

Assignment1-spreadsheet data analysis

This assignment makes use of the following two datasets. (Datasets can be downloaded in Ketangpai)

1. Titanic

Datasets:

Titanic.csv

Description: Data on passengers of the RMS Titanic. Entries include the name, age, class, fare, gender, and whether or not the passenger survived.

2. World Cup

Datasets:

Players.csv,

Teams.csv

Description: 2010 World Cup data including last name, team, position, minutes played, and game statistics for each player (Players.csv) as well as world ranking, games played in tournaments, and game statistics for each team (Teams.csv)

For each of these problems, **submit an answer** obtained by performing spreadsheet operations on the provided data **and a screenshot** of how you manipulated the spreadsheet(s) to obtain your answer. We do not require that you capture each step of your manipulation, just paste the main functions used to come to your answer.

World Cup Data

1. Problem 1. Which team has the highest ratio of goals for/goals against?
2. Problem 2. What is the average number of passes made by defenders? By forwards?
3. Problem 3. What player on a team with "ia" in the team name played less than 200 minutes and made more than 100 passes?
4. (Optional) Problem 4. Which team has the highest average number of passes per minute played (and what is that average)?
5. Problem 5. How many players on a team ranked under 10 (i.e. ranks 1 to 9) played more than 350 minutes?

Titanic Data

6. Problem 6. What characteristics are shared by all passengers whose fare is 0?
7. Problem 7. How many married women over age 50 embarked in Cherbourg? (Married women are denoted by "Mrs.")
8. Problem 8. Which embarkation city had the highest-paying passengers on average?
9. Problem 9. What is the most common last name among passengers? What is the average number of passengers per last name?
10. Problem 10. What's the survival rate for passengers in the three different classes, i.e., what fraction of passengers in each class survived? Find the answer using spreadsheet functions only - don't perform any arithmetic by hand!