**DATA HANDLING USING ‘Pandas’ and DATA VISUALIZATION USING ‘Seaborn’**

**Using the pandas function read\_csv(), read the given ‘iris’ data set.**

1. **Use appropriate functions in pandas to display**
2. **Shape of the data set**
3. **First 5 and last five rows of data set(head and tail)**
4. **Size of dataset**
5. **No:of samples available for each variety**
6. **Description of the data set( use describe**

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import pandas as pd

col=['sepal\_length','sepal\_width','petal\_length','petal\_width','type']

iris=pd.read\_csv("iris.csv",names=col)

print("shape:",iris.shape)

print("\*\*\*\*\*\*\*\*\*")

print("First five rows")

print(iris.head())

print("\*\*\*\*\*\*\*\*\*")

print("Last five rows")

print(iris.tail())

print("\*\*\*\*\*\*\*\*\*")

print("Size:",iris.size)

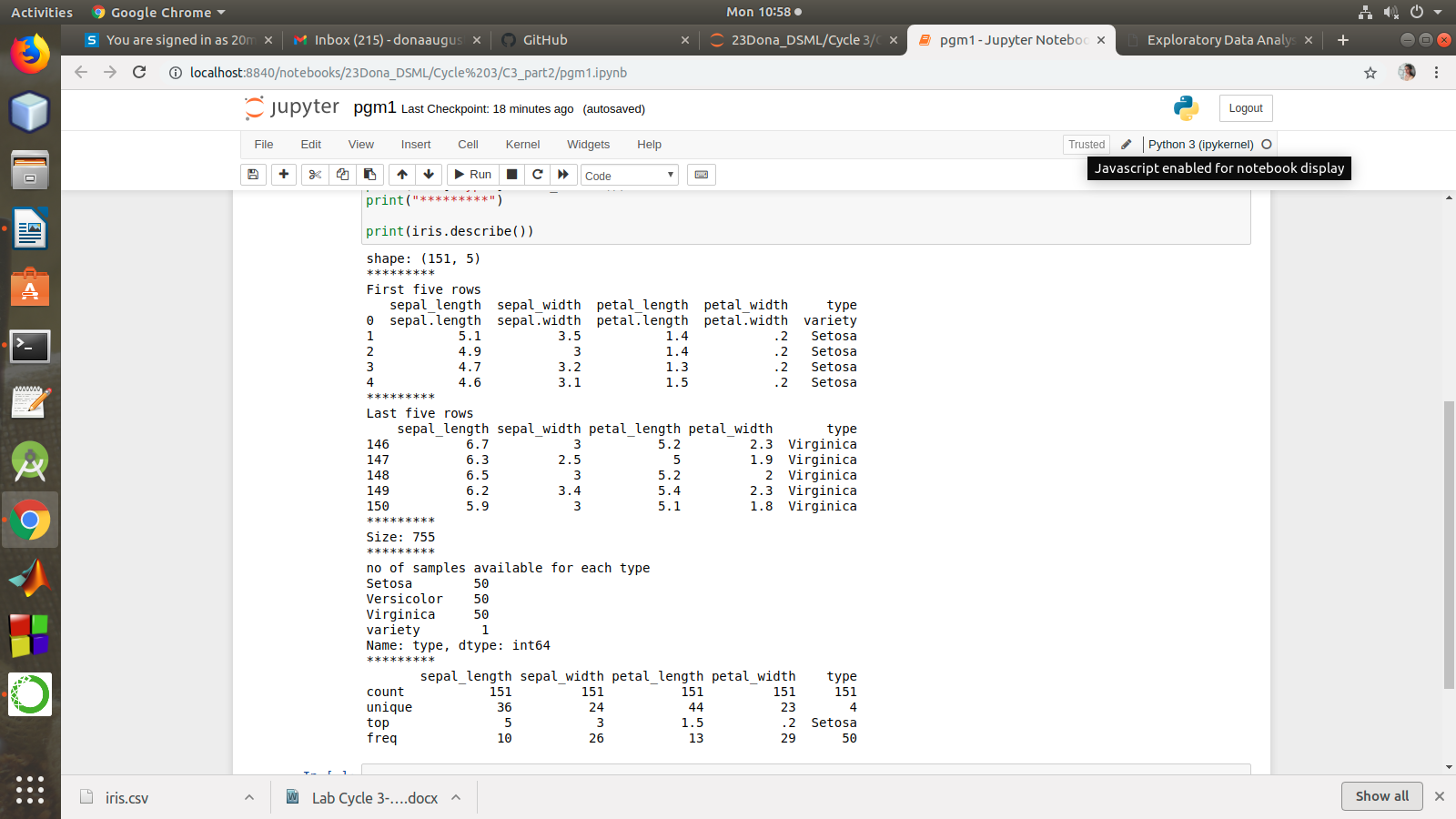
print("\*\*\*\*\*\*\*\*\*")

print("no of samples available for each type")

print(iris["type"].value\_counts())

print("\*\*\*\*\*\*\*\*\*")

print(iris.describe())



1. Use pairplot() function to display pairwise relationships between attributes. Try different kind of plots {***‘scatter’, ‘kde’, ‘hist’, ‘reg’}*** and different kind of markers

import numpy as np

import pandas as pd

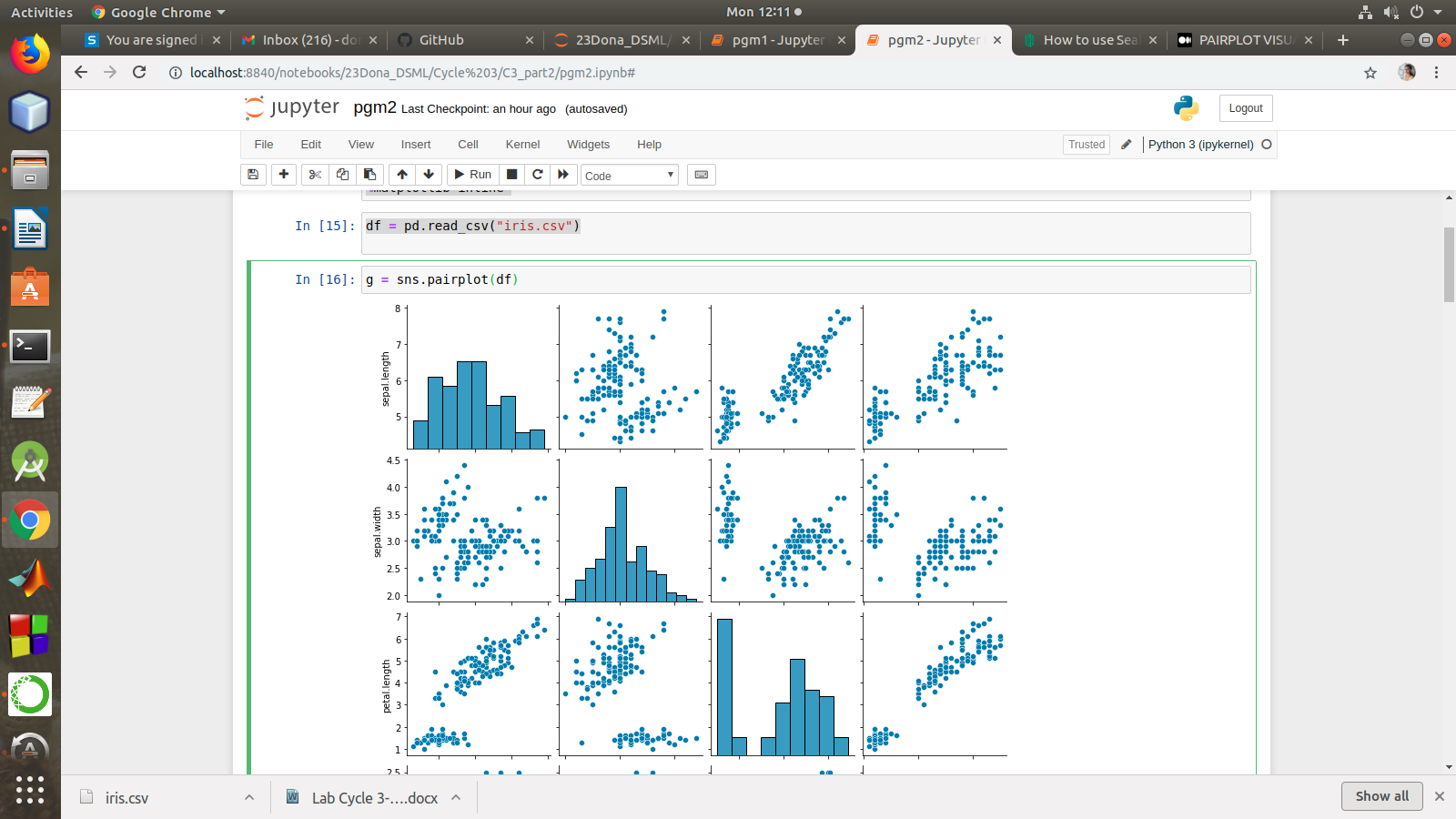
import seaborn as sns

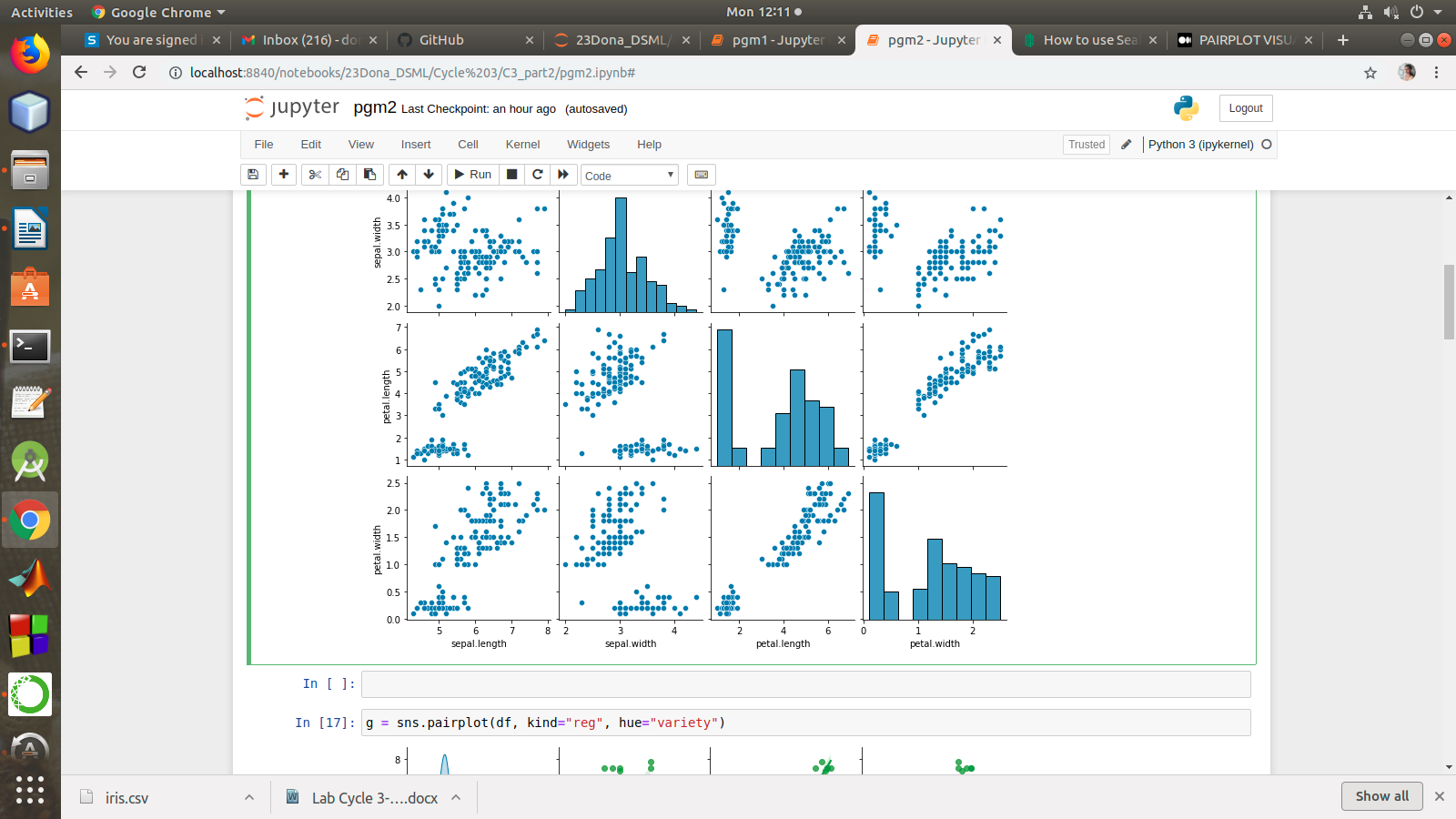
import matplotlib.pyplot as plt

%matplotlib inline

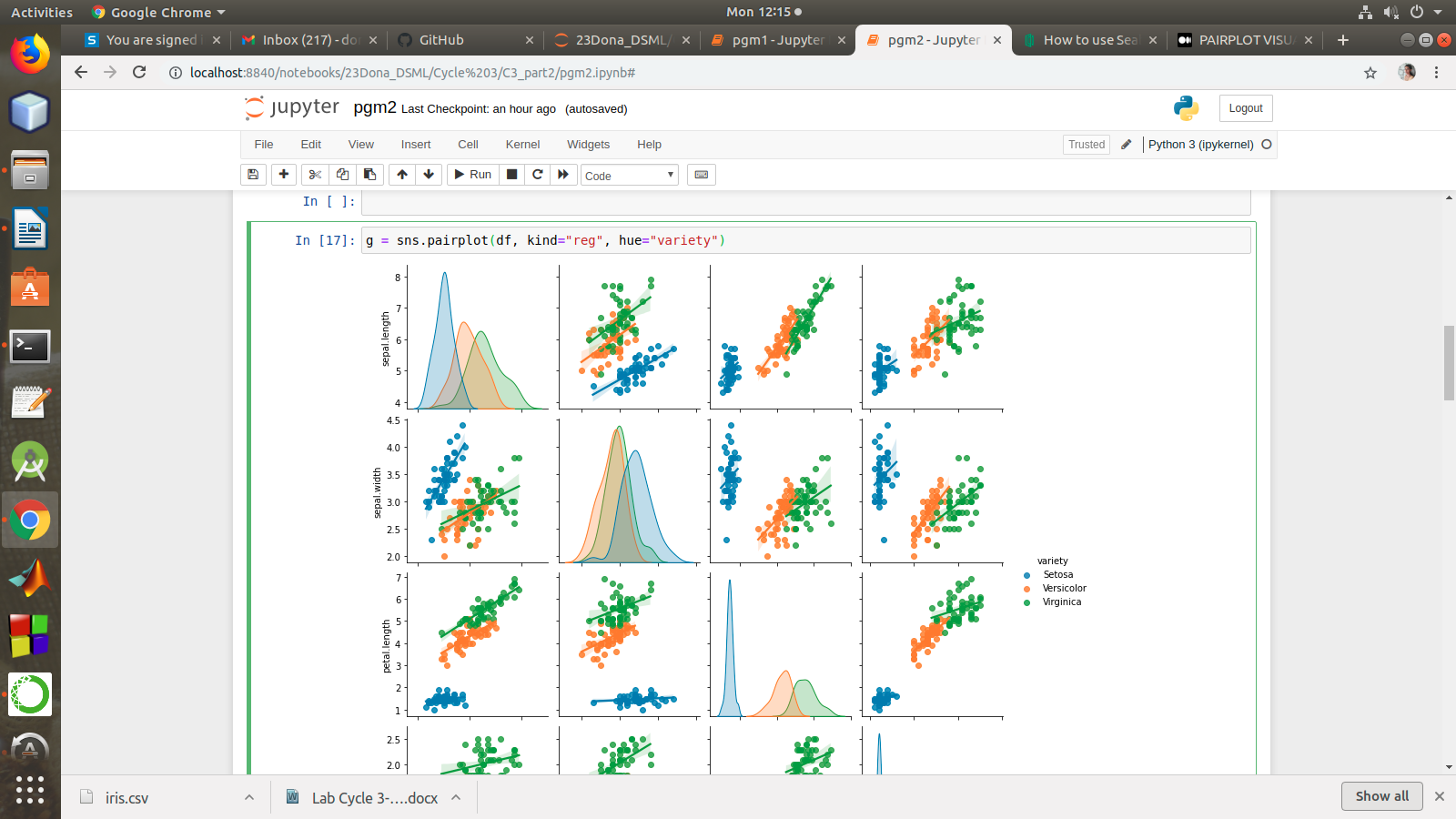
df = pd.read\_csv("iris.csv")

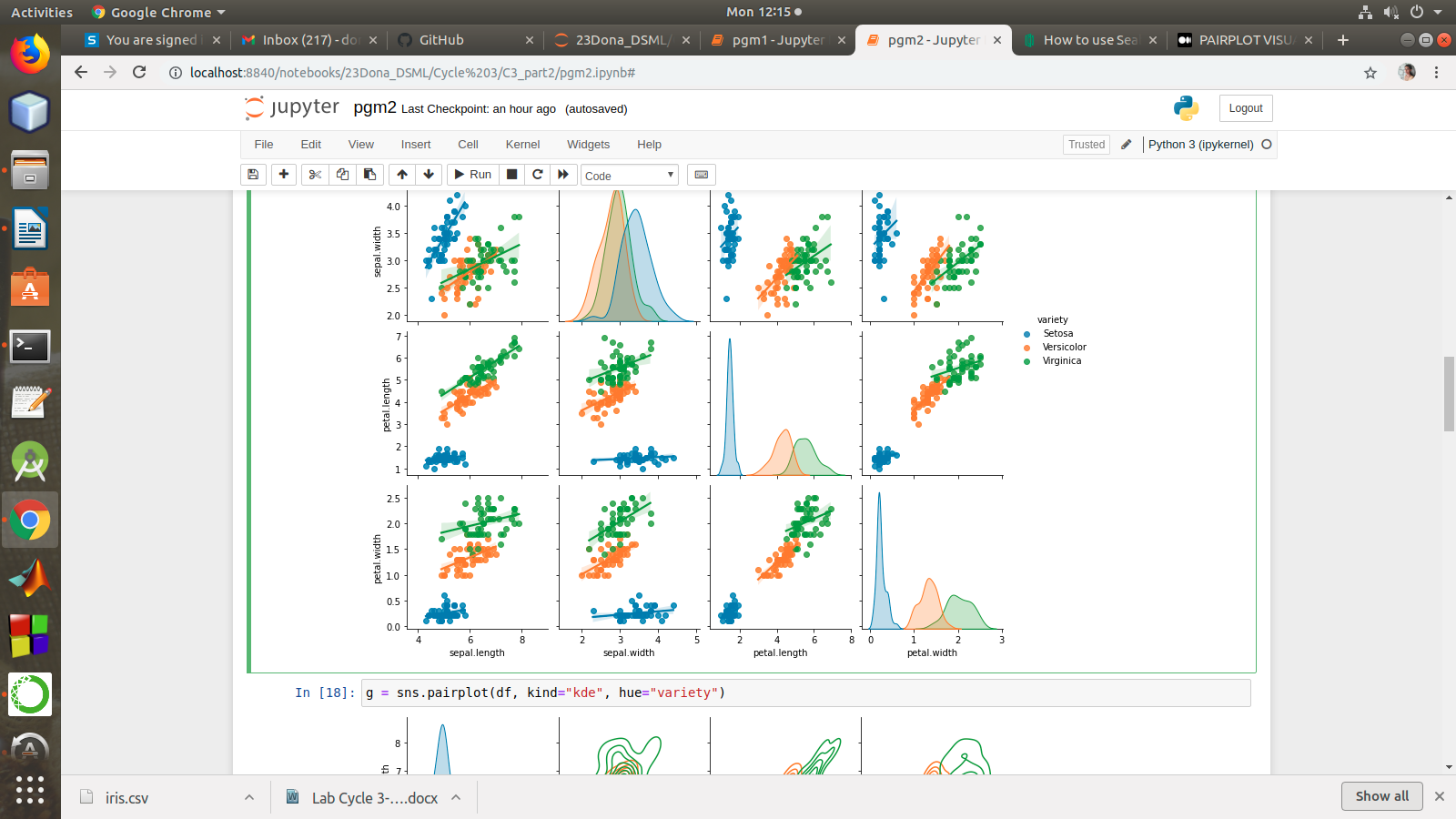
g = sns.pairplot(df)



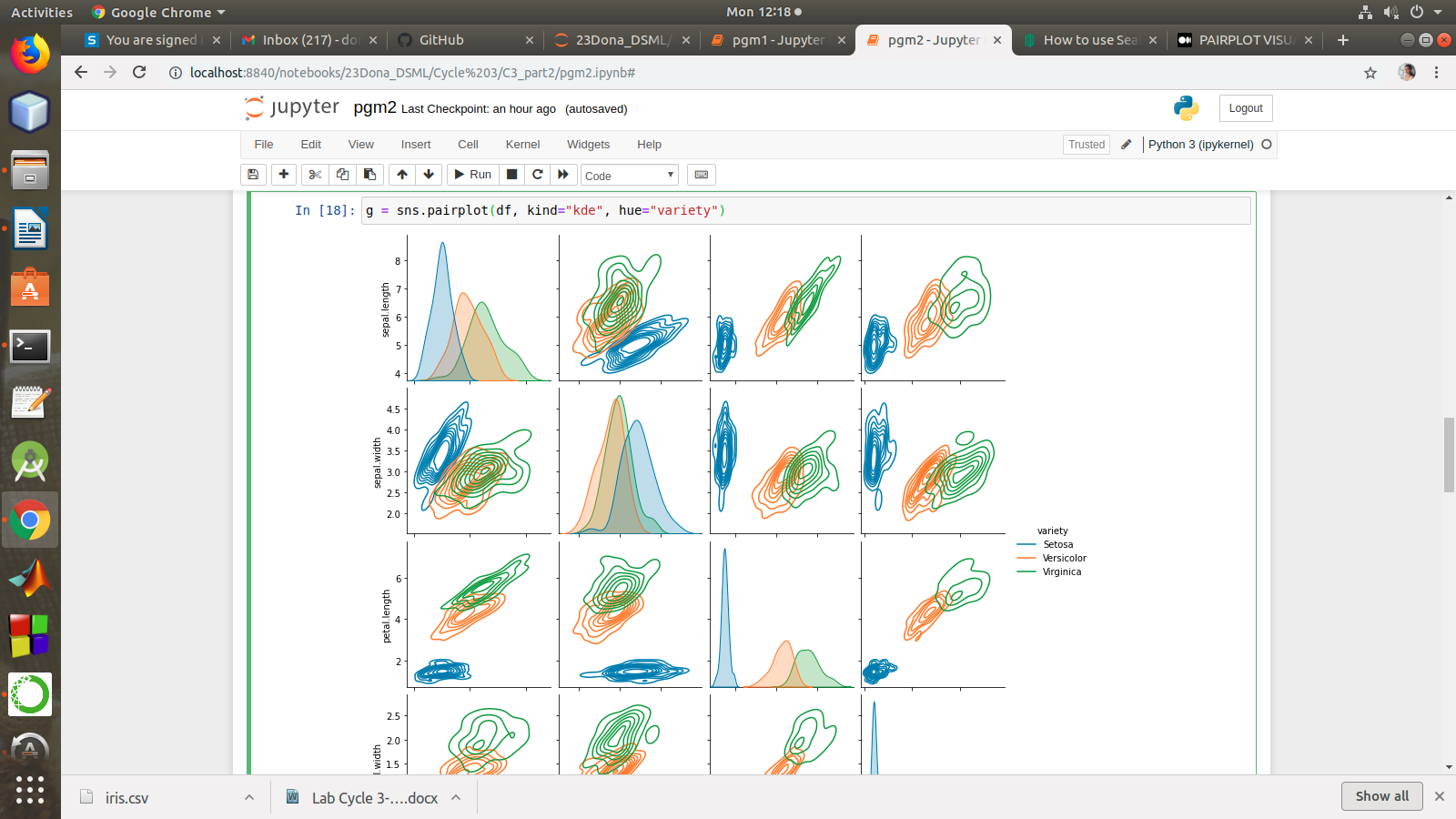


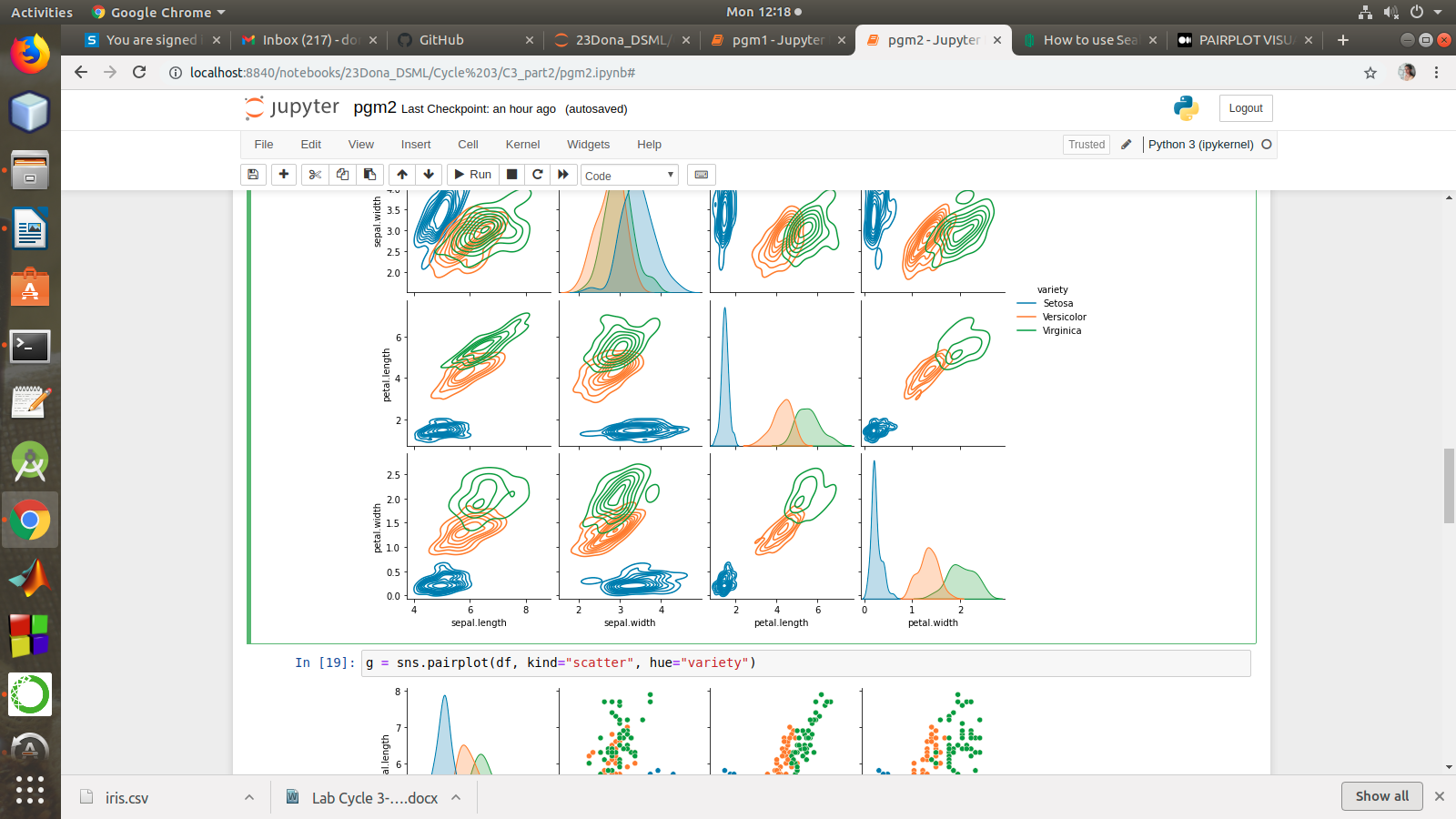
g = sns.pairplot(df, kind="reg", hue="variety")



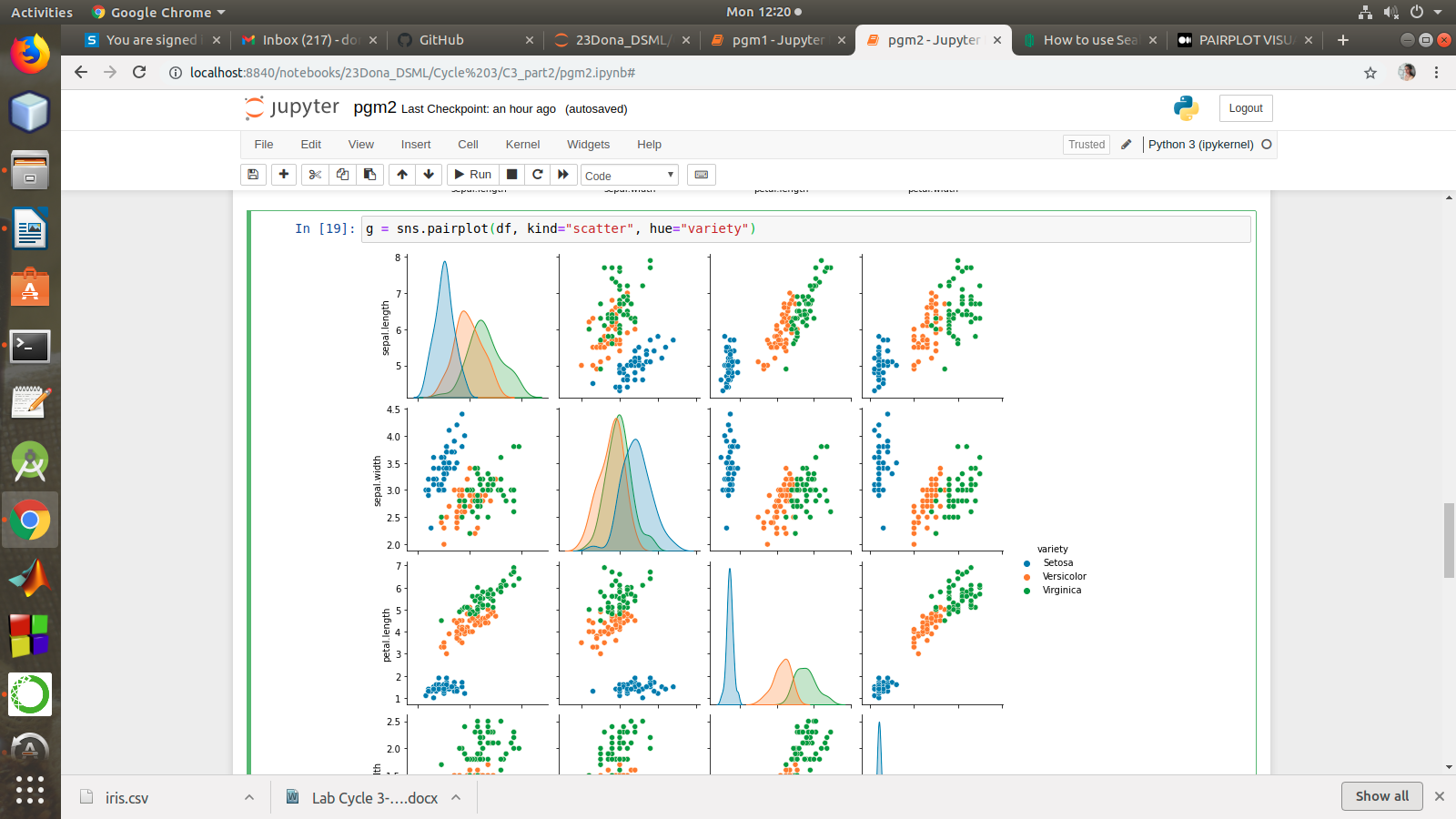


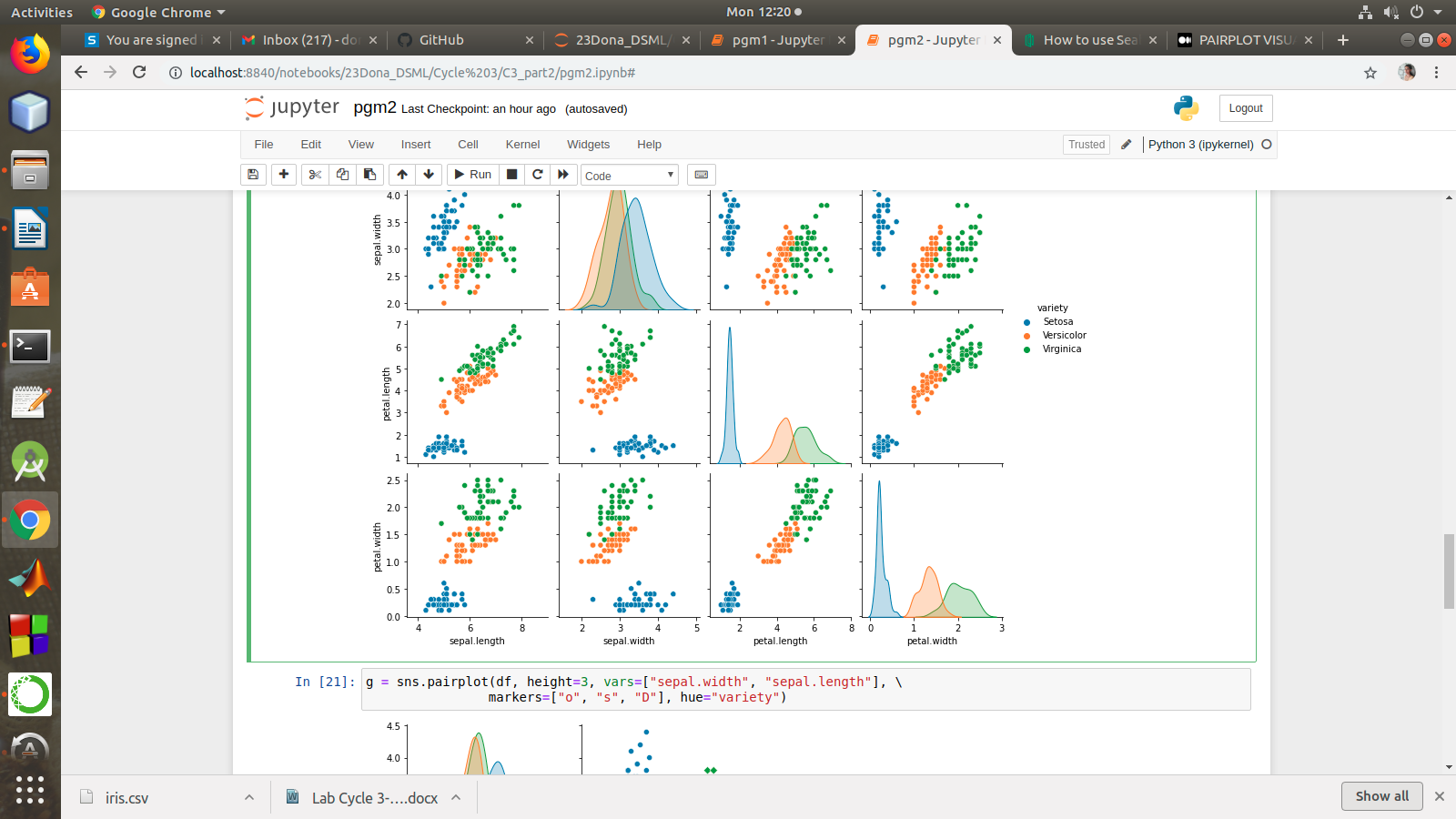
g = sns.pairplot(df, kind="kde", hue="variety")



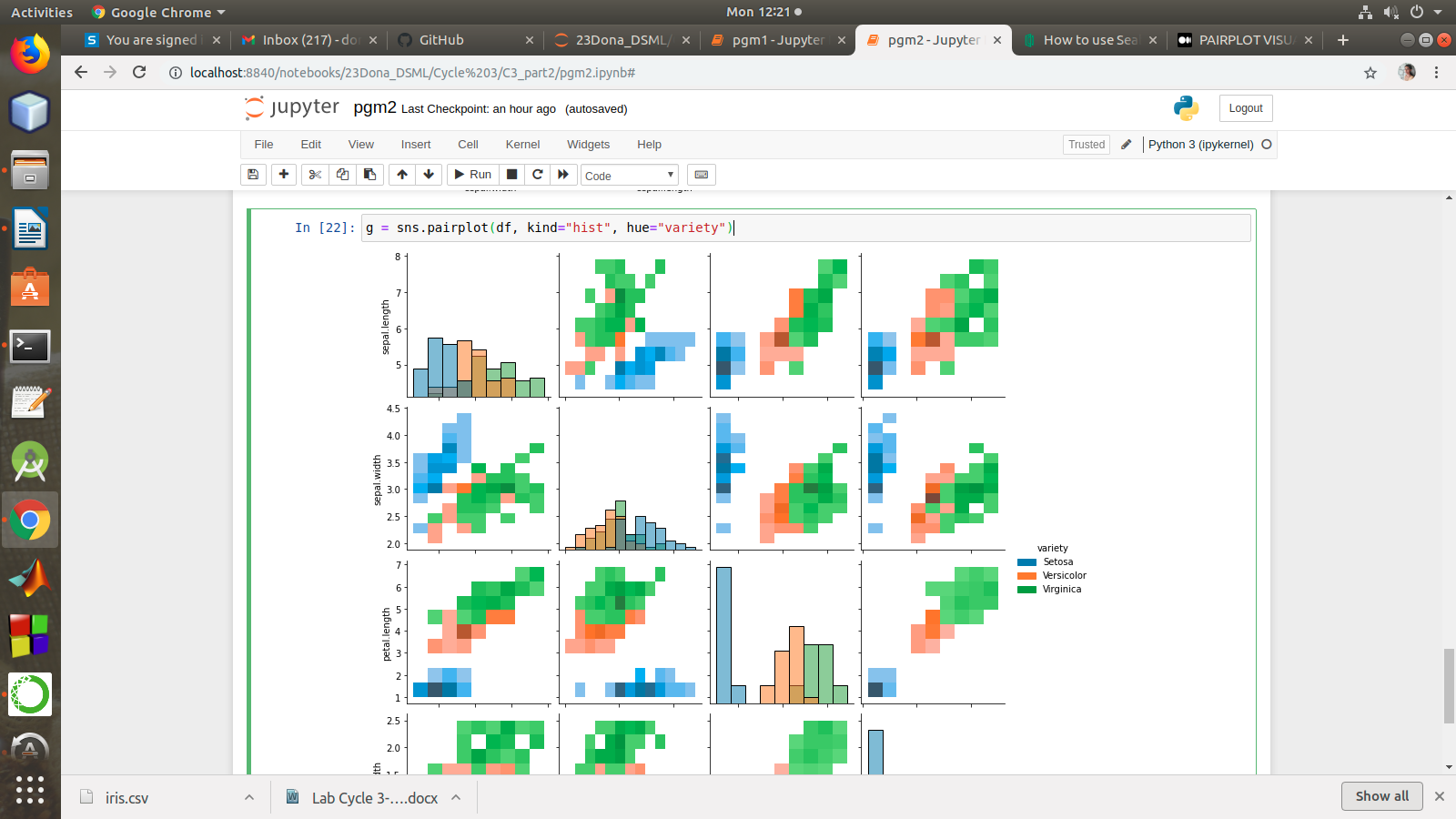


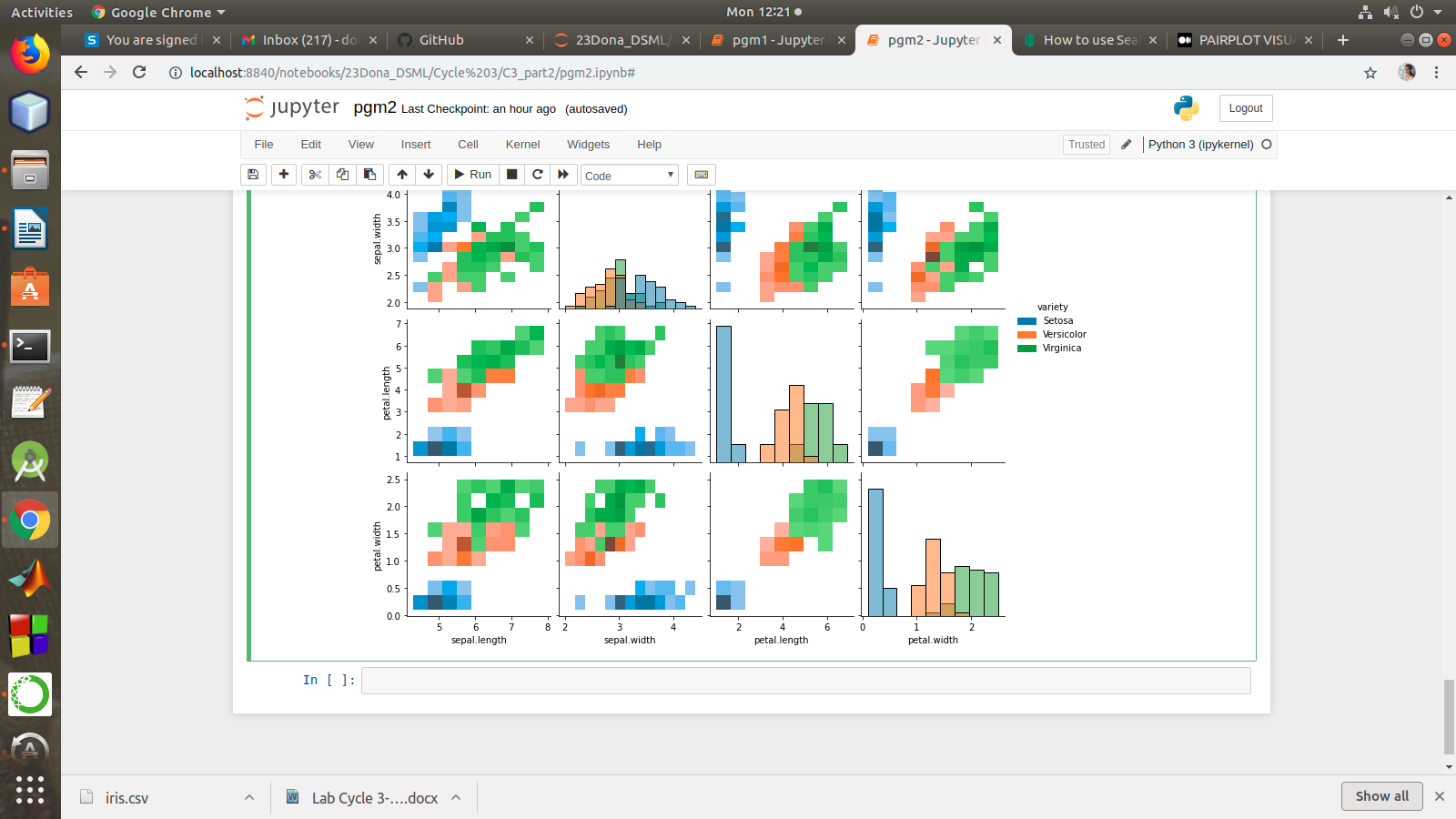
g = sns.pairplot(df, kind="scatter", hue="variety")





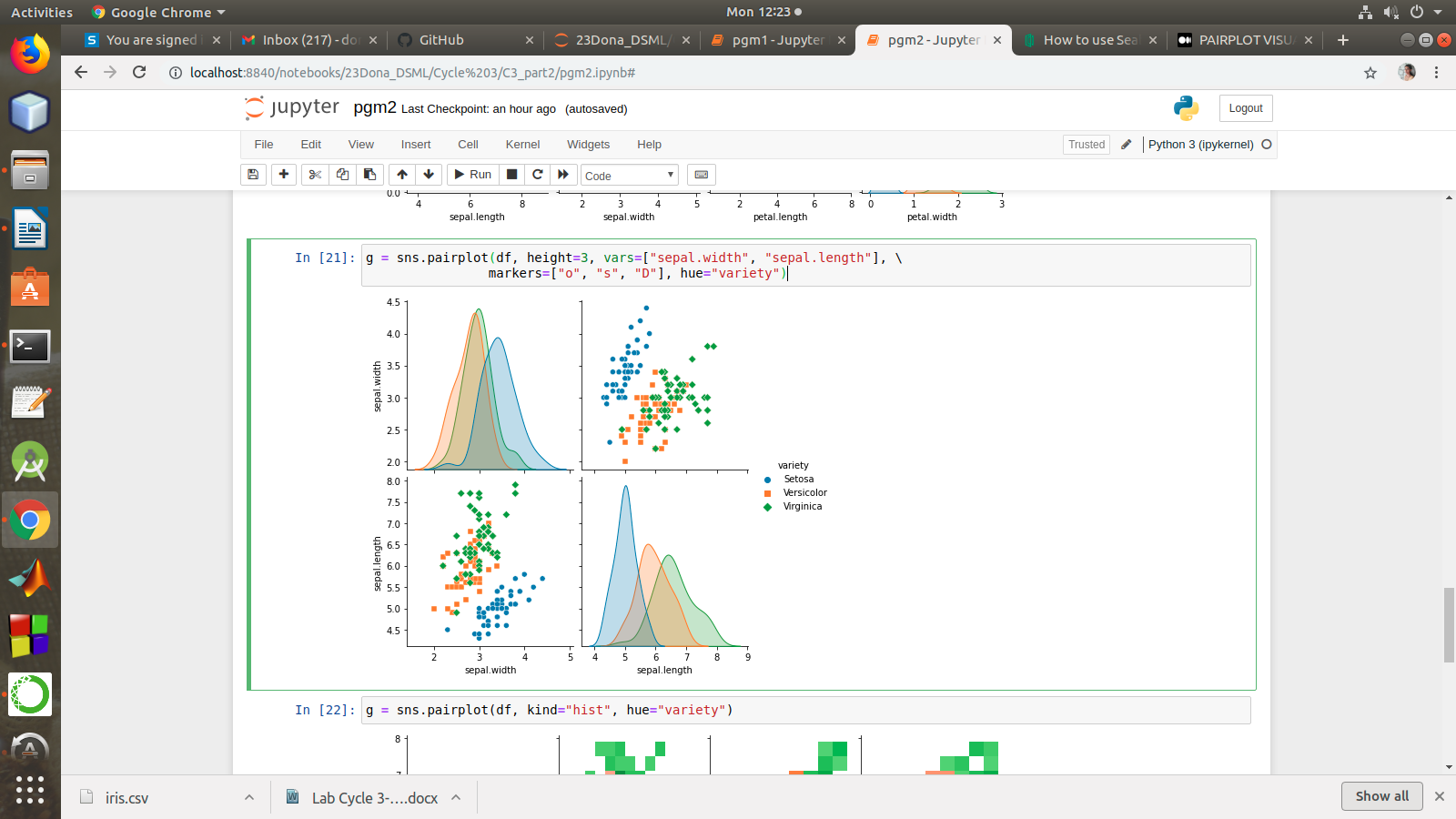
g = sns.pairplot(df, kind="hist", hue="variety")





g = sns.pairplot(df, height=3, vars=["sepal.width", "sepal.length"], \

markers=["o", "s", "D"], hue="variety")



#### using the iris data set,get familiarize with functions:

1)displot()

2) histplot()

3) relplot()

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

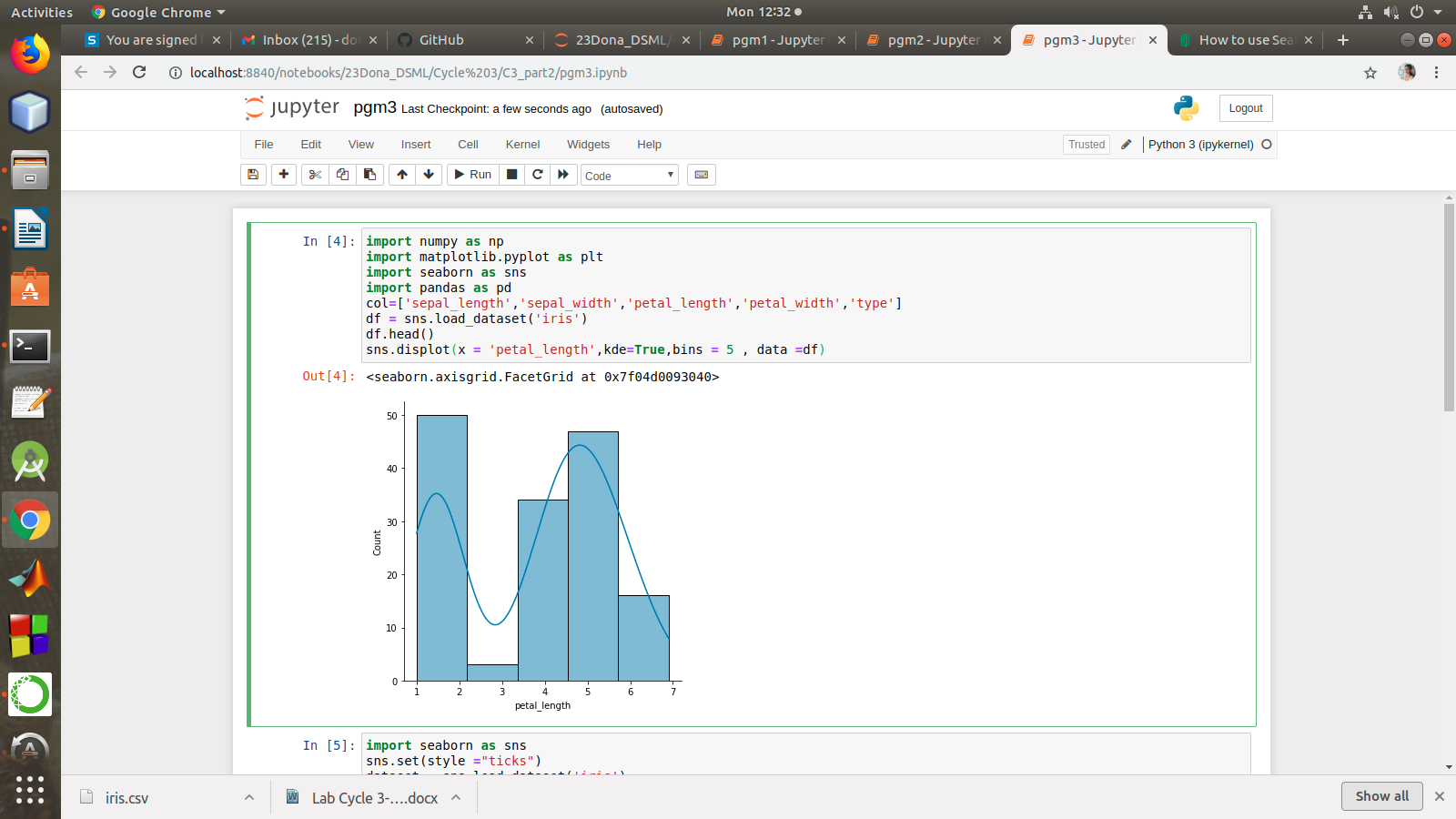
import pandas as pd

col=['sepal\_length','sepal\_width','petal\_length','petal\_width','type']

df = sns.load\_dataset('iris')

df.head()

sns.displot(x = 'petal\_length',kde=True,bins = 5 , data =df)



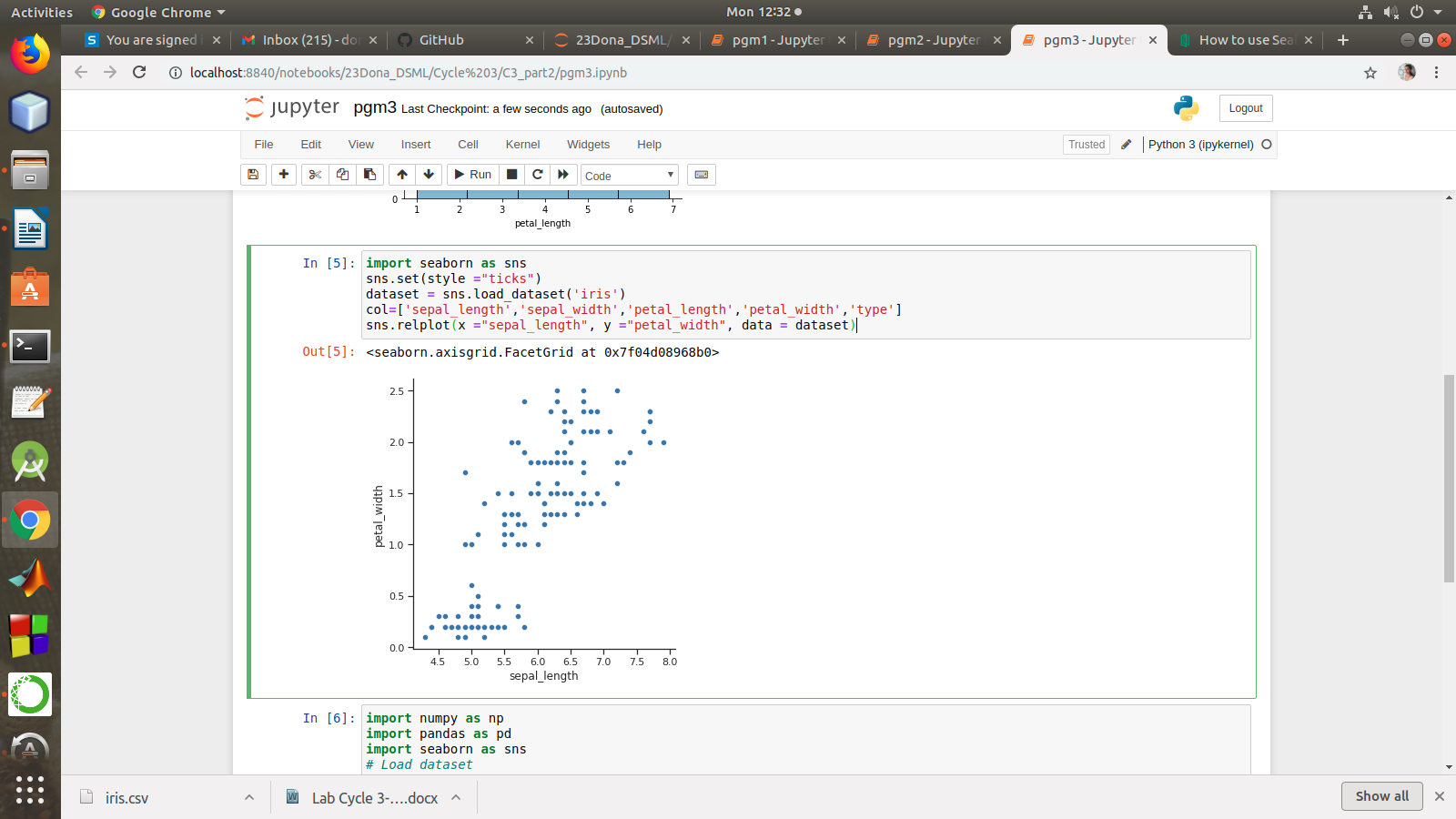
import seaborn as sns

sns.set(style ="ticks")

dataset = sns.load\_dataset('iris')

col=['sepal\_length','sepal\_width','petal\_length','petal\_width','type']

sns.relplot(x ="sepal\_length", y ="petal\_width", data = dataset)



import numpy as np

import pandas as pd

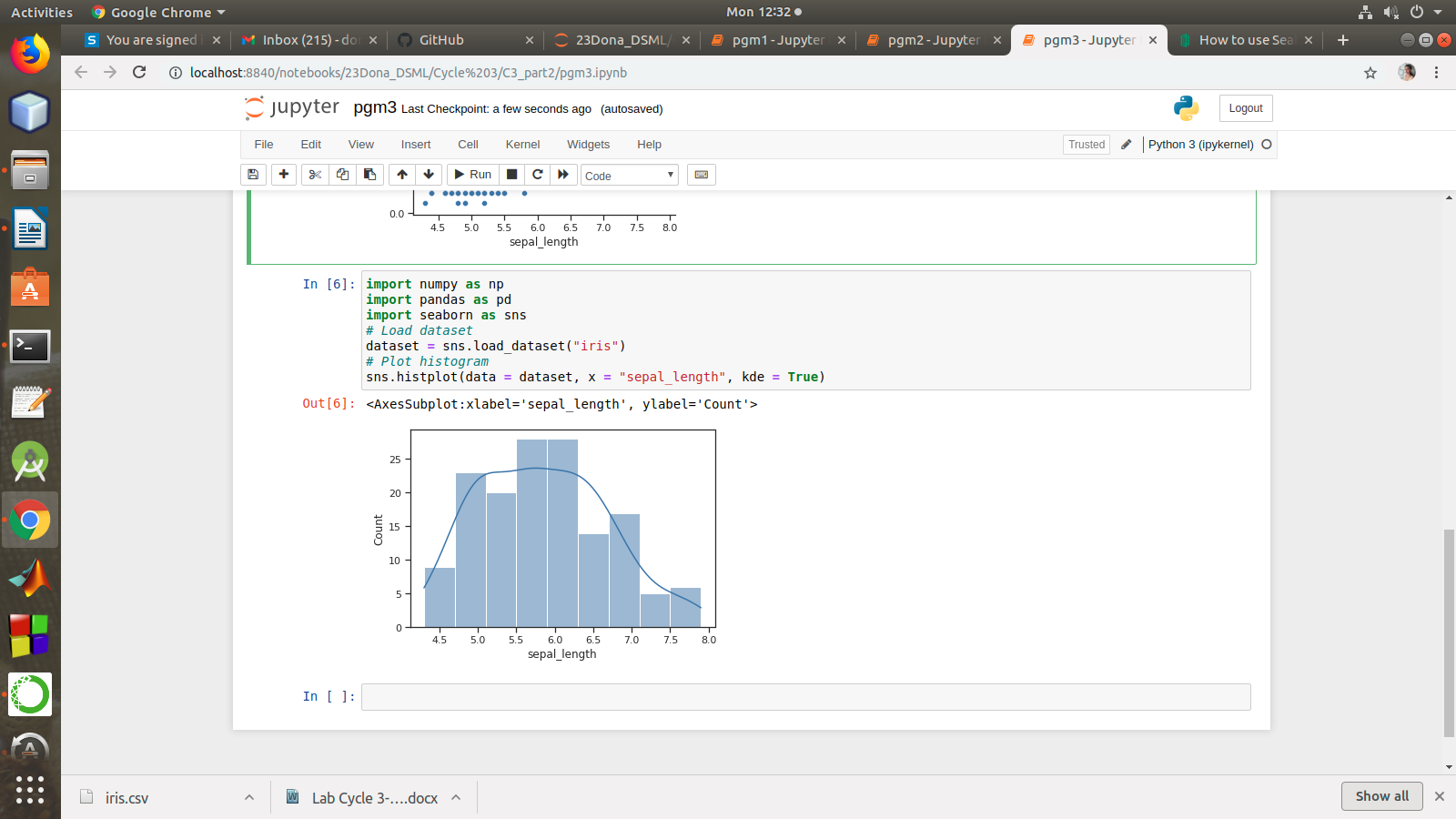
import seaborn as sns

# Load dataset

dataset = sns.load\_dataset("iris")

# Plot histogram

sns.histplot(data = dataset, x = "sepal\_length", kde = True)



Reference:

https://medium.com/@avulurivenkatasaireddy/exploratory-data-analysis-of-iris-data-set-using-python-823e54110d2d

https://web.ics.purdue.edu/~yrosokha/code/Seaborn\_Example\_1.html

https://www.section.io/engineering-education/seaborn-tutorial/

https://seaborn.pydata.org/introduction.html

https://towardsdatascience.com/seaborn-python-8563c3d0ad41