1. Introduction to forecasting

OTexts.com/fpp/1/
OTexts.com/fpp/2/3
The online environment

About this classroom

- It is online which is different from physical classroom.
- Features you should know (http://bit.ly/RTrainingCenterQRCard)
  - Toolbar at top
  - Mute / un-mute button
  - Chat
  - Raise hand
  - Quick poll
  - Mood icon
- Additional features we will be using
  - Breakout sessions – for the exercises
  - Polls – you will see one soon
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:
- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:
- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

**How my forecasting methodology is used:**

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

**How my forecasting methodology is used:**

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Director of Monash University’s Business & Economic Forecasting Unit

Editor-in-Chief, *International Journal of Forecasting*

**How my forecasting methodology is used:**

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales
Brief bio

- Director of Monash University’s Business & Economic Forecasting Unit
- Editor-in-Chief, *International Journal of Forecasting*

How my forecasting methodology is used:

- Pharmaceutical Benefits Scheme
- Cancer incidence and mortality
- Electricity demand
- Ageing population
- Fertilizer sales

[robjhyndman.com](http://robjhyndman.com)
Introductions

Please introduce yourself briefly using the chat box (be sure to send your message to everyone).

- What is your name?
- Where in the world are you?
- What time is it there?
- Why are you taking this course?

Example

Rob. Melbourne, Australia. 9am. To help me use R for forecasting sales for my company.
Introductions

Please introduce yourself briefly using the chat box (be sure to send your message to everyone).

- What is your name?
- Where in the world are you?
- What time is it there?
- Why are you taking this course?

Example

Rob. Melbourne, Australia. 9am. To help me use R for forecasting sales for my company.
Poll: How experienced are you in forecasting?

1. Guru: I wrote the book, done it for decades, now I do the conference circuit.

2. Expert: It has been my full time job for more than a decade.

3. Skilled: I have been doing it for years.

4. Comfortable: I understand it and have done it.

5. Learner: I am still learning.

6. Beginner: I have heard of it and would like to learn more.

7. Unknown: What is forecasting? Is that what the weather people do?
Key reference


[otexts.com/fpp/](otexts.com/fpp/)
Key reference


[otexts.com/fpp/](otexts.com/fpp/)

- Free and online
- Data sets in associated R package
- R code for examples

[otexts.com/fpp/](http://otexts.com/fpp/)

- Free and online
- Data sets in associated R package
- R code for examples

[ottexts.com/fpp/](http://ottexts.com/fpp/)

- Free and online
- Data sets in associated R package
- R code for examples

[otexts.com/fpp/](https://otexts.com/fpp/)

- Free and online
- Data sets in associated R package
- R code for examples
If you need more support on statistics or R, then I recommend:

*Introductory Statistics with R*

by Peter Dalgaard
Poll: How proficient are you in using R?

1. Guru: The R core team come to me for advice.
2. Expert: I have written several packages on CRAN.
3. Skilled: I use it regularly and it is an important part of my job.
4. Comfortable: I use it often and am comfortable with the tool.
5. User: I use it sometimes, but I am often searching around for the right function.
6. Learner: I have used it a few times.
7. Beginner: I’ve managed to download and install it.
8. Unknown: Why are you speaking like a pirate?
Which version of R are you using?

**Version:** (try `getRversion()` if you don’t know)

1. R 3.0.0 or higher
2. R 2.15.x
3. R 2.14.x
4. Something older.

**Edition**

1. Standard R (CRAN)
2. Standard R with RStudio
3. Revolution R: Community, Enterprise Workstation or Server
4. Something else?
Install required packages

install.packages("fpp", dependencies=TRUE)
Getting help with R

# Search for terms
help.search("forecasting")

# Detailed help
help(forecast)

# Worked examples
example("forecast.ar")

# Similar names
apropos("forecast")

# Help on package
help(package="fpp")
<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The forecaster’s toolbox</td>
<td>1,2</td>
</tr>
<tr>
<td>2</td>
<td>Seasonality and trends</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Exponential smoothing</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Stationarity, transformations and differencing</td>
<td>2,8</td>
</tr>
<tr>
<td>5</td>
<td>ARIMA models</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Time series cross-validation</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>Dynamic regression</td>
<td>9</td>
</tr>
</tbody>
</table>
Assumptions

This is not an introduction to R. I assume you are broadly comfortable with R code and the R environment.

This is not a statistics course. I assume you are familiar with concepts such as the mean, standard deviation, quantiles, regression, normal distribution, etc.

This is not a theory course. I am not going to derive anything. I will teach you forecasting tools, when to use them and how to use them most effectively.
Assumptions

This is not an introduction to R. I assume you are broadly comfortable with R code and the R environment.

This is not a statistics course. I assume you are familiar with concepts such as the mean, standard deviation, quantiles, regression, normal distribution, etc.

This is not a theory course. I am not going to derive anything. I will teach you forecasting tools, when to use them and how to use them most effectively.
Assumptions

- This is not an introduction to R. I assume you are broadly comfortable with R code and the R environment.
- This is not a statistics course. I assume you are familiar with concepts such as the mean, standard deviation, quantiles, regression, normal distribution, etc.
- This is not a theory course. I am not going to derive anything. I will teach you forecasting tools, when to use them and how to use them most effectively.
**Problem:** Want forecasts of each of hundreds of items. Series can be stationary, trended or seasonal. They currently have a large forecasting program written in-house but it doesn’t seem to produce sensible forecasts. They want me to tell them what is wrong and fix it.
CASE STUDY 1: Paperware company

**Problem:** Want forecasts of each of hundreds of items. Series can be stationary, trended or seasonal. They currently have a large forecasting program written in-house but it doesn’t seem to produce sensible forecasts. They want me to tell them what is wrong and fix it.

**Additional information**
- Program written in COBOL making numerical calculations limited. It is not possible to do any optimisation.
**CASE STUDY 1: Paperware company**

**Problem:** Want forecasts of each of hundreds of items. Series can be stationary, trended or seasonal. They currently have a large forecasting program written in-house but it doesn’t seem to produce sensible forecasts. They want me to tell them what is wrong and fix it.

**Additional information**
- Program written in COBOL making numerical calculations limited. It is not possible to do any optimisation.
- Their programmer has little experience in numerical computing.
**Problem:** Want forecasts of each of hundreds of items. Series can be stationary, trended or seasonal. They currently have a large forecasting program written in-house but it doesn’t seem to produce sensible forecasts. They want me to tell them what is wrong and fix it.

**Additional information**

- Program written in COBOL making numerical calculations limited. It is not possible to do any optimisation.
- Their programmer has little experience in numerical computing.
- They employ no statisticians and want the program to produce forecasts automatically.
CASE STUDY 1: Paperware company

Methods currently used

A  12 month average
C  6 month average
E  straight line regression over last 12 months
G  straight line regression over last 6 months
H  average slope between last year’s and this year’s values.
   (Equivalent to differencing at lag 12 and taking mean.)
I  Same as H except over 6 months.
K  I couldn’t understand the explanation.
The **Pharmaceutical Benefits Scheme** (PBS) is the Australian government drugs subsidy scheme.

- Many drugs bought from pharmacies are subsidised to allow more equitable access to modern drugs.
- The cost to government is determined by the number and types of drugs purchased. Currently nearly 1% of GDP.
- The total cost is budgeted based on forecasts of drug usage.
CASE STUDY 2: PBS

The **Pharmaceutical Benefits Scheme** (PBS) is the Australian government drugs subsidy scheme.

- Many drugs bought from pharmacies are subsidised to allow more equitable access to modern drugs.
- The cost to government is determined by the number and types of drugs purchased. Currently nearly 1% of GDP.
- The total cost is budgeted based on forecasts of drug usage.
The **Pharmaceutical Benefits Scheme** (PBS) is the Australian government drugs subsidy scheme.

- Many drugs bought from pharmacies are subsidised to allow more equitable access to modern drugs.
- The cost to government is determined by the number and types of drugs purchased. Currently nearly 1% of GDP.
- The total cost is budgeted based on forecasts of drug usage.
The **Pharmaceutical Benefits Scheme** (PBS) is the Australian government drugs subsidy scheme.

- Many drugs bought from pharmacies are subsidised to allow more equitable access to modern drugs.
- The cost to government is determined by the number and types of drugs purchased. Currently nearly 1% of GDP.
- The total cost is budgeted based on forecasts of drug usage.
Opp demands drug price restriction after PBS budget blow-out

The Federal Opposition has called for tighter controls on drug prices after the Pharmaceutical Benefits Scheme (PBS) budget blew out by almost $800 million.

The money was spent on two new drugs including the controversial anti-smoking aid Zyban, which dropped in price from $220 to $22 after it was listed on the PBS.
CASE STUDY 2: PBS

- In 2001: $4.5 billion budget, under-forecasted by $800 million.
- Thousands of products. Seasonal demand.
- Subject to covert marketing, volatile products, uncontrollable expenditure.
- Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.
- All forecasts being done with the FORECAST function in MS-Excel!

**Problem:** How to do the forecasting better?
CASE STUDY 2: PBS

- In 2001: $4.5 billion budget, under-forecasted by $800 million.
- **Thousands of products. Seasonal demand.**
- Subject to covert marketing, volatile products, uncontrollable expenditure.
- Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.
- All forecasts being done with the FORECAST function in MS-Excel!

**Problem:** How to do the forecasting better?
CASE STUDY 2: PBS

- In 2001: $4.5 billion budget, under-forecasts by $800 million.
- Thousands of products. Seasonal demand.
- Subject to covert marketing, volatile products, uncontrollable expenditure.
- Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.
- All forecasts being done with the FORECAST function in MS-Excel!

Problem: How to do the forecasting better?
CASE STUDY 2: PBS

- In 2001: $4.5 billion budget, under-forecasted by $800 million.
- Thousands of products. Seasonal demand.
- Subject to covert marketing, volatile products, uncontrollable expenditure.
- Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.
- All forecasts being done with the FORECAST function in MS-Excel!

**Problem:** How to do the forecasting better?
CASE STUDY 2: PBS

• In 2001: $4.5 billion budget, under-forecasted by $800 million.
• Thousands of products. Seasonal demand.
• Subject to covert marketing, volatile products, uncontrollable expenditure.
• Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.
• All forecasts being done with the FORECAST function in MS-Excel!

**Problem:** How to do the forecasting better?
In 2001: $4.5 billion budget, under-forecasted by $800 million.

Thousands of products. Seasonal demand.

Subject to covert marketing, volatile products, uncontrollable expenditure.

Although monthly data available for 10 years, data are aggregated to annual values, and only the first three years are used in estimating the forecasts.

All forecasts being done with the FORECAST function in MS-Excel!

**Problem:** How to do the forecasting better?
CASE STUDY 3: Airline
CASE STUDY 3: Airline

First class passengers: Melbourne–Sydney

Year
0.0 1.0 2.0

Business class passengers: Melbourne–Sydney

Year
0 2 4 6 8

Economy class passengers: Melbourne–Sydney

Year
0 10 20 30
CASE STUDY 3: Airline

First class passengers: Melbourne–Sydney

Business class passengers: Melbourne–Sydney

Economy class passengers: Melbourne–Sydney

Not the real data! Or is it?
CASE STUDY 3: Airline

**Problem:** how to forecast passenger traffic on major routes.

**Additional information**
- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the Grand Prix, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.
CASE STUDY 3: Airline

Problem: how to forecast passenger traffic on major routes.

Additional information

- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the Grand Prix, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.
CASE STUDY 3: Airline

**Problem:** how to forecast passenger traffic on major routes.

**Additional information**
- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the Grand Prix, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.
CASE STUDY 3: Airline

**Problem:** how to forecast passenger traffic on major routes.

**Additional information**

- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the Grand Prix, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.
Forecasting using R

Time series data
Time series consist of sequences of observations collected over time.

We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production

Forecasting is estimating how the sequence of observations will continue into the future.
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production

Forecasting is estimating how the sequence of observations will continue into the future.
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production

Forecasting is estimating how the sequence of observations will continue into the future.
Time series data

- Time series consist of sequences of observations collected over time.
- We will assume the time periods are equally spaced.

Time series examples

- Daily IBM stock prices
- Monthly rainfall
- Annual Google profits
- Quarterly Australian beer production

Forecasting is estimating how the sequence of observations will continue into the future.
Australian beer production

Forecasting using R

Time series data
Looking for stories

Peak Break-Up Times
According to Facebook status updates

- Spring Break
  - “spring clean”
- Valentine’s Day
- April Fool’s Day
- Mondays
- Summer holiday
- 2 weeks before winter holidays
- Christmas
  - “too cruel”
Looking for stories that make sense

**Total US Highway Fatality Rate**

- 1996
- 1997
- 1998
- 1999
- 2000

**Sources:**
- U.S. NHTSA, DOT HS 810 780
- U.S. Department of Agriculture

**Fresh Lemons Imported to USA from Mexico**

(Metric Tons)

**R² = 0.97**
Think about what you’re doing

My Hobby: Extrapolating

As you can see, by late next month you’ll have over four dozen husbands. Better get a bulk rate on wedding cake.

Number of Husbands

0

Yesterday Today
### Time series in R

#### Australian GDP

```r
ausgdp <- ts(scan("gdp.dat"), frequency=4, start=1971+2/4)
```

- **Class:** `ts`
- **Print and plotting methods available.**

```r
> ausgdp
  Qtr1 Qtr2 Qtr3 Qtr4
1971  4612  4651
1972  4645  4615  4645  4722
1973  4780  4830  4887  4933
1974  4921  4875  4867  4905
1975  4938  4934  4942  4979
1976  5028  5079  5112  5127
```
Time series in R

Australian GDP

ausgdp <- ts(scan("gdp.dat"), frequency=4, start=1971+2/4)

- Class: ts
- Print and plotting methods available.

> ausgdp

         Qtr1  Qtr2  Qtr3  Qtr4
1971        4612 4651
1972 4645 4615 4645 4722
1973 4780 4830 4887 4933
1974 4921 4875 4867 4905
1975 4938 4934 4942 4979
1976 5028 5079 5112 5127
Time series in R

Australian GDP

\[ \text{ausgdp} \leftarrow \text{ts}(\text{scan("gdp.dat")}, \text{frequency}=4, \text{start}=1971+2/4) \]

- Class: \text{ts}
- Print and plotting methods available.

\[
\begin{array}{cccc}
\text{Qtr1} & \text{Qtr2} & \text{Qtr3} & \text{Qtr4} \\
1971 & 4612 & 4651 \\
1972 & 4645 & 4615 & 4645 & 4722 \\
1973 & 4780 & 4830 & 4887 & 4933 \\
1974 & 4921 & 4875 & 4867 & 4905 \\
1975 & 4938 & 4934 & 4942 & 4979 \\
1976 & 5028 & 5079 & 5112 & 5127 \\
\end{array}
\]
> plot(ausgdp)
Residential electricity sales

> elecsales
Time Series:
Start = 1989
End = 2008
Frequency = 1

[1] 2354.34 2379.71 2318.52 2468.99 2386.09 2569.47
[7] 2575.72 2762.72 2844.50 3000.70 3108.10 3357.50
[13] 3075.70 3180.60 3221.60 3176.20 3430.60 3527.48
[19] 3637.89 3655.00
Main package used in this course

```r
> library(fpp)
```
Main package used in this course

```r
> library(fpp)
```

This loads:

- some data for use in examples and exercises
- `forecast` package (for forecasting functions)
- `tseries` package (for a few time series functions)
- `fma` package (for lots of time series data)
- `expsmooth` package (for more time series data)
- `lmtest` package (for some regression functions)
Main package used in this course

```r
> library(fpp)
```

This loads:

- some data for use in examples and exercises
- **forecast** package (for forecasting functions)
- **tseries** package (for a few time series functions)
- **fma** package (for lots of time series data)
- **expsmooth** package (for more time series data)
- **lmtest** package (for some regression functions)
Main package used in this course

```r
> library(fpp)
```

This loads:

- some data for use in examples and exercises
- **forecast** package (for forecasting functions)
- **tseries** package (for a few time series functions)
- **fma** package (for lots of time series data)
- **expsmooth** package (for more time series data)
- **lmtest** package (for some regression functions)
Main package used in this course

```r
> library(fpp)
```

This loads:

- some data for use in examples and exercises
- **forecast** package (for forecasting functions)
- **tseries** package (for a few time series functions)
- **fma** package (for lots of time series data)
- **expsmooth** package (for more time series data)
- **lmtest** package (for some regression functions)
Main package used in this course

`> library(fpp)`

This loads:

- some data for use in examples and exercises
- **forecast** package (for forecasting functions)
- **tseries** package (for a few time series functions)
- **fma** package (for lots of time series data)
- **expsmooth** package (for more time series data)
- **lmtest** package (for some regression functions)
Main package used in this course

```r
> library(fpp)
```
This loads:

- some data for use in examples and exercises
- `forecast` package (for forecasting functions)
- `tseries` package (for a few time series functions)
- `fma` package (for lots of time series data)
- `expsmooth` package (for more time series data)
- `lmtest` package (for some regression functions)
Forecasting using R

Some simple forecasting methods
Some simple forecasting methods

Australian quarterly beer production

megaliters

1995 2000 2005

Forecasting using R
Some simple forecasting methods

Australian quarterly beer production

Can you think of any forecasting methods for these data?
Some simple forecasting methods

Number of pigs slaughtered in Victoria

Number of pigs slaughtered in Victoria


Thousands

110

100

90

80

How would you forecast these data?
Some simple forecasting methods

Dow Jones index (daily ending 15 Jul 94)
How would you forecast these data?
Some simple forecasting methods

### Average method
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

### Naïve method (for time series only)
- Forecasts equal to last observed value.

### Seasonal naïve method
- Forecasts equal to last value from same season.
- Forecasts: \( \hat{y}_{T+h|T} = y_{T+h-km} \) where \( m \) is the seasonal period and \( k = \lfloor (h-1)/m \rfloor + 1 \).
Some simple forecasting methods

Average method

- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

Naïve method (for time series only)

- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \)

Seasonal naïve method

- Consequence of efficient market hypothesis.
Some simple forecasting methods

Average method

- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

Naïve method (for time series only)

- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

Seasonal naïve method
Some simple forecasting methods

**Average method**
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

**Naïve method** (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

**Seasonal naïve method**
Some simple forecasting methods

Average method
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

Naïve method (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

Seasonal naïve method
- Forecasts equal to last value from same season.
Some simple forecasting methods

**Average method**
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

**Naïve method** (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

**Seasonal naïve method**
- Forecasts equal to last value from same season.
- Forecasts: \( \hat{y}_{T+h|T} = y_{T-km} \), where \( m \) = seasonal period and \( k = \lfloor (h - 1)/m \rfloor + 1 \).
Some simple forecasting methods

Average method
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

Naïve method (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

Seasonal naïve method
- Forecasts equal to last value from same season.
- Forecasts: \( \hat{y}_{T+h|T} = y_{T+h-km} \) where \( m \) = seasonal period and \( k = \lfloor (h - 1)/m \rfloor + 1 \).
Some simple forecasting methods

**Average method**
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

**Naïve method** (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

**Seasonal naïve method**
- Forecasts equal to last value from same season.
- Forecasts: \( \hat{y}_{T+h|T} = y_{T+h-km} \) where \( m \) = seasonal period and \( k = \lceil (h - 1)/m \rceil + 1 \).
Some simple forecasting methods

**Average method**
- Forecast of all future values is equal to mean of historical data \( \{y_1, \ldots, y_T\} \).
- Forecasts: \( \hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T \)

**Naïve method** (for time series only)
- Forecasts equal to last observed value.
- Forecasts: \( \hat{y}_{T+h|T} = y_T \).
- Consequence of efficient market hypothesis.

**Seasonal naïve method**
- Forecasts equal to last value from same season.
- Forecasts: \( \hat{y}_{T+h|T} = y_{T+h-km} \) where \( m = \) seasonal period and \( k = \lfloor (h - 1)/m \rfloor + 1 \).
Some simple forecasting methods

Forecasts for quarterly beer production

Which method is which?
Some simple forecasting methods

Forecasts for quarterly beer production

- Mean method
- Naive method
- Seasonal naive method

Which method is which?
Drift method

- Forecasts equal to last value plus average change.

- Forecasts:

\[
\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \sum_{t=2}^{T} (y_t - y_{t-1}) = y_T + \frac{h}{T-1} (y_T - y_1).
\]

- Equivalent to extrapolating a line drawn between first and last observations.
Drift method

- Forecasts equal to last value plus average change.

- Forecasts:

\[
\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \sum_{t=2}^{T} (y_t - y_{t-1})
\]

\[
= y_T + \frac{h}{T-1} (y_T - y_1).
\]

- Equivalent to extrapolating a line drawn between first and last observations.
Forecasts equal to last value plus average change.

Forecasts:

\[
\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \sum_{t=2}^{T} (y_t - y_{t-1})
\]

\[
= y_T + \frac{h}{T-1} (y_T - y_1).
\]

Equivalent to extrapolating a line drawn between first and last observations.
Some simple forecasting methods

Dow Jones Index (daily ending 15 Jul 94)
Some simple forecasting methods

Dow Jones Index (daily ending 15 Jul 94)

- Mean method
- Naive method
- Drift model

Data points:
- Day 0: 3600
- Day 50: 3700
- Day 100: 3800
- Day 150: 3900
- Day 200: 3600
- Day 250: 3700
- Day 300: 3800

Graph showing the trend of the Dow Jones Index with different forecasting methods.
Some simple forecasting methods

- **Mean:** `meanf(x, h=20)`
- **Naive:** `naive(x, h=20)` or `rwf(x, h=20)`
- **Seasonal naive:** `snaive(x, h=20)`
- **Drift:** `rwf(x, drift=TRUE, h=20)`
Some simple forecasting methods

- **Mean**: `meanf(x, h=20)`
- **Naive**: `naive(x, h=20)` or `rfw(x, h=20)`
- **Seasonal naive**: `snaive(x, h=20)`
- **Drift**: `rfw(x, drift=TRUE, h=20)`
Some simple forecasting methods

- Mean: `meanf(x, h=20)`
- Naive: `naive(x, h=20)` or `rwf(x, h=20)`
- Seasonal naive: `snaive(x, h=20)`
- Drift: `rwf(x, drift=TRUE, h=20)`
Some simple forecasting methods

- Mean: `meanf(x, h=20)`
- Naive: `naive(x, h=20)` or `rwf(x, h=20)`
- Seasonal naive: `snaive(x, h=20)`
- Drift: `rwf(x, drift=TRUE, h=20)`
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
You should all have received an invitation to join the Piazza course page.

All discussion and distribution of information outside of classes will take place there.

Homework available there.

Session slides available there after each session.

Please ask questions there.

If you haven’t received an invitation, or are having trouble, please ask James to help (james@revolutionanalytics.com).
There will be a homework sheet available for each session to give you practice with the concepts introduced.

Please try to do each sheet before the next session.

I will post answers on Piazza.
Homework

- There will be a homework sheet available for each session to give you practice with the concepts introduced.
- Please try to do each sheet before the next session.
- I will post answers on Piazza.
Homework

- There will be a homework sheet available for each session to give you practice with the concepts introduced.
- Please try to do each sheet before the next session.
- I will post answers on Piazza.