HW 4: Financial Stressors and Cognitive Performance

Do changes in one's financial circumstances affect one's decision-making process and cognitive capacity? In a research study, researchers randomly selected a group of US respondents to be surveyed before their payday and another group to be surveyed after their payday. People working for different employers or in different jobs get paid at different times. So at any given time, some people will have just been paid and others will be waiting to be paid. Under this design, the researchers believed that the respondents of the `Before Payday` group would be more likely to be financially strained (which we will consider to be the treatment) than those of the `After Payday` group (which we will consider to be the alternative). The researchers were interested in investigating whether or not being financially strained affects people's decision making and cognitive performance. Other researchers have found that in a lab setting, scarcity induces an additional mental load that impedes cognitive capacity.

In this study, the researchers administered a number of decision-making and cognitive performance tasks to the `Before Payday` and `After Payday` groups. We focus on the *Numerical Stroop Task*, which measures cognitive control. An example of a Numerical Stroop Task would be to present two numbers side by side, and vary the font sizes so that sometimes the lower number is in a bigger font than the higher number and vice versa and test how long it takes the subject to correctly identify which number is larger and how many errors they make. In general, taking more time to complete this task indicates less cognitive control and reduced cognitive ability. They also measured the amount of cash the respondents have on hand, the amount of money in their checking and saving accounts, and whether or not they have a low annual income. The data set is in the CSV file `poverty.csv`. The names and descriptions of variables are given below:

Name	Description
`treatment`	Treatment conditions: `Before Payday` and `After Payday`
`cash`	Amount of cash respondent has on hand
`accts_amt`	Amount in checking and saving accounts
`stroop_time`	Log-transformed average response time for cognitive stroop test
`income_less20k`	Binary variable: `1` if respondent earns less than 20k a year and `0` otherwise

Question 1 (6 points)

1a (2 points)

What is the specific causal question the researchers would like to answer?

1b (2 points)

For the subjects who had not yet received their paycheck, what is the average missing counterfactual at the group level?

1c (2 points)

What strategy does this study propose using to estimate this missing counterfactual? What assumption is required for this estimate of the missing counterfactual to be unbiased?

Answer 1

Your Text answers for 1a - 1c go below:

Post only cross-sectional Study ### 1a What is the impact of financial circumstances of subjects in “Before Payday” group (intervention - Treatment Group) relative to the subjects in “After Payday” group (alternative - Control Group) on their respective decision making process and cognitive capacities for the randomly selected group of the US respondents.

1b

MCF: At the group level, for the subjects who had not yet received their paychecks, what their average decision making process and cognitive capacity would have been if they instead had the control group experience (i.e. had received their pays - After Payday) but all else remained the same.

1c

Clarification or Assumption: Although the first group of respondents has been randomly selected, yet those selected respondents are not randomly assigned to any of the two control and treatment groups, therefore, this study becomes an observational study after first stage. Further, as we are required to estimate the differences between intervention and alternative groups post the intervention, therefore, this study becomes the post only cross sectional and proposes to estimate this missing counter factual as per this study settings.

Estimated MCF at group level: the average factual outcome of the decision making process and cognitive capacity (stroop_time score) of the subjects in the “After payday” alternative group would be the average MCF for the “Before Payday” treatment group. Assumption Required to be unbiased Estimate of MCF: There are no systematic differences between the ‘Before Payday’ treatment group and the ‘After Payday’ control groups that could have impacted the outcome.

Question 2 (10 points)

2

Load the `poverty.csv` data set. Use histograms and boxplots to examine the univariate distributions of the two baseline covariates that are financial resources measures: `cash` and `accts_amt`. Calculate the mean and standard deviation, median, quartiles, and IQR for the two financial resources measures for all subjects in the dataset.

Describe these two variables' univariate distributions based on the figures and the summary statistics.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6     v purrr   0.3.4
## v tibble  3.1.8     v dplyr   1.0.9
## v tidyr   1.2.0     v stringr 1.4.1
## v readr   2.1.2     v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
# You will need ggplot2 to make plots.
require("ggplot2")
# Load any other packages

# Load the poverty data; remember to put the csv file in a data sub-folder!
poverty.full <- read.csv("data/poverty.csv")
poverty.dropped <- read.csv("data/poverty.csv") %>% na.omit()
summary(poverty.full)
```

```
##   treatment              cash           accts_amt      stroop_time
##  Length:2670        Min.   :   0.0   Min.   :    0   Min.   :5.356
##  Class :character   1st Qu.:  15.0   1st Qu.:  176   1st Qu.:7.436
##  Mode  :character   Median :  49.5   Median : 1000   Median :7.564
##                     Mean   : 169.0   Mean   : 6211   Mean   :7.545
##                     3rd Qu.: 136.2   3rd Qu.: 5000   3rd Qu.:7.692
##                     Max.   :9000.0   Max.   :95000   Max.   :8.517
##                     NA's   :226      NA's   :433
##  income_less20k
##  Min.   :0.0000
##  1st Qu.:0.0000
##  Median :0.0000
##  Mean   :0.4139
##  3rd Qu.:1.0000
##  Max.   :1.0000
##
```

```
summary(poverty.dropped)
```

```
##   treatment              cash           accts_amt      stroop_time
##  Length:2237        Min.   :   0.0   Min.   :    0   Min.   :5.356
```

```
##  Class :character   1st Qu.:  17.0   1st Qu.:  176   1st Qu.:7.436
##  Mode  :character   Median :  50.0   Median : 1000   Median :7.562
##                     Mean   : 174.3   Mean   : 6211   Mean   :7.543
##                     3rd Qu.: 148.0   3rd Qu.: 5000   3rd Qu.:7.686
##                     Max.   :9000.0   Max.   :95000   Max.   :8.517
##  income_less20k
##  Min.   :0.000
##  1st Qu.:0.000
##  Median :0.000
##  Mean   :0.392
##  3rd Qu.:1.000
##  Max.   :1.000
```

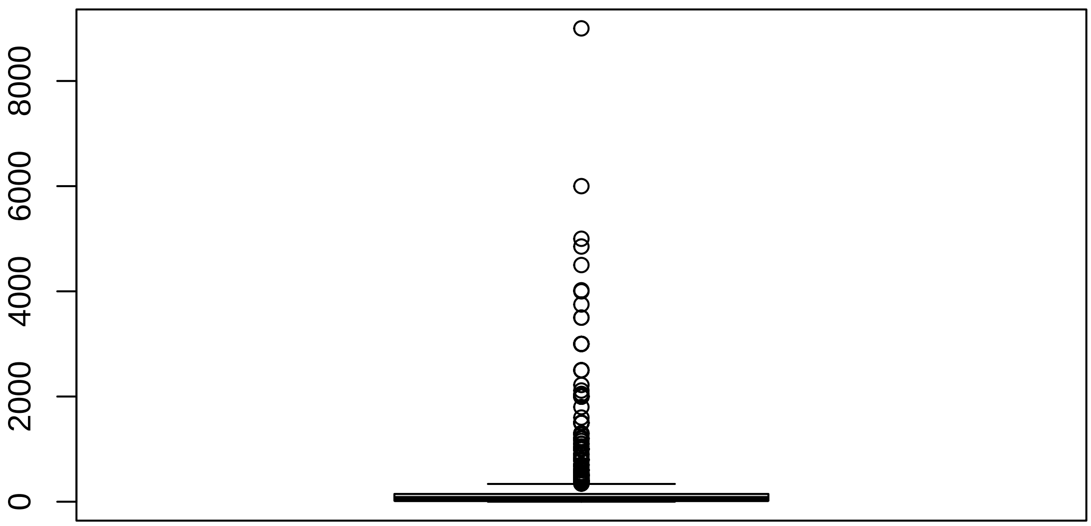
Answer 2

Your answers for Question 2 go here!

'Poverty.full' data set contains 2670 observations regarding unique subjects but there are null values for some of the fields which might create problems in calculation of summary statistics, therefore, to avoid such issues, the null values present in any of the fields have been omitted and the new data set is named 'Poverty.dropped' and it contains 2237 observations. In all further calculations, this new data set will be used in this assignment. Further, for sake of checking the difference created by this drop operation in two data sets, summary statistics for the two data frames are calculated and means of each field in two data sets are compared which are: poverty.full = (cash = 169, accts_amt = 6211, stroop_time = 7.545, and income_less20k = 0.4139) poverty.dropped = (cash = 174.3, accts_amt = 6211, stroop_time = 7.543, and income_less20k = 0.392) and it shows that the two data sets are not significantly different from each other.

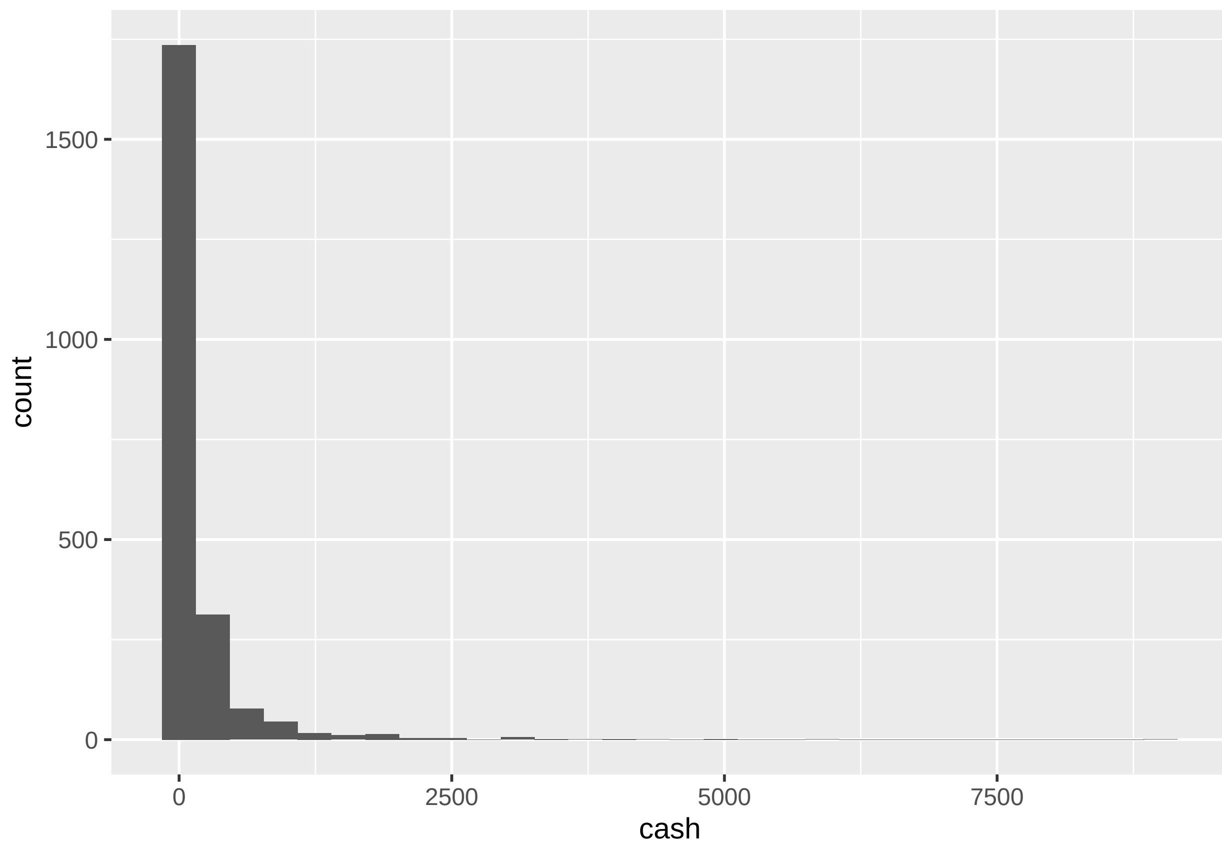
2 Cash Variable

```
# Your code for the cash variable goes here
# can remove x-axis values which are meaningless through this code
boxplot(poverty.dropped$cash)
```




```r
# boxplot
#poverty.dropped %>%
#  ggplot(aes(y = cash)) +
 # geom_boxplot()

poverty.dropped %>%
  ggplot(aes(x = cash)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
summary(poverty.dropped$cash)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     0.0    17.0    50.0   174.3   148.0  9000.0
```

```
mean(poverty.dropped$cash)
```

```
## [1] 174.3286
```

```
median(poverty.dropped$cash)
```

```
## [1] 50
```

```
sd(poverty.dropped$cash)
```

```
## [1] 464.0278
```

```
range(poverty.dropped$cash)
```

```
## [1]    0 9000
```

```r
quantile(poverty.dropped$cash, probs = c(0, 0.10, 0.5, 0.90, 1))
```

```
##   0%  10%  50%  90% 100% 
##    0    3   50  400 9000
```

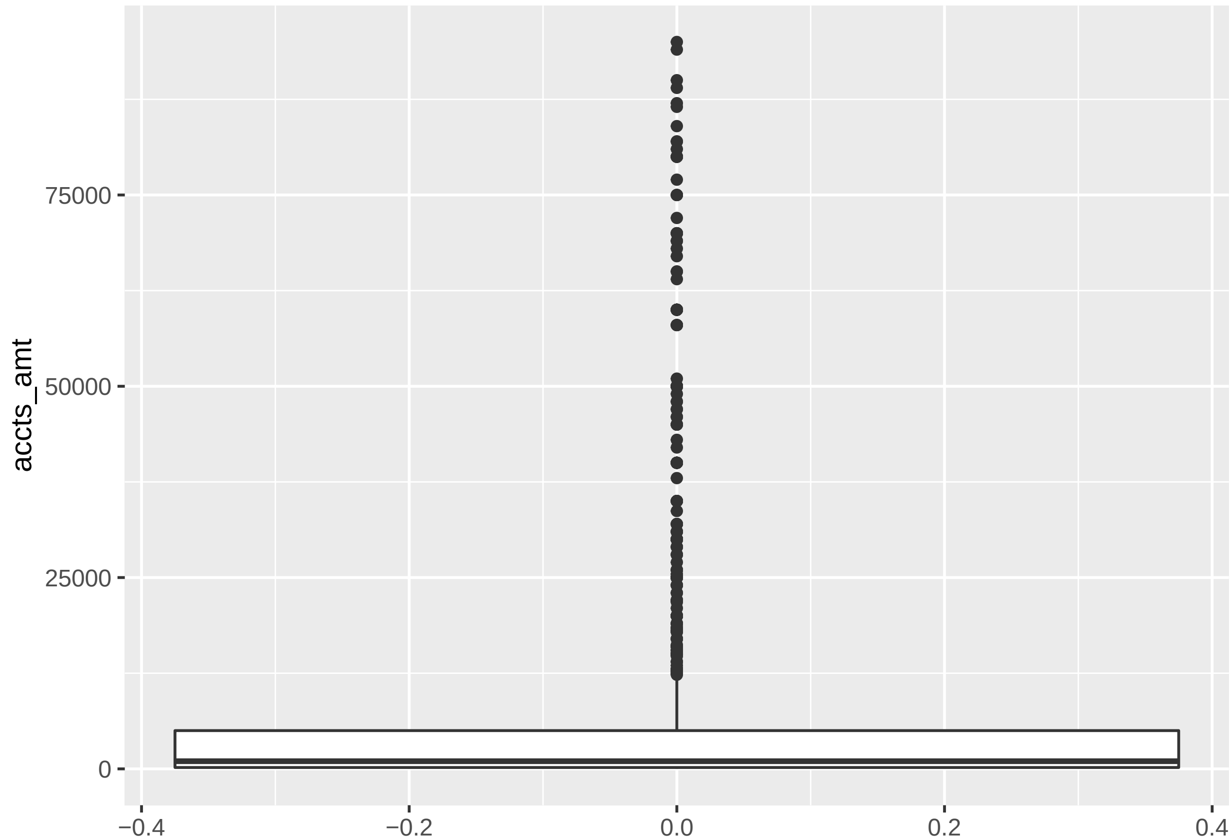
```r
IQR(poverty.dropped$cash)
```

```
## [1] 131
```

2 Cash Variable Text Answer:

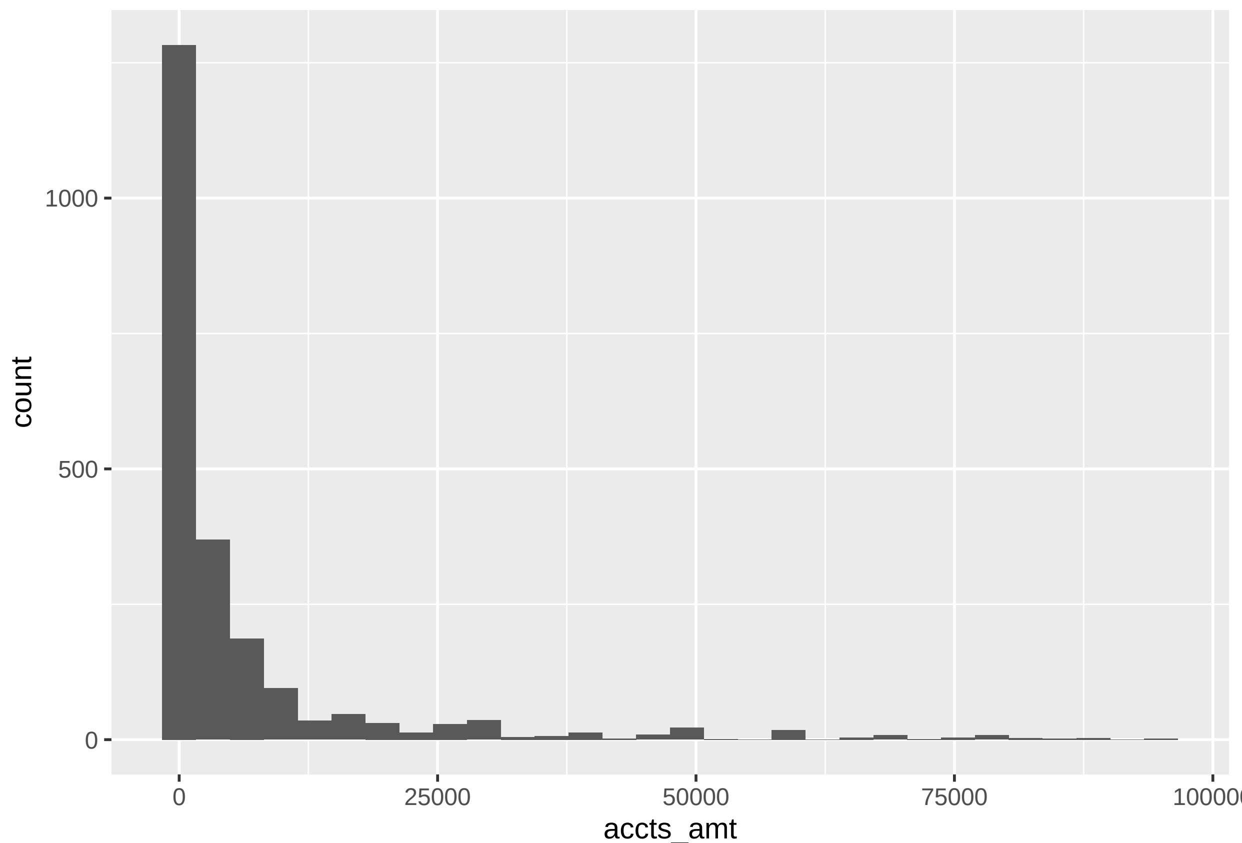
Boxplot for cash variable shows the first quartile, median and upper quartile and the lower whisker are all merged together around less than approximately 150 cash. There are many outliers above the upper whisker with gaps as we move away from the upper whisker. This cash variable distribution is highly right skewed. Histogram of the cash variable shows the distribution is strongly unimodal and highly right skewed with some gaps and outliers. There are small spikes on the right side. Summary statistics of the cash(in $) variable shows mean is around 174.3 and median is 50 while the 1st quartile is 17 and 3rd quartile is 148 with IQR 131, and the range is between 0 and 9000. Standard deviation is around 464.0278 which is understandable because the spread of variable cash is usually more within the group.

```r
# Your code for the accounts variable goes here
#boxplot(poverty.dropped$accts_amt)
# boxplot
poverty.dropped %>%
  ggplot(aes(y = accts_amt)) +
  geom_boxplot()
```




```
poverty.dropped %>%
  ggplot(aes(x = accts_amt)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
summary(poverty.dropped$accts_amt)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       0     176    1000    6211    5000   95000
```

```
mean(poverty.dropped$accts_amt)
```

```
## [1] 6210.834
```

```
median(poverty.dropped$accts_amt)
```

```
## [1] 1000
```

```
sd(poverty.dropped$accts_amt)
```

```
## [1] 13517.69
```

```
range(poverty.dropped$accts_amt)
```

```
## [1]     0 95000
```

```
quantile(poverty.dropped$accts_amt, probs = c(0, 0.10, 0.5, 0.90, 1))
```

```
##      0%     10%     50%     90%    100%
##     0.0    13.6  1000.0 18000.0 95000.0
```

```
IQR(poverty.dropped$accts_amt)
```

```
## [1] 4824
```

2 Accounts Variable Text Answer:

Boxplot for amount in checking and saving account - acctsd_amt variable is a little different than cash variable boxplot but overall both are same. The differnece here is in the size of the box and the first quartile, median, upper quartile and the lower whisker are not as merged as they were for the cash variables. lower whisker, 1st quartile, and median all are lying less than 500 while 3rd quartile is at around 5000. There are many outliers above the upper whisker with gaps as we move away. This accts_amt variable distribution is highly right skewed. Histogram of this variable shows the distribution is strongly unimodal and highly right skewed with some gaps and outliers. There are small spikes on the right side. Summary statistics of the acct_amt variable (in $) shows mean is around 6211 and median is 1000 while the 1st quartile is 176 and 3rd quartile is 5000 with IQR 4824, and the range is between 0 and 95000. Standard deviation is around 13517.69. Overall, both of the cash and amount in account for these subjects are almost similar and depict the same behavior of people regarding cash and keeping amount in the account.

Question 3 (9 points) ### 3a (7 points)

Now, use the code provided below to take the *natural logarithm* of the `cash` variable and summarize the transformed variable using summary statistics and figures: calculate the mean, median, standard deviation, and IQR of the log transformed `cash` variable. Create a boxplot and histogram for the transformed variable. Describe the log transformed variable's distribution based on the figures and the summary statistics. How does this distribution differ from the `cash` variable on the original scale?

3b (2 points)

State an advantage and a disadvantage of transforming the data in this way.

NOTE FOR INTERESTED STUDENTS: *Since the natural logarithm of 0 is undefined, researchers often add a small value (in this case, we will use $1 so that* $\log 1 = 0$*) to the 0 values for the variables being transformed (in this case, `cash` and `accts_amt`) in order to successfully apply the `log()` function to all values. Note that we do this recoding only for the purposes of taking the logarithmic transformation – we have kept the original variables the same.*

```
### Log transformation code for 3a
# Feel free to move this code where you need it

# Log-transform money in cash
# poverty$log_cash <- poverty$cash
# poverty$log_cash[poverty$log_cash == 0] <- 1
# poverty$log_cash <- log(poverty$log_cash)
```

3a

```
# Your code for 3a goes here
poverty.dropped$log_cash <- poverty.dropped$cash
poverty.dropped$log_cash[poverty.dropped$log_cash == 0] <- 1
poverty.dropped$log_cash <- log(poverty.dropped$log_cash)

summary(poverty.dropped$log_cash)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.833   3.912   3.742   4.997   9.105
```

```
mean(poverty.dropped$log_cash)
```

```
## [1] 3.742224
```

```
median(poverty.dropped$log_cash)
```

```
## [1] 3.912023
```

```
sd(poverty.dropped$log_cash)
```

```
## [1] 1.814489
```

```r
range(poverty.dropped$log_cash)
```

```
## [1] 0.00000 9.10498
```

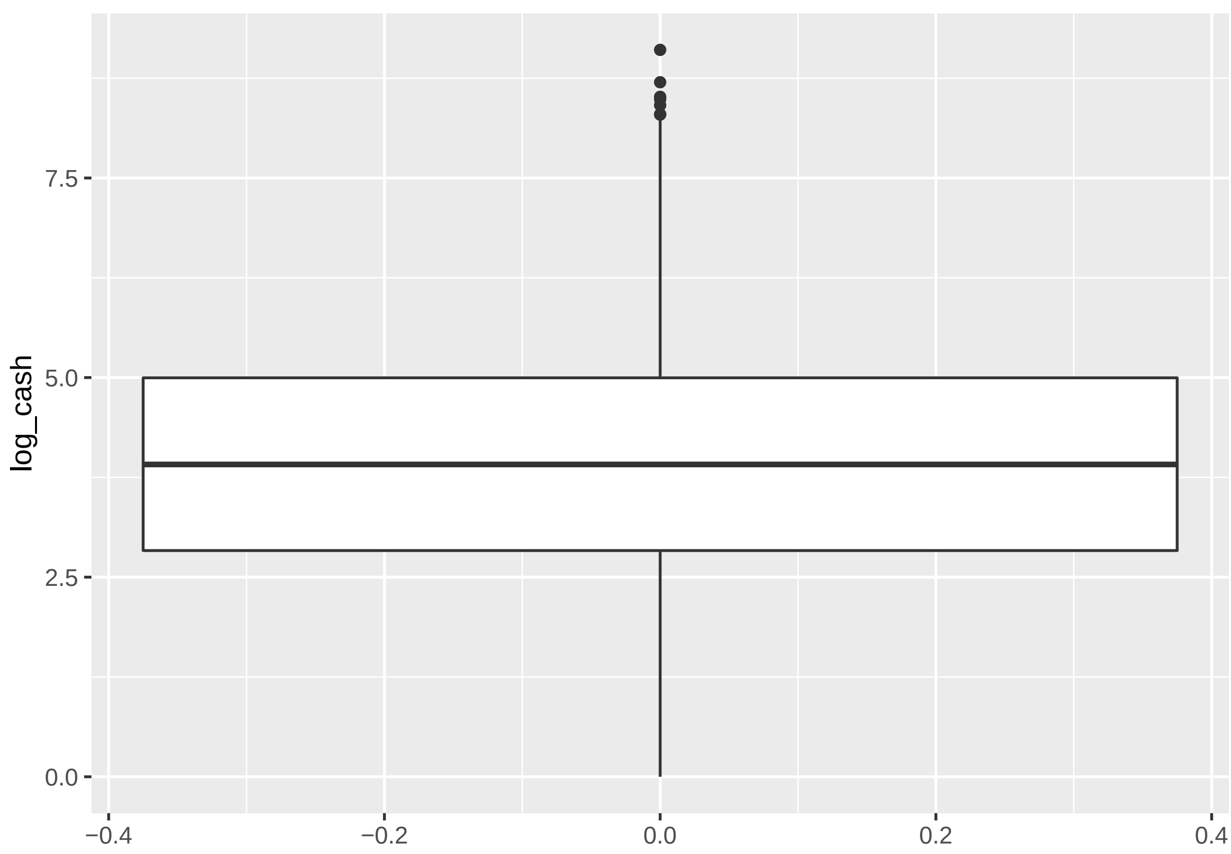
```r
quantile(poverty.dropped$log_cash, probs = c(0, 0.10, 0.5, 0.90, 1))
```

```
##       0%      10%      50%      90%     100%
## 0.000000 1.098612 3.912023 5.991465 9.104980
```

```r
IQR(poverty.dropped$log_cash)
```

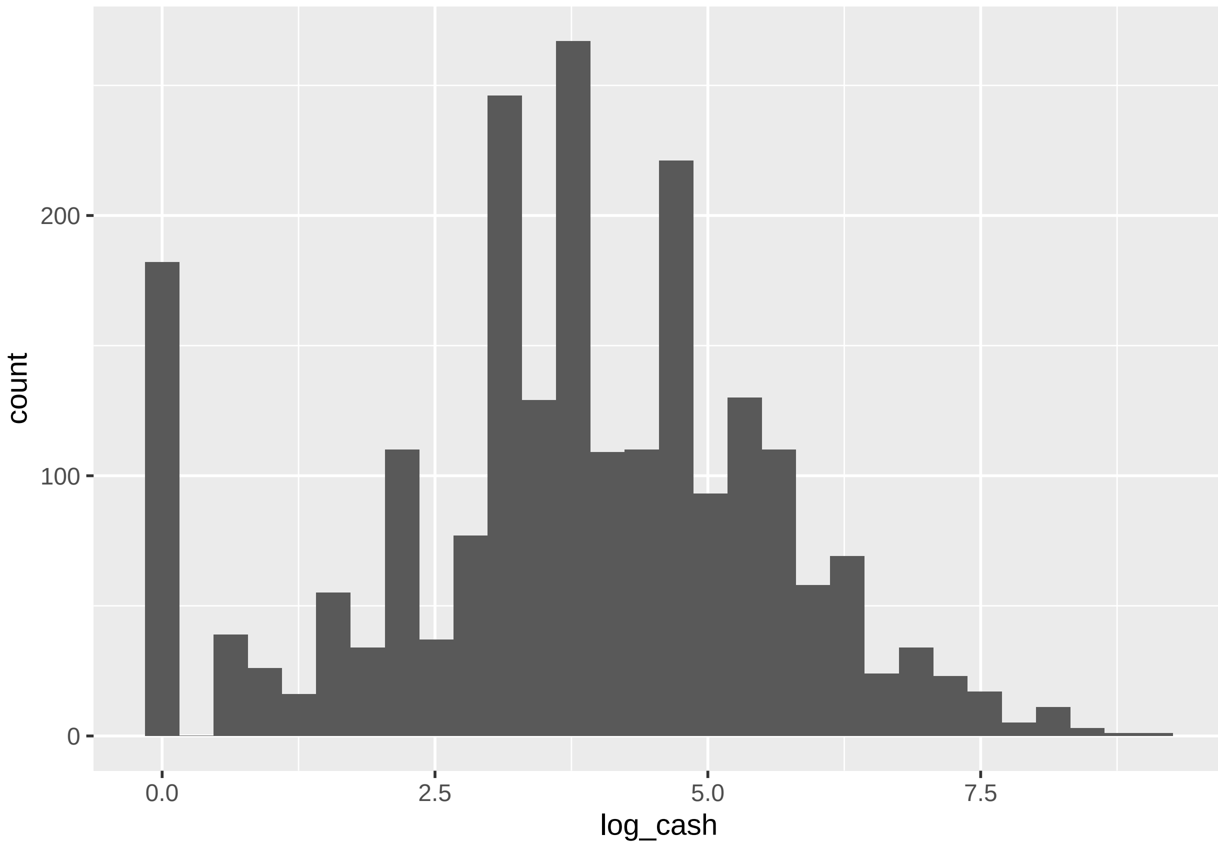
```
## [1] 2.163999
```

```r
#boxplot(poverty.dropped$log_cash)
# boxplot
poverty.dropped %>%
  ggplot(aes(y = log_cash)) +
  geom_boxplot()
```




```r
poverty.dropped %>%
  ggplot(aes(x = log_cash)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Text Answer for 3a:

Boxplot for log transformed cash variable now shows the distribution is quite uniform with a slight right skew. There is clear gap in the 1st quartile, meidan and the third quartile which was not as evident in the original cash distribution. Also, there are outliers above the upper whisker with some small gaps. However, on log transformed scale, interpretation of the quartiles and the median is not making much sense. Histogram of the log transformed cash variable shows the distribution is unimodal and slightly right skewed with a gap on the left and a strong spike along it. Summary statistics of the log transformed cash variable (in $) shows mean is around 3.742 and median is 3.912 while the 1st quartile is 2.833 and 3rd quartile is 4.997 with IQR 2.16, and the range is between 0.1 and 9.1049. Standard deviation is around 1.814.

This log transformed distribution differs from the original cash distribution in terms of symmetry and skewness as in opposition to the original cash distribution, it is relatively more symmetric and less skewed and helped in getting a better idea about identifying the differences lied in quartiles and meidan in boxplot and also varaitions in histogram which were not as clear in the original cash variable boxplot and histogram as these are in log transformed distribution. However, interpretation of the summary statistics is not making much sense here as compared to the original scale.

3b Text Answer:

The advantage of such transformation of data is that the summary statistics like mean, standard deviation and linear correlation which are sensitive to outliers are less likely to be misinterpreted as we have also witnessed above. The disadvantage, however, is the issue of interpretation of results as this log transformed scale is less interpretable scale and we have to back-transform the results for sake of better interpretation.

Question 4 (16 points)

4a (7 points)

Create a z-score version of the `cash` variable. Use a summary command and a boxplot and then describe the distribution of the z-scores for this variable. What proportion of the observations have a z-score greater than 2? What proportion of the observations have a z-score greater than 3?

4b (9 points)

Create a z-score version of the log transformed `cash` variable. Use a summary command and a boxplot and then describe the distribution of the z-scores for this log transformed variable. What proportion of the observations have a z-score greater than 2 for this log transformed variable? What proportion of the observations have a z-score greater than 3 for this log transformed variable? How does this distribution differ from the distribution you described in question 3a?

Answer 4

4a

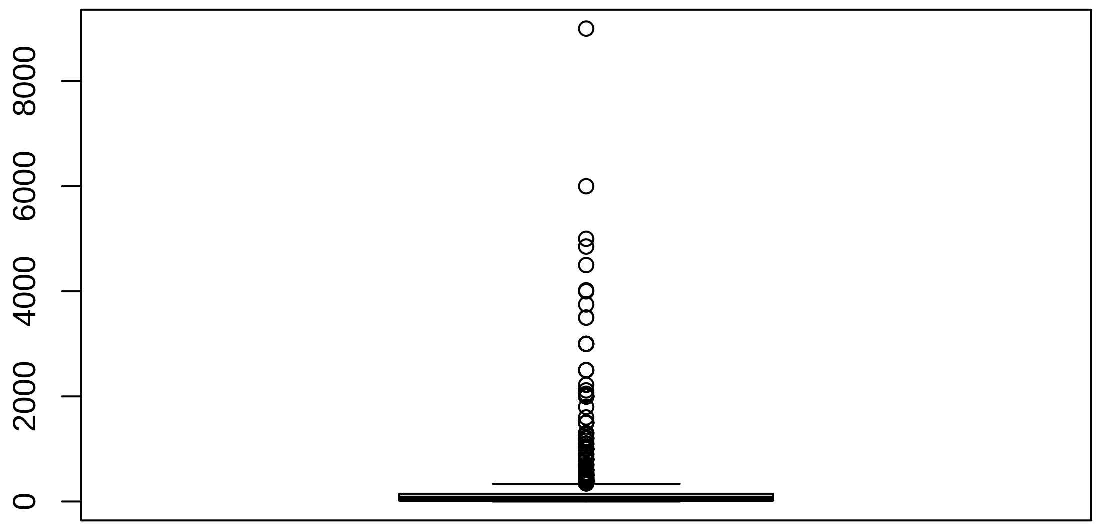
```
# Your code for 4a goes here
poverty.dropped$z.cash = (poverty.dropped$cash) - mean(poverty.dropped$cash)/sd(poverty.dropped$cash)

summary(poverty.dropped$z.cash)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##   -0.376   16.624   49.624  173.953  147.624 8999.624
```

```
sd(poverty.dropped$z.cash)
```

```
## [1] 464.0278
```

```
boxplot(poverty.dropped$z.cash)
```




```r
mean(poverty.dropped$z.cash > 2)
```

```
## [1] 0.901207
```

```r
mean(poverty.dropped$z.cash > 3)
```

```
## [1] 0.8895843
```

Text Answer for 4a:

For sake of comparison of different variables, standard z score is calculated. Here, the z-score version of cash variables shows a mean of 173.95 and standard deviation of 464.03. The fist quartile is 16.6, 3rd is 147.6 with IQR around 131 and range is between -0.38 and 8999.6. The boxplot of this version is almost similar in overall shape to the original cash variable boxplot with 75% of data falling under 200; there are outliers. Z score has rescaled the original distribution but has not impacted the shape and it is evident from the boxplot. 90% of the observations have a z-score greater than 2 while 89% observations have a z-score greater than 3.

4b

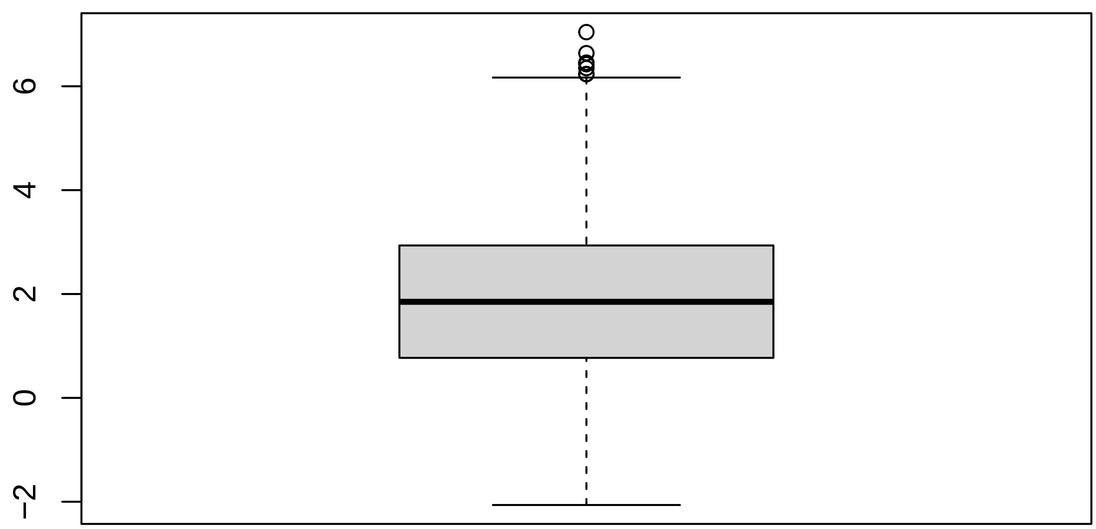
```r
# Your code for 4b goes here
poverty.dropped$z.log.cash = (poverty.dropped$log_cash) - mean(poverty.dropped$log_cash)/sd(poverty.drop

summary(poverty.dropped$z.log.cash)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -2.0624  0.7708  1.8496  1.6798  2.9348  7.0426
```

```r
sd(poverty.dropped$z.log.cash)
```

```
## [1] 1.814489
```

```r
boxplot(poverty.dropped$z.log.cash)
```




```r
mean(poverty.dropped$z.log.cash > 2)
```

```
## [1] 0.4456862
```

```r
mean(poverty.dropped$z.log.cash > 3)
```

```
## [1] 0.2239607
```

Text Answer for 4b:

the z-score version of log transformed cash variables shows a mean of 1.68 and standard deviation of 1.81. The fist quartile is .77, 3rd is 2.93 with IQR around 2.16 and range is between -2.06 and 7.04. The boxplot of this log transformed z version is almost similar in overall shape to the log transformed cash variable boxplot and its interpretaion is making little sense here; there are outliers. Z score has rescaled the log transformed distribution but has not impacted the shape and it is evident from the boxplot. For this log transformed variale, 45% of the observations have a z-score greater than 2 while 22% observations have a z-score greater than 3.

This z version of the log transformed cash variable is a bit different in terms of scale of the summary statistics as the z-score has tried switching mean towards 0 but it has not perfectly done it. The shape of the distribution is almost same and there is not much difference.

Question 5 (8 points)

5a (4 points)

Now, let's focus on the primary outcome of interest for this study - cognitive performance. Let's estimate the effect of a presumed change in financial situation (in this case, waiting to get paid relative to just having been paid) on cognitive performance. Estimate the treatment's effect on the `stroop_time` variable (a log-transformed variable of the average response time for the Stroop Cognitive Test), using first the mean and then the median. What do these results tell you about the answer to the causal question from Q1?

5b (4 points)

Now compare the time it took subjects to complete the cognitive test across the `Before Payday` and `After Payday` groups using overlaid or side-by-side histograms and side-by-side boxplots of the `stroop_time` variable for the before and after payday groups. Based on these plots, do you think there is a meaningful difference in the time it took participants from each group to complete the cognitive test?

Answer 5

5a

```r
# Your code for 5a goes here
means <- tapply(poverty.dropped$stroop_time, poverty.dropped$treatment, mean)
# look at those means
means
```

```
##  After Payday Before Payday
##      7.545226      7.540923
```

```r
# take the difference between elements of that vector of means
means[2] - means[1]
```

```
## Before Payday
##  -0.004302999
```

```r
medians <- tapply(poverty.dropped$stroop_time, poverty.dropped$treatment,median)
medians
```

```
##  After Payday Before Payday
##      7.570255      7.554988
```

```r
medians[2] - medians[1]
```

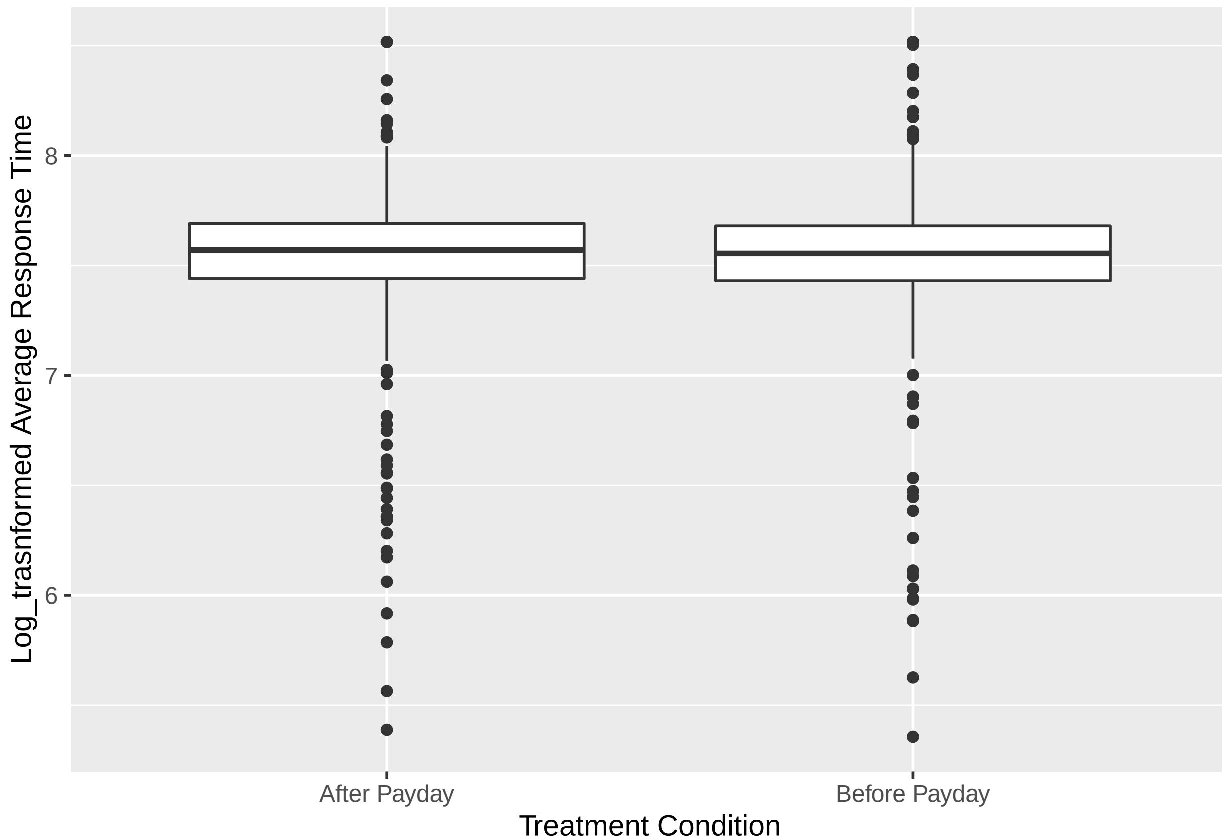
```
## Before Payday
##   -0.01526748
```

Text Answer for 5a:

The mean estimate of the treatment effect of 'Before Pay' day is around -0.004 while the median treatment effect is around -0.015. Both of these effects are quite minimal and it tells us that on average, there is very very slight impact (-0.4%) of financial situation on the cognitive performance of the people in the treatment group of 'Before Payday' in the group of US respondents.

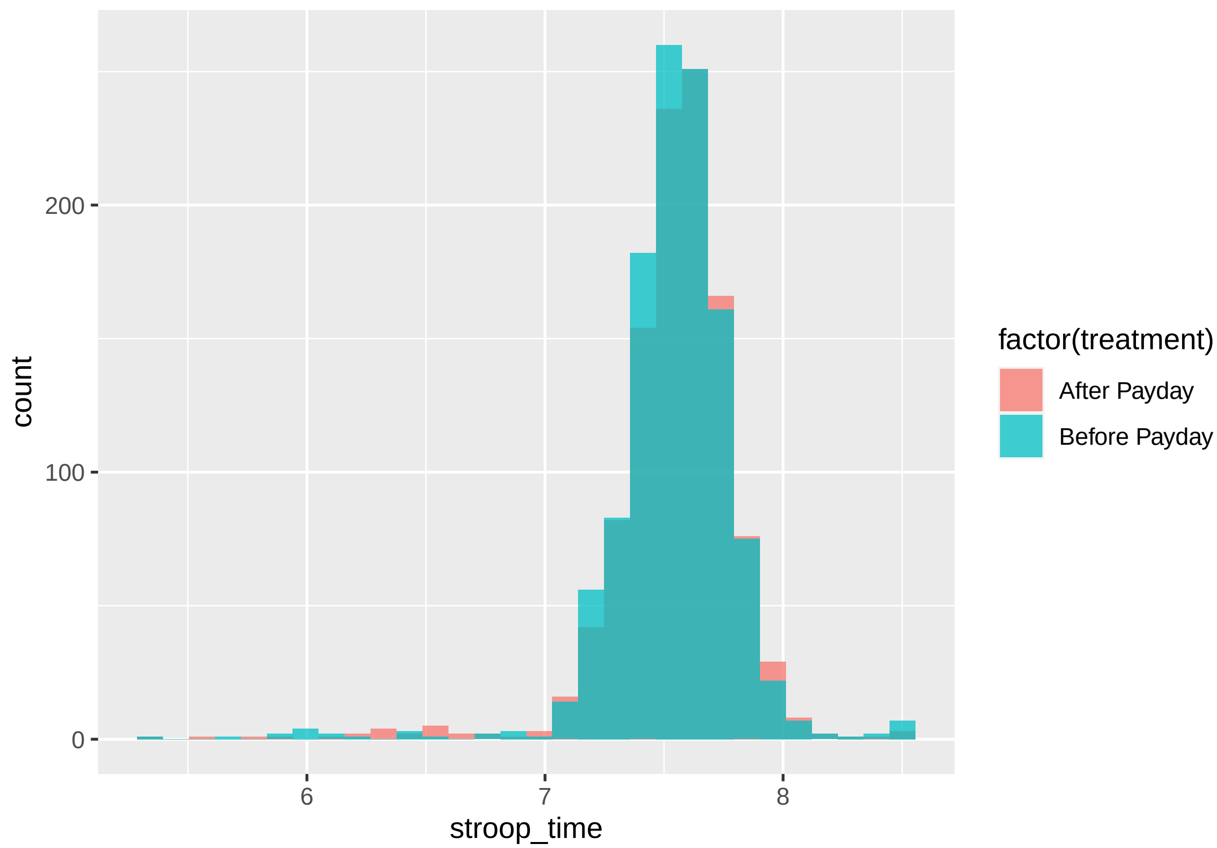
5b

```r
# Your code for 5b goes here
poverty.dropped %>%
  ggplot(aes(x = factor(treatment), y = stroop_time)) +
  geom_boxplot() +
  labs(x = 'Treatment Condition', y = 'Log_trasnformed Average Response Time')
```




```r
poverty.dropped %>%
  ggplot(aes(x = stroop_time, fill = factor(treatment))) +
  geom_histogram(position = 'identity', alpha = 0.75)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Text Answer for 5b: Comparing the two groups stroop_time variable, boxplot shows the median of the Before Payday group is slightly lower than that of After Payday and similarly, 1st quartile and 3rd quartile are also slightly lower than that of After Payday group while the outliers on both below lower whisker and above upper whiskers are also almost similar in their respective boxplots. Similarly, the overlaid histogram of Before Payday group has almost covered the After Payday group stroop_time variable distribution. There are minor differences at some points and both are left skewed with similar gaps and small spikes on right and left sides.
Analysis on the basis of side by side boxplot and the histogram of the Before Payday treatment group and the After Payday control group cognitive performance test time on average show no meaningful difference between the two groups.

Question 6 (9 points)

6a (6 points)

Now, we will look at the relationship between general financial circumstances and cognitive performance. Produce two scatter plots, one for each of the two treatment conditions, showing the bivariate relationship between your *log-transformed* `cash` variable and `stroop_time` (place the `stroop_time` variable on the *vertical axis*). Be sure to title your graphs to differentiate between the `Before Payday` and `After Payday` conditions. Calculate the linear correlation between your *log-transformed* `cash` variable and `stroop_time` for each of the two treatment conditions. Do the associations between the log-transformed `cash` variable and cognitive performance appear to be linear? If yes, what is the direction and strength of this linear association? If no, what kind of association do you see between these two variables?

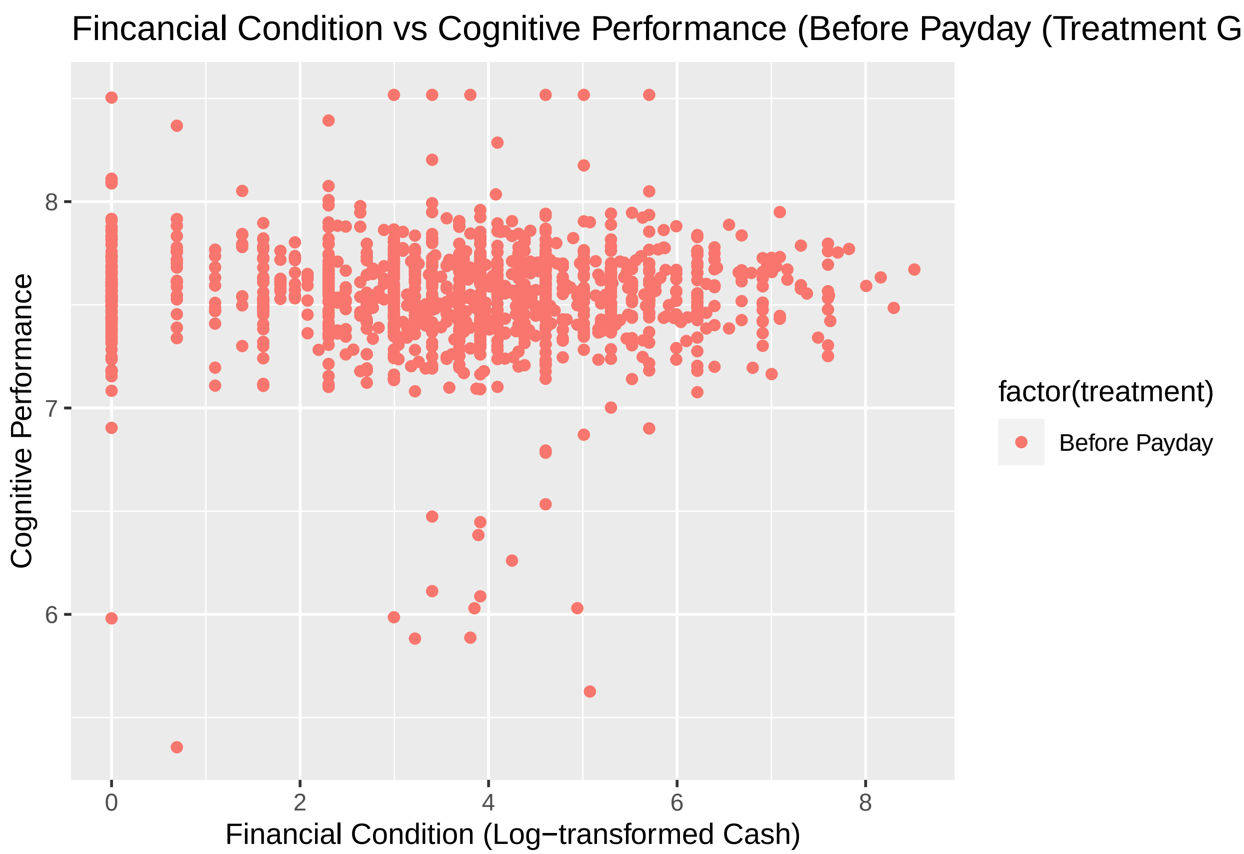
6b (3 points)

Using the scatterplots and the linear correlations as evidence, briefly comment on whether this evidence supports or contradicts the hypothesis that economic circumstances will influence cognitive performance.

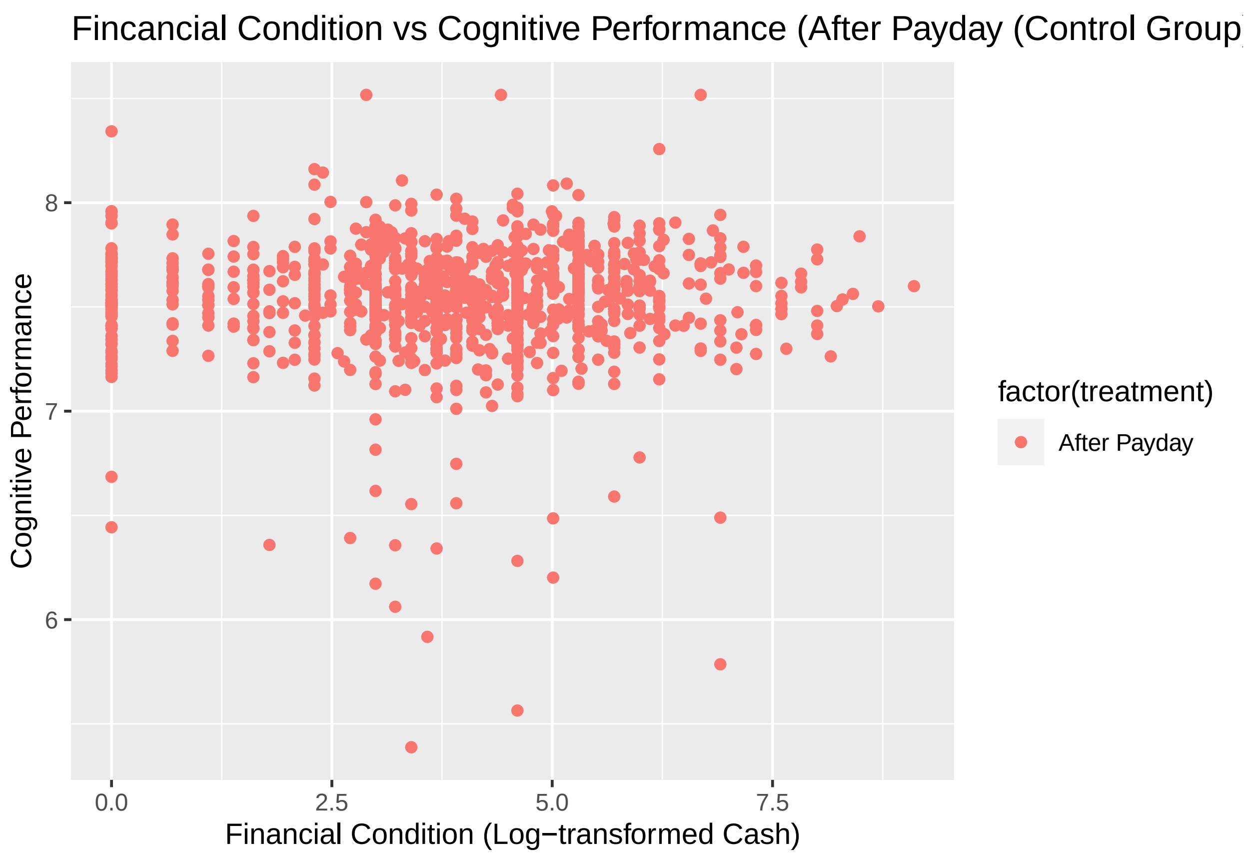
Answer 6

6a

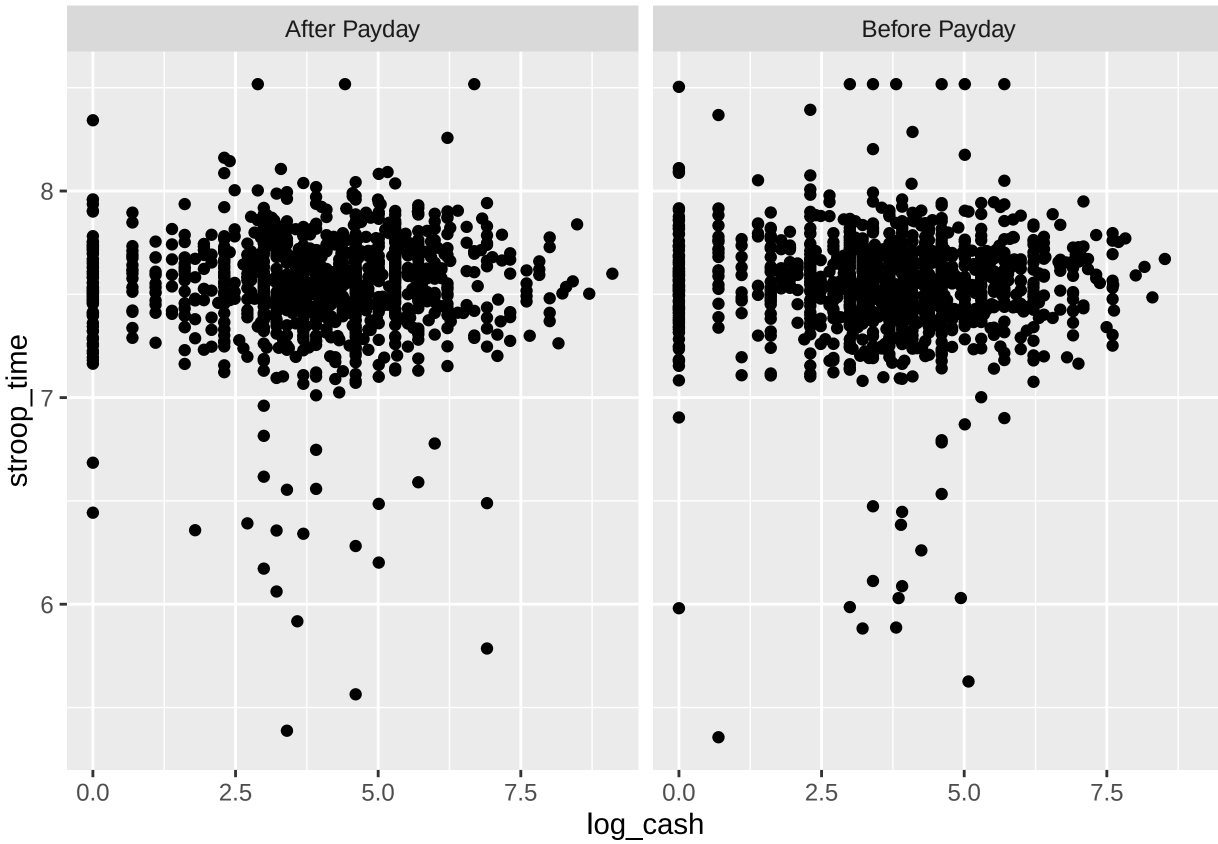
```
# Your code for 6a goes here
#Before Payday
povertyB.dropped <- poverty.dropped %>%
  filter(treatment == "Before Payday")
povertyB.dropped %>%
  ggplot(aes(x = log_cash, y = stroop_time, color = factor(treatment))) +
  geom_point() +
  labs(x = 'Financial Condition (Log-transformed Cash)', y = 'Cognitive Performance', title = 'Fincancia
```




```
# After Payday
povertyA.dropped <- poverty.dropped %>%
  filter(treatment == "After Payday")
povertyA.dropped %>%
  ggplot(aes(x = log_cash, y = stroop_time, color = factor(treatment))) +
  geom_point() +
  labs(x = 'Financial Condition (Log-transformed Cash)', y = 'Cognitive Performance', title = 'Fincancia
```




```
# side by side scatter plot for better comparison and interpretation
poverty.dropped %>%
  ggplot(aes(x = log_cash, y = stroop_time)) +
  geom_point() +
  facet_wrap(~treatment)
```




```
# Linear Correlation
cor(povertyB.dropped$log_cash, povertyB.dropped$stroop_time)
```

```
## [1] -0.006017568
```

```
cor(povertyA.dropped$log_cash, povertyA.dropped$stroop_time)
```

```
## [1] 0.02100648
```

Text Answer for 6a:

By looking at the scatter plots between logged transformed cash variable and log transformed stroop_test variable of both of the groups of Before Payday (treatment) and After Payday (control), it gets clear that there is no linear relation between the two variables for each of the groups. The correlation coefficient for the Before Payday group between logged cash and stroop_test variables is negative .006 while in After Payday group it is positive 0.02 but on the whole it does not signals any strong relationship between the variables of cash and cognitive performance. Therefore, overall, on average, it is concluded that there is very week association between presence of cash on the cognitive performance of these groups and there is no linear relation between cash and cognitive performance.
6b

```
# Your code for 6b goes here
```

Text Answer for 6b:

Based on the above results of scatter plot and linear correlation, there is no strong relation between the two variables, however, there is very very weak though. This result straight away neither contradicts the hypothesis nor strongly supports it and for safe interpretation further statistics might help as the sample size is very small and on the basis of results we get we are not in a position to safely conclude and take either of the side for the population as a whole. But, it is evident, on the basis of this sample result, hypothesis is not proved significantly right.