

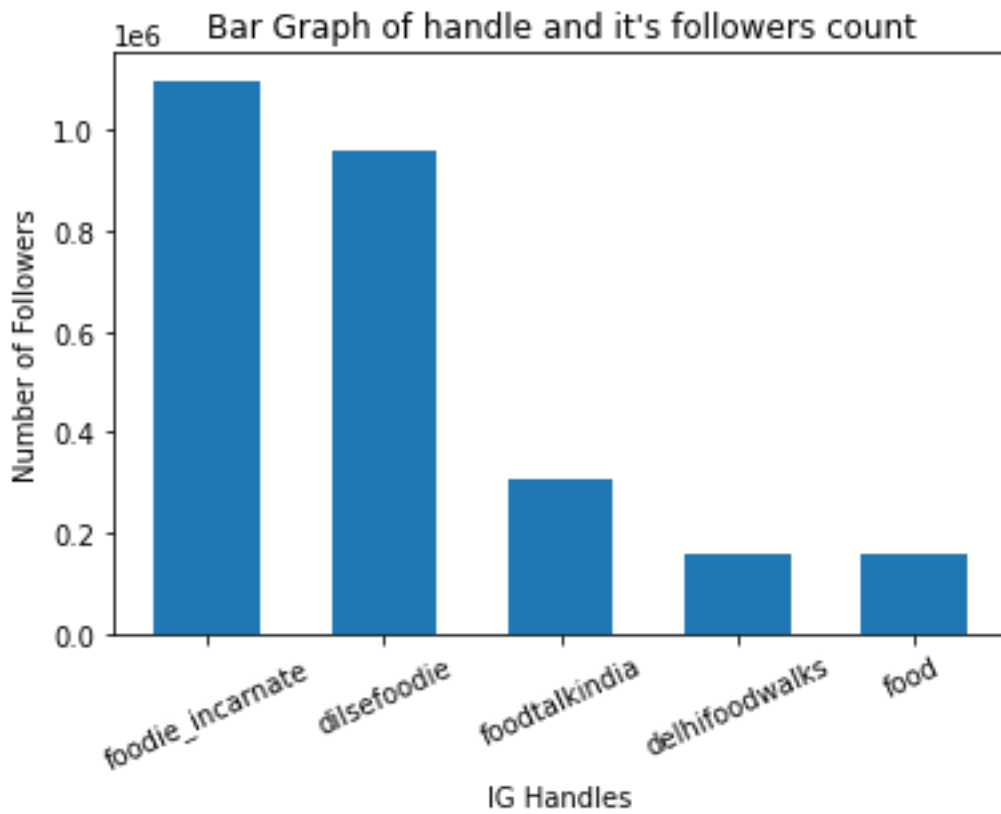
INSTA BOT-2

1) Now your friend has followed a lot of different food bloggers, he needs to analyze the habits of these bloggers.

- From the list of Instagram handles you obtained when you searched 'food' in a previous project. Open the first 10 handles and find the top 5 which have the highest number of followers
- Now Find the number of posts these handles have done in the previous 3 days.
- Depict this information using a suitable graph.

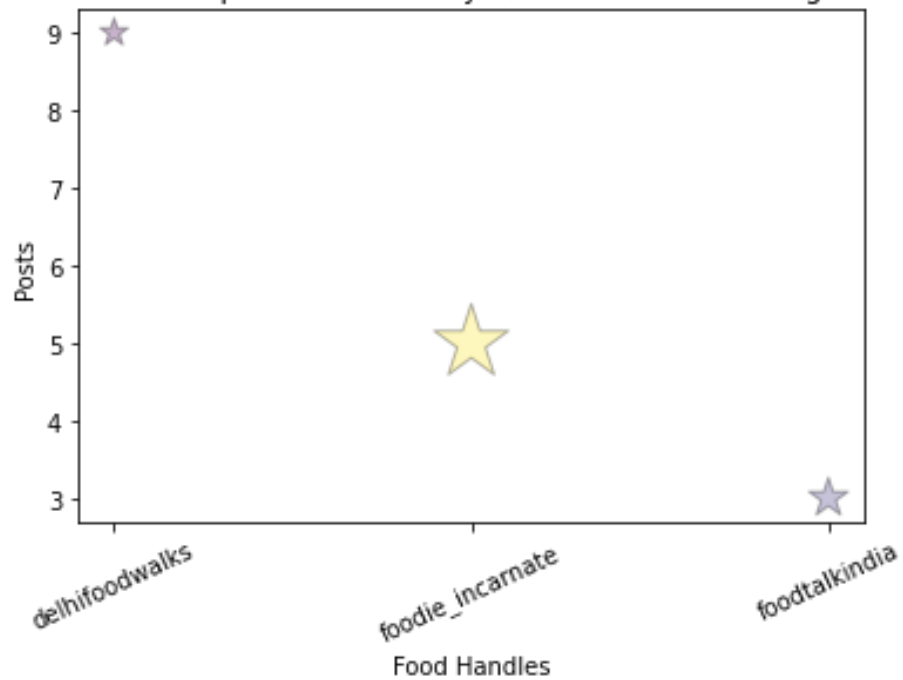
Solution Approach:

- Top 5 which have the highest number of followers –
 - i. Search 'food'.
 - ii. Extract links of the first 10 handle results.
 - iii. Extract handle names from links.
 - iv. Open each link.
 - v. Locate followers button.
 - vi. Extract followers count from the button and store it in an array.
 - vii. Sort the array and extract the Top 5 followers count and their handles.
- Top 3 which have the highest number of posts in last 3 days –
 - i. Open handle link of each handle.
 - ii. Load posts (scroll if required) and store them in a list.
 - iii. Open each post from the list.
 - iv. Extract time of post.
 - v. If time < 3 days → count it.
 - vi. Else → break
 - vii. Store the final post count of each handle in an array.
 - viii. Sort the array and extract the Top 3 posts count and their handles.
- Depict this information using a suitable graph –
 - i. Bar Graph of handles and their followers count.



- ii. Bubble chart of handle and their last 3 days posts (with followers count as weight).

Bubble chart --> No. of posts in last 3 days (follower count is weight of start-point)

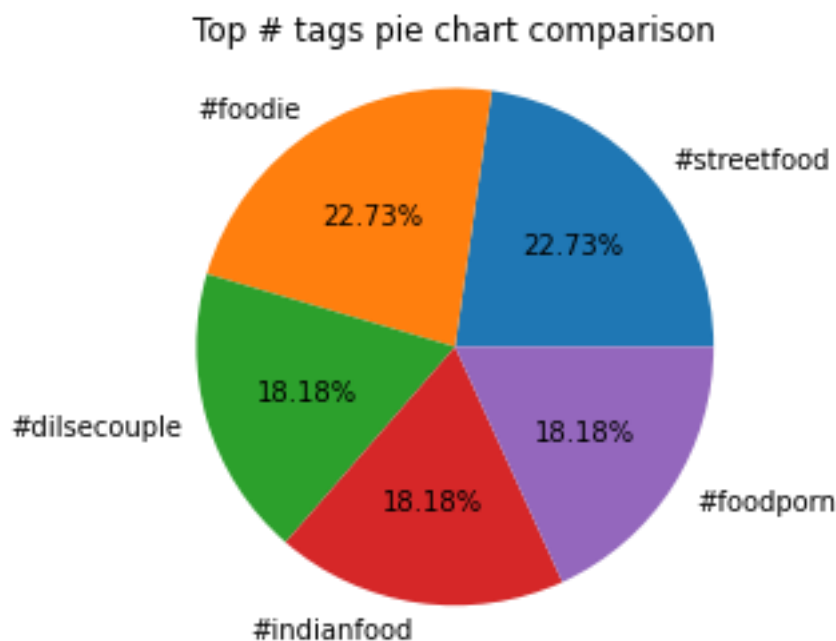


- 2) Your friend also needs a list of hashtags that he should use in his posts.
- Open the 5 handles you obtained in the last question, and scrape the content of the first 10 posts of each handle.
 - Prepare a list of all words used in all the scraped posts and calculate the frequency of each word.
 - Create a CSV file with two columns: the word and its frequency
 - Now, find the hashtags that were most popular among these bloggers
 - Plot a Pie Chart of the top 5 hashtags obtained and the number of times they were used by these bloggers in the scraped posts.

Solution Approach:

- Open the 5 handles you obtained in the last question, and scrape the content of the first 10 posts of each handle –
 - i. Load data from 'top5.csv' (that I created in the last notebook).
 - ii. Create a 'total_desc_string' string which stores all the posts text together.
 - iii. For each link in handle links.
 - iv. Open link.
 - v. Load the post description.
 - vi. Clean description.
 - vii. Add it to the 'total_desc_string'.
 - viii. Add '--sep_post--' after each post description is added (to separate posts later).
 - ix. Add '--sep_post--' after each all post description of each handle is added (to separate handles later).
 - x. Print the scrapped 'total_desc_string' (separated by '--sep_post--' and '--sep_handle--').
- Prepare a list of all words used in all the scraped posts and calculate the frequency of each word–
 - i. Create a 'word_dict' dictionary to store words and their counts.
 - ii. Iterate 'total_desc_string' and store each word as the key in the dictionary and increment their count.
- Create a CSV file with two columns: the word and its frequency–
 - i. Extract keys and values from the above-created dictionary and store them as a list separately.
 - ii. Create CSV with words, and count as counts and above list as values.

- Find the hashtags that were most popular among these bloggers–
 - i. Create hash_tags _dict dictionary which stores hashtags and their count.
 - ii. Iterate word_dict and store keys which start with '#' and count in hash_tags_dict.
- Plot a Pie Chart of the top 5 hashtags obtained and the number of times they were used by these bloggers in the scraped posts–



- 3) You need to also calculate the average followers: likes ratio for the obtained handles.

Followers: Likes ratio is calculated as follows:

- Find out the likes of the top 10 posts of the 5 handles obtained earlier.
- Calculate the average likes for a handle.
- Divide the average likes obtained from the number of followers of the handle to get the average followers: like ratio of each handle.
- Create a bar graph to depict the above-obtained information.

Solution Approach:

- Find out the likes of the top 10 posts of the 5 handles obtained earlier–

- i. Load data from 'top5.csv'.
 - ii. Create a 'total_likes_arr' to store the sum of likes of the post of each handle.
 - iii. Open each handle link.
 - iv. Collect the first 10 posts with likes.
 - v. Print their likes along with the handle name.
 - vi. Store sum of likes in 'total_likes_arr'.
- Calculate the average likes for each handle –
 - i. Convert 'total_likes_arr' to NumPy array and store as 'average_likes_arr'.
 - ii. Divide 'average_likes_arr' by 10.
 - Followers: Likes ratio –
 - i. Load followers from 'top5.csv' and store them as 'followers_counts'.
 - ii. Perform 'followers_counts' / 'average_likes_arr'.
 - Create a bar graph to depict the above-obtained information –

