Nitty-gritty of Data and Exploratory Analysis with Python 3

Dev Skill
Class 1

Jayanta Sarker Shuva

Course Plan:

- **Advanced python3 programming:**
  Python data structures, List comprehension, Lambda, Map, Filter, Reduce functions and handling files. (3 classes)

- **Data scrapping:**
  We will learn to scrape data from different websites using Urllib request and beautifulsoup. We will also learn to scrape data using JSON based API. (1/2 CLass)

- **Data Analysis with Pandas/Numpy :**
  We will learn all aspects of data analysis using pandas and numpy. We will cover topics such as Data input and validation, Series and Data Frames, indexing, Group by, stack-unstack, reshaping and Analysis. (3 classes)

- **Exploratory Data analysis:**
  Here we will learn how to convert data into insights by converting them attractive visuals using matplotlib, pandas, seaborn etc. Here we will not only learn how to make charts and visuals; we will also learn the principles of information visualization.

  We will be introduced to tools for thinking about design and graphical heuristics for thinking about creating effective visualizations. We will learn how to discard noises and misleading information from charts also increase the truthfulness property of any visual. (3 classes)

- **Introduction to Machine Learning:**
  In this section, we will introduce you to the world of machine learning and learn how to implement basic types of supervised learning algorithm (regression and classification) using pandas, numpy, scikit-learn and matplotlib. (5/6 classes)

Introduction to Advanced Python3:

Necessary software:

- Anaconda python distribution. (https://www.anaconda.com/download)
- Google Colab. (https://colab.research.google.com).

Python Data Structure:

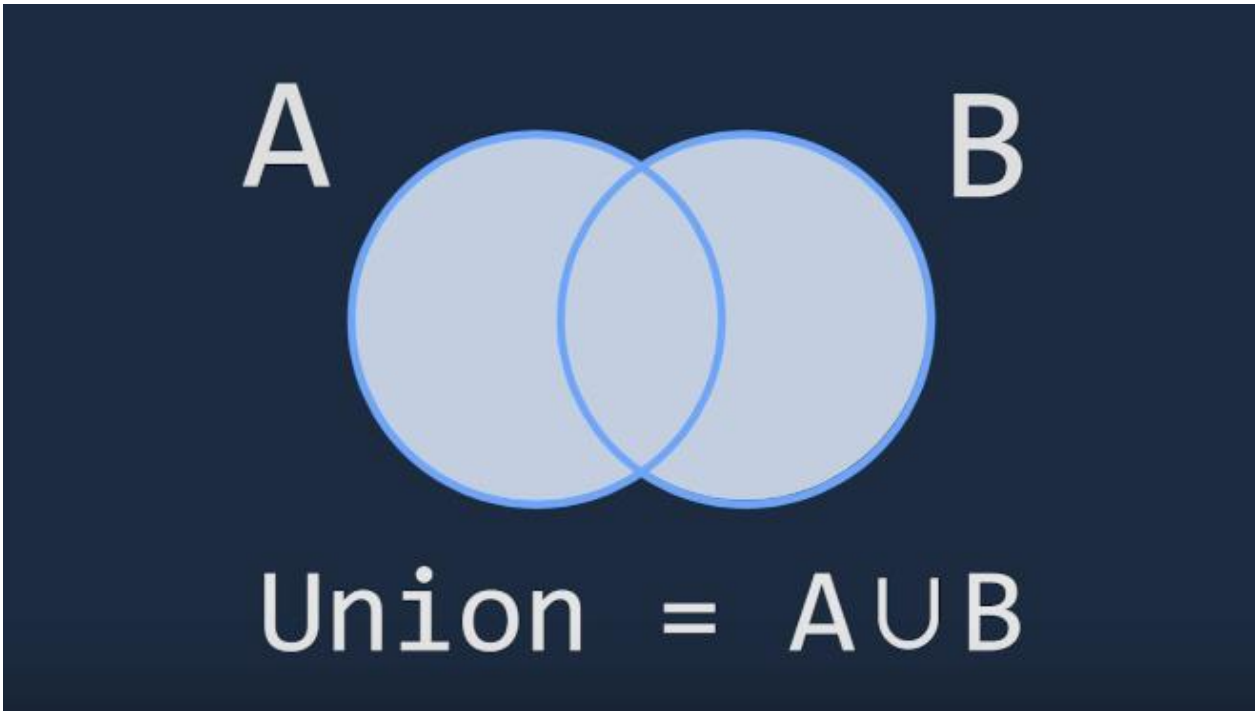
Sets:

A set is created by placing all the items (elements) inside curly braces {}, separated by comma or by using the built-in function set().

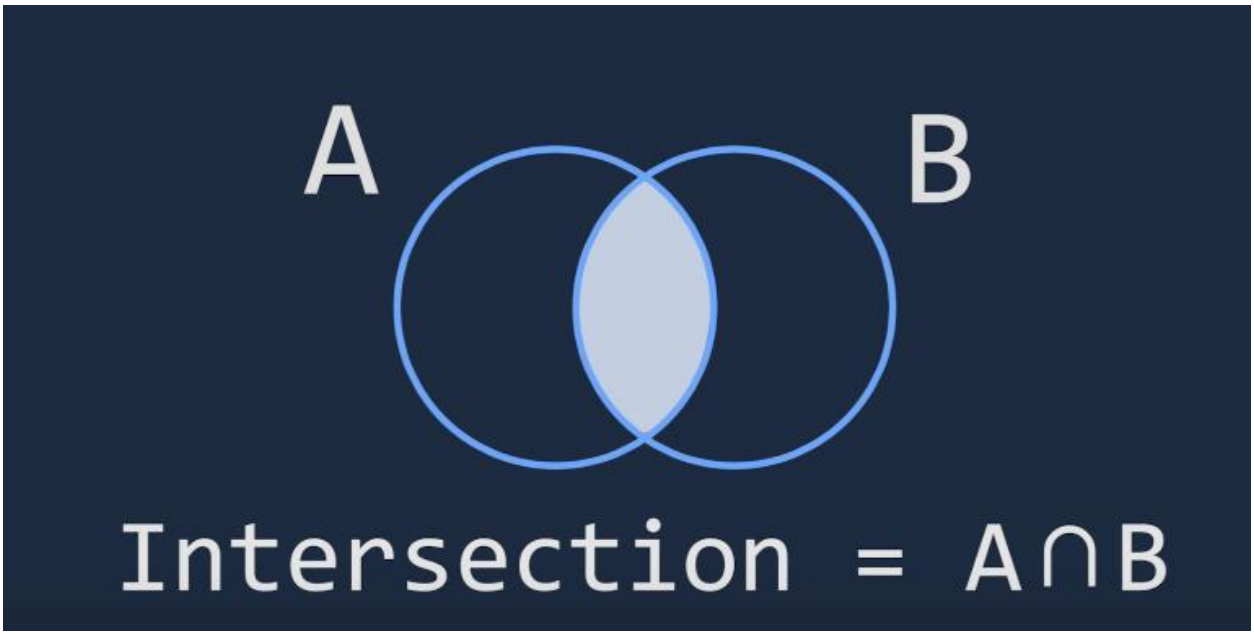
You can store all kinds of data types in same set like - in a single set you can store int, float, Boolean, string or another set whatever you want. In Set you cannot store duplicate value. For sets order does not matter. It means it doesn't matter in what order you put the values. In set it can be showed in different order than you inputted order.

There are 2 fundamental use of Sets:

Union:



Intersection:



List:

The list type is a container that holds a number of other objects, in a given order. (**Order matters**) Creating a list is as simple as putting different comma-separated values between square brackets or using the 'list ()' constructor. However, the most common way is using the square brackets. In python, you can store multiple data types. E.g. - L = [integer, float, Boolean, string, list]

['aam','jaam','kathal','komola','narikel','lichu']

0	1	2	3	4	5
Aam	Jam	kathal	komola	narikel	lichu
-6	-5	-4	-3	-2	-1

Slicing:

If we want to print the items of position 2nd, 3rd and 4th then we can slice the list as [2:5]. Note that the starting position is just as we write but the ending position we have to write the position we want + 1. So, though we want data until 4th position. We write until 5.

Dictionary:

A common structure in Computer science is associative array or Map. In python, it is called Dictionary. It is nothing but a key-value pair. You can imagine input as key and output as value. We can create dictionaries by two ways. By using Carli braces {} or using dict() constructor. Dictionary are not ordered data.

{A : 2, B : 'Jayanta', C : true}

Tuple:

Tuple is a sequenced data packed with first bracket (). Its behavior are almost same as list but it has lower memory size than because it has less available methods than lists.

Cannot be Changed (Immutable)
You can make it more quickly than List

Loops:

Loops are traditionally used when you have a block of code, which you want to repeat a fixed number of times. There are mainly two types of loops in python3.

1. For Loop
2. While Loop