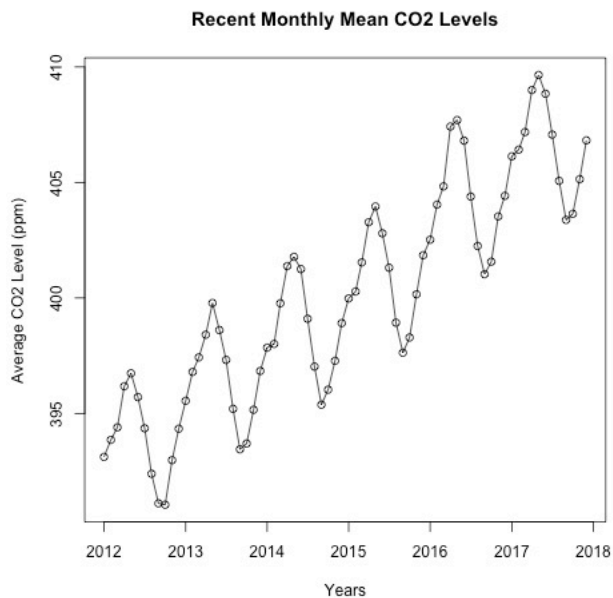# The Dataset

We are using the monthly mean Co2 ppm from Mauna Loa. This dataset

Analysis

We explore the dataset using triple exponential smoothing, ARIMA, and using Gaussian Processes.

```
df = read.table("Downloads/co2_mm_mlo.txt")
names(df) = c("Year", "Month", "Date", "average", "interpolated", "trend", "days")
df$Dates = as.Date(paste(df$Year, df$Month, 1, sep = "-"), format = "%Y-%m-%d")
plot(df[,8], df[,5], type = "l", xlab = "Year", ylab = "Yearly Average",
     main = "Mauna Loa Yearly Carbon Dioxide Concentration")
plot(1958:2018, tapply(df[,5], format(df[,8], "%Y"), mean), xlab = "Year",
     ylab = "Yearly Average", main = "Mean Annual CO2 Levels", type = "l")
plot(df[647:718, 8], df[647:718, 4], type = "o", xlab = "Years",
     ylab = "Average CO2 Level (ppm)", main = "Recent Monthly Mean CO2 Levels")

library(stats, forecast)
shortts = ts(df[647:718, 4], frequency = 12, start = c(2012, 1))
plot(HoltWinters(shortts))
MAPE(HoltWinters(shortts)$fitted[,1], df[659:718, 4])

fullts = ts(df[1:718,5], frequency = 12, start = c(1958, 3))
plot(HoltWinters(fullts))
MAPE(HoltWinters(fullts)$fitted[,1], df[13:718,5])

pred = data.frame(predict(stats::HoltWinters(fullts),
```

```
                                      n.ahead = 8, prediction.interval = 1))
pred$date = seq(as.Date("2018-01-01"), by = "month", length.out = 8)

plot(pred[,4], pred[,1], type = "o", main = "Predicted 2018 CO2 Levels",
     xlab = "Month of 2018", ylab = "Co2 Levels (ppm)", ylim = c(405, 414))

with(pred, polygon(c(date, rev(date)), c(lwr, rev(upr)), col = "grey", border = 1))
matlines(pred[,4], pred[,-4], lwd = c(2,1,1), lty = c(1,2,2),)
MAPE(forecasted[,1], df[718:725,5])
```