PROFESSIONAL

Single Sample z-test

WhiteHat Jr
Live Online Coding for Kids

INSTRUCTIONS:

Goal of the Project:

In Class 111, we learned about the single sample z test.

In this project you will have to write a program to do the z test of a given sample.
** This is a continuation of the project we did for Class 110. Please complete that project before attempting this project **

Story:

In our journey of analyzing the articlele's data, you also want to understand how the results are changing after the introduction of an intervention.

```
#finding the z score using the formula
z_score = (mean_of_sample1 - mean)/std_deviation
print("The z score is = ",z_score)
```

```
mean of sampling distribution:-  50.69924
Standard deviation of sampling distribution:-  2.879529182125215
Mean of sample1:-  50.41
The z score is =  -0.10044697646944323
```

***This is just for your reference. We expect you to apply your own creativity in the project.**

Getting Started:

1. Download the medium article published data from this link.
2. Download sample2.csv from this link.
3. Open your Code editor.

Specific Tasks to complete the Project:

1. Import statistics library , plotly.graph_objects, plotly. Figure_fatcory to the program file.
2. Read the medium_article.csv file using read csv.
3. Find the mean of the complete data (for column “reading_time” using statistics.mean() and assign it to the “population_mean” variable.

4. Write a function to take random 30 samples from the data and find their mean.
5. Write a setup function to repeat that process 100 times and store the means (mean of samples) in a list.
6. Write a **plot_graph()** function to take the list of means and plot it on the graph.
7. Find the first , second and third standard deviation and add traces of it on the graph.
8. Take the new intervention data and find it's mean (we'll call it new sample mean (**sample2.csv** in this case)).
9. Plot the new sample mean on the graph and check in which deviation it lies.
10. If the new mean is greater than the sampling mean then the intervention is successful.
11. Find the z score using the formula (new sample mean - sampling mean) / standard deviation.

Submitting the Project:

1. Upload your completed project to your own github account.
2. Create a new repository named **Project 111**.
3. **Upload** working code to this github repository.
4. Submit the published link of the project in the Student Dashboard.

Hints:

1. Code to find the mean of the 30 sample data. In place of the counter pass the number of samples you want to take.

```
def random_set_of_mean(counter):
    dataset = []
    for i in range(0, counter):
        random_index= random.randint(0,len(data))
        value = data[random_index]
        dataset.append(value)
    mean = statistics.mean(dataset)
    return mean
```

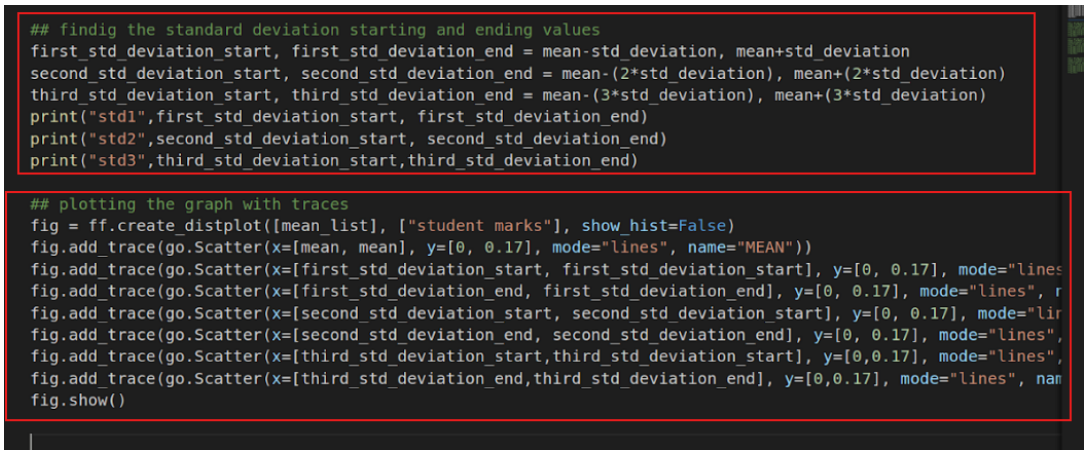
2. Code to repeat the process 100 times.

```
def setup():
    mean_list = []
    for i in range(0,100):
        set_of_means= random_set_of_mean(30)
        mean_list.append(set_of_means)
    show_fig(mean_list)
```

3. Function to plot the graph.

```
def show_fig(mean_list):
    df = mean_list
    fig = ff.create_distplot([df], ["temp"], show_hist=False)
    fig.show()
```

4. Code to find the 1st, 2nd and 3rd deviation and add traces of it.



```
## findig the standard deviation starting and ending values
first_std_deviation_start, first_std_deviation_end = mean-std_deviation, mean+std_deviation
second_std_deviation_start, second_std_deviation_end = mean-(2*std_deviation), mean+(2*std_deviation)
third_std_deviation_start, third_std_deviation_end = mean-(3*std_deviation), mean+(3*std_deviation)
print("std1",first_std_deviation_start, first_std_deviation_end)
print("std2",second_std_deviation_start, second_std_deviation_end)
print("std3",third_std_deviation_start,third_std_deviation_end)

## plotting the graph with traces
fig = ff.create_distplot([mean_list], ["student marks"], show_hist=False)
fig.add_trace(go.Scatter(x=[mean, mean], y=[0, 0.17], mode="lines", name="MEAN"))
fig.add_trace(go.Scatter(x=[first_std_deviation_start, first_std_deviation_start], y=[0, 0.17], mode="lines
fig.add_trace(go.Scatter(x=[first_std_deviation_end, first_std_deviation_end], y=[0, 0.17], mode="lines", r
fig.add_trace(go.Scatter(x=[second_std_deviation_start, second_std_deviation_start], y=[0, 0.17], mode="lir
fig.add_trace(go.Scatter(x=[second_std_deviation_end, second_std_deviation_end], y=[0, 0.17], mode="lines",
fig.add_trace(go.Scatter(x=[third_std_deviation_start,third_std_deviation_start], y=[0,0.17], mode="lines",
fig.add_trace(go.Scatter(x=[third_std_deviation_end,third_std_deviation_end], y=[0,0.17], mode="lines", nam
fig.show()
```

REMEMBER... Try your best, that's more important than being correct.
After submitting your project your teacher will send you feedback on your work.

PROFESSIONAL

Single Sample z-test

WhiteHat Jr
Live Online Coding for Kids

——— xxx ——— xxx ——— xxx ——— xxx ——— xxx ———