E01.2

Check the Tableau Public tutorial video "1. Tableau Public Overview (7:10)" from https://public.tableau.com/en-us/s/resources . Following the steps from the video, create a similar visualization using Tableau Public. You can load the Excel file which is the dataset either from the video page, or from Moodle (check the Examples and Data folder in Moodle)

Tableau data visualisation example here

E01.3

Check the Tableau Public tutorial video "13. Understanding the Logic of Charts (5:05)" from https://public.tableau.com/en-us/s/resources . Following the steps from the video, create a similar visualization using Tableau Public.

Tableau data visualisation example here

E01.4

Answer questions related to the Iris dataset.

1. What is the data all about?

   - The dataset was introduced by biologist and statistician Ronald Fisher, which contains data 50 samples from 3 species of the Iris flower. Iris setosa, Iris virginica and Iris versicolor

2. How big is the dataset? (Number of rows or instances.)

   - 150 Rows (instances)
   - 5 columns (attributes)

3. How was the data collected?

   - The data measures 4 features of the flower in cm.
     - Sepal Length
     - Sepal Width
     - Petal length
     - Petal width

4. What does an individual dataset row mean?

   - An individual dataset row comprises the features of the given species of the flower.

5. What do the attributes (columns) mean?

   - the attribute columns also the attributes shows the details of the measurements of the species.

6. (Based on this example:) How is data encoded in .csv files?

- The data is encoded using the UTF-8 unicode format.

7. Which use cases you can find for the dataset, from the wikipedia article?

   The data is used in cluster analysis to explain the differences between supervised and unsupervised techniques in data mining.

8. How precise you think the data is? What kinds of errors might occur in a dataset like this, collected like this?

   Although a correct estimate is hard to conclude but the data's precision is reliant on a few factors. Such as, quality of the flower or region in which it was harvested from as the physical attributes may differ depending on the location.

E01.5

Create a visualization using the Iris dataset.

Link to the comparison graphs of average attribute values and petal length distribution can be found here.

One of the challenges faced when creating the data was understanding the bin concept within tableau