

# STAT 443: Lab 1

Aronn Grant Laurel Y.

10 January, 2025

## Question 1

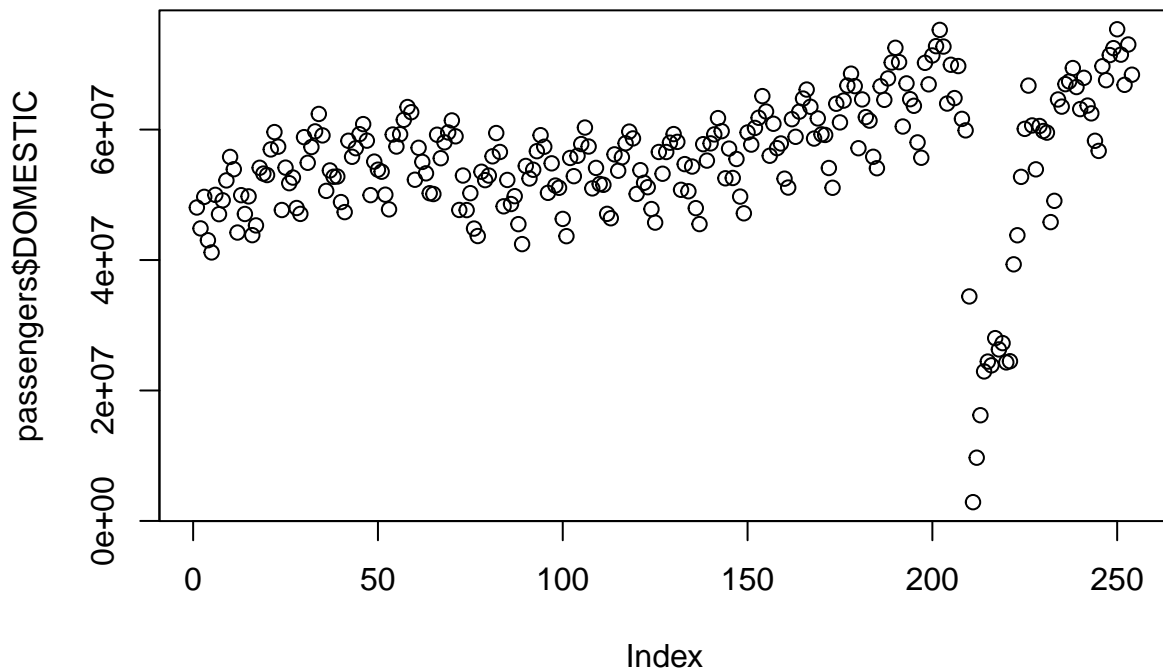
(a)

```
# this is where your R code goes
```

```
passengers <- read.csv("dat_Passengers.csv")
```

```
# head(passengers, 5)
```

```
plot(passengers$DOMESTIC)
```



Using the plot function on the DOMESTIC column, we see that it plots it against the index.

Ideally, I would like to have visualisation of Domestic Flights plotted against the Year or Month instead of the Index. Furthermore, I would extend the y-axis more as the range is a lot higher and I would prefer smaller points such that we have less overlapping data points.

(b)

```
# this is where your R code goes  
is.ts(passengers)
```

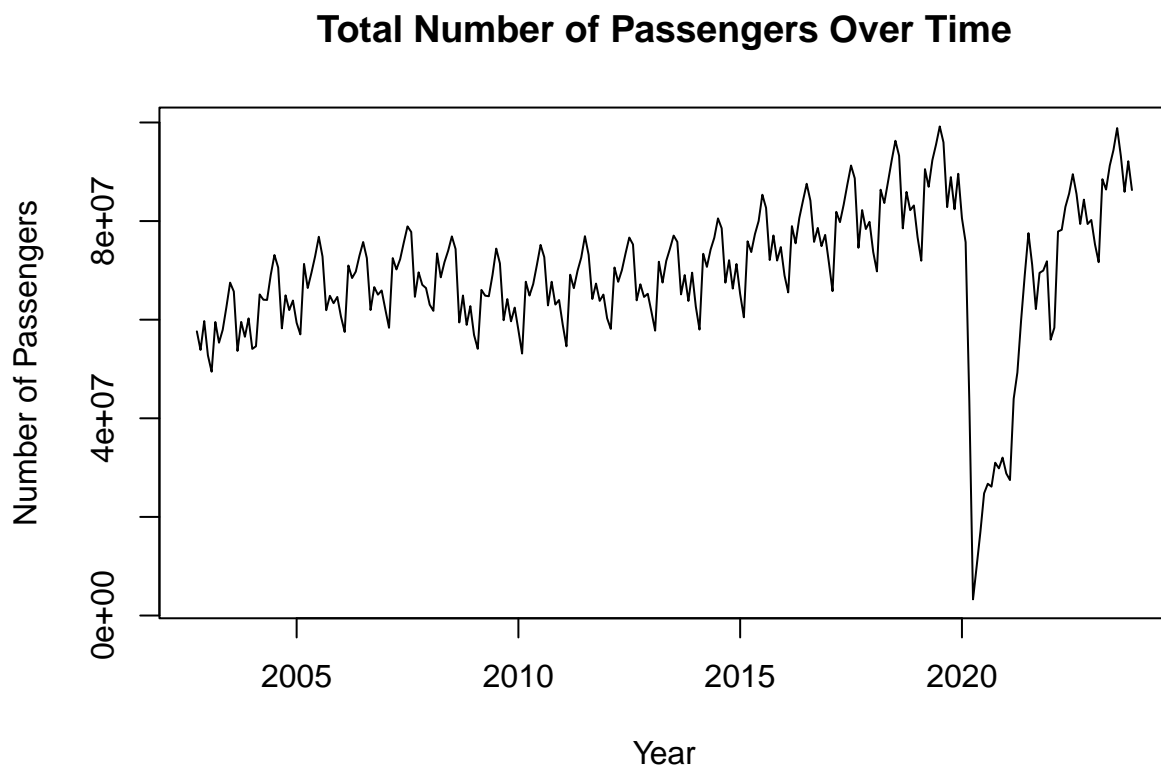
```
## [1] FALSE
```

```
x <- ts(passengers$TOTAL,  
        start = c(2002, 10),  
        frequency = 12)  
  
is.ts(x)
```

```
## [1] TRUE
```

(c)

```
# this is where your R code goes  
plot(x,  
      main = "Total Number of Passengers Over Time",  
      xlab = "Year",  
      ylab = "Number of Passengers")
```



If we plot time series object, we can see that `plot()` will use a line graph and it extends its y-axis range more. Now, we can easily identify if there's a clear trend, if there's seasonal variation.

## Question 2

(a)

```
# this is where your R code goes
set.seed(100)

x2 <- rnorm(500)

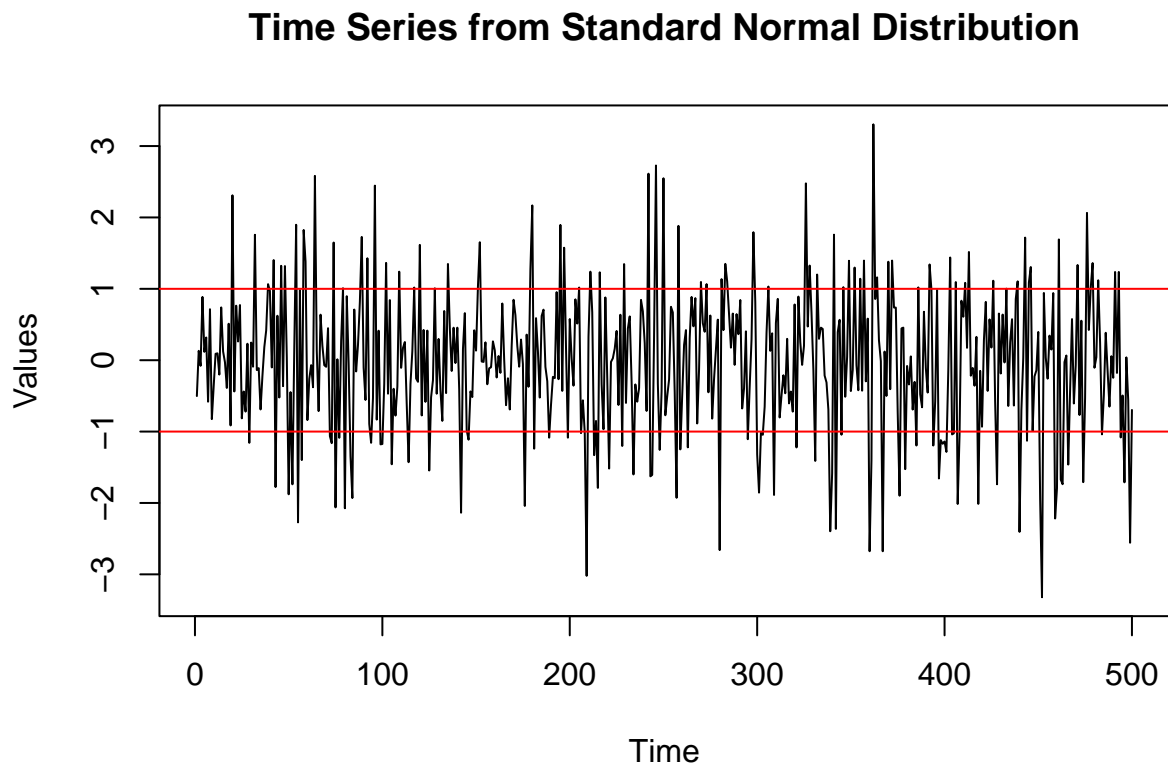
x2_ts <- ts(x2)
is.ts(x2_ts)
```

```
## [1] TRUE
```

(b)

```
# this is where your R code goes
plot(x2_ts,
      main = "Time Series from Standard Normal Distribution",
      ylab = "Values",
      xlab = "Time")

# Add lines + / - 1
abline(h = 1, col = "red")
abline(h = -1, col = "red")
```



```
# Expected outside + / - 1 range
2* (1 - pnorm(1)) * 500
```

```
## [1] 158.6553
```

```
sum(abs(x2_ts) > 1)
```

```
## [1] 159
```

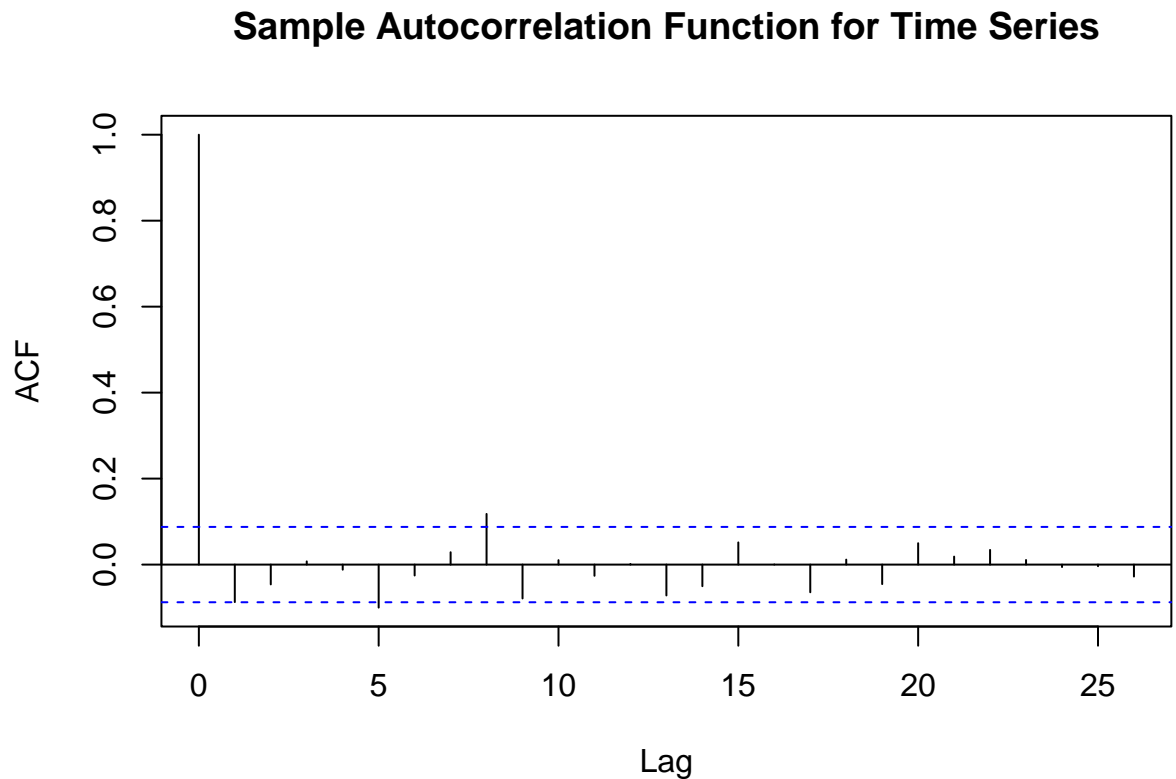
```
159/500 # Or 2* (1 - pnorm(1))
```

```
## [1] 0.318
```

I would expect 31.8% or 159 out of the 500 observsation to be outside of the + / - 1 range

(c)

```
# this is where your R code goes
acf(x2_ts, main="Sample Autocorrelation Function for Time Series")
```



The lag 0 autocorrelation will be 1, with majority of all the other lag values close to value 0 and staying between the confidence boundary line. We can say lag 0, 5, and 8 is outside the boundary line, therefore statistically significant.

## More information on R Markdown

This is an R Markdown document, which can be used as a template for STAT 443 labs and assignments. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean    : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.    :120.00
```

Using the function *kable*, it produces a nicer table

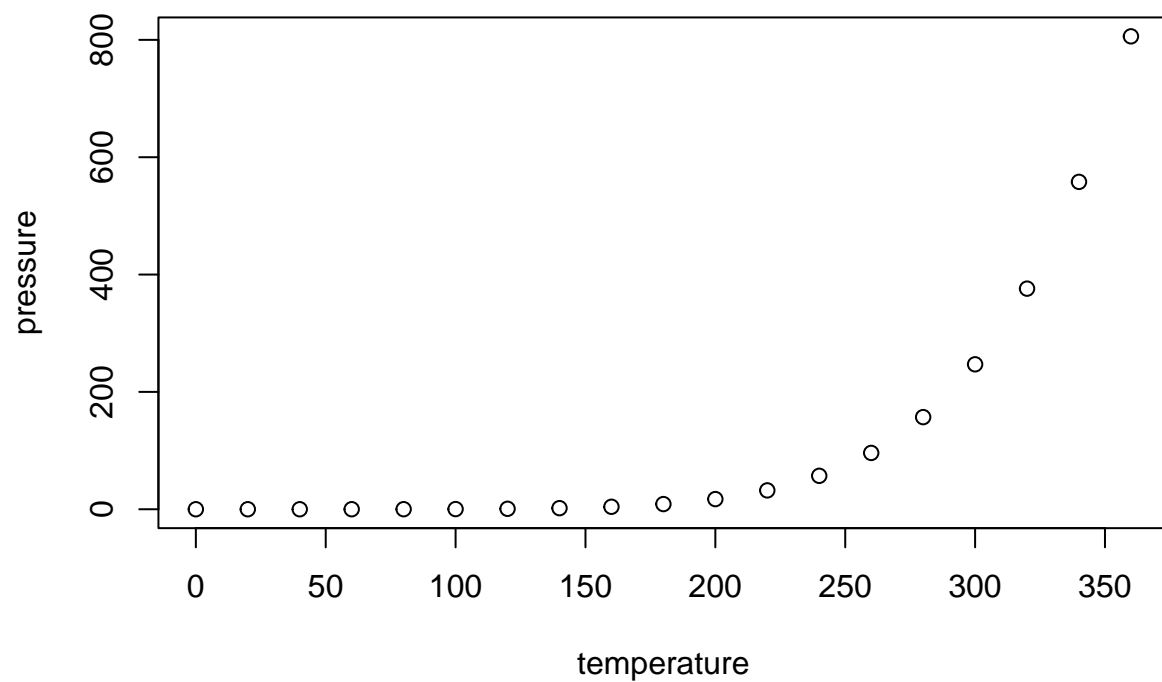
```
kable(summary(cars))
```

speed	dist
Min. : 4.0	Min. : 2.00
1st Qu.:12.0	1st Qu.: 26.00
Median :15.0	Median : 36.00
Mean :15.4	Mean : 42.98
3rd Qu.:19.0	3rd Qu.: 56.00
Max. :25.0	Max. :120.00

## Including Plots

You can also embed plots, for example:

```
plot(pressure)
```



Note that specifying `echo = FALSE` parameter would prevent printing of the R code that generated the plot. This is something you may want to do for larger reports that would not require display of the R code.

You can also modify the size and alignment of the figure.

```
plot(pressure)
```

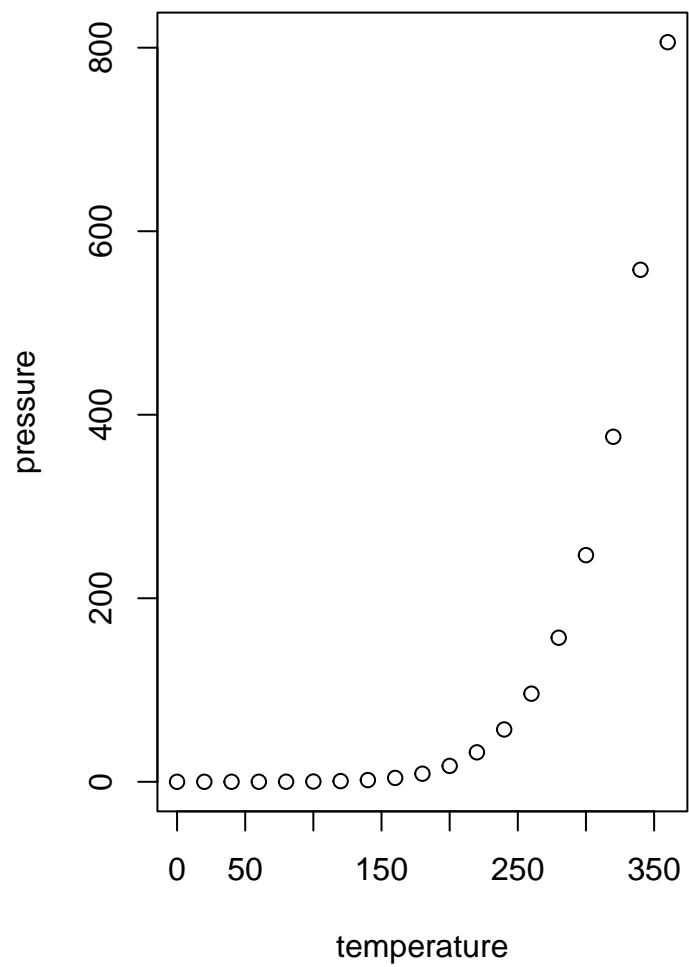


Figure 1: title