ECON219 Project

Sara Bolf

3/4/2019

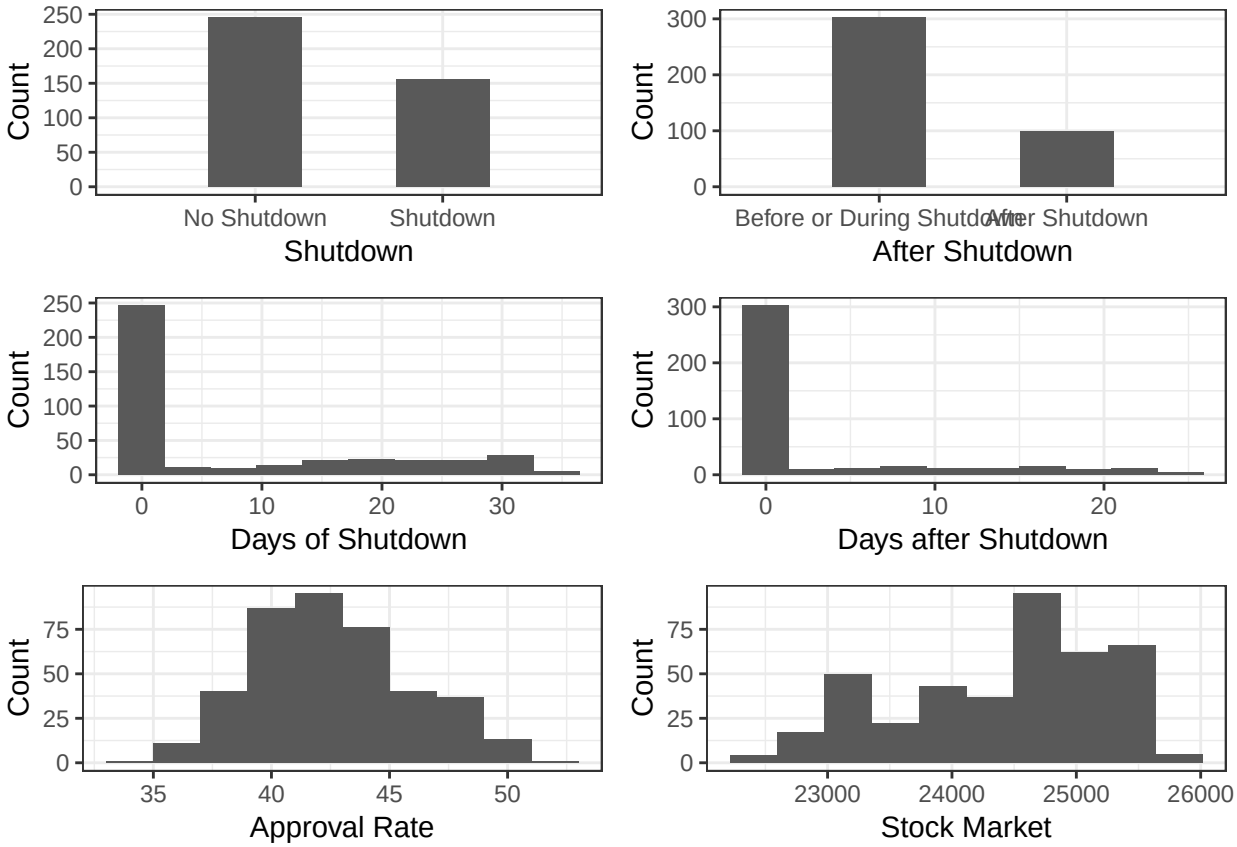
```
library(readxl)
poll_data <- read_excel("Downloads/trump-approval-ratings 2/polldatafinal.xltx")
pollvar <- c("approve","strike","poststrike","strikedays","daysafter","stockmarket")
poll_data <- poll_data[pollvar]
colnames(poll_data)<-c("approve","sd","postsd","sddays","daysafter","stockmarket")
attach(poll_data)
```

```
# create table of summary statistics for poll data

poll_summary <- do.call(cbind, lapply(poll_data, summary))

sum_table <- kable(round(poll_summary,3),col.names = c("Approval Rate","Shutdown","After Shutdown","Numb

kabtab <- kable(round(poll_summary,3),col.names = c("Approval Rate","Shutdown","After Shutdown","Number

kable_as_image(kabtab, filename = "poll_table",file_format = "png")
```

	Approval Rate	Shutdown	After Shutdown	Number of Days of Shutdown	Number of Days after Shutdown	Stock Market
Min.	34.000	0.000	0.000	0.000	0.000	22493.90
1st Qu.	40.600	0.000	0.000	0.000	0.000	23839.16
Median	42.900	0.000	0.000	0.000	0.000	24558.43
Mean	42.961	0.389	0.247	7.852	3.183	24390.58
3rd Qu.	45.000	1.000	0.000	17.000	0.000	25126.54
Max.	52.000	1.000	1.000	34.500	24.500	25911.83

```
# histogram plots for each variable in the poll data set

b1 <- ggplot(poll_data,aes(x= sd))+
  geom_bar(width=.5)+
  theme_bw() +
  scale_x_discrete(limits=c(0, 1),labels = c("No Shutdown","Shutdown"))+
  labs(x="Shutdown",y="Count")


b2 <- ggplot(poll_data,aes(x= postsd))+
  geom_bar(width=.5)+
  theme_bw() +
  scale_x_discrete(limits=c(0, 1),labels = c("Before or During Shutdown","After Shutdown"))+
  labs(x="After Shutdown",y="Count")



h1 <- ggplot(poll_data,aes(x=sddays))+
  geom_histogram(bins=10)+
  theme_bw()+
  labs(x="Days of Shutdown",y="Count")


h2 <- ggplot(poll_data,aes(x=daysafter))+
```

```
  geom_histogram(bins=10)+
  theme_bw()+
  labs(x="Days after Shutdown",y="Count")



h01<- ggplot(poll_data,aes(x=approve))+
  geom_histogram(bins=10)+
  theme_bw()+
  labs(x="Approval Rate",y="Count")


h02 <- ggplot(poll_data, aes(x=stockmarket))+
  geom_histogram(bins=10)+
  theme_bw()+
  labs(x="Stock Market",y="Count")

grid.arrange(b1,b2,h1,h2,h01,h02)
```




```
ggsave("approvedata.png", plot =grid.arrange(b1,b2,h1,h2,h01,h02) )
```

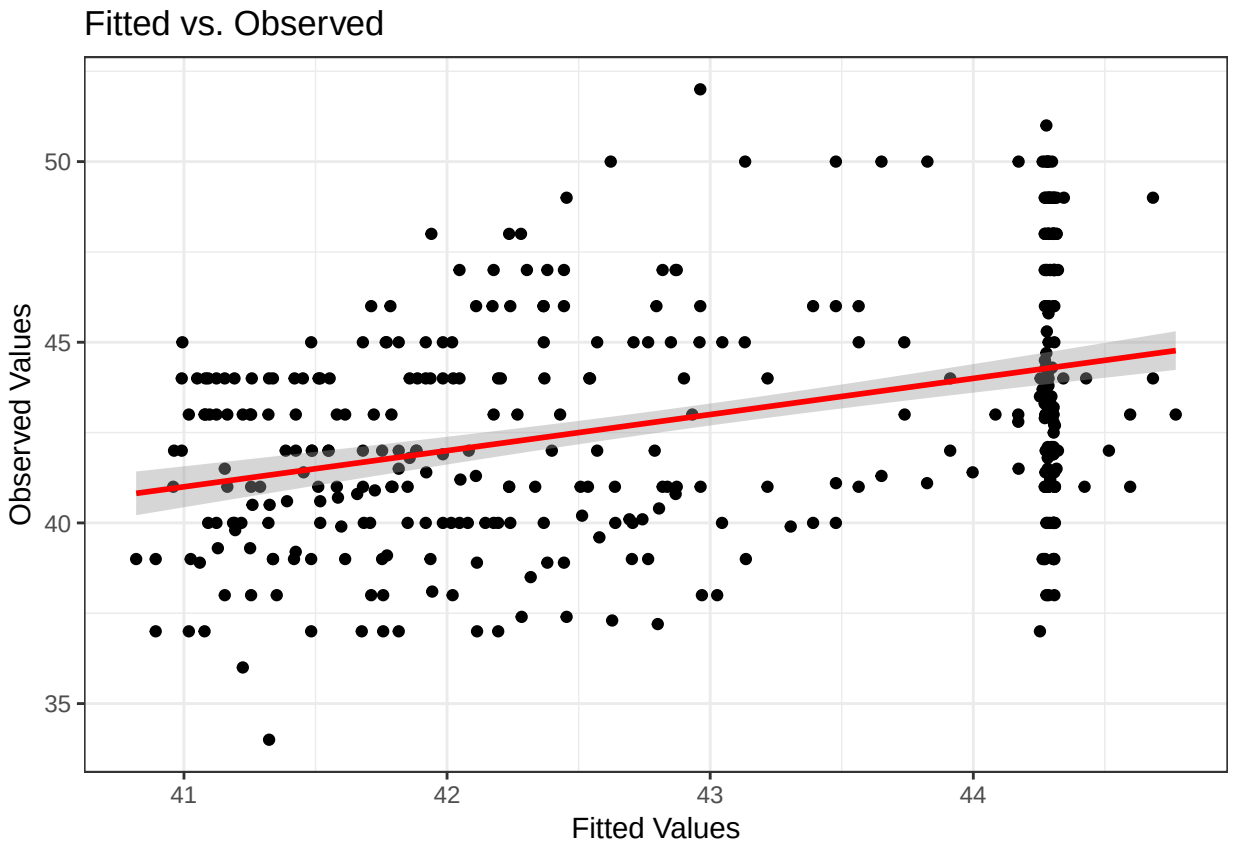
```
## Saving 6.5 x 4.5 in image
```

```
# linear model with dependent variable approve and independent variables
# sd, postsd, sddays, daysafter, and stockmarket.

pollmod <- lm(approve ~ sd+postsd+sddays+daysafter+stockmarket)
```

```r
summary(pollmod)
```

```
##
## Call:
## lm(formula = approve ~ sd + postsd + sddays + daysafter + stockmarket)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.3236 -2.3054 -0.5104  2.6089  9.0376
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.523e+01  1.391e+01   3.251  0.00125 **
## sd          -1.235e+00  1.580e+00  -0.782  0.43468
## postsd      -3.746e+00  8.131e-01  -4.607 5.53e-06 ***
## sddays      -6.314e-02  4.325e-02  -1.460  0.14514
## daysafter    1.733e-01  5.609e-02   3.089  0.00215 **
## stockmarket -3.776e-05  5.585e-04  -0.068  0.94612
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.111 on 395 degrees of freedom
## Multiple R-squared:  0.1403, Adjusted R-squared:  0.1294
## F-statistic: 12.89 on 5 and 395 DF,  p-value: 1.263e-11
```

```r
# Create Fitted vs. Observed plot

ggplot(pollmod, aes(x = pollmod$fitted.values, y = approve)) + geom_point() +
  stat_smooth(method = "lm", col = "red") +
  labs(title="Fitted vs. Observed",x = "Fitted Values",y = "Observed Values")+
  theme_bw()
```




```
ggsave(file="fittedvobserved.png", last_plot())
```

```
## Saving 6.5 x 4.5 in image
```

```
# create ANOVA table for pollmod

poll_anova <- as.data.frame(anova(pollmod))
row.names(poll_anova) <- c("Shutdown","After Shutdown","Days of Shutdown","Days after Shutdown","Stock M

anova_tab <- kable(poll_anova,format = "latex")

kable_as_image(anova_tab, filename = "poll_anova",file_format = "png")
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Shutdown	1	324.5171002	324.5171002	33.5270735	0.0000000
After Shutdown	1	135.2888464	135.2888464	13.9771959	0.0002124
Days of Shutdown	1	49.3412441	49.3412441	5.0976282	0.0245038
Days after Shutdown	1	114.5736197	114.5736197	11.8370285	0.0006426
Stock Market	1	0.0442606	0.0442606	0.0045727	0.9461208
Residuals	395	3823.3058017	9.6792552	NA	NA

```
# 95% confidence intervals for each of the regression coefficients and intercept.

ci <- (confint(pollmod,conf.level=0.95))

ci_df <- as.data.frame(ci)
```

```
row.names(ci_df) <- c("(Intercept)","Shutdown","After Shutdown","Days of Shutdown","Days after Shutdown

ci_tab <- kable(round(ci_df,3),format = "latex")

kable_as_image(ci_tab, filename = "sd_ci",file_format = "png")
```

	2.5 %	97.5 %
(Intercept)	17.877	72.587
Shutdown	-4.340	1.870
After Shutdown	-5.344	-2.147
Days of Shutdown	-0.148	0.022
Days after Shutdown	0.063	0.284
Stock Market	-0.001	0.001

```
mean(pollmod$residuals)
```

```
## [1] 1.175559e-16
```

```
# find fourth moment for each variable in poll_data

kurtosis(poll_data$approve)
```

```
## [1] 2.541077
```

```
kurtosis(poll_data$sd)
```

```
## [1] 1.207248
```

```
kurtosis(poll_data$postsd)
```

```
## [1] 2.37832
```

```
kurtosis(poll_data$sddays)
```

```
## [1] 2.424664
```

```
kurtosis(poll_data$daysafter)
```

```
## [1] 5.235354
```

```
kurtosis(poll_data$stockmarket)
```

```
## [1] 2.225409
```

```
# creating table of VIF values for each independent variable in pollmod.

vif_df <- as.data.frame((vif(pollmod)))
```

```r
row.names(vif_df) <- c("Shutdown","After Shutdown","Days of Shutdown","Days after Shutdown", "Stock Mar

vif_tab <- kable(round(vif_df,3),col.names = c("VIF"),format = "latex")

kable_as_image(vif_tab, filename = "vif_poll",file_format = "png")
```

	VIF
Shutdown	24.567
After Shutdown	5.093
Days of Shutdown	9.739
Days after Shutdown	5.295
Stock Market	8.676

```r
# correlation table for variables in poll_data

sum_table <- kable(round(cor(poll_data),3),col.names = c("Approval Rate","Shutdown","After Shutdown","Nu

kabtab <- kable(round(cor(poll_data),3),col.names = c("Approval Rate","Shutdown","After Shutdown","Numbe

kable_as_image(kabtab, filename = "pollcor",file_format = "png")
```

	Approval Rate	Shutdown	After Shutdown	Number of Days of Shutdown	Number of Days after Shutdown	Stock Market
approve	1.000	-0.270	-0.032	-0.288	0.051	0.209
sd	-0.270	1.000	-0.457	0.878	-0.399	-0.843
postsd	-0.032	-0.457	1.000	-0.401	0.872	0.416
sddays	-0.288	0.878	-0.401	1.000	-0.350	-0.559
daysafter	0.051	-0.399	0.872	-0.350	1.000	0.449
stockmarket	0.209	-0.843	0.416	-0.559	0.449	1.000