TV Watching Habits Study

Cameron Bell, Alyssa Eisenberg, Sarah Cha

4/26/2018

Abstract

Television is consumed through a growing number of channels today (cable, streaming networks like Hulu, Amazon, Netflix, and Youtube) while increasingly the mould for what looks like television is becoming more fluid (e.g. video clips on Youtube). We suspect some individuals don't actually realize how much television they're watching on a daily basis and we set out to understand how social pressures around television watching would alter or impact individual behavior. In our study, we randomly assigned treatment emails to half of the participants that highlights how much their peers are watching TV on average. XXX

Introduction

Nielsen Total Audience report for the fourth quarter of 2016 discussed the steady trend of decline in live TV watching alongside increased use of smartphones to view video and the increasing penetration of subscription video on demand. While live viewing decreased to 4 hours and 23 minutes a day among adults 18 years and up (from 4 hours and 27 minutes in the prior year), smartphone usage rose to 2 hours and 32 minutes a day from 1 hour and 15 minutes a year ago. Needless to say, individuals are consuming video content in more formats today and given the rate at which online and on-demand content is growing, consumers can access an endless amount of content right at their fingertips. This points to a trend of increased video/screen watching and could mean that consumers are watching more TV than they actually realize or desire.

In other contexts, it has been found that providing peer comparisons can cause changes in people's behavior. For example, in the Opower study by Ian Ayres, Sophie Raseman, and Alice Shih, they find that providing mailed peer feedback reports resulted in reductions in energy consumption. We are interested in whether a similar mechanism can be applied to the new context of TV watching habits.

We set up an experimental study where we randomly assigned treatment emails to half of the participants that summarizes how much their peers are watching television on average. We wanted to understand if social/peer pressure has the ability to impact how much television a person watches (e.g. "Wow, my peers are watching only 2 hours of TV a week and I'm watching at least 10. Shame on me.").

We begin by discussing the high level experimental design we developed to test whether social pressures/influence can actually impact individual behavior as it relates to TV watching, including changes based on an initial pilot study. Next, we discuss the details of our real experiment - the participants, the nature and frequency of the treatment, and our outcome measures. Statistical analysis from our study suggests that XXXXXXX taking into account non-compliance and attrition aspects of the experiment. We conclude by discussing the implications for generalizability to a broader population at large and future potential avenues for exploration.

Experimental Design

Our aim was to develop a means for studying the influence of social pressures on TV watching, recognizing that the means by which Americans consume TV is very diverse today (cable, online video streaming, YouTube, etc) and what people classify as television can vary quite a bit (e.g. television shows, news, sports, movies, YouTube clips, etc). We hypothesized that when we carefully defined what we classify as "television" watching and participants could compare apples to apples their consumption vs. peers, they'd be impacted if their peers were watching less than they were. We also hypothesized that we'd see heterogeneous treatment effect from our participants based on amount of TV watched prior to experiment (ie. those in 1st quartile are not likely to reduce their amount of TV and may actually see an increase instead).

$$\begin{array}{ccccc} R & O_{1,2} & X & O_{1,2} \\ R & O_{1,2} & & O_{1,2} \end{array}$$
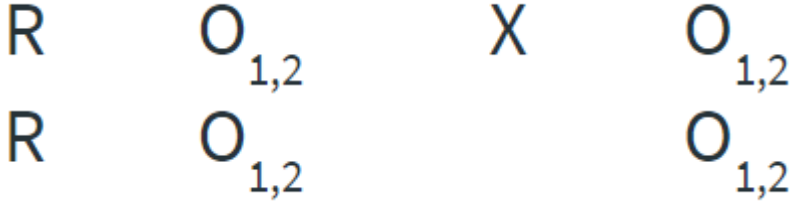
Figure 1: Experimental Design

We decided on surveys as the means by which we would collect all the necessary information including outcome measures from our participants. Our experimental design follows a difference in difference approach. Our experiment consisted of two surveys - one pre and one post treatment. In the pre-survey, we would measure pre-survey TV behavior for all participants along with what we thought could be key covariates of interest (age, gender, marital/employment status, etc). Then we would randomly assign individuals to treatment vs. control, where the individuals in the treatment group would be aware of how much their peers were watching. The post-survey measures TV behavior during the experiment. This allows us to compare the hours of TV watched between treatment and control group for causal estimate conditioning on pre-experiment TV watching measure as a covariate.

In designing our study, we were very mindful of how to maximize participation to get a large enough sample size, get accurate reportings, and how we could extract the most information in the least intrusive way. On the latter, conversations with our peers made us realize that there was clearly a balance between incentives for participants and the time commitment involved. We offered a drawing for 5 Amazon gift cards for $20 each as a reward for participating, with the stipulation that participants had to complete a 3 minute pre-survey and a 1 minute post-survey in order to be entered into the drawing.

Pilot Study We implemented a pilot study with 9 individuals to test our initial design framework and help identify potential problems and inform our power analysis. The pilot took place over the course of 1 week (Feb 26 - Mar 4) where friends and family answered a pre-survey, were randomly assigned to control or treatment, treatment group was sent an email about peer usage, and then everyone filled out a post-survey on their TV consumption. See the Appendix for full questionaires.

Learnings and adjustments in the experimental design:

1. Attrition is likely to be high. Of the 9 close family and friends that we targeted for the survey, we ended up with full observation sets for only 7 of 9. Of course the short observation time had an impact in the case of the pilot. To reduce attrition as much as possible in order to maximize the inference ability of the experiment, we incorporated more time for individuals to fill out both the pre- and post-survey and planned to send several reminder emails emphasizing the Amazon gift card drawing for completing participation in the study.

2. The quality of the data coming out of the Chrome extension reading was low - 3 of 7 individuals had 0 or NA output even though they indicated they watched TV during that time. This emphasized that we needed to incorporate more optional language around the Chrome Netflix extension. We didn't want to make installation of that extension and its output such a strong requirement with the fear that it would reduce our subject pool too severely.

3. The data from the self-reported average amount of TV watched did not have much variability. This was in part due to the low number of participants, but also because the response options were a limited set of ranges of hours. This made the treatment less effective in showing variation, so we changed the question about pre-experiment self-reported TV to have participants enter their own value instead of selecting a radio button.

4. We emailed the pre- and post-surveys out via Qualtrics with the sender name as "UCBerkeley TV Habits Study" but the treatment email came from Alyssa's gmail, raising the question if this created

confusion and/or potential for individuals to ignore Alyssa's emails (non-compliance). To resolve this, we created a gmail account with the same title as the Qualtric email sender which is "UCBerkeley TV Habits Study."

Calculating Power

The pilot also provided reference information for a power calculation that would help us understand what level of treatment effect we would be able to accurately identify. We ran a simulation regressing outcome TV on treatment and using the pre-experiment amount of TV watched as a covariate to reduce our SE and increase the power, and checked the proportion that found a statistically significant coefficient for treatment. We made the following simplifying assumptions for our simulation:

- set mean and variance for self-reported pre-experiment TV hours watched
- each individuals potential outcomes to control are normally distributed with mean of that individuals TV prior to experiment, and a SD value provided as an input
- treatment effect is constant for everyone

Since the pilot was so short, we did not get a good estimate of the treatment effect and we have no data for individual SD for TV watched. Instead, we run across a variety of treatment effects and SD values and check the power for that combination of parameters.

First, we ran this analysis using the pilot mean and variance for pre-experiment hours of TV (7.6 and 5.25 respectively). We found that for 100 participants and an individual SD of 2.5, we would need a treatment effect of -1.5 hrs/week to have 80% power. We updated and re-ran this analysis after gathering the pre-survey data for the real experiment, leading to the chart below. Here we only use 80 participants, and the mean and variance for pre-experiment hours of TV from the pre-survey (8.6 and 6.25). If the individual SD was about half that of the overall population (3), we would need a treatment effect of -2 hrs/week to have 80% power.

Power Level by Treatment Effect Size and Individual SD
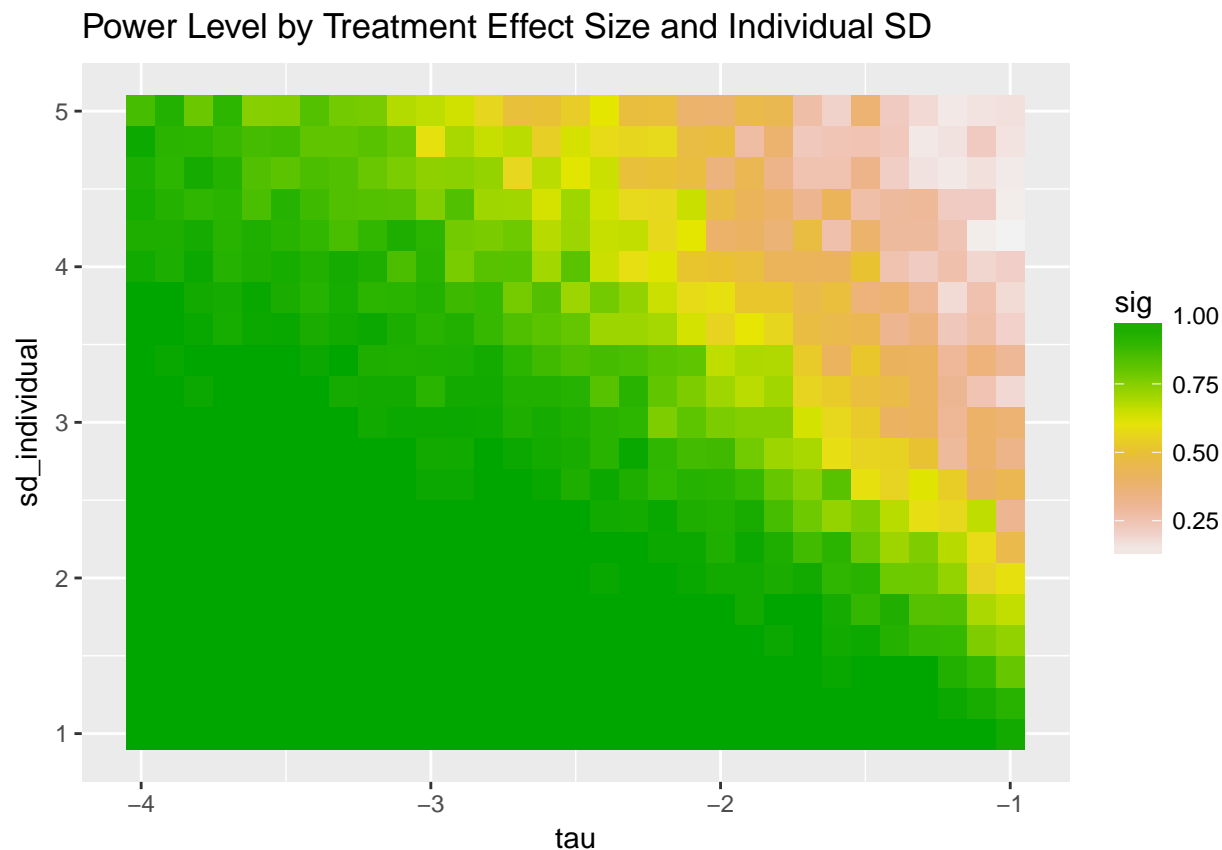
The Actual Experiment

Subjects

In an ideal scenario, we would survey a large sample of the US adult population in order to get a representative sample for our experiment. However, with the key constraints being time and resources (we have only about 10 weeks and $500), we implemented a much more limited study where we, ourselves, were tasked with finding our respondents by word of mouth and social media.

We recruited on college campuses and posted on social media to get participants. We chose to limit participants to one per household to avoid problems with non-interference. Even though we used different methods of outreach including word of mouth, social media, and email, we used standardized pitch with key talking points. This ensured that all three of us were giving out the same details about the experiment and, more importantly, that none of us provided too much detail on what we were aiming to do/analyze to avoid participants skewing their answers. See appendix for the standardized pitch used.

Initially we took down email addresses of all interested participants. There were 136 email addresses of potential participants of which 42 came from BYU, 38 from social media websites like Facebook and Instagram, 37 from UCBerkeley Slack, and 19 friends of Alyssa, Cameron, and Sarah.

Everyone on that list was emailed a pre-survey that they had to fill out to be a part of the actual experiment. Just as with the pilot, the pre-survey asked all participants about their initial TV watching habits and covariates at the beginning of the experiment. Useful potential covariates included age, martial status, and pre-existing TV watching habits. There was about 40% drop out rate with only 81 of the 136 individuals filling out the pre-survey by the deadline for the experiment (another 2 completed the survey after treatment/control was already randomized).

Random assignment of treatment and controls

In terms of randomization once we had our set list of participants from the pre-survey, we used a simple random assignment to get half of our subjects into control and the other half in treatment in an effort to prevent against any selection bias. Those assigned to control received an email thanking them for their participation, while those in treatment received the same email with a chart comparing their amount of TV watched with their peers (see Appendix for examples).

For a check on our randomization before sending these emails, we did a covariate balance check and found that the randomization had worked. There were no statistically significant differences between the control and treatment group on key variables. Also, the F-test of joint significance on a regression of treatment assignment on all our key covariates did not return anything of significance (p-value of 67%). For the full model results, see the Appendix.

variable	control	treatment	diff	se
hours TV	8.74	8.41	0.33	1.40
male	0.48	0.41	0.07	0.11
married	0.43	0.54	-0.11	0.11
no children	0.81	0.77	0.04	0.09
watched more than wanted: couple times a year	0.29	0.38	-0.09	0.11
watched more than wanted: once a month	0.14	0.15	-0.01	0.08
watched more than wanted: couple times a month	0.24	0.18	0.06	0.09
watched more than wanted: once a week	0.14	0.13	0.01	0.08

Outcome Measures

We carefully define in our communication with participants what we consider to be TV consumption (movies, tv shows, or video clips through cable, streaming online, or Youtube videos). Our primary outcome measure asks survey respondents to estimate and record how many hours they watched during the two weeks of the experiment.

For our secondary outcome measure, we created a Chrome Netflix extension that would track and display all Netflix usage an individual had during a set period for the individual to then input the numbers directly into the post-survey. The goal was to avoid measurement error from self-reporting due to issues in memory or miscounting.

Results

```
#Post-survey questions
str(all_data)
```

```
## 'data.frame':    81 obs. of  29 variables:
##  $ linkedEmail            : Factor w/ 81 levels "ace9312@gmail.com",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ enteredEmail           : Factor w/ 81 levels "ace9312@gmail.com",..: 1 2 3 4 5 17 6 7 8 18 ...
##  $ gender                 : Factor w/ 2 levels "male","female": 1 2 1 2 2 2 1 2 2 1 ...
##  $ age                    : Factor w/ 4 levels "21-","22-34",..: 2 1 3 4 2 2 2 2 2 2 ...
##  $ region                 : Factor w/ 6 levels "midwest","northeast",..: 5 5 2 5 2 2 5 5 5 1 ...
##  $ employment             : Factor w/ 6 levels "full","part",..: 4 2 4 2 5 1 1 1 4 1 ...
##  $ maritalStatus          : Factor w/ 3 levels "single","married",..: 2 1 3 2 2 2 2 1 1 1 ...
##  $ children               : Factor w/ 2 levels "yes","no": 2 2 1 1 2 2 2 2 2 2 ...
##  $ hoursTV                : num  10 5 15 3 5 10 4.5 12 8 14 ...
##  $ binge                  : Factor w/ 5 levels "once a week",..: 3 5 5 5 3 2 4 2 3 5 ...
##  $ primaryChannel         : Factor w/ 7 levels "netflix","HBO",..: 1 1 1 1 1 1 1 1 1 1 1 ...
##  $ allMethods             : Factor w/ 36 levels "1","1,2","1,2,3,4",..: 25 24 13 18 23 18 11 9 24 19
##  $ moreTimeThanWanted     : Factor w/ 5 levels "once a year",..: 4 5 5 2 2 3 4 4 5 4 ...
##  $ watchAlone             : Factor w/ 2 levels "alone","withOthers": 2 2 1 2 2 1 1 2 1 1 ...
##  $ shareProfile           : Factor w/ 2 levels "yes","no": 1 1 2 2 1 1 1 1 2 2 ...
##  $ netflixDays            : Factor w/ 12 levels "","0","1","2",..: 2 NA 5 1 3 11 NA 3 5 2 ...
##  $ netflixHours           : Factor w/ 19 levels "","0","10","11",..: 15 NA 8 1 10 19 NA 3 4 9 ...
##  $ netflixMin             : Factor w/ 28 levels "","10","11","17",..: 17 NA 21 1 4 27 NA 22 11 12 ..
##  $ netflixAccountAndChrome: Factor w/ 4 levels "both","noNetflix",..: 1 2 1 3 1 1 1 1 1 1 ...
##  $ source                 : Factor w/ 4 levels "BYU","Friend",..: 1 1 4 3 3 2 4 4 3 3 ...
##  $ post_survey_ID         : Factor w/ 65 levels "ace9312@gmail.com",..: 1 2 3 4 5 6 NA 7 8 9 ...
##  $ reportedhrs_P          : num  20 10 20 14 9 10 NA 8 13 30 ...
##  $ netflixDays_P          : num  1 NA 1 3 1 NA NA 1 1 NA ...
##  $ netflixHours_P         : num  1 NA 6 15 11 NA NA 10 2 NA ...
##  $ netflixMin_P           : num  22 NA 5 23 5 NA NA 18 4 NA ...
##  $ Treated                : Factor w/ 3 levels "","No","Yes": 1 2 3 1 2 2 NA 3 3 2 ...
##  $ Finished               : num  1 1 1 1 1 1 0 1 1 1 ...
##  $ NA                     : num  1 1 1 1 1 1 NA 1 1 1 ...
##  $ Assigned               : num  0 1 1 0 1 1 0 1 1 1 ...
```

```
#Self-reported TV watching
t9 <- round(mean(all_data$reportedhrs[all_data$Assigned==1 & all_data$Finished==1], na.rm=T),2)
c9 <- round(mean(all_data$reportedhrs[all_data$Assigned==0 & all_data$Finished==1], na.rm=T),2)
diff9 <- t9-c9
se9 <- se_diff_means(as.numeric(complete.cases(all_data$reportedhrs[all_data$Assigned==1])),as.numeric(
```

```
#Netflix extension readings
all_data$NFLXreading <- (all_data$netflixDays_P*24) + all_data$netflixHours_P + (all_data$netflixMin_P/(
t10 <- round(mean(all_data$NFLXreading[all_data$Assigned==1], na.rm=T),2)
c10 <- round(mean(all_data$NFLXreading[all_data$Assigned==0], na.rm=T),2)
diff10 <- t10-c10
se10 <- se_diff_means(as.numeric(complete.cases(all_data$NFLXreading[all_data$Assigned==1])),as.numeric
```

```
#Received treatment email
#recode Yes = 1, No = 0, Blank/NA = NA
levels(all_data$Treated) <- c("Yes", "No", "")
all_data$Treated <- as.numeric(all_data$Treated)
all_data$Treated[is.na(all_data$Treated)] <- 0
all_data$Treated <- ifelse(all_data$Treated == 3, 1, 0)
all_data$Treated
```

```
##  [1] 0 0 1 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0
## [36] 0 0 0 1 0 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0
## [71] 1 0 0 0 1 0 0 0 0 0 0
```

```
#% of assigned that were actually treated
compl <- round(mean(all_data$Treated[all_data$Assigned == 1]), 2) #compliance rate is low --> 36% of th
compl2 <- round(mean(all_data$Treated[(all_data$Assigned == 1) & (all_data$Finished == 1)]), 2) #actual
compl
```

```
## [1] 0.36
```

```
compl2
```

```
## [1] 0.43
```

```
# Put into a table for display
outcomes <- data.frame(variable = c("Self-ReportedTV", "NetflixExt"),
                control = c(c9, c10),
                treatment = c(t9, t10),
                diff = c(diff9, diff10),
                se = c(se9, se10))
knitr::kable(outcomes)
```

variable	control	treatment	diff	se
Self-ReportedTV	17.48	17.73	0.25	0.09
NetflixExt	70.39	67.16	-3.23	0.11

Summary of initial findings

Why are the netflix readings so much bigger than the self-reported #s?
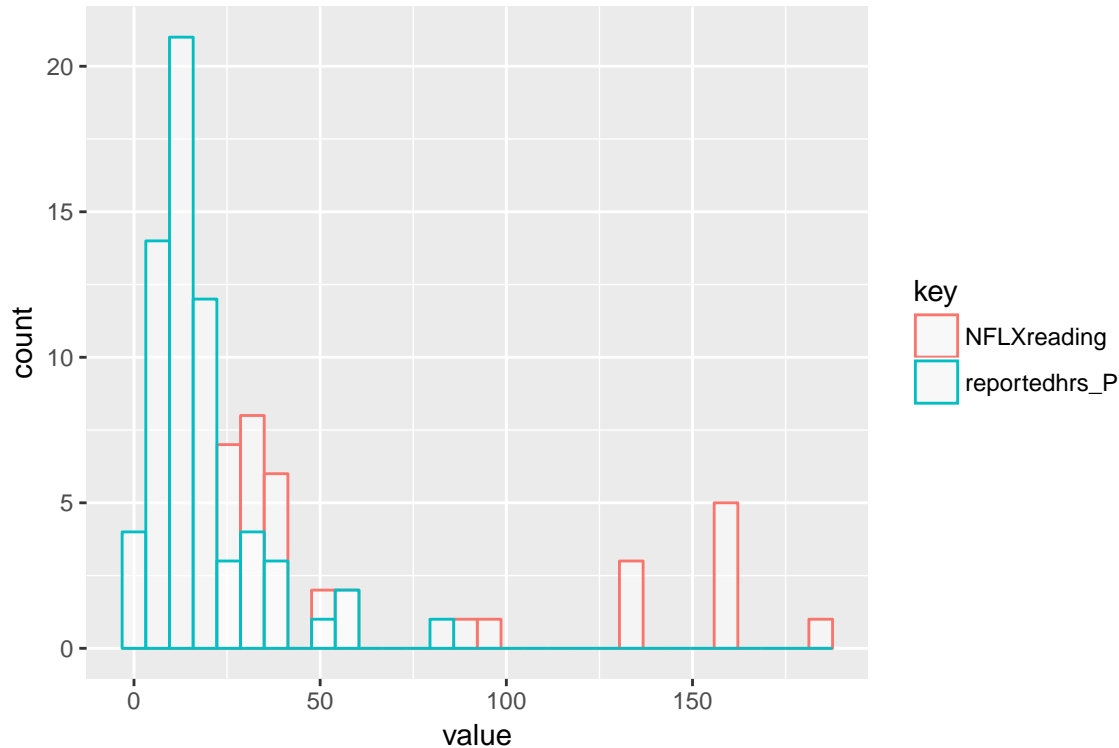
```
hours_watched <- all_data[ , c("reportedhrs_P", "NFLXreading")]

t <- hours_watched %>%
  gather(key, value)

ggplot(t, aes(x = value, color = key)) +
  geom_histogram(fill="white", alpha=0.5, position="identity")
```

The below shows us that there were only 5 participants who used the netflix extension and do not share a profile. This is simply not enough data, so it was unusable.

```
all_data$extension <- is.na(as.numeric(all_data$netflixMin))
NFmod <- lm(reportedhrs_P/2 ~ Assigned + source + NFLXreading + extension, data=all_data[all_data$shareP
summary(NFmod)
```

```
##
## Call:
## lm(formula = reportedhrs_P/2 ~ Assigned + source + NFLXreading +
##     extension, data = all_data[all_data$shareProfile == "no",
##     ])
##
## Residuals:
##         3          4          9         11         71
## -1.971e+00  1.917e-01 -4.441e-16 -1.917e-01  1.971e+00
##
## Coefficients: (1 not defined because of singularities)
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           26.08596    5.38721   4.842    0.130
## Assigned             -13.83538    4.50905  -3.068    0.201
## sourceUC Berkeley Slack  6.35665    3.43161   1.852    0.315
## NFLXreading           -0.22061    0.07989  -2.762    0.221
## extensionTRUE              NA         NA      NA       NA
##
## Residual standard error: 2.8 on 1 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.9177, Adjusted R-squared:  0.671
## F-statistic: 3.719 on 3 and 1 DF,  p-value: 0.3601
```

```
hist(na.omit(all_data$NFLXreading[all_data$shareProfile=="no"]))
```

## Histogram of na.omit(all_data$NFLXreading[all_data$shareProfile == "ı
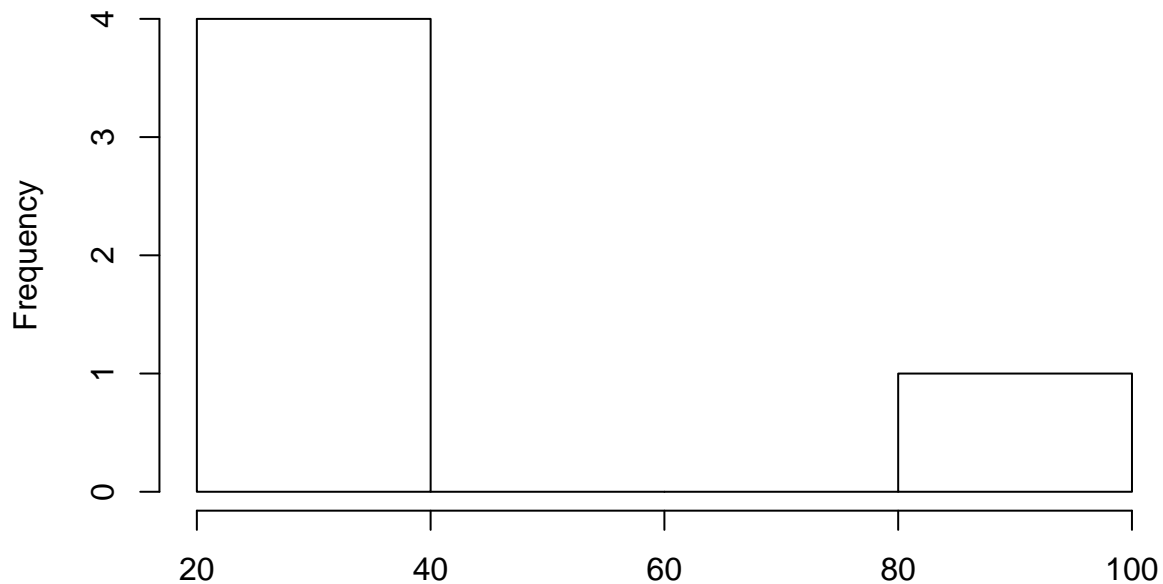
na.omit(all_data$NFLXreading[all_data$shareProfile == "no"])

In the below plots we see that participating in the study did not affect the control group or the non-compliers in treatment. Though we recognize that asking the participants in the final survey if they ever received the treatment email acts as a poor proxy for identifying compliers, it allows us to calculate some somewhat-useful values for alpha and CACE (see below). We control for the only statistically significant covariate, previous hours/week.

```
boxplot(all_data$reportedhrs_P[all_data$Assigned==0]/2,
        all_data$hoursTV[all_data$Assigned==0],
        names=c('After', 'Before'), main='Hours Watched (Control)', xlab='Hours of TV Watched per Week'
        ylim=c(0,40), horizontal=T)
```

Hours Watched (Control)
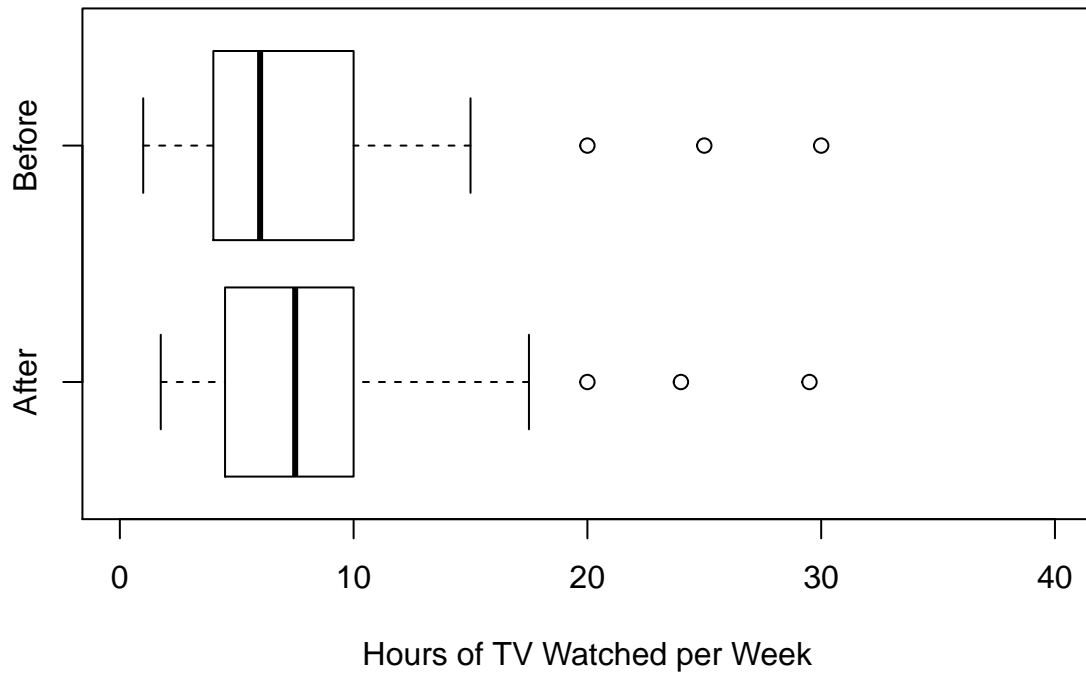

```r
boxplot(all_data$reportedhrs_P[all_data$Assigned==1 & all_data$Treated==1]/2,
        all_data$hoursTV[all_data$Assigned==1 & all_data$Treated==1],
        all_data$reportedhrs_P[all_data$Assigned==1 & all_data$Treated!=1]/2,
        all_data$hoursTV[all_data$Assigned==1 & all_data$Treated!=1],
        names=c('After', 'Before', 'After', 'Before'), main='Hours Watched (Treatment) per Week', xlab=
        ylim=c(0,40), horizontal=T)
```

Hours Watched (Treatment) per Week
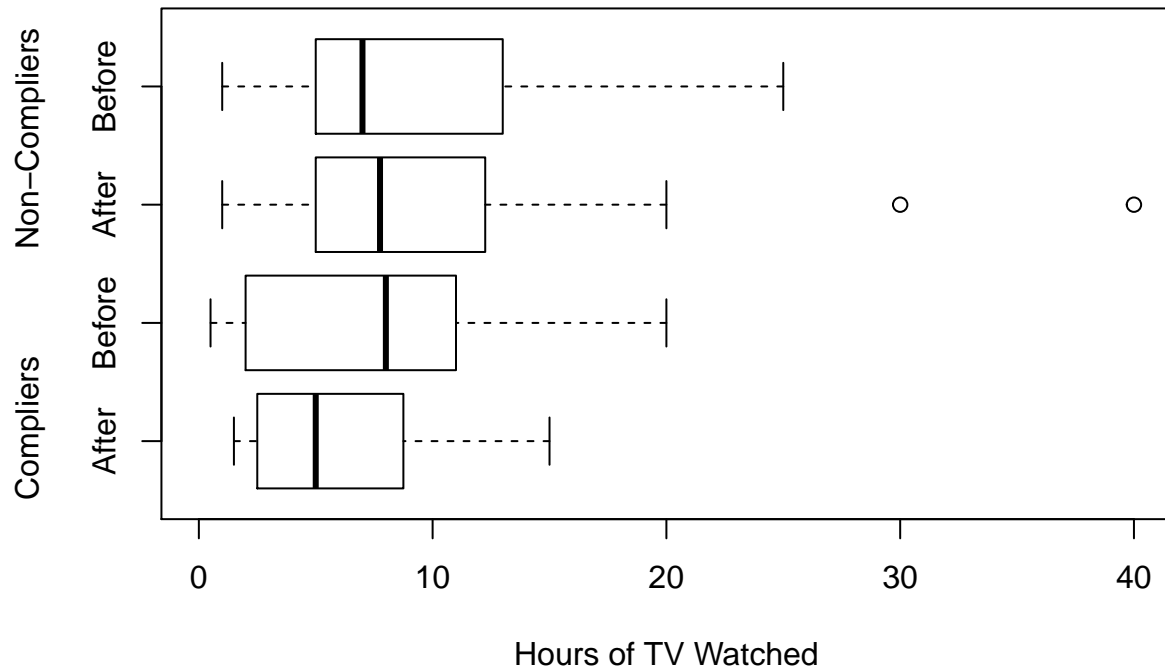

```r
# The below section considers before/after analysis for compliers and non-compliers,
# which produces an inaccurate view.
# BNC <- na.omit(all_data$hoursTV[all_data$Assigned==1 & all_data$Treated!=1])
# BC <- na.omit(all_data$hoursTV[all_data$Assigned==1 & all_data$Treated==1])
# ANC <- na.omit(all_data$reportedhrs_P[all_data$Assigned==1 & all_data$Treated!=1]/2)
# AC <- na.omit(all_data$reportedhrs_P[all_data$Assigned==1 & all_data$Treated==1]/2)
# print("ATE for Non-Compliers:")
# (ATENC <- mean(ANC)-mean(BNC))
# print("ATE for Compliers:")
# (ATEC <- mean(AC)-mean(BC))

print("Overall ATE:")
```

```
## [1] "Overall ATE:"
```

```r
(ATE <- lm(reportedhrs_P/2 ~ Assigned + hoursTV, all_data)$coefficients[2])
```

```
##   Assigned
## -0.5933588
```

```r
print("Alpha (rate of compliance):")
```

```
## [1] "Alpha (rate of compliance):"
```

```r
(alpha <- length(all_data$hoursTV[all_data$Assigned==1 & all_data$Treated==1]) / length(all_data$hoursTV
```

```
## [1] 0.3571429
```

10

```
print("CACE:")
```

```
## [1] "CACE:"
```

```
(CACE <- ATE/alpha)
```

```
##  Assigned
## -1.661405
```

When we create the full model, we see that the only statistically significant covariates are hoursTV, which tells us the hours of TV watched per week previous to the study, and source, where we found the study participant. Below we discuss a few heterogeneous treatment effects we expected to see (but did not), and the treatment effect itself: Assigned: This is the variable that indicates assignment to treatment. There was no statistically significant effect of receiving treatment as a whole. Treated: Reporting that a participant was a complier also had no effect. Assigned*Treated: We also saw no effect for the interaction term. We suspect that source (online/social media) may reflect a statistically significant heterogeneous treatment effect because it ts as a better proxy than Treated for identifying compliers. those who joined the study after recruitment on social media are likely close friends of those conducting the study and were more likely to pay attention to the process. extension: We expected to see that those who had installed the Chrome extension would have been shocked by how much Netflix they were watching (we were!) and reduced their watching habits due to that observation alone. However, there was no such effect. first_quartile: Similar to the OPower study, we expected that those who already watch very little TV would be the least affected by the study. However, there was no such statistically significant effect (likely because the treatment itself was not statistically significant).

```
all_data$first_quartile <- all_data$hoursTV < quantile(all_data$hoursTV, 0.25)
m <- lm(reportedhrs_P/2 ~ Assigned * Treated + gender + age + region + employment + maritalStatus + chil
stargazer(m, type='text', omit=c('region', 'age', 'employment', 'maritalStatus', 'binge', 'children', 'C
```

```
##
## =========================================================
##                            Dependent variable:
##                      ------------------------------
##                              reportedhrs_P/2
## ---------------------------------------------------------
## Assigned                         -1.155
##                                  (2.132)
##
## Treated                          -3.201
##                                  (2.635)
##
## genderfemale                     -2.379
##                                  (2.126)
##
## watchAlonewithOthers             -0.218
##                                  (2.113)
##
## shareProfileno                   -1.143
##                                  (3.575)
##
## sourceFriend                      1.412
##                                  (5.252)
##
## sourceOnline (social media)      6.772*
##                                  (3.758)
##
## sourceUC Berkeley Slack           3.523
```

```
##                                               (4.450)
##
## hoursTV                                       0.540**
##                                               (0.237)
##
## extension                                     -0.339
##                                               (2.492)
##
## first_quartile                                -1.486
##                                               (2.976)
##
## Assigned:Treated
##
##
## ----------------------------------------------------------
## Observations                                  65
## Residual Std. Error                    5.864 (df = 34)
## ==========================================================
## Note:                          *p<0.1; **p<0.05; ***p<0.01
```

Attrition We witnessed high levels of attrition through out the experiment, starting with 136 potential participants and ending up with 65 actually completing both pre and post surveys. From the time the participants were assigned to treatment and control groups, we saw 17% attrition in the treatment group and 23% in the control. When we regressed completion on treatment group, we found there was so statistical evidence to suggest that attrition was related to treatment group (p-value of .475). Further, ex-ante covariate balance check fails to reject null of no joint significance of covariates predicting treatment assignment.

```
attrit_cases <- all_data[all_data$Finished == 0 ,]
ggplot(attrit_cases, aes (x=Assigned)) +
  geom_histogram(binwidth = 0.5, fill="darkblue", col="grey") +
  scale_x_continuous(breaks = seq(0,1,1)) +
  ggtitle("Comparing attrition for control(0) vs. treatment (1)")
```

Comparing attrition for control(0) vs. treatment (1)
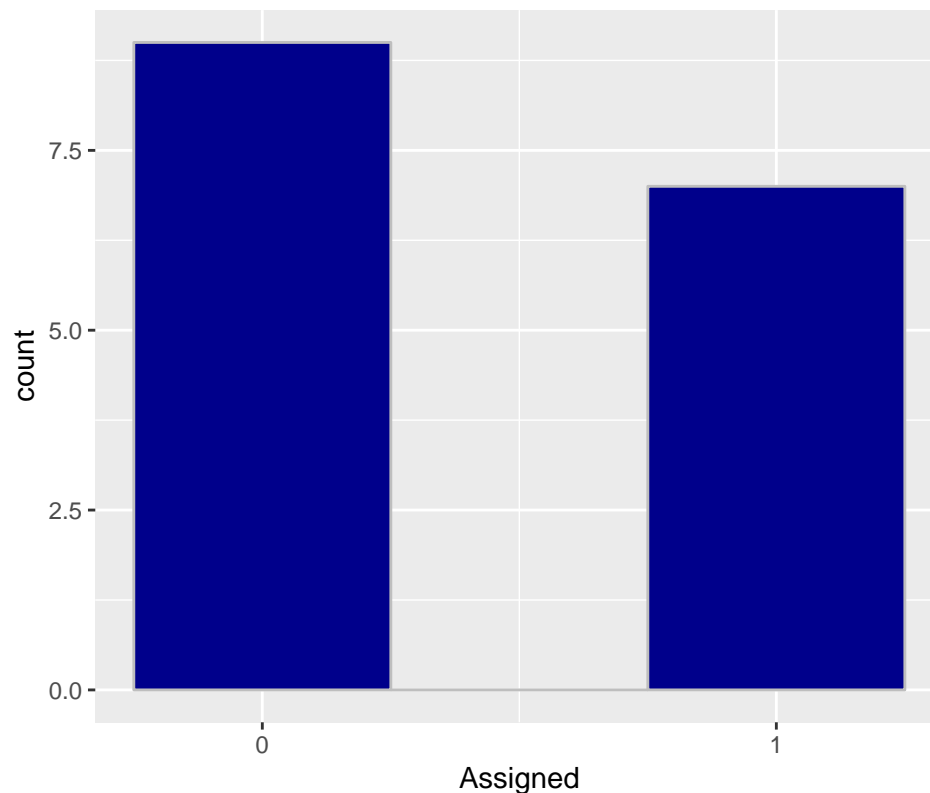

```
#percent attrition by assignment
xtabs(~ Finished + Assigned, data=all_data)
```

```
##         Assigned
## Finished  0  1
##        0  9  7
##        1 30 35
```

```
round(prop.table(xtabs(~ Finished + Assigned, data=all_data)),2)
```

```
##         Assigned
## Finished    0    1
##        0 0.11 0.09
##        1 0.37 0.43
```

```
attrit_by_assigned <- lm(Finished ~ Assigned, data=all_data)
summary(attrit_by_assigned)
```

```
##
## Call:
## lm(formula = Finished ~ Assigned, data = all_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8333  0.1667  0.1667  0.2308  0.2308
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)   0.76923     0.06435  11.955   <2e-16 ***
## Assigned       0.06410     0.08936   0.717    0.475
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4018 on 79 degrees of freedom
## Multiple R-squared:  0.006472,   Adjusted R-squared:  -0.006104
## F-statistic: 0.5146 on 1 and 79 DF,  p-value: 0.4753
```

```r
attrit_covariate_check <- lm(Assigned ~ gender + age + region + employment + maritalStatus + children +

summary(attrit_covariate_check)
```

```
##
## Call:
## lm(formula = Assigned ~ gender + age + region + employment +
##     maritalStatus + children + hoursTV + binge + primaryChannel +
##     moreTimeThanWanted + watchAlone + shareProfile + source,
##     data = all_data[all_data$Finished == 1, ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.7925 -0.2648  0.0000  0.1926  0.7656
##
## Coefficients:
##                                 Estimate Std. Error t value
## (Intercept)                     0.780280   0.742524   1.051
## genderfemale                    0.011888   0.222420   0.053
## age22-34                       -0.730372   0.347550  -2.101
## age35-44                       -0.617413   0.501151  -1.232
## age45-54                       -1.488938   0.612322  -2.432
## regionnortheast                -0.088633   0.443604  -0.200
## regionsoutheast                -1.473855   0.716649  -2.057
## regionsouthwest                -0.249889   0.976121  -0.256
## regionwest                     -0.099272   0.418580  -0.237
## regionoutsideUS                -0.368271   0.788755  -0.467
## employmentpart                  0.230664   0.406877   0.567
## employmentstudent              -0.067821   0.366886  -0.185
## employmenthomemaker             0.616241   0.701882   0.878
## employmentself                  0.306065   0.533251   0.574
## maritalStatusmarried            0.357057   0.214726   1.663
## maritalStatusdivorced           0.122356   0.851637   0.144
## childrenno                      0.108830   0.296727   0.367
## hoursTV                         0.010582   0.019846   0.533
## bingeonce a month              -0.104093   0.400094  -0.260
## bingeonce every couple months   0.014922   0.398400   0.037
## bingeonce a year               -0.261701   0.442231  -0.592
## bingeno                         0.100424   0.411776   0.244
## primaryChannelHBO              -0.181192   0.689853  -0.263
## primaryChannelhulu             -0.017222   0.332850  -0.052
## primaryChannelamazon            0.994517   0.740814   1.342
## primaryChannelyoutube          -0.004435   0.445836  -0.010
## primaryChannelcable             0.142984   0.539320   0.265
## primaryChannelother             0.412018   0.328701   1.253
## moreTimeThanWantedcouple times a year  -0.079962   0.337992  -0.237
```

```
## moreTimeThanWantedonce a month             0.146948   0.351999   0.417
## moreTimeThanWantedcouple times a month    0.156936   0.342818   0.458
## moreTimeThanWantedonce a week              0.021999   0.401548   0.055
## watchAlonewithOthers                      -0.266044   0.218361  -1.218
## shareProfileno                             0.117118   0.374235   0.313
## sourceFriend                               0.062644   0.504726   0.124
## sourceOnline (social media)                0.222885   0.332739   0.670
## sourceUC Berkeley Slack                    0.472965   0.396673   1.192
##                                           Pr(>|t|)
## (Intercept)                                 0.3023
## genderfemale                                0.9578
## age22-34                                    0.0447 *
## age35-44                                    0.2282
## age45-54                                    0.0217 *
## regionnortheast                             0.8431
## regionsoutheast                             0.0492 *
## regionsouthwest                             0.7998
## regionwest                                  0.8143
## regionoutsideUS                             0.6442
## employmentpart                              0.5753
## employmentstudent                           0.8547
## employmenthomemaker                         0.3874
## employmentself                              0.5706
## maritalStatusmarried                        0.1075
## maritalStatusdivorced                       0.8868
## childrenno                                  0.7165
## hoursTV                                     0.5981
## bingeonce a month                           0.7966
## bingeonce every couple months               0.9704
## bingeonce a year                            0.5587
## bingeno                                     0.8091
## primaryChannelHBO                           0.7947
## primaryChannelhulu                          0.9591
## primaryChannelamazon                        0.1902
## primaryChannelyoutube                       0.9921
## primaryChannelcable                         0.7929
## primaryChannelother                         0.2204
## moreTimeThanWantedcouple times a year       0.8147
## moreTimeThanWantedonce a month              0.6795
## moreTimeThanWantedcouple times a month      0.6506
## moreTimeThanWantedonce a week               0.9567
## watchAlonewithOthers                        0.2333
## shareProfileno                              0.7566
## sourceFriend                                0.9021
## sourceOnline (social media)                 0.5084
## sourceUC Berkeley Slack                     0.2431
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5406 on 28 degrees of freedom
## Multiple R-squared:  0.4935, Adjusted R-squared:  -0.1577
## F-statistic: 0.7578 on 36 and 28 DF,  p-value: 0.7853
```
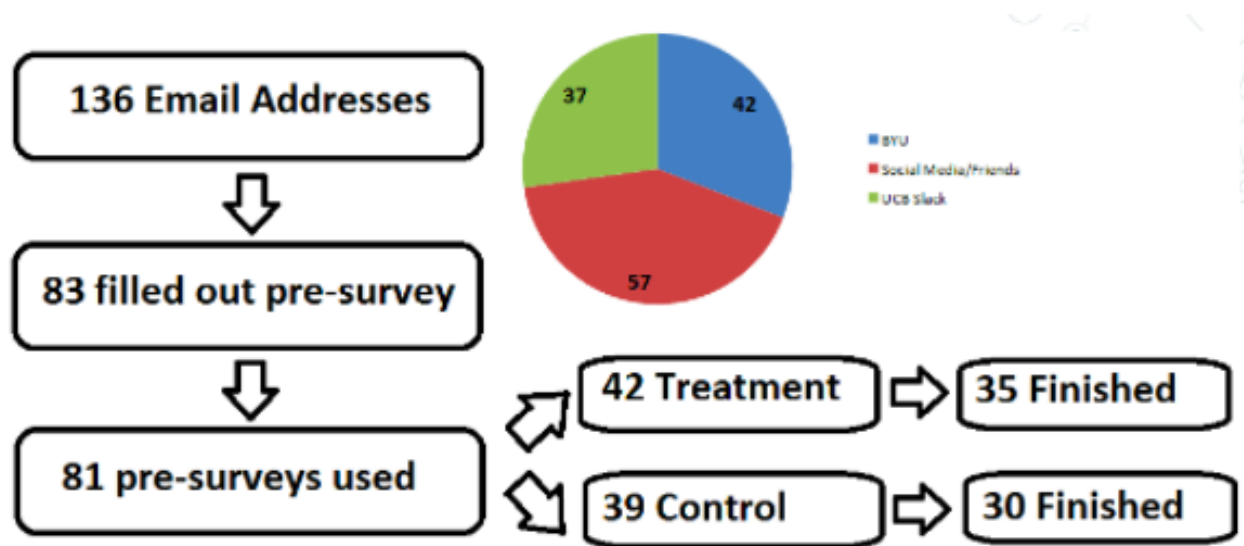
Model Building

Figure 2: Flow Diagram for Participants

In the exploratory data analysis portion of the model building, a few things stood out to us in terms of high level trends. First of all, the pre-survey "TV hours watched" covariate appeared most negative correlated with tendency to binge (ie. those who indicated a greater tendency to binge ended up logging less pre-survey hours) and tendency to watch TV alone (vs. with someone else or a group). On the other hand, the pre-survey "TV hours watched" appears more positively correlated with feelings of regret or that they had spent more time watching TV than they would have liked. In terms of the key outcome measure of reported hours of TV watched during the survey period, as we suspected the correlation matrix below seems to suggest there is a strong positive correlation between pre-survey TV watching and survey TV watching. We also again see a reasonably high negative correlation between the outcome measure and the tendency to binge.

```r
#reordering some of the factor variables to make more intuitive sense for EDA
levels(all_data$binge) = c("no", "once a year", "once every couple months", "once a month", "once a wee
levels(all_data$employment) =c("looking", "homemaker", "student", "part", "self", "full")
levels(all_data$children) = c("no", "yes")
levels(all_data$shareProfile) =c("no", "yes")
levels(all_data$watchAlone) = c("withOthers", "alone")
```

```r
# Make correlation plot
cols <- c(3:15, 20, 22, 30)
all_data_numeric <- all_data[cols]
colnames(all_data_numeric) =c("gender", "age", "region", "employment", "martial status", "children", "h
DF <- data.frame(lapply(all_data_numeric,as.numeric))
corrmatrix = round(cor(DF[, sapply(DF, is.numeric)], use = "na.or.complete"),
2)
corrplot(corrmatrix)
```
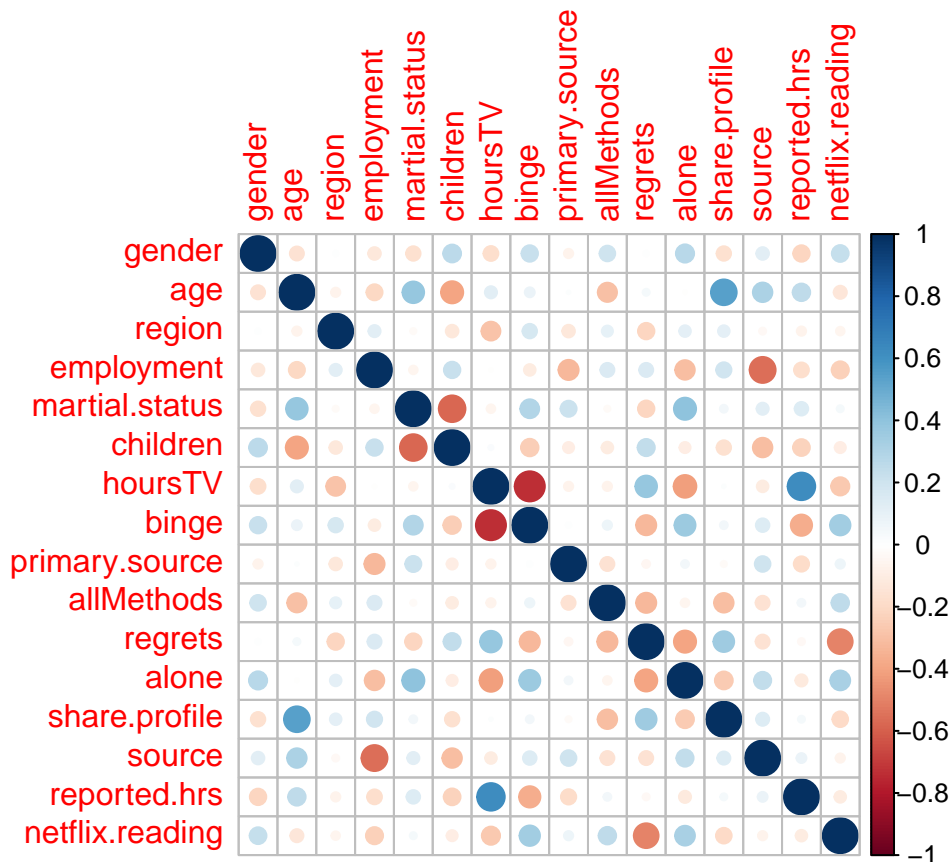
We conducted further EDA for other variables like age, region, and marital status to see if there were any noteworthy patterns. We immediately ruled out age and employment because there were some pretty big clusters in the data that made it hard to get sizable groups (for example, more than 75% of our participants were in the 22-34 age and over 60% of participants said they were currently looking for jobs). By including them in the equation we'd likely end up with some pretty small groups which could skew the comparison quite a bit. Further, most of our participants were from the Northeast and West Coast which wasn't a huge surprise to us given how we targeted participants but the average TV_hours watched during the survey period by Northeast vs. West Coast doesn't appear to be all that different upon exploring the data. As such, we decided not to include region in the model.

```r
all_data$TV_bin<- cut(all_data$reportedhrs_P, breaks=c(0,10,20,30,40,50,60,70,80),
                      labels=c("0-10","10-20", "20-30", "30-40", "40-50", "50-60", "60-70","70-8
xtabs(~all_data$employment + all_data$TV_bin)
```

```
##                    all_data$TV_bin
## all_data$employment 0-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80
##          looking      11    15     7     2     1     1     0     1
##          homemaker     2     2     1     0     0     0     0     0
##          student       0     0     0     0     0     0     0     0
##          part          7     9     0     1     0     1     0     0
##          self          2     0     0     0     0     0     0     0
##          full          2     0     0     0     0     0     0     0
```

```r
#Region variable
regions <- data.frame(aggregate(all_data$reportedhrs_P, by=list(all_data$region), FUN=mean, na.rm=TRUE)
colnames(regions) <- c("region", "average_TV_hours")
```

```r
ggplot(data=regions, aes(x=region, y=average_TV_hours))+
  geom_bar(stat="identity")
```
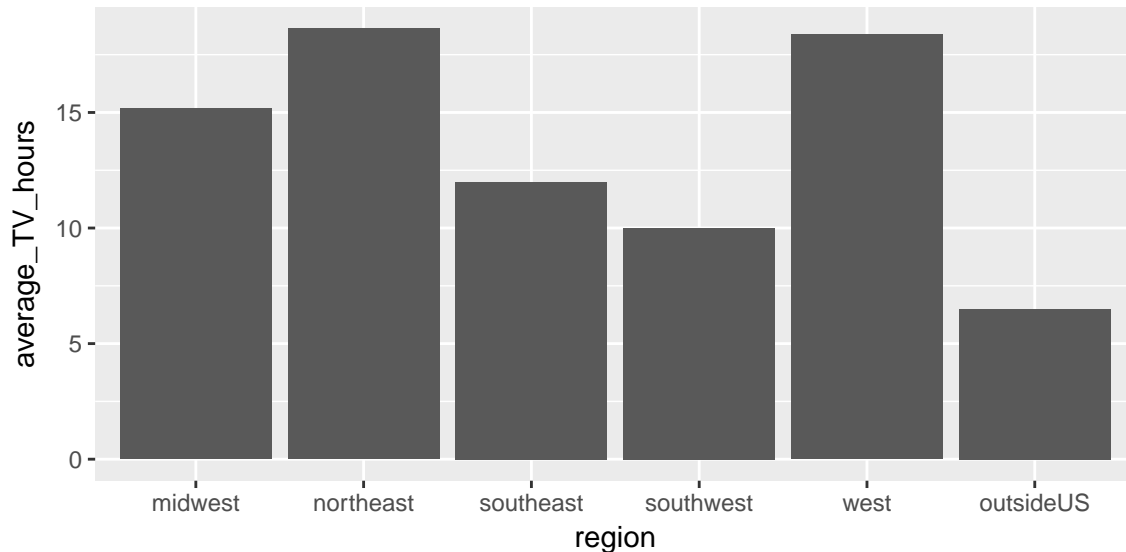


```r
table(all_data$region)
```

```
## 
##   midwest northeast southeast southwest      west outsideUS
##         4        26         4         2        42         3
```

Marital status and gender were two variables that seemed like it could make some sense to include in model. There are roughly an equal proportion of single vs. married individuals in our study (41 vs. 39) and married individuals in the study on averaged watched a little over 3 hours more of television during the 2 week study than their single counterparts did. It may make sense to use an indicator variable for 'married' as a result. There were also some modest signs to suggest that females had much more clustered values toward the low end of the TV watching spectrum while most of the outliers were male, making that a potentially interesting variable to include.
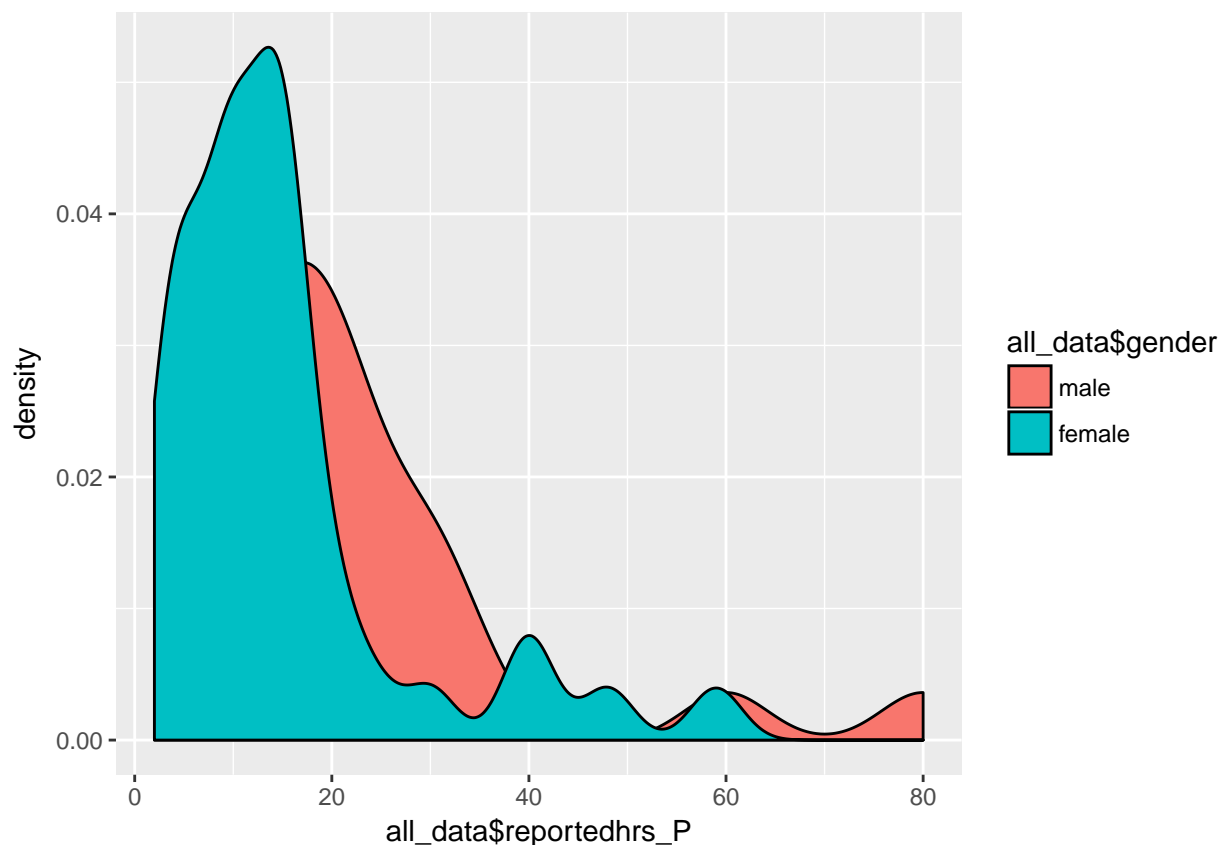
```r
#Marital Status variable
married <- data.frame(aggregate(all_data$reportedhrs_P, by=list(all_data$maritalStatus), FUN=mean, na.r
colnames(married) <- c("marital_status", "average_TV_hours")

ggplot(data=married, aes(x=marital_status, y=average_TV_hours))+
  geom_bar(stat="identity")
```
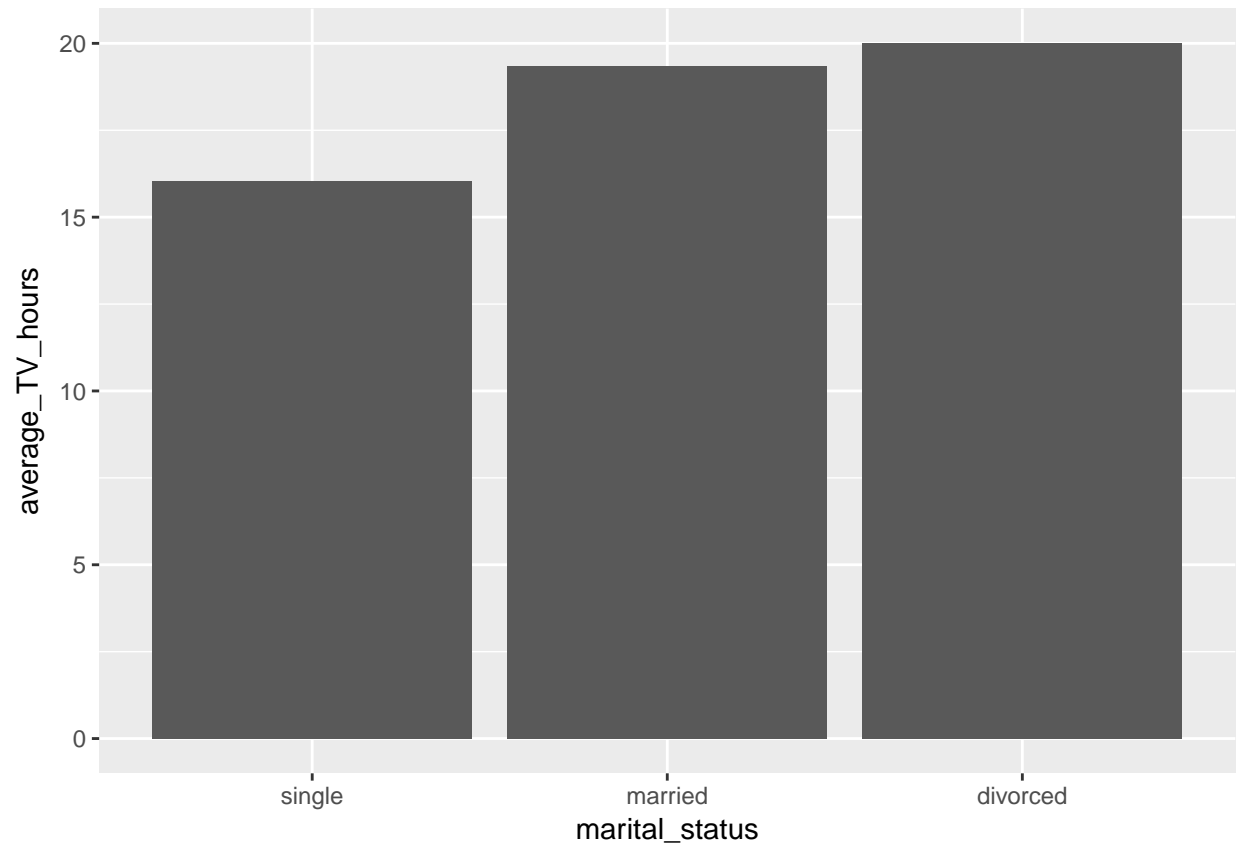
```
ggplot(all_data) +
  geom_density(aes(x=all_data$reportedhrs_P, fill=all_data$gender))
```

Below is a model using all of the variables that seemed to be noteworthy from our exploratory data analysis. Interestingly the only statistically significant variable (at a p-value of 5% or less) was the pre-survey TV hours. The treatment effect was -0.27 suggesting a small decrease in TV watching for the treatment groupu but it was not statistically significant.

```
all_data$married = ifelse(all_data$maritalStatus == "married", 1, 0)
mod_EDA <- lm(reportedhrs_P ~ Assigned + hoursTV +  binge + moreTimeThanWanted + watchAlone+ married + g
stargazer(mod_EDA, type = "text", title="Model X - Based on EDA")
```

```
##
## Model X - Based on EDA
## ================================================================
##                                         Dependent variable:
##                                     ----------------------------
##                                             reportedhrs_P
## ----------------------------------------------------------------
## Assigned                                       -0.270
##                                                (3.206)
##
## hoursTV                                       1.234***
##                                                (0.345)
##
## bingeonce a year                               -2.578
##                                                (7.285)
##
## bingeonce every couple months                  -6.438
##                                                (7.002)
```

```
##
## bingeonce a month                                   -0.753
##                                                     (7.825)
##
## bingeonce a week                                    -4.343
##                                                     (7.233)
##
## moreTimeThanWantedcouple times a year               -8.515*
##                                                     (5.079)
##
## moreTimeThanWantedonce a month                      -7.798
##                                                     (5.465)
##
## moreTimeThanWantedcouple times a month              -5.788
##                                                     (5.903)
##
## moreTimeThanWantedonce a week                       -5.098
##                                                     (6.369)
##
## watchAlonealone                                      2.604
##                                                     (3.791)
##
## married                                              4.093
##                                                     (3.642)
##
## genderfemale                                        -3.098
##                                                     (3.402)
##
## Constant                                            15.265*
##                                                     (8.599)
##
## -------------------------------------------------------------------
## Observations                                           65
## R2                                                    0.464
## Adjusted R2                                           0.327
## Residual Std. Error                            12.058 (df = 51)
## F Statistic                                3.395*** (df = 13; 51)
## ===================================================================
## Note:                                   *p<0.1; **p<0.05; ***p<0.01
```

Conclusion

Our study on TV watching behavior was ultimately inconclusive as there were no signs that being assigned to the treatment group and being shown how much TV peers are watching had any bearing on actual watching behavior. Our strongest model based on XXX showed XXX. The most statistically significant predictor we found was previous TV watching behavior. That said we would question the generalizaility of this study. Given the limited resources we had we sampled from a pretty narrow constituent base. Most of our participants were aged 22-34 and living in either the Northeast or West Coast. Further, the fact that 60% of our participants were "looking" for employment suggested to us that students were potentially a large part the base here and that's likely not that indicative of the broader population.

Appendix

The pre-survey in the pilot consisted of the following questions:

1. Provide us with an email address that will be used to contact you throughout the course of this survey.
2. Are you male or female?
3. How old are you?
4. What region do you currently live in?
5. What is your current employment status?
6. What is your marital status?
7. Do you have children under the age of 18?
8. How much television do you watch on a weekly basis? This includes cable, online streaming, Youtube, and can consist of movies, tv shows, and video clips.
9. Do you tend to binge watch TV shows (more than 4 hours of consecutive TV watching)?
10. What is your primary method of watching television?
11. Check off all the ways in which you consume television - Netflix, HBO, Hulu, Amazon, Youtube, Cable, Other.
12. How often do you find that after watching television you spent more time on it than you would have liked to?
13. Do you tend to watch television alone or with others?
14. Do you share your Netflix profile with someone else?
15. A link to installing the Netflix Chrome instruction as well as a text box for inputting what the extension outputs as usage over the last two week.

The post-survey in the pilot consisted of the following questions:

1. Provide us with your email address. This is the same identifier you used on the first survey and how we have been contacting you.
2. How much television did you watch during the study period from Feb 28 through March 1, inclusive? This includes cable, online streaming, and YouTube, and can consist of movies, tv shows, and video clips.
3. Input the time reading from your Chrome Netflix extension.

Summary of pilot pre-survey takeaways:

1. 8 of 9 individuals filled out the pre-survey in the designated time. Given the quick turnaround time (2 days), we expected some level of non-response as work, travel, etc can get in the way.

2. of the 8 individuals who did respond, 5 were Female and 3 male; 7 of the 8 were in the 22-34 age group, 7 of the 8 were in the Northeast region, 7 of 8 were full-time employed; 6 of 8 were married and 2 were single; and none of the 8 individuals had children under the age of 18.

3. At the mid-point of the ranges, individuals in the pilot reported watching 7.75 hours of TV a week. Netflix was the primary method of watching TV for half of the individuals while cable and other were the primary method for the other half while 7 of the 8 individuals said they consumed television through Netflix.

4. 3 of 8 individuals were unable to install the extension either because they don't have a Netflix account or they don't use Chrome as their web browser. Of the individuals that did input a reading from the extension, the average Netflix consumption over the prior two weeks was 787.2 minutes or just over 13 hours of TV.

Over the next couple of days individuals randomly assigned to the treatment group were sent an email with a bar chart showing peer consumption trends while those in the control group did not receive any email. Then all 8 individuals were asked to fill out a *post-survey* on how much TV they watched during the pilot period.

Standard Pitch
(in person)
Would you be willing to participate in a study about TV-watching habits? It will only involve taking one 5-minute survey, and then a one-minute survey at the end of the study period (2 weeks later). You'll be entered in a drawing for one of five $20 Amazon gift cards if you finish the study.
Check you junk mail folder in the next few days.
(screen roommates out)

(in writing - for social media, emails)

A few classmates and I are conducting a study about TV-watching habits - if you can participate, we would really appreciate it! It will require about 7 minutes of your time total (one survey at the beginning, and a very quick one at the end of the study period 2 weeks later), and upon completion you'll be entered in a drawing for one of five $20 Amazon gift cards. Please only one participant per household, in the US.

If you want to participate, enter your email address https://goo.gl/forms/sdyF6WA7fFMoOy9M2 by March 11th and we will send you the link to the first survey. Thank you so much in advance!

Figure 3: Standardized pitch

The experiment timeline
>Mar 8-11 - gather participants
>Mar 12 - send out pre-survey early in the morning
>Mar 15 - close pre-survey, randomize intervention, and covariance imbalance check
>Mar 18 - treatment/control email [completed 11pm EST]
>Apr 2 - start final data collection
>Apr 6 - end final data collection
>Apr 7 - start data analysis

CONTROL EMAIL BODY:

Thank you for completing the pre-survey to enter into our study. We really appreciate you taking the time out of your day to help us out.
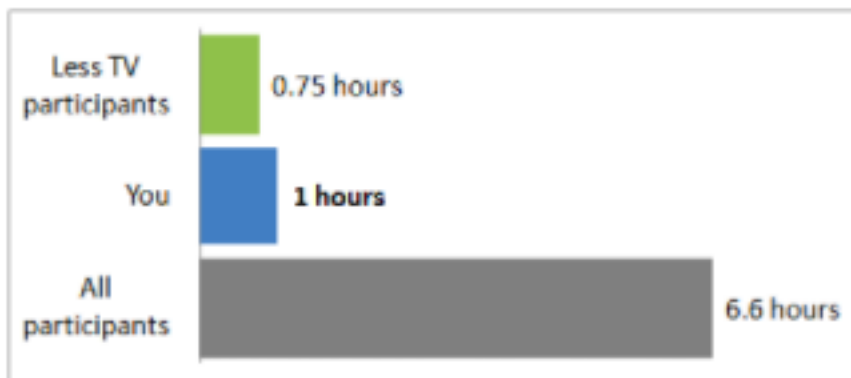
As a reminder, we will be sending a final brief survey on April 2nd asking you to report your amount of television watched during the study period. If you are using the extension, it tracks Netflix usage regardless of the medium - whether you watch it on a smart TV, your phone, or a computer, it will show up. Fill out that final survey by April 6th and we will enter you into the drawing for five $20 Amazon gift cards!

Thanks again,
Alyssa, Cameron, and Sarah

Figure 4: Sample Control email

TREATMENT EMAIL BODY (V1):

Thank you for completing the pre-survey to enter into our study. We really appreciate you taking the time out of your day to help us out, and wanted to share some interesting tidbits about TV watching habits from our data collection so far.

As a reminder, we will be sending a final brief survey on April 2nd asking you to report your amount of television watched during the study period. If you are using the extension, it tracks Netflix usage regardless of the medium - whether you watch it on a smart TV, your phone, or a computer, it will show up. Fill out that final survey by April 6th and we will enter you into the drawing for five $20 Amazon gift cards!

Thanks again,
Alyssa, Cameron, and Sarah

Figure 5: Sample Treatment email

Covariate balance check model results

```
stargazer(cov_balance_check1)
```

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Wed, Apr 25, 2018 - 10:40:22 PM

Table 3:

	Dependent variable:
	treat
genderfemale	−0.016
	(0.170)
age22-34	−0.539*
	(0.308)
age35-44	−0.601
	(0.425)
age45-54	−1.306**
	(0.528)
regionnortheast	0.268
	(0.342)
regionsoutheast	−0.519
	(0.471)
regionsouthwest	−0.295
	(0.875)
regionwest	0.034
	(0.350)
regionoutsideUS	−0.152
	(0.528)
employmentpart	0.304
	(0.343)
employmentlooking	−0.067
	(0.682)
employmentstudent	0.073
	(0.295)
employmenthomemaker	0.318
	(0.625)
employmentself	0.271
	(0.447)
maritalStatusmarried	0.159
	(0.181)
maritalStatusdivorced	0.151
	(0.759)
childrenno	0.079
	(0.237)
hoursTV	0.004
	(0.016)
bingeonce a month	−0.341