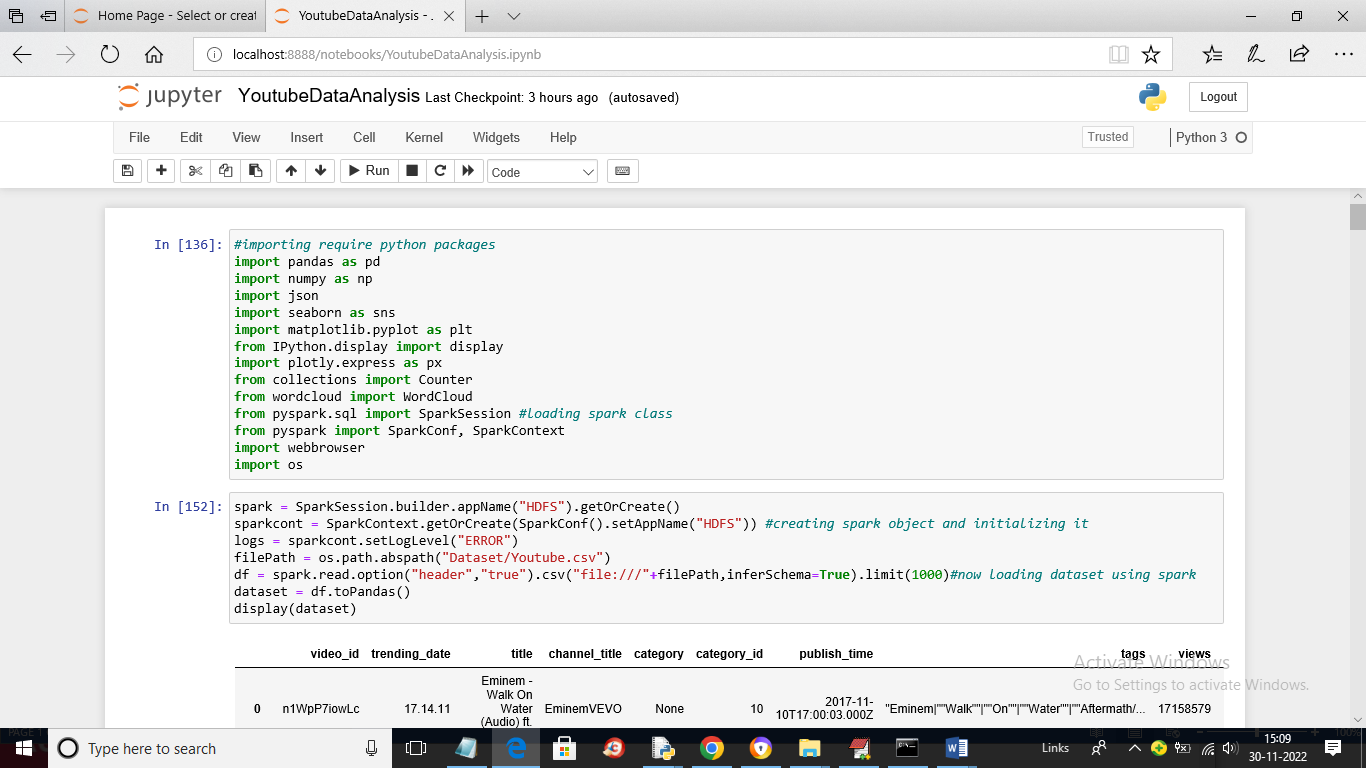
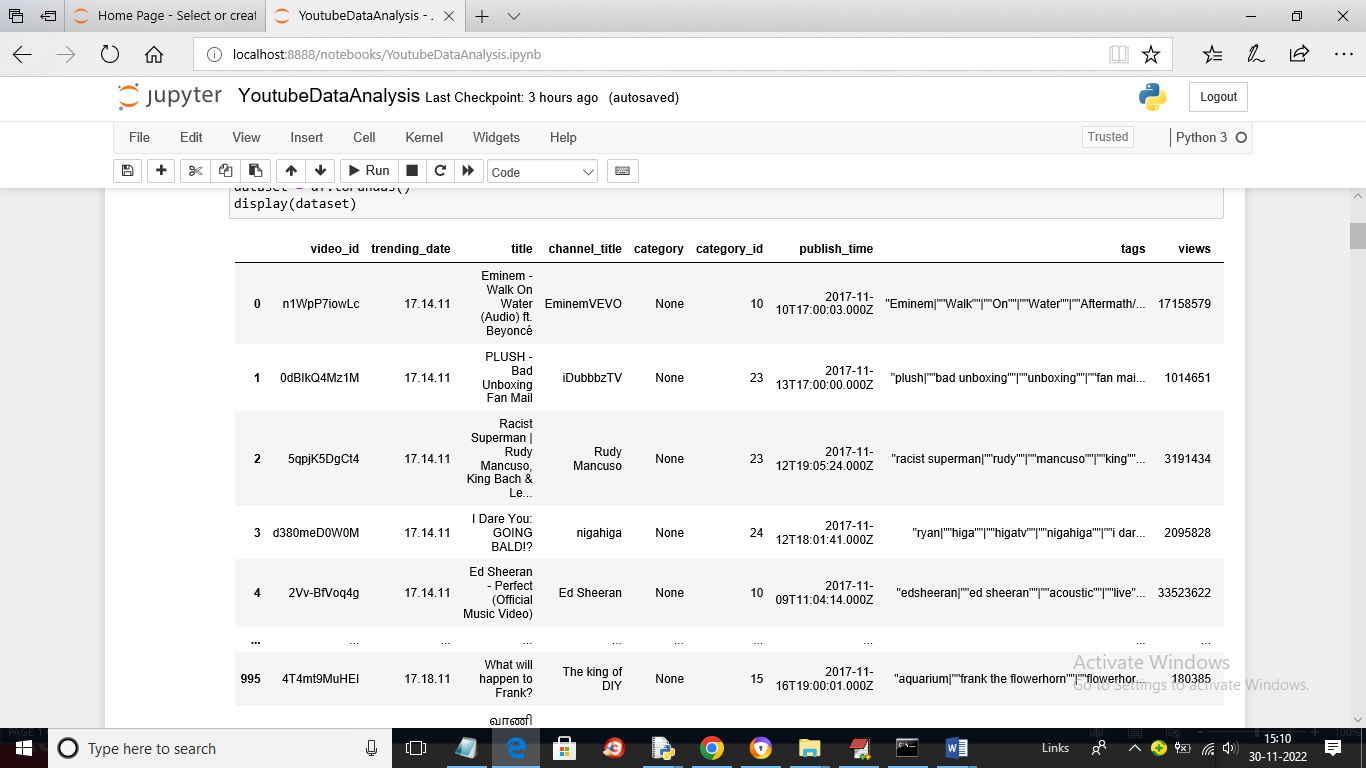
YouTube Data Analysis using Hadoop and Spark

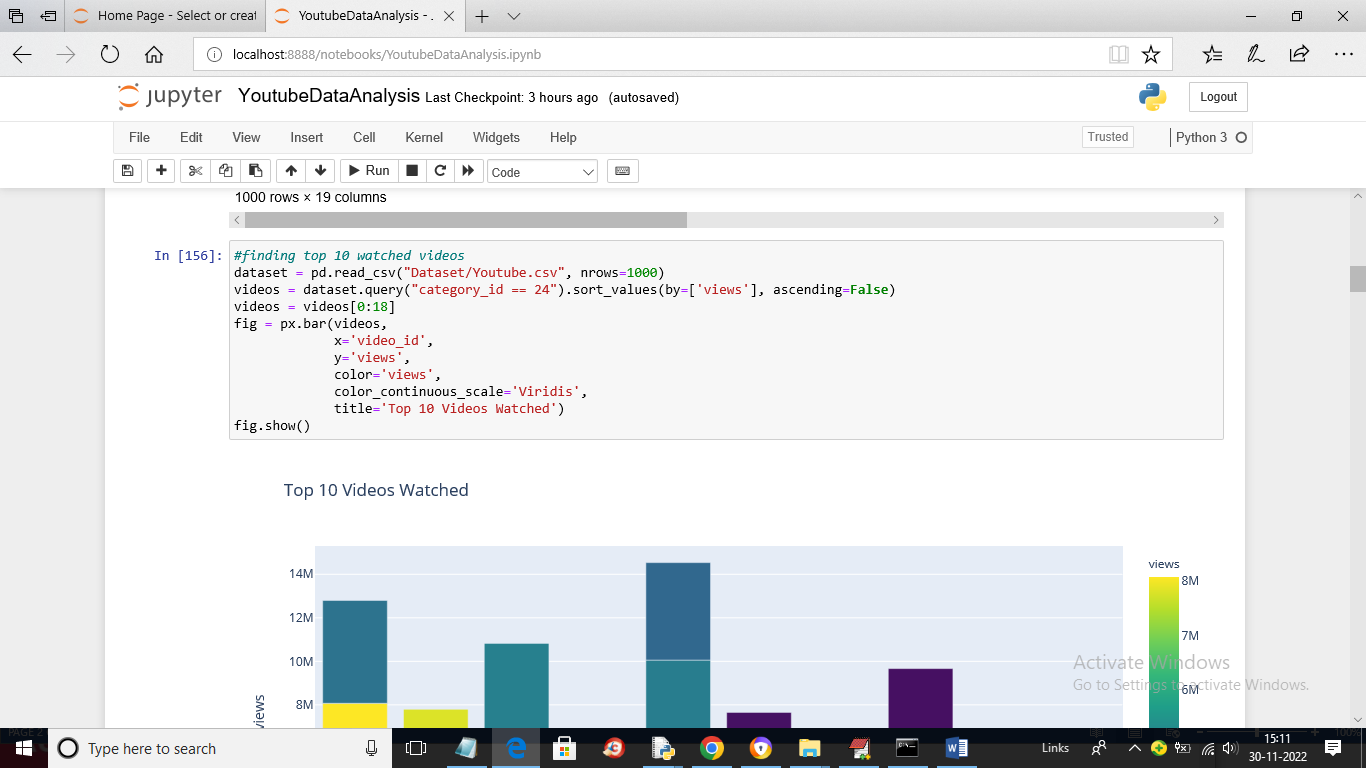
In this project we are using YouTube dataset of more than 250 MB for various analysis such as finding top watch videos, top trending videos, top videos upload in each category. To process huge data we are using SPARK packages and below are the output screen. This project consists of so many graphs so we coded using JUPYTER notebook

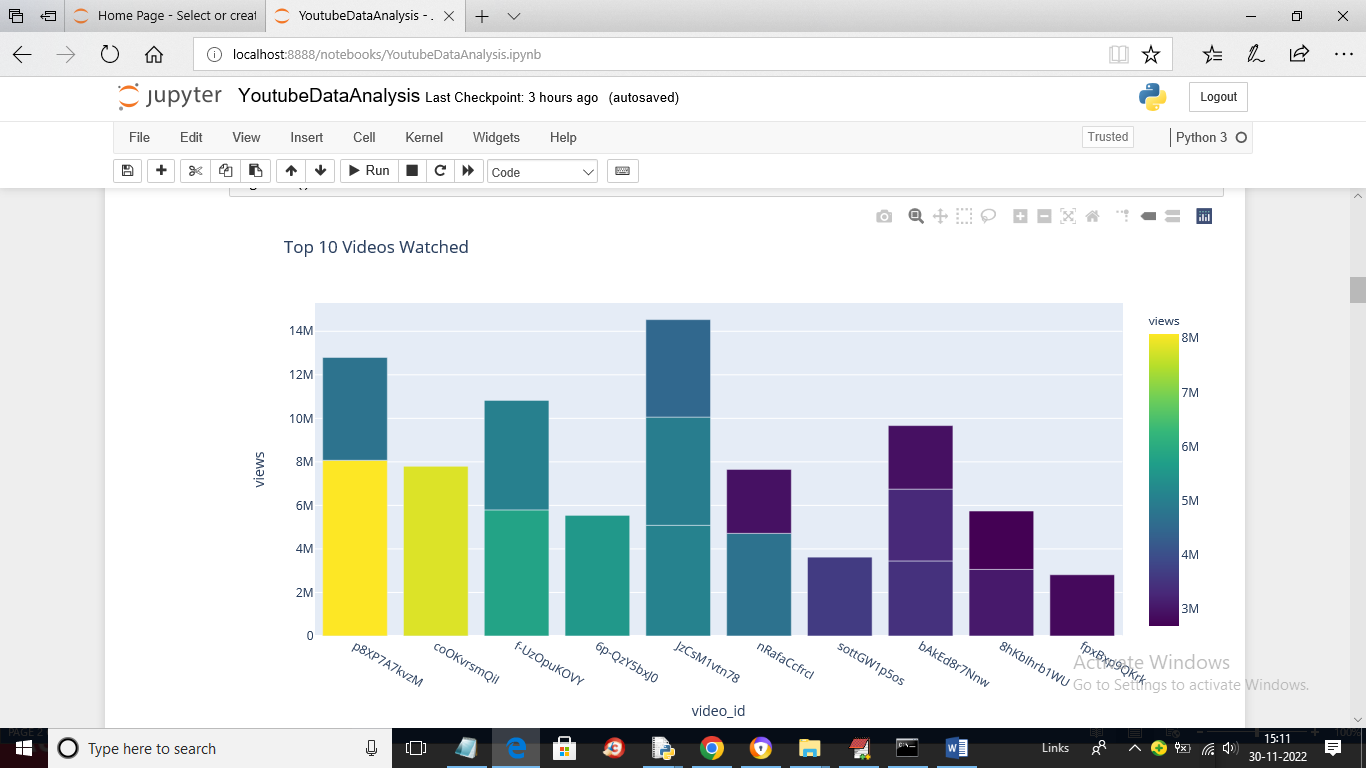


In above screen we are importing require python packages and then using SPARK class we are loading dataset and then in below screen we are displaying loaded dataset values

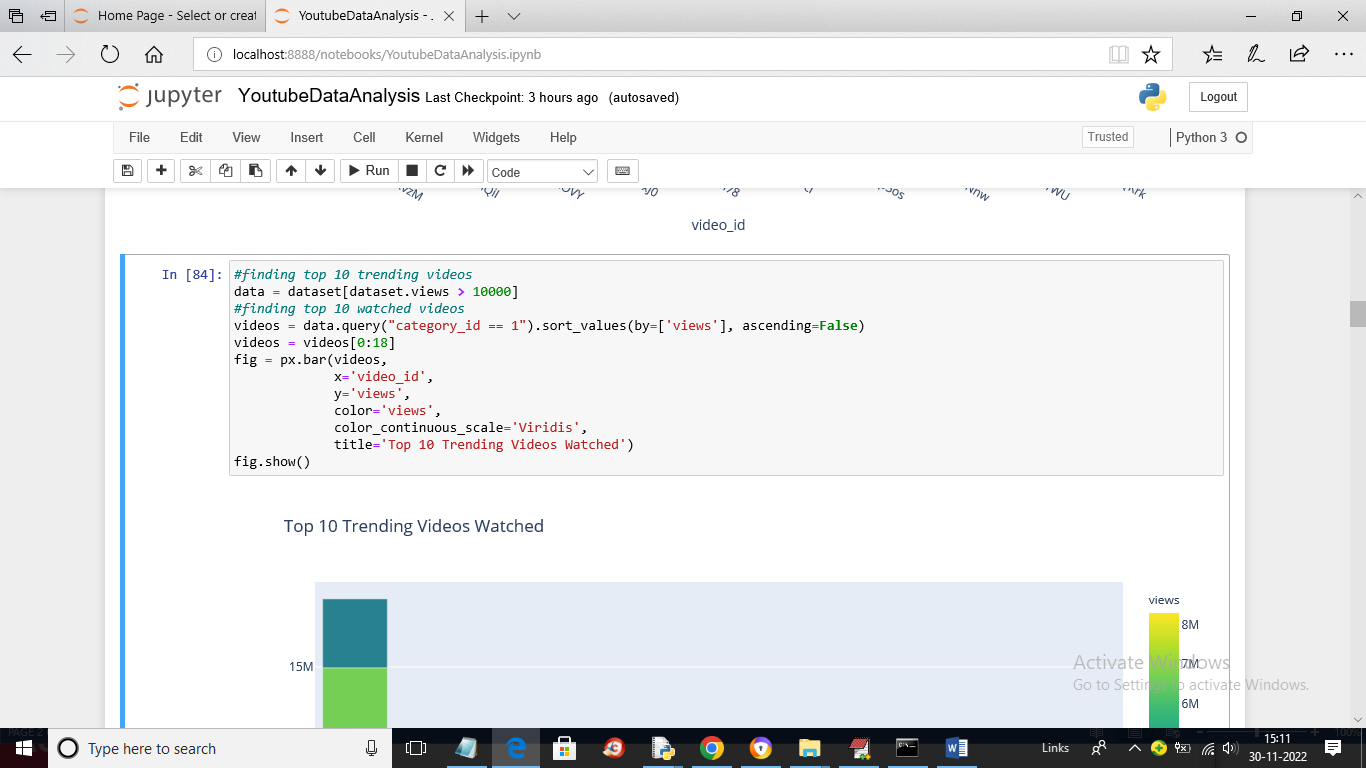


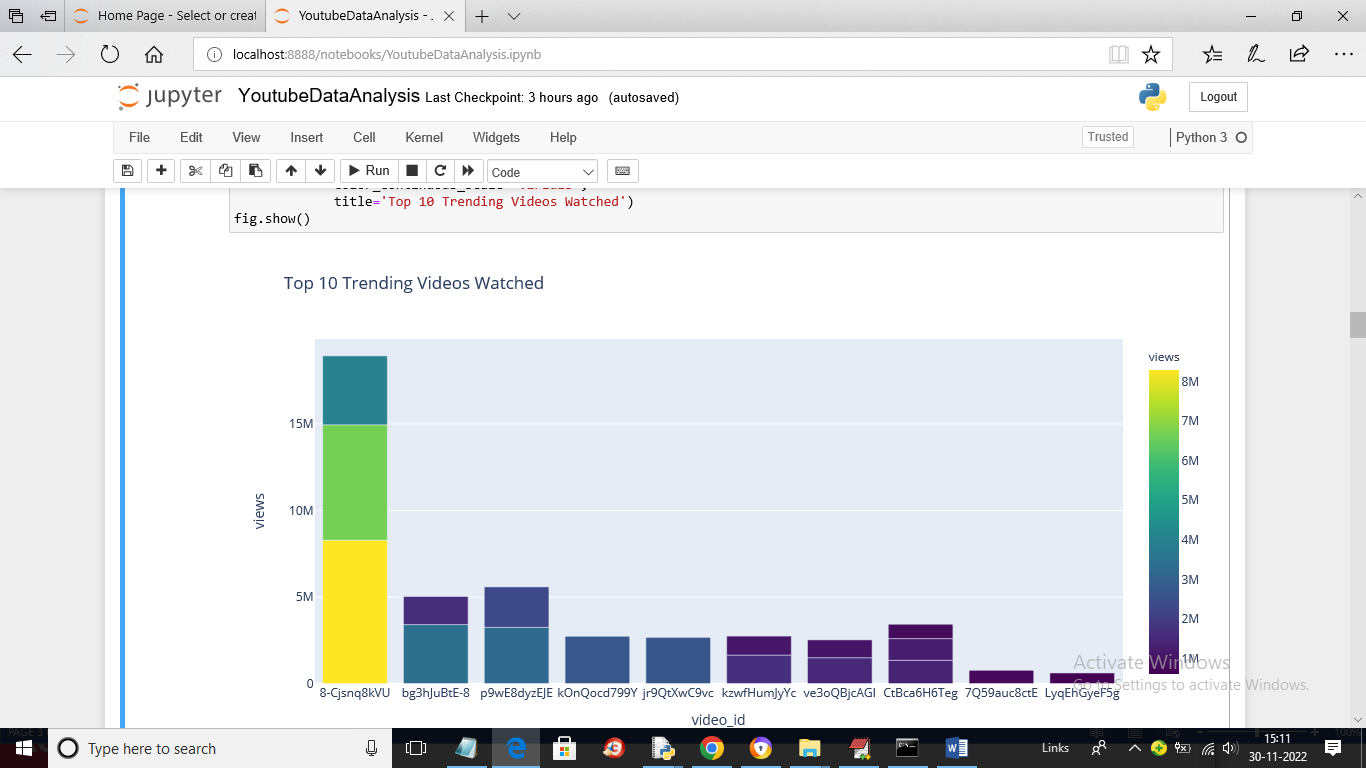
In above screen dataset loaded and displaying



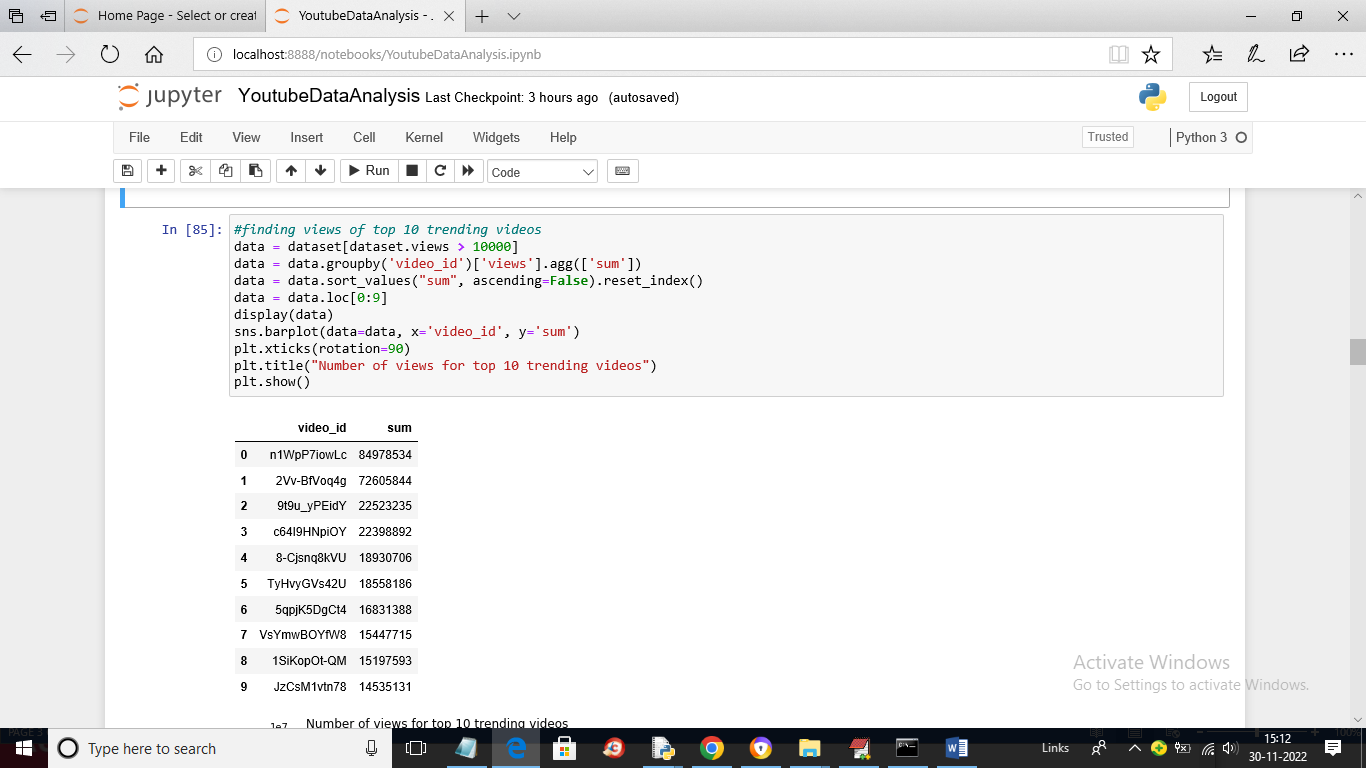


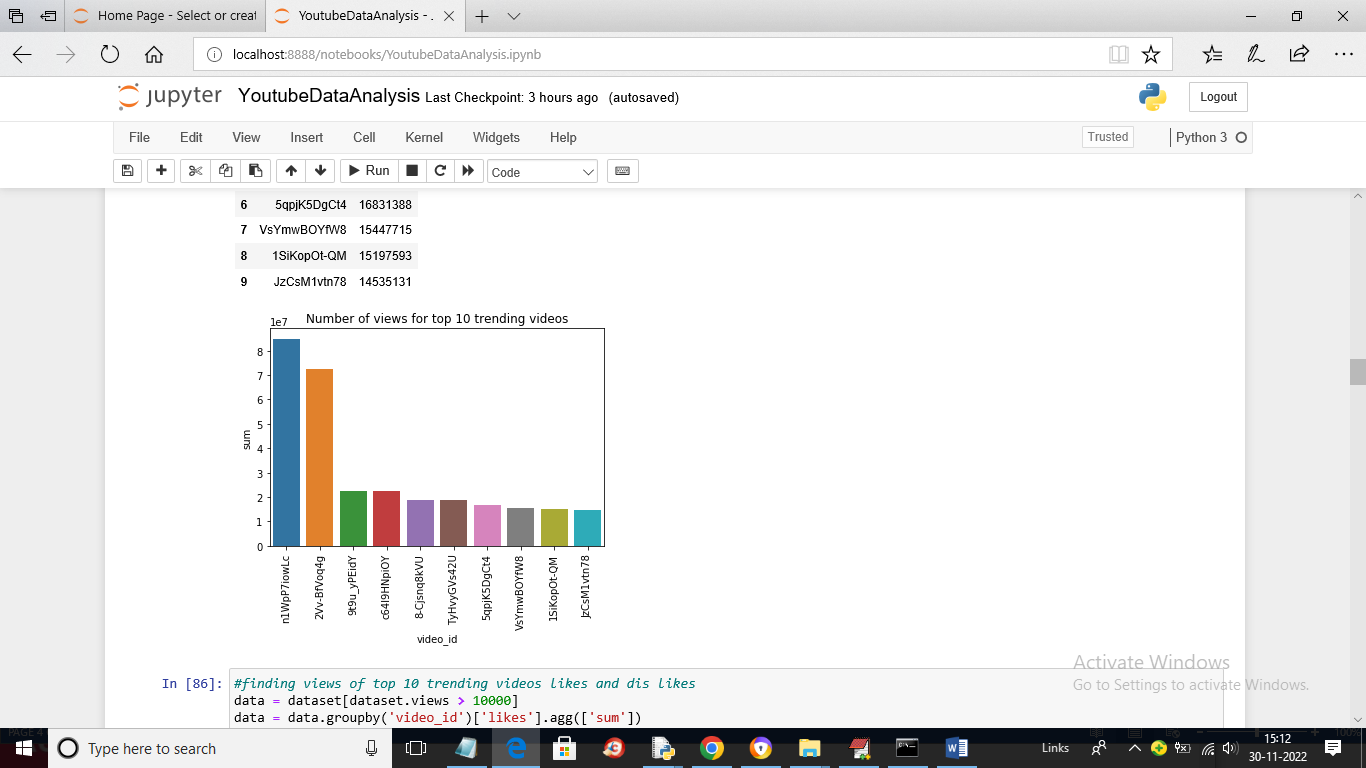
In above 2 screen we are showing code and output of top 10 watched videos where x-axis contains video ID and y-axis represents COUNT



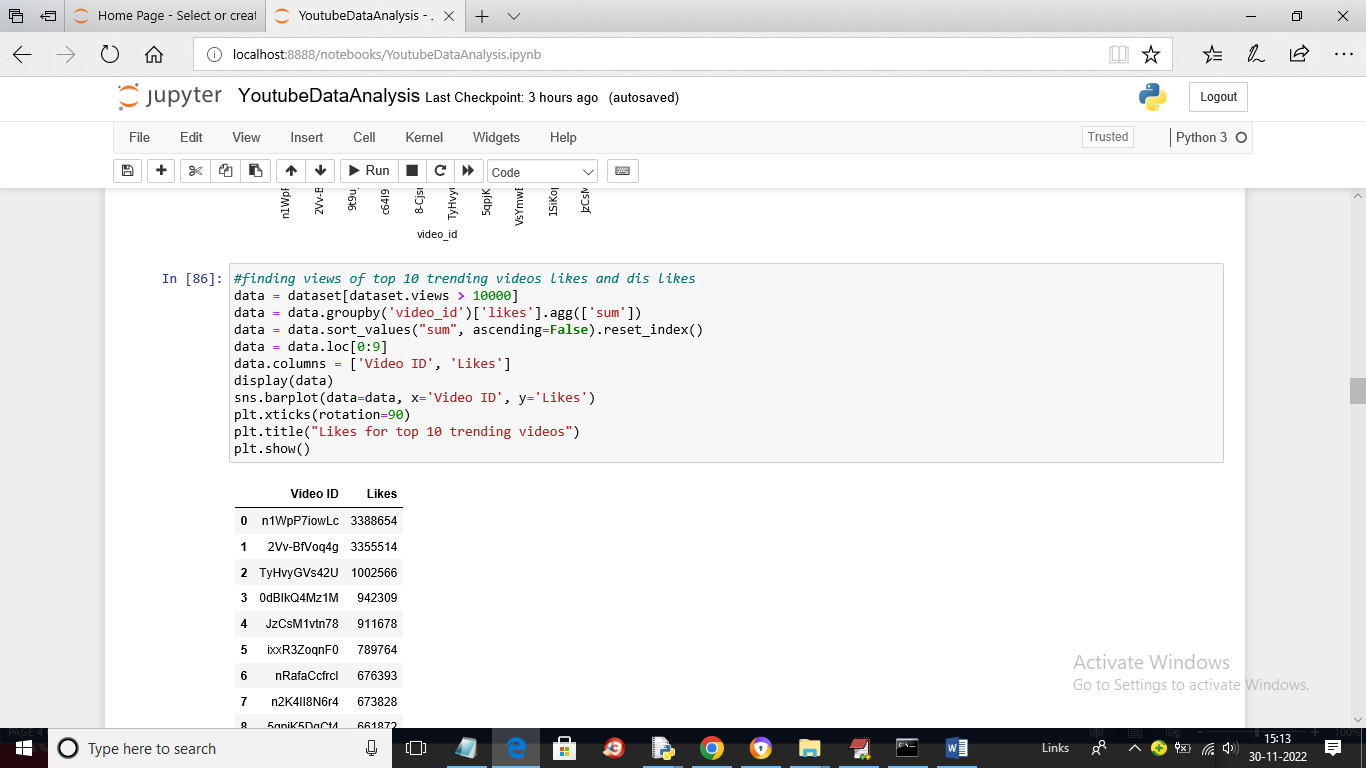


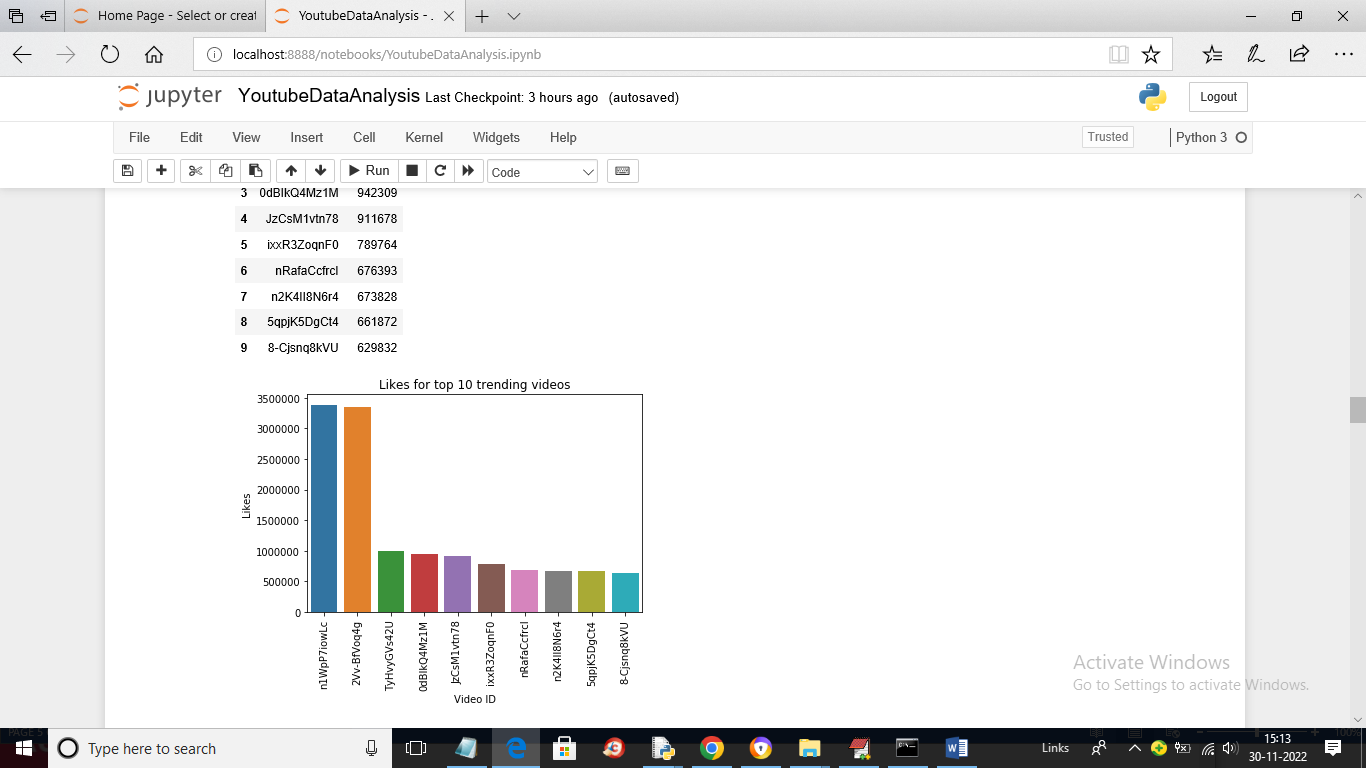
In above screen we are showing TOP 10 trending videos



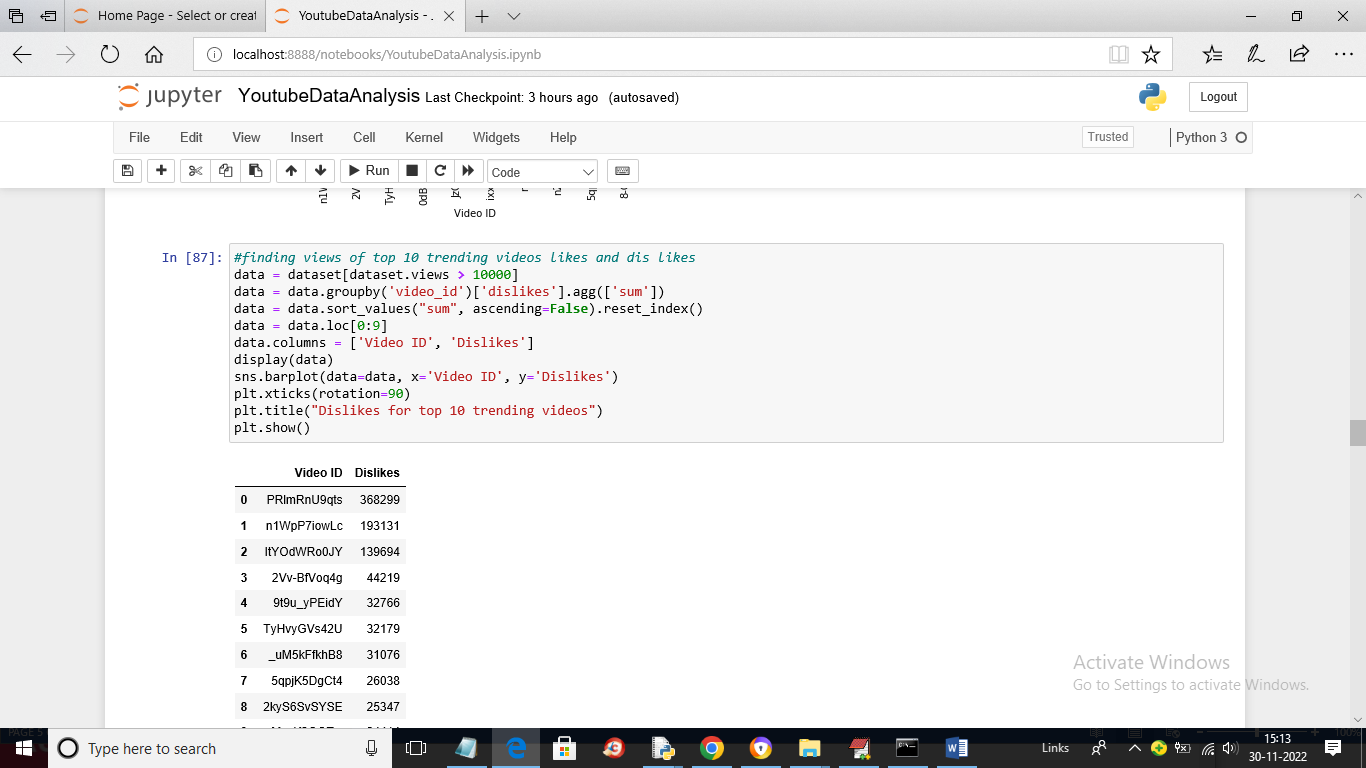


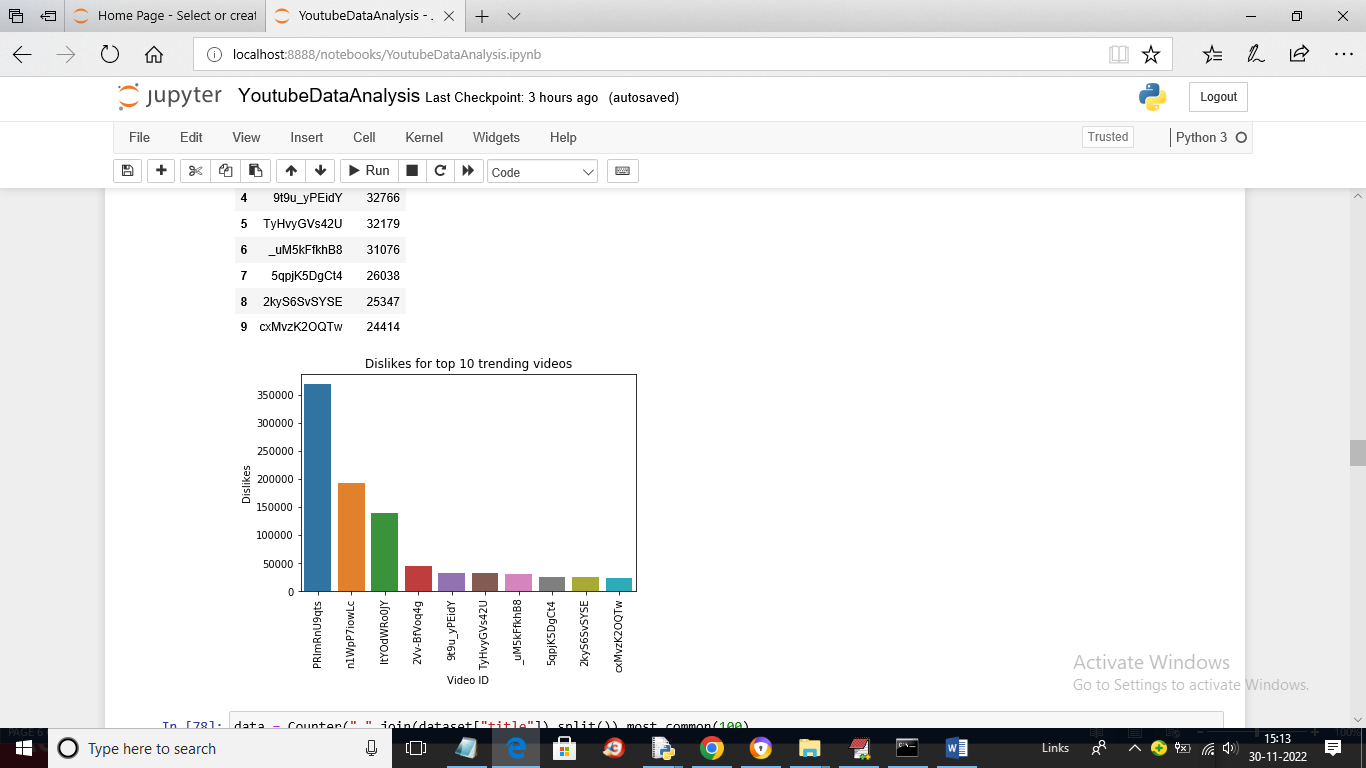
In above two screens we are finding views of top 10 trending videos



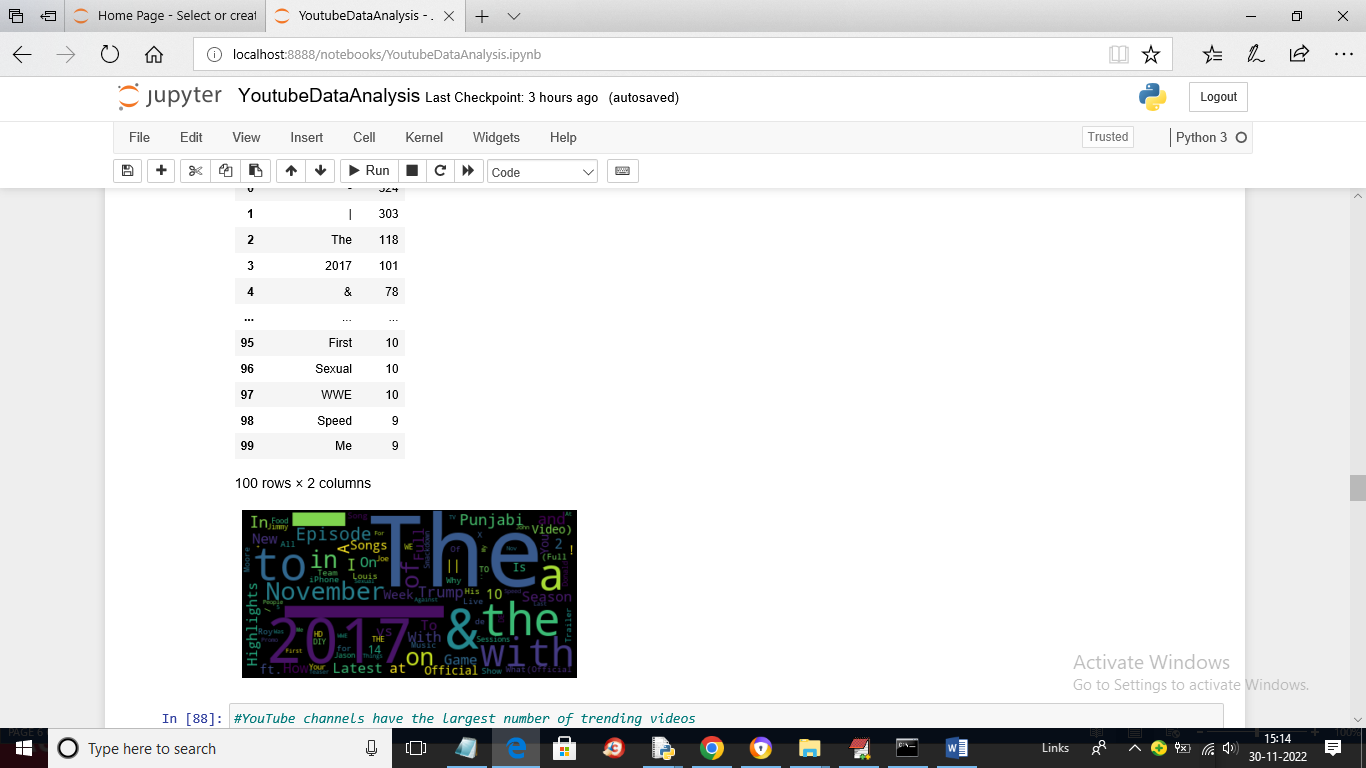


In above screen we are showing graph and code for top 10 trending Videos LIKES

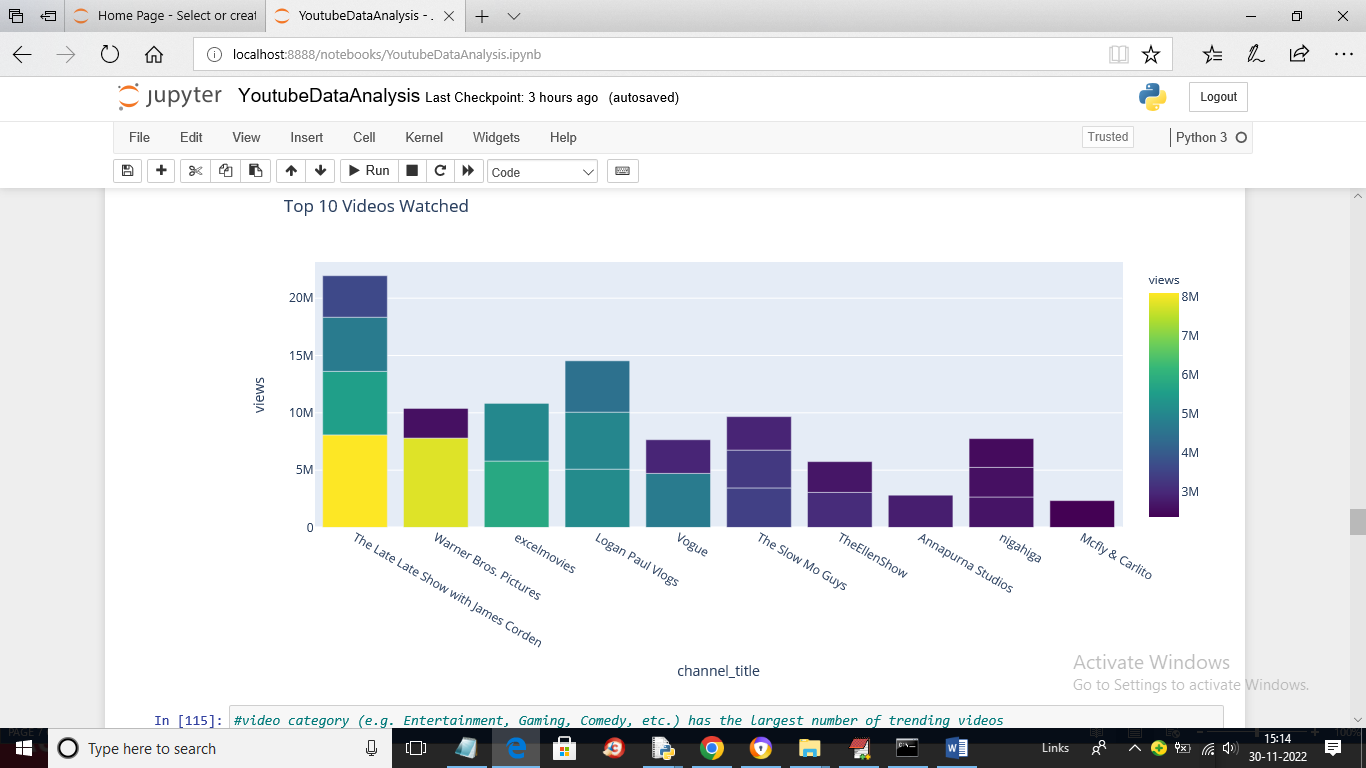




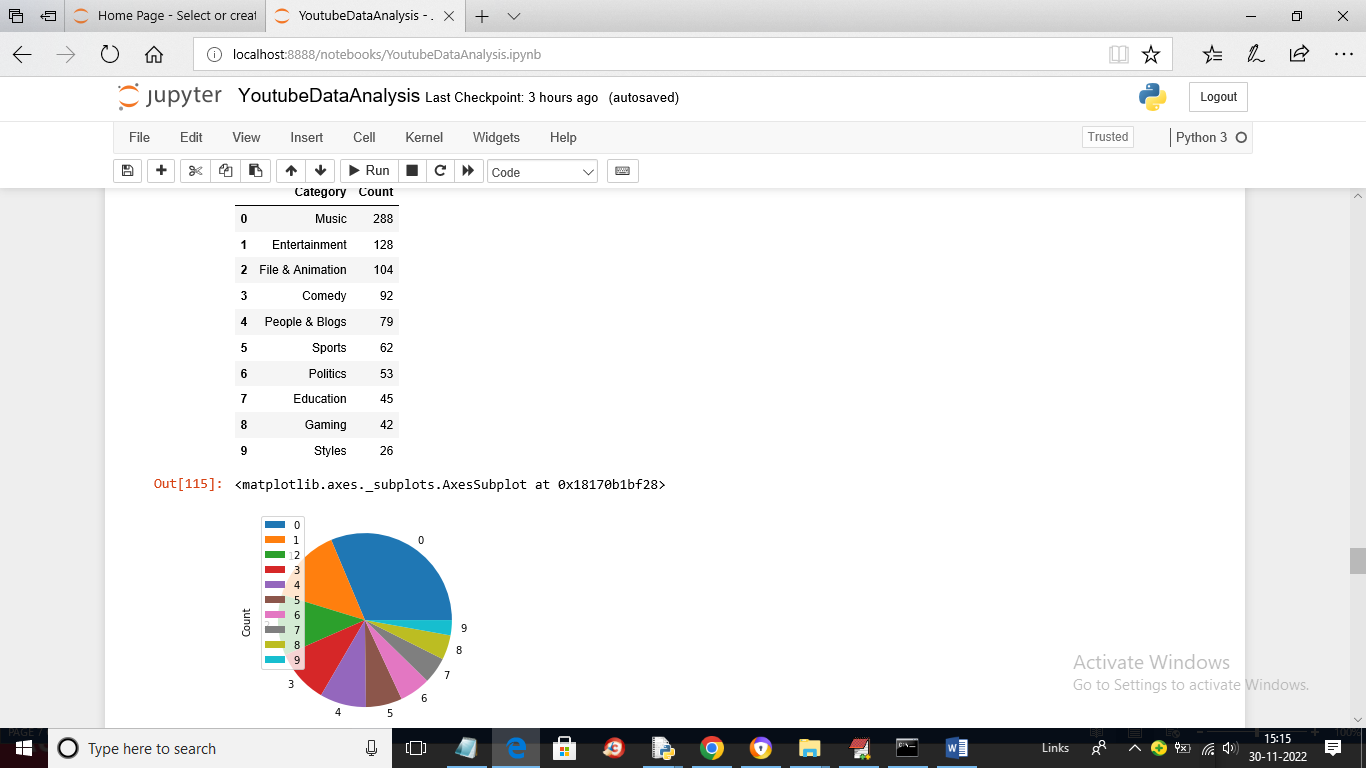
In above screen we are plotting DISLIKES graph for trending videos



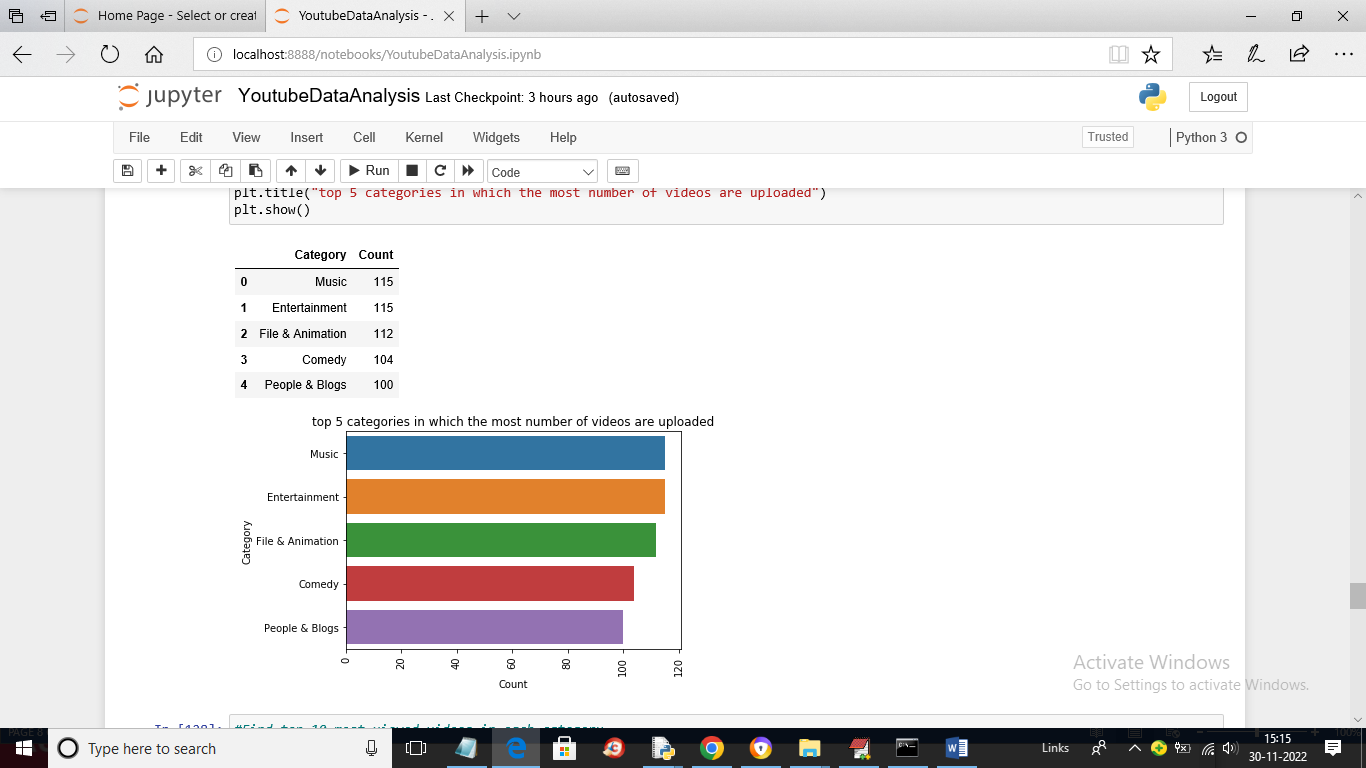
In above screen we are plotting and displaying count of most common words



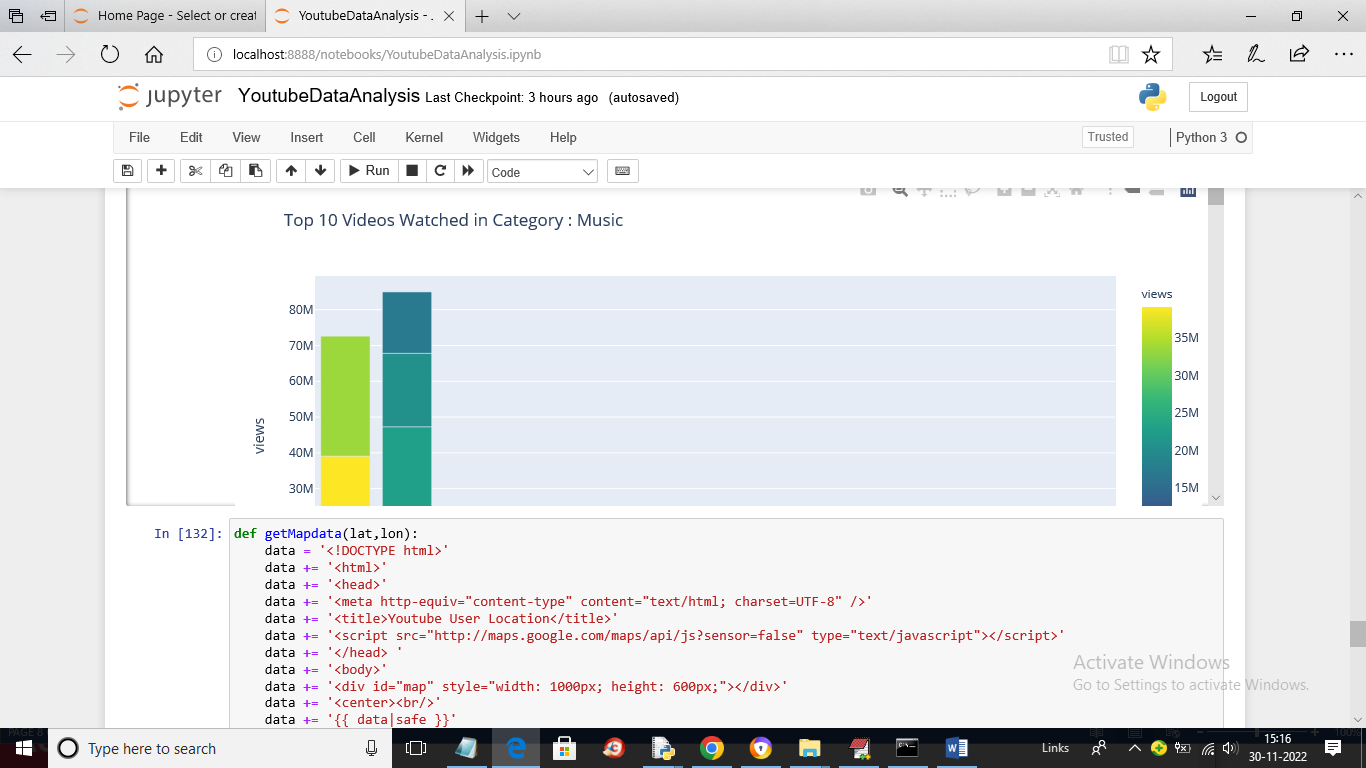
In above graph we are showing TOP 10 channels watched



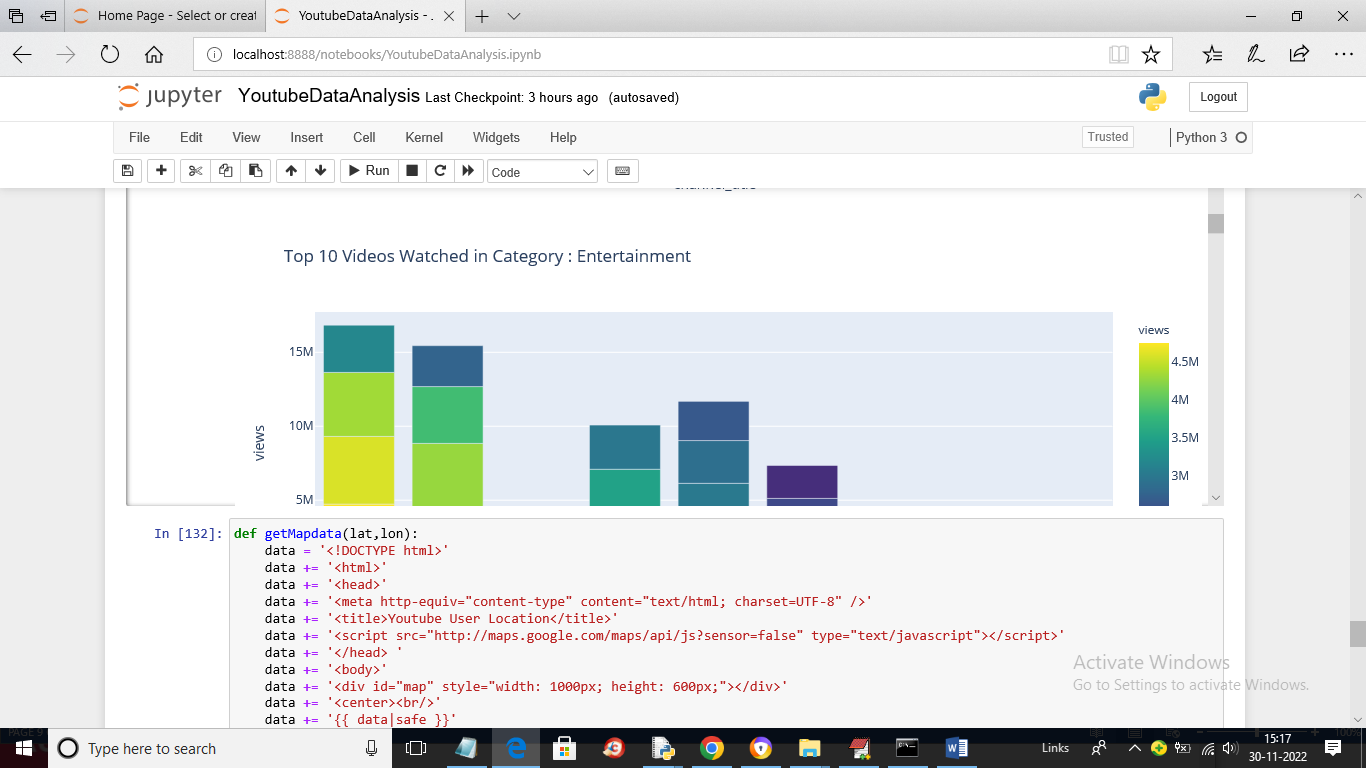
In above graph we are showing largest trending videos based on category



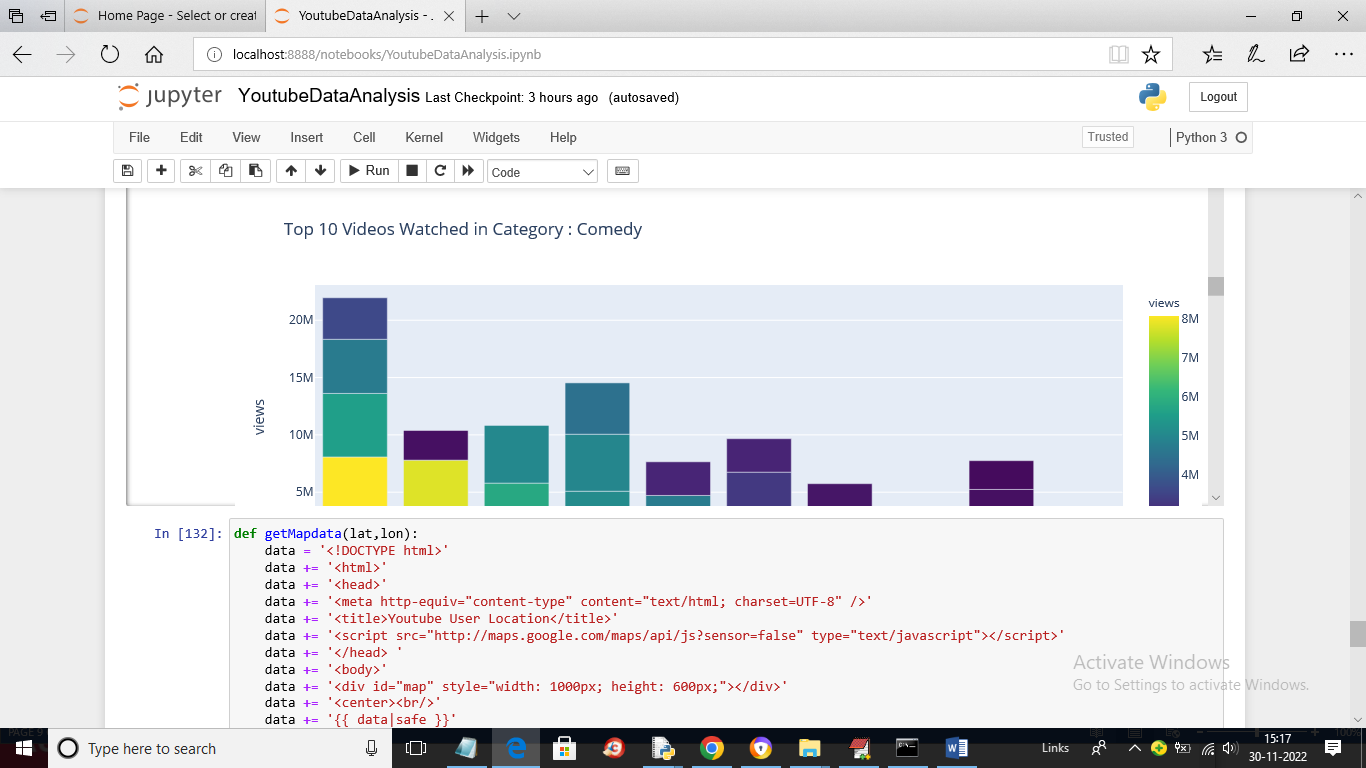
In above screen showing top 5 categories videos uploaded



In above graph showing Top 10 videos watch in music category

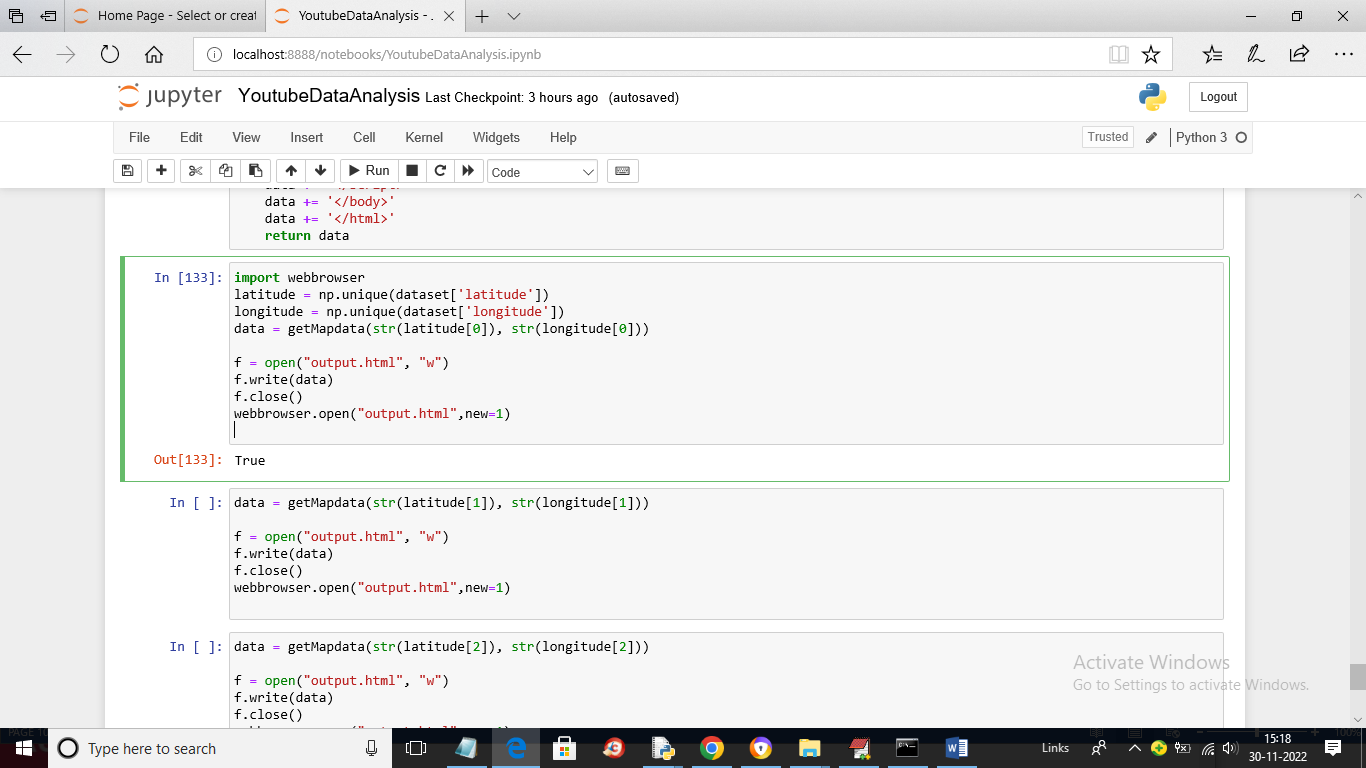


Above graph for Entertainment category



Above graph for comedy category and similarly you can see graph for each category

In below screen run block to view user location in MAP



In above screen run each block to view user location MAP like below screen

