DATA 624: PREDICTIVE ANALYTICS

Gabriel Campos

Last edited February 04, 2024

```
library('fpp3')
library('tsibble')
library('ggplot2')
library('USgas')
```

INSTRUCTIONS

Please submit exercises 2.1, 2.2, 2.3, 2.4, 2.5 and 2.8 from the Hyndman online Forecasting book. Please submit both your Rpubs link as well as attach the .pdf file with your code.

#2.1

1. Explore the following four time series: `Bricks` from `aus_production`, `Lynx` from `pelt`, `Close` from `gafa_stock`, `Demand` from `vic_elec`.

    i. Use ? (or help()) to find out about the data in each series.
    ii. What is the time interval of each series?
    iii. Use autoplot() to produce a time plot of each series.
    iv. For the last plot, modify the axis labels and title.

```
data("aus_production")
data("pelt")
data("gafa_stock")
data("vic_elec")
```

Bricks

i

Details Quarterly estimates of selected indicators of manufacturing production in Australia.

Bricks: Clay brick production in millions of bricks.

ii

Quarterly

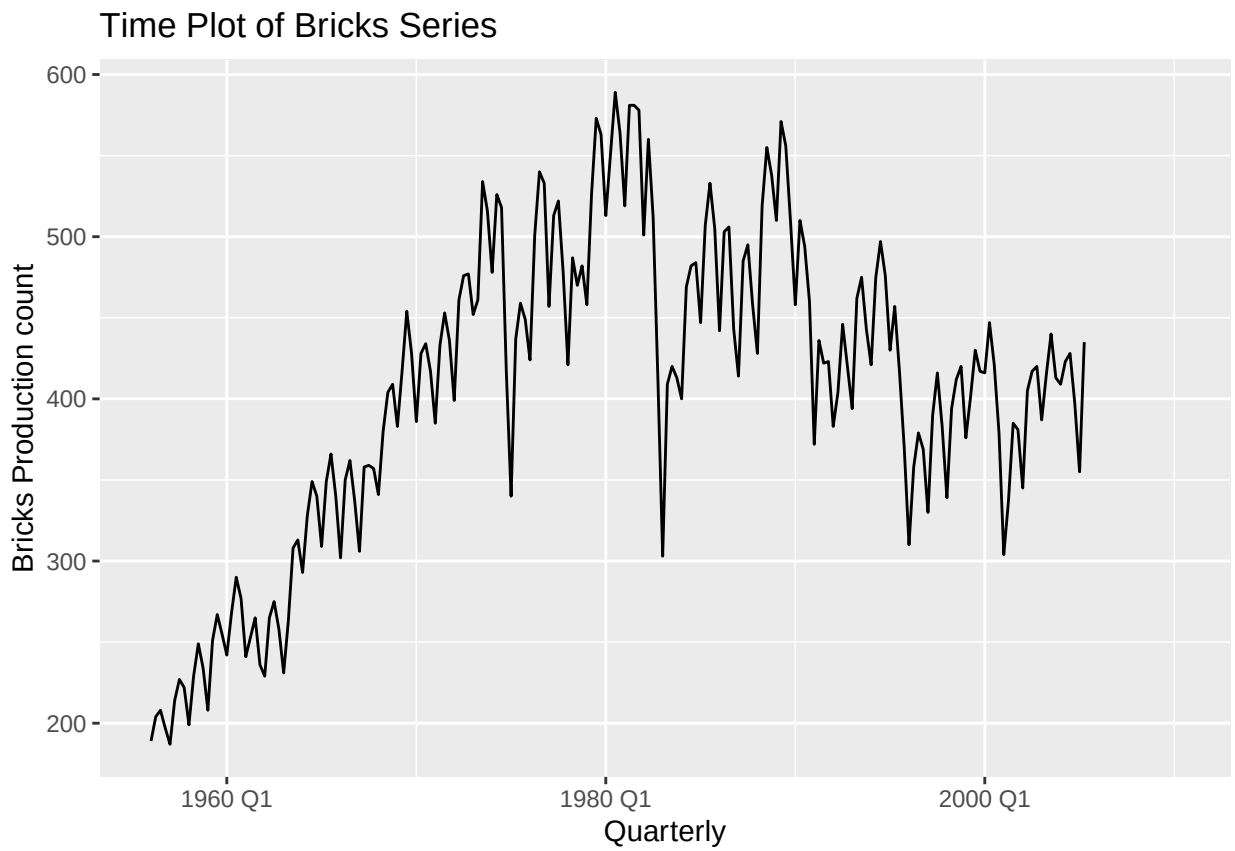
```
aus_production%>%
  select(Bricks)
```

```
## # A tsibble: 218 x 2 [1Q]
##    Bricks Quarter
##     <dbl>   <qtr>
##  1    189 1956 Q1
##  2    204 1956 Q2
##  3    208 1956 Q3
##  4    197 1956 Q4
##  5    187 1957 Q1
##  6    214 1957 Q2
##  7    227 1957 Q3
##  8    222 1957 Q4
##  9    199 1958 Q1
## 10    229 1958 Q2
## # i 208 more rows
```

iii

```
autoplot(aus_production,Bricks) +
  labs(title = "Time Plot of Bricks Series",
       x = "Quarterly",
       y = "Bricks Production count")
```

```
## Warning: Removed 20 rows containing missing values (`geom_line()`).
```



Lynx

i

pelt is an annual tsibble with two values:

Hare: The number of Snowshoe Hare pelts traded. Lynx: The number of Canadian Lynx pelts traded.

ii

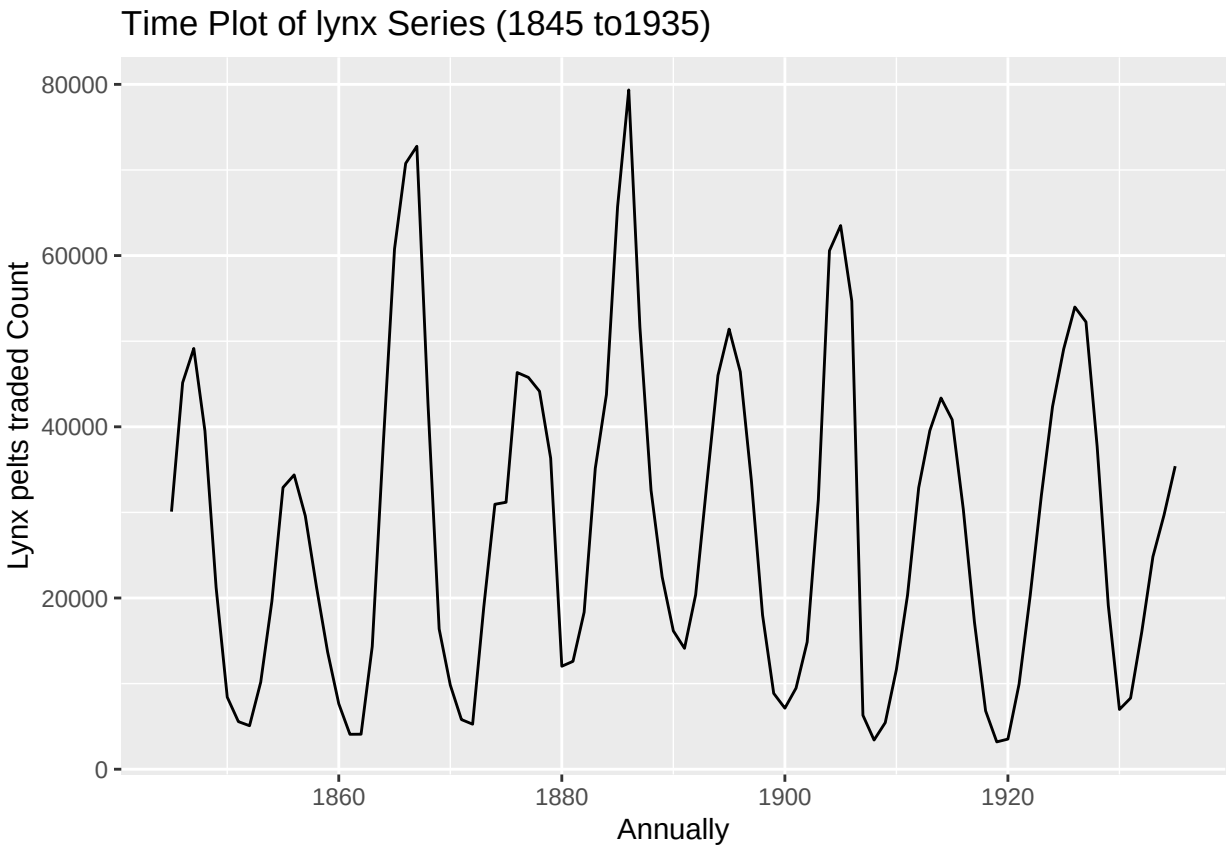
```
pelt %>%
  select(Lynx)
```

```
## # A tsibble: 91 x 2 [1Y]
##     Lynx  Year
##    <dbl> <dbl>
##  1 30090  1845
##  2 45150  1846
##  3 49150  1847
##  4 39520  1848
##  5 21230  1849
##  6  8420  1850
##  7  5560  1851
##  8  5080  1852
```

```
##  9 10170  1853
## 10 19600  1854
## # i 81 more rows
```

iii

```
autoplot(pelt,Lynx) +
  labs(title = "Time Plot of lynx Series (1845 to1935)",
       x = "Annually",
       y = "Lynx pelts traded Count")
```



Close

i

Details gafa_stock is a tsibble containing data on irregular trading days:

Open: The opening price for the stock. High: The stock's highest trading price. Low: The stock's lowest trading price. Close: The closing price for the stock. Adj_Close: The adjusted closing price for the stock. Volume: The amount of stock traded. Each stock is uniquely identified by one key:

Symbol: The ticker symbol for the stock.

ii

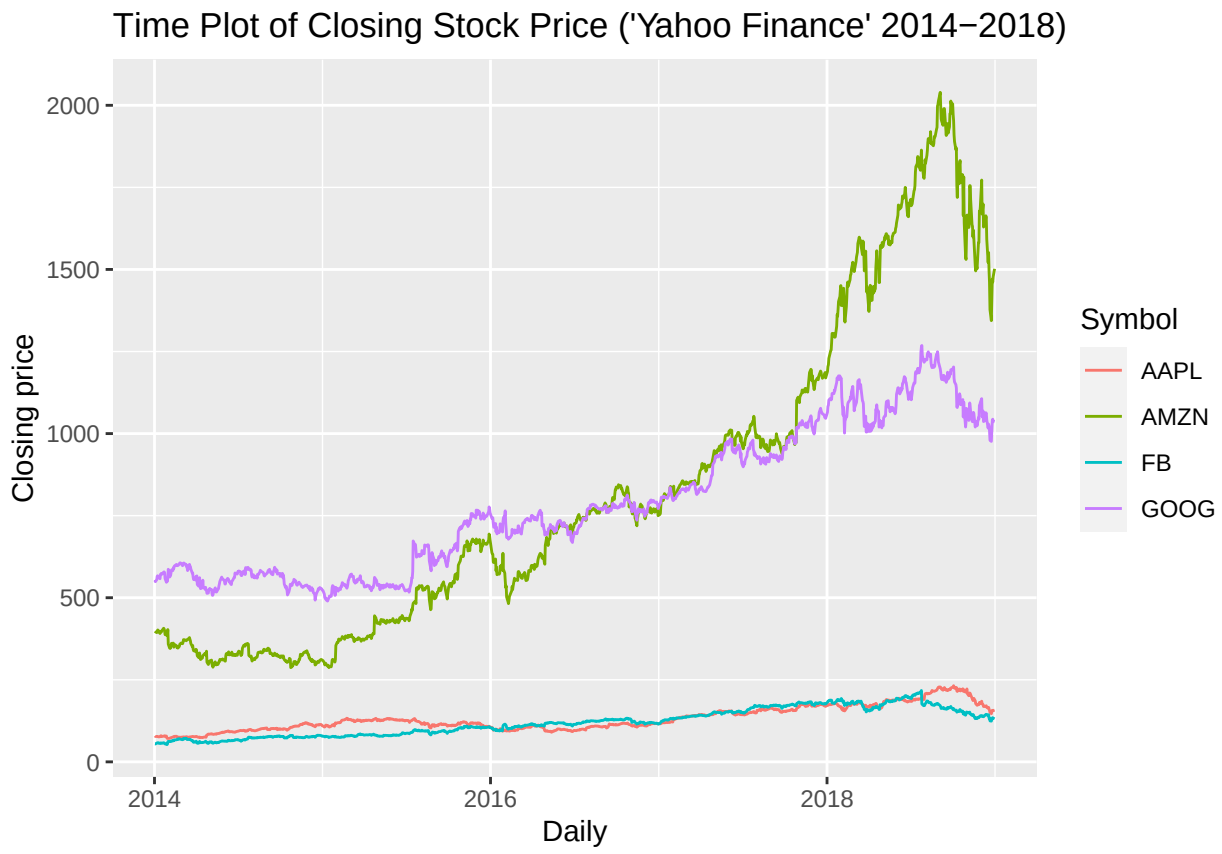
```
gafa_stock%>%
  select(Close)
```

```
## # A tsibble: 5,032 x 3 [!]
## # Key:       Symbol [4]
##    Close Date       Symbol
##    <dbl> <date>     <chr>
##  1  79.0 2014-01-02 AAPL
##  2  77.3 2014-01-03 AAPL
##  3  77.7 2014-01-06 AAPL
##  4  77.1 2014-01-07 AAPL
##  5  77.6 2014-01-08 AAPL
##  6  76.6 2014-01-09 AAPL
##  7  76.1 2014-01-10 AAPL
##  8  76.5 2014-01-13 AAPL
##  9  78.1 2014-01-14 AAPL
## 10  79.6 2014-01-15 AAPL
## # i 5,022 more rows
```

The `gafa_stock` is daily data

iii

```
autoplot(gafa_stock,Close) +
  labs(title = "Time Plot of Closing Stock Price ('Yahoo Finance' 2014-2018)",
       x = "Daily",
       y = "Closing price")
```



Demand

i

Description*

`vic_elec` is a half-hourly `tsibble` with three values:

Demand: Total electricity demand in MWh. Temperature: Temperature of Melbourne (BOM site 086071). Holiday: Indicator for if that day is a public holiday.

ii

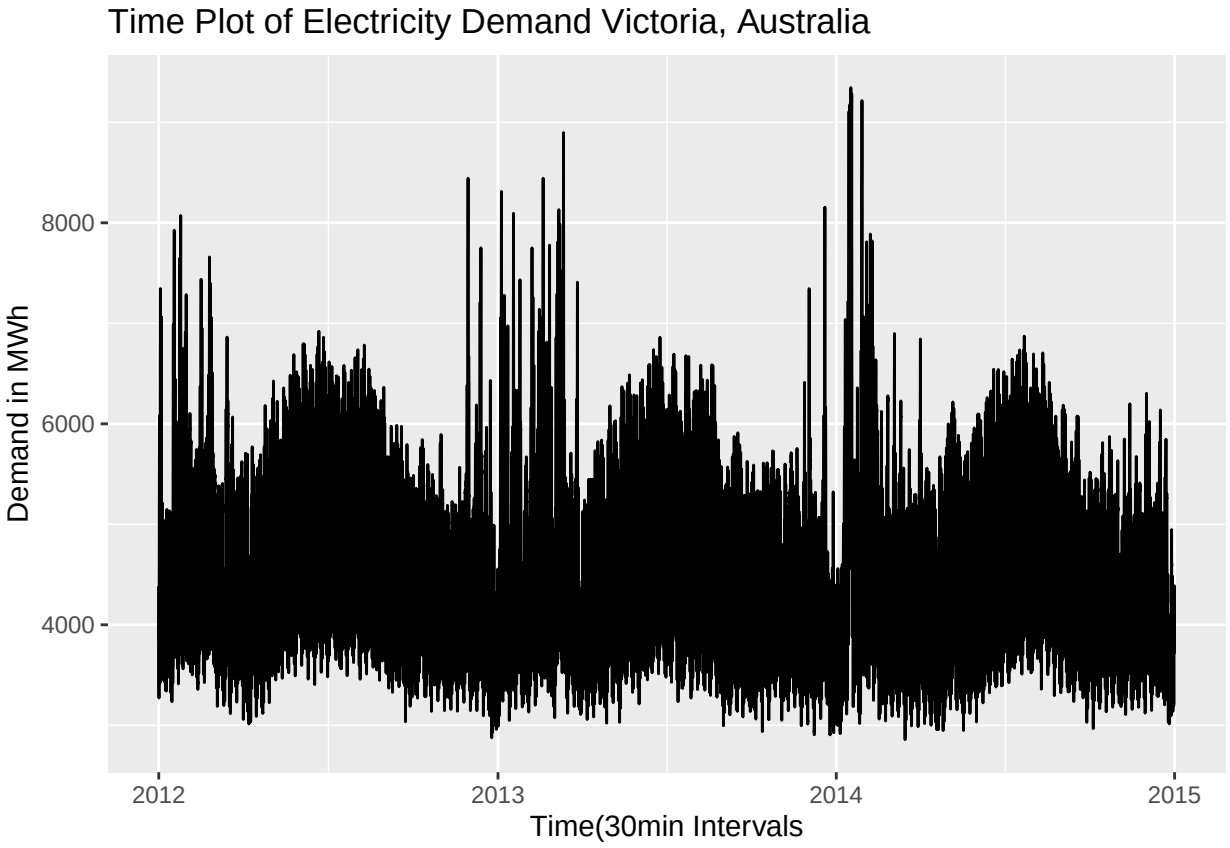
```
vic_elec %>%
  select(Demand)
```

```
## # A tsibble: 52,608 x 2 [30m] <Australia/Melbourne>
##    Demand Time
##     <dbl> <dttm>
##  1  4383. 2012-01-01 00:00:00
##  2  4263. 2012-01-01 00:30:00
##  3  4049. 2012-01-01 01:00:00
##  4  3878. 2012-01-01 01:30:00
##  5  4036. 2012-01-01 02:00:00
```

```
##  6  3866. 2012-01-01 02:30:00
##  7  3694. 2012-01-01 03:00:00
##  8  3562. 2012-01-01 03:30:00
##  9  3433. 2012-01-01 04:00:00
## 10  3359. 2012-01-01 04:30:00
## # i 52,598 more rows
```

iii & vi

```
autoplot(vic_elec,Demand) +
  labs(title = "Time Plot of Electricity Demand Victoria, Australia",
       x = "Time(30min Intervals",
       y = "Demand in MWh")
```



2.2

Use filter() to find what days corresponded to the peak closing price for each of the four stocks in gafa_stock.

```
colnames(gafa_stock)
```

```
## [1] "Symbol"    "Date"      "Open"      "High"      "Low"       "Close"
## [7] "Adj_Close" "Volume"
```

```
gafa_stock %>%
 group_by(Symbol) %>%
 filter(Close == max(Close))
```

```
## # A tsibble: 4 x 8 [!]
## # Key:       Symbol [4]
## # Groups:    Symbol [4]
##   Symbol Date        Open  High   Low Close Adj_Close   Volume
##   <chr>  <date>     <dbl> <dbl> <dbl> <dbl>     <dbl>    <dbl>
## 1 AAPL   2018-10-03  230.  233.  230.  232.      230. 28654800
## 2 AMZN   2018-09-04 2026. 2050. 2013  2040.     2040.  5721100
## 3 FB     2018-07-25  216.  219.  214.  218.      218. 58954200
## 4 GOOG   2018-07-26 1251  1270. 1249. 1268.     1268.  2405600
```

2.3