

Survival Analysis Project, Math 150

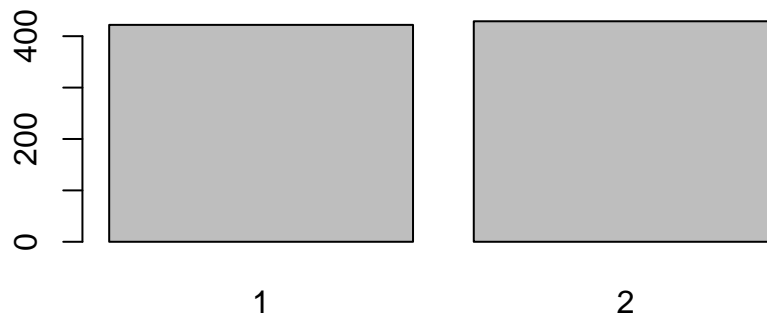
Siddhartha Aggarwal

4/8/2019

After uploading the data, the first visualization I wanted to create was to see how many individuals are part of each treatment group out of the 4 we have outlined. This would be interesting to get a sense of the variety of treatments in the study. The graph below shows that only Treatments 1 and 2 are administered in the study. This is odd, since not a single participant got treatments 3 or 4.

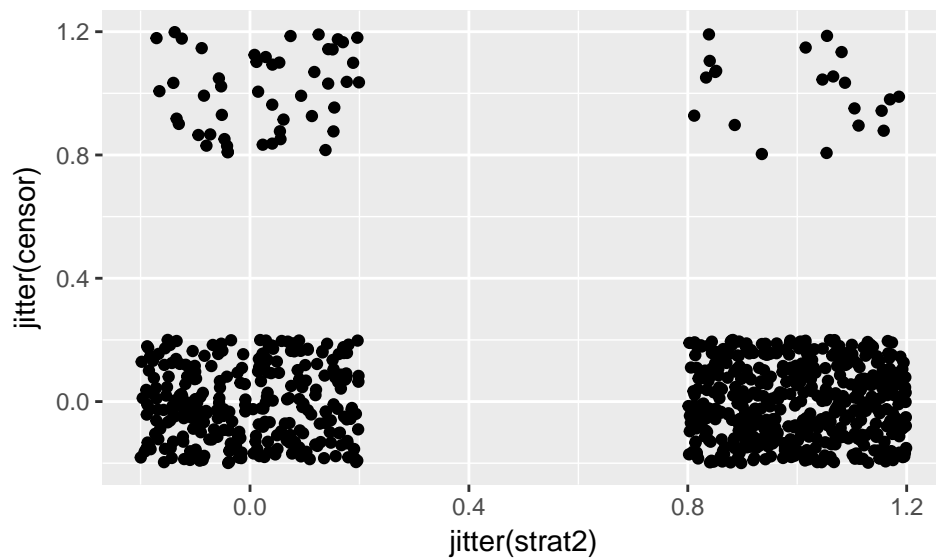
```
tbl <- with(AIDSdata, table(txgrp))
```

```
barplot(tbl)
```



The next visualization I made was to plot the CD4 stratum at screening against the event indicator for AIDS defining diagnosis or death (censor variable). This was to get a sense of whether CD4 stratum levels in any way had a strong correlation with the censor variable.

```
ggplot(AIDSdata, aes(x=jitter(strat2), y=jitter(censor))) + geom_point()
```



I made a few other visualizations but didn't really feel like they were useful enough to include.
