# Problem 2: Simple EDA

*Problem Statement:* Performing EDA on a simple data set

**Data:**

The Given dataset represent one day’s Ad’s response from User. Each row is 1 user response.   
The five columns are the features of the user that we are interested in to analyze with the response.

**Part-I Observations:**

**Simple Data analysis on single Dataset:**

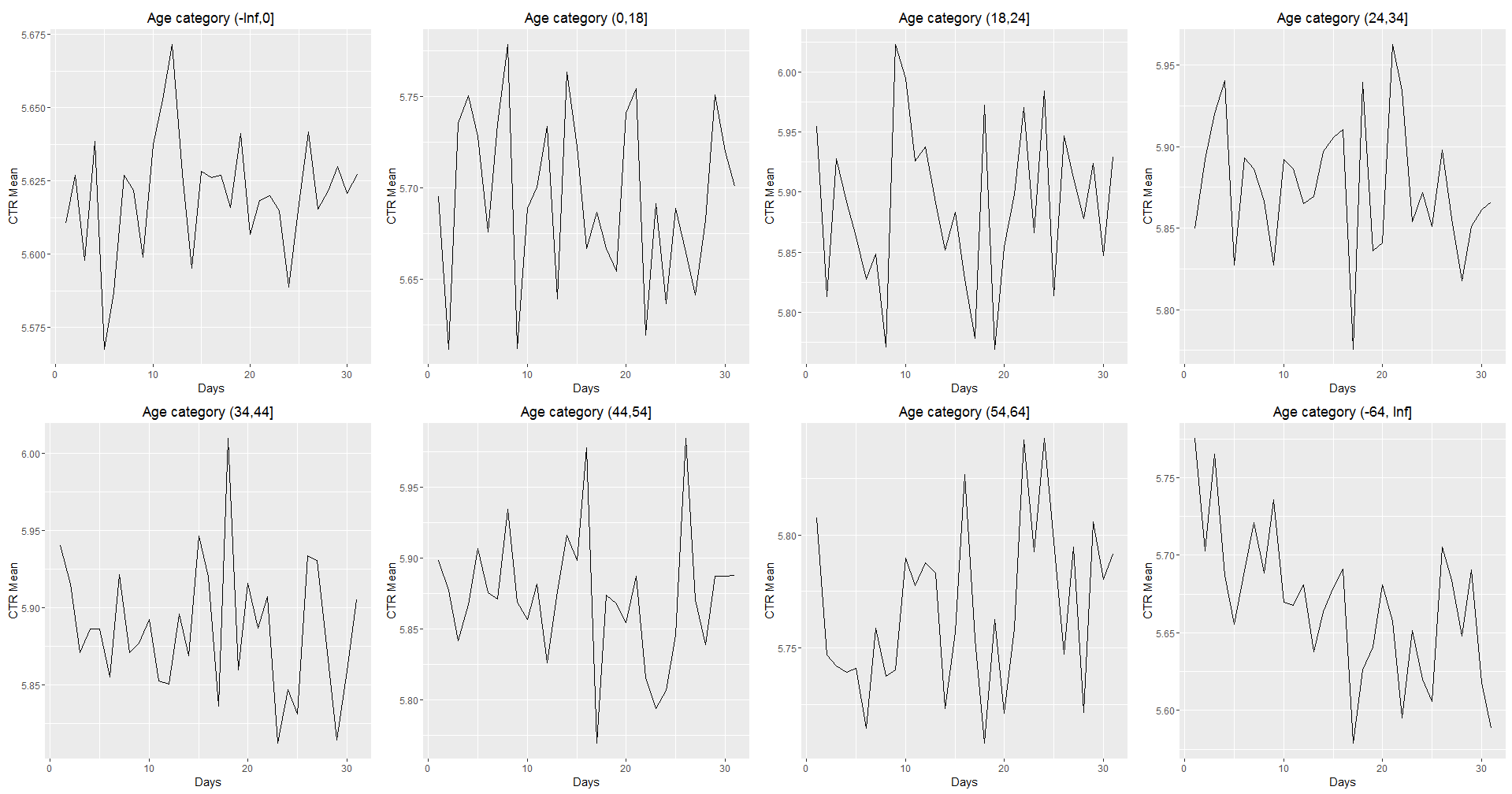
After categorizing data based on Age categories, we can leverage the gender factor to observe how the different gender impressions differ. The below graph demonstrates that feature. The Click through Rate is high in males for the age category [0 to 18]. Gender (0=female, 1=male)

****

**Part-II Observations:**

Now we can extend our summary along 30 days of click through data. We can summarize the average Click through rate for each category separately and plotted then to compare and find out the target audience for the ads based on impressions and click rate.

Graph below represents CTR for different age categories:



CTR variation in different Age groups for 30 Days

References:

References:

My submission with for this project includes references from below links:

1. http://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html

2. http://stackoverflow.com/questions/11370323/learning-to-understand-plyr-ddply

3. http://ww2.coastal.edu/kingw/statistics/R-tutorials/dataframes.html

4. R-Cookbook

5. http://stackoverflow.com/questions/19655431/reading-multiple-csv-files-in-r