

CASE STUDY 006

[R]

Titanic Passengers



Please use the following code to access the data

```
# install package
install.packages('titanic')

# load package
require(titanic)

# create data frame of the data
titanic_data = data.frame(titanic_train)
```

Here are some clues in case you are stuck with the case study:

1. Make sure you use the code above to access the data
2. See this link for the explanation of variables <https://www.kaggle.com/c/titanic/data>
3. Rename the variables to make them more understandable
4. Change the Pclass variables to be 1st class, 2nd class and 3rd class so they are easier to understand
5. You will need to use the aggregate function for questions 3 & 4
6. For question 5 you will need to use the grepl function