

DataAnalysis

AmrutaC

November 2, 2016

```
library(tidyverse)

## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr

## Conflicts with tidy packages -----

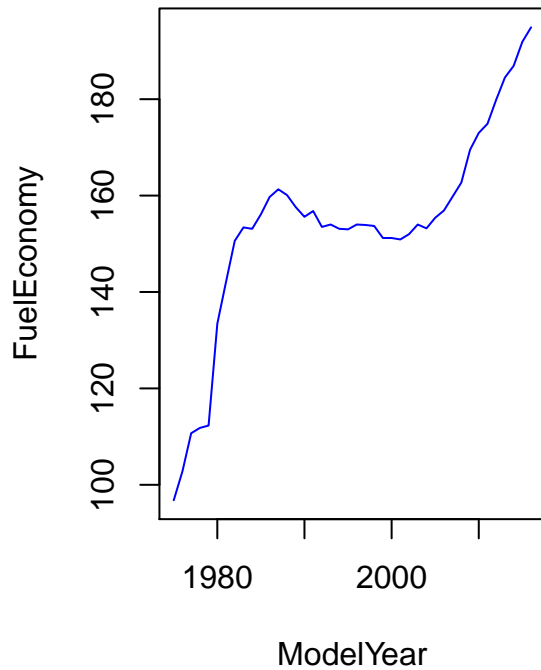
## filter(): dplyr, stats
## lag():      dplyr, stats

vehicle<-read.csv("D:/Fall2016/BDAA/PRO_DAT/vehicles.csv",header=T)
emission<-read.csv("D:/Fall2016/BDAA/PRO_DAT/emissions.csv",header=T)
economy<-read.csv("D:/Fall2016/BDAA/PRO_DAT/FuelEconomyData.csv",header=T)
smmry<-read.csv("D:/Fall2016/BDAA/PRO_DAT/co2_fuel_aggr.csv",header=T)
par(mfrow=c(1,2))

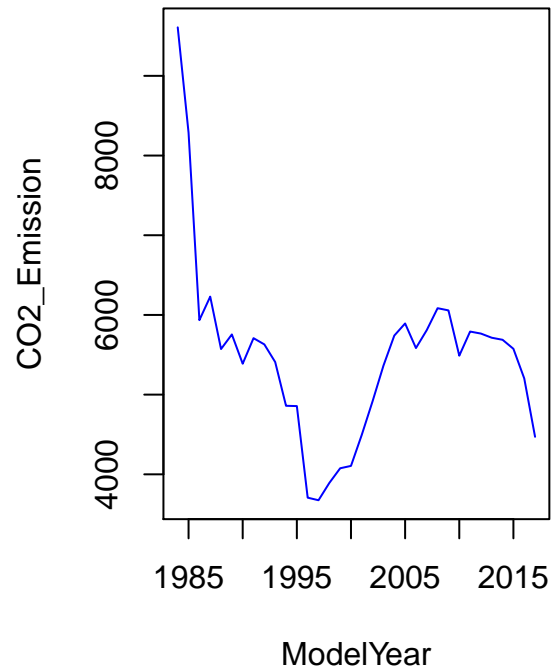
#Model Year Vs Fuel Economy
FuelYr<-aggregate(economy$FE.Adj.Comb~economy$Model.Year,economy,sum)
colnames(FuelYr)<-c("ModelYear","FuelEconomy")
plot(FuelYr,type="l",col="blue",main="Fuel Economy VS Model Year")

#Model Year Vs CO2 Emission
co2_FT<-aggregate(vehicle$co2TailpipeGpm~vehicle$co2TailpipeAGpm~vehicle$year,vehicle,sum)
colnames(co2_FT)<-c("ModelYear","CO2_Emission")
co2_FT$CO2_Emission<-(co2_FT[,2])/100
plot(co2_FT,type="l",col="blue",main="CO2 Emission VS Model Year")
```

Fuel Economy VS Model Year

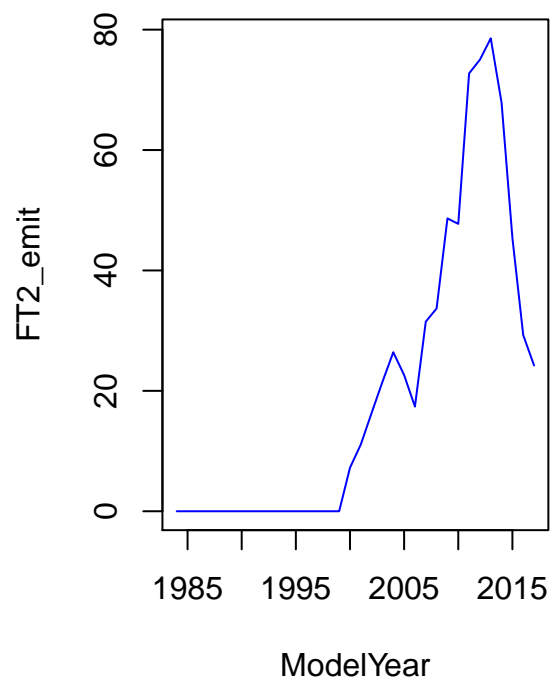
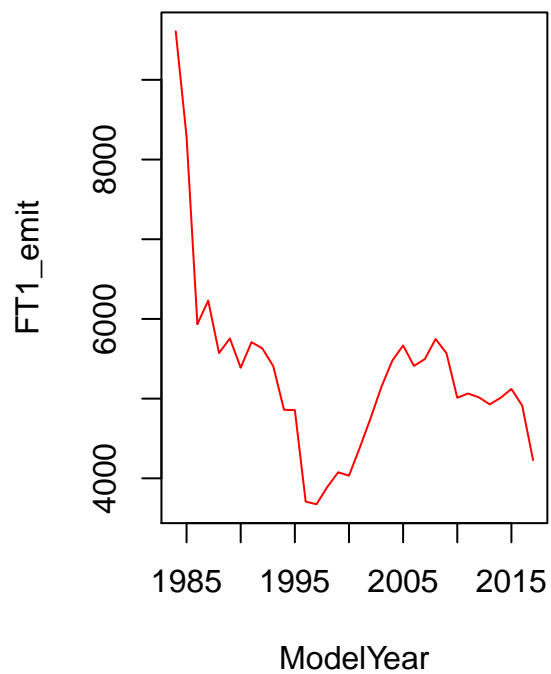


CO2 Emission VS Model Year

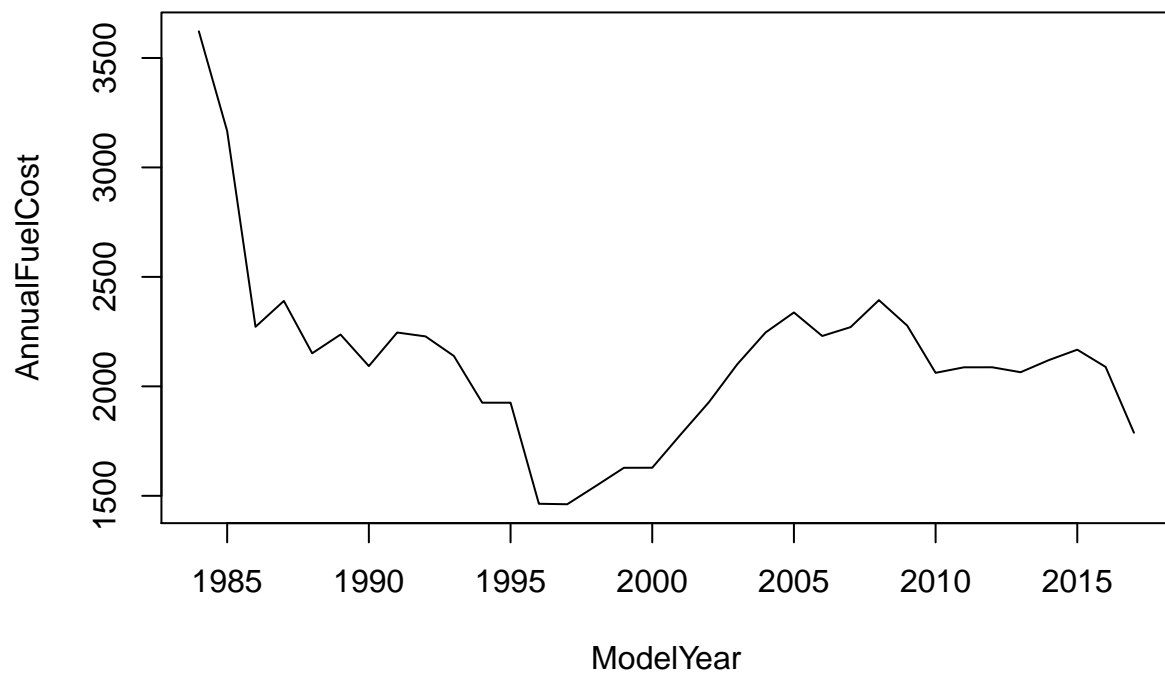


```
par(mfrow=c(1,2))
#FuelType1-Co2 Emission
co2_FT1<-aggregate(vehicle$co2TailpipeGpm~vehicle$year,vehicle,sum)
colnames(co2_FT1)<-c("ModelYear","FT1_emit")
co2_FT1$FT1_emit<-(co2_FT1[,2])/100
#co2_FT1<-co2_FT1[!co2_FT1$ModelYear > "2015",]
plot(co2_FT1,type="l",col="red")

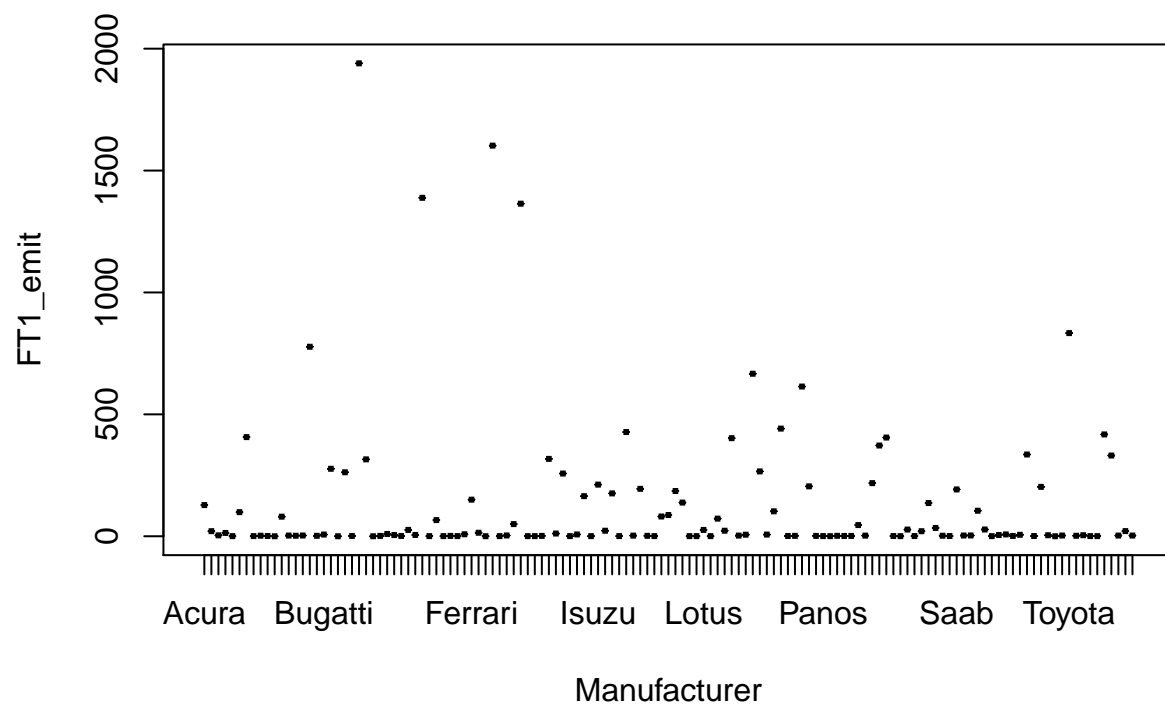
#FuelType2-Co2 Emission
co2_FT2<-aggregate(vehicle$co2TailpipeAGpm~vehicle$year,vehicle,sum)
colnames(co2_FT2)<-c("ModelYear","FT2_emit")
co2_FT2$FT2_emit<-co2_FT2[,2]/1000
#co2_FT2<-co2_FT2[!co2_FT2$ModelYear > "2015",]
plot(co2_FT2,type="l",col="blue")
```



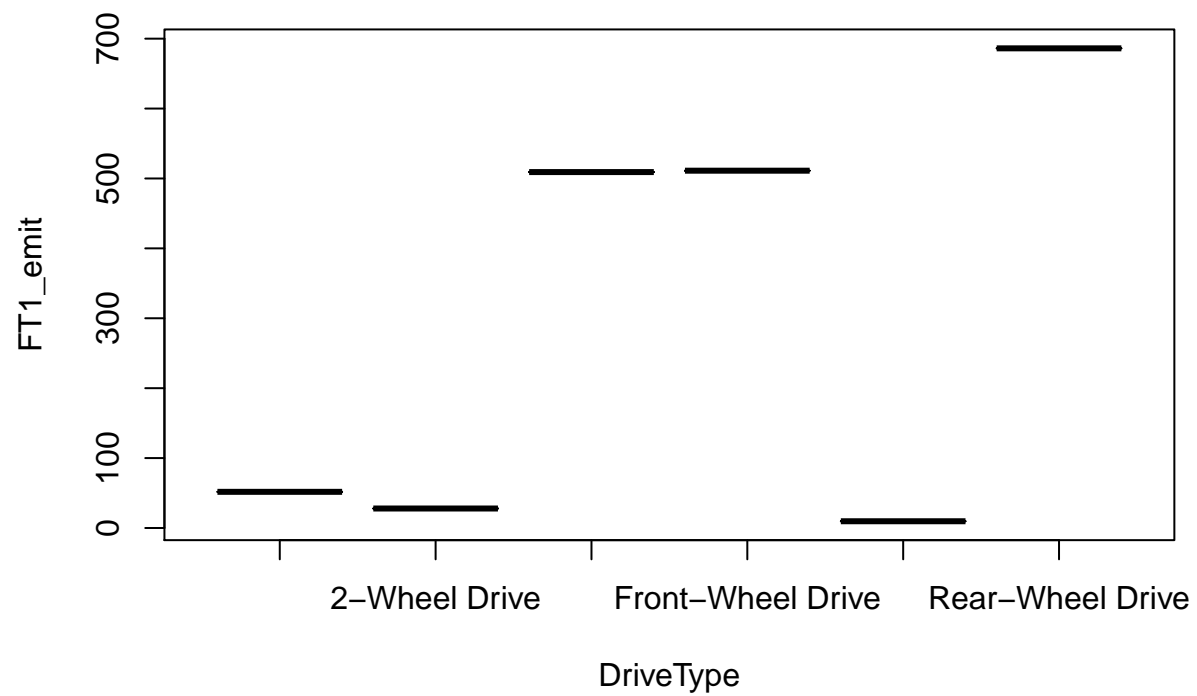
```
par(mfrow=c(1,1))
# Fuel Economy Vs Model Year
ecofuel<-aggregate(vehicle$fuelCost08~vehicle$year,vehicle,sum)
colnames(ecofuel)<-c("ModelYear","AnnualFuelCost")
ecofuel$AnnualFuelCost<-ecofuel[,2]/1000
plot(ecofuel,type='l')
```



```
#CO2 Emission Vs Make
Make<-aggregate(vehicle$co2TailpipeGpm~vehicle$make,vehicle,sum)
colnames(Make)<-c("Manufacturer","FT1_emit")
Make$FT1_emit<-Make[,2]/1000
Make<-Make[order(Make$FT1_emit),]
plot(Make,type="l")
```

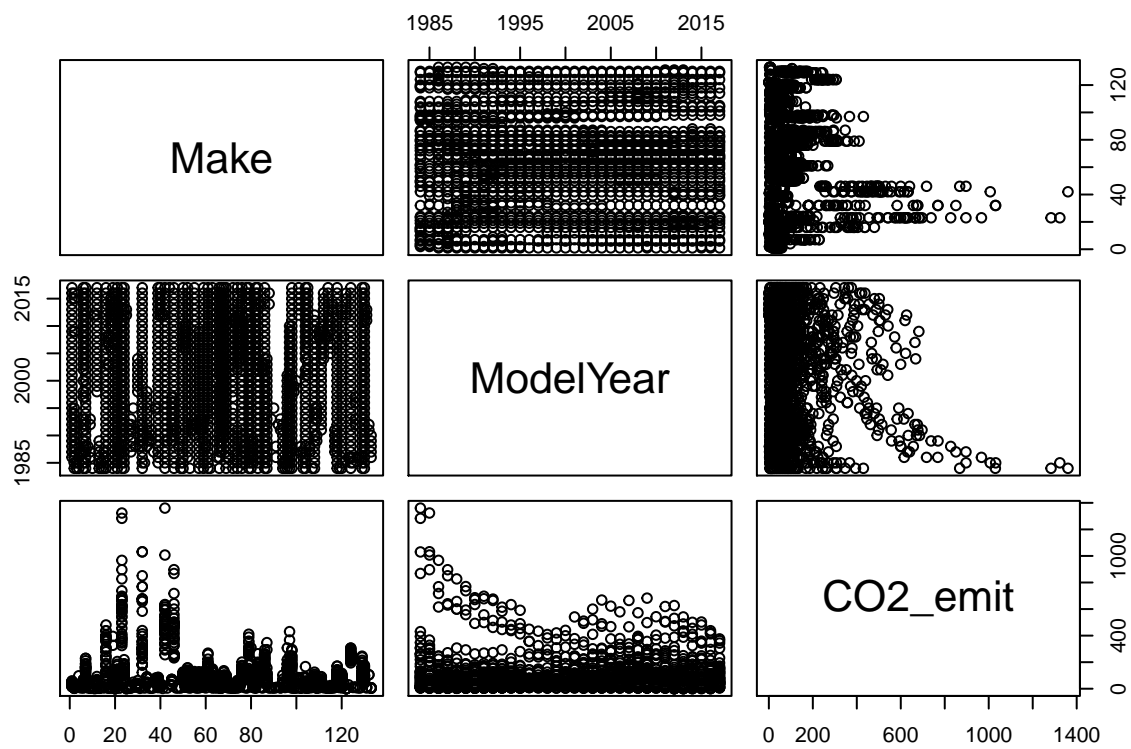


```
#CO2 Emission based on Drive type
co2_drv<-aggregate(vehicle$co2TailpipeGpm~vehicle$drive,vehicle,sum)
colnames(co2_drv)<-c("DriveType","FT1_emit")
co2_drv$FT1_emit<-co2_drv[,2]/10000
plot(co2_drv)
```

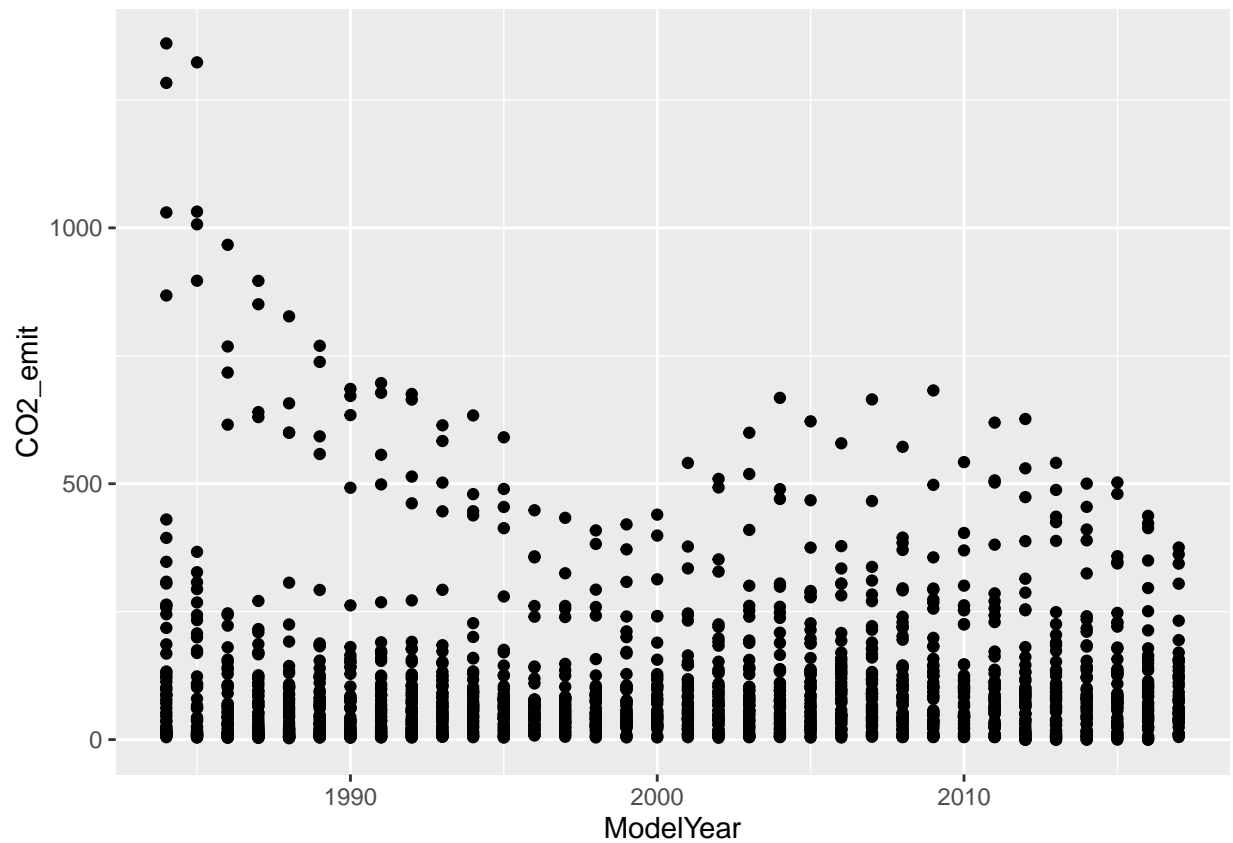


#CO2 Emission Vs Make and Year

```
MakeYr<-aggregate(vehicle$co2TailpipeGpm+vehicle$co2TailpipeAGpm~vehicle$make+vehicle$year,vehicle,sum)
colnames(MakeYr)<-c("Make", "ModelYear", "CO2_emit")
MakeYr$CO2_emit<-MakeYr[,3]/100
plot(MakeYr)
```



```
MakeYr<-MakeYr[order(-MakeYr$CO2_emit),]
ggplot(data = MakeYr) +
  geom_point(mapping = aes(x = ModelYear, y = CO2_emit))
```



```
#Drive Type CO2 Emission
```

```
DriveCO2<-aggregate(vehicle$co2TailpipeGpm+vehicle$co2TailpipeAGpm~vehicle$drive+vehicle$year,vehicle,s
```

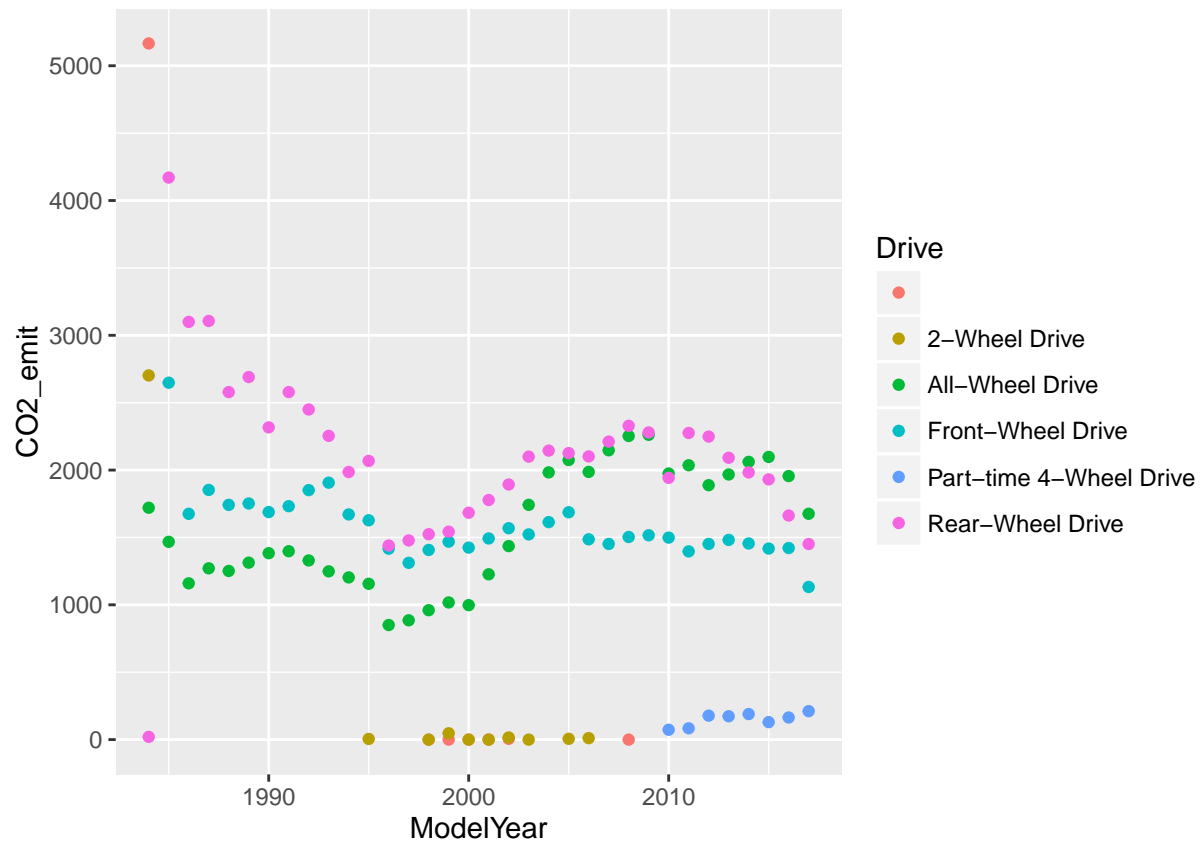
```
colnames(DriveCO2)<-c("Drive","ModelYear","CO2_emit")
```

```
DriveCO2$Drive<-na.omit(DriveCO2$Drive)
```

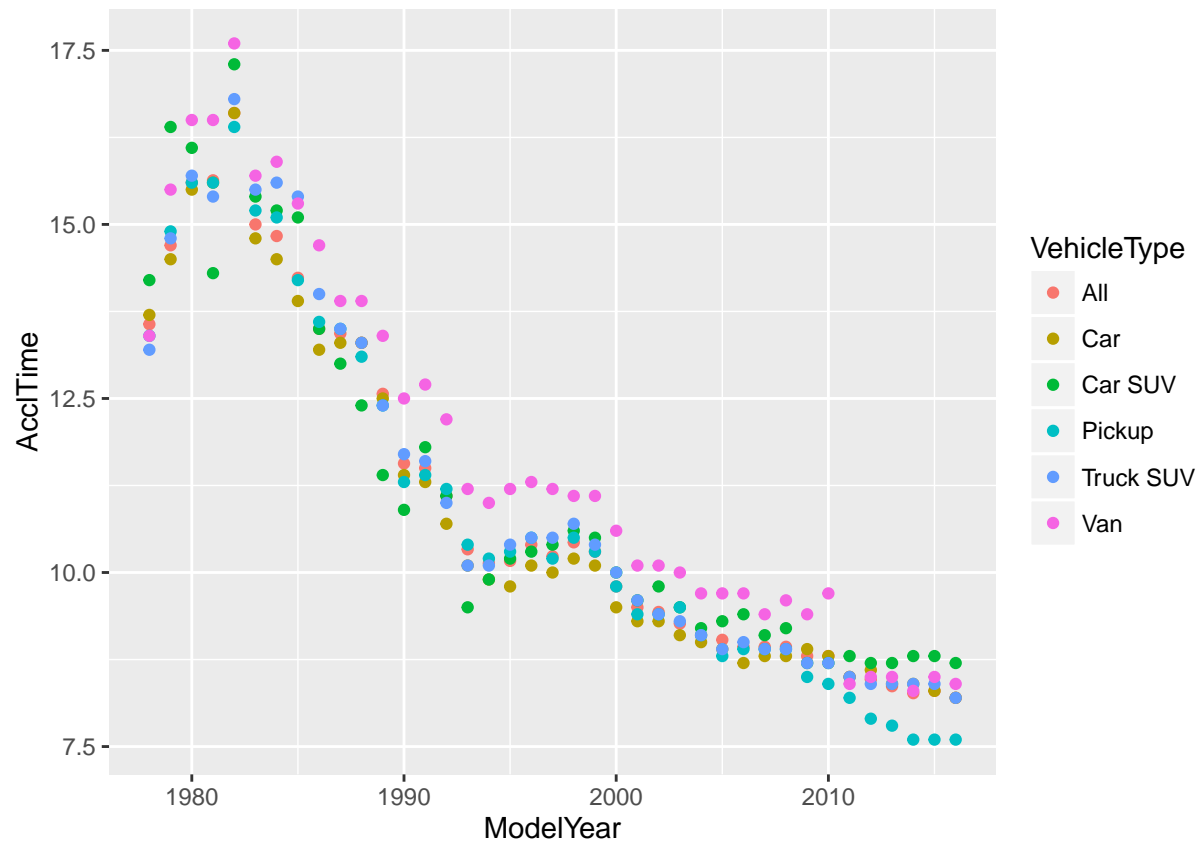
```
DriveCO2$CO2_emit<-DriveCO2[,3]/100
```

```
ggplot(data = DriveCO2) +
```

```
  geom_point(mapping = aes(x = ModelYear, y = CO2_emit,color=Drive))
```

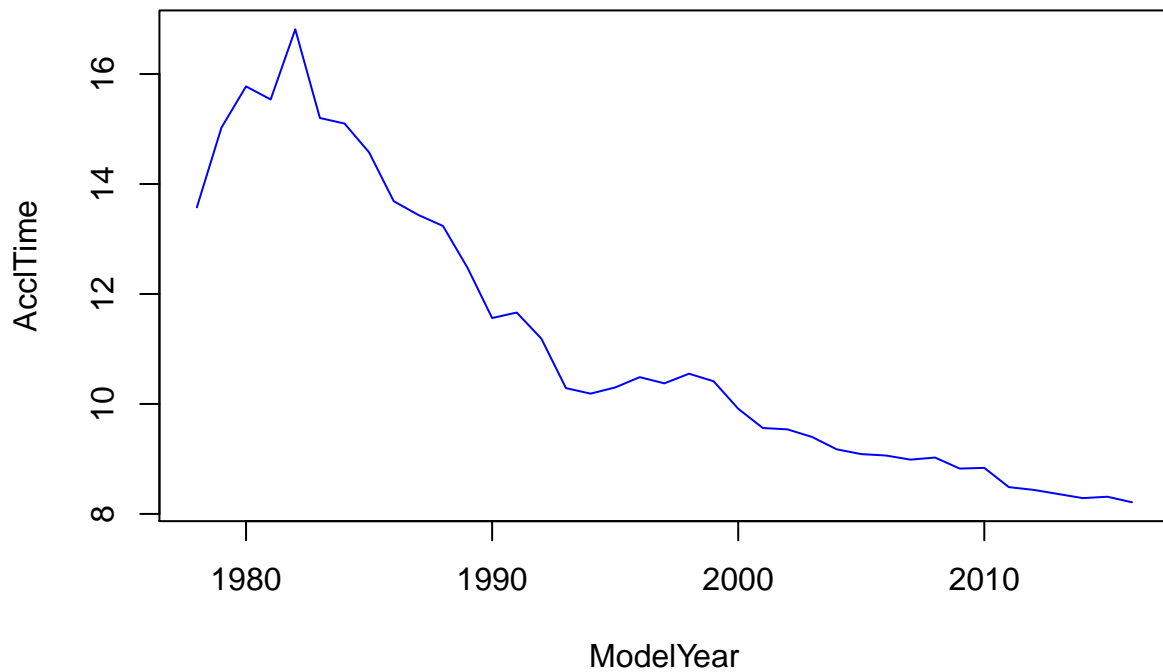



```
#ModelYear VS Acceleration
par(mfrow=c(1,1))
acclVStyp<-aggregate(economy$Acc.0.60.Time~economy$Model.Year+economy$Vehicle.Type,economy,mean)
colnames(acclVStyp)<-c("ModelYear","VehicleType","AcclTime")
ggplot(data = acclVStyp) +
  geom_point(mapping = aes(x = ModelYear, y = AcclTime,color=VehicleType))
```

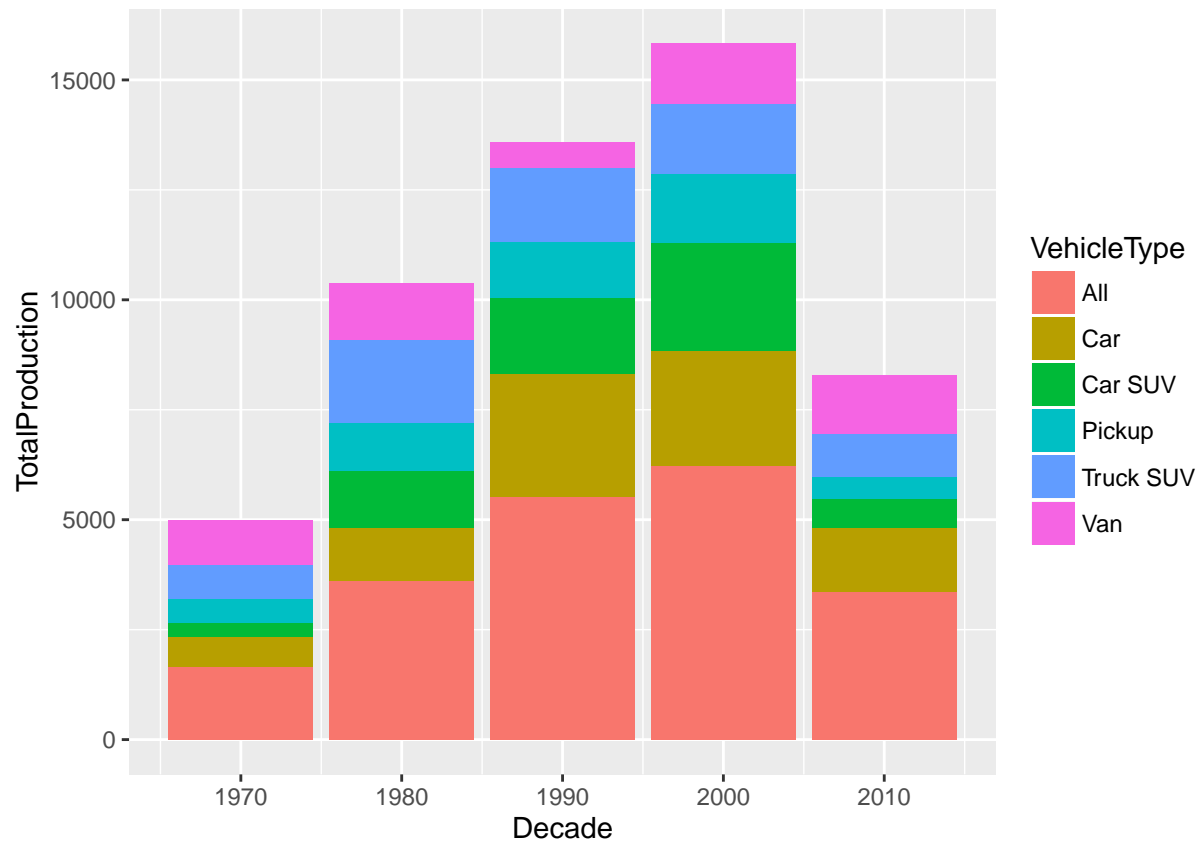


```
accl<-aggregate(economy$Acc.0.60.Time~economy$Model.Year,economy,mean)
colnames(accl)<-c("ModelYear","AcclTime")
plot(accl,type="l",main="Calculated 0-to-60 Acceleration Performance",col="blue")
```

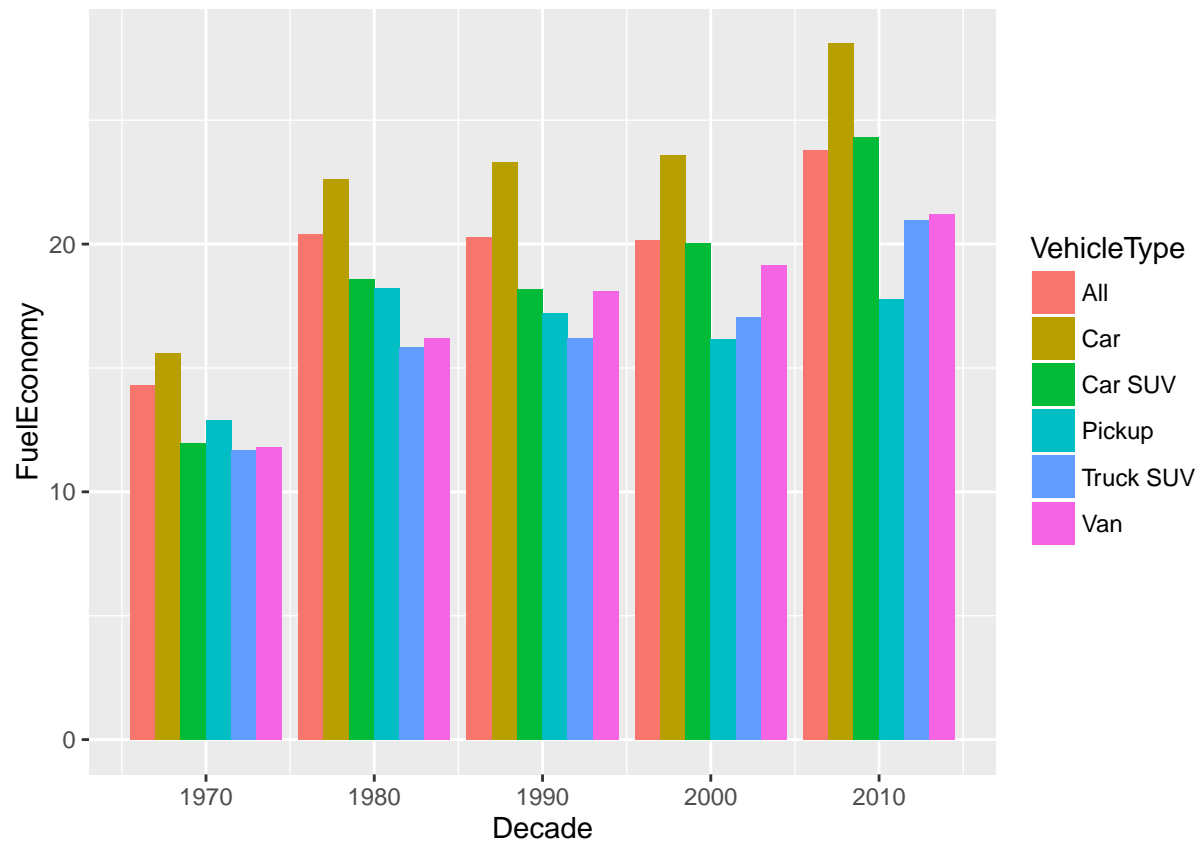
Calculated 0-to-60 Acceleration Performance



```
#Vehicle Choice  
#convert production into numeric data  
economy$Prod<-as.numeric(economy$Prod..000.)  
copy_eco<-economy  
  
#Floor Years into Decades  
copy_eco$Decade<-(copy_eco$Model.Year %/% 10) * 10  
prod<-aggregate(copy_eco$Prod~copy_eco$Decade+copy_eco$Vehicle.Type,copy_eco,sum)  
colnames(prod)<-c("Decade","VehicleType","TotalProduction")  
ggplot(data = prod, aes(x = Decade, y = TotalProduction, fill = VehicleType)) +  
  geom_bar(position = "stack",stat="identity")
```



```
#FuelEfficient VS Vehicle Type
typeco<-aggregate(copy_eco$FE.Adj.Comb~copy_eco$Decade+copy_eco$Vehicle.Type,copy_eco,mean)
colnames(typeco)<-c("Decade","VehicleType","FuelEconomy")
ggplot(data = typeco, aes(x = Decade, y = FuelEconomy, fill = VehicleType )) +
  geom_bar(position = "dodge",stat="identity")
```



```
prod<-aggregate(copy_eco$Prod~copy_eco$Decade+copy_eco$Vehicle.Type,copy_eco,sum)
colnames(prod)<-c("Decade","VehicleType","TotalProduction")
ggplot(data = prod, aes(x = Decade, y = TotalProduction, fill = VehicleType)) +
  geom_bar(position = "dodge",stat="identity")
```

