

In Class Exercise 2

Joshua L Eubanks

Contents

1	Wealth of 1 Percent	1
1.1	Time Series Plot	1
1.2	Time Series Regression	2
1.3	Prediction	3
2	Covid cases NYT	3
2.1	Plot	3
2.2	Regression	4
2.3	Prediction	4
2.4	Thing 2	4

1 Wealth of 1 Percent

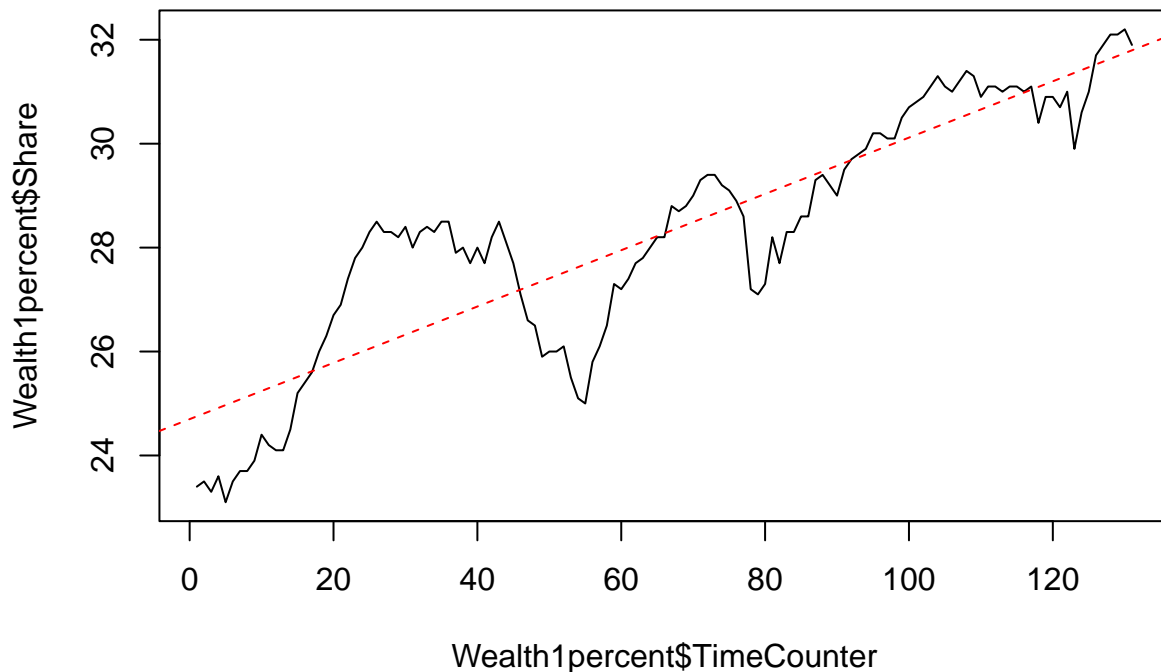
This dataset was obtained from FRED which shows the share of total net worth held by the top 1%. It is represented over time.

```
Wealth1percent <- read_excel("../Data/Wealth1percent.xlsx",  
col_types = c("date", "numeric", "numeric"))
```

This gives us a range of data from the third quarter of 1989 to the first quarter of 2021

1.1 Time Series Plot

```
ts_model <- lm(Share ~ TimeCounter, data = Wealth1percent)  
  
plot(Wealth1percent$TimeCounter, Wealth1percent$Share, type = "l")  
abline(ts_model, lty = 2, col = "red")
```



1.2 Time Series Regression

We can see the trend in an upwards manner, let's see what the quarterly change in slope is:

```
summary(ts_model)
```

```
##
## Call:
## lm(formula = Share ~ TimeCounter, data = Wealth1percent)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.67883 -0.80840  0.08339  0.69964  2.39238
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 24.698943   0.197183   125.3   <2e-16 ***
## TimeCounter  0.054180   0.002592    20.9   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.122 on 129 degrees of freedom
## Multiple R-squared:  0.772, Adjusted R-squared:  0.7702
## F-statistic: 436.8 on 1 and 129 DF, p-value: < 2.2e-16
```

in this case, the share of wealth held by the top 1% is increasing by 0.05% each quarter.

1.3 Prediction

Suppose we want to predict the share of wealth 2 quarters from now:

```
newForecast <- data.frame(TimeCounter = max(Wealth1percent$TimeCounter) + 2)

predict(ts_model, newdata = newForecast)

##          1
## 31.90486
```

2 Covid cases NYT

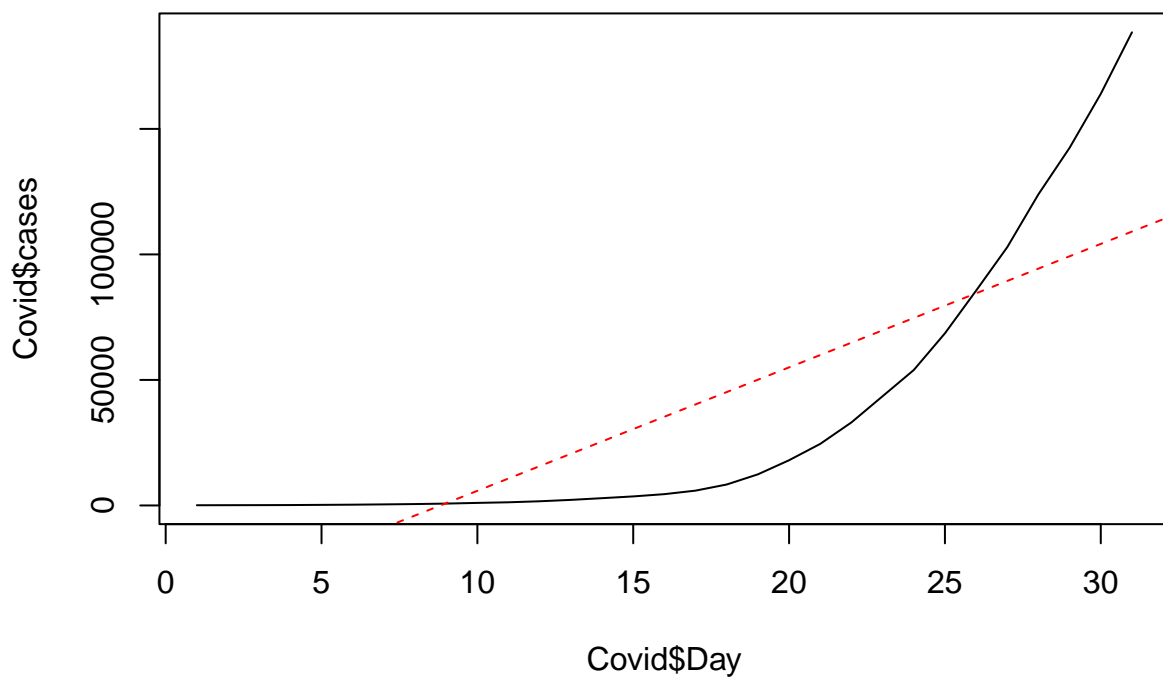
This data was pulled for March 2020 for covid cases in the US

```
Covid <- read_excel("../Data/Covid.xlsx")
```

2.1 Plot

```
ts_model <- lm(cases ~ Day, data = Covid)

plot(Covid$Day, Covid$cases, type = "l")
abline(ts_model, lty = 2, col = "red")
```



2.2 Regression

```
summary(ts_model)

##
## Call:
## lm(formula = cases ~ Day, data = Covid)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -37699 -26540  -4817   21406   79359
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -43339.7     11665.6  -3.715 0.000862 ***
## Day          4917.5       636.4    7.727 1.61e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 31690 on 29 degrees of freedom
## Multiple R-squared:  0.6731, Adjusted R-squared:  0.6618
## F-statistic: 59.7 on 1 and 29 DF,  p-value: 1.607e-08
```

2.3 Prediction

What would be the predicted number of cases by April 15, 2020?

```
newForecast <- data.frame(Day = 46)

predict(ts_model, newdata = newForecast, level = "confidence")

##           1
## 182863.7
```

2.4 Thing 2

On average, how many new cases of covid are there daily?

```
ts_model$coefficients[2]

##      Day
## 4917.466
```