HW 2

February 18, 2021

1 IST 387 HW 2

[1]:
```
# Enter your name here: Connor Hanan
```

1.0.1 Attribution statement: (choose only one and delete the rest)

[2]:
```
# 1. I did this homework by myself, with help from the book and the professor.
```

1.0.2 Reminders of things to practice from last week:

Assignment arrow <- The combine command c( ) Descriptive statistics mean( ) sum( ) max( ) Arithmetic operators + - * / Boolean operators > < >= <= == !=

This Week: Explore the **quakes** dataset (which is included in R). Copy the **quakes** dataset into a new dataframe (call it **myQuakes**), so that if you need to start over, you can do so easily (by copying quakes into myQuakes again). Summarize the variables in **myQuakes**. Use these commands to get started:

[1]:
```
myQuakes <- quakes                # Copy into new data frame
summary(myQuakes)         # Summarize data in the console
head(myQuakes)                # View the data in a new tab in RStudio
```

```
      lat              long           depth            mag       
 Min.   :-38.59   Min.   :165.7   Min.   : 40.0   Min.   :4.00  
 1st Qu.:-23.47   1st Qu.:179.6   1st Qu.: 99.0   1st Qu.:4.30  
 Median :-20.30   Median :181.4   Median :247.0   Median :4.60  
 Mean   :-20.64   Mean   :179.5   Mean   :311.4   Mean   :4.62  
 3rd Qu.:-17.64   3rd Qu.:183.2   3rd Qu.:543.0   3rd Qu.:4.90  
 Max.   :-10.72   Max.   :188.1   Max.   :680.0   Max.   :6.40  
    stations    
 Min.   : 10.00  
 1st Qu.: 18.00  
 Median : 27.00  
 Mean   : 33.42  
 3rd Qu.: 42.00  
 Max.   :132.00  
```

A data.frame: 6 × 5

	lat <dbl>	long <dbl>	depth <int>	mag <dbl>	stations <int>
1	-20.42	181.62	562	4.8	41
2	-20.62	181.03	650	4.2	15
3	-26.00	184.10	42	5.4	43
4	-17.97	181.66	626	4.1	19
5	-20.42	181.96	649	4.0	11
6	-19.68	184.31	195	4.0	12

Step 1: Explore the earthquake magnitude variable called **mag** (To address these items, add a comment after the command that produces the result in your code.)

A. What is the average magnitude? Use mean() or summary():

[2]:
```
mean(myQuakes$mag) #use the column mag in myQuakes as the argument for mean()
```

4.6204

B. What is the magnitude of the largest earthquake? Use max() or summary() and save the result in a variable called **maxQuake**:

[4]:
```
maxQuake <- max(myQuakes$mag) #find max value of column mag, save as variable
maxQuake #print to see value
```

6.4

C. What is the magnitude of the smallest earthquake? Use min() or summary() and save the result in a variable called **minQuake**:

[6]:
```
minQuake <- min(myQuakes$mag) #find max value of column mag, save as variable
minQuake #print to see value
```

4

D. Create a **sorted dataframe** based on magnitude and store it in **quakeSorted**. **Hint:** Use order( )

[8]:
```
quakeSorted <- myQuakes[order(myQuakes$mag),] #sort dataframe by magnitude and␣
↪save as new dataframe
quakeSorted #print to check the sort
```

	lat	long	depth	mag	stations
	<dbl>	<dbl>	<int>	<dbl>	<int>
5	-20.42	181.96	649	4	11
6	-19.68	184.31	195	4	12
26	-17.94	181.49	537	4	15
34	-23.55	180.80	349	4	10
52	-19.26	184.42	223	4	15
58	-22.06	180.60	584	4	11
71	-15.31	185.80	152	4	11
85	-17.70	181.70	450	4	11
96	-19.73	182.40	375	4	18
113	-19.06	182.45	477	4	16
142	-20.65	181.40	582	4	14
150	-17.90	181.50	573	4	19
202	-17.70	182.20	445	4	12
236	-23.54	179.93	574	4	12
284	-17.70	185.00	383	4	10
298	-17.94	181.51	601	4	16
299	-30.64	181.20	175	4	16
362	-16.90	185.72	135	4	22
389	-10.72	165.99	195	4	14
433	-18.55	182.23	563	4	17
483	-22.70	183.30	180	4	13
533	-21.00	183.20	296	4	16
598	-17.02	182.93	406	4	17
637	-19.51	183.97	280	4	16
698	-15.43	185.19	249	4	11
722	-17.91	181.48	555	4	17
727	-17.10	182.80	390	4	14
733	-30.30	180.80	275	4	14
750	-25.60	180.30	440	4	12
A data.frame: 1000 × 5 770	-20.70	186.30	80	4	10
531	-15.77	167.01	64	5.5	73
541	-15.90	167.42	40	5.5	86
663	-18.14	180.87	624	5.5	105
893	-13.80	166.53	42	5.5	70
948	-27.89	182.92	87	5.5	67
952	-35.94	178.52	138	5.5	78
167	-26.00	182.12	205	5.6	98
297	-24.57	178.40	562	5.6	80
570	-34.68	179.82	75	5.6	79
636	-18.82	182.21	417	5.6	129
649	-37.03	177.52	153	5.6	87
653	-11.40	166.07	93	5.6	94
712	-15.93	167.91	183	5.6	109
920	-17.85	181.44	589	5.6	115
935	-20.25	184.75	107	5.6	121
109	-22.55	185.90	42	5.7	76
151	-23.34	184.50	56	5.7	106
176	-32.22	180.20	216	5.7	90
275	-22.13	180.38	577	5.7	104
376	-15.33	186.75	48	5.7	123

Step 2: Explore the **stations** variable

E. **Write a comment:** Does there appear to be a relationship between magnitude and the number of reporting stations?

[ ]:
```
#Generally, it seems that there is a relationship between magnitude and number␣
↪of stations where
#the larger the magnitude of the quake, the more stations report the quake.
```

F. What are the latitude and longitude of the quake reported by the largest number of stations? **Hint:** Use which.max( )

[24]:
```
quakeSorted[which.max(quakeSorted$stations),1:2] #take row of dataframe whose␣
↪index matches the result from which.max() and return with columns 1 and 2␣
↪(lat and long)
```

A data.frame: 1 × 2

	lat	long
	<dbl>	<dbl>
870	-12.23	167.02

G. What are the latitude and longitude of the quake reported by the smallest number of stations? **Hint:** Use which.min( )

[25]:
```
quakeSorted[which.min(quakeSorted$stations),1:2] #take row of dataframe whose␣
↪index matches the result from which.min() and return with columns 1 and 2␣
↪(lat and long)
```

A data.frame: 1 × 2

	lat	long
	<dbl>	<dbl>
34	-23.55	180.8

Step 3: Using conditional if statements H. Test if **maxQuake** is greater than 7 (output "yes" or "no") **Hint:** Try modifying the following code in R:

[34]:
```
if  (maxQuake>7) "yes" else "no"
maxQuake
```

'no'

6.4

I. Following the same logic, test if minQuake is less than 3 (output "yes" or "no"):

[35]:
```
if (minQuake<3) "yes" else "no"
minQuake
```

'no'

4

[ ]:
```

```