



Sheet 2

Python Libraries - Pandas

Questions

1. How can you create a DataFrame in Pandas from a Numpy array or a Python list?
2. How can you perform operations on DataFrame columns and rows in Pandas?
3. How can you sort a DataFrame in Pandas?
4. How can you drop missing values from a DataFrame in Pandas?
5. How can you perform groupby operations on a DataFrame in Pandas?
6. How can you merge two DataFrames in Pandas based on a common column or index?
7. How can you calculate the mean, median, and mode of a Pandas DataFrame column?
8. What is the difference between `iloc` and `loc` in Pandas?

1 Coding Questions

1.
 - Generate a pandas data frame contains 10 entries with the following columns: Name [A - Z], Age [7 - 90], Gender [M / F], and City [A - Z] and fill it with randomly generated *Numbers* or *Alphabetical Characters* as shown with brackets.
 - Create a new column called Category that assigns each person to a category based on their age: "Child" for ages under 18, "Adult" for ages 18-65, and "Senior" for ages over 65. Then, count the number of people in each category and display the result.
2.
 - Generate a pandas data frame that contains 10 rows of information about different movies, including the title [A - Z], genre["Comedy" / "Drama" / "Action"], year [1990 - 2010], and rating [1 - 10].
 - Create a new data frame that contains only the movies from the year 2000 or later and with a rating of 7.0 or higher. Then, sort the data frame by rating in descending order and display the top 10 movies.



Deliverable

- This sheet is to be solved **Individually**.
- You are required to submit a ZIP file named **ID_FirstName_LastName_sheet2.zip**, any other naming format will not be accepted and file will be discarded.
- Inside the ZIP file there should be the PDF file with solutions for the questions and the Downloaded Colab Notebook for the Coding questions
- Any copied sheets will be immediately zeroed and other penalties may be applied.

Good Luck