Readings

All in one readings:

- https://www.tutorialspoint.com/python_pandas/index.htm
- https://www.geeksforgeeks.org/pandas-tutorial/
- https://www.javatpoint.com/python-pandas
- https://data-flair.training/blogs/python-pandas-tutorial/
- *https://pandas.pydata.org/pandas-docs/stable/getting_started/intro_tutorials/index.html*

Video tutorials:

- https://www.youtube.com/watch?v=ZyhVh-qRZPA&list=PL-osiE80TeTsWmV9i9c58mdDCSskIFdDS

For books consult this site:
I recommend Python for Data Analysis by Wes Mickenny

- https://data-flair.training/blogs/python-pandas-books/

Specific Readings:

Difference Between NaN and None:

- https://stackoverflow.com/questions/17534106/what-is-the-difference-between-nan-and-none#:~:text=NaN can be used as a numerical value,or "empty" than "numerically invalid" in this context.

Reading excel file in python :

- https://www.journaldev.com/33306/pandas-read_excel-reading-excel-file-in-python

Understanding map() function to manipulate series

- https://towardsdatascience.com/understand-map-function-to-manipulate-pandas-series-8ac340d514f7

Assignment

1. Import pandas and alias it as pd.
2. Create a series from a list, numpy array and dict containing at least 10 elements.
3. Provide index to all the rows of the series created.
4. Perform operations on the created series
    a. Add a row to the series
    b. Select a 4th row from the series
    c. Slice the series
        i. Select first three elements
        ii. Select the last three elements
        iii. Select the elements from 5th row to the 7th row
5. Create a DataFrame consisting of 5 rows and 5 columns using series and dict
6. Perform operations on the created DataFrame
    a. Select the 3rd row of the DataFrame using both index value and integer value
    b. Select the 4th column of the DataFrame using both index value and integer value
    c. Slice the DataFrame
        i. Select first three columns
        ii. Select first three rows
        iii. Select first three columns and first three rows
        iv. Select from the 3rd row to the fifth row
        v. Select from the 3rd column to fifth column
        vi. Create a new DataFrame consisting of two particular columns by their column names.

Manipulation of data in the provided csv files

1. Read the wine-quality(white) csv file
2. Select the chlorides and density column
3. Select the columns from density to quality using both loc and iloc
4. Select the last three rows index using both loc and iloc
5. Drop Columns chlorides and total sulfur dioxide
6. *Make a new column named Total acids and fill its values with the sum of fixed acidity, volatile acidity and citric acid*

7. *Merge two csv files provided using pandas.merge(Check this part of documentation incase of confusion : https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.merge.html)*

Read the excel file provided and perform all the operation indicated in the 6th number.

Use map() and perform mapping to the contents of the dataframe you created.