1. Create a class and name it Z-test:

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

2. Create a class and name it T-test this class includes:
   - One-sample: **df = n - 1**

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

   - Two-sample: **df = $n_1 + n_2 - 2$**

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

   - Paired t-test: **df = n - 1**

$$t_{calc} = \frac{\bar{d}}{s_d / \sqrt{n}}$$

   d bar: is the difference between the two samples
   S_d is the standard deviation of the difference between samples

3. Create a class and name it ANOVA it includes

- One way

$$\boxed{F = \frac{MS_{between}}{MS_{within}}}$$

$$\boxed{MS_{between} = \frac{SS_{between}}{df_{between}}} \quad \boxed{MS_{within} = \frac{SS_{within}}{df_{within}}}$$

$$\boxed{SS_{between} = \Sigma\frac{(\Sigma x)^2}{n} - \frac{(\Sigma\Sigma x)^2}{n_T}} \quad \boxed{SS_{within} = \Sigma\Sigma(x^2) - \Sigma\frac{(\Sigma x)^2}{n}}$$

$$\boxed{df_{between} = k - 1} \quad \boxed{df_{within} = n_T - k}$$

- Two ways Using the **scipy** library

4. Create a class and name it Chi-Square

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

5. Create a class and name it AB testing that inherits the two sample t-test from the class T-test

Use these classes to solve the following problems:

1. Z-test:
   Suppose we are testing a new model of cell-phone and want to compare its bandwidth to the previous model.

   Sample = [4.186, 4.439, 4.781, 4.388, 4.947, 4.853, 4.889, 4.682, 4.428, 4.533, 4.557, 4.761, 4.491, 4.334, 4.83 , 4.268, 4.68 , 4.437, 5.382, 5.111, 5.096, 5.232, 5.033, 5.57 , 4.474, 4.789, 4.725, 4.84 , 4.817, 4.438, 4.754, 4.966, 4.285, 4.482, 4.396, 4.418, 4.514, 5.383, 5.264, 4.309, 5.058, 4.392, 4.788, 4.934, 4.967, 4.554, 4.42 , 5. , 5.126, 5.082, 4.944, 4.658]

   - State the null and alternative hypotheses.
   - Use the Z-test to determine whether the new model has a significantly different bandwidth 4.5 GHz from the previous model with a standard deviation of 0.6 GHz, using a significance level of 0.05.
   - Interpret the results

2. T-test
    A. **One sample T-test:** Suppose we have a new manufacturing process for producing aluminum cans, and we want to test whether the mean weight of the cans produced using this new process is significantly different from the target weight of 15 grams. We randomly sample 30 cans produced using the new process and measure their weights in grams. We obtain the following data:

        Sample = [14.8, 15.2, 15.1, 15.3, 15.0, 14.9, 15.2, 14.8, 15.1, 15.0, 14.9, 14.8, 15.2, 14.9, 15.0, 14.9, 15.1, 15.3, 15.0, 15.1, 14.8, 15.0, 15.2, 15.1, 15.3, 15.1, 15.0, 14.8, 15.2, 15.0]

        - State the null and alternative hypotheses.
        - Use the appropriate T-test to determine whether the mean weight of the produced cans is equal to the target weight of 15 grams, using a significance level of 0.05.
        - Interpret the results

    B. **Two Sample T-test:** Suppose a food company has developed a new flavor of potato chips and wants to compare it to the current best-selling flavor. The company randomly selects two groups of 20 customers each. The first group is given the new flavor of potato chips, while the second group is given the best-selling flavor. After trying the potato chips, each customer rates the flavor on a scale of 1 to 10. The following are the flavor ratings for the two groups:

        New flavor = [8, 7, 9, 6, 7, 8, 9, 7, 8, 7, 6, 8, 7, 9, 8, 7, 6, 9, 8, 7]
        Best selling flavor = [6, 7, 8, 6, 7, 6, 7, 6, 8, 7, 6, 7, 6, 8, 7, 6, 7, 8, 6, 7]

        - State the null and alternative hypotheses.
        - Use the appropriate T-test to determine whether there is a significant difference between the two groups or not, using a significance level of 0.05.
        - Interpret the results

    C. **Paired T-test:** Suppose a company wants to evaluate a new training program for its employees. The company selects 20 employees and measures their productivity before and after the training program. The following are the productivity scores (number of tasks completed per hour) for each employee before and after the training:

        Before = [15, 18, 12, 10, 17, 16, 12, 14, 19, 18, 11, 13, 16, 17, 19, 14, 16, 13, 15, 12]
        After = [18, 20, 15, 13, 19, 18, 14, 16, 21, 20, 14, 16, 19, 20, 22, 16, 18, 15, 17, 14]

- State the null and alternative hypotheses.
- Use the appropriate T-test to determine whether the new training program has had an effect on employee productivity, with a significance level of 0.05.

3. ANOVA Test
   A. **One-Way:** Suppose a company has three departments (A, B, and C) and wants to test whether there is a significant difference in salaries between the departments. The company selects 10 employees randomly from each department and records their salaries.

      Department A = [55, 60, 50, 58, 63, 62, 57, 56, 61, 59]
      Department B = [50, 52, 48, 49, 55, 53, 51, 54, 47, 50]
      Department C = [45, 43, 48, 50, 42, 47, 49, 46, 44, 48]

      - State the null and alternative hypotheses.
      - Use One-Way ANOVA-test to determine whether there is a significant difference in salaries between the three departments.
      - Interpret the results

   B. **Two-Way:** Suppose a company has three departments (A, B, and C) and wants to test whether there is a significant difference in salaries between the departments, while also considering the effect of gender. The company selects 10 employees randomly from each department and records their salaries and gender.

      Department A:
        Male: [$55k, $60k, $50k, $58k, $63k]
        Female: [$62k, $57k, $56k, $61k, $59k]
      Department B:
        Male: [$50k, $52k, $48k, $49k, $55k]
        Female: [$53k, $51k, $54k, $47k, $50k]
      Department C:
        Male: [$45k, $43k, $48k, $50k, $42k]
        Female: $47k, $49k, $46k, $44k, $48k]

      - State the null and alternative hypotheses.
      - Use Two-Way ANOVA-test to determine whether there is a significant difference in salaries between the three departments.
      - Interpret the results

4. **Chi-Square:** Suppose we have a hypothesis that a six-sided die is fair, with each side having an equal probability of landing face up. We roll the die 100 times and record the number of times each side appears:

Side	Observed Frequency
1	18
2	20
3	16
4	22
5	14
6	10

Given that the expected probabilities are equal to 1/6, Use the class CHi-Square to estimate the goodness-of-fit through chi-square test to determine whether the observed frequencies are consistent with the expected frequencies under the null hypothesis of a fair die.

5. **A/B testing:** Suppose a company has launched a new flavor of soda and wants to test if it has a higher preference than the old flavor. The company conducts a survey with a sample of 30 customers, randomly split into two groups of 15. One group is given the old flavor of soda and the other group is given the new flavor of soda. Each participant rates their preference on a scale of 1 to 10.

Old Flavor = [6, 7, 8, 5, 6, 7, 5, 8, 6, 7, 5, 6, 7, 6, 5]
New Flavor = [8, 9, 7, 8, 9, 6, 7, 8, 7, 8, 7, 8, 9, 6, 8]

- State the null and alternative hypotheses.
- Use two sample t-test to determine whether there is a significant difference in preference between the old and new flavors of soda, with a significance level of 0.05
- Interpret the results