

**1. The school sample dataset has the height of school children. Determine whether the sample mean is statistically different from a known population mean = 5.7 for boys to a significance level of 0.05?**

**CODE:**

```
height=[5.6,5.7,5.4,5.8,5.9,6,5.7,5.8,5.4,5.5,5.5,5.7,5.8,5.8,5.8,5.8,5.9,6,5.7,5.8,6.1,6.1,6.2,6.1,5.6,5.7,5.7]

from scipy.stats import ttest_1samp

import numpy as np

height_mean = np.mean(height)

print(height_mean)

tset, pval = ttest_1samp(height, 5.7)

print(pval)

if pval < 0.05:

    print("we are rejecting null hypothesis i.e,average height is not 50")

else:

    print("we are accepting null hypothesis i.e,average height is 50")
```

**OUTPUT**

```
5.78148148148148
0.05561173145472786
we are accepting null hypothesis i.e,average height is 50
```

**2. The school sample dataset has the height of school children. Determine whether the sample mean is statistically different from a known population mean = 5.7 for girls (group-1) to a significance level of 0.05?**

**CODE:**

```
height=[5.2,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.7,5.4,5.2,5.6,5.7,5.4,5.7,5.4,5.2,5.6,5.7,5.4]
from scipy.stats import ttest_1samp
import numpy as np
height_mean = np.mean(height)
print(height_mean)
tset, pval = ttest_1samp(height, 5.7)
print(pval)
if pval < 0.05:
    print("we are rejecting null hypothesis i.e,average height is not 50")
else:
    print("we are accepting null hypothesis i.e,average height is 50")
```

**OUTPUT:**

```
5.47037037037037
2.2489303609065707e-06
we are rejecting null hypothesis i.e,average height is not 50
```

**3. The school sample dataset has the height of school children. Determine whether the sample mean is statistically different from a known population mean = 5.7 for girls (group-2) to a significance level of 0.05?**

**CODE:**

```
height=[5.6,5.7,5.4,5.8,5.9,6,5.6,5.7,5.4,5.8,5.9,6,5.6,5.7,5.4,5.8,5.9,6,5.4,5.5,5.5,5.7,5.8,5.8,5.8,5.8,5.9]
from scipy.stats import ttest_1samp
import numpy as np
height_mean = np.mean(height)
print(height_mean)
tset, pval = ttest_1samp(height, 5.7)
print(pval)
if pval < 0.05:
    print("we are rejecting null hypothesis i.e,average height is not 50")
else:
    print("we are accepting null hypothesis i.e,average height is 50")
```

**OUTPUT:**

```
5.71851851851852
0.6208793777493744
we are accepting null hypothesis i.e,average height is 50
```

**4. The school sample dataset has the height of school children but also their gender. Compare the height of boys and girls (Group-1) and see if the difference is statistically significant to a significance level of 0.05?**

**CODE:**

```
from scipy.stats import ttest_ind
height1=[5.6,5.7,5.4,5.8,5.9,6,5.7,5.8,5.4,5.5,5.5,5.7,5.8,5.8,5.8,5.8,5.9,6,5.7,5.8,6.1,6.1,6.2,6.1,5.6,5.7,5.7]
height2=[5.2,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.2,5.6,5.7,5.4,5.7,5.4,5.2,5.6,5.7,5.4]
stat, p = ttest_ind(height1,height2)
print('P Value is :',p)
if p < 0.05:
    print('same distribution')
else:
    print('different distributions')
```

**OUTPUT:**

```
P Value is : 8.584228837236856e-07
same distribution
```

**Q5. The school sample dataset has the height of school children but also their gender. Compare the height of boys and girls (Group-2) and see if the difference is statistically significant to a significance level of 0.05?**

**CODE:**

```
from scipy.stats import ttest_ind
height1=[5.6,5.7,5.4,5.8,5.9,6,5.7,5.8,5.4,5.5,5.5,5.7,5.8,5.8,5.8,5.8,5.9,6,5.7,5.8,6.1,6.1,6.2,6
.1,5.6,5.7,5.7]
height2=[5.6,5.7,5.4,5.8,5.9,6,5.6,5.7,5.4,5.8,5.9,6,5.6,5.7,5.4,5.8,5.9,6,5.4,5.5,5.5,5.7,5.8,5.8
,5.8,5.8,5.9]
stat, p = ttest_ind(height1,height2)
print('P Value is :',p)
if p < 0.05:
    print('same distribution')
else:
    print('different distributions')
```

**OUTPUT**

```
P Value is : 0.2573127348673761
different distributions
```