# Exercise 5 - Introduction to Pandas

## Pandas exercise

# import necessary libraries  
import numpy as np  
import pandas as pd

### Series

Create a series from the below list

attempts = [1, 3, 2, 3, 2, 3, 1, 1, 2, 1]

Find the first person’s number of attempts

Find the first 5 people’s number of attempts

Find the last 3 peoples number of attempts

### DataFrames

Create and display a DataFrame from the below dictionary and list of index labels.

exam\_data = {'name': ['Anne', 'Kofi', 'Ridhwaan', 'Zara', 'Muhammad', 'Donald', 'Isabella', 'Sarah', 'Paula', 'John'],  
'score': [12.5, 9, 16.5, None, 9, 20, 14.5, None, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
  
labels = ['Student ' + letter for letter in ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']]

Display the first 3 rows of the data frame.

Display the name and score columns of the data frame.

Display the name and score columns and rows 1, 3, 4 and 8.

Display the rows where the number of attempts is greater than 2.

Display the rows where score is greater or equal than a pass rate of 10.

Display the rows where the score is above 15 and the student has only made one attempt.

### Reading in Data

Read the file mortgage\_applicants.csv, which sits in the data folder, into a variable called mortgage.

**Hint: Your path will need to reflect the location of the file**

The same data sits in an excel file called mortgage\_applicants.xlsx in a sheet called PrevYear. Read that into another DataFrame called mortgage\_excel.

Some weather data is held in a file called weather\_data.json. Read it into a dataframe called weather.

### Changing Types & Parsing Times

What type has each column in mortgage been read in as? Use a method to find out.

Convert the ID column so that it is instead represented as a Unicode string

Convert the day column of the weather Dataframe to an appropriate type

### Retrieving Rows and Columns

Select the Score column of the mortgage DataFrame

Select the Balance and Income columns

Select the first row in the mortgage DataFrame

What is the Debt of the first applicant?

What is are the Balance and Income of the 10th to 20th applicants

What are the final 5 peoples Score

Using weather, display the weather on the most recent day

### Querying DataFrames

Compute the monthly earnings of mortgage applicants, just using Income

Compute the ratio of debts to assets for the mortgage applicants

Display all mortgage applicants who have a Balance greater than £1000

Display all mortgage applicants who have a Balance greater than £1000 and a Debt below £50

Display all loan applicants who have an Income greater than 30,000 who have a 10-year loan, and those with an Income greater than 20,000 who have a 20-year loan (together!)

### Aggregating DataFrames

Compute the average Balance of mortgage applicants depending on the Term of their loan

Compute the average Income of mortgage applicants depending on whether they defaulted or not

Compute the average Debt, Income, and Balance of mortgage applicants based on whether they defaulted or not, and the term of their loan

Add a column to the DataFrame called MeanTermIncome, which contains the mean Income of mortgage applicants based on their Term.