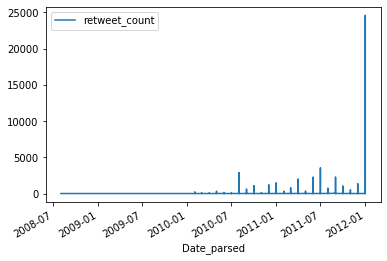
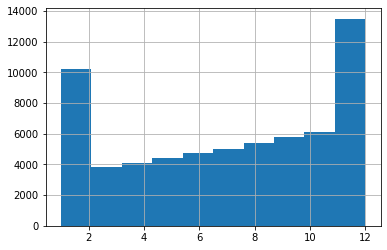
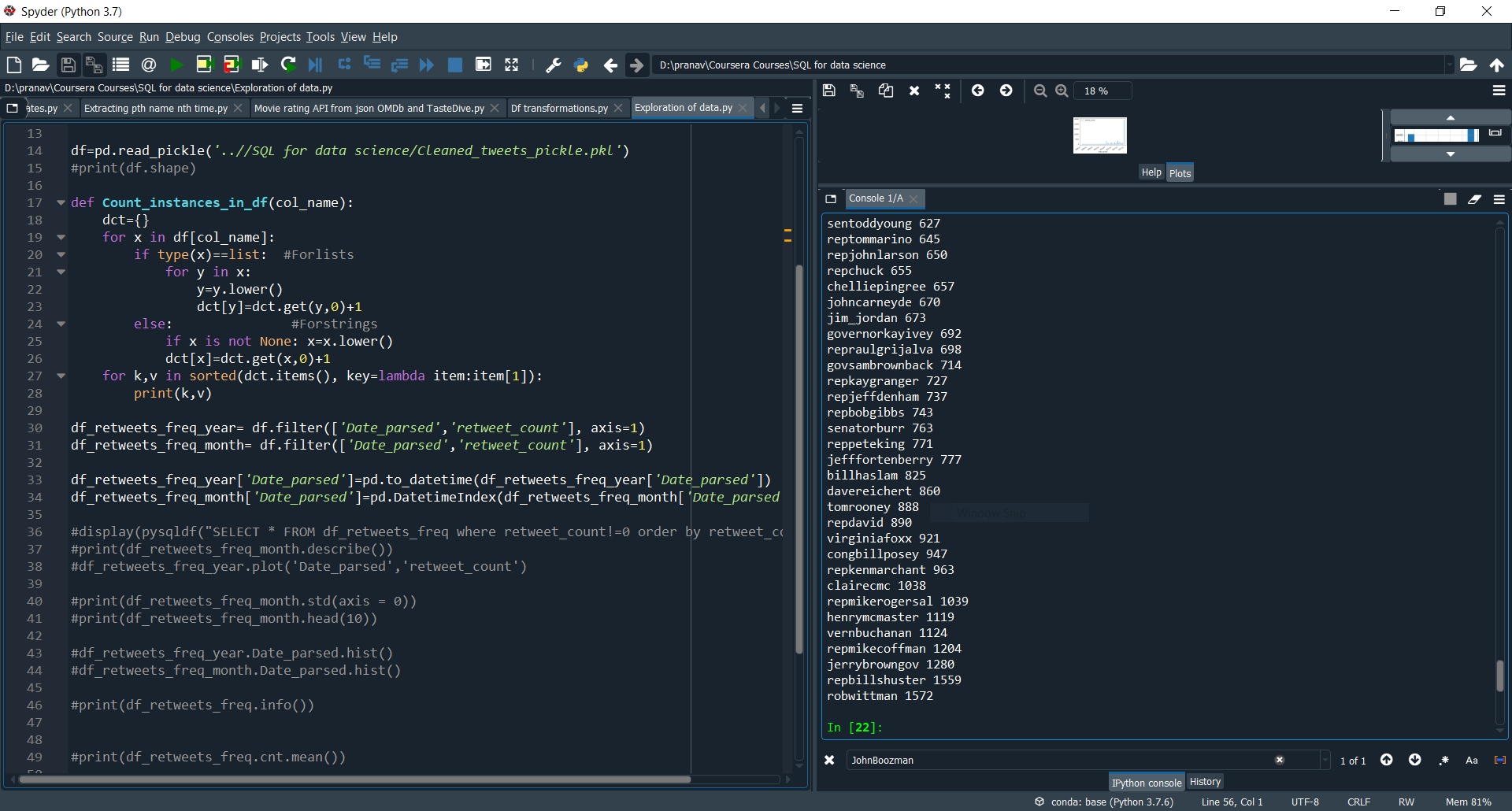
1. Provide a summary of the different descriptive statistics you looked at and WHY.

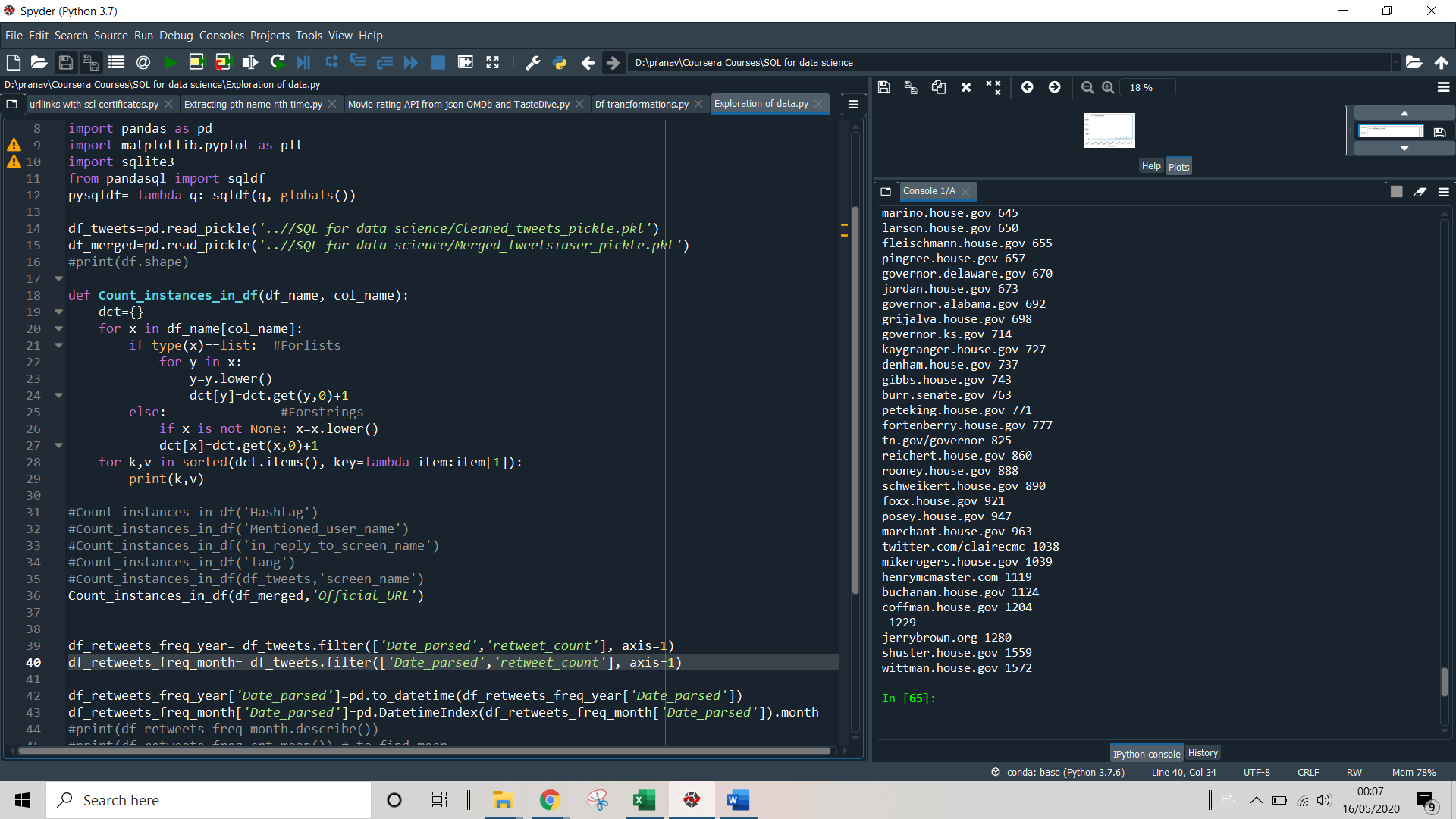
I looked at the retweet count with respect to the years and months as per below in order to gauge if the response from people and politician is on the increase. I noticed that in the month of Jan and December we have the maximum no. of rewteets and that interest has increased since 2008.



I also made a dictionary to plot which politician has been most active in the initiative of Lobbyists4America. We can see certain people are taking an active interest in the cause, with the most active ones, robwittman, repbillshuster having tweeted around 1500 tweets over 4 years which amounts to more than 1 tweet/day. But we need to see how many are politicians out of these. We can make use of the official URL registered on twitter and see if it is a .gov URL.



So now merging dataframes, i.e of tweets and users, we are able to plot the same nos. as above against the URL:



We see that except jerrybrown.org all other twitter handles are official government handles and therefore the politicians are the ones who are taking an active interest in the issues.

1. Submit 2-3 key points you may have discovered about the data, e.g. new relationships? Aha's! Did you come up with additional ideas for other things to review?

I have found that retweets and user mentions are increasing with time. Also that Jan and Dec are favourite months in which max tweets are received. Some frequently mentioned hashtags are #tcot, #gob, #4jobs. The maximum tweets have been received from Wshington DC, followed by Montross and Holidaysburg. This is where politicians who are ardent supports are from.

There was one Aha! Moment- looking at the initial data I had thought that the tweets data is for both lobbyists (general public) and the politicians but turns out that the data is only for politicians. Will have to change the data analysis according to that..

I did come up with the idea to check the .gov URL to ascertain if the tweets were from an official verified account.

1. Did you prove or disprove any of your initial hypotheses? If so, which one and what do you plan to do next?

I proved the hypothesis that the no. of tweets and interest of the public would increase as time went by. Next, I plan to do a sentiment analysis of the tweet text to determine the attitude of the Congressmen towards the initiative.

1. What additional questions are you seeking to answer?

What the specific hashtags mean. I could not find what specifically #tcot, #gob mean in the Lobbyists4America and how they relate to the cause.