

# Week 14 Self Assessment 9

**Due** Apr 22 at 11:59pm**Points** 10**Questions** 10**Available** Apr 10 at 8am - Apr 22 at 11:59pm 13 days**Time Limit** None**Allowed Attempts** Unlimited

## Instructions

Please provide answers to the following 10 questions. Each question is worth 1 point.

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
KEPT	<a href="#">Attempt 4</a>	2 minutes	10 out of 10
LATEST	<a href="#">Attempt 4</a>	2 minutes	10 out of 10
	<a href="#">Attempt 3</a>	less than 1 minute	6 out of 10
	<a href="#">Attempt 2</a>	less than 1 minute	3 out of 10
	<a href="#">Attempt 1</a>	1 minute	1 out of 10

❗ Correct answers are hidden.

Score for this attempt: **10** out of 10

Submitted Apr 10 at 5:10pm

This attempt took 2 minutes.

### Use the following for Questions 1-2:

Quickest Trippy is a local gas station. They want to predict demand for gasoline and have the following historical data.

Month	Demand (in thousands of gallons)	Forecast
April	12	

May	17	
June	20	
July	19	
August	24	

**Question 1****1 / 1 pts**

Using  $\alpha = 0.2$  and  $\delta = 0.4$  as well as  $F_1=11,000$  and  $T_1=2,000$  what would be the Trend Component predicted for week 2 ( $T_2$ )?

- ☒ (a) 1.92
- ☐ (b) 12.8
- ☐ (c) 2.10
- ☐ (d) 15.18

**Question 2****1 / 1 pts**

Using  $\alpha = 0.2$  and  $\delta = 0.4$  as well as  $F_1=11,000$  and  $T_1=2,000$  what would be the Forecast Including Trend for week 2 ( $FIT_2$ )?

- ☐ (a) 15.18
- ☐ (b) 12.8
- ☒ (c) 14.72

● (d) 17.28

You require “householdpower.csv”.

<https://www.dropbox.com/s/qc2bo8ok2dazjmg/householdpower.csv?dl=0>  
[\\_ \(https://www.dropbox.com/s/qc2bo8ok2dazjmg/householdpower.csv?dl=0\)](https://www.dropbox.com/s/qc2bo8ok2dazjmg/householdpower.csv?dl=0)

We will explore applying simple exponential smoothing in R. Some important information of the data is:

**Source:**

<http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption>

[\\_ \(http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption\)](http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption)

**Data Set Information:**

This archive contains 2075259 measurements gathered in a house located in Sceaux (7km of Paris, France) between December 2006 and November 2010 (47 months).

**Notes:**

- 1.(global\_active\_power\*1000/60 - sub\_metering\_1 - sub\_metering\_2 - sub\_metering\_3) represents the active energy consumed every minute (in watt hour) in the household by electrical equipment not measured in sub-meterings 1, 2 and 3.
- 2.The dataset contains some missing values in the measurements (nearly 1,25% of the rows). All calendar timestamps are present in the dataset but for some timestamps, the measurement values are missing: a missing value is represented by the absence of value between two consecutive semi-colon attribute separators. For instance, the dataset shows missing values on April 28, 2007.

**Attribute Information:**

- 1.date: Date in format dd/mm/yyyy
- 2.time: time in format hh:mm:ss
- 3.global\_active\_power: household global minute-averaged active power (in kilowatt)
- 4.global\_reactive\_power: household global minute-averaged reactive power (in kilowatt)

- 5.voltage: minute-averaged voltage (in volt)
- 6.global\_intensity: household global minute-averaged current intensity (in ampere)
- 7.sub\_metering\_1: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered).
- 8.sub\_metering\_2: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light.
- 9.sub\_metering\_3: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.

First we need to load and clean up the data set. We want daily data with the total active energy consumed (in watt hour). The steps to do this are

1. Find the total active energy consumed per minute (in watt hour) for each row (using the formula  $(\text{total\_active\_energy} = \text{global\_active\_power} * 1000 / 60 - \text{sub\_metering\_1} - \text{sub\_metering\_2} - \text{sub\_metering\_3})$ )
2. Group by Date
3. Find the total power in a day (by summing up value of active energy consumed per minute in each group)

### Question 3

1 / 1 pts

Observe the dataset. What is the beginning and end date of the dataset?

- ☐ (a) 12th December 2005, 19th September 2009
- ☒ (b) 16th December 2005, 10th September 2009
- ☐ (c) 22th December 2005, 20th September 2009
- ☐ (d) 30th December 2005, 16th September 2009

**Question 4****1 / 1 pts**

What is the average value of daily data consumed?

(NOTE: There are NA values present, so while finding the mean, use the attribute `na.rm=TRUE`)

☐ (a) 20152

☐ (b) 11029

☒ (c) 13294

☐ (d) 9269

Convert the dataframe to an xts object.

You may have noticed there are NA values in the dataset. We can take care of them by simply filling them with previous value ( use function `na.locf(df)` from the xts package)

Load the fpp2 package, which will help us with exponential smoothing.

Using an alpha of 0.35 and h=100, use the `ses()` function to fit the xts object.

- Plot the model using `autoplot()`
- Print the accuracy using the `accuracy()` function

**Question 5****1 / 1 pts**

What do you observe in the graph?

☐ (a) It is continuously decreasing

☐ (b) It is continuously increasing

☒ (c) It is moving up and down

☐ (d) It is a straight line

### Question 6

1 / 1 pts

What is the RMSE (Root mean squared error) of this fitted dataset?

☐ (a) 1637

☐ (b) 9626

☐ (c) 4592

☒ (d) 3974

### Question 7

1 / 1 pts

What is the MAE (Mean absolute error) of this fitted dataset?

☐ (a) 3104

☒ (b) 2608

☐ (c) 1826

☐ (d) 2048

Now you have the code to find the RMSE for a value of a particular alpha.

Now create a loop to find the RMSE of each alpha from 0.01, 0.02, 0.03 ... to 0.99

**Hint:** Use this code

```
alpha <- seq(.01, .99, by = .01)
```

```
RMSE <- NA
```

```
for(i in seq_along(alpha)) {
```

```
  accuracy_table = # add code to find the accuracy_table using accuracy()
```

```
  RMSE[i] <- accuracy_table[1,2]
```

```
}
```

```
alpha.fit <- data_frame(alpha, RMSE)
```

Now you have a dataset with values of RMSE for different value of alpha

### Question 8

1 / 1 pts

Find the minimum RMSE. What is the minimum value?

☐ (a) 4103

☐ (b) 3179

☐ (c) 3340

☒ (d) 3973

### Question 9

1 / 1 pts

What is the optimal value of alpha (which has minimum RMSE)?

☐ (a) 0.35

☐ (b) 0.12☐ (c) 0.56☒ (d) 0.22**Question 10****1 / 1 pts**

Plot the graph of RMSE vs alpha. Which of the following is **always** true?

☐ (a) Lower the alpha, lower the RMSE☐ (b) Higher the alpha, Higher the RMSE☐ (c) Lower the alpha, Higher the RMSE☒ (d) None of the above**Quiz Score: 10** out of 10