QSCI 381 Homework #1 - Vinsensius

Number 1

1a

Given the information shown in the table, and your reading of the information about this data, is travel distance a quantitative or qualitative variable, AND what level of measurement does it belong to AND explain why (4 pts)

Answer: Travel distance of bird is a quantitative data because it is numerics. The level of measurement is ratio because there is an absolute zero within the measurement which means the bird does not travel to anywhere and stay where they are.

1b

Separately for each of the three birds find the shortest daily travel distance. This is known as the minimum of the data (3 pts)

Answer:

- The minimum travel distance for BirdID_1 is 13.90 km.
- The minimum travel distance for BirdID_2 is 9.00 km.
- The minimum travel distance for BirdID_3 is 10.80 km.

1c

Separately for each of the three birds find the longest daily travel distance. This is known as the maximum of the data (3 pts)

Answer:

- The maximum travel distance for BirdID_1 is 162.30 km.
- The maximum travel distance for BirdID_2 is 218.20 km.
- The maximum travel distance for BirdID_3 is 245.60 km.

1d

The difference between the maximum and minimum values in a data set is known as the range. What is the range in distances travelled for each bird (3 pts)?

Answer:

- The range travel distance for BirdID_1 is 148.40 km.
- The range travel distance for BirdID_2 is 209.20 km.
- The range travel distance for BirdID_3 is 234.80 km.

1e

For each bird calculate the mean, the standard deviation, and coefficient of variation of distance travelled by each bird (Note: round your answer to 2 decimal places) (9 pts)

Answer:

BirdID	**Mean** (in km)	**Standard Dev.** (in km)	**Coeff. Variation**
1	53.16	29.17	0.55
2	49.28	40.84	0.83
3	65.95	43.11	0.65

1f

Given your answer in (1e) which bird had the most variable travel distances? (2 pts)

Answer:

The second bird has the most variable travel distances because it has the highest coefficient of variation.

Number 2

2a

Create a line graph of travel distances through time for bird ID_2, plotting daily travel distance against date. When creating the plot add the following plot attributes by specifying the necessary arguments in the plot function (for more help with the plot function type ?plot in R):

- The default plot type in R is a scatterplot, we want a line graph instead. To do this, use the type="l" option.
- Label the x axis, using the xlab argument, "Date"
- Label the y axis, using the ylab argument, "Daily travel distance (km)"
- Change the y-axis range using the ylim argument so that the axis ranges from 0 to 250
- Add a plot title using the main argument. Set the plot title to be a brief description of the information plotted followed by your name in parentheses
- Change the line width using the lwd function to 2

Take a screenshot of your graph and paste it below (6 pts)
Answer:

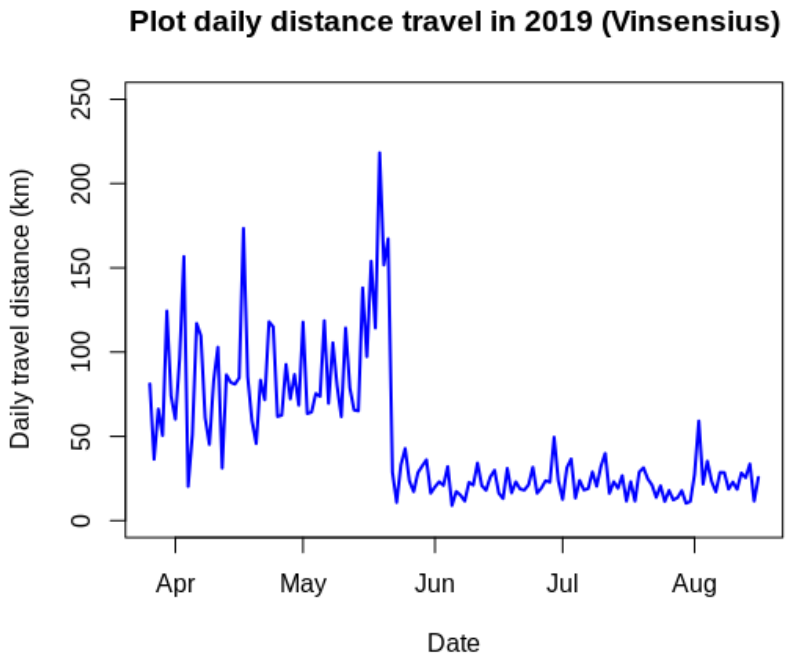


Figure 1: Plot of travel distance of BirdID_2

2b

Now overlap the distances travelled by Bird ID_3 onto this figure using the lines() function, changing the color of the Bird ID_3 line to distinguish it from Bird ID_2, but keeping the same linewidth as before.

Once the 2nd line is added, add a legend to tell the viewer which line belongs to which bird using the legend() function. (4 points)

Answer:

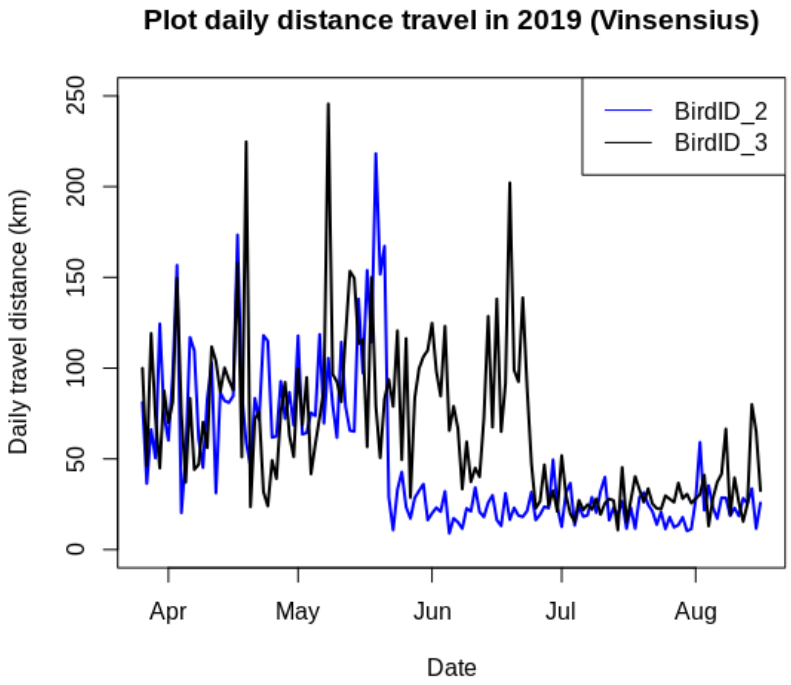


Figure 2: Plot of travel distance of BirdID_2 and BirdID_3

2c

Comparing the distances travelled, provide a brief description of any major differences you see between the two patterns through time (2 pts)

Answer:

BirdID_2 in general have shorter travel distances to BirdID_3 except in very few instances in which they are flying much farther than BirdID_3. This is supported from the answer in number 1 which shows that the mean daily distance flown by third bird is higher than the second bird. The maximum daily distance that the second bird flew was at the end of May while the maximum daily distance that the third bird flew was at the start of May. Furthermore, the second bird has longer period of short travel distance (below 50 km) than the third bird. The second bird daily distance travel starting June and it did not seem to increase to increase at the end of measurement after August, except at the start of August. For the third bird, the daily travel distance began to drop in July, but began to climb back up in August.

Number 3

3a

Use Sturge's Rule to come up with class widths for the histogram of travel distances recorded for Bird ID_1.

Sturge's rule: Bin width = $\frac{x_{\max} - x_{\min}}{1+1.44 \cdot ln(n)}$

Note: try the nrow function to tell you the size of the dataset, and the log function in R returns the natural log ln.

Write your answer below, including your working. Round your answer to the nearest whole unit (4 pts)

Answer:

- $x_{\max} = 162.30$ km
- $x_{\min} = 13.90$ km
- n = 144

Thus, the bin's width is 18 km.

3b

Using the class width (nearest whole unit) you got in 3a calculate the frequency distribution of travel distances made by Bird ID_1

Step 1: Find the class limits (set the lowest limit at 0) using the class width

Step 2: Based on those classes, find the frequency of observations in each class using the following snippet, replacing LOWERLIMIT and UPPERLIMIT with the values for each class

nrow(BirdMove[BirdMove$BirdID_1 >= LOWERLIMIT & BirdMove$BirdID_1 < UPPERLIMIT,])

Fill in the table below with information on class limits, midpoints and frequencies (make sure to check that all the frequencies sum to the total size of the dataset) (5 pts)

Answer:

Lower Limit (in km)	**Upper Limit** (in km)	**Midpoint** (in km)	**Frequency**
0	18	9	5
18	36	27	43
36	54	45	38
54	72	63	25
72	90	81	15
90	108	99	10
108	126	117	4
126	144	135	2
144	162	153	1

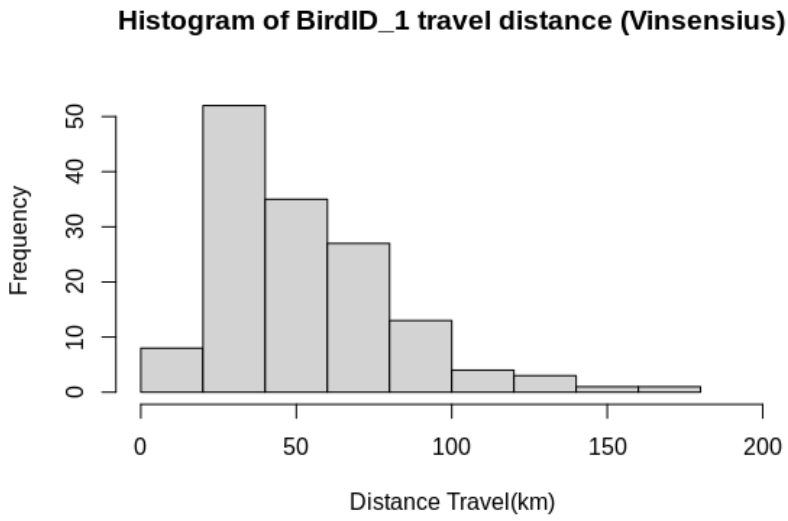
3c

Now, make a histogram of the travel distances made by Bird ID_1 using the hist function.

Label the x-axis label with the appropriate variable name and include your name as the histogram title

Paste your graph below (4 pts)

Answer:



3d

Based on the frequency distribution and the histogram, would you say the distribution of daily travel distances made by Bird ID_1 was symmetric, uniform, left-skewed or right-skewed, justifying your answer based on observed features. (3 pts)

Answer:

Based on the histogram, the distribution of travel distances Bird ID_1 is right-skewed because of the low frequency at high values which creates "tail". Based on the frequency distribution, we know that the mode is centered at value of 27 km, which is lower than the mean of the dataset, which is 53.16 km.