DSC 534–Lab 1

Yiolanda Englezou

Quarterly Earnings of J & J

Quarterly earnings per share for the U.S. company Johnson & Johnson. There are 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980. One of the most important steps in time series analysis is to visualize the data, i.e. create a time series plot, where the quarterly earnings are ploted against time.

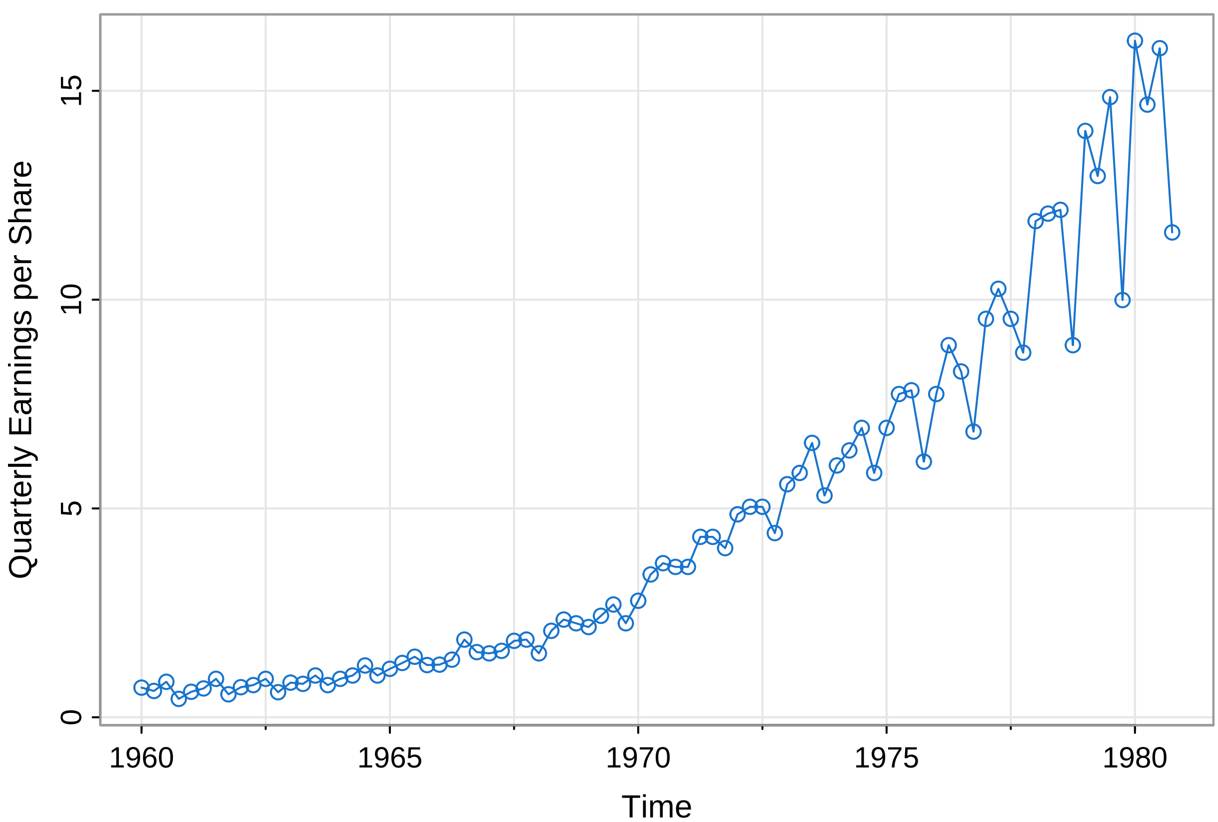
Main observations from the plot:

1. Gradually increasing trend. In general, a systematic change in the mean level of a time series that does not appear to be periodic is known as a trend. The simplest model for a trend is a linear increase or decrease.
2. Regular variation superimposed on the trend that appears to be repeated over quarters. This is known as a seasonal effect, or seasonality.

```r
library(astsa)
```

```
## Warning: package 'astsa' was built under R version 4.3.3
```

```r
tsplot(jj, col=4, type="o", ylab="Quarterly Earnings per Share")
```



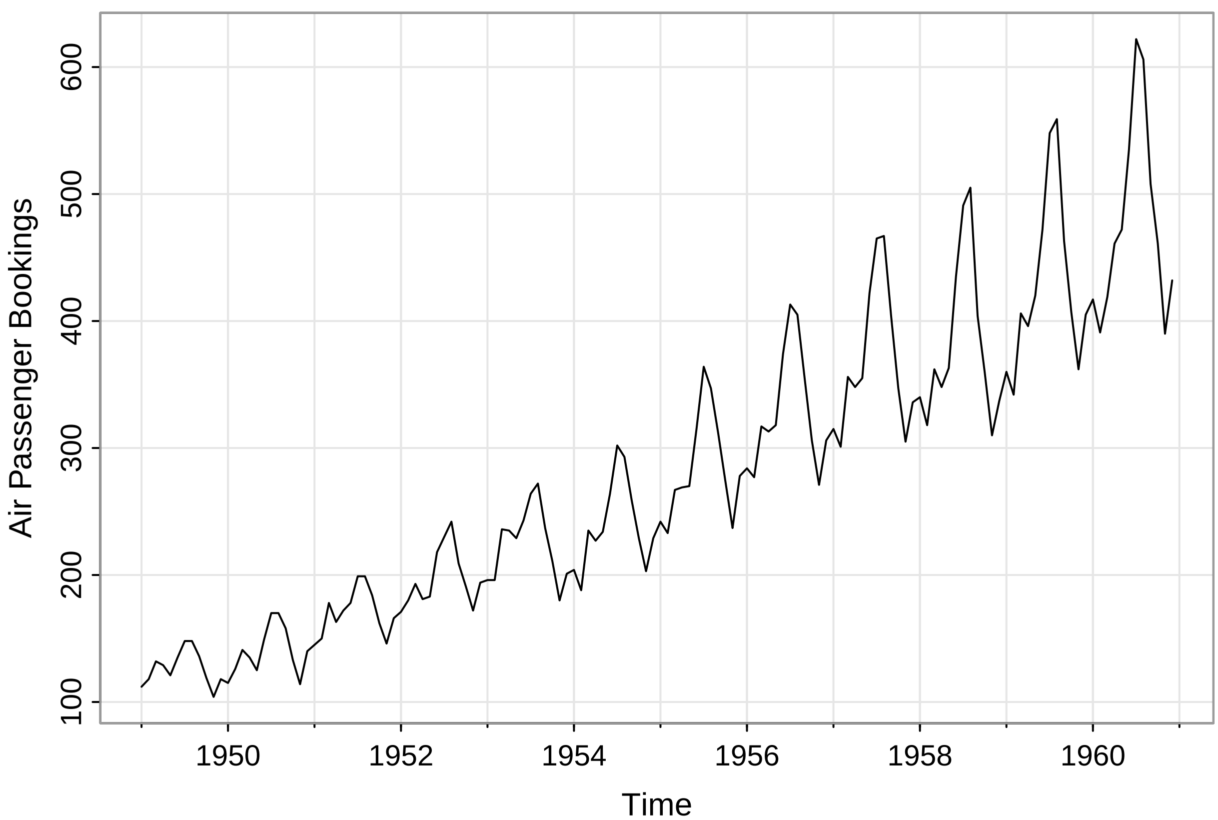
Air Passenger Bookings

The numbers of international passenger bookings (in thousands) per month on an airline in the United States were obtained from the Federal Aviation Administration for the period 1949-1960. The company used the data to predict future demand before ordering new aircraft and training aircrew. The data are available as a time series in R. These data share similar characteristics with the previous example.

```
data(AirPassengers)
AirPassengers
```

```
##      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
## 1949 112 118 132 129 121 135 148 148 136 119 104 118
## 1950 115 126 141 135 125 149 170 170 158 133 114 140
## 1951 145 150 178 163 172 178 199 199 184 162 146 166
## 1952 171 180 193 181 183 218 230 242 209 191 172 194
## 1953 196 196 236 235 229 243 264 272 237 211 180 201
## 1954 204 188 235 227 234 264 302 293 259 229 203 229
## 1955 242 233 267 269 270 315 364 347 312 274 237 278
## 1956 284 277 317 313 318 374 413 405 355 306 271 306
## 1957 315 301 356 348 355 422 465 467 404 347 305 336
## 1958 340 318 362 348 363 435 491 505 404 359 310 337
## 1959 360 342 406 396 420 472 548 559 463 407 362 405
## 1960 417 391 419 461 472 535 622 606 508 461 390 432
```

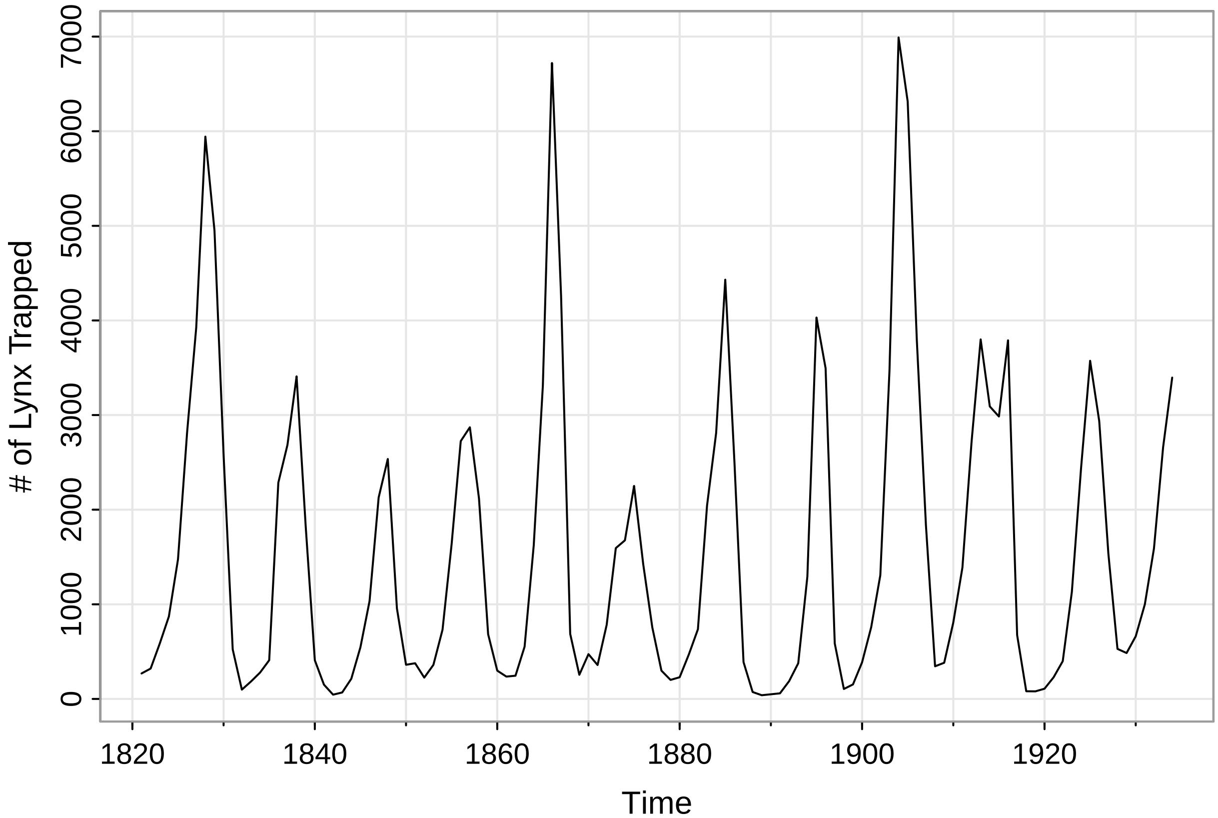
```
tsplot(AirPassengers, ylab="Air Passenger Bookings")
```



Lynx Trappings

Annual number of lynx trappings for the years 1821-1934 in the Mackenzie River District in Canada.

```
data(lynx)
tsplot(lynx, ylab="# of Lynx Trapped")
```

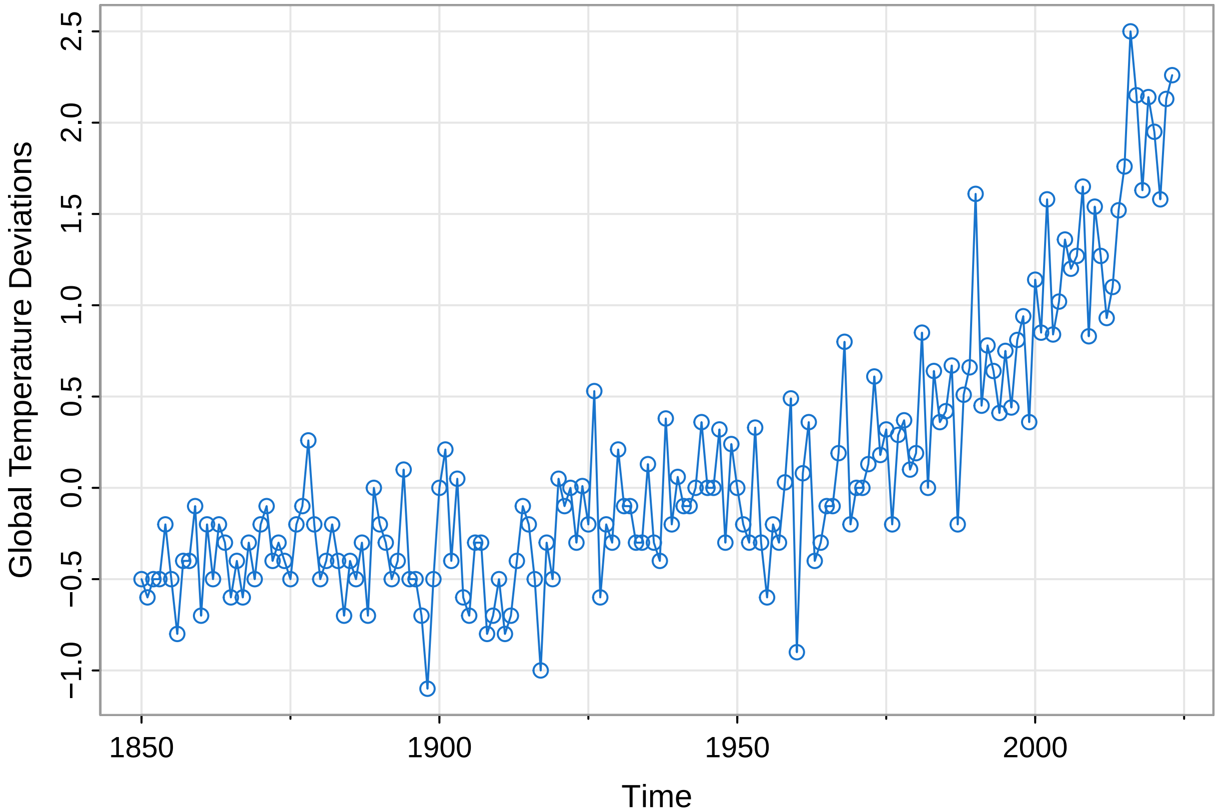


Observe that number of trapped lynx reaches high and low values every about 10 years, and some even larger figure every about 40 years. This suggests that the prominent periodicity is to be interpreted as random, but not deterministic. Understanding and modeling trend and seasonal variation is a very important aspect, much of the time series methodology is aimed at stationary series, i.e. data which do not show deterministic, but only random (cyclic) variation.

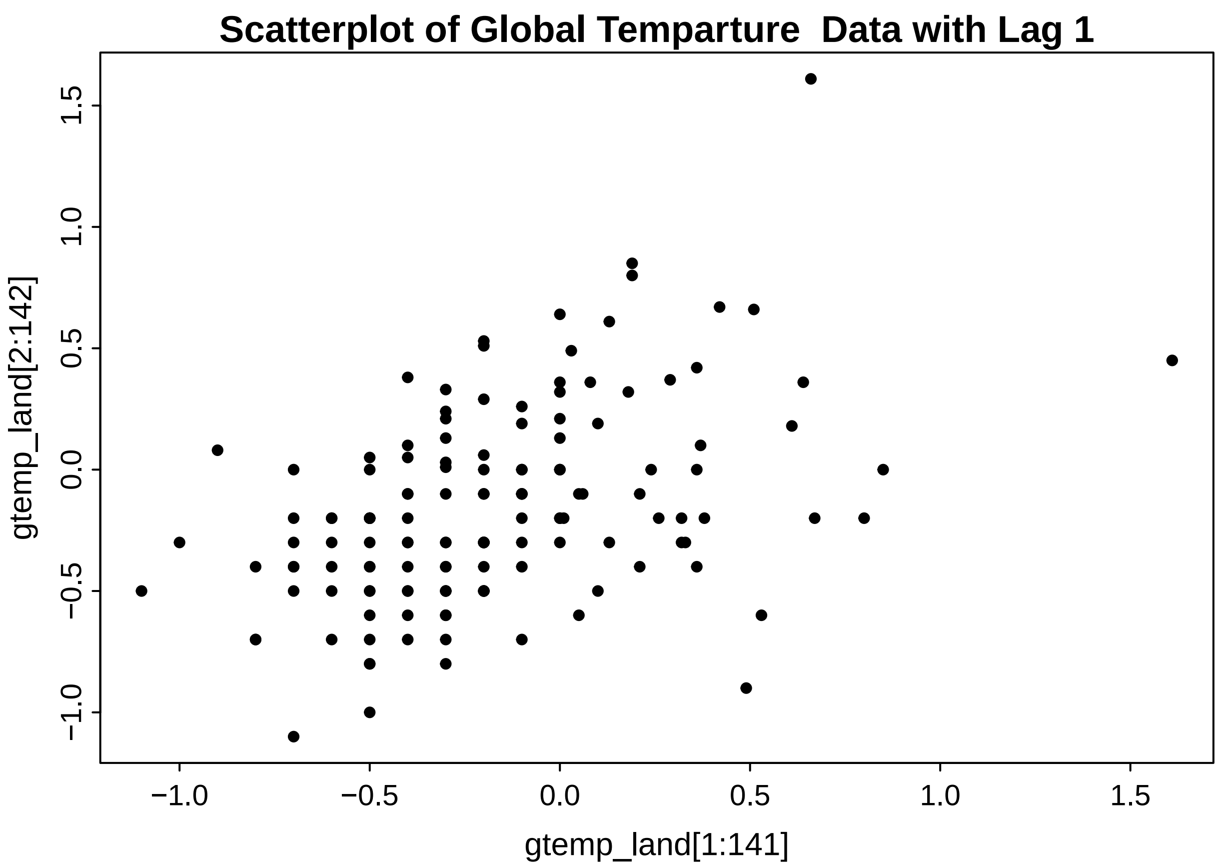
Global Temperature

These data correspond to global mean land–ocean temperature index from 1880 to 2015, with the base period 1951–1980. In particular, the data are deviations, measured in degrees centigrade, from the 1951–1980 average. Observe an apparent upward trend in the series during the latter part of the twentieth century that has been used as an argument for the global warming hypothesis. Note also the leveling off at about 1935 and then another rather sharp upward trend at about 1970. The question of interest for global warming proponents and opponents is whether the overall trend is natural or whether it is caused by some human-induced interface. The question of trend is of more interest than particular periodicities. Instead, we address the simpler question of analyzing the correlation of subsequent records, called auto correlations. The autocorrelation for lag 1 can be easily examined by producing a scatter plot of adjacent observations:

```
tsplot(gtemp_land, col=4, type="o", ylab="Global Temperature Deviations")
```




```r
plot(gtemp_land[1:141], gtemp_land[2:142], pch=20)
title("Scatterplot of Global Temparture  Data with Lag 1")
```



```
cor(gtemp_land[1:141], gtemp_land[2:142])
```

```
## [1] 0.441967
```

Hence we observe positive correlation between consecutive meausrements. So if the previous observation is below/above the mean, then the next is more likely to have the same direction.

Financial returns

The figure below shows daily returns (or percent change) of the Dow Jones Industrial Average (DJIA) from April 20, 2006 to April 20, 2016. Observe the financial crisis of 2008 in the figure. In addition, observe that the mean of the series appears to be stable with an average return of approximately zero, however, highly volatile (variable) periods tend to be clustered together.

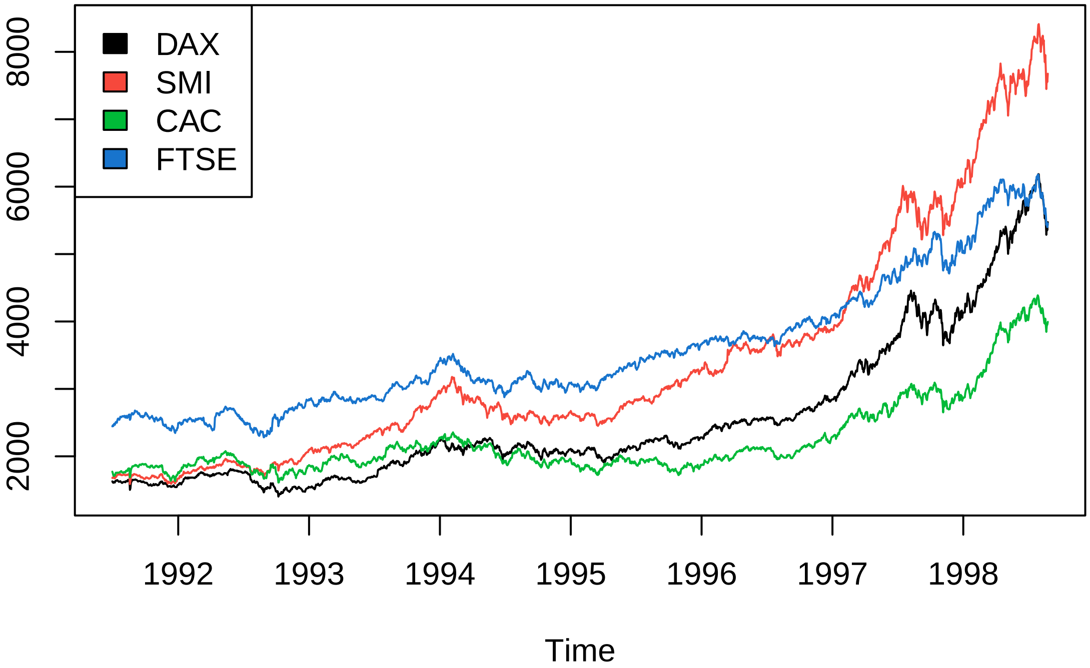
```
library(xts)          # install it if you don't have it

djiar = diff(log(djia$Close))[-1]
tsplot(as.vector(djiar$Close), xlab="Time", ylab="Returns",col=4, main="DJIA Returns")
```

More than one time series (multivariate)

The data EuStockMarkets contains the daily closing prices of major European stock indices: Germany DAX (Ibis), Switzerland SMI, France CAC, and UK FTSE. The data are sampled in business time, i.e., weekends and holidays are omitted.

```
data(EuStockMarkets)
ts.plot(EuStockMarkets, col=1:4)
leg.txt <- c("DAX", "SMI", "CAC", "FTSE")
legend("topleft", leg.txt, fill=1:4)
```


DAX
SMI
CAC
FTSE
8000
6000
4000
2000
1992
1993
1994
1995
1996
1997
1998
Time