

Visualizing Data

Model 1 Simple Plot

When analyzing data, it's helpful to create charts, plots, and other visualizations. Doing so allows you to see important numerical relationships. Enter the following code into a Python Editor, and run the program.

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 def model_one():
5     x = np.arange(0.0, 2.0, .01)
6     y = np.sin(2 * np.pi * x)
7     plt.plot(x, y)
8     plt.xlabel('time (s)')
9     plt.ylabel('volts (mV)')
10    plt.show()
11
12 model_one()
```

Questions (15 min)

Start time:

1. Identify in the source code which line numbers:

- | | |
|-----------------------------|-----------------------------|
| a) generated the data? | c) displayed the window? |
| b) set the axes properties? | d) plotted the actual data? |

2. Describe in your own words what is being plotted.

3. Modify the code to plot only one cycle of the sine wave (instead of two). Write the edited line of code below.

4. Change the third argument of `np.arange` from 0.01 to 0.15. What is the result?

5. Add "o" as a third argument to the plot function. What is the result?
6. How does the third parameter of `np.arange` affect how the plot looks?
7. How would you modify the code to plot the function $y = x^2 - 1$ instead? Show the results from -2 to +2.
8. Which three Python libraries are used in Model 3? Quickly search the Internet and find their websites. Write a one-sentence description about each library.

Model 2 Histograms

Recall that you can generate a sequence of numbers using the `random` module. Run the following program, and view the output.

```
1 import matplotlib.pyplot as plt
2 import random
3
4 def model_two(npts):
5     numbers = []
6     for _ in range(npts):
7         numbers.append(random.random())
8     plt.hist(numbers)
9     plt.show()
10
11 model_two(100)
```

Questions (10 min)

Start time:

9. Based on the Python code:

- a) What is the range of values generated by the random function?
- b) How many random values are generated?

10. Based on the figure plotted:

- a) How many bars are displayed?
- b) What is the width of each bar?
- c) What is the sum of the heights of the bars?

11. Based on your answers above, what are appropriate labels for the x and y axes?

12. Increase the argument of `model_two` to 1000, 10000, and 100000. Describe how the output plot changes when you run the program.

13. Add the number 50 as second argument to the `hist` function. What is the meaning of the result?

14. In general, describe what the `hist` function does with the list of random numbers to create this type of plot.

Model 3 CSV Data

Recalls that “Comma Separated Values” is a common file format when exporting data from spreadsheets and databases. Each line of the file is a row, and each column is separated by a comma. Cells that contain commas are wrapped in quote marks.

data.csv file contents:

```
Name,Location,URL,Students
Westminster College,"Salt Lake City, UT",westminstercollege.edu,2135
Muhlenberg College,"Allentown, PA",muhlenberg.edu,2330
University of Maine,"Orono, ME",umaine.edu,8677
James Madison University,"Harrisonburg, VA",jmu.edu,19019
Michigan State University,"East Lansing, MI",msu.edu,38853
```

Python includes a csv module (<https://docs.python.org/3/library/csv.html>) that makes it easy to read and write CSV files.

```
import csv
```

```
infile = open("data.csv")
data = csv.reader(infile)
names = next(data) # column names
for row in data:
    print(row[1]) # 2nd column
```

Program output:

```
Salt Lake City, UT
Allentown, PA
Orono, ME
Harrisonburg, VA
East Lansing, MI
```

Questions (20 min)

Start time:

15. In the example data.csv file above:

- a) In what way is the first line different?
- b) How many rows of data are there? How many columns?

16. Compare data.csv with the program output:

- a) Are quote marks included in data.csv? In the program output?
- b) What is the purpose of the quote marks?

17. In the Python code above:

- a) Which line of code reads the first line of the file?
- b) What type of data does the variable row contain?