Week 7 Assignment

(Estimated Time To Complete: **2 - 5 hrs**)

1. **Data Generation**
   1. **Generate Normal Data**:
      1. Create a dataset of 1000 samples drawn from a normal distribution with mean = 50 and standard deviation = 5, named normal\_data.
   2. **Generate Uniform Data:**
      1. Generate 1000 samples from a uniform distribution ranging between 10 and 20, stored as uniform\_data.
2. **Conducting Statistical Tests**
   1. **Independent Samples t-Test:** 
      1. Conduct an independent samples t-test comparing normal\_data and uniform\_data.
      2. Extract and print the t-statistic and p-value only.
   2. **Paired Samples t-Test:** 
      1. Create two datasets, sample1 and sample2, each with 100 samples from a normal distribution (mean=10, std=2).
      2. Conduct a paired t-test on sample1 and sample2, and print relevant results.
3. **Linear Regression and Correlation**
   1. **Linear Regression:** 
      1. Generate x as an array from 0 to 99
      2. Create y as a linear function of x with added noise.
         1. Hint: **y = m\*x + b + noise**
            1. For m, define m as an integer
            2. For b, define b as an integer
            3. For noise, generate normal data with the same length as x
      3. Perform a regression analysis and print the true and estimated slope and intercept.
   2. **Pearson Correlation:**
      1. Calculate the Pearson correlation coefficient between x and y
      2. Print the correlation coefficient and p-value.
4. **Applying Statistical Tests in a Loop**
   1. **Independent Samples t-Test Across Multiple Datasets**:
      1. Generate five datasets, each containing 1000 samples drawn from a normal distribution with a mean of 50 and a standard deviation of 5.
      2. **Store all 5 arrays into a single array** called datasets.
      3. Use a for loop to **perform an independent samples t-test between each dataset** in datasets and uniform\_data from 1bi.
      4. **Extract and print only the p-values** for each test, labeling them by dataset (e.g., "Dataset 1 p-value:", "Dataset 2 p-value:", etc.).