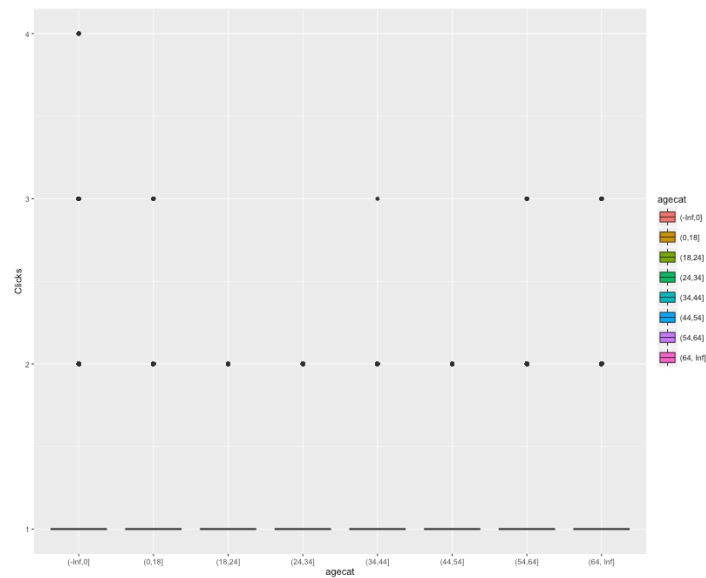SUDHARCHITH SONTY
UBID: 50169912
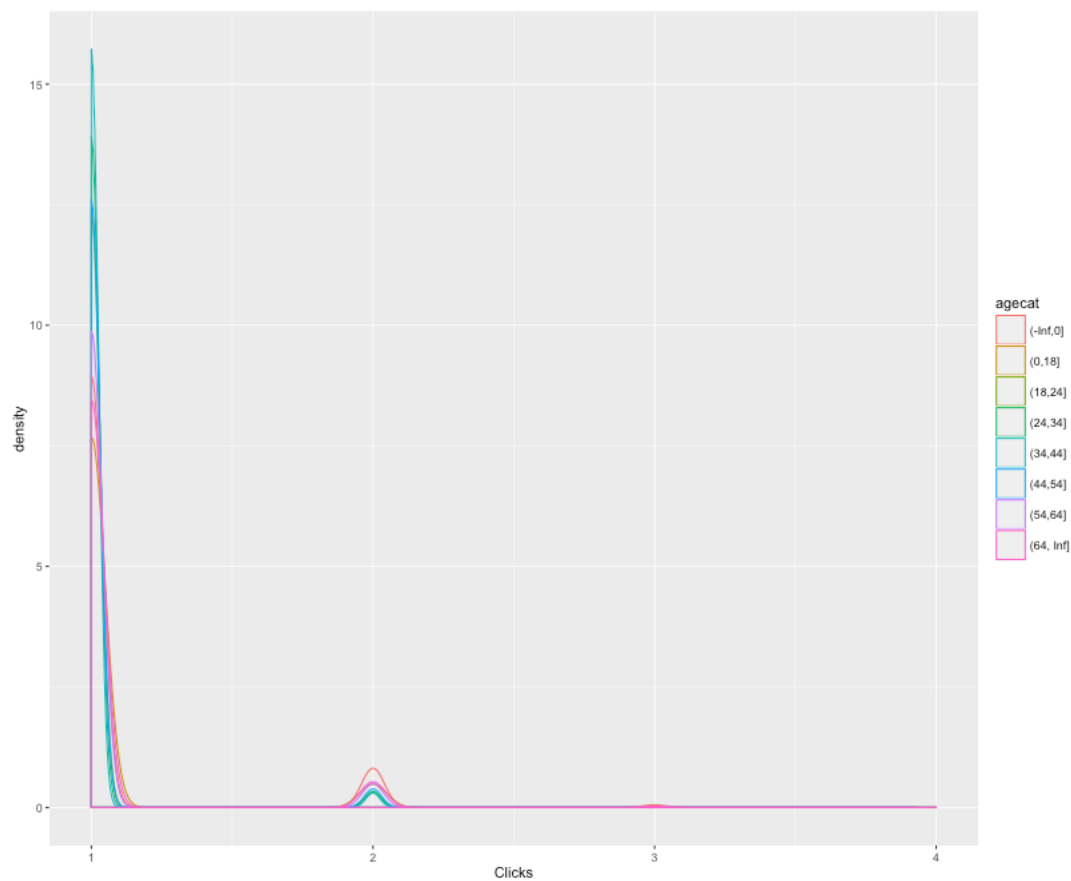Mail: sudharch@buffalo.edu

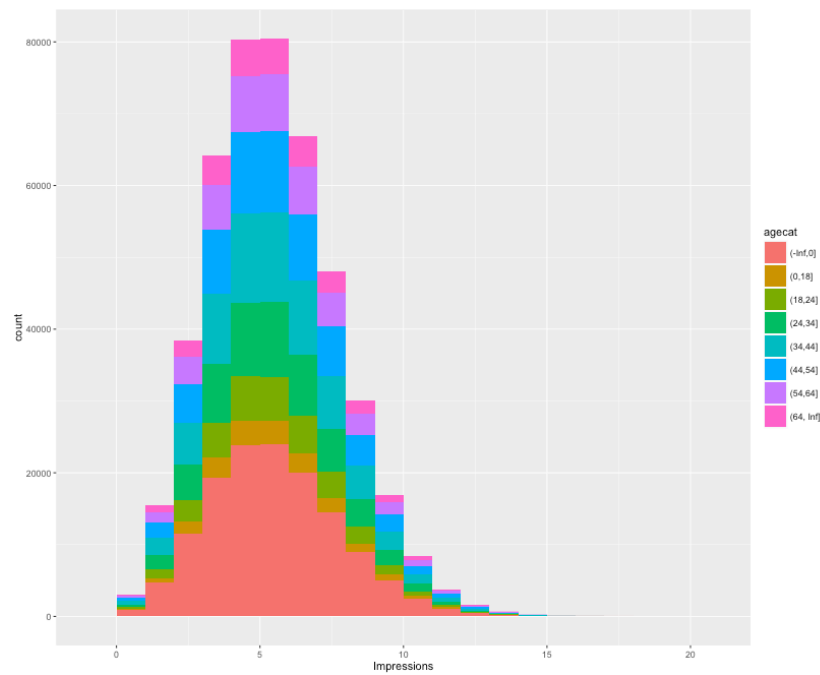CSE -587

DATA INTENSIVE COMPUTING

PROJECT 1 – PART 2 – REPORT

We have plotted a graph of the age category against the Clicks and we note that the number of clicks is mostly uniform across the age categories


Here we plotted a graph of the Clicks against the density of the clicks for each age category

A histogram showing the the Impressions made by each age category against the count of the number of Impressions and we can note that the age cat 54-64 has the highest count of the Impressions
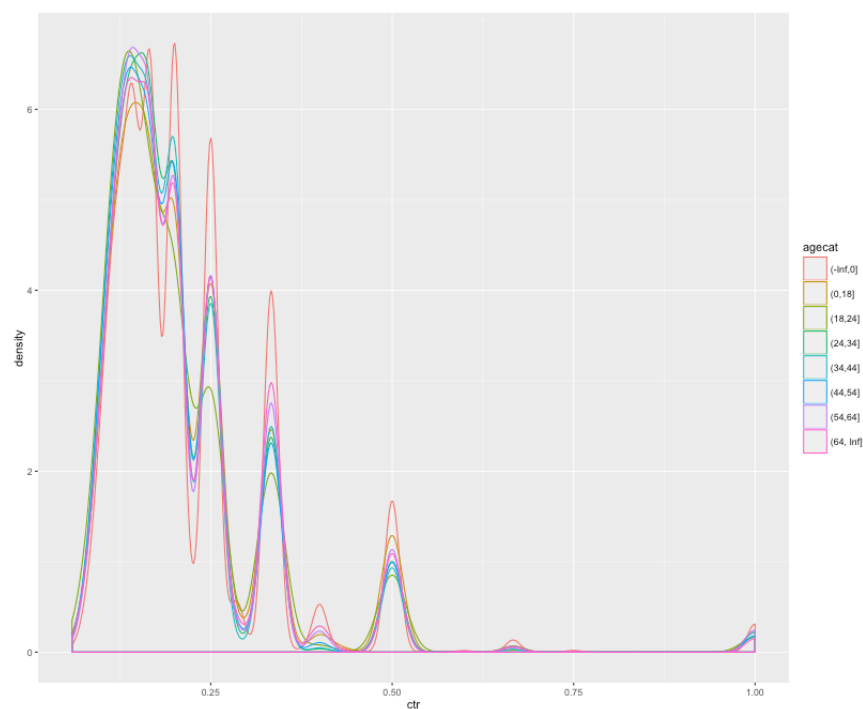
A graph plotting the Click Through rate which is defined the number of Clicks divided by the Number of Impressions against the Density of the ctr.
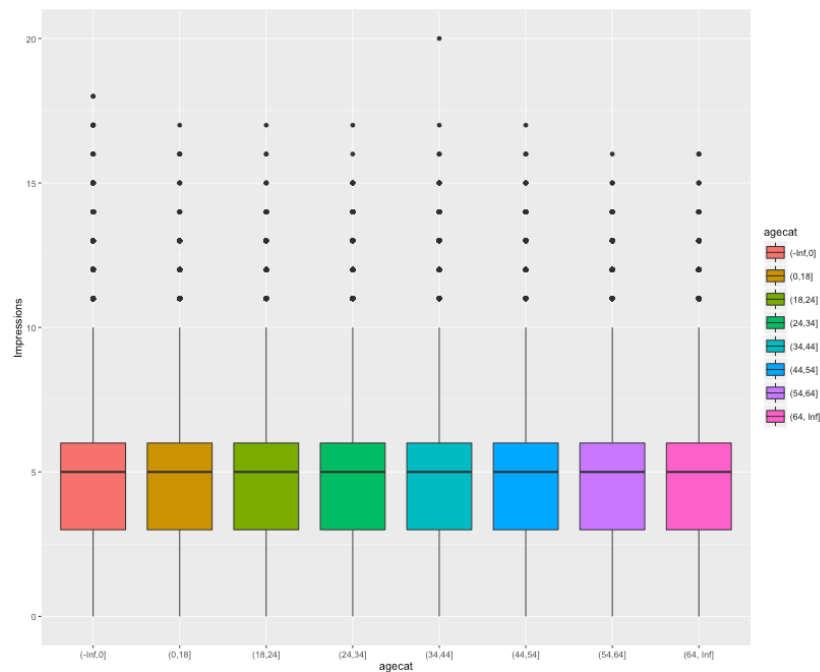
A box plot plotted with the x axis as the age category and the y axis as the number of impressions. And we can see that the number of impressions is highest for the users whose age category is unknown and leaving that side we have the impressions to be uniform across most of the age categories
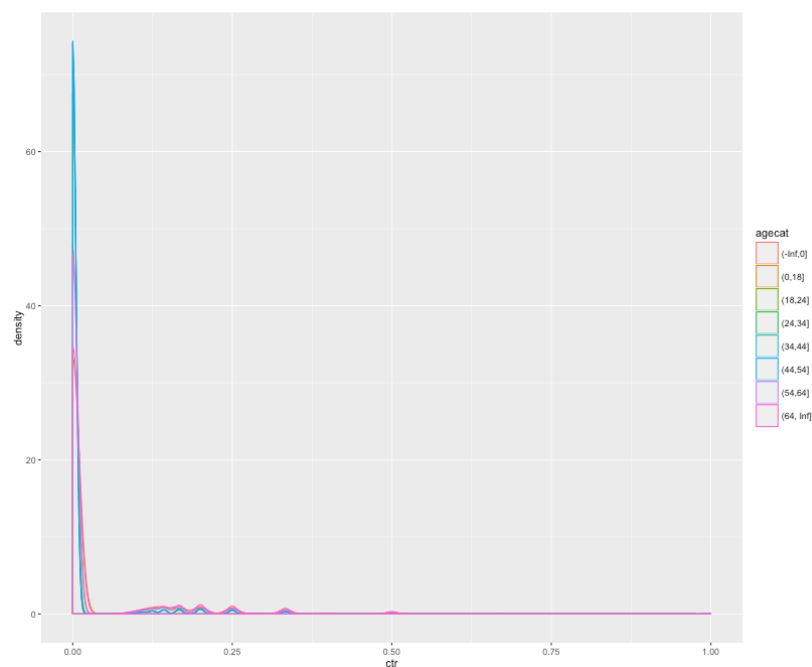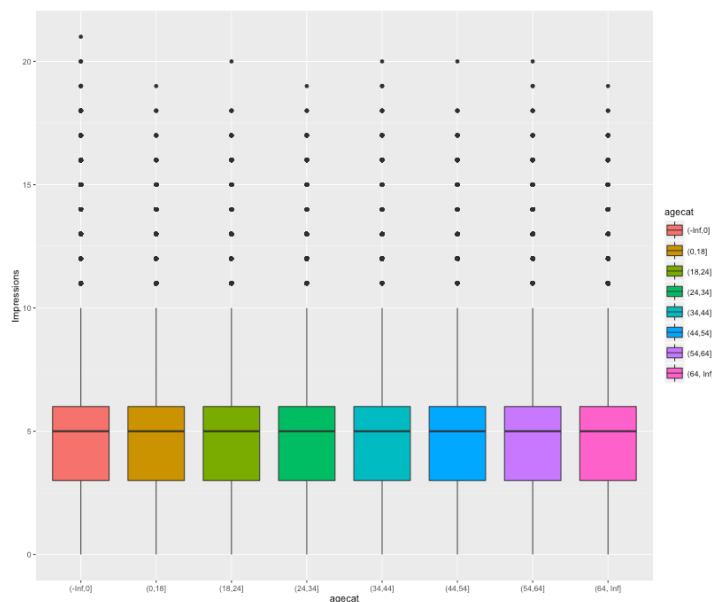

A graph that we have plotted using the ctr and the density when the Impressions are strictly greater than zero.
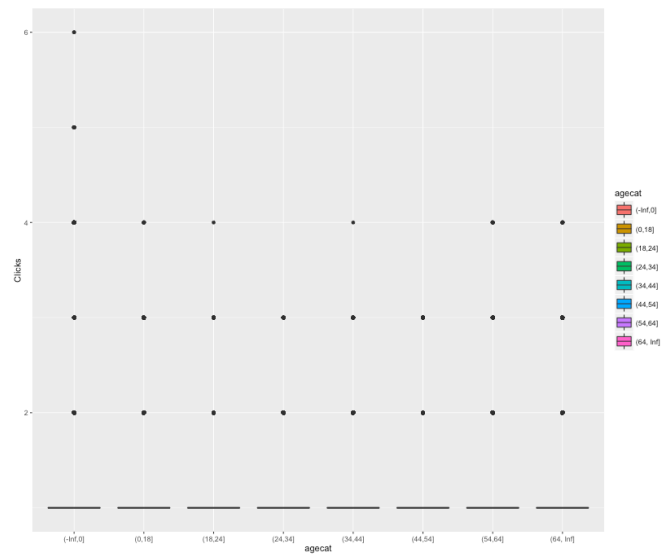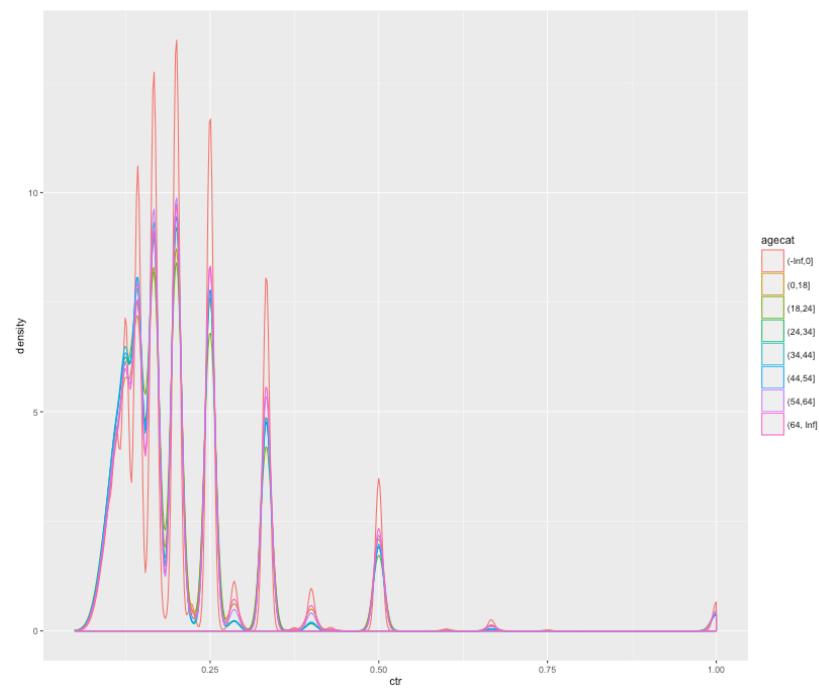
ANALYSIS EXTENDED ACROSS 30 DAYS
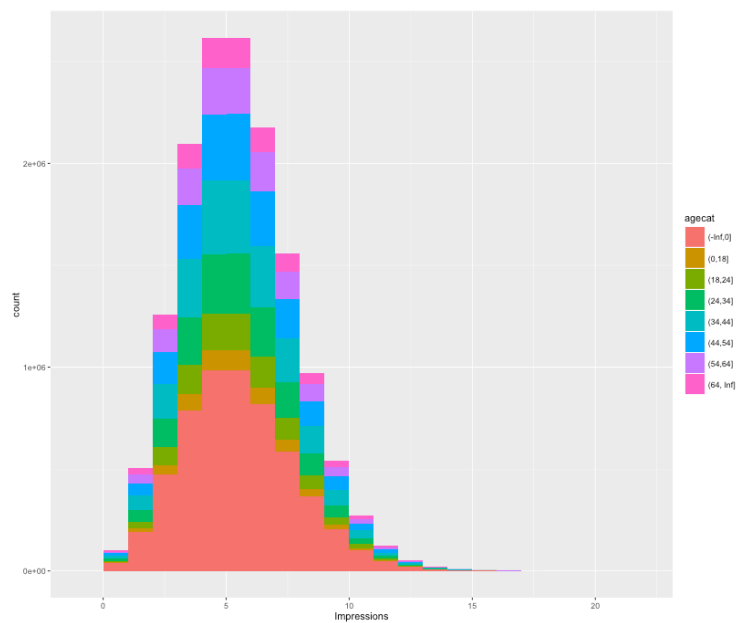
The analysis extended across 30 days shows that the age category above 64 has the highest count of Impressions contrary to the analysis we did on a single day's data.


The box plot also shows us that the Impressions of people who have not listed their age is higher but that being said the number of Impressions in the age category 54-64 has increased and the Impressions between the age category 34-44 is significantly more consistent

The number of clicks per age category again shows that the age category not listed i.e ;
people who have not entered their age have the highest amount of clicks

Plotting the density vs the Click Through Rate extended across 30 days shows a significant
spike the CTR
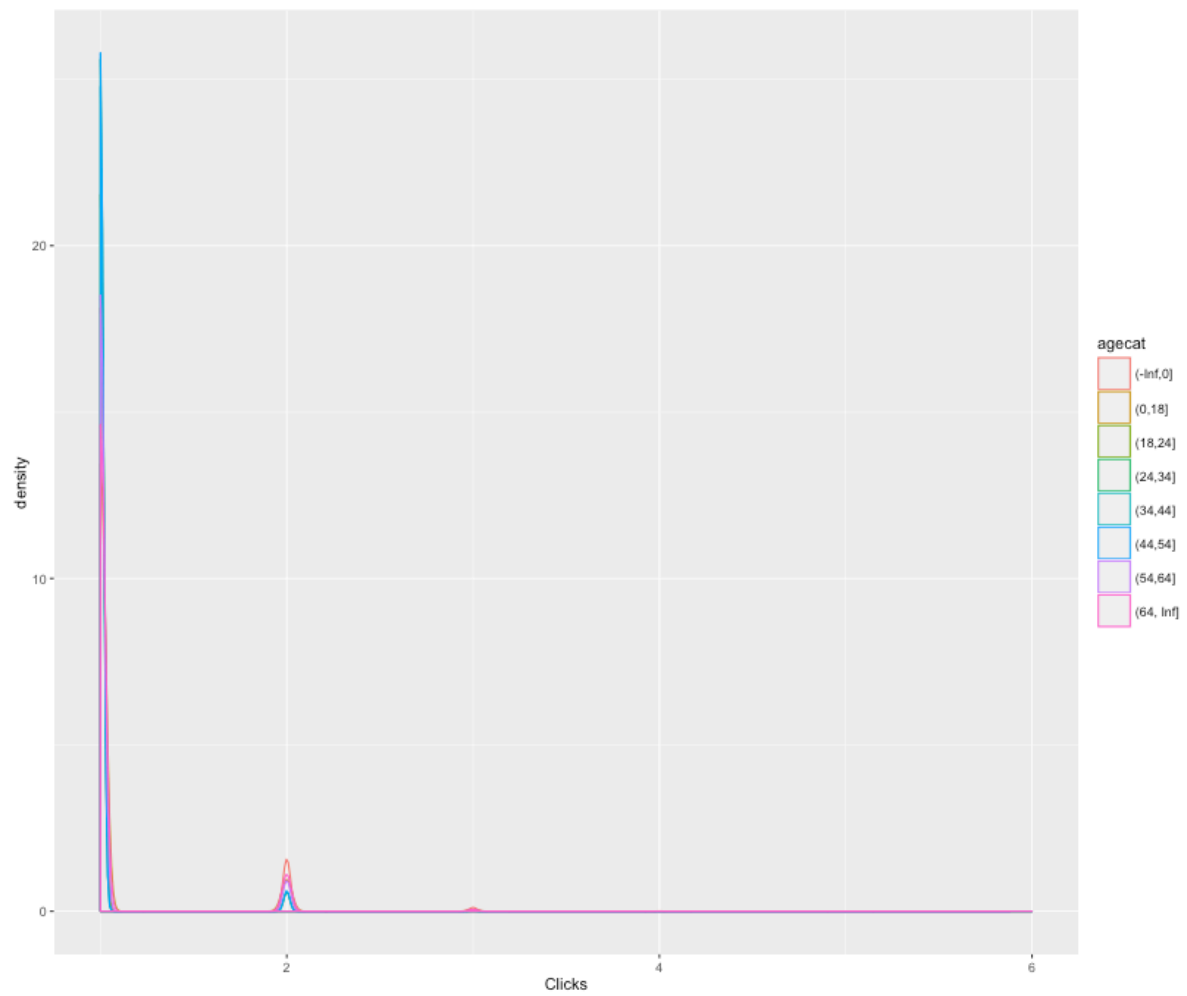
The number of clicks vs density graph extended shows that the density of clicks in higher for the age category34-44 and 44-54