

Assignment 5 - Question 1



≡ Machine Learning / AI 15 points

DESCRIPTION

Assignment 5 is a scenario-based assignment which uses Titanic Dataset and consists of 3 different questions. Read and understand the requirements and answer the questions carefully.

Dataset: Titanic disaster.

Data Dictionary:

Variable	Definition	Key
----------	------------	-----

- survival | Survival | 0 = No, 1 = Yes
- pclass | Ticket class | 1 = 1st, 2 = 2nd, 3 = 3rd
- sex | Sex | M or F
- Age | Age in years
- sibsp | # of siblings / spouses aboard the Titanic
- parch | # of parents / children aboard the Titanic
- ticket | Ticket number
- fare | Passenger fare
- cabin | Cabin number
- embarked | Port of Embarkation | C = Cherbourg, Q = Queenstown, S = Southampton

Variable Notes:

- pclass: A proxy for socio-economic status (SES)
 - 1st = Upper
 - 2nd = Middle
 - 3rd = Lower
- age: Age is fractional if less than 1. If the age is estimated, is it in the form of xx.5
- sibsp: The dataset defines family relations in this way...
- Sibling = brother, sister, stepbrother, stepsister
- Spouse = husband, wife (mistresses and fiancés were ignored)

- parch: The dataset defines family relations in this way
- Parent = mother, father
- Child = daughter, son, stepdaughter, stepson. Some children travelled only with a nanny, therefore parch=0 for them.

Dataset Path:

The dataset `Titanic_train.csv` is present at the location

```
/data/training/Titanic_train.csv
```

The dataset `Titanic_test.csv` is present at the location

```
/data/test/Titanic_test.csv
```

Problem Statement:

You are provided with the datasets about people from the Titanic disaster. Use the dataset resolve the following issues:

Q1: Find the relation of the following columns (having discrete values) with the “Survived” columns and answer the below questions:

- Pclass
- Sex
- Embarked

1. Find the total number of survivors from the 3rd PClass
(`Titanic_train.csv`)

Example:

```
If Total number of suvivor from Pclass(3): 100
```

```
Output: 100
```

2. Find the total number of male who died in the accident
(`Titanic_train.csv`)

3. Find the total number of the survivor who embarked the ship from "Southampton" (Titanic_train.csv)

Hint:

```
Pclass relation with Survived Column:
Group | Total | Survived: 1 | 189 | 116

Sex relation with Survived Column:
Group | Total | Survived: female | 262 | 194

Embarked relation with Survived Column:
Group | Total | Survived: C | 146 | 78
```

Final Output Sample:

	A
1	100
2	200
3	300

NOTE: Here, 100, 200 and 300 are the answer of 1st, 2nd and 3rd question respectively.

Output Format:

- Perform the above operations and write (written above as **print**) your output to a file named **output.csv**, which should be present at the location **/code/output/output.csv**
- **output.csv** should contain the answer to **each question** on consecutive rows.

DATASETS

- [Training dataset](#) ⓘ

EXECUTION TIME LIMIT

Default.