Exploratory Anamytics

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.3
```

```
library(plyr)
```

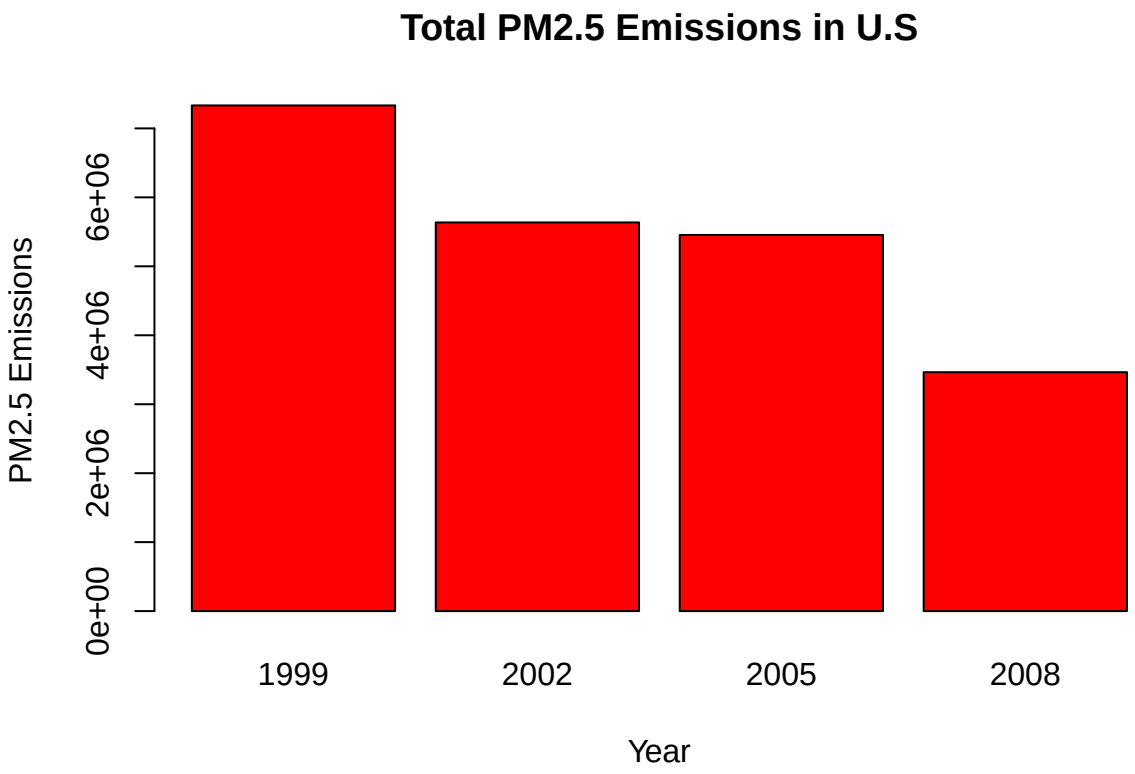
```
## Warning: package 'plyr' was built under R version 3.3.3
```

```
setwd("H:/Data Science Johns Hopkins/exploratory-data-analysis/Week4/data")
NEI <- readRDS("summarySCC_PM25.rds")
SCC <- readRDS("Source_Classification_Code.rds")
```

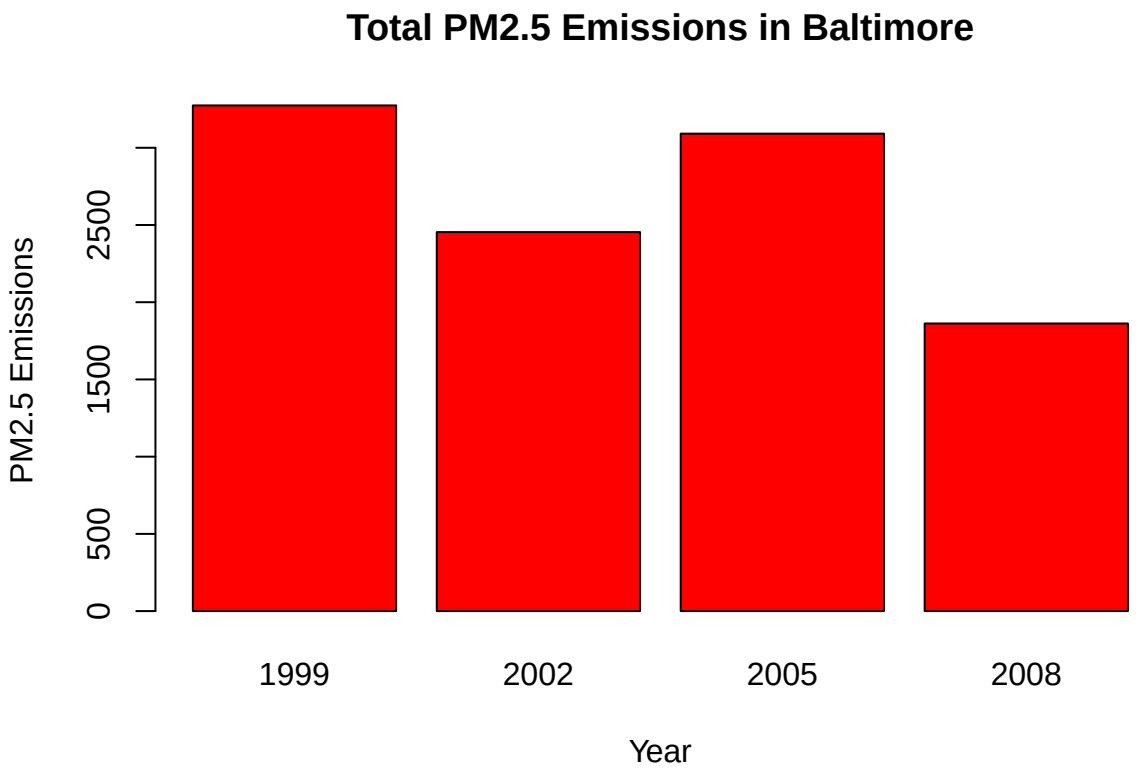
Including Plots

You can also embed plots, for example:

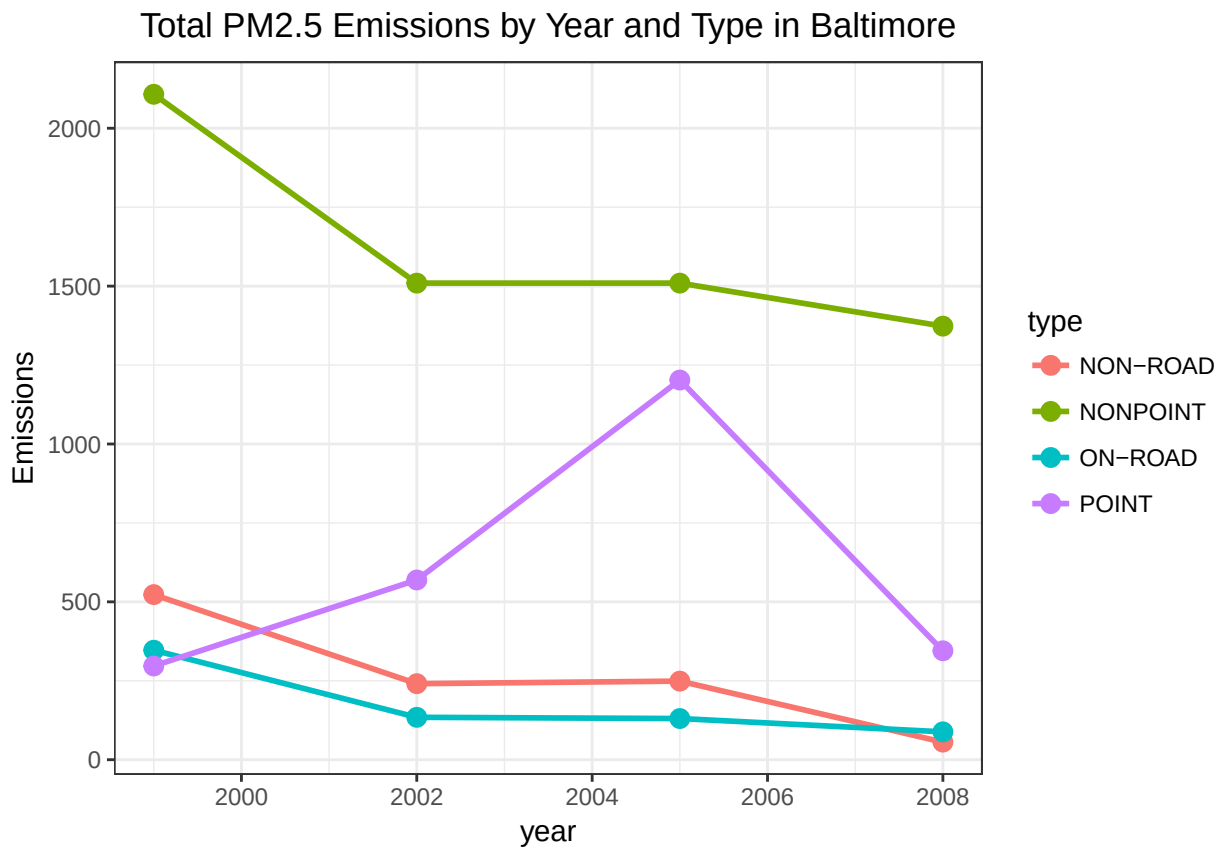
```
totEmissions <- aggregate(Emissions ~ year,NEI, sum)
barplot(
  totEmissions$Emissions,
  names.arg=totEmissions$year,
  xlab="Year",
  ylab="PM2.5 Emissions",
  col= "red",
  main="Total PM2.5 Emissions in U.S"
)
```



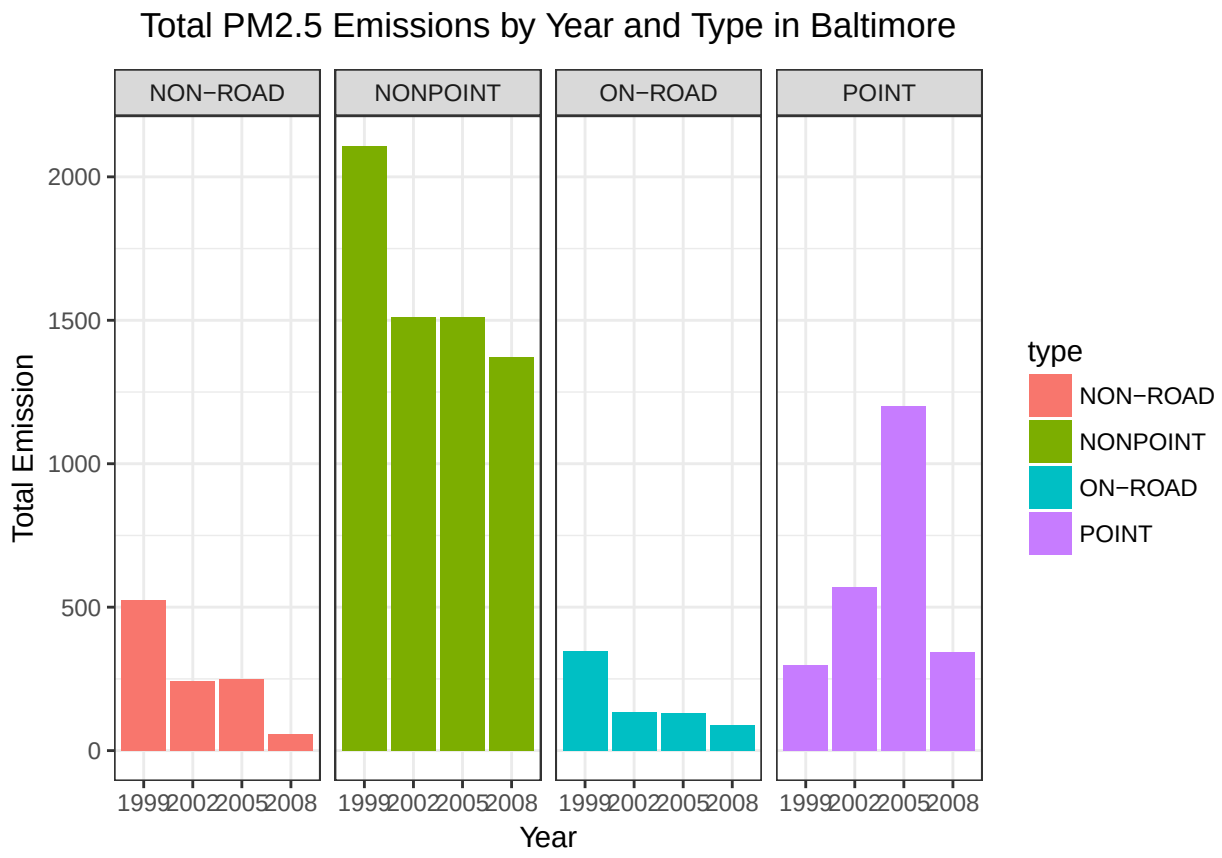

```
totEmBaltimore <- aggregate(Emissions ~ year,NEI[NEI$fips=="24510",], sum)
barplot(
  totEmBaltimore$Emissions,
  names.arg=totEmBaltimore$year,
  xlab="Year",
  ylab="PM2.5 Emissions",
  col= "red",
  main="Total PM2.5 Emissions in Baltimore"
)
```




```
totEmBaltimore <- aggregate(Emissions ~ year + type,NEI[NEI$fips=="24510",], sum)
ggplot(data=totEmBaltimore,
       aes(x=year, y=Emissions, colour=type)) +
  geom_line(size=1) +
   geom_point(size=3) +
  ggtitle("Total PM2.5 Emissions by Year and Type in Baltimore") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```



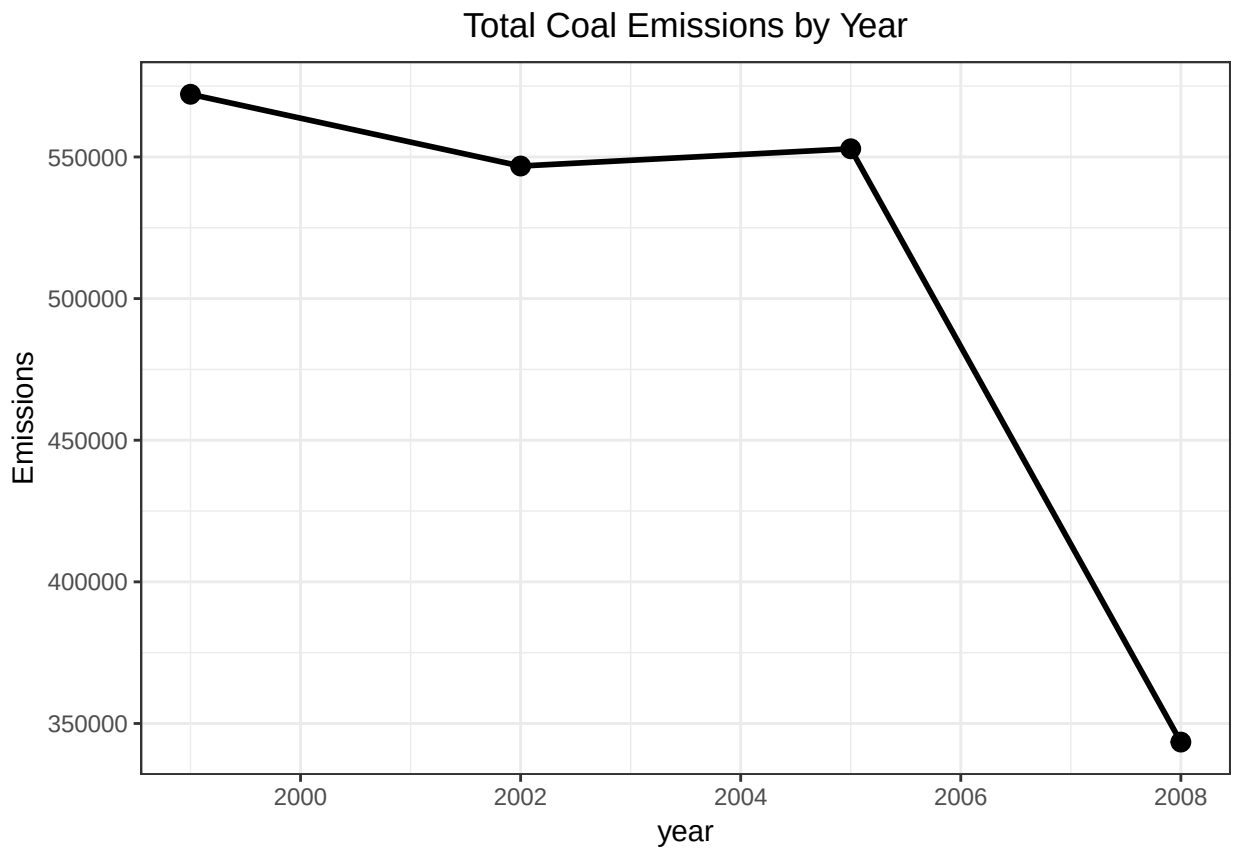

```
ggplot(totEmBaltimore,aes(factor(year),Emissions,fill=type)) +
  geom_bar(stat="identity") +
  facet_grid(.~type) +
  labs(x="Year", y="Total Emission") +
  labs(title=expression("Total PM2.5 Emissions by Year and Type in Baltimore"))+
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```



Find all coal combustion-related sources

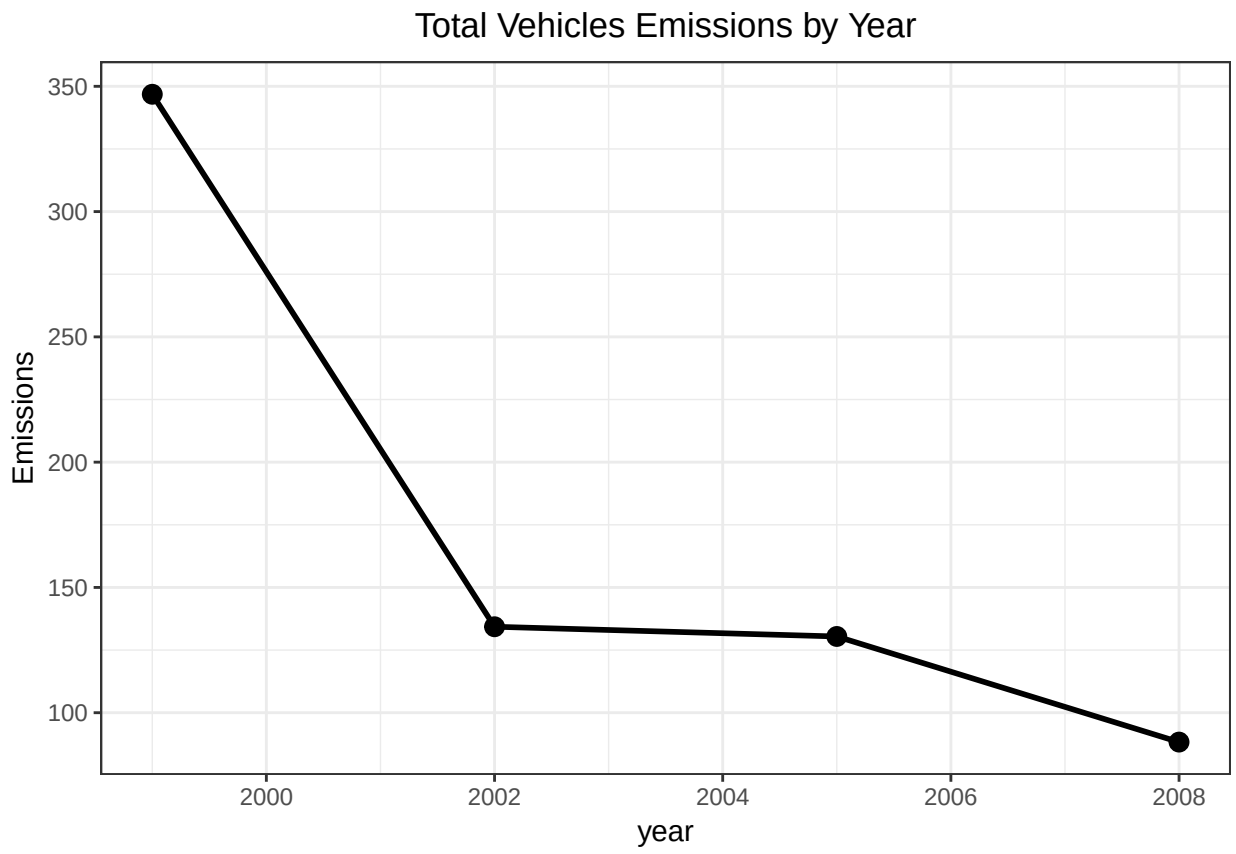
```
coalRow <- grep("coal", SCC$EI.Sector, ignore.case=TRUE)
coalRelated <-SCC[coalRow,"SCC"]
emFromCoal<-NEI[NEI$SCC %in% coalRelated,]
totEmCoal <- aggregate(Emissions ~ year ,emFromCoal, sum)

ggplot(data=totEmCoal,
       aes(x=year, y=Emissions)) +
  geom_line(size=1) +
  geom_point(size=3) +
  ggtitle("Total Coal Emissions by Year") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```

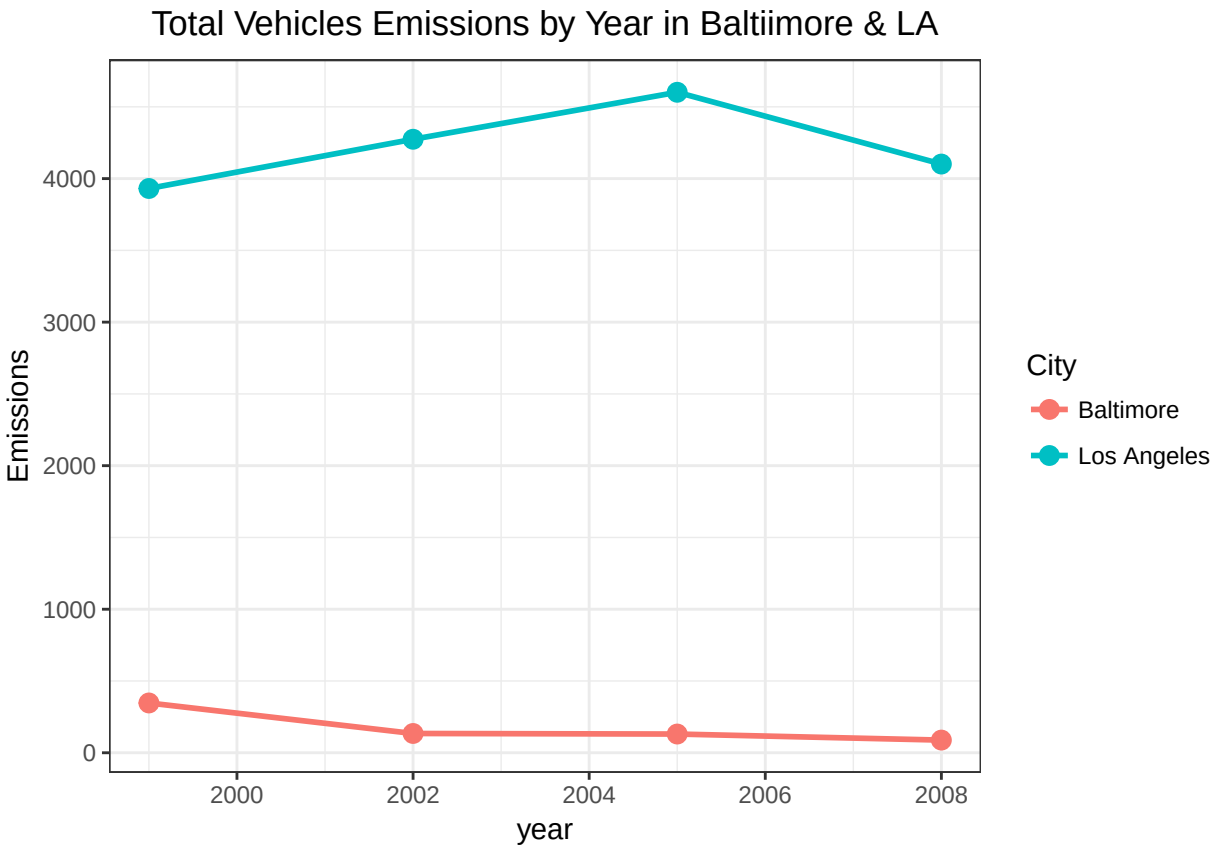



```
vehicleRow<- grep("Vehicle", SCC$EI.Sector, ignore.case = TRUE)
vehicleRelated<-SCC[vehicleRow,"SCC"]
emFromBaltVehicles<-NEI[NEI$SCC %in% vehicleRelated & NEI$fips=="24510",]
totEmBaltVehicles <- aggregate(Emissions ~ year ,emFromBaltVehicles, sum)

ggplot(data=totEmBaltVehicles,
       aes(x=year, y=Emissions)) +
  geom_line(size=1) +
  geom_point(size=3) +
  ggtitle("Total Vehicles Emissions by Year") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```




```
emFromBaltLAVehicles<- NEI[NEI$SCC %in% vehicleRelated & NEI$fips %in% c("24510","06037"),]
totEmBaltLAVehicles <- aggregate(Emissions ~ year + fips ,emFromBaltLAVehicles, sum)
totEmBaltLAVehicles <- rename(totEmBaltLAVehicles, c("fips"="City"))
totEmBaltLAVehicles$City <- factor(totEmBaltLAVehicles$City, levels=c( "24510","06037"), labels=c( "Bal
ggplot(data=totEmBaltLAVehicles,
       aes(x=year, y=Emissions,col=City)) +
  geom_line(size=1) +
  geom_point(size=3) +
  ggtitle("Total Vehicles Emissions by Year in Baltiimore & LA") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```


Total Vehicles Emissions by Year in Baltiimore & LA
Emissions
4000
3000
2000
1000
0
2000
2002
2004
2006
2008
year
City
Baltimore
Los Angeles