Titanic Case Study

Dear Applicant,

Thank you very much for your application!

With this case-study we would like to test both your analytical and programming skills. For this exercise we use the data from the Titanic Kaggle competition (https://www.kaggle.com/c/titanic). In this challenge, the participants analyze what sorts of people were likely to survive the Titanic tragedy.

We prepared the data for you. In the "data" folder you can find 6 .csv files, including:

• train_base: base data for model training
• train_passenger : passenger data for model training
• train_ticket : ticket data for model training
• test_base: base data for model testing
• test_passenger : passenger data for model testing
• test_ticket : ticket data for model testing

You can find a description of the variables on https://www.kaggle.com/c/titanic/data . Please use the data we have sent you and not the data from Kaggle.

Please answer the following 5 questions:

1. What is the average age of the female passengers in the trainings data?
2. Do you expect passenger "Bowen, Miss. Grace Scott" and "Willard, Miss. Constance" (in the test data) to have survived? Why (not)?
3. How accurate do you think your predictions are?
4. What can you do to improve your predictions?

Please use the Python programming language for this exercise and present your code in a Jupyter Notebook.

Have fun!