DL1_C5_S6_Challenge

In [1]:
```
import statistics as st
import math
import numpy as np
import seaborn as sns
import matplotlib.pyplot  as plt
from scipy.stats import norm
```

Task 3 : Stockholder Age Analysis

In [2]:
```
x_samp= 45.1
mu_popmean = 44
sd = 8.7
samp = 68
alpha=0.05
```

Task 3.A: What is null and alternate hypothesis

H0 = Avg age of female shareholders is 44 years H1 = Avg age of female shareholders is higher so both side tail test

Task 3.B : Calculate critical value

In [3]:
```
z_critical = norm.ppf(alpha/2)
print("The z-critical is", z_critical)
```

```
The z-critical is -1.9599639845400545
```

Task 3.C : Test Statistics(z_statistic)

In [4]:
```
z_statistic = (x_samp- mu_popmean)/(sd/(samp)**0.5)
print("The Z statistics is ", abs(z_statistic))
```

```
The Z statistics is  1.042624411075732
```

Task 3.d : p-value

In [5]:
```
p_value = norm.sf(abs(z_statistic))
print("The p_value is ",p_value*2)
```

```
The p_value is  0.29712227883324194
```

Task 3.E : Draw the Conclusion

i. If p-value < alpha : Rejection of Null Hypothesis(H0) p_value = 1.44e-20 p_value < alpha which means not to acceptance of Null Hypothesis. so, inconclusion H1 is True that is price of house is not 33.88 sqft.

ii. If -(z-critical) > z-statistics > +(z-critical) : Rejection of Null Hypothesis(H0) z-critical = 1.95 Z statistics = 9.29 z-critical < z-statistics which means not to acceptance of Null Hypothesis. so, inconclusion null Hypothesis is false and it can be rejected.

Task 2 : Health Drink Price Survey Analysis

In [6]:
```
import pandas as pd
Drink_df=pd.read_csv("DS1_C5_S6_CityPrice_Data_Challenge.csv")
Drink_df
```

Out[6]:

	California	Florida
0	2.55	2.25
1	2.36	2.40
2	2.38	2.19
3	2.43	2.39
4	2.73	2.39
5	2.67	2.50
6	2.54	2.33
7	2.24	2.33
8	2.43	2.40
9	2.80	2.49
10	2.50	2.19
11	2.64	2.29
12	2.25	2.59
13	2.38	2.23
14	2.61	2.41
15	2.71	2.56
16	2.80	2.48
17	2.49	2.29
18	2.69	2.28
19	2.43	2.39
20	2.23	2.59
21	2.61	2.58
22	2.57	2.53
23	2.36	2.26
24	2.57	2.39
25	2.56	2.58
26	2.38	2.29
27	2.43	2.27
28	2.71	2.45
29	2.25	2.56
30	2.64	NaN
31	2.27	NaN

	California	Florida
32	2.18	NaN

In [9]:
```
x_Florida  = Drink_df['Florida'].mean()
x_Florida
```

Out[9]: 2.396

In [10]:
```
x_California = st.mean(Drink_df['California'])
x_California
```

Out[10]: 2.4966666666666666

In [11]:
```
n1=33
n2=30

var_c=0.028
var_f=0.015

sd_c = var_c**0.5
sd_f = var_f**0.5
```

Task 2.A : What is null and alternate hypothesis

```
H0 = The mean of entire price is same
H1 = The price is decrease or decrease.
so, two tail test
```

Task 2.B : using alpha = 0.1 calculate critical value

In [17]:
```
z_critical = norm.ppf(1-alpha/2)
print("The z-critical is", z_critical)
```

```
The z-critical is 1.959963984540054
```

Task 2.C : Test Statistics(z_statistic)

In [13]:
```
z_statistic = (x_Florida-x_California)/(sd_c/(n2)+sd_f/(n1))**0.5
print("The Z statistics is ", abs(z_statistic))
```

```
The Z statistics is  1.0444780056077585
```

Task 2.d : p-value

In [14]:
```
p_value = norm.sf(abs(z_statistic))
print("The p_value is ",p_value*2)
```

```
The p_value is  0.2962642899686494
```

In [16]:
```
rs =  (x_Florida-x_California) + 2.58*((var_c/n1 + var_f/2))
ls =  (x_Florida-x_California) - 2.58*((var_c/n1 + var_f/2))
print("99% confidance level ", rs-ls)
```

```
99% confidance level  0.04307818181818182
```

Task 2.e : Conclusion

In [ ]:
```
i. If p-value < α : Rejection of Null Hypothesis(H0)

ii. If -z-critical > z-statistic > +z-critical : Rejection of Null Hypothesis(H0)
```

```
p_value=0.29 > alpha=0.1, which means not  to rejection of Null Hypothesis.

z_statistics=1.04 < z-critical= 1.95, which means  not rejection of Null
Hypothesis.
```

Task 1 : Customer Service Analysis

In [ ]:
```
x_samp = [3,4,5,5,4,5,5,4,4,4,4,4,4,4,4,5,4,4,4,3,3,3,4,3,5,4,4,5,4,4,4,5 ]
x_samp = st.mean(x_samp)
x_samp
```

In [ ]:
```
x_samp= 4.09
mu_popmean = 4.30
sd = 0.574
samp = 32
alpha=0.10
```

Task 1.A: What is null and alternate hypothesis

H0 = The mean of the survey was 4.30 H1 = The mean of the survey was not 4.30 We have to proves that would not rate highly , so its both side tail hypothesis

Task 1.B : Calculate critical value

In [ ]:
```
z_critical = norm.ppf(alpha/2)
print("z_critical is ", z_critical)
```

Task 1.C: Test Statistics(z_statistic)

In [ ]:
```
z_statistic = (x_samp / mu_popmean)/(sd/(samp)**0.05)
print("The z_statistic is ", z_statistic)
```

Task 1.D : P_value

In [ ]:
```
p_value = norm.sf(abs(z_statistic))
print("The p_value is ", p_value)
```

Task 1.E : Draw a Conclusion

i. If p-value < alpha : Rejection of Null Hypothesis(H0) p_value = 0.024 p-value < 0.10 which means, we can reject null Hypothesis

ii. If -z-critical > z-statistics > +z-critical : Rejection of Null Hypothesis(H0) z_critical = 1.97