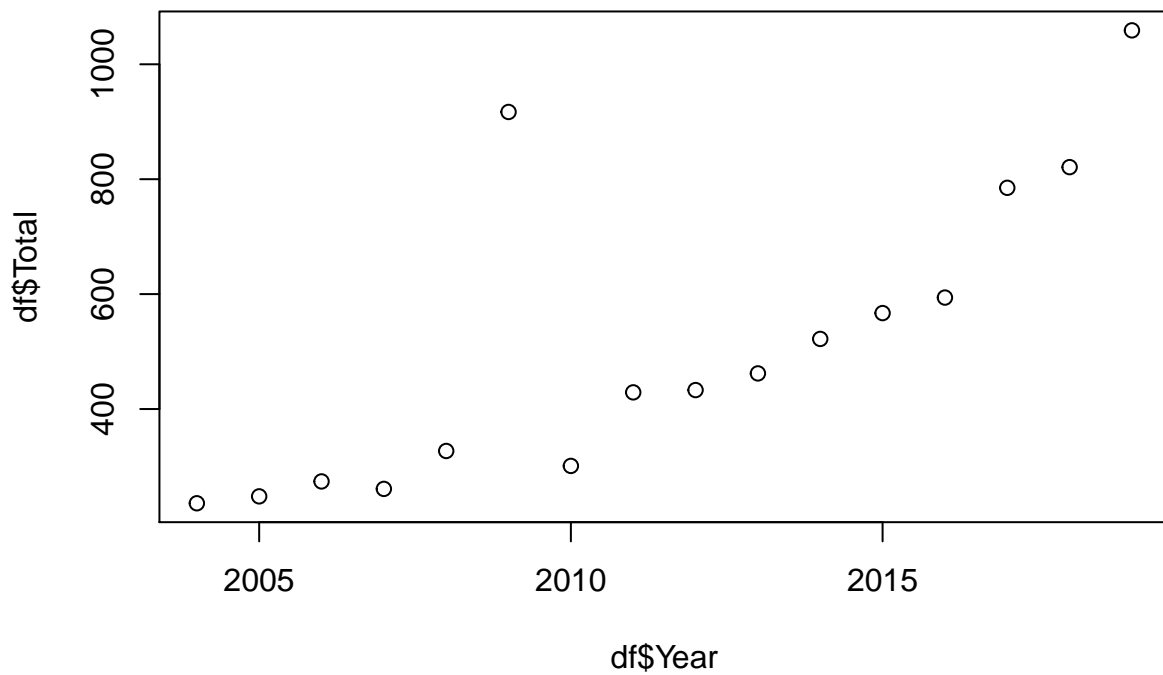# Question 3

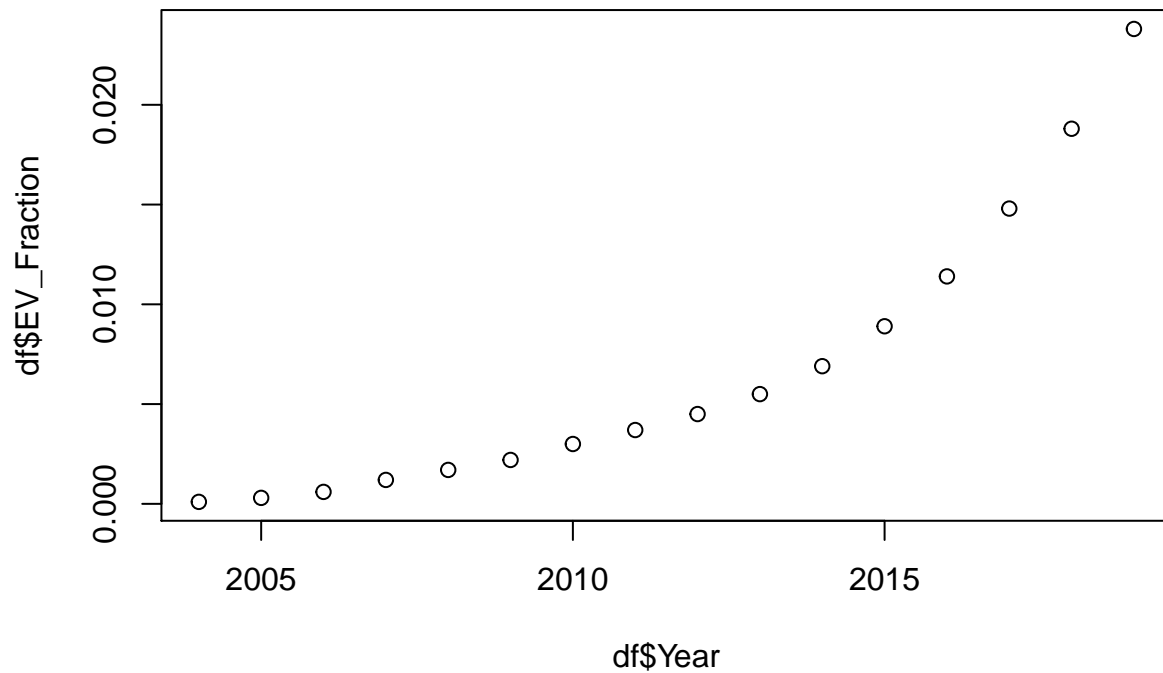## 1. Load the data and plot Google trend and EV fraction trend

```r
df <- read.csv("UK_pivotGoogle.csv")
colnames(df)[1] <- "Year"
df$Year <- as.character(df$Year)

# create function to normalize data
normalize <- function(x) {
  return ((x - min(x)) / (max(x) - min(x)))
}


#df <- df %>% mutate_if(is.numeric, normalize)

plot(df$Year, df$Total)
```

```r
plot(df$Year, df$EV_Fraction)
```


The goal of this is to show that as Google Trends increases so does EV Fraction. The data set only has up to 2019 cause in 2020 google trends wasn't as high (feel free to change).

2. Check correlation Ev_Fraction ~ Google Trend

```r
# Total predicts EV_Fraction
m1 <- lm(EV_Fraction ~ Total -1, df)
m1$coefficients
```

```
##        Total
## 1.484171e-05
```

```r
# Seem very correlated
cor(df[-1])
```

```
##                Total EV_Fraction
## Total      1.0000000   0.8170563
## EV_Fraction 0.8170563   1.0000000
```
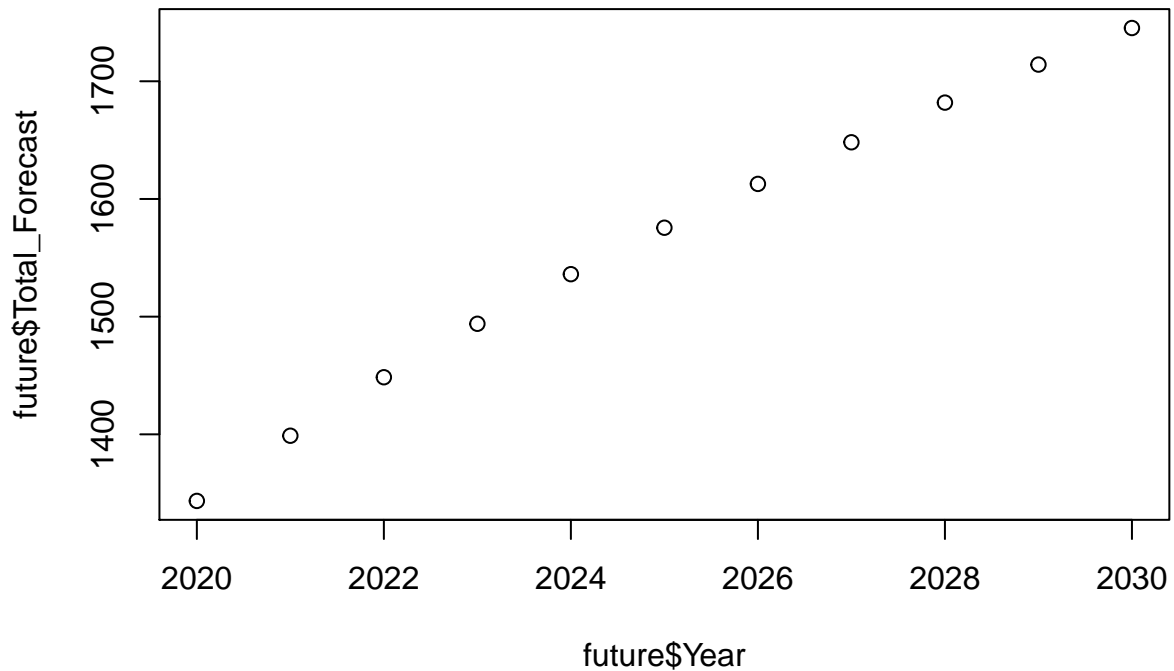
3. Forecast and then plot forecast

```
# type of forecast is not good so try other type
a1 <- forecast(df$Total, 11)


# DF for results
future <- NULL
future$Year <- c(2020:2030)
future <- data.frame(future)
future$Total_Forecast <- a1$upper[,2]

future$coef <- m1$coefficients
future <- future %>%
  mutate(Future_EV_fraction = Total_Forecast * coef)

# Plot the future
plot(future$Year, future$Total_Forecast)
```
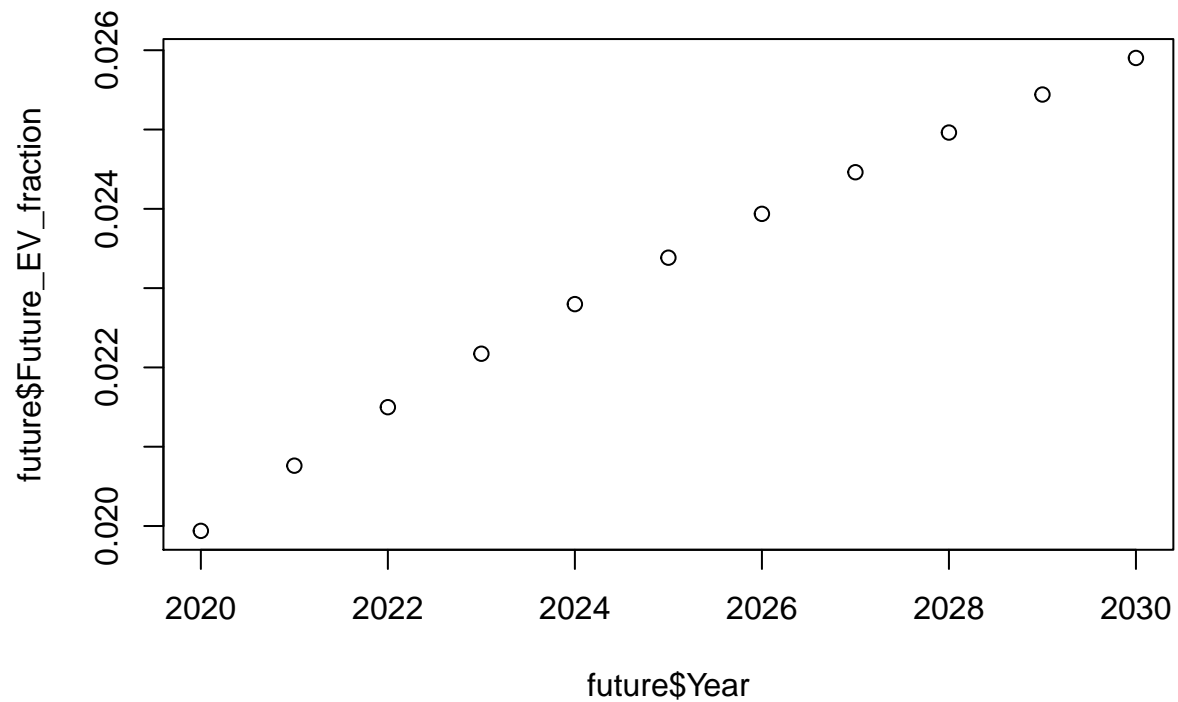
```r
plot(future$Year, future$Future_EV_fraction)
```


4. This is what you put into excel

```r
print(future$Future_EV_fraction)
```

```
## Time Series:
## Start = 17
## End = 27
## Frequency = 1
##  [1] 0.01993891 0.02076098 0.02149840 0.02217292 0.02279834 0.02338403
##  [7] 0.02393672 0.02446143 0.02496202 0.02544155 0.02590246
```