Problem Statement 1:

You survey households in your area to find the average rent they are paying. Find the

standard deviation from the following data:

$1550, $1700, $900, $850, $1000, $950.

Code:

import numpy as np

import pandas as pd

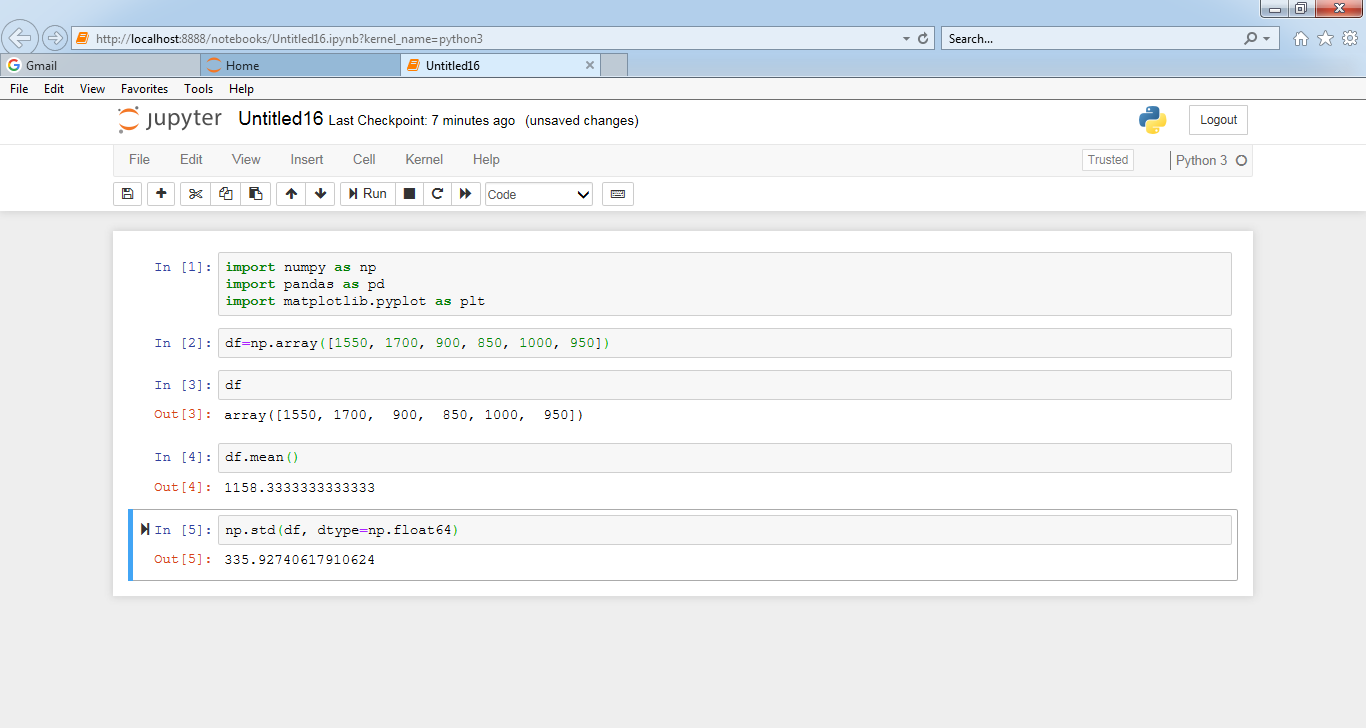
import matplotlib.pyplot as plt

df=np.array([1550, 1700, 900, 850, 1000, 950])

df

df.mean()

np.std(df, dtype=np.float64)



Problem Statement 2:

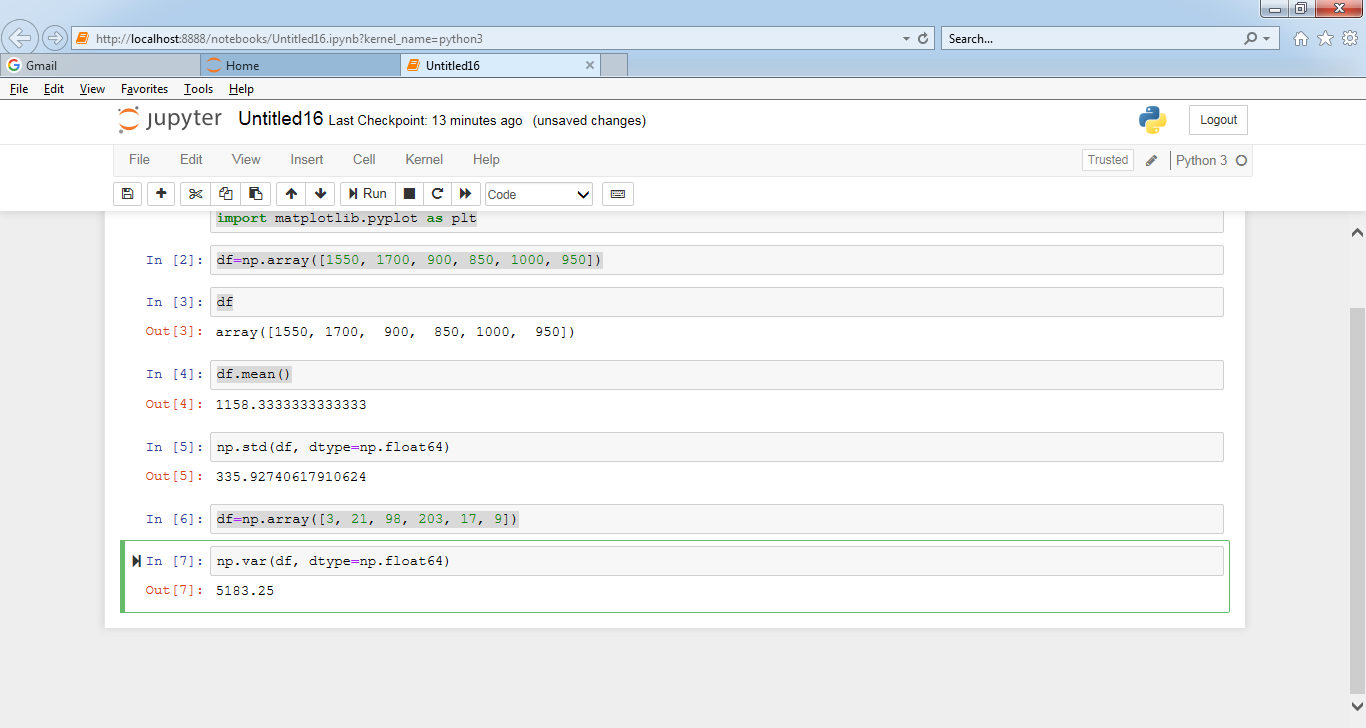
Find the variance for the following set of data representing trees in California (heights in

feet):

3, 21, 98, 203, 17, 9

df=np.array([3, 21, 98, 203, 17, 9])

np.var(df, dtype=np.float64)



**Problem Statement 3:**

In a class on 100 students, 80 students passed in all subjects, 10 failed in one subject, 7

failed in two subjects and 3 failed in three subjects. Find the probability distribution of

the variable for number of subjects a student from the given class has failed in.

Code:

from scipy.stats import norm

Tot\_students= 100

failNOsubject= 80

failONEsubect=10

failTWOsubject=7

failTHREEsubject=3

def fail\_probability(fail\_subect, students):

probability = (fail\_subect / students)

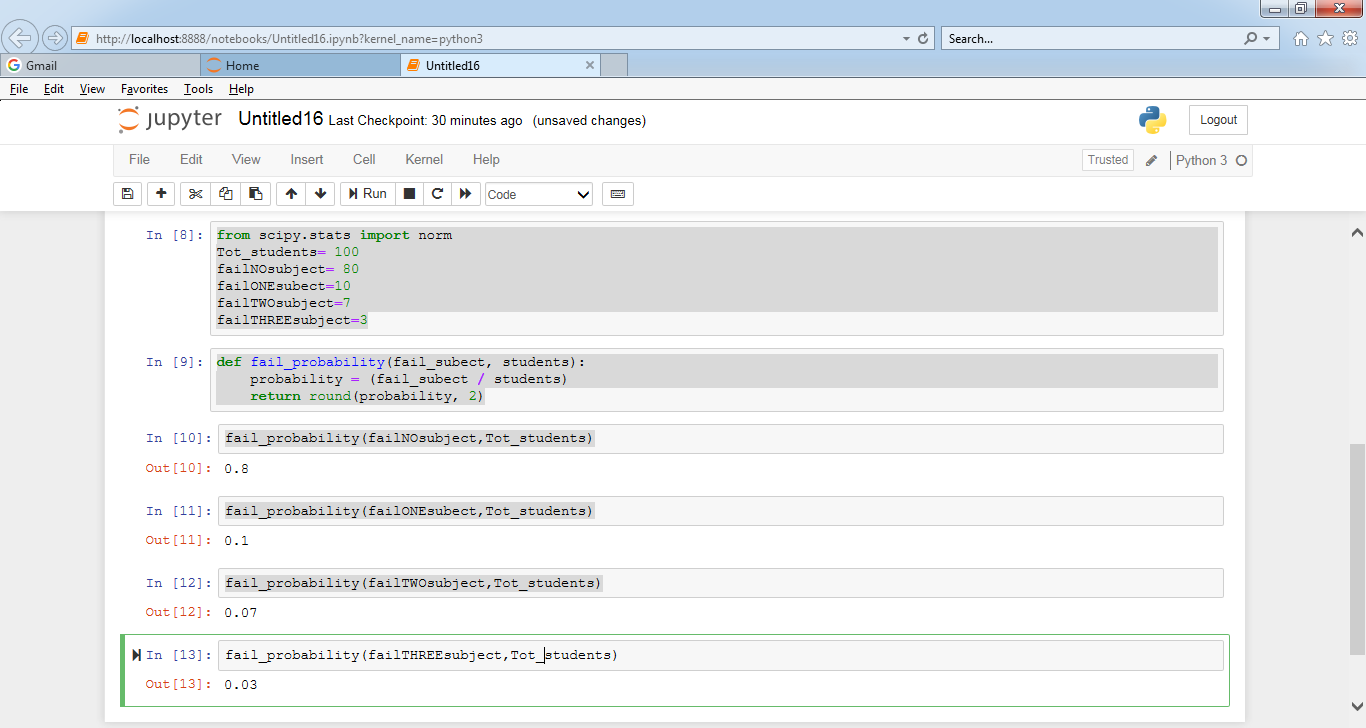
return round(probability, 2)

fail\_probability(failNOsubject,Tot\_students)

fail\_probability(failONEsubect,Tot\_students)

fail\_probability(failTWOsubject,Tot\_students)

fail\_probability(failTHREEsubject,Tot\_students)



Note: Solution submitted via github must contain all the detailed steps.