# M-3 Lab: Gathering Information for Better Data Understanding

=>Write a python script that uses padas data frame to find the following

Total number of purchase per product type

1.) There are three product types sold by this etsy store: Framed Posters, Canvas Paintings & Poster Prints.

Find the total number of purchases for each product type.

import pandas as pd   
  
df\_product = pd.read\_csv("/home/swetha/Documents/sales.csv");  
df\_product

purchase\_date product\_type  
0 01/01/2020 Canvas Painting  
1 01/01/2020 Framed Poster  
2 01/01/2020 Framed Poster  
3 01/01/2020 Framed Poster  
4 01/01/2020 Canvas Painting  
... ... ...  
3964 12/31/2020 Framed Poster  
3965 12/31/2020 Canvas Painting  
3966 12/31/2020 Framed Poster  
3967 12/31/2020 Framed Poster  
3968 12/31/2020 Framed Poster  
  
[3969 rows x 2 columns]

totalPurchasesOfProductType = df\_product['product\_type'].value\_counts();

totalPurchasesOfProductType

Framed Poster 2351  
Canvas Painting 1215  
Poster Print 403  
Name: product\_type, dtype: int64

Which product has the highest purchase count?

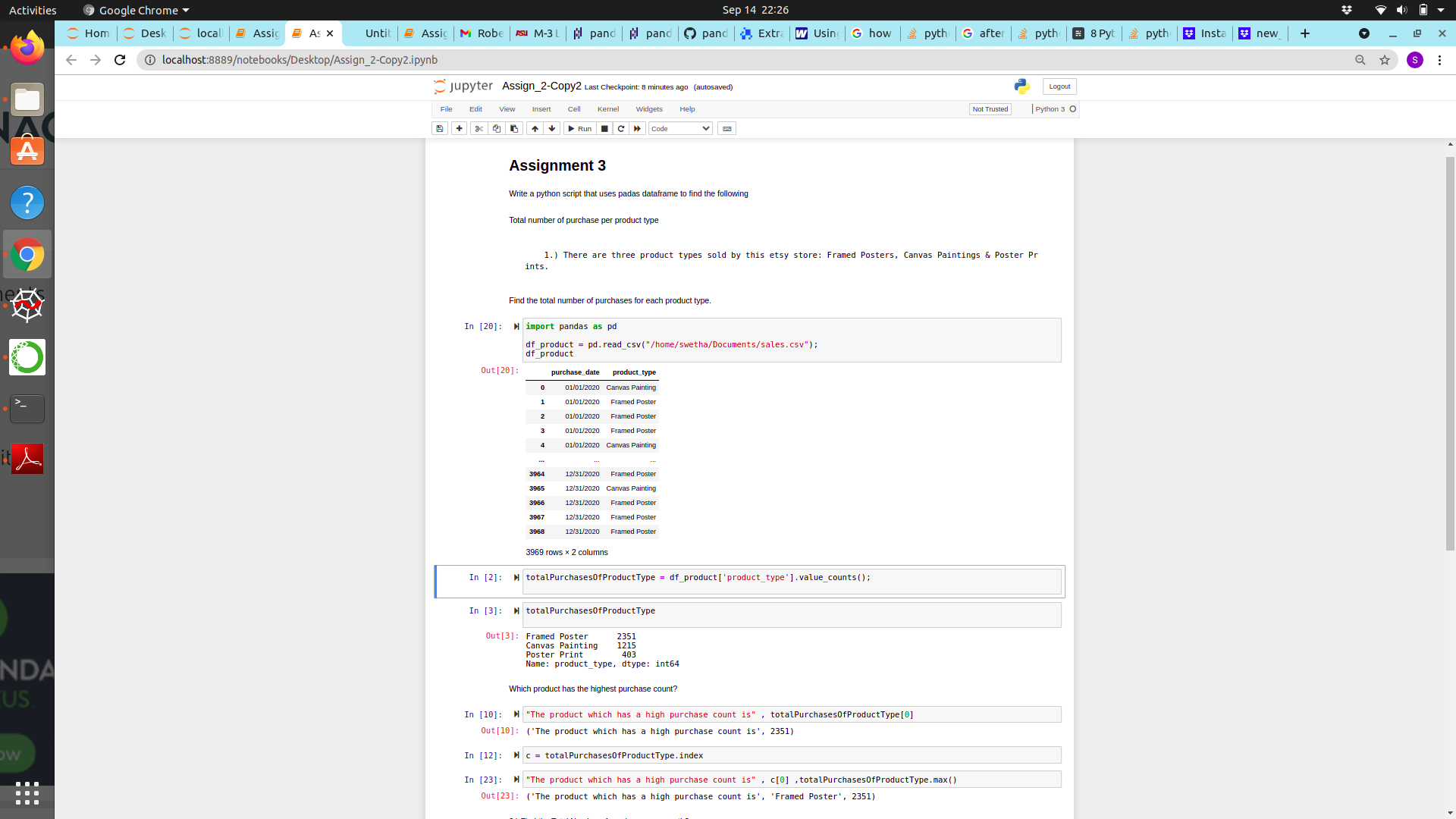
"The product which has a high purchase count is" , totalPurchasesOfProductType[0]

('The product which has a high purchase count is', 2351)

c = totalPurchasesOfProductType.index

"The product which has a high purchase count is" , c[0] ,totalPurchasesOfProductType.max()

('The product which has a high purchase count is', 'Framed Poster', 2351)



2.) Find the Total Number of purchases per month?

import pandas as pd   
  
df\_products = pd.read\_csv("/home/swetha/Documents/sales.csv")

purchase\_date = df\_product['purchase\_date']  
purchase\_date

0 01/01/2020  
1 01/01/2020  
2 01/01/2020  
3 01/01/2020  
4 01/01/2020  
 ...   
3964 12/31/2020  
3965 12/31/2020  
3966 12/31/2020  
3967 12/31/2020  
3968 12/31/2020  
Name: purchase\_date, Length: 3969, dtype: object

totalPurchaseCountPerMonth = pd.DatetimeIndex(purchase\_date).month\_name().value\_counts();  
totalPurchaseCountPerMonth

April 382  
March 365  
October 346  
July 345  
May 345  
December 342  
June 323  
February 323  
January 321  
September 319  
August 288  
November 270  
Name: purchase\_date, dtype: int64

