Question 1

Complete

Points out of 10.00

The forecast model for the monthly sales of new one-family houses in the USA was built using the Random Walk Model as described below:

```
require(fpp)
data(hsales)
plot(hsales, xlab = "Time",
     ylab = "Sales",
     main = "Monthly house sales in US (Jan/1973-Nov/1995)")
train = window(hsales, end = c(1993,12))
test = window(hsales, start = c(1994,1))
help(rwf)
rwfForecast = rwf(train, h = 23)$mean
accuracy(rwfForecast, test)
```

Build comparable models (i.e., the same training and testing data sets, as well as the parameter h) for the following baseline models: Average, Seasonal Naive, and Drift; compute their performance metrics, and submit the R code.

```
> #Average
> rwfForecastAvg = meanf(train, h = 23)$mean
> print(accuracy(rwfForecastAvg, test))
               ME     RMSE      MAE     MPE     MAPE      ACF1 Theil's U
Test set 4.051587 9.216133 7.850759 5.07499 13.75973 0.5095178   1.13105



> #Seasonal naive
> rwfForecastsnaive = snaive(train, h = 23)$mean
> print(accuracy(rwfForecastsnaive, test))
                ME     RMSE MAE        MPE    MAPE     ACF1 Theil's U
Test set 0.3043478 6.160886   5 -0.7312374 9.12828 0.224307 0.8031005



> #Drift
> rwfForecastdrift = rwf(train,drift = TRUE ,h = 23)$mean
> print(accuracy(rwfForecastdrift, test))
```

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
Test set	5.191235	9.761548	8.393037	7.159951	14.50303	0.5083059	1.188562

Q1.R

Question 2

Correct

6.00 points out of 6.00

Compare the accuracy metrics for the forecast models for the monthly housing sales prices in US that you built in the previous assignment. Go over the slides (Part-6) on TS Forecast Validation and Evaluation and select all the answers that apply:

Select one or more:

- [x] a. According to RMSE (Root Mean Square Error) the best forecast is provided by the Seasonal Naive method. ✔
- [x] b. If MAE (Mean Absolute Error) or MAPE (Mean Absolute Percentage Error) are considered, they give the same result. ✔
- [ ] c. Only the metrics that are independent on the scale of the data should be considered when doing this comparative analysis of model performances.
- [ ] d. Due to the disadvantages of the percentage of error, the metrics that used this values should be ignored in this comparison.
- [ ] e. Due to dependency on the data scale, the MAPE but not RMSE and MAE should be considered for this model inter-comparison.
- [ ] f. The accuracy should have been computed for both the training and the testing data to make any comparative analysis of baseline models appropriate.

Your answer is correct.

The correct answers are: According to RMSE (Root Mean Square Error) the best forecast is provided by the Seasonal Naive method. , If MAE (Mean Absolute Error) or MAPE (Mean Absolute Percentage Error) are considered, they give the same result.

Question 3

Complete

Points out of 6.00

Write an R script that generates and plots ACFs for the following three random samples: (a) 30 random numbers; (b) 300 random numbers and (c) 1000 random numbers (hint: rnorm() and acf()). Explain the similarities and differences among the obtained figures.

```
require(fpp)
random30 <- rnorm(30)
Acf(random30)
random300 <- rnorm(300)
Acf(random300)
random1000 <- rnorm(1000)
Acf(random1000)
```

Similarities: All the plots show that the data is white noise

Dissimilarities: As the size of data increases the critical value for Acf plot decreases. All the plots show different critical Values.

* All figures indicate that the data is white noise.

* They show different critical values (blue dashed lines).

The critical values are at different distances from zero because the data sets have different number of observations. The more observation are in a data set representing white noise the less noise appears in the correlation estimations (spikes). Therefore the critical values for bigger data sets can be smaller in order to check if the data is not white noise.

Question 4

Complete

Points out of 5.00

The monthly number of people on unemployed benefits have been visually examined and the different variations at difference levels of the time series data have been observed. To stabilize this variance the Box.Cox transformation has been applied as depicted in the R snippet code below:

```
require(fpp)
plot(dole, xlab = "Time", ylab = "No. Unemployed",
     main = "Monthly Unemployed Benefit Usage, Astralia (01/56-07/92)")
lambda = BoxCox.lambda(dole)
plot(BoxCox(dole, lambda), xlab = "Time",
     ylab = paste("BoxCox(# people,", round(lambda, 2), ")"))
```

Examine the following data sets and determine if they require a similar variance stabilization. Submit the supporting R code:

- *usdeath*: Monthly total of accidental deaths in the United States (January 1973-December 1978)
- *bricksq*: Quarterly production of bricks (in millions of units) at Portland, Australia (March 1956-□□September 1994)

usdeath: doesn't require transformation

bricksq: requires similar variance stabilization

Q4.R

No transformation is needed for **usdeath but *bricksq* must be transferred similarly.**

Question 5

Complete

Points out of 5.00

Examine the time series data on the number of pigs slaughter in Victoria:

```
require(fpp)
data(pigs)
```

Submit the R code that checks whether the data is white noise or not. Use both the Acf and Portmanteau Tests.

Highlight your answer with the print statement that displays "White noise" if it is or "Not a white noise" if it is not.

This is not a white noise

```
require(fpp)

data(pigs)


#using acf

snaivedata <- snaive(pigs)

res <- residuals(snaivedata)

Acf(res)

#Portmanteau Tests

test<- Box.test (res, lag = 1, type = "Ljung")

if( test["p.value"] <= 0.05){

  print("Not a white noise")

}else{

  print("White noise")

}
```

Q5.R

This is NOT a white noise.

Question 6

Complete

Points out of 5.00

Examine the time series data on the number of pigs slaughter in Victoria:

```
require(fpp)
data(ausbeer)
beer <- window(ausbeer,start=1992)
```

Submit the R code that builds the seasonal naive forecast model, calculates the residuals from the **forecast applied to the quarterly Australian beer production data from 1992, and tests if the residuals are white noise using the Ljung-Box test (hint: type="Lj" in Box.test()).**

Highlight your answer with the print statement that displays "White noise" if it is or "Not a white noise" if it is not.

This is not a white noise

```
require(fpp)
data(ausbeer)
beer <- window(ausbeer,start=1992)
snaivedata <- snaive(beer)
res <- residuals(snaivedata)
#acf
Acf(res)
#Portmanteau Tests
test<- Box.test (res, lag = 1, type = "Lj")
if( test["p.value"] <= 0.05){
  print("Not a white noise")
}else{
  print("White noise")
}
```

Q6.R

beer <- window(ausbeer,start=1992)

fc <- snaive(beer)

res <- residuals(fc)

Acf(res)

Box.test(res, lag=8, fitdf=0, type="Lj")

Question 7

Correct

5.00 points out of 5.00

Sam visually examined the Dow-Jones index to assess whether it is a stationary time series data or not.

```
require(fpp)
data(dj)
plot(dj)
```

Select all that he should have confirmed to hold true about the time series to conclude it is stationary.

Select one or more:

- [x] a. Time series is almost horizontal ✔
- [x] b. Variance does not vary differently across different levels (i.e. stable variance) ✔
- [x] c. The means is stable ✔
- [ ] d. The ACF decreases slowly
- [x] e. No patterns predictable in the long-term ✔

Your answer is correct.

The correct answers are: Time series is almost horizontal, Variance does not vary differently across different levels (i.e. stable variance), The means is stable, No patterns predictable in the long-term

Question 8

Correct

3.00 points out of 3.00

Ann was asked to test if the data of interest is white noise. She first attempted to visually examine the auto-corellation plot and check is the autocorellation values are within the critical values for the given time series. She used the following snippet of the R code.

```
require(fpp)
data(econsumption)
acf(econsumption[, "Mwh"])
```

Compute the critical values and report the absolute value for this simple white noise diagnostics. Round the answer up to the fourth digit after the period.

Answer: 0.5658

critical.value <- 1.96 / sqrt(length(econsumption[, "Mwh"]))

The correct answer is: 0.5658

Question 9

Correct

5.00 points out of 5.00

Ann now built a simple linear regression model and decided to confirm that the residuals of the model are indeed white noise. Here is an initial snippet of R code she did.

```
require(fpp)
data(econsumption)
plot(Mwh ~ temp, data = econsumption)
fit = lm(Mwh ~ temp, data = econsumption)
errors <- residuals(fit)
plot(residuals(fit) ~ temp, data = econsumption)
abline(0, 0, col="grey")
```

Complete the rest of her code to validate both by using the Acf plot and by performing the Portmanteau test, namely Box-Pierce test (Box.test()).

Report the p-value for the χ^2 statistics; round your answer to the first decimal after the period.

Answer: 0.9

```
critical.value <- 1.96 / sqrt(length(econsumption[, "Mwh"]))
require(fpp)
data(econsumption)
plot(Mwh ~ temp, data = econsumption)
fit = glm(Mwh ~ temp, data = econsumption)
errors <- residuals(fit)
plot(residuals(fit) ~ temp, data = econsumption)
abline(0, 0, col="grey")
acf(errors)
Box.test(errors, lag=10, fitdf=0)
```

The correct answer is: 0.9

◄ (DUE: 04/26/2019): SUBMIT: QUIZ: BONUS: Time Series Decomposition and Smoothing

Jump to...

(DUE: 04/26/2019): SUBMIT: QUIZ: BONUS: TS Forecasting with ARIMA ►