**DECISION TREE WORKSHOP INSTRUCTIONS**

**STEP 1:** View attached presentation. Make sure you understand the following:

* Relationship between Information Gain and Entropy
* How Random Forests Reduce Bias in Feature Selection
* Pros and Cons on Random Forests
* When to Use a Random Forest vs. a Neural Network

**STEP 2:** Follow the Links below to access the Google Collab Notebook and download the attached dataset:

* **Random Forest Sample Problem**
  + [Random Forest Example](file:///C:\Users\Chase\OneDrive\Documents\Career%20Development\Data%20Science%20Club\Workshops\RandomForestWorkshop\Random-Forest-Example.ipynb)
    - [Sample Dataset](../../Career%20Development/Data%20Science%20Club/Workshops/RandomForestWorkshop/clean_data.csv)
      * **NOTE: Make sure to perform some Exploratory Data Analysis (EDA) to gain a basic understanding of the dataset.**

**STEP 3:** Download and complete the following challenges:

* **Random Forest Challenge**
  + **Description:** Using the dataset provided below, create a Regression Tree to predict salaries. Optimize your model using a Random Forest and identify which features are more important. ***If you do this challenge in Python, make sure to follow the Sample Problem with the Label Encoding and Data Preparation Tasks.***
    - [Random Forest Challenge](../../Career%20Development/Data%20Science%20Club/Workshops/RandomForestWorkshop/Random-Forest-Challenge.ipynb)
    - [Salaries Dataset](file:///C:\Users\Chase\OneDrive\Documents\Career%20Development\Data%20Science%20Club\Workshops\RandomForestWorkshop\salary_data_cleaned.csv)