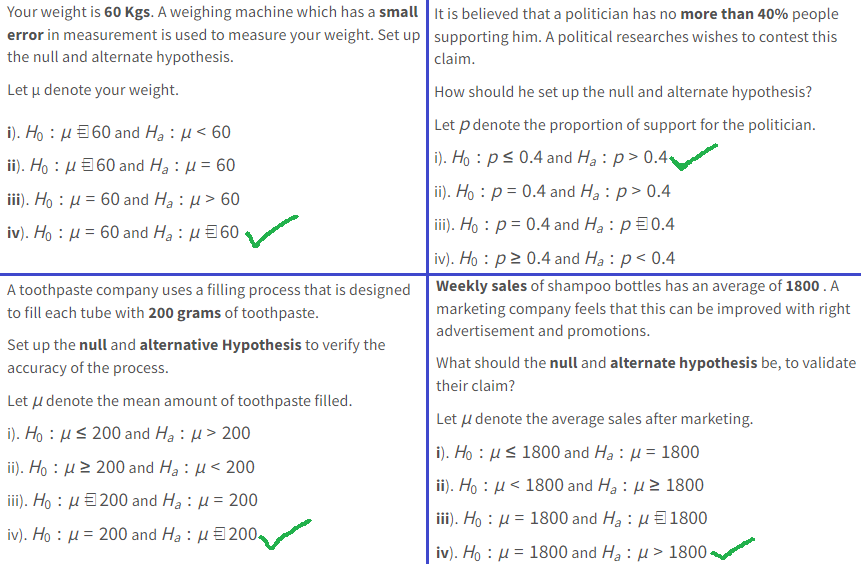
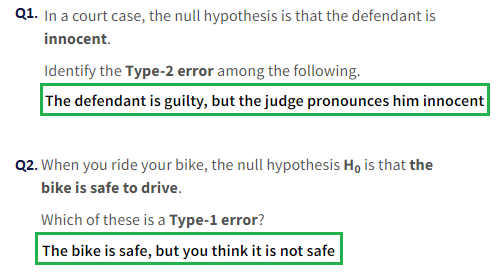
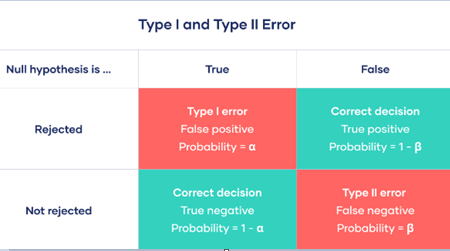
**HYPOTHESIS TESTING PROBLEMS**



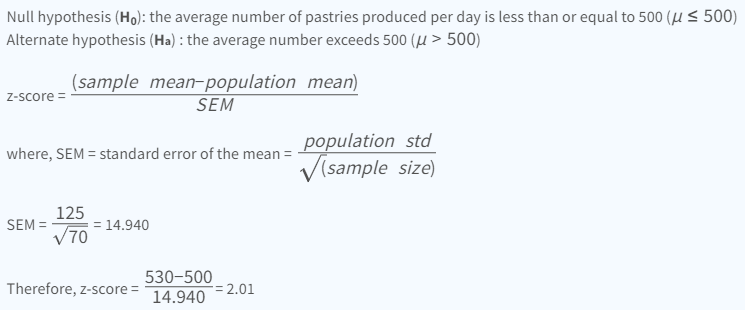




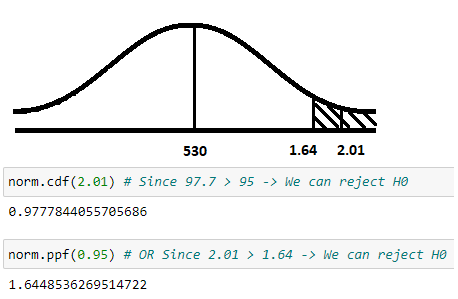
**Q1.** A French cake shop claims that the **average number of pastries** they can produce in a day **exceeds 500**.

The average number of pastries produced per day over a **70-day period** was found to be **530**. Assume that the population **standard deviation** for the pastries produced per day is **125**.

Test the claim using a **z-test** with the critical z-value = 1.64 at the alpha (significance level) = 0.05, and state your interpretation.



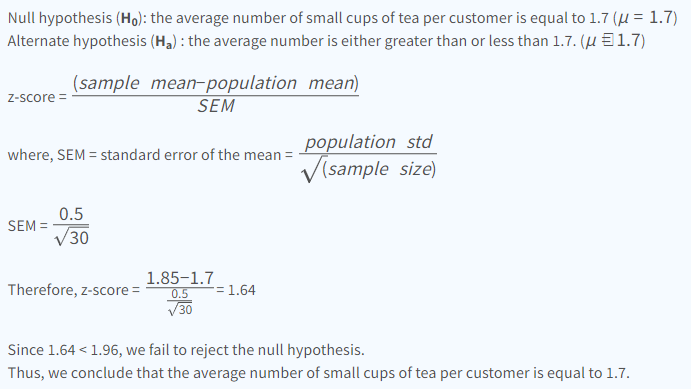
\*Right Tailed Test



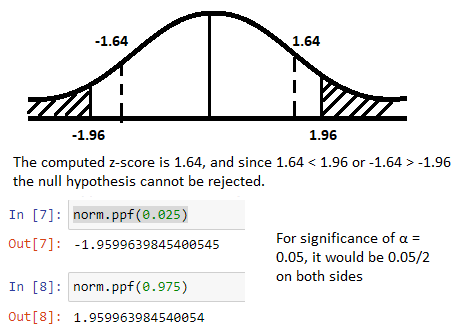
**Q2.** The Chai Point stall at Bengaluru airport estimates that each person visiting the store drinks an average of **1.7 small cups** of tea.

Assume a population **standard deviation** of **0.5** small cups. A sample of **30 customers** collected over a few days averaged **1.85** small cups of tea per person.

Test the claim using a **z-test** at an alpha = 0.05 significance value, with a critical z-score value of ±1.96.



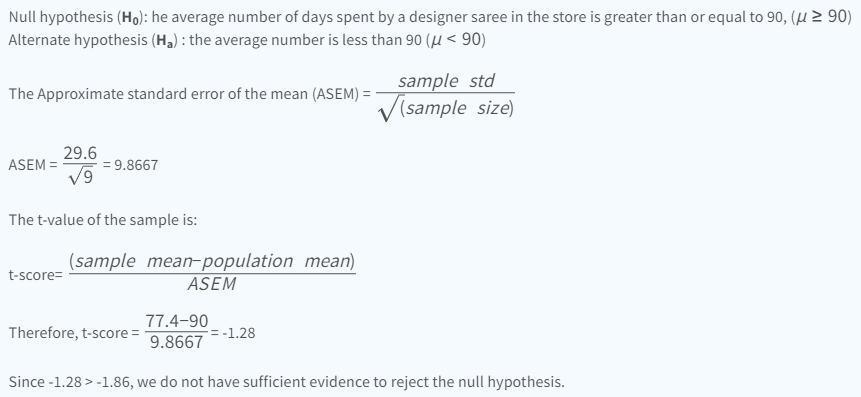
\*Two Tailed Test



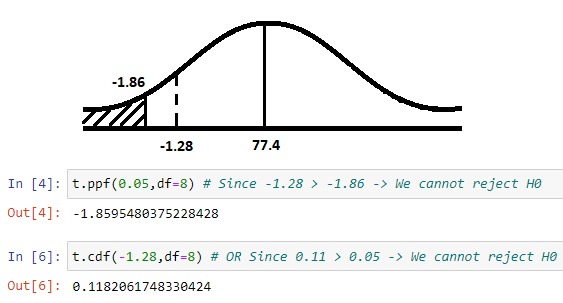
Q3. A luxury women's clothing store claims that designer grade sarees are sold **within 90 days** from the day they are introduced in their store.

Last month, they sold all **9 sarees** in a certain catalogue at an average of **77.4 days** spent in the store. The **standard deviation** of the sample was **29.6 days**.

Assuming the time spent in store is **normally distributed**, test the claim made by the store at an **alpha = 0.05** significance value, for which the one-sided **critical t-value** is **-1.86**.



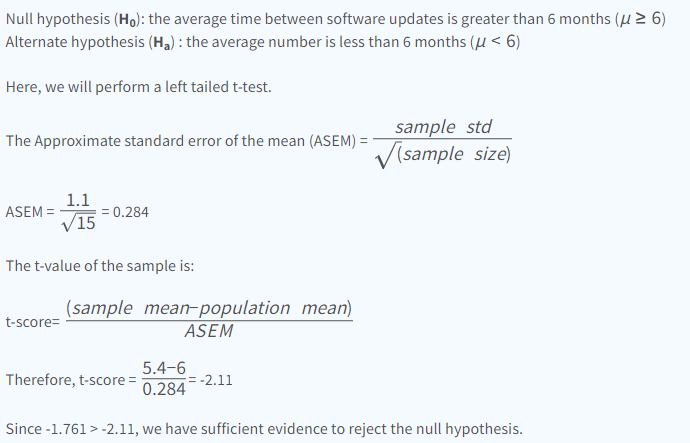
\*Left tailed test



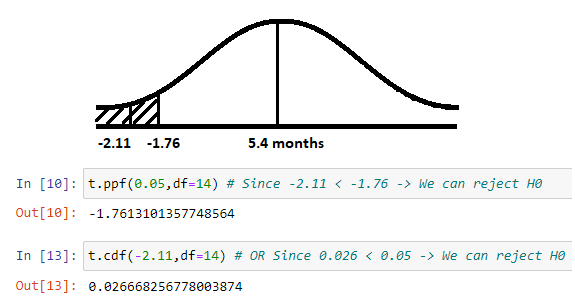
Q4. An accounting software company claims that they provide regular software updates with the time between updates being **less than 6 months**.

A survey covering all **15 client firms** showed that the actual average time between software updates is **5.4 months** with a **standard deviation** of **1.1 months**.

Assume the population to be **normally distributed**. Test the claim made by the company using a **t-test** at an **alpha = 0.05** significance level, for which the critical t-value is given as **-1.761**.



\*Left Tailed test



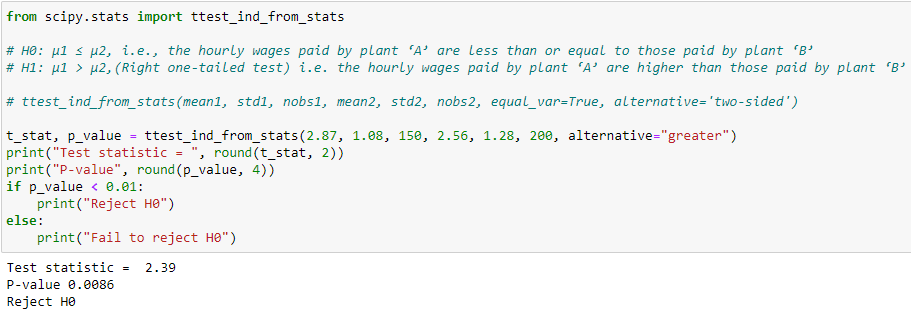
Q5. The **average** hourly wage of a sample of **150** workers in plant 'A' was Rs.**2·87** with a **standard deviation** of Rs. **1·08**.

The **average** wage of a sample of **200** workers in plant 'B' was Rs. **2·56** with a standard deviation of Rs. **1·28**.

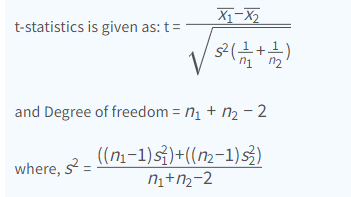
Can an applicant safely assume that the hourly wages paid by plant 'A' are higher than those paid by plant 'B' at a 1% significance level?

NOTE: When mean, standard deviation & nobs of both datasets are given: use “ttest\_ind\_from\_stats”

Solutions: -



Here,

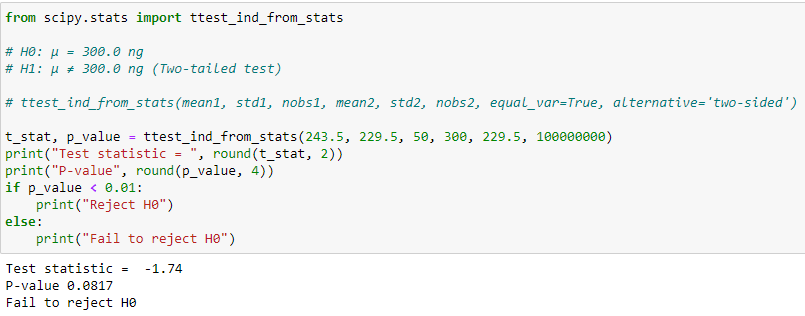


Q6. When smokers smoke, nicotine is transformed into cotinine, which can be tested.

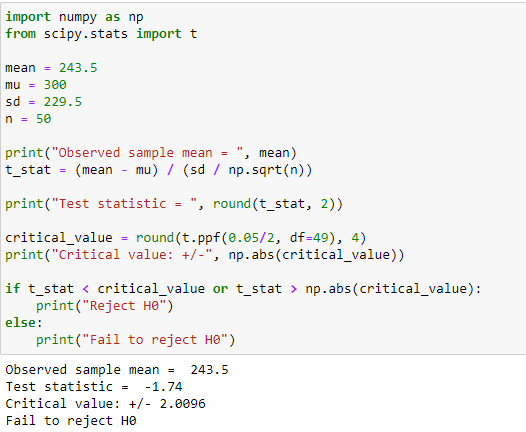
The average **cotinine level** in a group of **50 smokers** was **243.5 ng ml**.

Assuming that the **standard deviation** is known to be **229.5 ng ml**, at **95 %** confidence, test the assertion that the mean **cotinine level** of all smokers is equal to **300.0 ng ml**.

NOTE: When mean, standard deviation & nobs of only 1 dataset is given: use “ttest\_ind\_from\_stats” with nobs = very large number like 1M

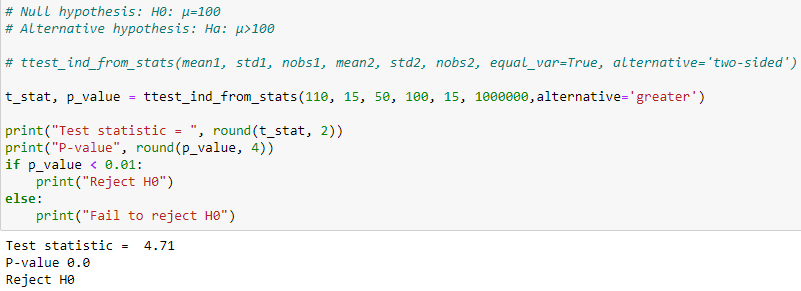


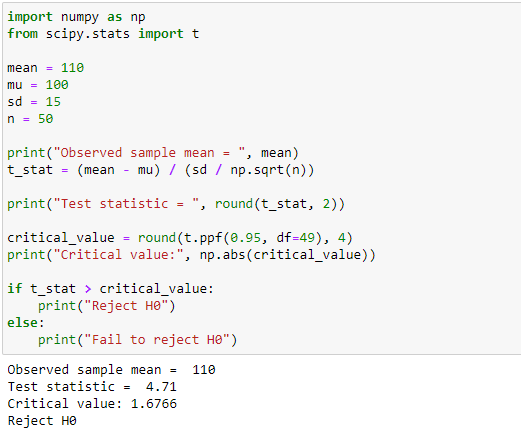
**OR**

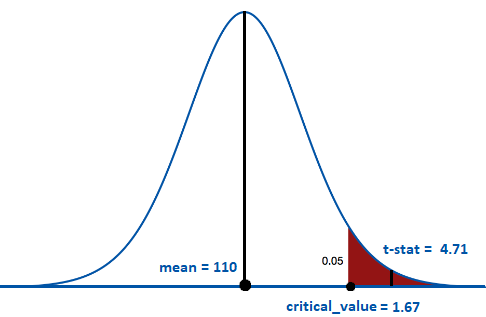


Q7. An institution claims that the candidates who study there have more IQ than the average institution. When the IQ of **50** students was calculated, the average turned out to be **110**.

The mean of the population IQ is **100** and the standard deviation is **15**. Find the validity of the claim with a significance level of **5%**.





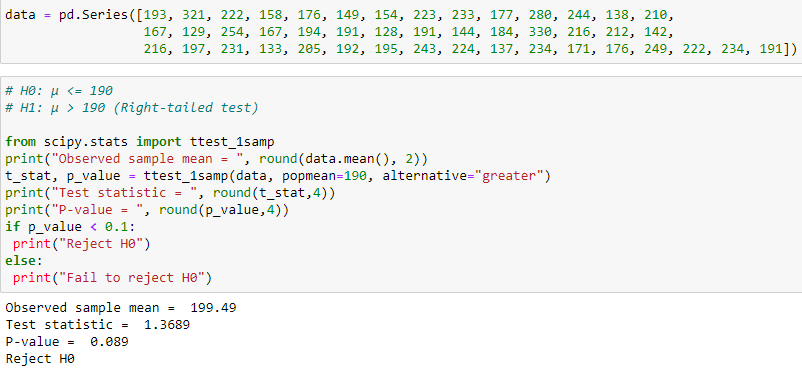


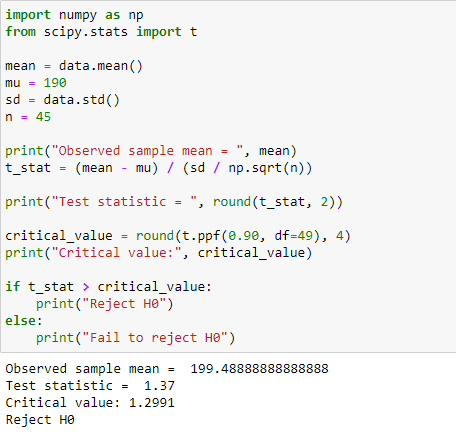
Q8. The quality assurance department claims that on average the non-fat milk contains **more than 190 mg** of Calcium per 500 ml packet.

To check this claim **45** packets of milk are collected and the content of calcium is recorded. perform the appropriate test with **90% confidence**.

data = [193, 321, 222, 158, 176, 149, 154, 223, 233, 177, 280, 244, 138, 210, 167, 129, 254, 167, 194, 191, 128, 191, 144, 184, 330, 216, 212, 142, 216, 197, 231, 133, 205, 192, 195, 243, 224, 137, 234, 171, 176, 249, 222, 234, 191]

NOTE: When only 1 dataset are given: use ttest\_1samp



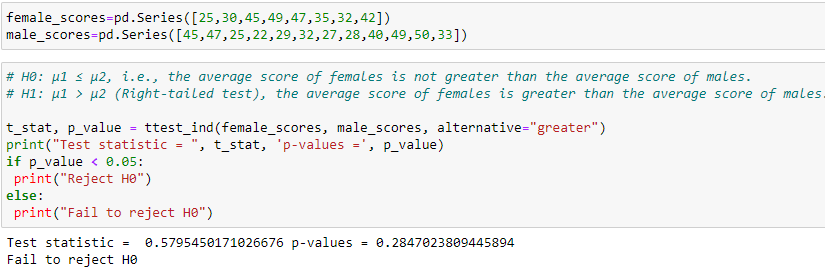


Q9. There are **8 females** and **12 males** in a coaching class. After a practice test, the coach wants to know whether the average score of females is **greater than** the average score of males.

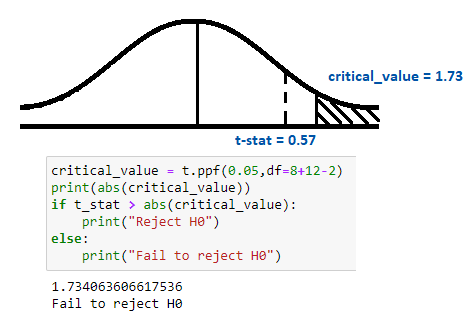
Given data describes the **scores** of females and males. Check whether the assumption of the coach is significant or not at a **5%** of **significance level**?

**Female scores = [25,30,45,49,47,35,32,42]  
male scores = [45,47,25,22,29,32,27,28,40,49,50,33]**

NOTE: When 2 datasets are given: use ttest\_ind

****

**OR**

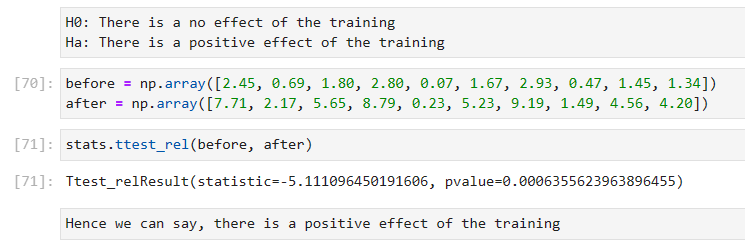
****

**Paired T-Test**

Q10. You are appointed as a Data Analyst for a training program deployed by the Government of India. The participants’ skills were tested **before** and **after** the training using some metrics on a scale of **10**.

before = [2.45, 0.69, 1.80, 2.80, 0.07, 1.67, 2.93, 0.47, 1.45, 1.34]

after = [7.71, 2.17, 5.65, 8.79, 0.23, 5.23, 9.19, 1.49, 4.56, 4.20]



**IMP NOTE: -**

