DATA 604 Project - Core Calcs

Getting ready for the data...starting the core calcs...

Kissock's Calcs...

All Acronyms:

- 3PC = 3 Param Cooling (ELECTRICITY)
- 3PH = 3 Param Heating (FUEL)
- Heating: Fi = Fuel Use INDEPENDENT (function of hot water, process heating, etc.), HS = Heat Slope, Tcph = Heating change-point, Toa = Temp Outdoor Air
- Cooling: Ei = Electric Use INDEPENDENT (function of lights, plug loads, etc.), CS = Cooling Slope, Tcpc = Cooling Change-point, Toa = Temp Outdoor Air

3pc is 3 param cooling (electricity), 3ph is 3 param heating (fuel)

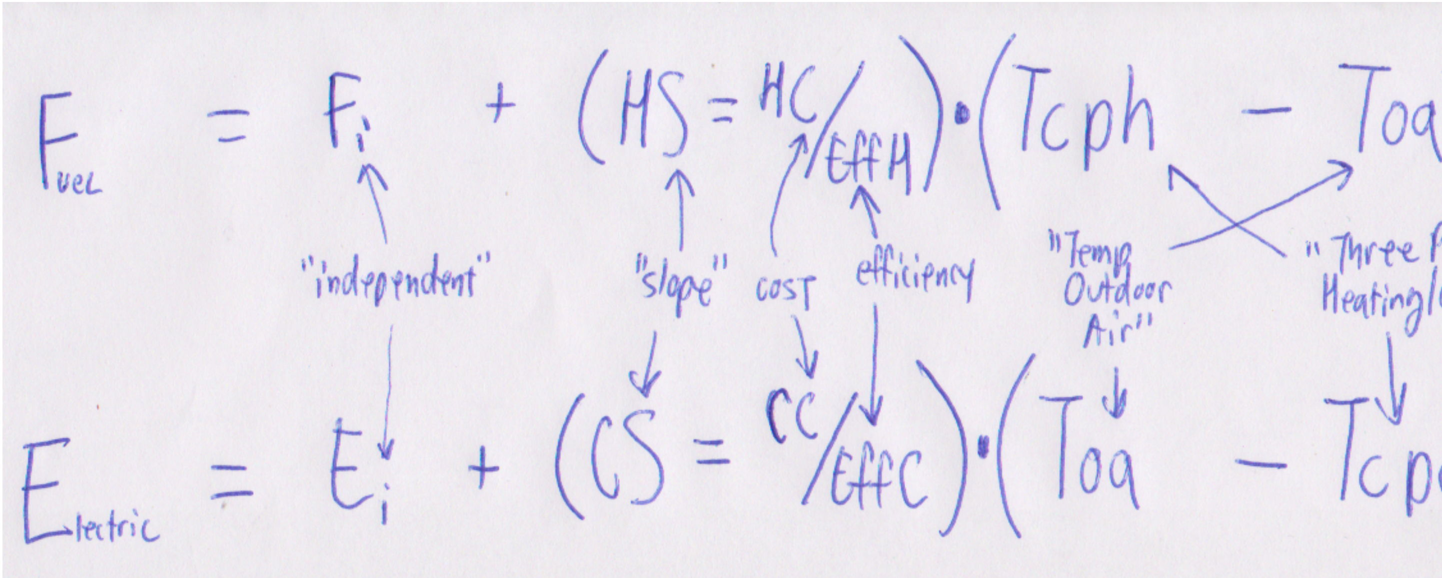
- HS and CS: heating and cooling slopes: heating/cooling energy per degree of temperature.
- Slopes function of:
- Envelope Heat Loss/Gain HC = CC UA + V p cp
- Efficiency of heating/cooling equipment: Effh and Effc
- HS = HC / Effh
- CS = CC / Effc
- TBH and TBC are BALANCED TEMPERATURES (temps where heating/cooling begins), its a function of
- TSP = Thermostat Set Point
- Qi = Internal Heat Gain
- TBH = Tsp - Qi/HC
- TBC = Tsp - Qi/CC

(We want to predict energy use based solely on 1) monthly time scale and 2) outdoor air temp)

Other (not used) acronyms: * Solar Radiation = S * Humidity = H * Wind Speed = W

NOTE: Important to normalize by square footage, so that you get energy per degree of temperature per square foot.

```r
# so ONLY toa param is in both funcs...

three_param_heating_3PH_fuel_model <- function(fuel_in_FI, heat_slope_HS, tcph, toa){
  the_fuel <- fuel_in_FI + heat_slope_HS * (tcph - toa)
}

three_param_cooling_3PC_electricity_model <-  function(elec_in_EI, cool_slope_CS, tcpc, toa){
  the_fuel <- fuel_in_FI + heat_slope_HS * (tcpc - toa)
}

knitr::include_graphics("basic_formula.png")
```



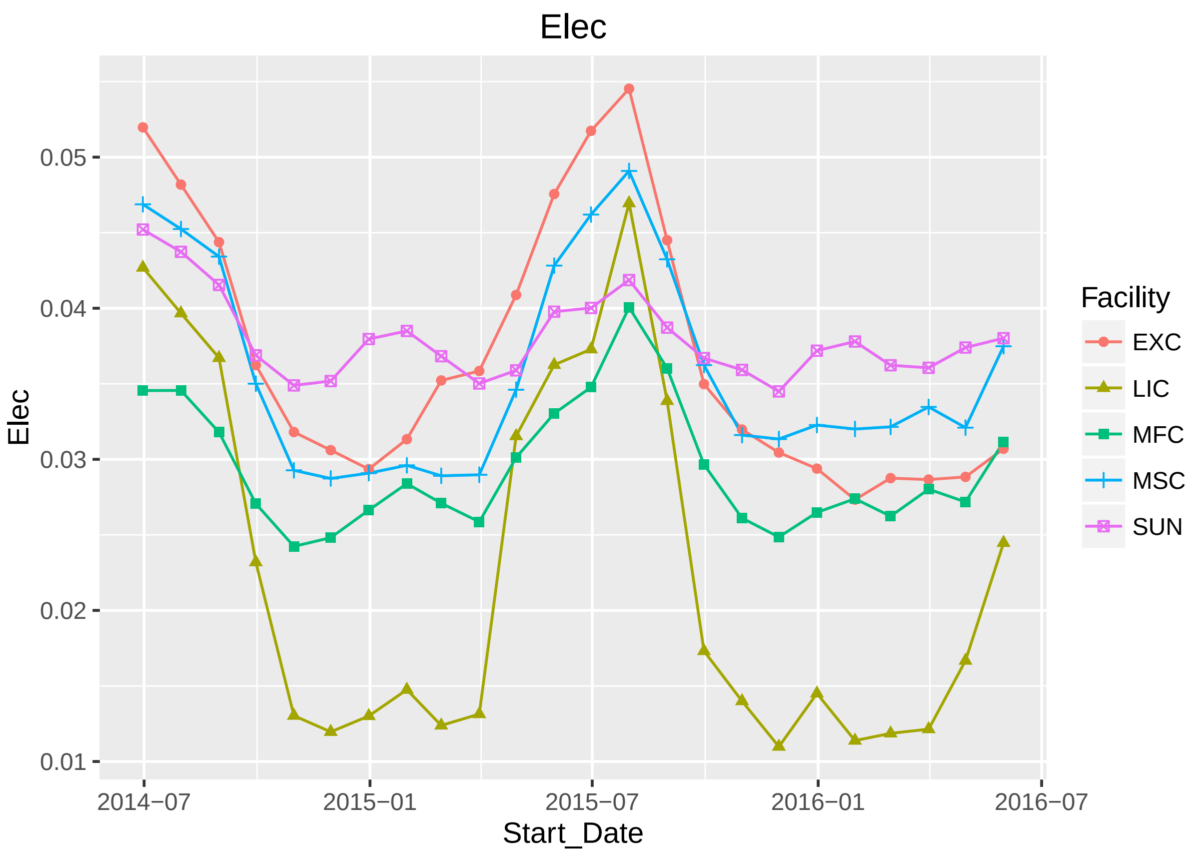
Try to apply to the data...

```
files <- c("EXC","LIC","MFC","MSC","SUN")

df_exc <- read.csv("data/EXC.csv", stringsAsFactors = FALSE)
df_exc$Facility <- "EXC"

df_lic <- read.csv("data/LIC.csv", stringsAsFactors = FALSE)
df_lic$Facility <- "LIC"

df_mfc <- read.csv("data/MFC.csv", stringsAsFactors = FALSE)
df_mfc$Facility <- "MFC"

df_msc <- read.csv("data/MSC.csv", stringsAsFactors = FALSE)
df_msc$Facility <- "MSC"

df_sun <- read.csv("data/SUN.csv", stringsAsFactors = FALSE)
df_sun$Facility <- "SUN"

df <- df_exc
df <- rbind(df, df_lic)
df <- rbind(df, df_mfc)
df <- rbind(df, df_msc)
df <- rbind(df, df_sun)

df$Start_Date <- as.Date(df$Start_Date, format = "%d-%b-%y")
df$End_Date <- as.Date(df$End_Date, format = "%d-%b-%y")

df$Facility <- as.factor(df$Facility)

df$MonthYear <- format(as.Date(df$Start_Date), "%Y-%m")
```

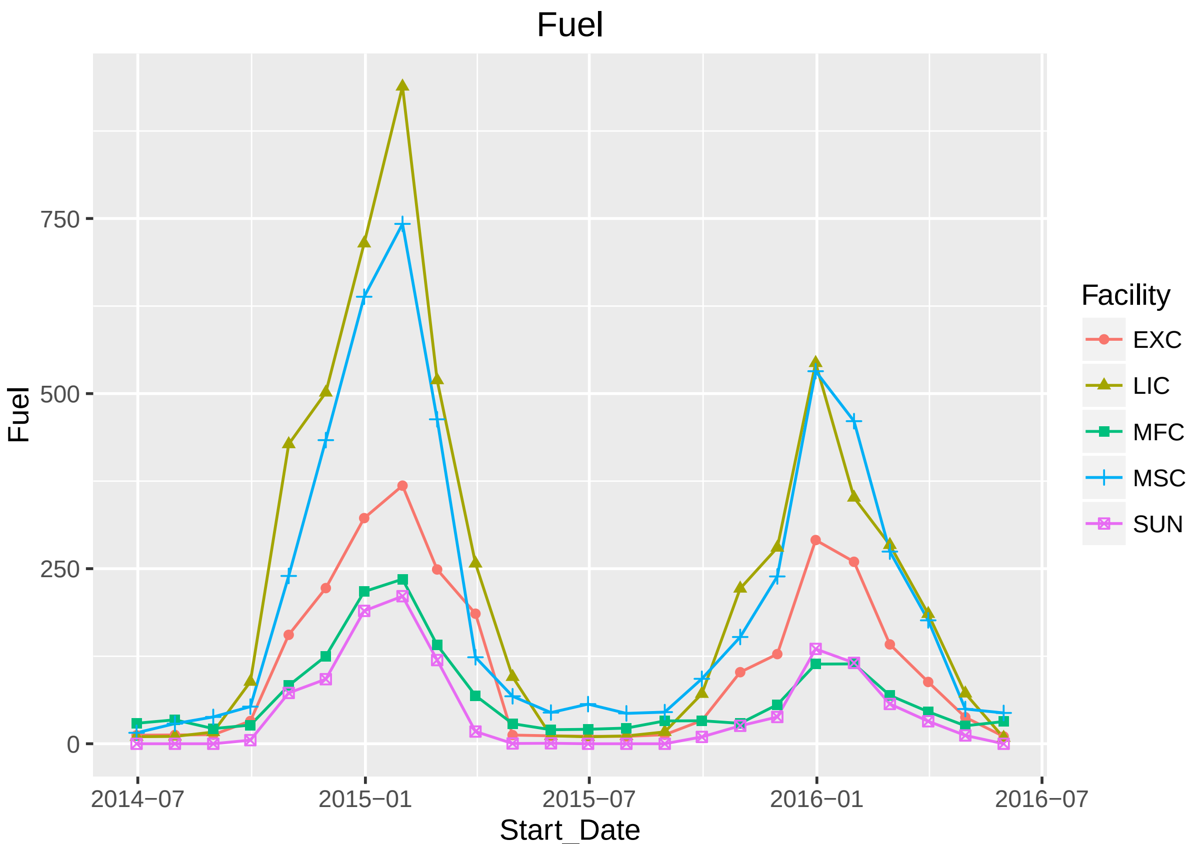
```r
kable(head(df, n=10))
```

Start_Date	End_Date	Elec	Fuel	Retrofit	Facility	MonthYear
2014-06-30	2014-07-31	0.0519721	12.17714	1	EXC	2014-06
2014-07-31	2014-08-31	0.0481829	12.57455	1	EXC	2014-07
2014-08-31	2014-09-30	0.0443738	12.92961	1	EXC	2014-08
2014-09-30	2014-10-31	0.0362404	32.44577	1	EXC	2014-09
2014-10-31	2014-11-30	0.0318080	155.52824	1	EXC	2014-10
2014-11-30	2014-12-31	0.0306075	222.26696	1	EXC	2014-11
2014-12-31	2015-01-31	0.0293494	322.16330	1	EXC	2014-12
2015-01-31	2015-02-28	0.0313302	368.56577	1	EXC	2015-01
2015-02-28	2015-03-31	0.0352210	248.82754	1	EXC	2015-02
2015-03-31	2015-04-30	0.0358475	185.69894	1	EXC	2015-03

```r
ggplot(data=df, aes(x=Start_Date, y=Elec, group=Facility, shape=Facility, color=Facility)) + ggtitle("El
```




```r
ggplot(data=df, aes(x=Start_Date, y=Fuel, group=Facility, shape=Facility, color=Facility)) + ggtitle("Fu
```



Floor area:

```
Facility <- c("MSC", "MFC", "EXC", "SUN", "LIC")
Gross.floor.area <- c(212500, 481000, 59000, 242062, 59300)
Floor.area <- cbind(Facility, Gross.floor.area) %>% as.data.frame()
Floor.area$Facility <- as.factor(Floor.area$Facility)

df <- left_join(df, Floor.area, by = "Facility")
```

Get temp data:

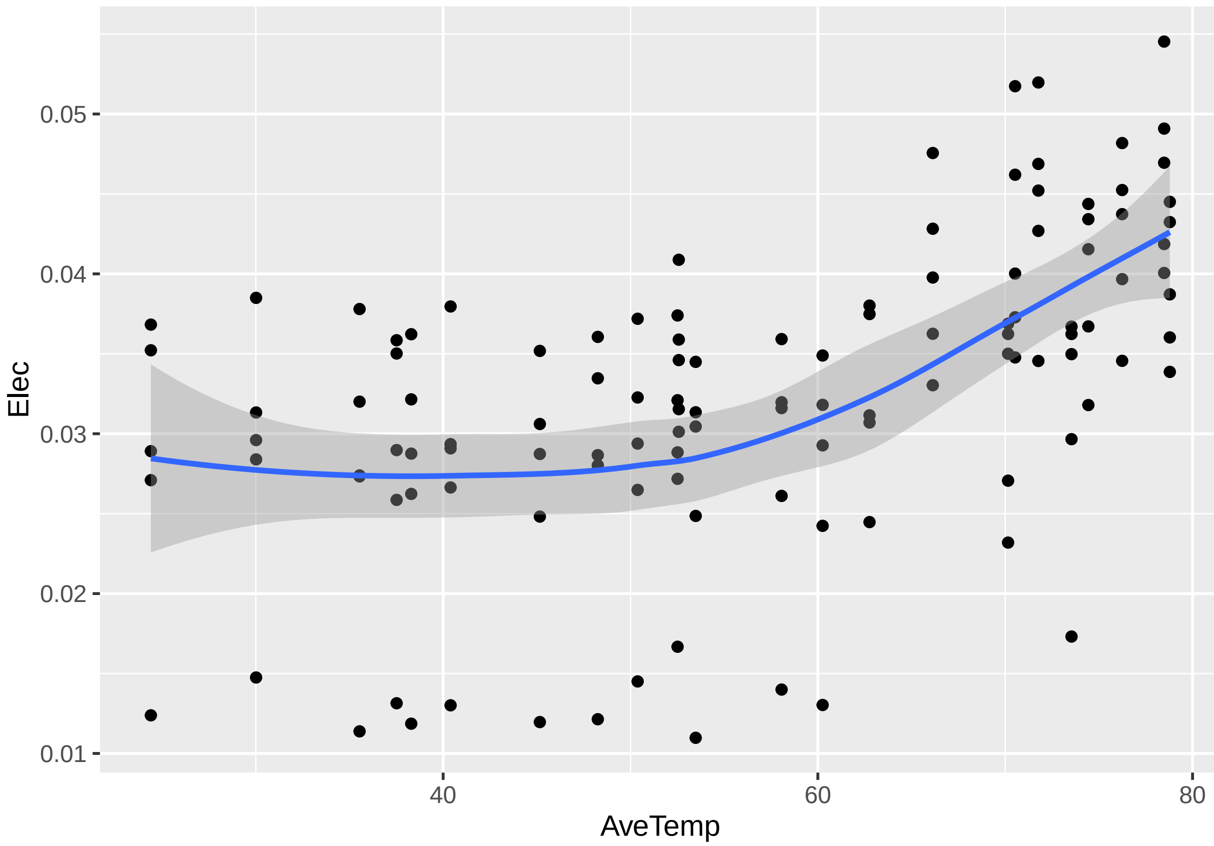
```
NYC_weather <- read.csv("data/NYNEWYOR.csv", header=FALSE)
colnames(NYC_weather) <- c("Month", "Day", "Year", "AveTemp")
NYC_weather$Month <- str_pad(NYC_weather$Month, 2, pad = "0")
```

```
NYC_month_ave <-aggregate(AveTemp ~ Month + Year, NYC_weather, mean)
NYC_month_ave$MonthYear = paste(NYC_month_ave$Year, "-", NYC_month_ave$Month, sep = "")
```

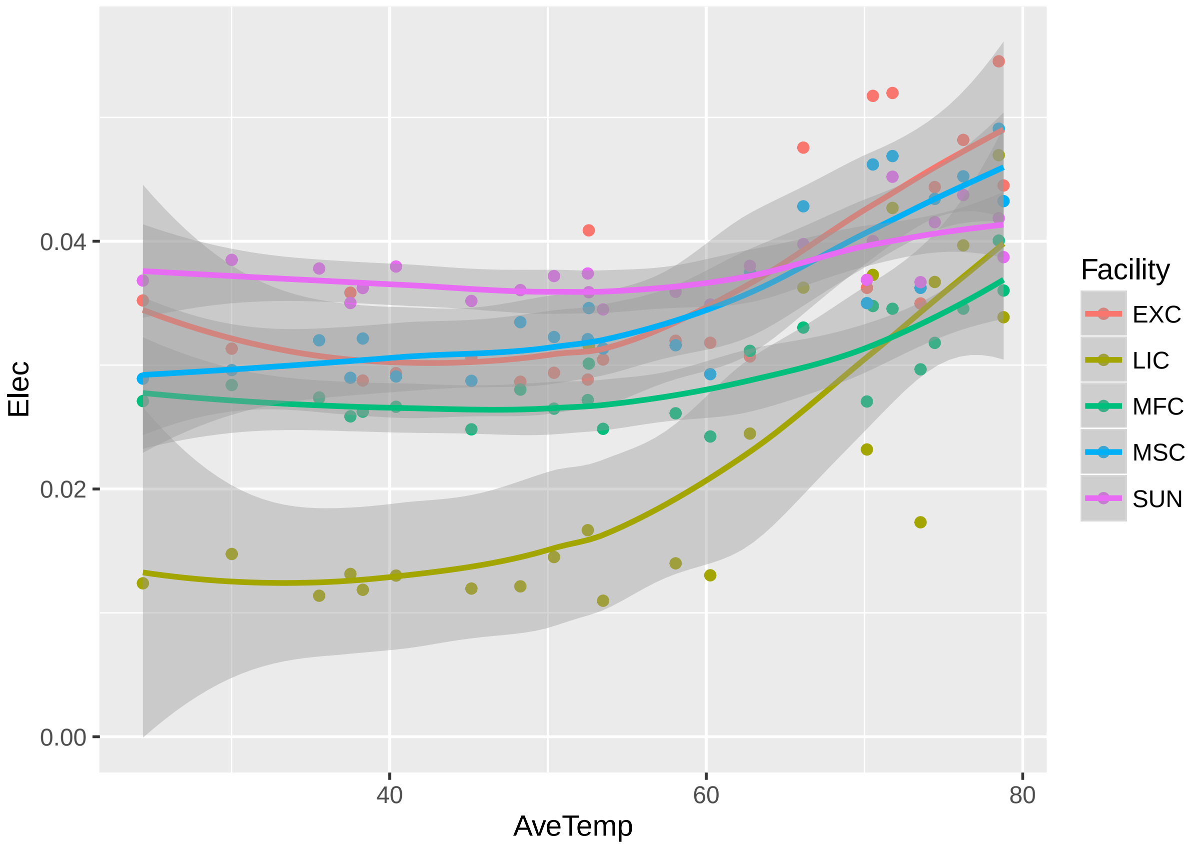
```
df <- left_join(df, NYC_month_ave, by = "MonthYear") %>%
  select(MonthYear, Elec, Fuel, Facility, Gross.floor.area, AveTemp)
```

EDA

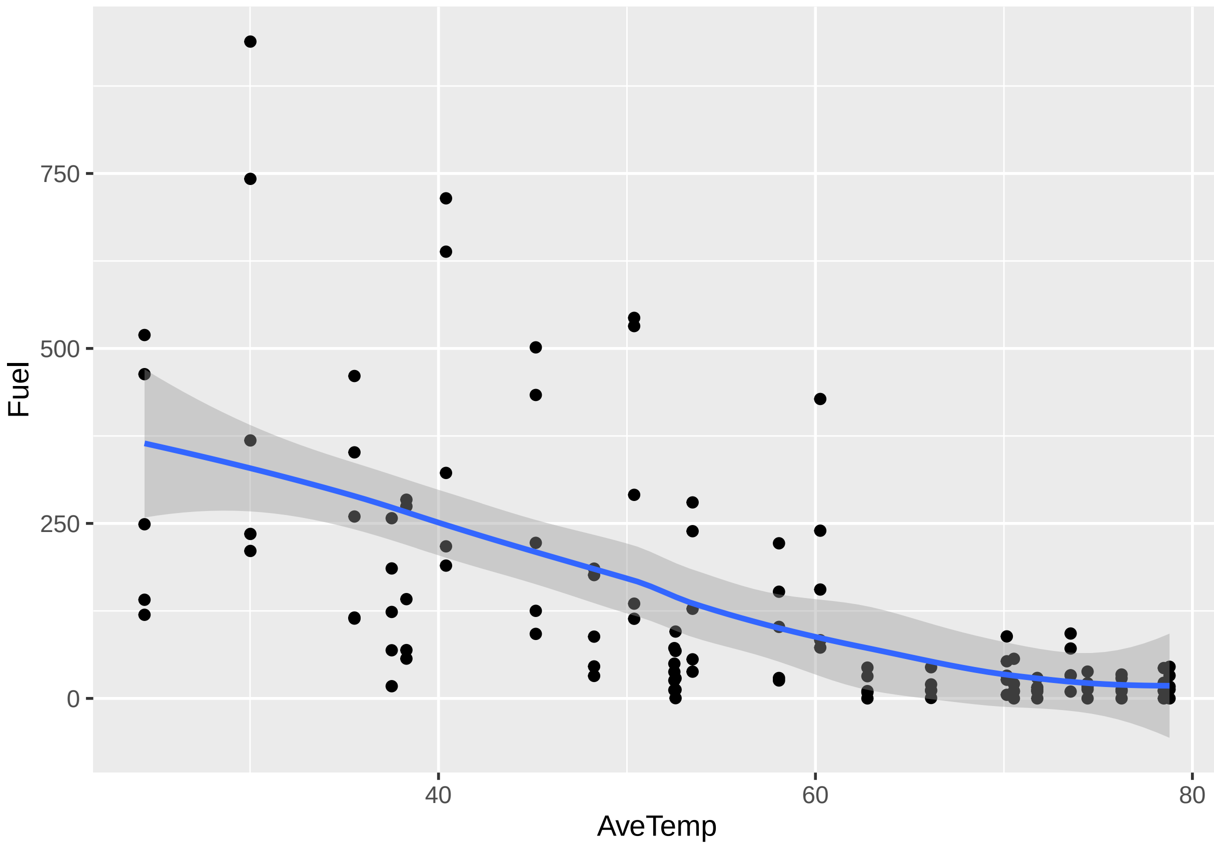
```
ggplot(df, aes(x = AveTemp, y = Elec)) + geom_point() + stat_smooth()
```



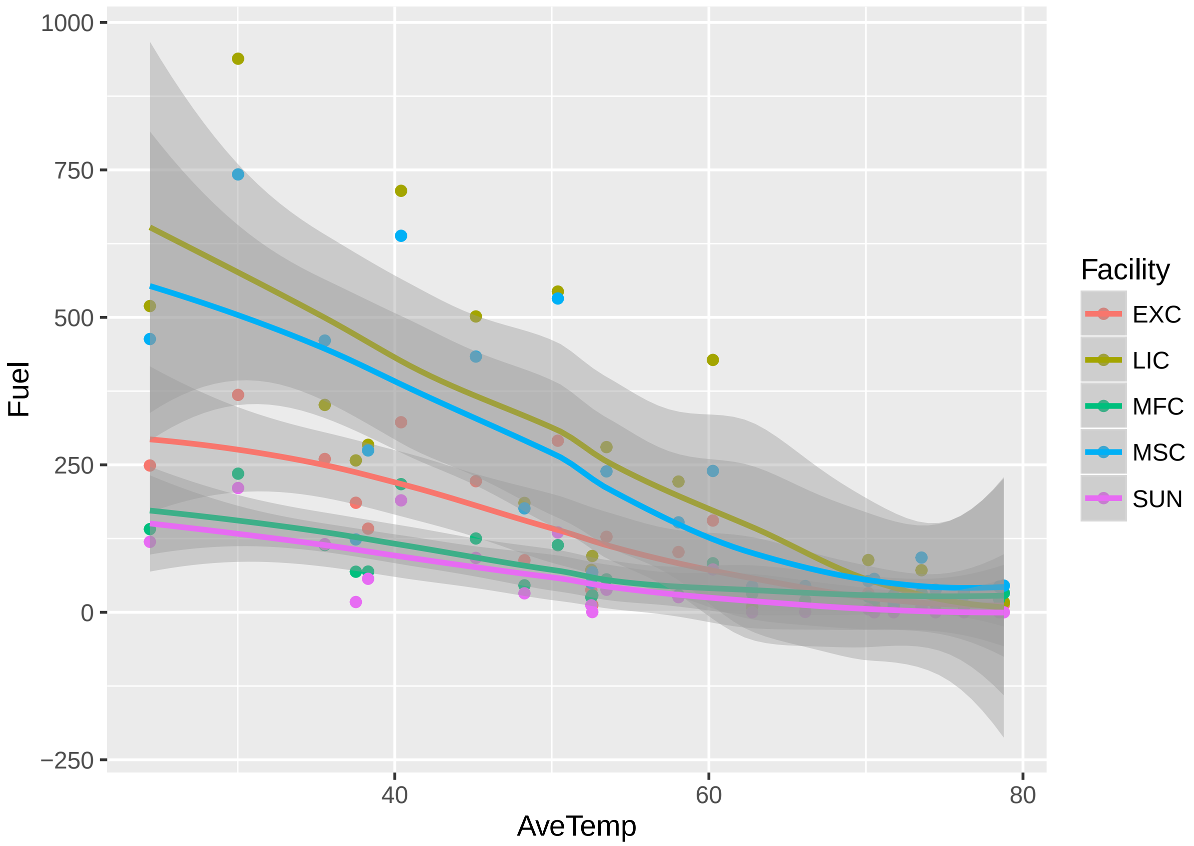

```
ggplot(df, aes(x = AveTemp, y = Elec, color = Facility)) + geom_point() + stat_smooth()
```




```
ggplot(df, aes(x = AveTemp, y = Fuel)) + geom_point() + stat_smooth()
```

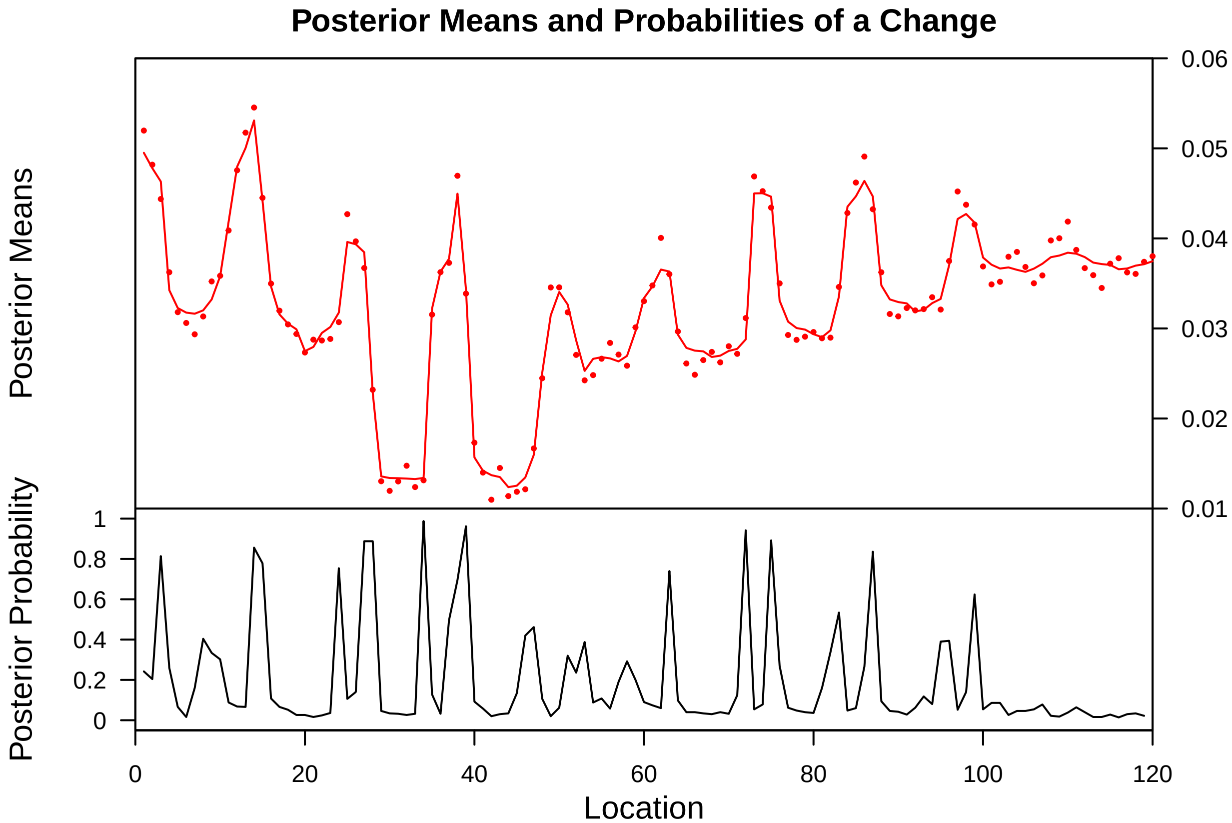



```
ggplot(df, aes(x = AveTemp, y = Fuel, color = Facility)) + geom_point() + stat_smooth()
```

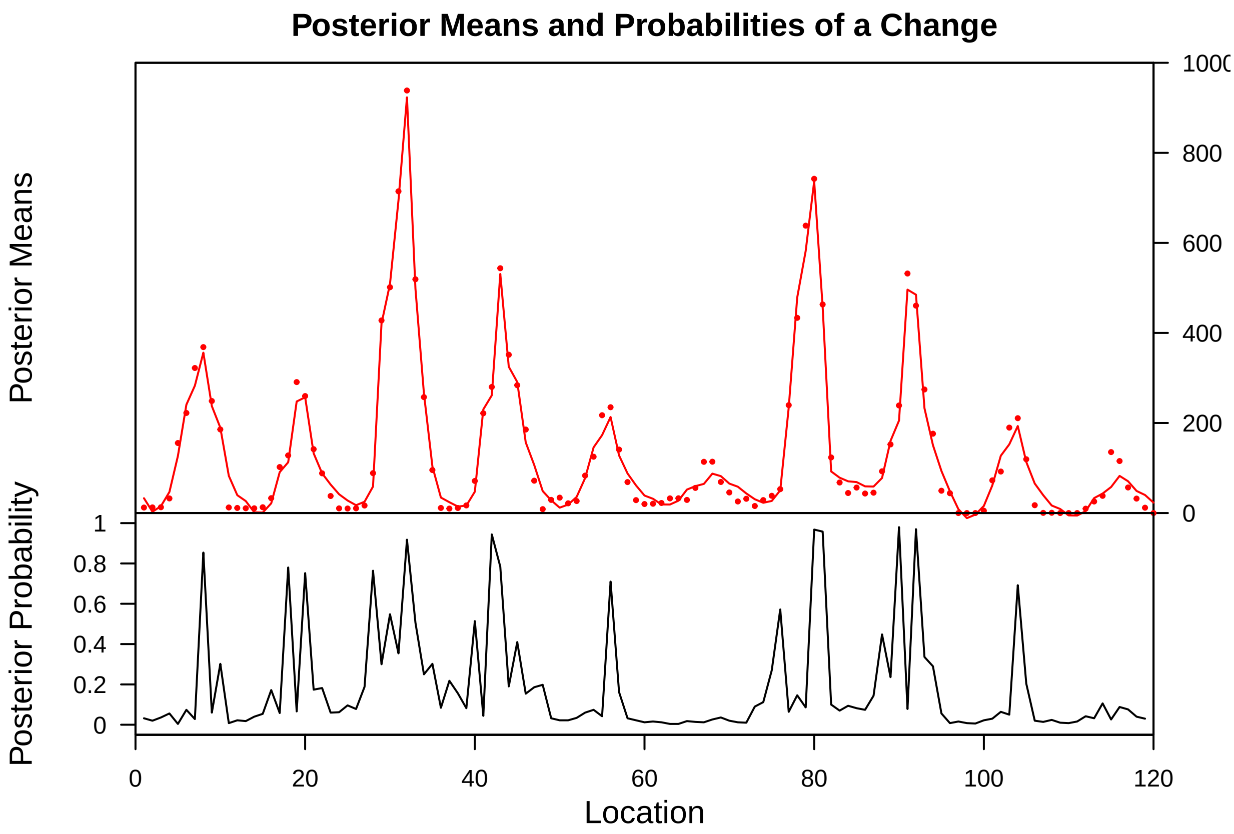


Changepoint Analysis

```
elec.bcp <- bcp(y = df$Elec, x = df$AveTemp)
plot(elec.bcp)
```




```
fuel.bcp <- bcp(y = df$Fuel, x = df$AveTemp)
plot(fuel.bcp)
```


Posterior Means and Probabilities of a Change
Posterior Means
Posterior Probability
Location
0
20
40
60
80
100
120
1
0.8
0.6
0.4
0.2
0
1000
800
600
400
200
0