

# Twitter newspapers

**Hassan Abdullah**

**website: <https://hassan6633.github.io/mis492/>  
(<https://hassan6633.github.io/mis492/>)**

## newspapers:

- Alwatan
- Alseyasseyah
- Aldaar
- Alrai
- Alanbaa
- Alqabas
- Alnahar
- Aljareda

## 1- newspapers followers

## 2- newspapers tweets

## Tools of data collection:

- tweepy
- collected through API

## Tools of analysis:

- matplotlib.pyplot
- pandas
- seaborn
- numpy
- matplotlib.dates

## Followers:

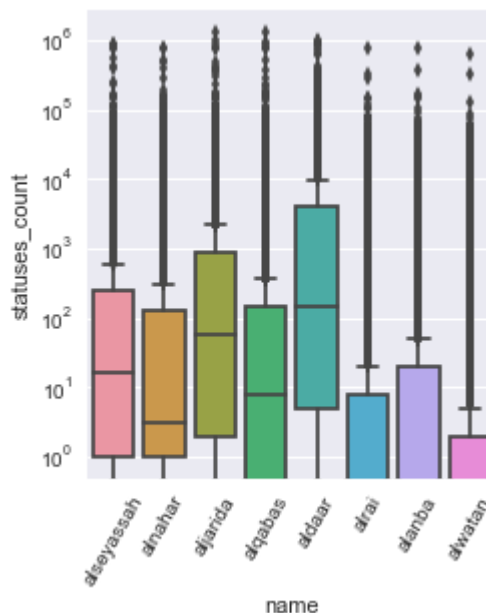
Variables	Description	Data type	Notes
created_at	date the account was created at	datetime64, Categorical	converted from object to datetime64
default_profile_image	shows the type of the profile image if its the default image or not	object, Categorical	
description	bio of the account	object, Categorical	
favourites_count	hows the number of the account likes on tweets	float64, Nominal	
followers_count	followers of the account	float64, Nominal	
friends_count	people that the account is following	float64, Nominal	converted from object to float64
geo_enabled	location of the account user	object, Categorical	
id	id of the account	object, Nominal	
listed_count	number of public lists the user have or member of	float64, Nominal	converted from object to float64
location	the location of the user	object, Categorical	
profile_background_color	color of the background	object, Categorical	
profile_background_image_url	url of the background image	object, Nominal	
profile_text_color	color of the profile text	object, Nominal	
profile_use_background_image	the type of the backgrnd image if its the original or changed	object, Categorical	
protected	is the account private or not	object, Categorical	
screen_name	username of the account	object, Categorical	
statuses_count	number of tweets	float64, continuous	
url	the url account has in his profile	object, Categorical	
name	name of the newspaper	object, Categorical	i added this column
tweets_mean	average number of tweets	float64, continuous	i added this column

- **contributors\_enabled** : this column was not used in the analysis because its definition was not known
- **default\_profile** : this column was not used in the analysis because its definition was not known
- **profile\_background\_tile** : this column was not used in the analysis because its definition was not known
- **verified** : this column was not used in the analysis because its definition was not known

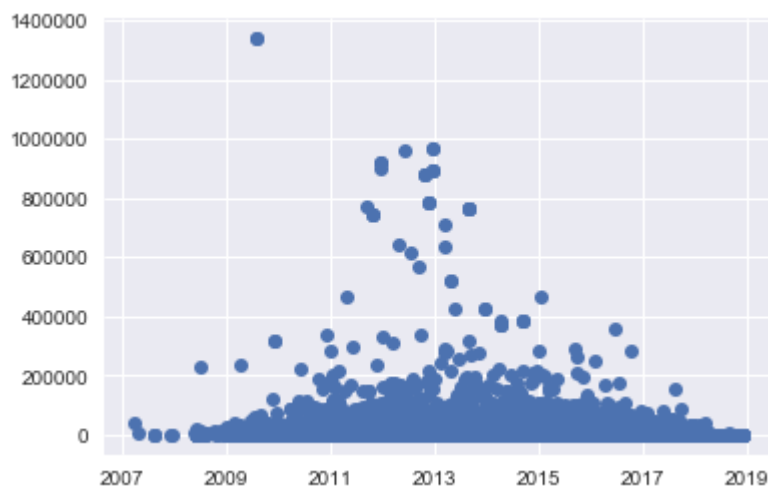
In [ ]: *#removed any duplicates between the followers*

```
followers_df2 = followers_df[~followers_df[["screen_name","name"]].duplicated()]
```

- aldaar followers are the more active in twitter over other newspapers followers



- followers of the newspapers that created there account between 2012 and 2013 have the most number of tweets (a political problems at period may pushed people to open twitter accounts and follow the newspapers to follow up news and interact with it)

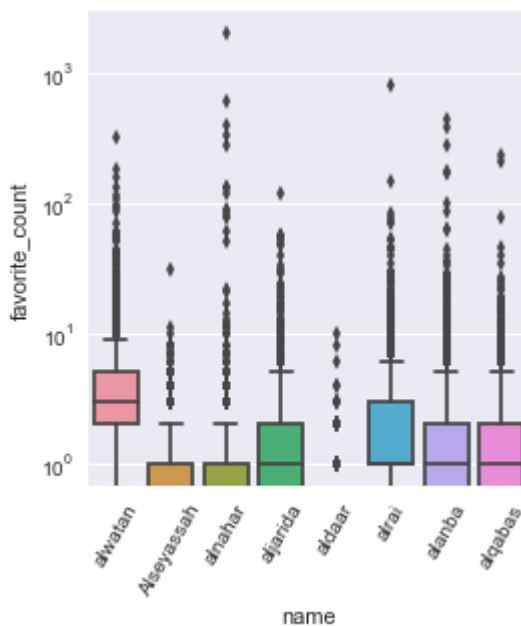


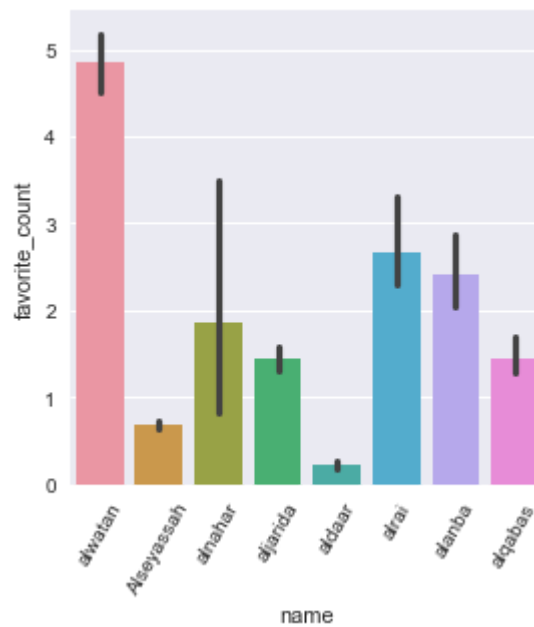
## Tweets:

Variables	Description	Data type	Notes
created_at	date the tweet was created	datetime, continuous	converted from object to datetime

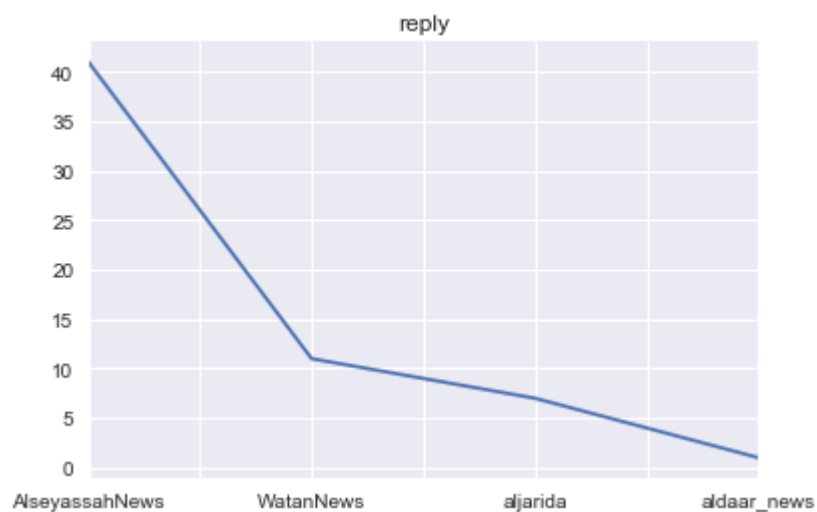
Variables	Description	Data type	Notes
entities	any entity the tweet has (hashtags, symbols, mentions)	object, categorical	
favorite_count	number of likes the tweet has	float64, continuous	
id	id of the tweet	float64, continuous	
in_reply_to_screen_name	if the tweet has reply (show screen name)	object, categorical	
in_reply_to_status_id	if the tweet has reply (show tweet id)	float64, continuous	
in_reply_to_user_id	if the tweet has reply (show user id)	float64, continuous	
place	place the tweet was tweeted from	object, categorical	all the data was NaN so i excluded the column
possibly_sensitive	if there are any sensitive content in the tweet	object, categorical	all the data was False so i excluded the column
retweet_count	number of retweets	float64, continuous	
text	the text of the tweet	object, categorical	
name	the name of the newspaper	object, categorical	i added this column
retweet_mean	average number of retweets	float64, continuous	i added thid column

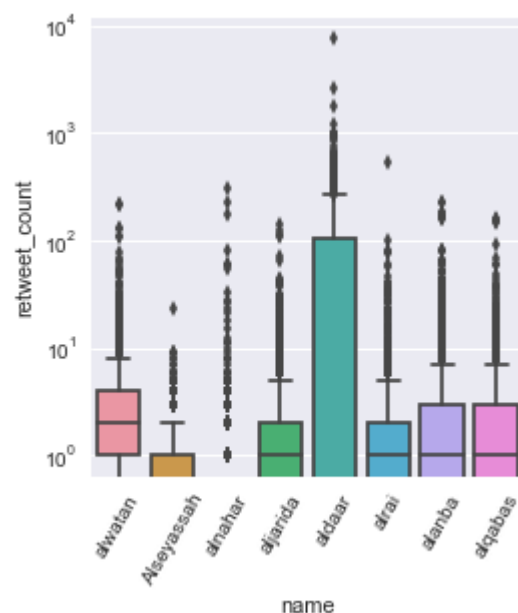
- alwatan have the highest average of tweets likes (is it because people like the tweets more than other newspapers tweets or because it have more follwers?)



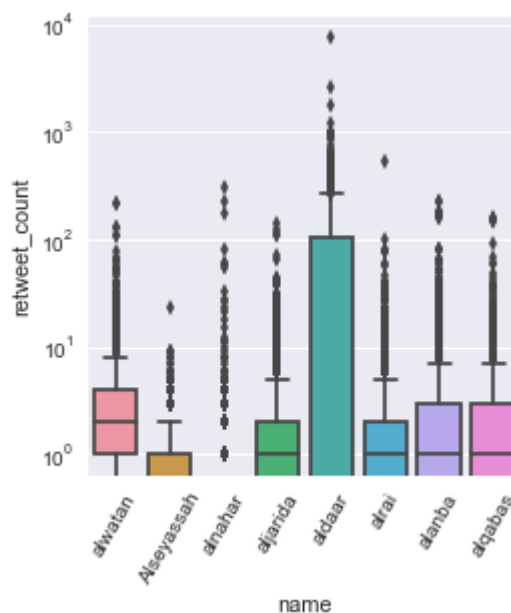


- alseyaseyah have people reply more than other newspapers but retweets less than other people (the newspaper tweets may not reflect followers thoughts and opinions or the people is replying with something they didn't like about the tweet)

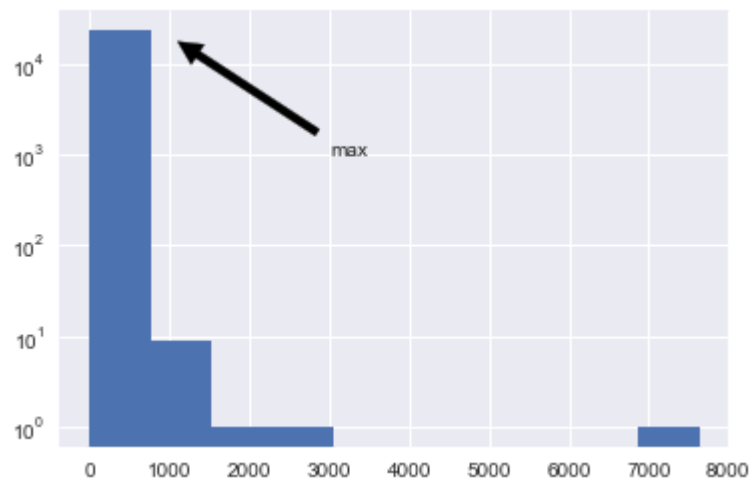




- alwatan have above average number of retweets and alseyassah have below average number of retweets
- although aldaar is below average, it has the most retweets as outliers (that may be because it can get the interesting news that people interested in faster than other newspapers for some types of news)



- most retweets for all the newspapers are between 0 to 800



- if i had more time to get more data like the followers of the newspapers accounts and to explore more there may be more intresting and accurate results to find.

In [ ]: