Line Graphs and Time Series: Takeaways

Syntax

- Importing the pyplot submodule:

```
import matplotlib.pyplot as plt
```

- Plotting a line graph:

```
plt.plot(x_coordinates, y_coordinates)
plt.show()
```

- Changing the scientific notation to plain notation:

```
plt.ticklabel_format(axis='both', style='plain')
```

- Adding a title and axes labels:

```
plt.title('New Reported Cases By Month (Globally)')
plt.xlabel('Month Number')
plt.ylabel('Number Of Cases')
```

- Plotting three graphs that share the same axes:

```
plt.plot(x_coordinates_1, y_coordinates_1)
plt.plot(x_coordinates_2, y_coordinates_2)
plt.plot(x_coordinates_3, y_coordinates_3)
plt.show()
```

- Adding a legend:

```
plt.plot(x_coordinates_1, y_coordinates_1, label='label')
plt.plot(x_coordinates_2, y_coordinates_2, label='label')
plt.legend()
plt.show()
```

- Plotting two graphs separately:

```
plt.plot(x_coordinates_1, y_coordinates_1)
plt.show()

plt.plot(x_coordinates_2, y_coordinates_2)
plt.show()
```

Concepts

- There are two kinds of data visualization:
    - Exploratory data visualization: We build graphs for *ourselves* to explore data and find patterns.
    - Explanatory data visualization: We build graphs for *others* to communicate and explain the patterns we've found through exploring data.

- The horizontal and vertical lines that intersect to make a graph are called **axes**. The horizontal line at the bottom is the **x-axis**, and the vertical line on the left is the **y-axis**. The point where the two lines intersect is called the **origin**.
- The two numbers that represent the distances of a point from the x- and y-axis are called **coordinates**. Point *A* above has two coordinates: seven and two. Seven is the **x-coordinate**, and two is the **y-coordinate**.
- A series of data points that is listed in time order is called a **time series**. To visualize a time series, we can use a **line graph**.
- We learned four types of growth associated with time series:
    - Linear: The growth occurs at a constant rate.
    - Exponential: The growth starts off slow but accelerates over time, becoming faster and faster.
    - Logistic: The growth starts slow, increases rapidly for a period, and then slows down again, eventually plateauing at a 'carrying capacity.'
    - Logarithmic: The growth starts off at a moderate pace and then decelerates over time. The quantity continues to increase indefinitely, but the increments become increasingly smaller.
- These four types of change can also decrease:
    - Linear: The decrease occurs at a constant rate.
    - Exponential: The decrease starts off slow but accelerates over time, becoming faster and faster.
    - Logistic: Initially, the decrease is rapid, but it eventually slows down and approaches a lower limit or 'carrying capacity.'
    - Logarithmic: The decrease is rapid initially and then starts to slow down considerably. While the quantity continues to decrease, the rate of decrease becomes almost negligible over time.
- In practice, most of the line graphs we plot don't show any clear pattern. We need to pay close attention to what we see, and try to extract meaning from the graphs without forcing the data into common patterns we already know.

Resources

- Anatomy of a graph
- A pyplot tutorial from matplotlib
- A short article on line graphs by The Data Visualization Catalogue
- A fun and useful tutorial on graphing equations