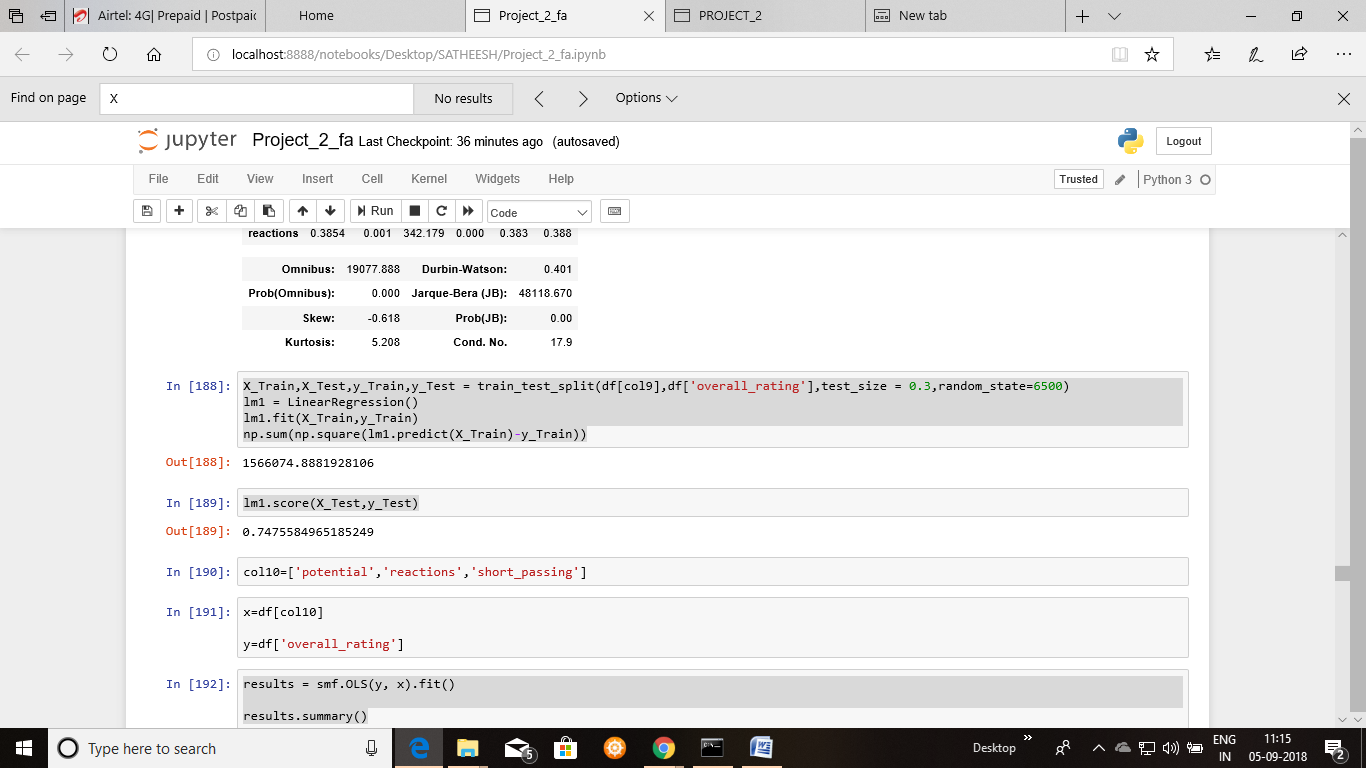
****

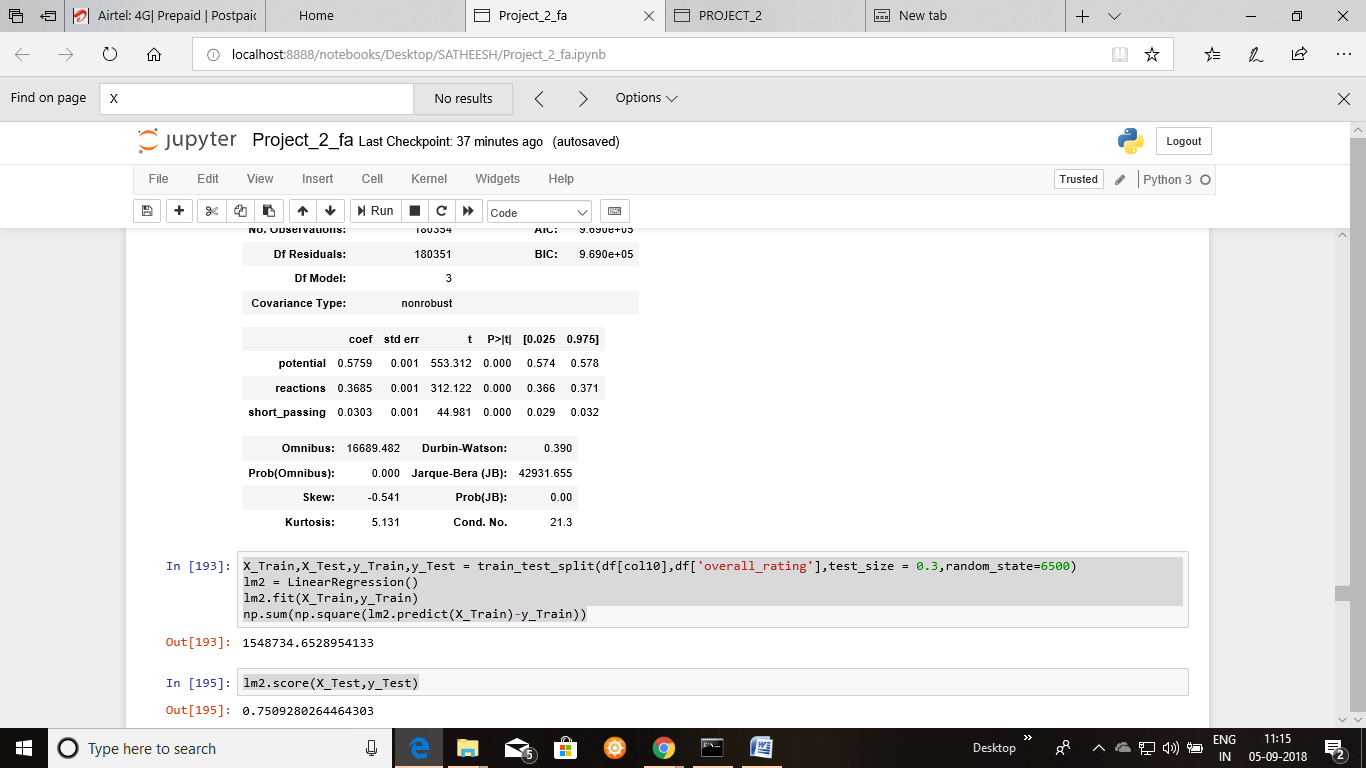
**LINEAR REGRESSION WITH FEATURES ONE BY ONE:**

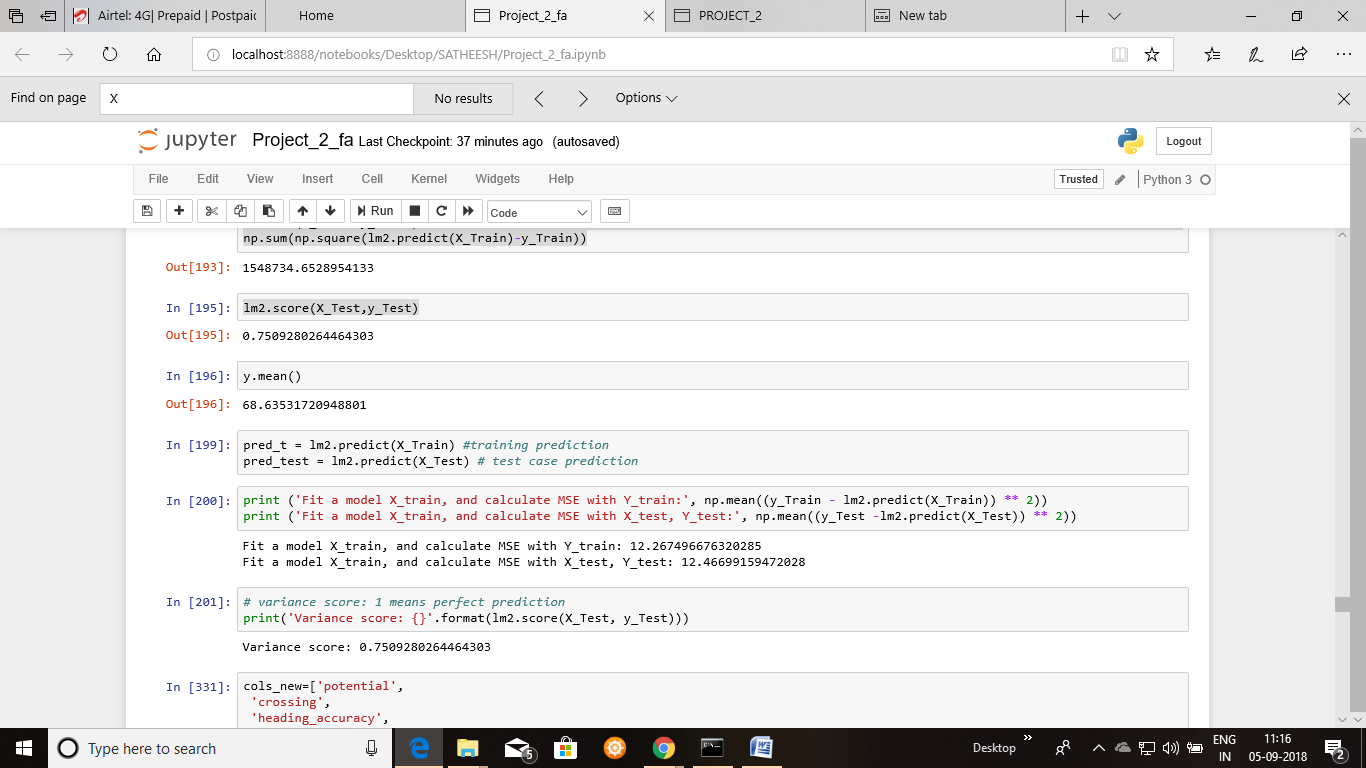
****

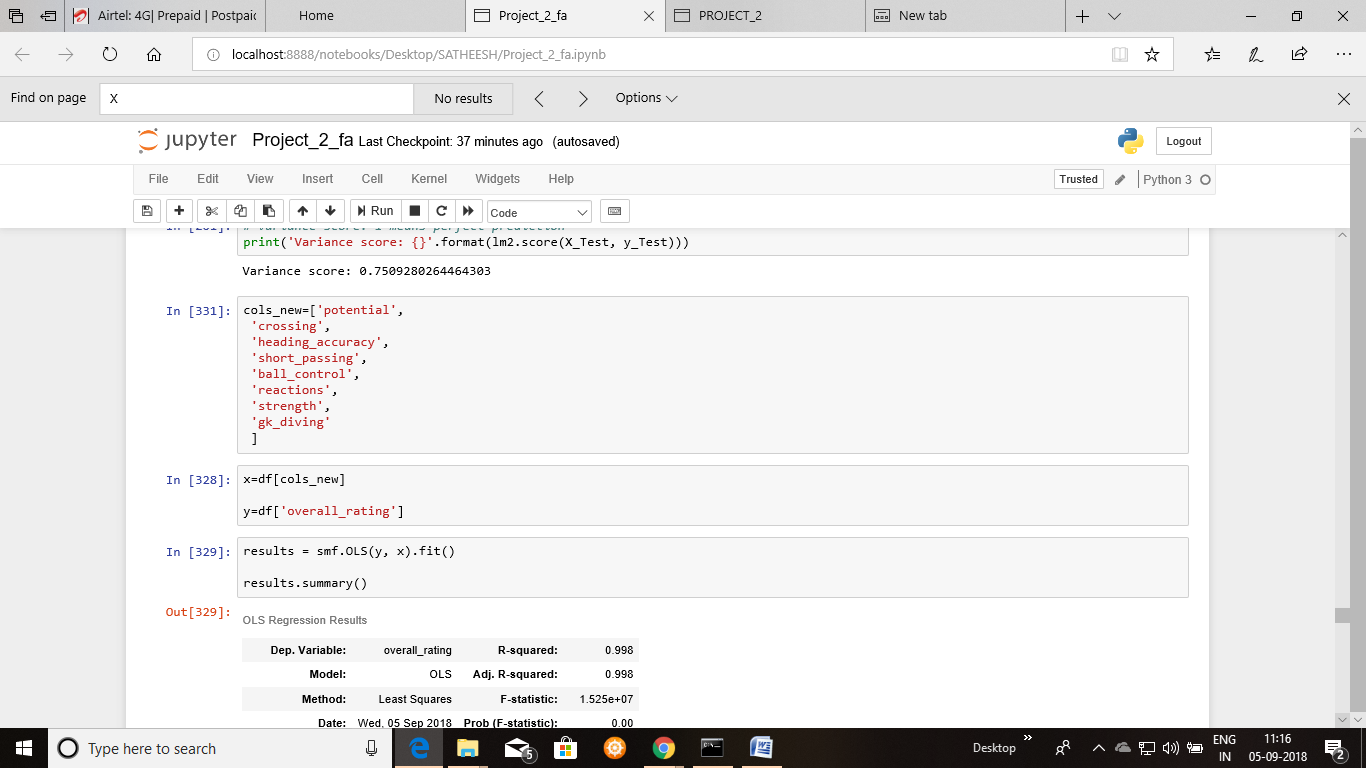
****

****

****

****

****

****

****

**FROM ABOVE ALL ANALYSIS I CHOSE BELOW FEATURES AS FINAL ONE**

**cols\_new=['potential',**

**'crossing',**

**'heading\_accuracy',**

**'short\_passing',**

**'ball\_control',**

**'reactions',**

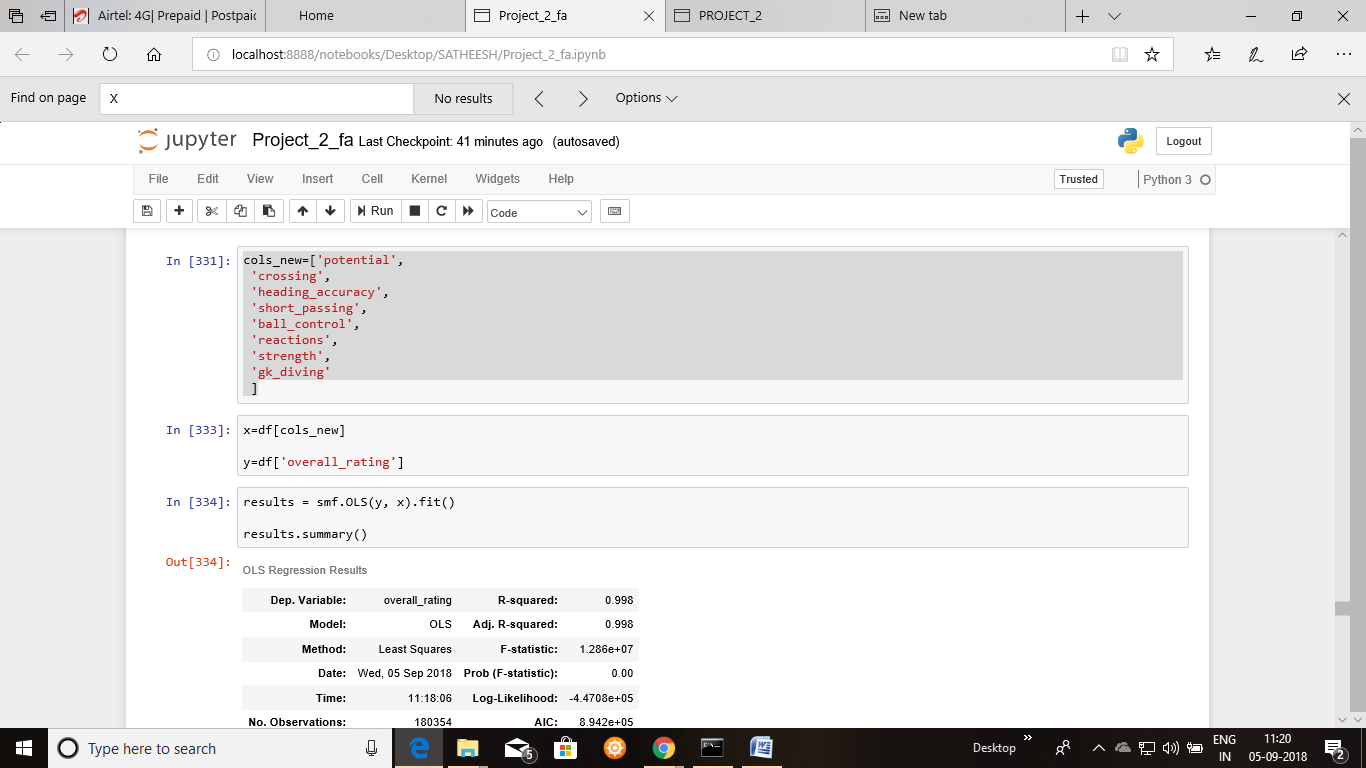
**'strength',**

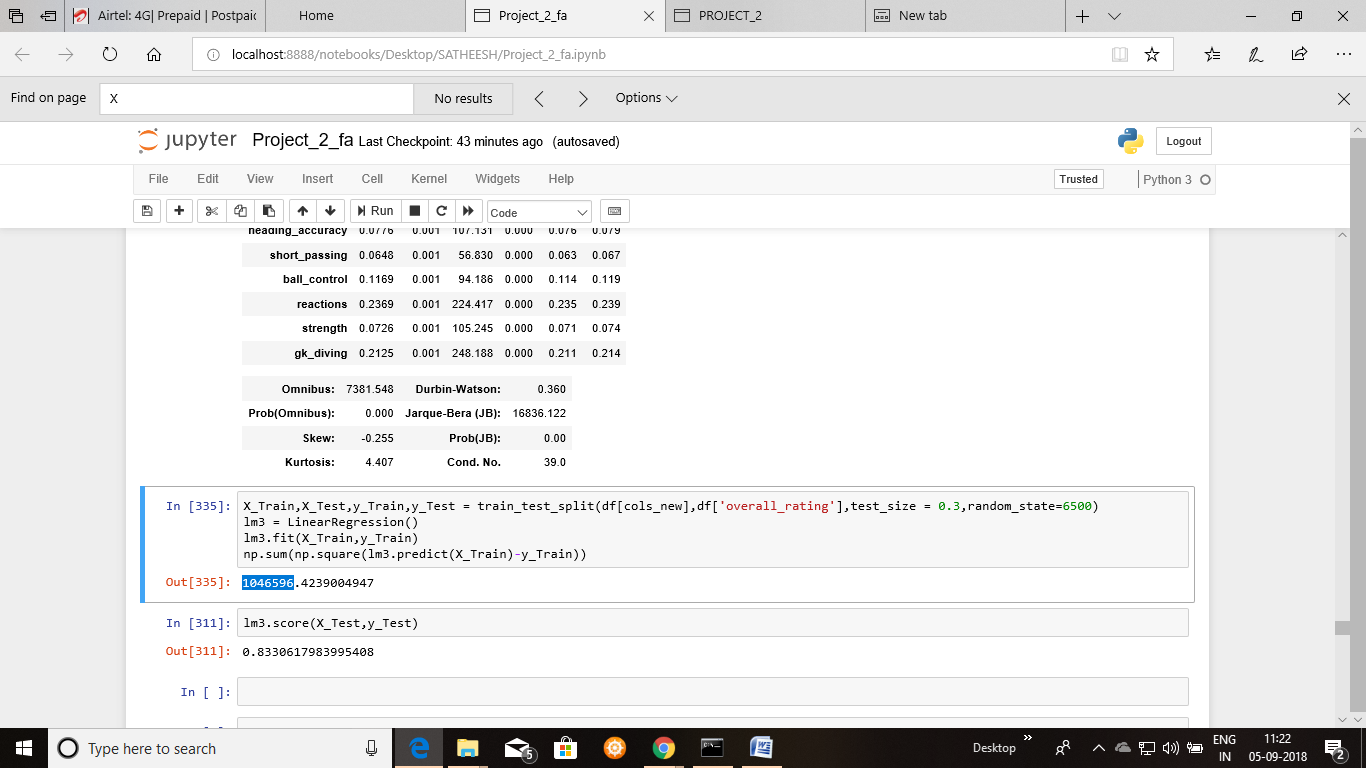
**'gk\_diving'**

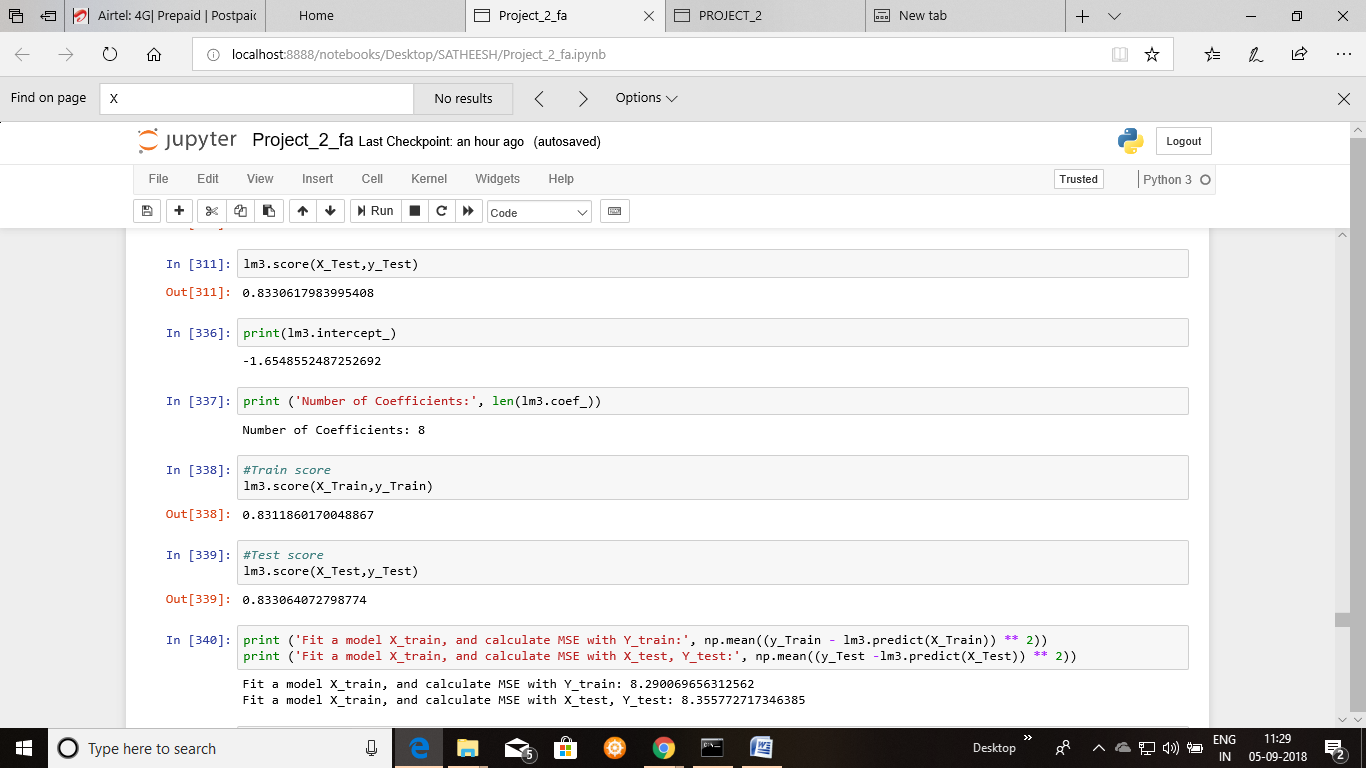
**]**

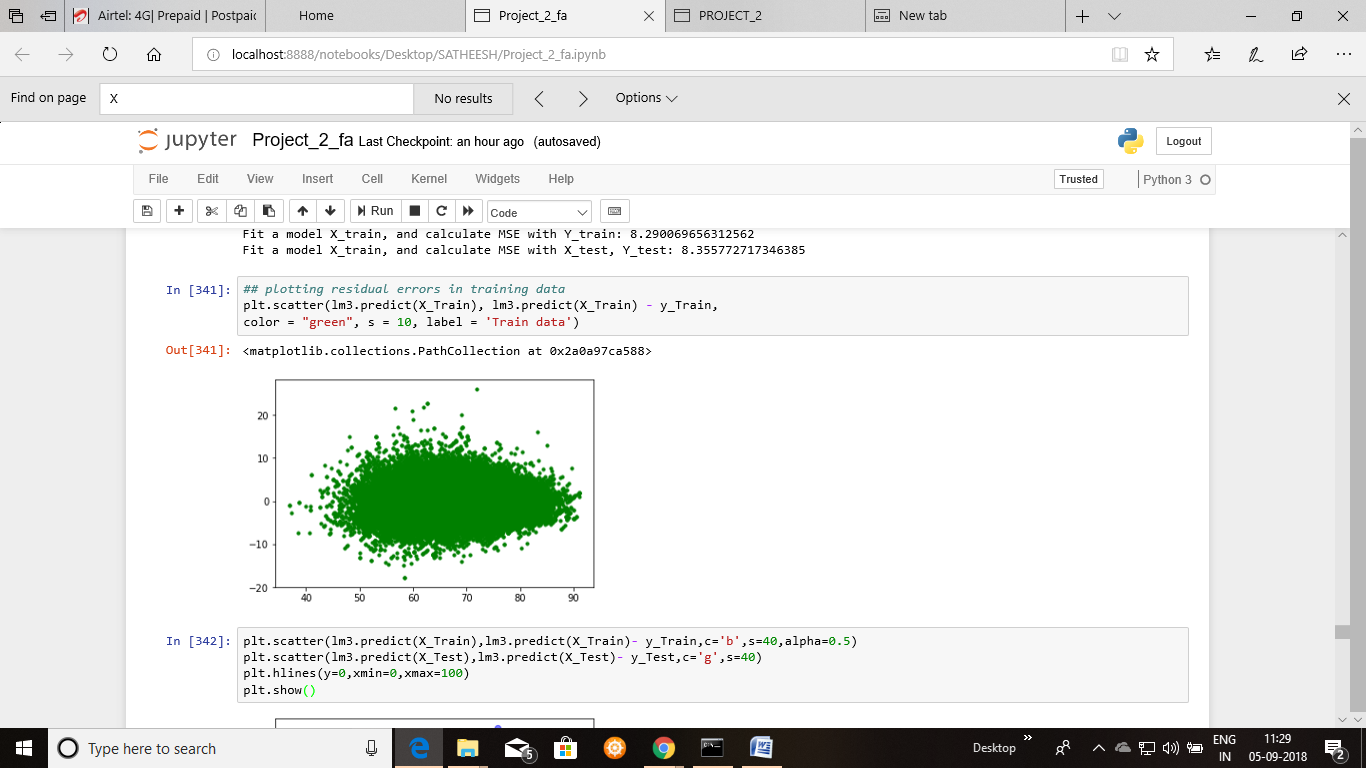
**If i remove the gk\_diving feature,error rate is increase from** 1046596 to 1446596.So i decided to keep this feature

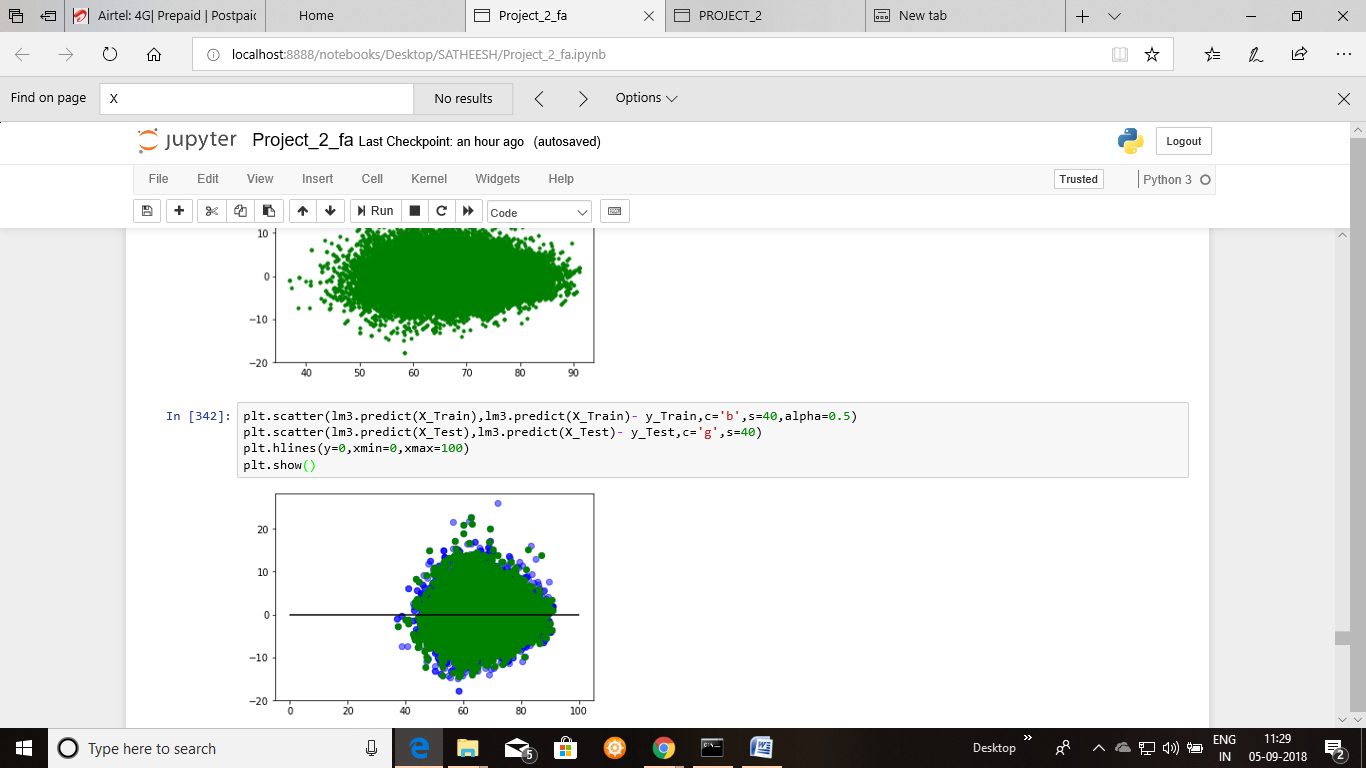
**LINEAR REGRESSION WITH SELECTED FEATURES:**

****

****

****

****

****