

In this exercise sheet you'll be using some random data that I've generated for carrot and onion price variations over one year. Please be aware that these data points are just simulated and, although the prices ranges are meant to be comparable to the actual ranges, they don't actually show any seasonal trends you may expect.

The units of the price can be interpreted as either Euros/Dollars (any type of Dollar), meaning that a value of 1 corresponds to one of that unit, and a value of 0.01 corresponds to 1 cent.

## 1 Exercise 1 - Histograms

Histograms provide a nice visual representation of the range covered by the data as well as how noisy it may be. Looking at histograms can be a very useful way to get a quick understanding of the nature of your data.

For this task choose which dataset (onion vs. carrot) you want to use to create a histogram out of.

1. Create a figure of desired size and add a subplot to it.
2. For your dataset create a new list for rounded values, which contains the vegetable's price for each date rounded to the nearest cent (i.e. 0.01).
3. Plot a histogram distribution of the range of prices for your vegetable.
4. Adjust the bins so that the lowest bin is 1 cent below the minimum price and the highest bin is 1 cent above the maximum price, with bin sizes being equal to one cent.
5. Align the bins so that the center of the bin is directly above the corresponding price
6. Repeat step 2. & 3. from here for the other vegetable's price, and add its price histogram onto your already existing histogram.
7. Change the bin ranges to capture the price ranges of both vegetables, and change the bin sizes to be 5 cents per bin.
8. Change the color of each histogram to reflect the color of the vegetable used in the previous exercise sheet.
9. Add a title and labels to your graph and perform any extra formatting you may want to.

## 2 Exercise 2 - Advanced Histograms

This task is going to build up on the previous exercise, but we're going to add some more customizations to it which will allows us to easier point out important features

1. To the histograms created in Exercise 1, add a legend for each histogram.
2. For each histogram, change the color of the bar that contains the mean value to blue/green for the carrot/onion dataset respectively.
3. Add to the legend a label for each of the mean values shown above, so that these values can easily be understood.

### 3 Exercise 3 - Bar Graphs

Bar graphs are a great way to represent single variables, and plotting multiple bar graphs allow for very fast and easy comparisons.

1. Choose one of the two vegetables and create a new list where each element is the average price of that vegetable for that month.
2. Create a bar graph that plots each of the averaged prices next to each other, in order of increasing month from left to right.
3. Adjust the width of the bars to make them a bit thinner, and change the alignment of them so that each bar is right on top of the major axis tick.
4. Now add on the other vegetable's average price for each month onto the same bar graph, stacking them so that each bar ends at the average price for that vegetable for that month. Make sure that both bar graphs have the same width.
5. Adjust the color of each bar graph to represent the color of each vegetable.
6. Add an appropriate title, axes labels, axis tick names, and a legend to the graph.

### 4 Exercise 4 - Advanced Bar Graphs

To make the bar graphs even better visualizations, we will add further developments to them. Exercise 4 builds upon the results produced in Exercise 3.

1. For each month for each vegetable, find the minimum and maximum price the vegetable took on that month, and save each appropriately.
2. For each vegetable's bar graph in each month, add in error bars that reflect the minimum and maximum price of that vegetable for that month. Color the error bars black.
3. To reduce clutter, instead of plotting the bar graphs stacked, adjust the graphs so that the bar graph for each vegetable is plotted next to each other symmetrically around the major axis tick.

### 5 Exercise 5 - Box and Whisker Plots

Box and whisker plots can give us a statistical overview of our data similar to how histograms can, but instead of visualization the distribution of the data points it gives us the visualizations of some of the main statistical variables of the data.

1. For each vegetable, create a box and whisker plot and show them next to each other on the same graph.
2. Change the whiskers so that the ends of the whiskers go to the 10th and 90th percentile of each vegetable's price.
3. Add a notch to the box and whisker plot for each vegetable.
4. Adjust the confidence interval of each vegetable's box and whisker plot to be 5 cents in either direction around the median value.