**Exploring Titanic Dataset with Pandas**

**1. Import the Required Libraries**

You should import Pandas (import pandas as pd) and Matplotlib (import matplotlib.pyplot as plt) for data manipulation and visualization.

Cimport pandas as pd

import matplotlib.pyplot as plt

Python 3.12.2 (tags/v3.12.2:6abddd9, Feb 6 2024, 21:26:36) [MSC v.1937 64 bit (AMD64)] on win32

Type "help", "copyright", "credits" or "license" for more information.

To import package, in this case, need to import pandas and matplotlib.

The output indicates that the version of python is 3.12.2

**2. Load the Dataset**

Load the Titanic dataset from a CSV file named 'titanic.csv' into a Pandas DataFrame.

Comment:

import pandas as pd

df\_csv = pd.read\_csv("C:/Users/mingy/Downloads/titanic.csv")

display(df\_csv)

Output:

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

Load csv file, we have to use jupyter notebook to read it.

**3. Data Exploration**

Display the first few rows of the dataset to understand its structure.

Print information about the columns and their data types.

Show summary statistics for numerical columns (e.g., mean, min, max, etc.).

**4. Data Cleaning**

Identify and handle missing values in the dataset (e.g., fill missing age values with median).

We missing some Age



We missing some Cabin



Perform any necessary data transformations or cleaning steps.

We must handle missing values, outliers, and inconsistencies in the dataset.

**5. Data Analysis**

Calculate and display the count of passengers by gender.

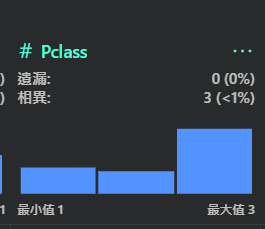


Compute the average age of passengers.

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

Determine the survival rate by passenger class.

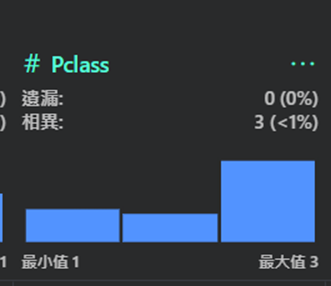


I use ipynb.(df\_csv), to show the data.

**6. Data Visualization**

Create a bar chart to visualize the survival rate by passenger clas

import matplotlib.pyplot as plt



Add appropriate labels and a title to the chart.

We can create various types of plots using libraries like Matplotlib and Seaborn.

**7. Conclusion**

Summarize your findings or insights gained from analyzing the Titanic dataset.

Reflect on the importance of data cleaning and exploration in data analysis.

Based on the exploratory data analysis, we can summarize the key findings, insights, and potential areas for further investigation. This could include patterns, trends, outliers, or relationships observed during the analysis.