# Using R

Forecasting: principles and practice 3rd edition
(https://otexts.com/fpp3/)

# Installing R and RStudio

1. Download and install R.
2. Download and install RStudio Desktop.
3. Run RStudio. On the "Packages" tab, click on "Install packages" and install the package "fpp3" (make sure "install dependencies" is checked).

That's it! You should now be ready to go.

# Getting started with R

- Load the `fpp3` package using the **Packages** tab. (This needs to be done at the start of every R session.) This can also be done by typing `library(fpp3)` in the **Console** panel or in your script.
- Download the file http://OTexts.com/fpp3/extrafiles/tute1.csv.
- Using the **Tools** menu, choose "Import Dataset" and import the data from the `tute1.csv` file.
- The data is now saved as an object in your **Global Environment** workspace. Clicking the name of the object will cause it to be viewed. Typing the name of the object in the **Console** tab will cause it to be printed to the console.

# Use R

See what the following commands do:

```
tute1[,2]
tute1[,"Sales"]
tute1[5,]
tute1[1:10,3:4]
tute1[1:10,2] <- 0
tute1[1:20,]
```

Notice that <- means to assign the value on the right to the object on the left.

Now try using RStudio as a calculator. Figure out what each of the following is doing.

```
(100+2)/3
5*10^2
1/0
0/0
(0i-9)^(1/2)
sqrt(2 * max(-10, 0.2, 4.5))
x <- sqrt(2 * max(-10, 0.2, 4.5)) + 100
x
log(100)
log(100, base=10)
```

Save the workfile, and exit RStudio.

# R examples in the online fpp3 book

```r
# Load required packages
library(fpp3)

# Plot one time series
aus_retail |>
  filter(`Series ID`=="A3349640L") |>
  autoplot(Turnover)

# Produce some forecasts
aus_retail |>
  filter(`Series ID`=="A3349640L") |>
  model(ETS(Turnover)) |>
  forecast(h = "2 years")
```

- Once `fpp3` package is installed, its associated library `library(fpp3)` should be loaded at the beginning of the session, i.e., simply put `library(fpp3)` in the first line of your `R` code/syntax/program like above.
- Note that the pipe function `|>` in the book is equivalent to `%>%` in any `R` packages.

# Time series in R

# `tsibble` Objects (Section 2.1)

- A time series can be thought of as a list of numbers (the measurements), along with some information about what times those numbers were recorded (the index).
- This information can be stored as a `tsibble` object in R.
- `fable` refers to the forecast table in the `fpp3` package.

Suppose you have annual observations for the last few years:

| Year | Observation |
|---|---|
| 2015 | 123 |
| 2016 | 39 |
| 2017 | 78 |
| 2018 | 52 |
| 2019 | 110 |

- A `tsibble` allows storage and manipulation of multiple time series in R.
- It contains:
  - An index: time information about the observation
  - Measured variable(s): numbers of interest
  - Key variable(s): optional unique identifiers for each series
- It works with tidyverse functions.

# The tsibble index

## Example

```
mydata <- tsibble(
    year = 2015:2019,
    y = c(123, 39, 78, 52, 110),
    index = year
)
mydata
```

```
## # A tsibble: 5 x 2 [1Y]
##    year     y
##   <int> <dbl>
## 1  2015   123
## 2  2016    39
## 3  2017    78
## 4  2018    52
## 5  2019   110
```

# The `tsibble` index

- `tsibble` objects extend tidy data frames (`tibble` objects) by introducing temporal structure.
- We have set the time series `index` to be the `Year` column, which associates the measurements (`Observation`) with the time of recording (`Year`).

# The tsibble index

## Example

```
mydata <- tibble(
    year = 2015:2019,
    y = c(123, 39, 78, 52, 110)
  ) %>%
  as_tsibble(index = year)
mydata
```

```
## # A tsibble: 5 x 2 [1Y]
##    year     y
##   <int> <dbl>
## 1  2015   123
## 2  2016    39
## 3  2017    78
## 4  2018    52
## 5  2019   110
```

# The tsibble index

For observations more frequent than once per year, we need to use a time class function on the index.

```
z <- tibble(Month = paste(2019, month.abb[1:5]),
            Observation = c(50, 23, 34, 30, 25))
# knitr::kable(z, booktabs=TRUE)
z
```

```
## # A tibble: 5 x 2
##   Month    Observation
##   <chr>          <dbl>
## 1 2019 Jan          50
## 2 2019 Feb          23
## 3 2019 Mar          34
## 4 2019 Apr          30
## 5 2019 May          25
```

# The tsibble index

This can be converted to a tsibble object using the following code:

```
z %>%
  mutate(Month = yearmonth(Month)) %>%
  as_tsibble(index = Month)
```

```
## # A tsibble: 5 x 2 [1M]
##      Month Observation
##      <mth>       <dbl>
## 1 2019 Jan          50
## 2 2019 Feb          23
## 3 2019 Mar          34
## 4 2019 Apr          30
## 5 2019 May          25
```

# The `tsibble` index

- First, the `Month` column is being converted from text to a monthly time object with `yearmonth()`.
- We then convert the data frame to a `tsibble` by identifying the `index` variable using `as_tsibble()`.
- Note the addition of "[1M]" on the first line indicating this is monthly data.

# The `tsibble` index

Common time index variables can be created with these functions:

| Frequency | Function |
|---|---|
| Annual | `start:end` |
| Quarterly | `yearquarter()` |
| Monthly | `yearmonth()` |
| Weekly | `yearweek()` |
| Daily | `as_date()`, `ymd()` |
| Sub-daily | `as_datetime()` |

# Example: Australian prison population

# Read a csv file and convert to a tsibble

```
prison <- readr::read_csv(
  "https://OTexts.com/fpp3/extrafiles/prison_population.csv")

prison <- prison %>%
  mutate(Quarter = yearquarter(Date)) %>%
  select(-Date) %>%
  as_tsibble(key = c(State, Gender, Legal, Indigenous),
             index = Quarter)
```

# Read a csv file and convert to a tsibble

```
prison
```

```
## # A tsibble: 3,072 x 6 [1Q]
## # Key:       State, Gender, Legal, Indigenous [64]
##    State Gender Legal    Indigenous Count Quarter
##    <chr> <chr>  <chr>    <chr>      <dbl>   <qtr>
##  1 ACT   Female Remanded ATSI           0 2005 Q1
##  2 ACT   Female Remanded ATSI           1 2005 Q2
##  3 ACT   Female Remanded ATSI           0 2005 Q3
##  4 ACT   Female Remanded ATSI           0 2005 Q4
##  5 ACT   Female Remanded ATSI           1 2006 Q1
##  6 ACT   Female Remanded ATSI           1 2006 Q2
##  7 ACT   Female Remanded ATSI           1 2006 Q3
##  8 ACT   Female Remanded ATSI           0 2006 Q4
##  9 ACT   Female Remanded ATSI           0 2007 Q1
## 10 ACT   Female Remanded ATSI           1 2007 Q2
## # i 3,062 more rows
```

# More tutorials

There are dozens of R tutorials available on the web. Some of the best of them are listed below:

- Try R Code School
- DataCamp Introduction to R
- R tutorial (Clarkson University)
- Coursera R Programming