

DTSC-580: Data Manipulation

Assignment: Survivor

Directions

In this assignment, you will test all the skills that you have learned during this course to manipulate the provided data to find the answers to questions about the TV show *Survivor*. If you are not familiar with this show, start by watching this short clip that briefly explains it: [Survivor Explained](#)

See the “Data” section in the notebook for more information on the data set and the show.

Please note that your notebook should be named “survivor.ipynb” when submitting to CodeGrade for the automatic grading to work properly. You should also have the Excel file in the same location as your notebook as this is the way CodeGrade is expecting your file to be coded.

Show Work

Remember that you must show your work. Students submissions are checked manually to verify that they are not hard coding the answer from looking only in the file or in CodeGrade's expected output. If this is seen, the student's answer will be manually marked wrong and their grade will be changed to reflect this.

For example, if the question is: “Who is the contestant who has received the most tribal votes to be voted out. Select their record from the `castaway_details` DataFrame.”

You would show your work and code similar to this:

```
### incorrect way ###

Q1 = castaway_details[castaway_details['castaway_id'] == 333]

### correct way - showing your work ###

# get index

idx = vote_history.groupby('vote_id').size().sort_values(ascending=False).index[0]

# select row based on index

Q1 = castaway_details[castaway_details['castaway_id'] == idx]
```

Below is a simple rubric indicating how you will be graded:

Correct Answers 100%

Are your answers to the questions correct?