

Term Project

Tanveer S Bariana

December 7, 2017

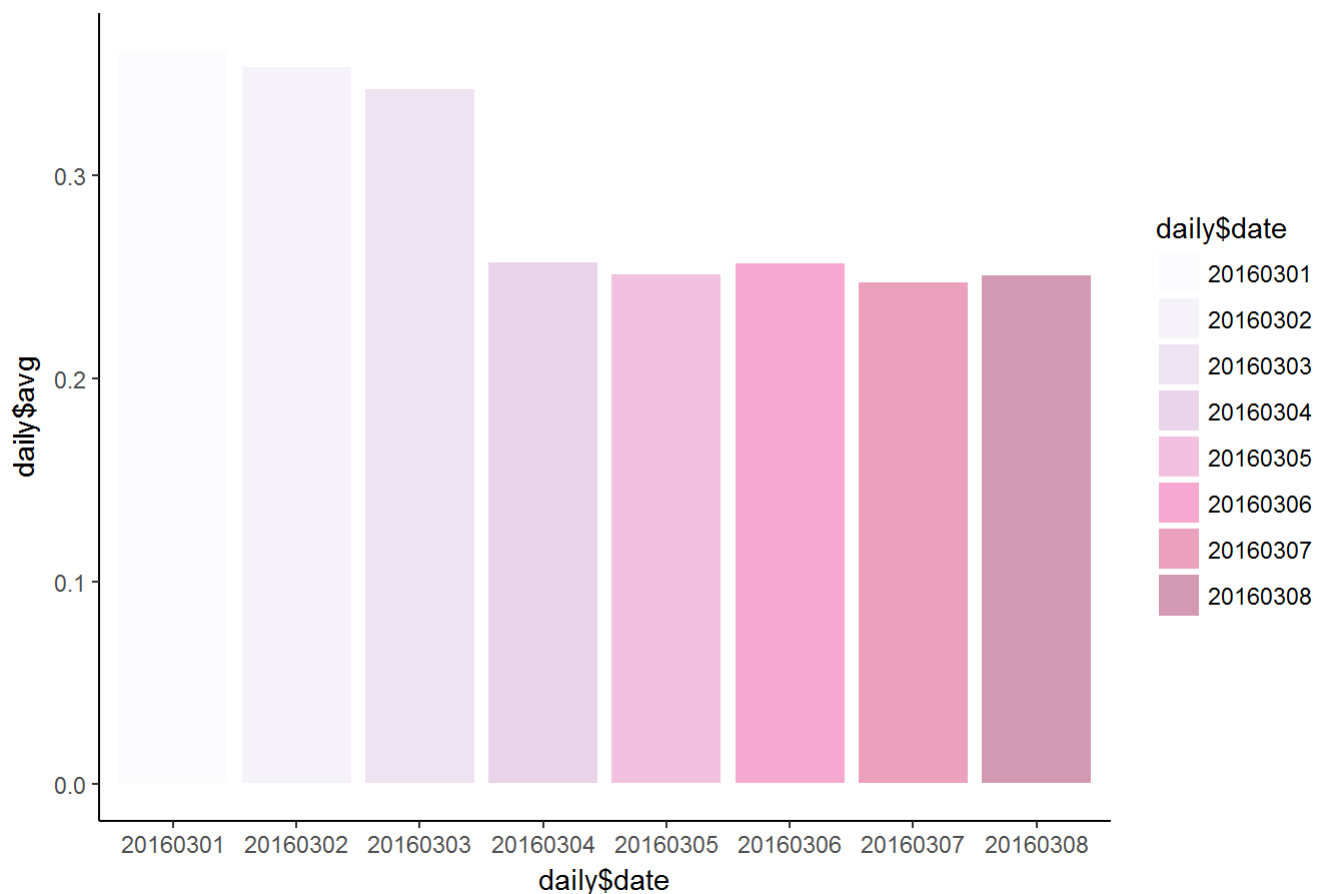
In this R markdown file i will go through my results for the data science interveiw question for the prestigious wikimedia foundation

here are the questions

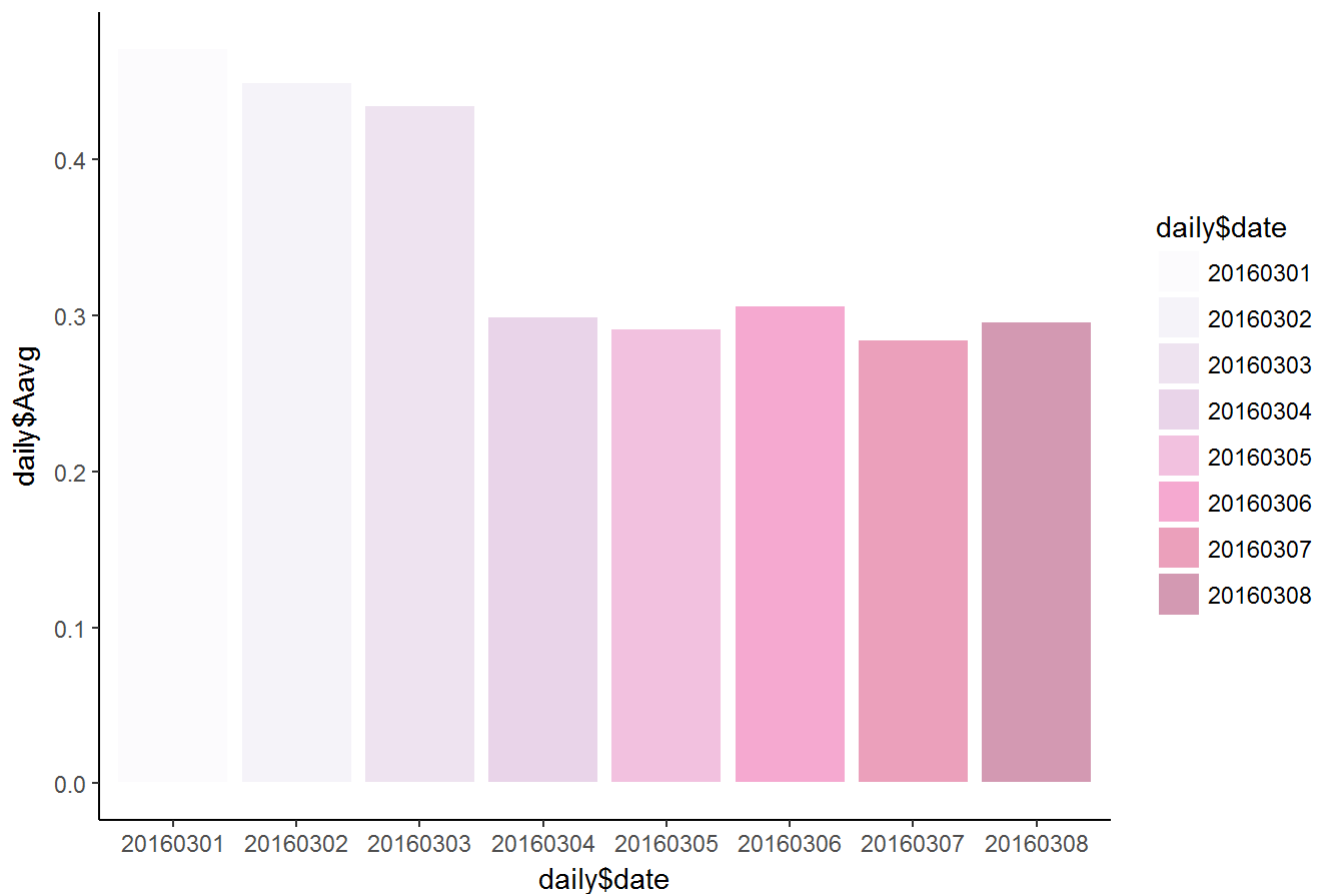
1. What is our daily overall clickthrough rate? How does it vary between the groups?
2. Which results do people tend to try first? How does it change day-to-day?
3. What is our daily overall zero results rate? How does it vary between the groups?
4. Let session length be approximately the time between the first event and the last event in a session.
Choose a variable from the dataset and describe its relationship to session length. Visualize the relationship.

so the folowing are my grpahs showing the overall daily click through rate and teh daily rates of Group A and B respectivly

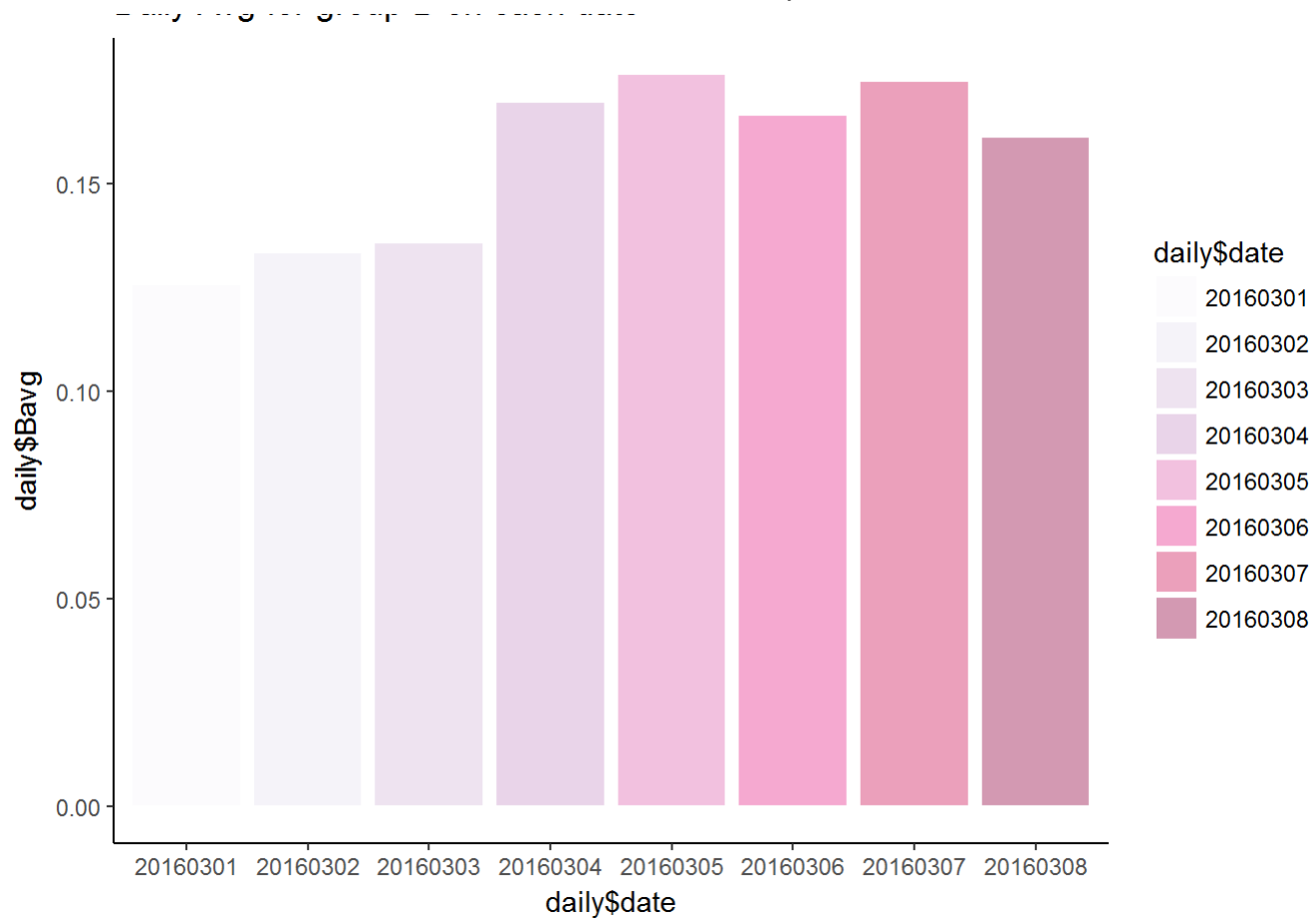
Over all daily Avg for each date



Daily Avg for group A on each date



Daily Ava for aroup B on each date

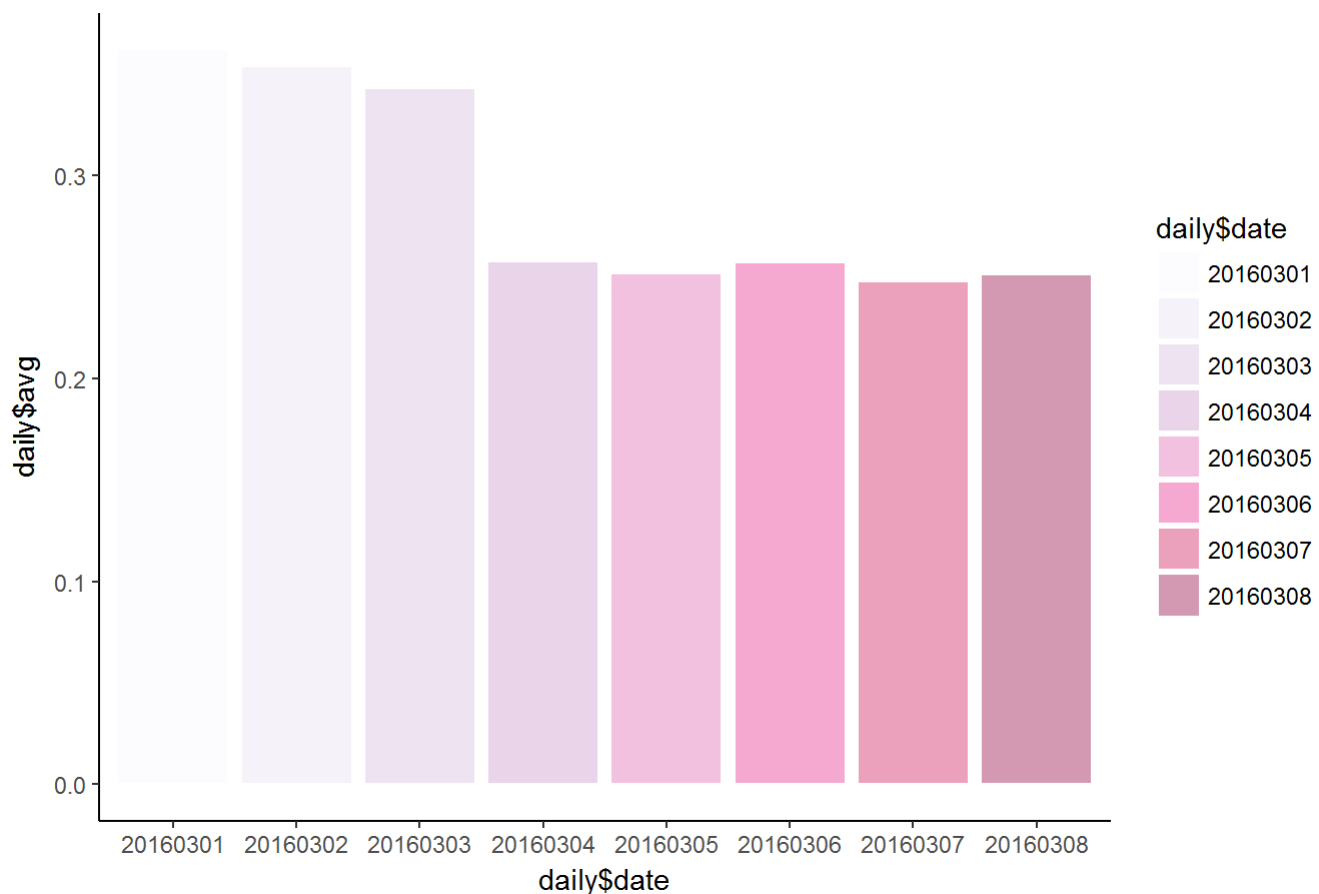


one can observe how the overall click rate and click rate for group A both decreases as time progresses but Group B increases

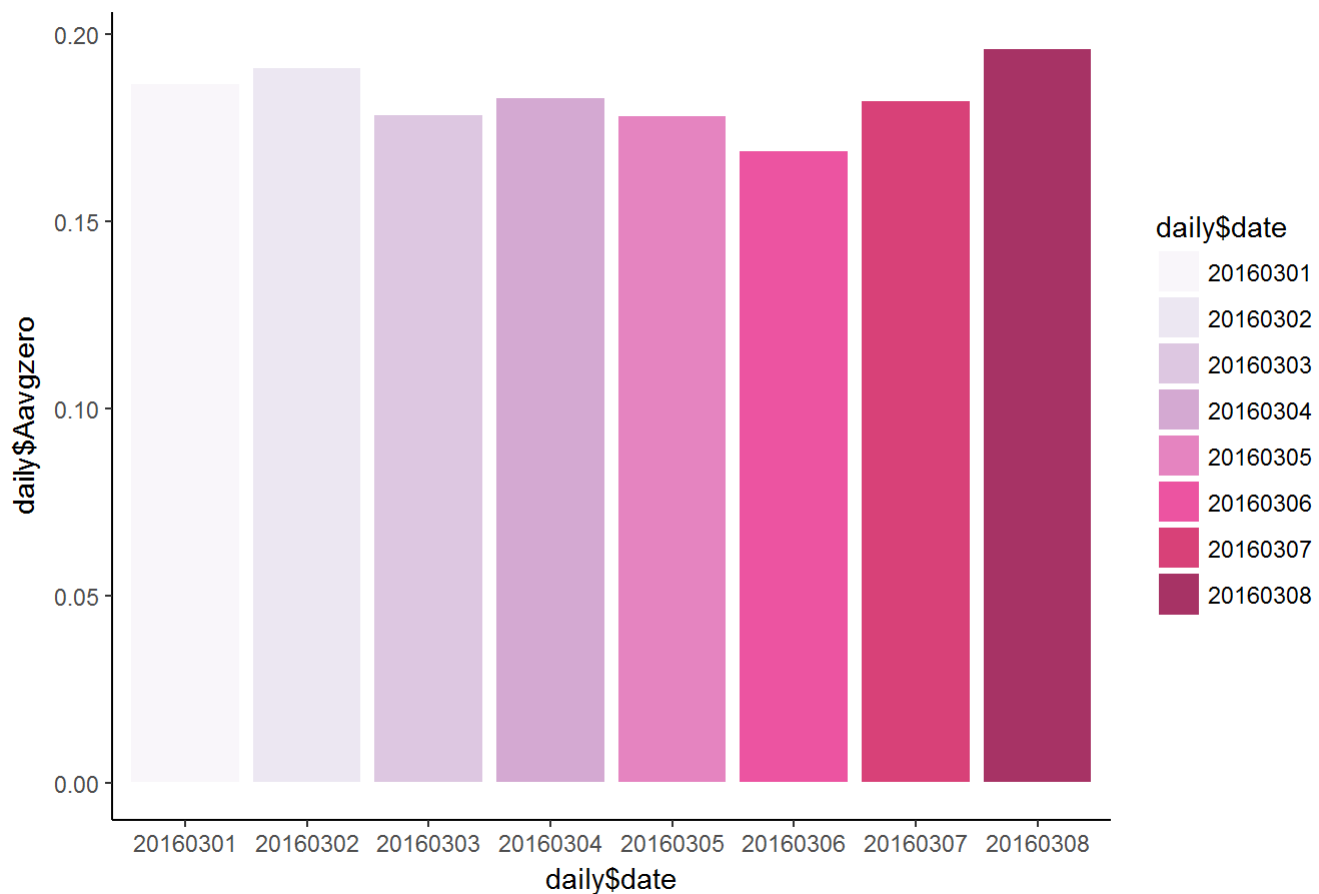
now we move onto graphs for the Zero result rates, and the zero result rates of groups A and B respectively

```
## Scale for 'fill' is already present. Adding another scale for 'fill',
## which will replace the existing scale.
```

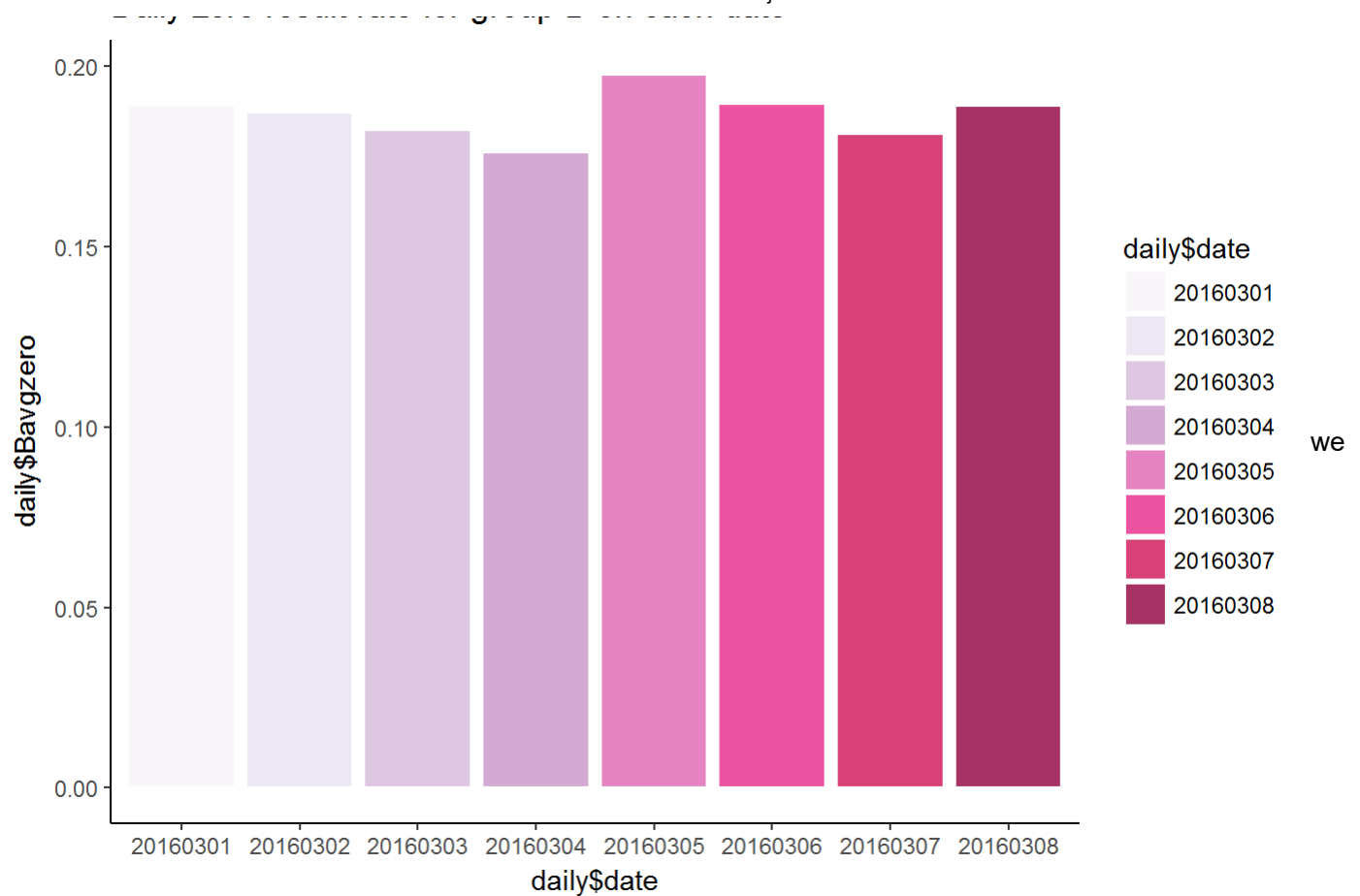

Over all daily zero result rate for each date



Daily zero result rate for group A on each date



Daily zero result rate for group B on each date



can observe that group A and B both have stable Zero Result rates very day with no major changes between days. WE can also note how the overall Zero Result rate decreases daily much liek how the overall daily click rate decreased