

Spartz Data Exercise
dcl, April 2015

1. Review and describe the data provided. Note any problems, characterize any distributions of interest, include anything you think is interesting.
2. Recommend ways to improve the performance of the @OMGFacts twitter account, and justify your answers with data and graphics.
3. Show all of your work! We want to see as much as we can of your programming and analysis style.
4. Describe your favorite machine learning algorithm and tell us why you think it's great :-)

NOTES:

- you are trying to increase the value of `ga_twitter_pageviews`
- the field `post_id` in `TwitterData.csv` joins to the `id` field in `Lists.csv` if the `list` field is 1, and to the `id` field in `OMGPosts.csv` when the `list` field is 0
- there are many possible features to be derived from the data provided, please be creative