14.750 Problem Set 3

Wonjae Lee

April 2021

Problem 1

Answer 1.a

```
library(tidyverse)
```

```
## -- Attaching packages -------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.0     v dplyr   1.0.5
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

```
setwd("C:/Users/wonja/Documents/GitHub/14.750/PS3")
rm(list = ls())
load("votingData.RData")

model1 <- lm(formula =  votePercent ~ campaignDonation, data = votingData)
summary(model1)
```

```
##
## Call:
## lm(formula = votePercent ~ campaignDonation, data = votingData)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.124117 -0.022814 -0.000893  0.023577  0.097810
##
```

```
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0078643  0.0043062   234.1   <2e-16 ***
## campaignDonation -0.0120381  0.0001168  -103.1   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.035 on 998 degrees of freedom
## Multiple R-squared:  0.9141, Adjusted R-squared:  0.914
## F-statistic: 1.062e+04 on 1 and 998 DF,  p-value: < 2.2e-16
```

```
print(paste("standard deviation of campaign donation = ", sd(votingData$campaignDonation)))
```

```
## [1] "standard deviation of campaign donation =  9.48050088642655"
```

```
stargazer(model1, type = "latex", out = "PS3-1a.tex", covariate.labels = "Campaign donation
", dep.var.labels = "Percent of votes")
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac
 at fas.harvard.edu
## % Date and time: Thu, Apr 01, 2021 - 5:10:18 AM
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{1}{c}{\textit{Dependent variable:}} \\
## \cline{2-2}
## \\[-1.8ex] & Percent of votes \\
## \hline \\[-1.8ex]
##  Campaign donation & $-$0.012$^{***}$ \\
##   & (0.0001) \\
##   & \\
##  Constant & 1.008$^{***}$ \\
##   & (0.004) \\
##   & \\
## \hline \\[-1.8ex]
## Observations & 1,000 \\
## R$^{2}$ & 0.914 \\
## Adjusted R$^{2}$ & 0.914 \\
## Residual Std. Error & 0.035 (df = 998) \\
## F Statistic & 10,624.630$^{***}$ (df = 1; 998) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{1}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \
\
## \end{tabular}
## \end{table}
```

Answer 1.b

```
ggplot(data = votingData, aes(x = campaignDonation, y = votePercent)) +
  labs(x = "Campaign donation", y = "Percent of votes") +
  geom_point() +
  geom_smooth(method = lm, se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
```
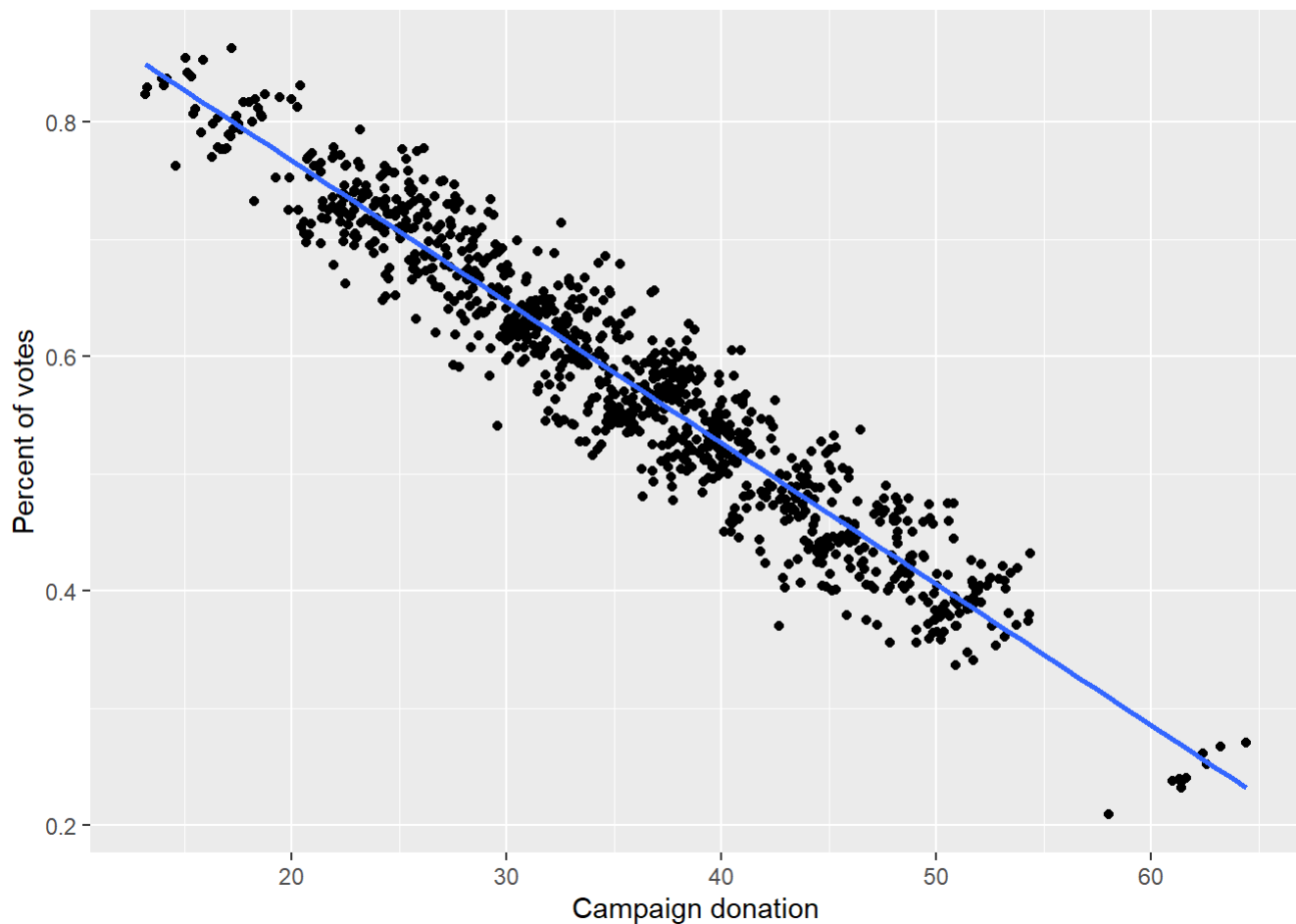


```
ggsave(filename = "PS3-1b.png")
```

```
## Saving 7 x 5 in image
## `geom_smooth()` using formula 'y ~ x'
```

Answer 1.c

```
model2 <- lm(formula = votePercent ~ campaignDonation + factor(county), data = votingData)
summary(model2)
```

```
##
## Call:
## lm(formula = votePercent ~ campaignDonation + factor(county),
##     data = votingData)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0224805 -0.0050592  0.0000658  0.0050019  0.0209900
##
```

```
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.2441272  0.0095853 -25.469  < 2e-16 ***
## campaignDonation   0.0124727  0.0001816  68.688  < 2e-16 ***
## factor(county)2    0.1285640  0.0035017  36.714  < 2e-16 ***
## factor(county)3    0.2914779  0.0039934  72.990  < 2e-16 ***
## factor(county)4    0.4309809  0.0046496  92.692  < 2e-16 ***
## factor(county)5    0.8140589  0.0069285 117.494  < 2e-16 ***
## factor(county)6    0.8226434  0.0068636 119.855  < 2e-16 ***
## factor(county)7    0.2944241  0.0040161  73.312  < 2e-16 ***
## factor(county)8    0.3602820  0.0043065  83.660  < 2e-16 ***
## factor(county)9    0.6748510  0.0058621 115.122  < 2e-16 ***
## factor(county)10   0.8401808  0.0071061 118.234  < 2e-16 ***
## factor(county)11   0.4839082  0.0048347 100.092  < 2e-16 ***
## factor(county)12   0.0239089  0.0033665   7.102 2.50e-12 ***
## factor(county)13   0.1267844  0.0035325  35.891  < 2e-16 ***
## factor(county)14   0.2771629  0.0039981  69.324  < 2e-16 ***
## factor(county)15   0.0506873  0.0033903  14.951  < 2e-16 ***
## factor(county)16   0.0799933  0.0034379  23.268  < 2e-16 ***
## factor(county)17   0.2440442  0.0038626  63.181  < 2e-16 ***
## factor(county)18   0.4370725  0.0046466  94.063  < 2e-16 ***
## factor(county)19   0.2730612  0.0039055  69.916  < 2e-16 ***
## factor(county)20   0.3592189  0.0043221  83.112  < 2e-16 ***
## factor(county)21   0.6813466  0.0061601 110.606  < 2e-16 ***
## factor(county)22   0.1659084  0.0035950  46.149  < 2e-16 ***
## factor(county)23   0.4930417  0.0049433  99.739  < 2e-16 ***
## factor(county)24   0.5512330  0.0052455 105.087  < 2e-16 ***
## factor(county)25   0.2532727  0.0039130  64.726  < 2e-16 ***
## factor(county)26   0.6601260  0.0059486 110.972  < 2e-16 ***
## factor(county)27  -0.0132241  0.0033633  -3.932 9.07e-05 ***
## factor(county)28   0.0914006  0.0034619  26.402  < 2e-16 ***
## factor(county)29   0.6165760  0.0056016 110.071  < 2e-16 ***
## factor(county)30   0.1803439  0.0036191  49.831  < 2e-16 ***
## factor(county)31   0.7502708  0.0064643 116.065  < 2e-16 ***
## factor(county)32   0.3577691  0.0043013  83.176  < 2e-16 ***
## factor(county)33   0.1954368  0.0036508  53.533  < 2e-16 ***
## factor(county)34   0.4284043  0.0045853  93.430  < 2e-16 ***
## factor(county)35   0.3420038  0.0041473  82.465  < 2e-16 ***
## factor(county)36   0.6553146  0.0058777 111.493  < 2e-16 ***
## factor(county)37   0.7333306  0.0062988 116.423  < 2e-16 ***
## factor(county)38   0.6370149  0.0058046 109.742  < 2e-16 ***
## factor(county)39   0.5903661  0.0054878 107.578  < 2e-16 ***
## factor(county)40   0.1893713  0.0036181  52.340  < 2e-16 ***
## factor(county)41   0.6605118  0.0059137 111.692  < 2e-16 ***
## factor(county)42   0.6041473  0.0056576 106.784  < 2e-16 ***
## factor(county)43   0.6961131  0.0061709 112.806  < 2e-16 ***
## factor(county)44   0.1202113  0.0035040  34.307  < 2e-16 ***
## factor(county)45  -0.0052659  0.0033635  -1.566   0.1178
## factor(county)46   0.6974481  0.0060928 114.470  < 2e-16 ***
## factor(county)47   0.3855514  0.0044425  86.787  < 2e-16 ***
## factor(county)48   0.4643899  0.0048572  95.608  < 2e-16 ***
## factor(county)49   0.1160533  0.0034623  33.519  < 2e-16 ***
## factor(county)50  -0.2814186  0.0038733 -72.656  < 2e-16 ***
## factor(county)51   0.3650265  0.0042137  86.628  < 2e-16 ***
## factor(county)52   0.1969044  0.0036547  53.877  < 2e-16 ***
## factor(county)53   0.2789016  0.0039401  70.786  < 2e-16 ***
## factor(county)54   0.3833207  0.0042753  89.658  < 2e-16 ***
## factor(county)55   0.6531805  0.0058855 110.981  < 2e-16 ***
```

```
## factor(county)56    0.6025159  0.0055209 109.133  < 2e-16 ***
## factor(county)57    0.2985239  0.0041147  72.550  < 2e-16 ***
## factor(county)58    0.1206805  0.0035049  34.432  < 2e-16 ***
## factor(county)59    0.5362908  0.0052433 102.280  < 2e-16 ***
## factor(county)60    0.4928407  0.0049526  99.511  < 2e-16 ***
## factor(county)61    0.7053387  0.0062693 112.507  < 2e-16 ***
## factor(county)62   -0.0036800  0.0033595  -1.095   0.2736
## factor(county)63    0.3419146  0.0041770  81.857  < 2e-16 ***
## factor(county)64    0.2062565  0.0037515  54.980  < 2e-16 ***
## factor(county)65    0.3454754  0.0043308  79.771  < 2e-16 ***
## factor(county)66    0.6358214  0.0056006 113.526  < 2e-16 ***
## factor(county)67    0.3892577  0.0044948  86.602  < 2e-16 ***
## factor(county)68    0.9011430  0.0074152 121.527  < 2e-16 ***
## factor(county)69    0.4187145  0.0045971  91.082  < 2e-16 ***
## factor(county)70    0.4276991  0.0046449  92.080  < 2e-16 ***
## factor(county)71    0.4903638  0.0049593  98.878  < 2e-16 ***
## factor(county)72    0.1585673  0.0036226  43.772  < 2e-16 ***
## factor(county)73    0.4617505  0.0047438  97.339  < 2e-16 ***
## factor(county)74    0.5908258  0.0055210 107.014  < 2e-16 ***
## factor(county)75   -0.0534664  0.0033736 -15.848  < 2e-16 ***
## factor(county)76    0.2533737  0.0038756  65.377  < 2e-16 ***
## factor(county)77    0.1070097  0.0034500  31.018  < 2e-16 ***
## factor(county)78    0.4811104  0.0048780  98.629  < 2e-16 ***
## factor(county)79    0.4649951  0.0047151  98.619  < 2e-16 ***
## factor(county)80    0.4936900  0.0049807  99.120  < 2e-16 ***
## factor(county)81    0.4963765  0.0049801  99.672  < 2e-16 ***
## factor(county)82    0.5665021  0.0053746 105.403  < 2e-16 ***
## factor(county)83    0.3369118  0.0041608  80.972  < 2e-16 ***
## factor(county)84    0.3510203  0.0042813  81.989  < 2e-16 ***
## factor(county)85   -0.0080160  0.0033606  -2.385   0.0173 *
## factor(county)86    0.3384538  0.0041823  80.925  < 2e-16 ***
## factor(county)87    0.2607051  0.0038653  67.447  < 2e-16 ***
## factor(county)88    0.5281148  0.0051233 103.082  < 2e-16 ***
## factor(county)89    0.4546656  0.0047526  95.667  < 2e-16 ***
## factor(county)90    0.0558967  0.0034076  16.404  < 2e-16 ***
## factor(county)91    0.5248775  0.0052007 100.925  < 2e-16 ***
## factor(county)92    0.2938758  0.0040386  72.767  < 2e-16 ***
## factor(county)93    0.3813104  0.0044905  84.915  < 2e-16 ***
## factor(county)94    0.3625517  0.0043268  83.792  < 2e-16 ***
## factor(county)95    0.2082754  0.0037329  55.794  < 2e-16 ***
## factor(county)96    0.6904761  0.0060121 114.847  < 2e-16 ***
## factor(county)97    0.3261767  0.0040676  80.189  < 2e-16 ***
## factor(county)98    0.2742659  0.0039729  69.035  < 2e-16 ***
## factor(county)99    0.5626790  0.0052944 106.278  < 2e-16 ***
## factor(county)100   0.0804135  0.0033915  23.710  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.007512 on 899 degrees of freedom
## Multiple R-squared:  0.9964, Adjusted R-squared:  0.996
## F-statistic:  2514 on 100 and 899 DF,  p-value: < 2.2e-16
```

```
stargazer(model2, type = "latex", out = "PS3-1c.tex", covariate.labels = "Campaign donation
", dep.var.labels = "Percent of votes", add.lines = list(c("County fixed effect", "Yes"), c
("","")), keep = "campaignDonation", table.layout = "=ld-t-a-s=n")
```

```
##
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac
 at fas.harvard.edu
## % Date and time: Thu, Apr 01, 2021 - 5:10:21 AM
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \\[-1.8ex]\hline
## \hline \\[-1.8ex]
##  & \multicolumn{1}{c}{\textit{Dependent variable:}} \\
## \cline{2-2}
## \\[-1.8ex] & Percent of votes \\
## \hline \\[-1.8ex]
##  Campaign donation & 0.012$^{***}$ \\
##    & (0.0002) \\
##    & \\
## \hline \\[-1.8ex]
## County fixed effect & Yes \\
##   & \\
## \hline \\[-1.8ex]
## Observations & 1,000 \\
## R$^{2}$ & 0.996 \\
## Adjusted R$^{2}$ & 0.996 \\
## Residual Std. Error & 0.008 (df = 899) \\
## F Statistic & 2,513.537$^{***}$ (df = 100; 899) \\
## \hline
## \hline \\[-1.8ex]
## \textit{Note:}  & \multicolumn{1}{r}{$^{*}$p$<$0.1; $^{**}$p$<$0.05; $^{***}$p$<$0.01} \
\
## \end{tabular}
## \end{table}
```

Answer 1.e

```
county10 <- votingData %>% filter(county == 10)
county50 <- votingData %>% filter(county == 50)
ggplot(data = NULL, aes(x = campaignDonation, y = votePercent)) +
  labs(x = "Campaign donation", y = "Percent of votes") +
  geom_point(data = county10) +
  geom_smooth(data = county10, method = lm, se = FALSE) +
  geom_point(data = county50) +
  geom_smooth(data = county50, method = lm, se = FALSE)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```
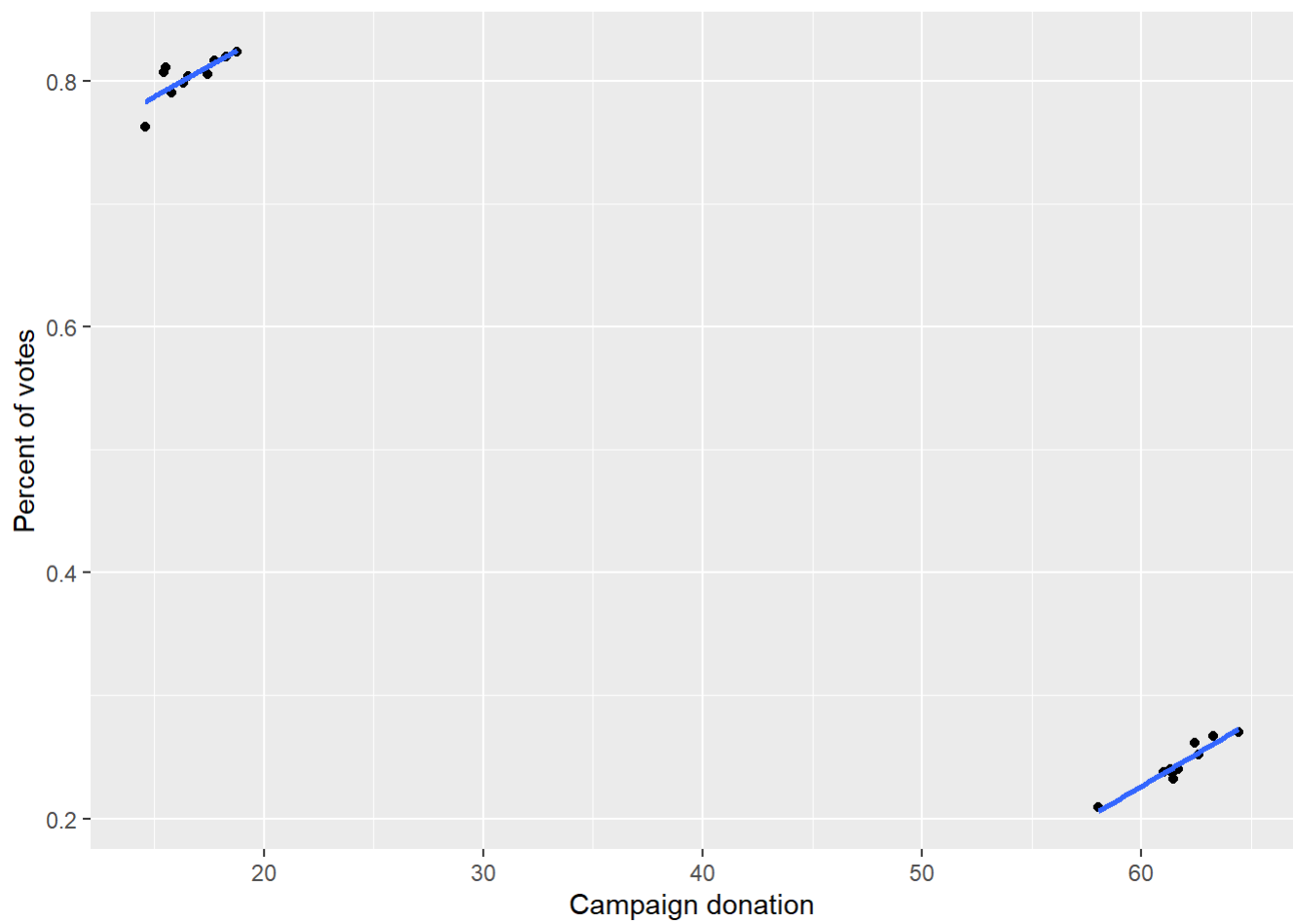
```
ggsave(filename = "PS3-1e.png")
```

```
## Saving 7 x 5 in image
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```