HW 4

Arjun Ganesh

1. [4] Write a code to extract data (subset) on the donors in the PVA_data data frame with more than
   \$50 in Current.Gift donations AND have more than one-year lapse in between gift donations. [Note: be
   sure to include the output of your code.]

```
PVA_data<-read.csv("/cloud/project/PVA.csv")
PVA_data[PVA_data$Current.Gift>(50.00) & PVA_data$Time.Between.Gifts>1,]
```

```
##    Age Own.Home Sex Children Income Smallest.Gift Largest.Gift Current.Gift
## 1   64       H   F        0      1             3          300          200
## 2   91       U   M        0      1            25           50          200
## 3   82       H   M        0      5             2          250          200
## 4   61       H   F        0      1            20           25          200
## 5   71       U   M        0      5            25           25          150
## 6   90       U   F        0      1             5          133          102
## 7   57       H   F        0      1            25          100          101
## 8   38       U   F        0      7            50           50          100
## 9   43       H   M        0      6           200          200          100
## 10  79       U   M        0      1            10           50          100
## 11  81       H   F        0      6             5          100          100
## 12  50       H   M        0      1            10          100          100
## 13  73       H   M        0      4             2          100          100
## 14  48       U   F        0      2            20          100          100
## 15  83       U   F        0      5            10           16          100
## 16  90       U   F        0      1             5           75          100
## 17  42       U   M        0      3            20           32          100
## 18  61       U   M        0      1             3          100          100
## 19  69       H   M        0      7             5           30          100
## 20  61       U   M        0      1            50           50          100
## 21  48       H   M        0      6            25          200          100
## 22  62       H   M        0      1            25           30          100
## 23  82       U   F        0      2            15          200          100
## 24  66       H   M        0      5            43           50          100
## 25  74       H   M        0      2             5          100          100
## 26  62       H   M        0      5             3           25          100
## 28  54       H   F        0      4            25           25          100
## 29  71       U   M        0      5             4           10          100
## 30  68       U   M        0      5            50           50           75
## 31  49       U   M        0      1             5           70           75
## 32  82       U   F        0      6             5           50           75
## 34  68       H   F        0      6            25          200           75
## 35  74       H   M        0      7             5           50           75
## 37  69       H   M        0      6             5           50           60
## 38  60       H   M        0      1            10           50           60
## 39  46       H   F        0      5             2           45           60
## 40  47       H   M        0      6            25           40           55
```

```
## 41 64          U   M        0        2               10              58              53
## 42 79          H   M        0        7                3              60              53
## 43 69          H   M        0        4               15              36              51
## 44 68          H   F        0        1                5              15              51
## 45 50          H   F        0        4               25              36              51
##      Other.Gifts Time.Between.Gifts
## 1              0                  3
## 2              1                 13
## 3             42                  6
## 4              0                  5
## 5              2                  5
## 6              0                  5
## 7              0                  3
## 8              0                  9
## 9              7                 10
## 10             0                 13
## 11             0                 10
## 12             0                  7
## 13             7                 12
## 14             0                 18
## 15             0                 24
## 16             0                  3
## 17             0                  8
## 18             0                 21
## 19             0                  7
## 20             0                  7
## 21            37                  6
## 22             0                  6
## 23            61                 14
## 24             6                 10
## 25            13                 10
## 26            14                  3
## 28             4                 11
## 29             0                 13
## 30             0                  2
## 31             0                  7
## 32             0                 10
## 34             7                  3
## 35             0                  7
## 37             2                  5
## 38             0                  5
## 39             2                  3
## 40             4                 11
## 41             1                  7
## 42             1                  8
## 43             0                  3
## 44             0                 22
## 45             0                  8
```

2. [3] Write a one line code to simultaneously compute the standard deviations of the Smallest.Gift,Largest.Gift and Current.Gift variables. What can you tell the CEO of the organization about the standard deviations of these three variables? [Note: be sure to include the output of your code.]

```
apply(PVA_data[,6:8],2,sd)
```

```
## Smallest.Gift  Largest.Gift   Current.Gift
##      8.599788     20.004166      12.405962
```

Mr. CEO, I would say that the Largest Gifts have a high standard deviation meaning that there is a higher variability on our distribution of gifts. That is probably what can be predicted. The Current Gifts are inbetween small and large gifts as predicted as well. If we wanted more analysis, I would need specific variables you would like me to test for so that we can make even better speculations.

3. [3] Write a one line code to change the Sex, Income and Current.Gift variables to gender,income_level and current_gift, respectively. [Note: just your code will suffice, no output necessary.]

```
names(PVA_data)[c(3,5,8)]<-c("gender","income_level","current_gift")
```

4. [4] Find the mean and median current_gift of donors grouped by their income_level. [Note: be sure to include the output of your code.]

```
aggregate(x=PVA_data$current_gift,list(PVA_data$income_level),mean)
```

```
##   Group.1         x
## 1       1  9.566821
## 2       2  8.437417
## 3       3  8.768559
## 4       4  9.508008
## 5       5 10.632430
## 6       6 11.666407
## 7       7 11.313636
```

```
aggregate(x=PVA_data$current_gift,list(PVA_data$income_level),median)
```

```
##   Group.1  x
## 1       1  7
## 2       2  5
## 3       3  6
## 4       4  9
## 5       5 10
## 6       6 10
## 7       7 10
```

5. [5] Repeat part 4 but restrict it to current gift donors who gave $50 or less. [Note: be sure to include the output of your code.]

```
sub_pva<-subset(x=PVA_data, PVA_data$current_gift<= 50)
aggregate(sub_pva$current_gift, list(sub_pva$income_level),mean)
```

```
##   Group.1         x
## 1       1  8.939638
## 2       2  8.123811
## 3       3  8.635569
## 4       4  9.242714
## 5       5  9.938764
## 6       6 10.906447
## 7       7 10.648318
```

```
aggregate(sub_pva$current_gift, list(sub_pva$income_level),median)
```

```
##   Group.1   x
## 1       1 7.0
```

```
## 2        2  5.0
## 3        3  6.0
## 4        4  8.0
## 5        5  9.5
## 6        6 10.0
## 7        7 10.0
```

6. [4] Subset the PVA_data data frame to remove the Age, Children and Time.Between.Gifts variables and save the resulting data frame as revised_pva. [Note: just your code will suffice, no output necessary.]

```
revised_pva <- PVA_data[-c(1,4,10)]
```

7. [4] Combine the first 100 observations and the last 50 observations of the revised_pva data frame and save the resulting data frame as Homework4_Out. [Note, just your code will suffice, no output necessary.]

```
first_obv<-revised_pva[1:100,]
second_obv<-revised_pva[7551:7600,]
Homework4_Out<-rbind(first_obv,second_obv)
```

8. [3] Export the data frame created in question 7 in a .csv file named Homework4_Out.csv. [Note, upload Homework4_Out.csv file as part of your homework submission.]

```
write.csv(Homework4_Out,"Homework4_Out.csv")
```