

## Technical Programming II

### Practical Assessment: Medical Costs Case Study

#### Data files:

- insurance.csv

#### Scenario:

You are acting as an analyst employed by a healthcare company. Your boss has asked you to produce a report that answers the following questions and completes the following tasks.

#### Tasks and questions:

1. How is age related to medical costs?
2. How is number of children related to medical costs?
3. How is the number of people distributed by region?
4. How is the number of people distributed by age?
5. How are the charges distributed?
6. Bin the bmi column. To do that, search the internet to determine how you should bin and label the data.
7. How are the charges related to the bmi?
8. How is the smoker status related to the charges?
9. How are the charges related to the region?
10. Which region has the highest obesity percentage?
11. Create a simple regression to show the relationship between charges and age.
12. How does this relationship change with smoking status?
13. How does this relationship change with BMI?
14. Create a multiple regression model to predict charges. To do that, you'll need to dummy encode and rescale the data.
15. Make predictions with your multiple regression model and evaluate how well your model is working.
16. What is the optimal number of parameters for the multiple regression?

#### Instructions:

1. Create a Notebook for the study.
2. Create a heading for the question or task.
3. If necessary, perform cleaning or preparation for the question or task.

4. Create a table or plot that completes the task or answers the question. Make sure to use a heading or raw text to explain what the table or plot tells you.
5. Repeat steps 2-4 for each question or task.

-END-