Time Series Analysis and Forecasting with R

Mohammad Alfan Alfian Riyadi (Data Scientist)
alfan@warungpintar.co
Meet 2

About Me

- Education
  - Bachelor Degree Statistics (2012-2016)
  - Master Degree Statistics (2016-2018)
- Activity
  - DATA SCIENCE INDONESIA - Advisor DSI East Java Chapter
- Working Experience
  - Warung PINTAR Sept 2018 - Now

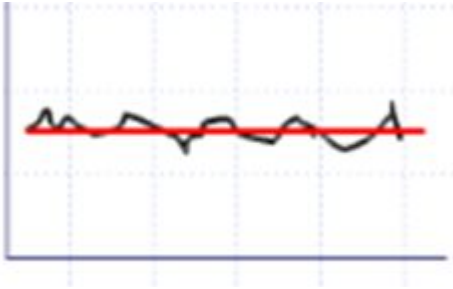
Outline

- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive Models
- Forecasting with Moving Average Models
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Outline

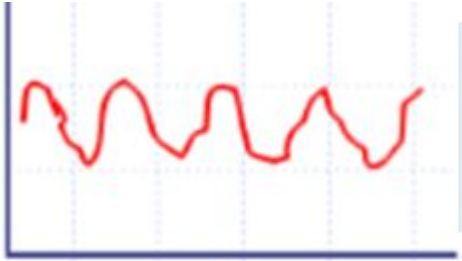
- **Introduction**
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Introduction : General Time Series Pattern (Review)



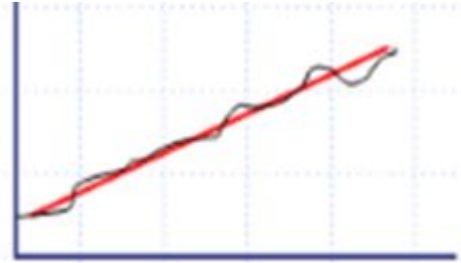
Stationary

- ✓ Naive Model
- ✓ Simple Exponential Smoothing
- ✓ Simple Average



Seasonal

- ✓ Naive Model
- ✓ Seasonal Time Series Regression
- ✓ "Holt-Winters" Exponential Smoothing



Trend

- ✓ Naive Model
- ✓ "Holt" Exponential Smoothing
- ✓ Double Moving Average
- ✓ Trend Analysis Time Series Regression

Introduction : Library Materials

- forecast
- smooth
- ggplot2

Outline

- Introduction
- **Exploratory Time Series Data Analysis**
- Forecasting with Naive
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Exploratory Time Series Data Analysis

- **ts** object

```
y <- ts(cumsum(1 + round(rnorm(100), 2)), start = c(1990, 7), frequency = 12)
y

y <-  ts(cumsum(1 + round(rnorm(100), 2)), start = c(1990, 7), frequency = 4)
y
```

Exploratory Time Series Data Analysis

- Frequency in time series

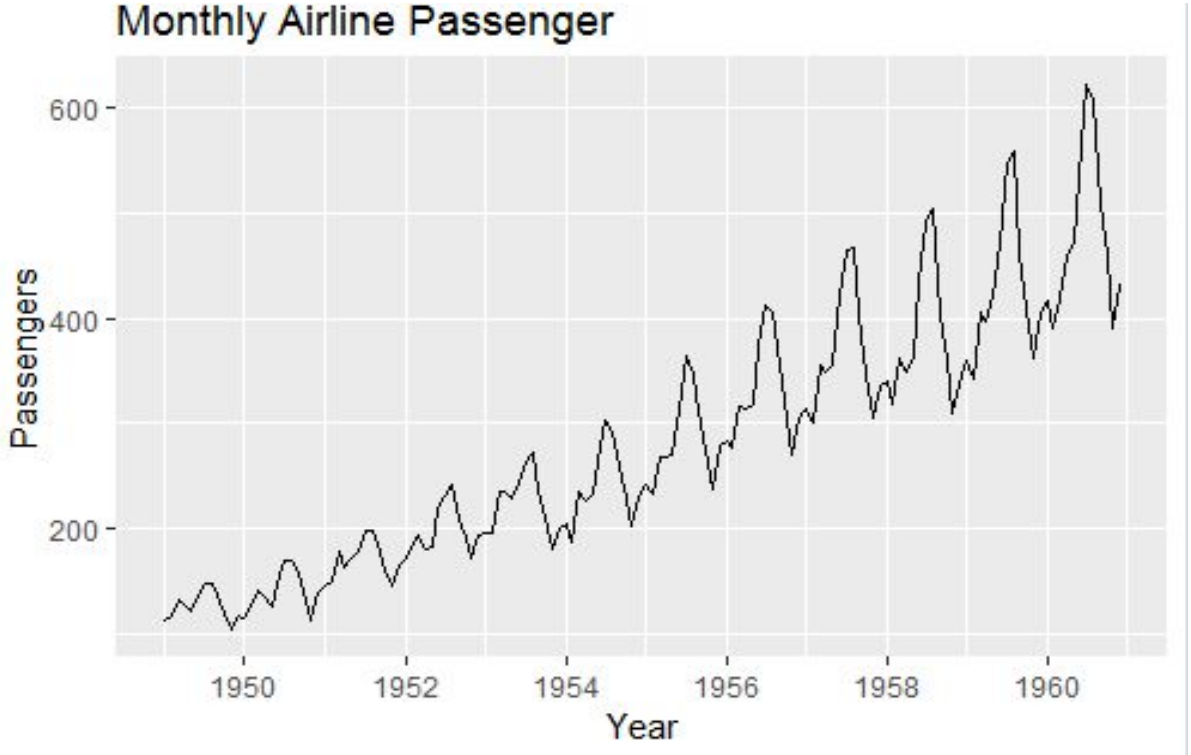
Data	Frequency
Annual	1
Quarterly	4
Monthly	12
Weekly	52

Exploratory Time Series Data Analysis

- Time Series Plot

```
autoplot(AirPassengers)
    + ggtitle("Monthly Airline Passenger")
    + xlab("Year")
    + ylab("Passengers")
```

Exploratory Time Series Data Analysis

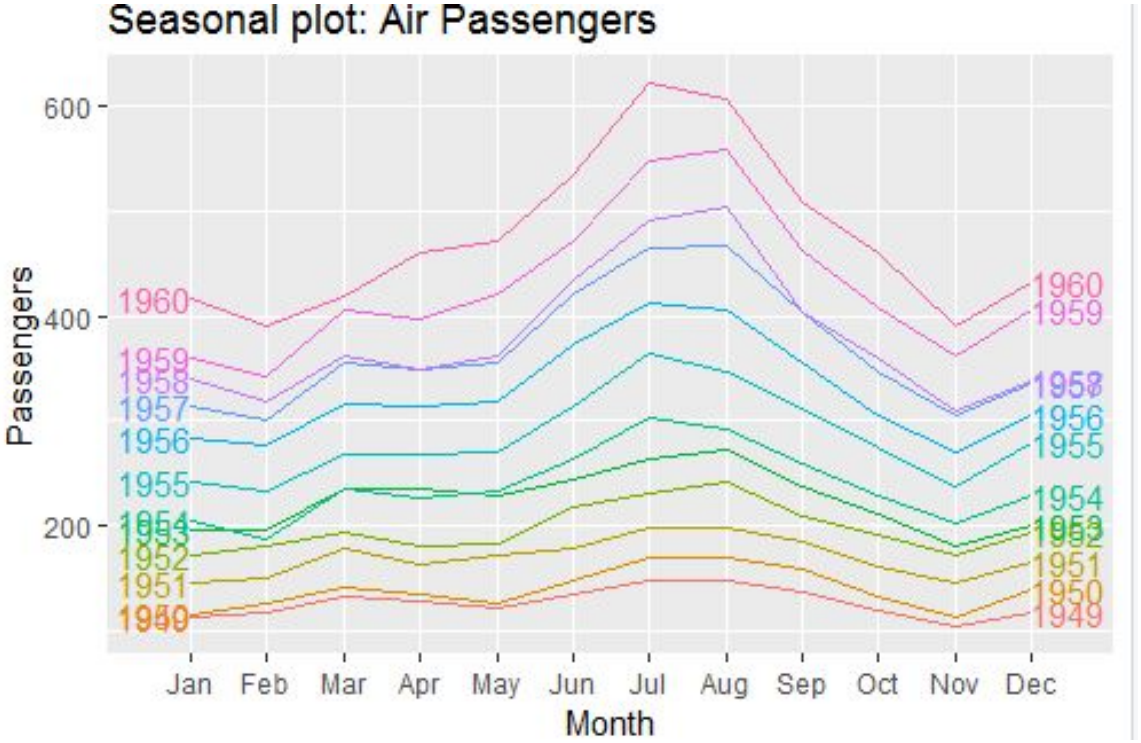


Exploratory Time Series Data Analysis

- Time Series Plot

```
ggseasonplot(AirPassengers, year.labels=TRUE, year.labels.left=TRUE)
    + ylab("Passengers")
    + ggtitle("Seasonal plot: Air Passengers")
```

Exploratory Time Series Data Analysis



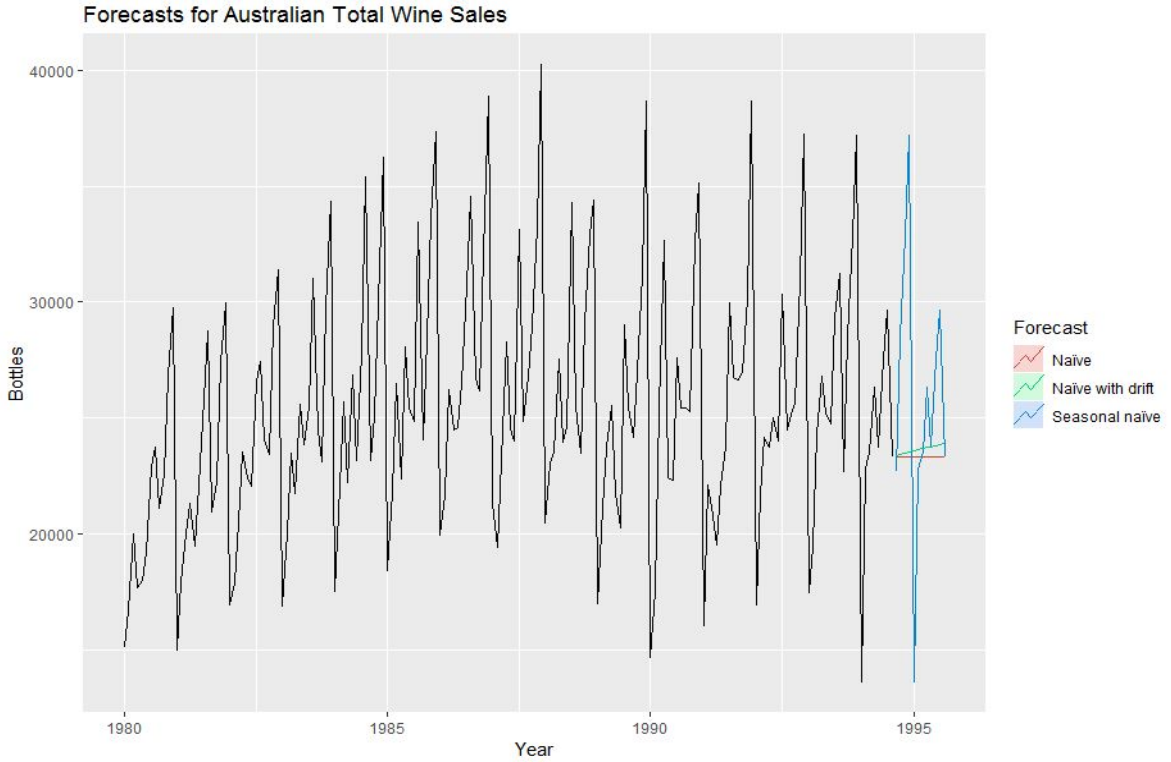
Outline

- Introduction
- Exploratory Time Series Data Analysis
- **Forecasting with Naive**
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Forecasting with Naive

```
autoplot(wineind) +
      autolayer(naive(wineind, h=12),
                series="Naïve", PI=FALSE) +
      autolayer(snaive(wineind, h=12),
                series="Seasonal naïve", PI=FALSE) +
      autolayer(rwf(wineind, h=12,drift = TRUE),
                series="Naïve with drift", PI=FALSE) +
      ggtitle("Forecasts for Australian Total Wine Sales") +
      xlab("Year") +
      ylab("Bottles") +
      guides(colour=guide_legend(title="Forecast"))
```

Forecasting with Naive



Outline

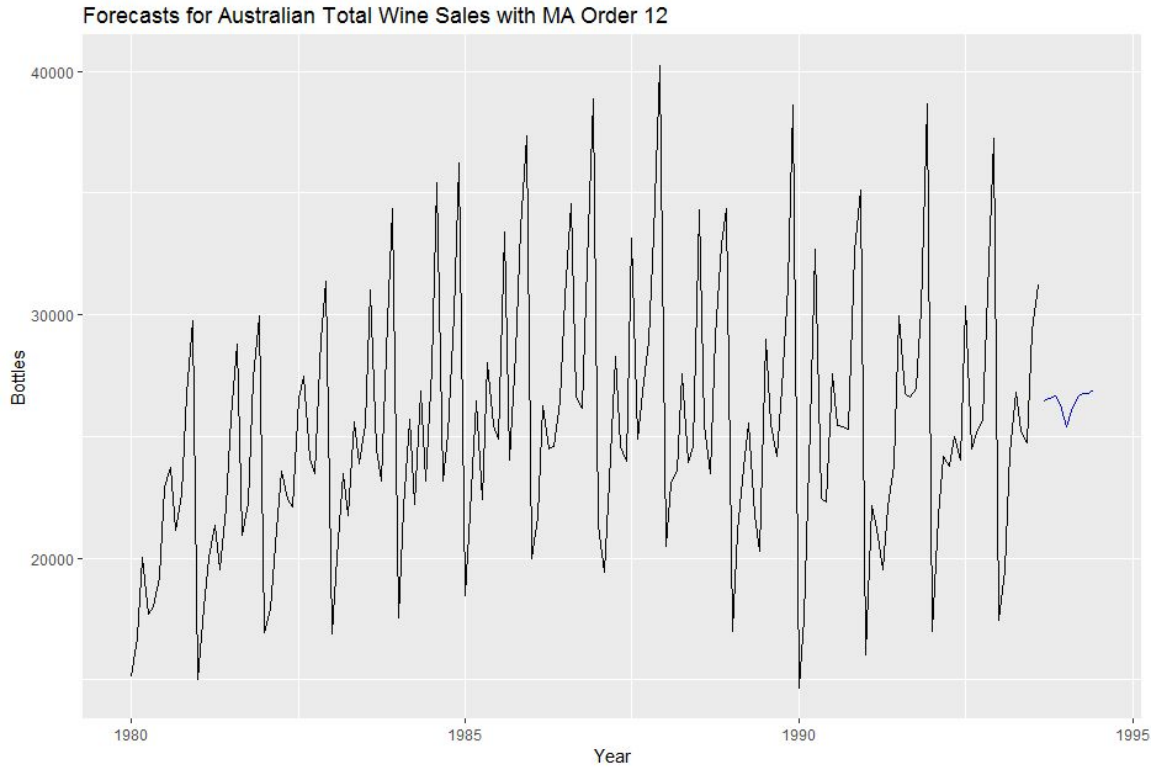
- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- **Forecasting with Moving Average**
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Forecasting with Moving Average

```
sma12 <- sma(wineind,order=12,h=12,holdout=TRUE,interval="none")

autoplot(forecast(sma12,interval=FALSE))+
    ggtitle("Forecasts for Australian Total Wine Sales with MA Order 12") +
    xlab("Year") +
    ylab("Bottles") +
    guides(colour=guide_legend(title="Forecast"))
```

Forecasting with Moving Average



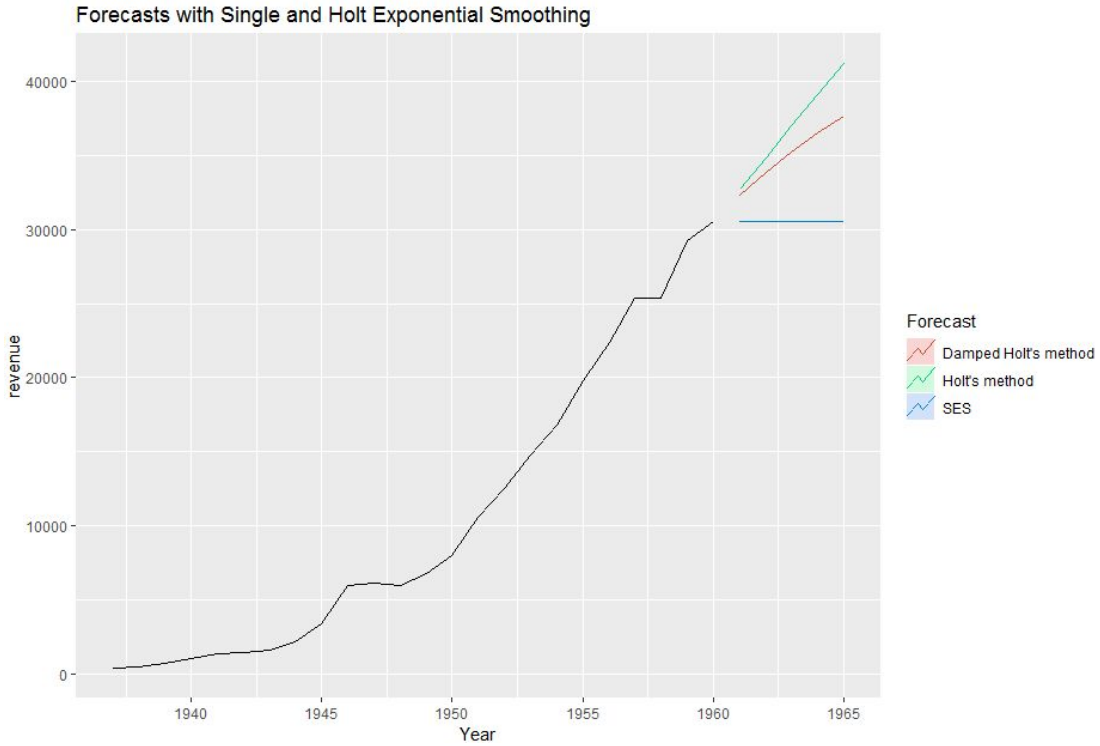
Outline

- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- Forecasting with Moving Average
- **Forecasting with Exponential Smoothing**
- Forecasting with Time Series Regression
- Evaluation Models
- Exercise

Forecasting with Exponential Smoothing

```
s_exp <- ses(airmiles, h=5)
holt1 <- holt(airmiles, h=5)
holt2 <- holt(airmiles, damped=TRUE, phi = 0.9, h=5)
autoplot(airmiles) +
      autolayer(s_exp, series="SES",PI=FALSE) +
      autolayer(holt1, series="Holt's method", PI=FALSE) +
      autolayer(holt2, series="Damped Holt's method", PI=FALSE) +
      ggtitle("Forecasts with Single and Holt Exponential Smoothing") +
      xlab("Year") +
      ylab("revenue") +
      guides(colour=guide_legend(title="Forecast"))
```

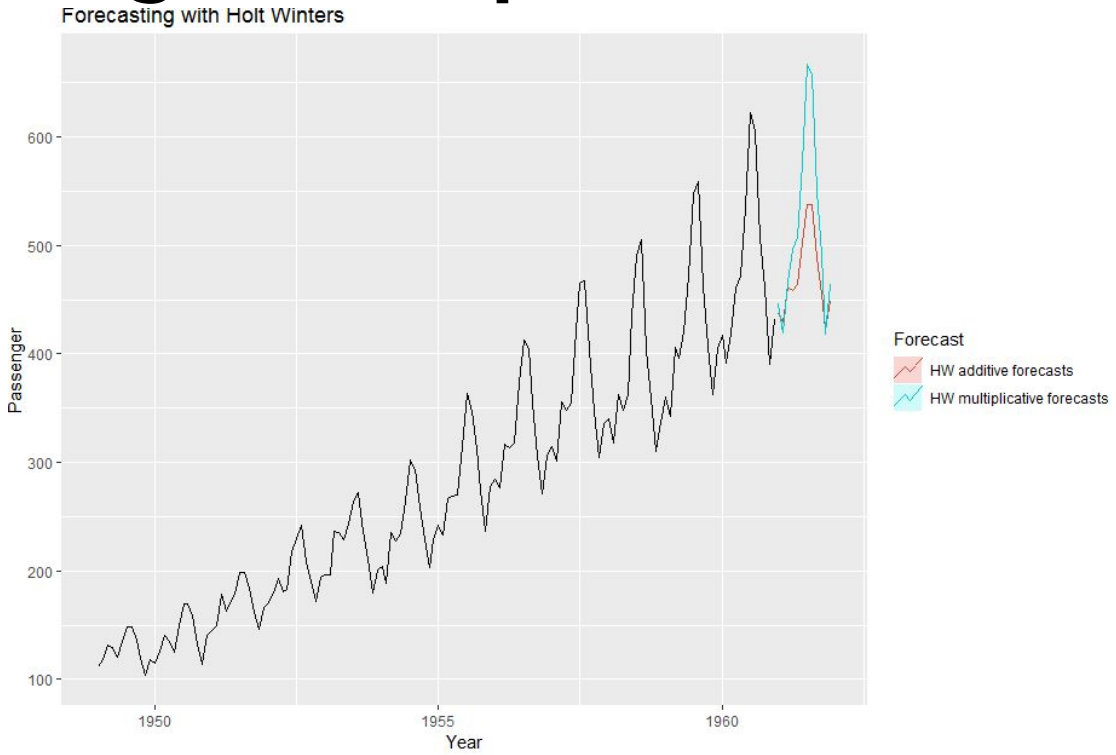
Forecasting with Exponential Smoothing



Forecasting with Exponential Smoothing

```
hw_1 <- hw(AirPassengers,seasonal="additive",h=12)
hw_2 <- hw(AirPassengers,seasonal="multiplicative",h=12)
autoplot(AirPassengers) +
    autolayer(hw_1, series="HW additive forecasts", PI=FALSE) +
    autolayer(hw_2, series="HW multiplicative forecasts", PI=FALSE) +
    xlab("Year") +
    ylab("Passenger") +
    ggtitle("Forecasting with Holt Winters") +
    guides(colour=guide_legend(title="Forecast"))
```

Forecasting with Exponential Smoothing



Outline

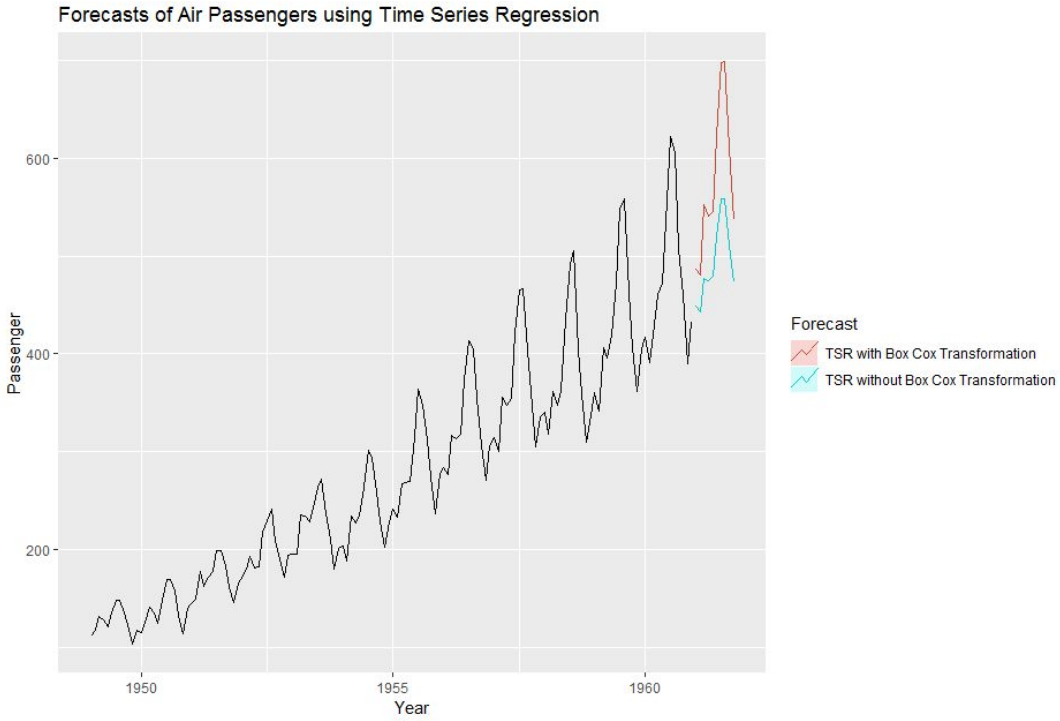
- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- **Forecasting with Time Series Regression**
- Evaluation Models
- Exercise

Forecasting with Time Series Regression

```
fit1 <- tslm(AirPassengers ~ trend + season,lambda = 0)
fit2 <- tslm(AirPassengers ~ trend + season)
fcast1 <- forecast(fit1,h=12)
fcast2 <- forecast(fit2,h=12)

autoplot(AirPassengers) +
    autolayer(fcast1,series="TSR with Box Cox Transformation", PI=FALSE) +
    autolayer(fcast2,series="TSR without Box Cox Transformation", PI=FALSE) +
    xlab("Year") +
    ylab("Passenger") +
    ggtitle("Forecasts of Air Passengers using Time Series Regression") +
    guides(colour=guide_legend(title="Forecast"))
```

Forecasting with Time Series Regression



Box Cox Transformation Formula

$$w_t = \begin{cases} \log(y_t) & \text{if } \lambda = 0; \\ (y_t^{\lambda} - 1)/\lambda & \text{otherwise.} \end{cases}$$

Outline

- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- **Evaluation Models**
- Exercise

Evaluation Models

```
AirPassengers_train <-window(AirPassengers,start=1949,end=c(1959,12))
AirPassengers_test <-window(AirPassengers,start=1960)
fit1 <- tslm(AirPassengers_train ~ trend + season,lambda = 0)
fit2 <- tslm(AirPassengers_train ~ trend + season)
fcast1 <- forecast(fit1,h=12)
fcast2 <- forecast(fit2,h=12)

accuracy(AirPassengers_test,fcast1$mean)
accuracy(AirPassengers_test,fcast2$mean)
```

Evaluation Models

```
hw_1 <- hw(AirPassengers_train,seasonal="additive",h=12)
hw_2 <- hw(AirPassengers_train,seasonal="multiplicative",h=12)

accuracy(AirPassengers_test,hw_1$mean)
accuracy(AirPassengers_test,hw_2$mean)
```

Outline

- Introduction
- Exploratory Time Series Data Analysis
- Forecasting with Naive
- Forecasting with Moving Average
- Forecasting with Exponential Smoothing
- Forecasting with Time Series Regression
- Evaluation Models
- **Exercise**

Practice Makes Right
Repetition Makes Perfect

https://intip.in/latihanMFM