Homework 5

Max Ryoo (hr2ee)

library(ggplot2) library(car) library(gcookbook) library(MASS) library(Hmisc)

Problem 1

```
library(ggplot2)
library(car)
```

```
## Loading required package: carData
```

```
library(gcookbook)
library(MASS)
library(Hmisc)
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
##
## Attaching package: 'Hmisc'
```

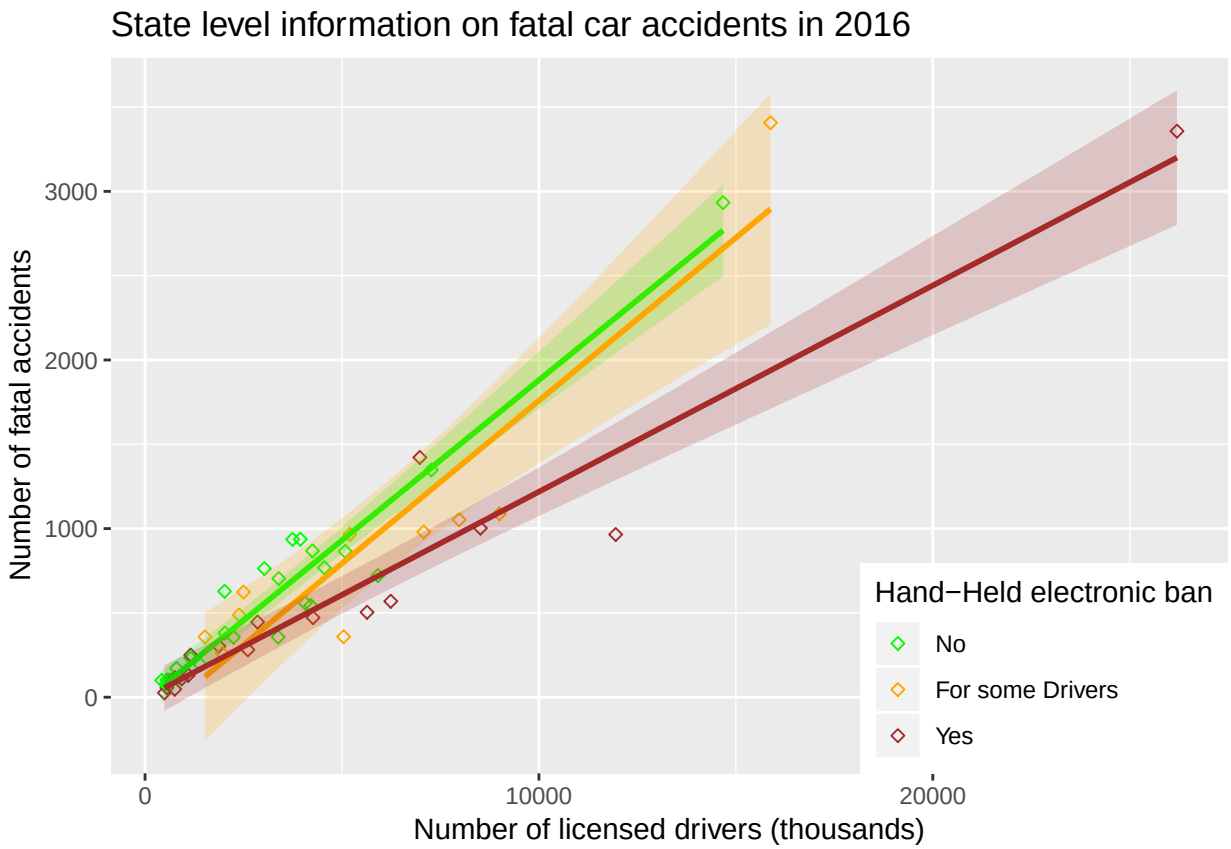
```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```
setwd("/Users/maxryoo/Documents/Fall 2018/STAT3080/HW3")
crash <- read.csv("state crashes.csv")
crash_plot <- ggplot(crash, aes(x=Licensed.drivers,
y=Fatal.crashes)) + geom_point(shape=23)
crash_plot <- ggplot(crash, aes(x=Licensed.drivers, y=Fatal.crashes,
colour=as.character(Hand.held.ban)))

crash_handheld <- crash_plot + geom_point(
shape=23)
smooth_crash1 <- crash_handheld +
  geom_smooth(aes(fill=as.character(Hand.held.ban)
),method=lm,alpha=0.2, show.legend =FALSE) + scale_fill_manual(
values=c("1"="green","2"="orange", "3"="brown"))

# Recreated the graphic #
Graphics1 <- smooth_crash1
```

```
ans1 <- Graphics1 + scale_colour_manual(labels=c("No","For some Drivers","Yes"),
                                values=c("green", "orange", "brown")) +
  labs(title="State level information on fatal car accidents in 2016",
                    x="Number of licensed drivers (thousands)",
                y="Number of fatal accidents",
       color="Hand-Held electronic ban") +
  theme(legend.position = c(1,0),
        legend.justification = c(1,0))
print(ans1)
```

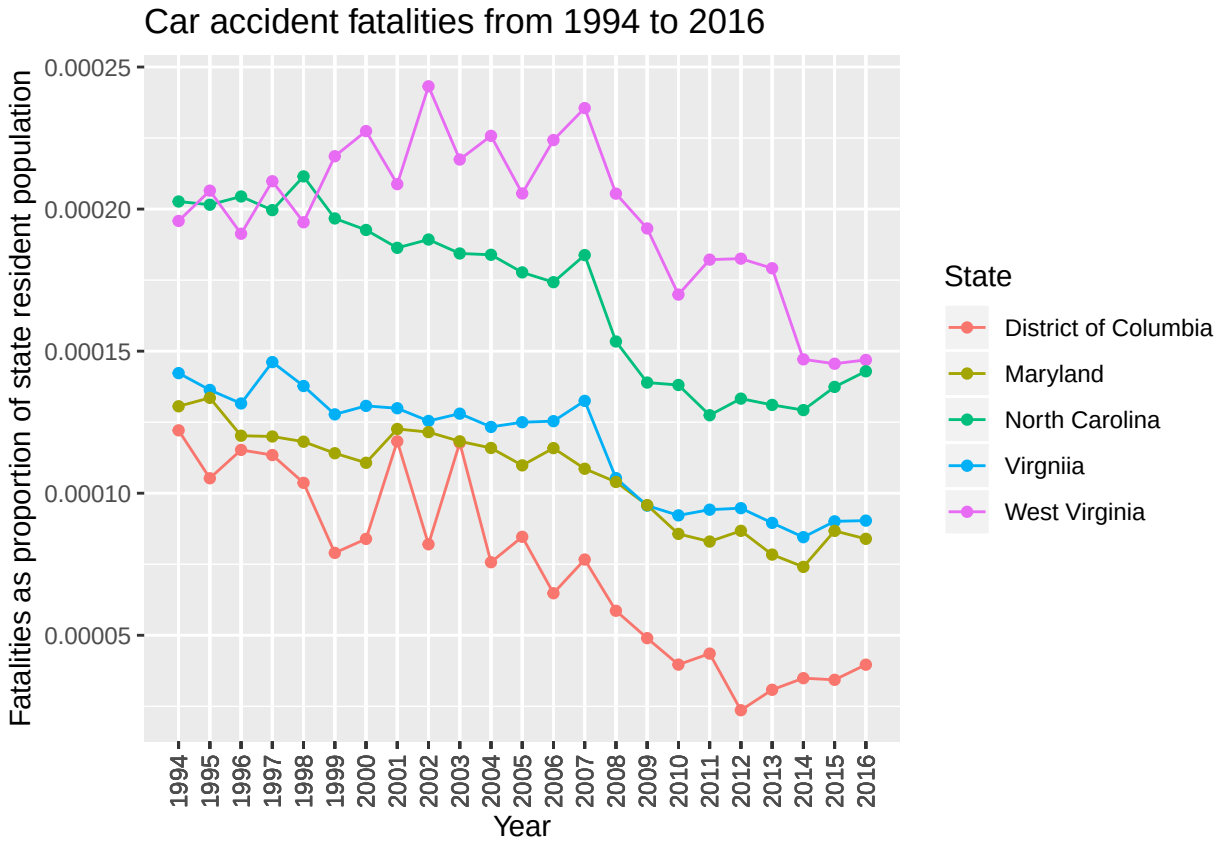


I imported the code for making the base graph from previous homework. I then changed the labels to "NO", "for some drivers", and "Yes" for the hand.held.ban column. I then added the title x axis title, y axis title, and distinction title (legend). I positioned the legend at the bottom right corner with position and justification function.

Problem 2

```
setwd("/Users/maxryoo/Documents/Fall 2018/STAT3080/HW4/")
fatalities <- read.csv("fatalities.csv")
lsmooth <- ggplot(fatalities,
                  aes(x=Year, y=Fatalities,
                      color=State)) + geom_line()
```

```
fatalitiesp <- fatalities
fatalitiesp$proportion <-
  fatalities$Fatalities /
  fatalities$Resident.Population/1000
lsmoothp <- ggplot(fatalitiesp,
                   aes(x=Year,
                       y=proportion,
                       color=State)) +
  geom_line()
pointslsmoothp1 <- lsmoothp + geom_point()
# created graph #
Graphic2 <- pointslsmoothp1
ans2 <- Graphic2 +
  labs(title="Car accident fatalities from 1994 to 2016",
                      y="Fatalities as proportion of state resident population") +
  scale_x_continuous(breaks=fatalities$Year) +
  theme(axis.text.x=element_text(angle=90,vjust=0.5)) +
  theme(panel.grid.minor.x=element_blank())
print(ans2)
```



I imported the code for making the base graph from previous homework. I then added the title and y axis title. The x axis title was already Years so it was okay to leave it. I added the years inbetween with the sacle_x_continuous function with breaks as the years in the

dataframe. I aligned the x axis accordingly to the instruction and left out the x minor grid lines since they weren't representing years.

References

1. http://r.789695.n4.nabble.com/ggplot-not-showing-all-the-years-on-the-x-axis-td4654925.html