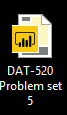
# DAT 520 Problem Set 5 Conditional Probability in Decision Trees

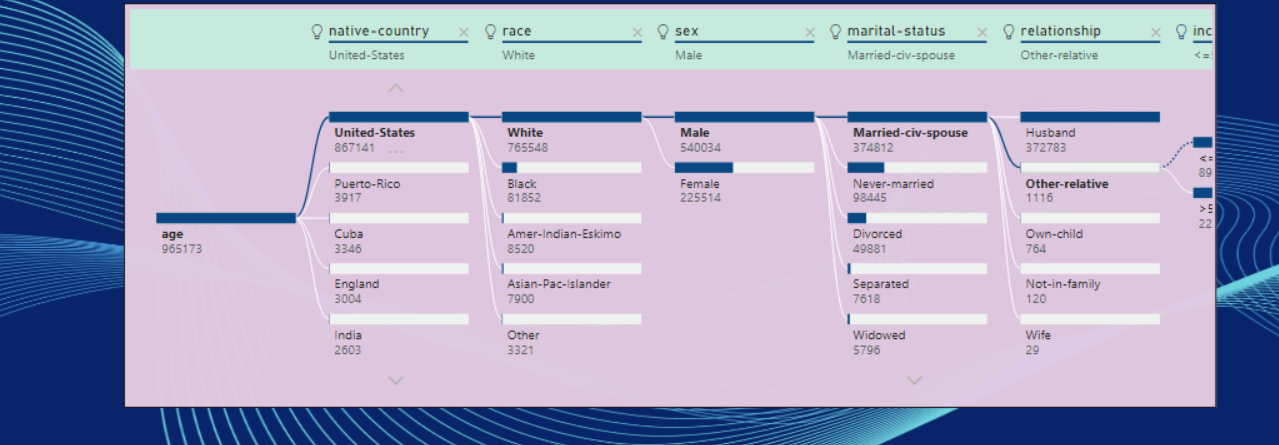
In Module Five, our exercise is building on what you have learned in prior modules. This problem set is going to introduce you to PowerBI, which is a powerful data analytics visualization tool. PowerBI gives you the ability to create decision trees by using point and click capabilities, which is beneficial for non-coding professionals. You will need to open the PowerBI file and answer the problem set questions. If you’re having difficulty, you’ll need to use Excel to answer the questions. The goal of this assignment is to be introduced to a new technology, learn how to read decision trees, and answer questions about them. Conditional probabilities will be calculated for you! Just point and click.

*Note:* Notice that you can view the entirety of the decision tree by clicking and dragging the tree left to right as needed. Also, while you are practicing, the layout dashboard may become altered to the point it is no longer useful. If this happens, simply close it, DO NOT save it, and reopen it. The dashboard will default back to its original state.

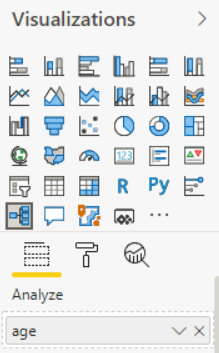
1. Login to Apporto.
2. On the desktop you should see the file (screenshot below) titled DAT-520 Problem Set 5. Double-click on the icon:

.

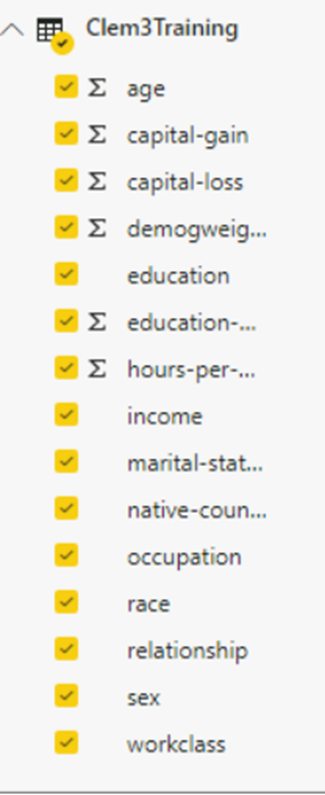
1. The decision tree should look like below:



1. What do you think the decision tree is predicting when you first open the file?
2. Let’s change what the decision tree is predicting.
   1. Enlarge the dashboard by clicking the enlarge iconEnlarge icon. It is located on the lower left side of the dashboard.
   2. Under **Visualizations** you should see an **Analyze** drop-down appear. It will look like this:



* 1. Click on the drop-down caret next to the word *age*.
  2. Select **Average of Age**.
  3. Under the **Fields** category, deselect all but age, capital gain, capital loss, education fields (2), hours-per, occupation, race, and sex. It should look something like this:



Questions:

**#1**: What does your new decision tree look like? Take a screenshot.

[Insert screenshot of decision tree here]

# #2:What is the average age with these new selections?

[Insert text]

**#3**:You can select variables from the decision tree by clicking on them; the output of the tree will change based upon each selection. Change the inputs so only the following are selected, then take a screenshot of the revised decision tree.

* Weighted percentage of white
* Female
* Occupation of prof-specialty
* With a doctorate degree

[Insert screenshot of decision tree here]

**#4**: As you change the inputs, the numbers change. Why do you think this is?

[Insert text]

**#5**: Using the skill you’ve learned above, create 3 different decision trees with the data and explain what each one is calculating.

1. [Insert screenshot of first decision tree here]
2. [Insert explanation of calculation]
3. [Insert screenshot of second decision tree here]
4. [Insert explanation of calculation]
5. [Insert screenshot of third decision tree here]
6. [Insert explanation of calculation]