Assignment 5: kaggle.com: Titanic

Direction: The assignment should be done in IPython notebook.
(1) Go to kaggle.com > Titanic: Machine Learning from Disaster and download all csv files.
(2) You will submit a single ipynb file.

(3) Make sure to include visual aids (e.g. graphs, tables, etc):

  (a) for exploring data to understand them

  (b) to present the steps toward your solution

(4) Answer the below questions at the end of your the ipynb file. Please repeat the question sentences.

(5) If you use ideas from any Kernels, you must disclose them by including the links to the Kernels in the ipynb file and you also explain how you improved their approaches. You may not simply use their approaches without improvement.

(6) Include one link to the best of your submissions in the ipynb file.
All submission links can be found
at https://www.kaggle.com/c/titanic/submissions?sortBy=date&group=all

Question:

1. How did you do to wrangle, prepare, cleanse the data?

2. How did you analyze, identify patterns, and explore the data?

3. Improve prediction from the lab and pick the best to submit to kaggle.com:
   a. What are the models you ran (you should run at least 5)?
   b. What are the features/variables in the data you chose to build your models?
   c. How and why did you pick the features/variables?
   d. What is the model you picked to submit and its score rated by kaggle.com (it should be higher than 0.75598)?
   e. What is/are the parameter value/values as a result of your parameter tuning?
   f. How did you pick the value(s)?

4. Answer the below questions about SettingWithCopyWarning in pandas (https://www.dataquest.io/blog/settingwithcopywarning)
   a. What is SettingWithCopyWarning? Please summarize.
   b. How did you make sure that the warning is not applicable in your cases?
   c. What can you do to avoid the warning according to the web? Please summarize.
   d. (extra points) What else can you do to avoid the warning? Please explain.