**Your Task - 7:** You your is to learn and understand about the countplot and Heatmap functions and apply it on any dataset you want and sent me a report and what have you learned and what have you concluded.

To

Rishabh Goyal,

From

Pampana Jai Kiran.

Sir, I have successfully completed the task 7.

Code:

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

data = pd.read\_csv('/content/Practice dataset 5.csv')

df = data

df

sns.countplot(x='gender',data=df,hue='diabetes',palette="Set2")

plt.show()

sns.countplot(y='gender',data=df,hue='smoking\_history',palette="Set1")

plt.show()

correlation\_matrix = data.corr()

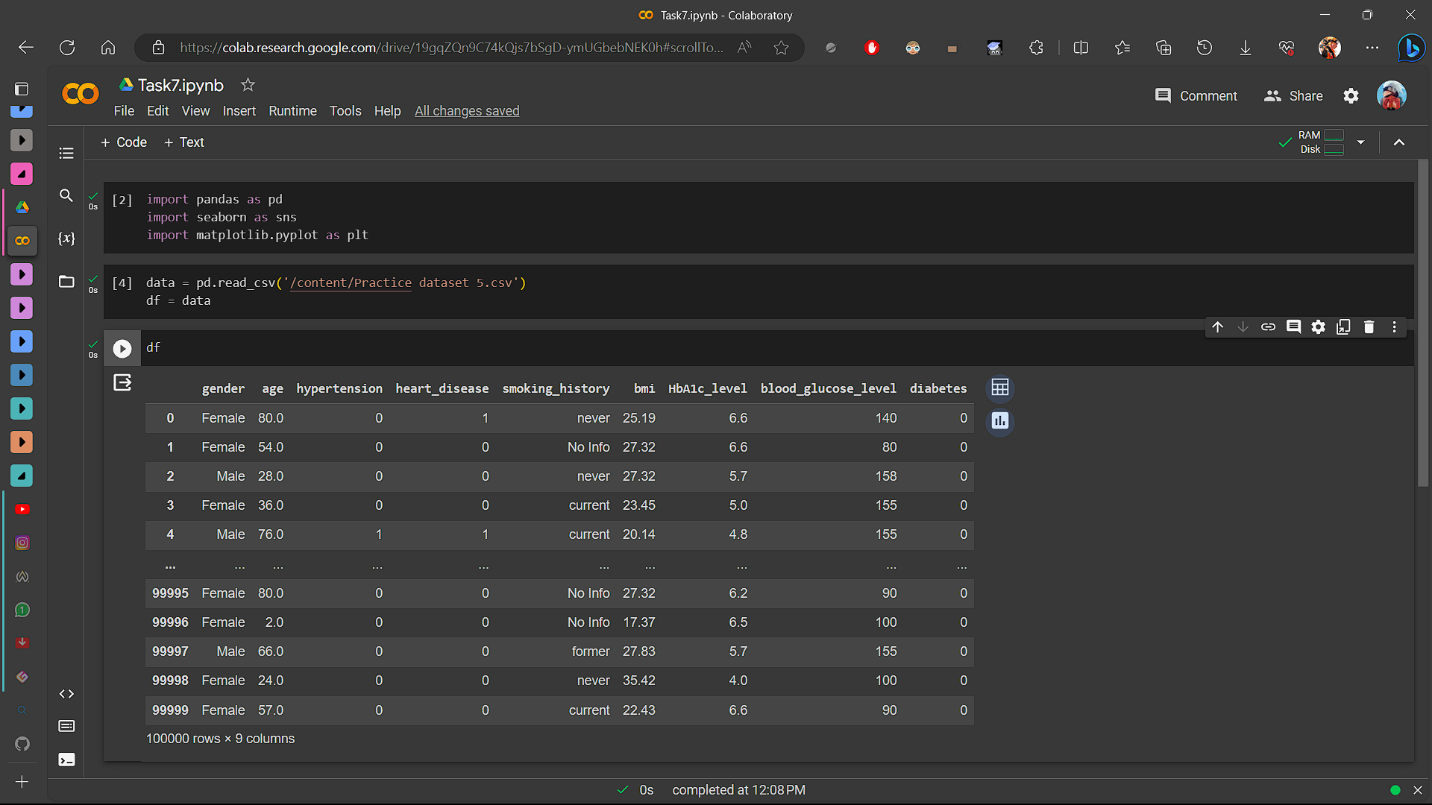
print(correlation\_matrix)

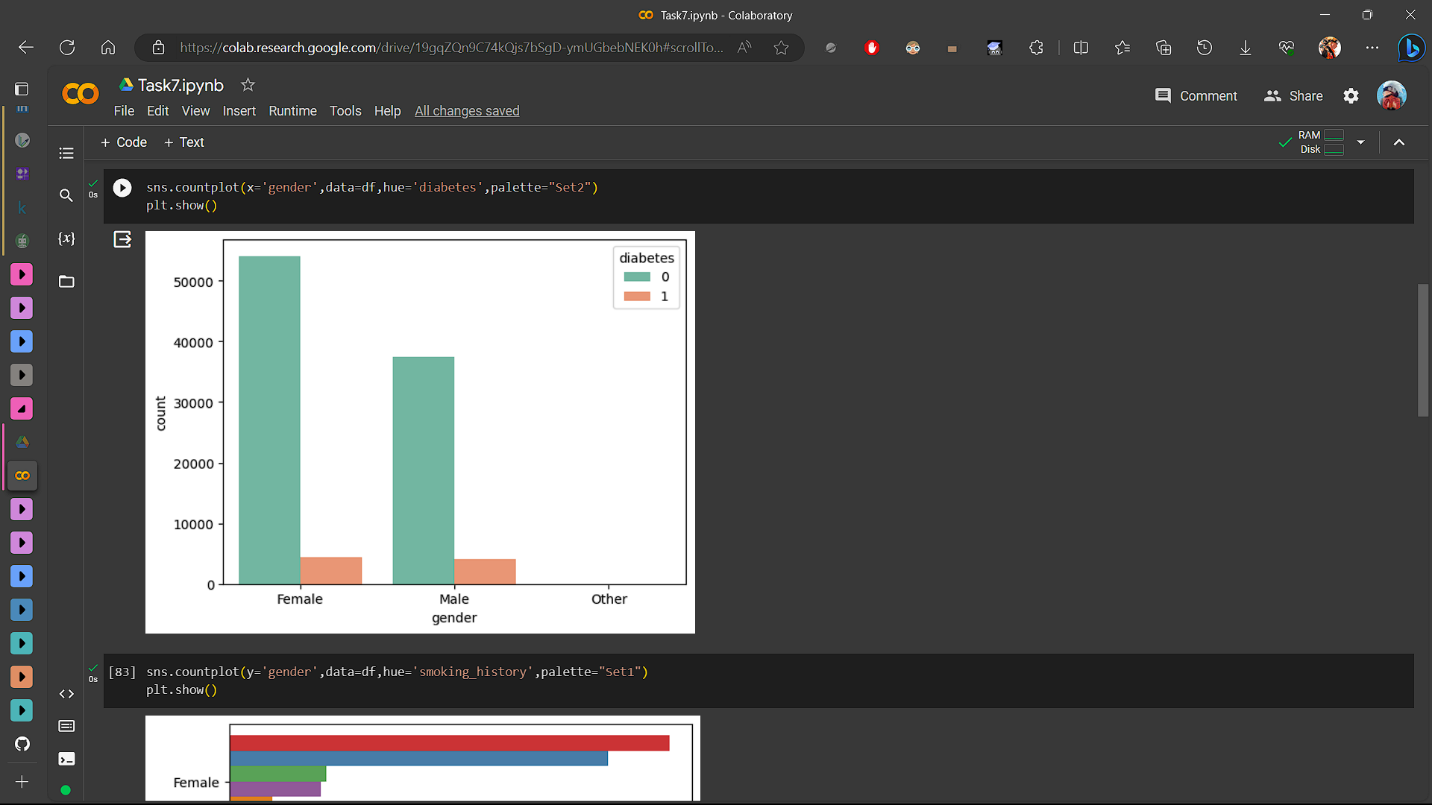
sns.heatmap(data.corr(),annot=True, cmap="coolwarm")

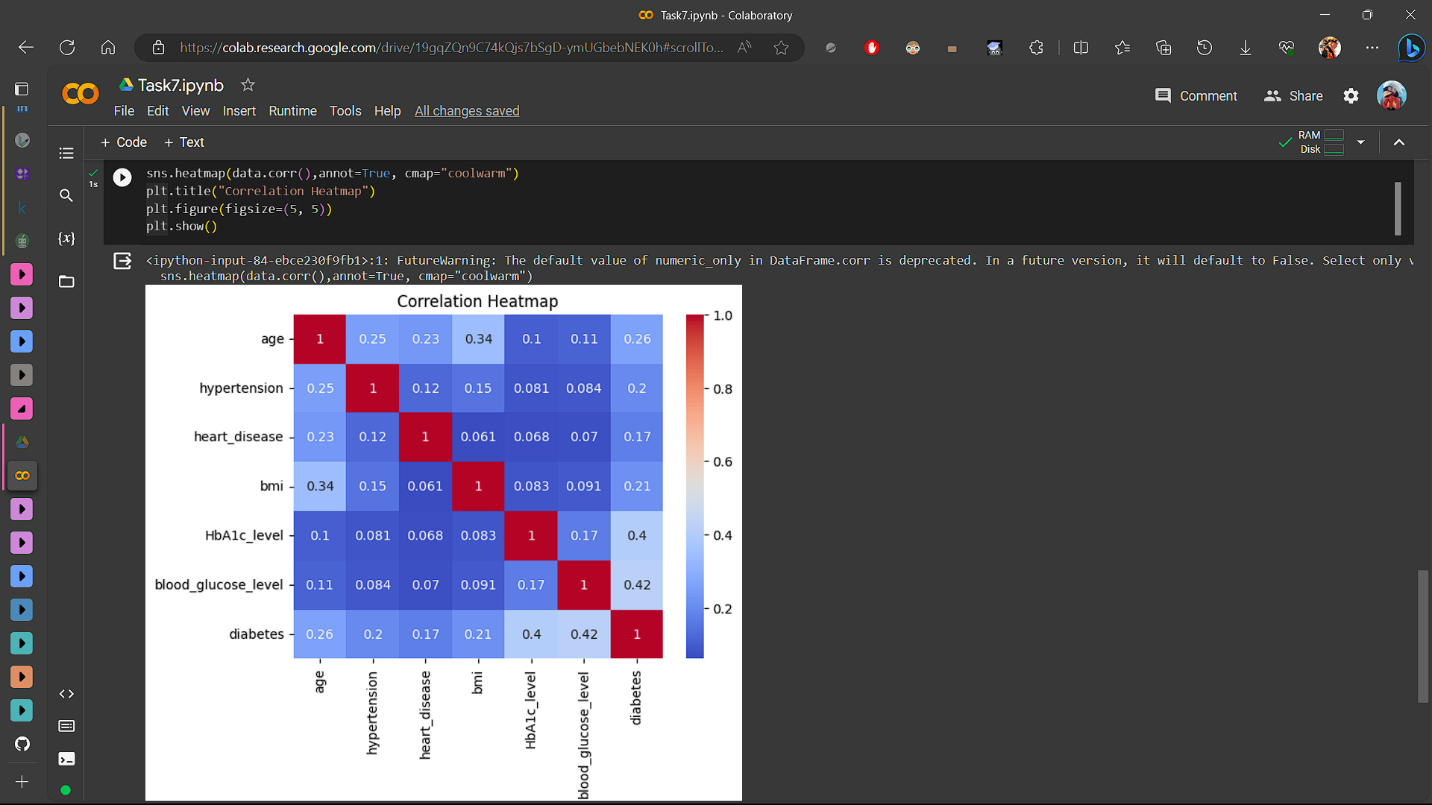
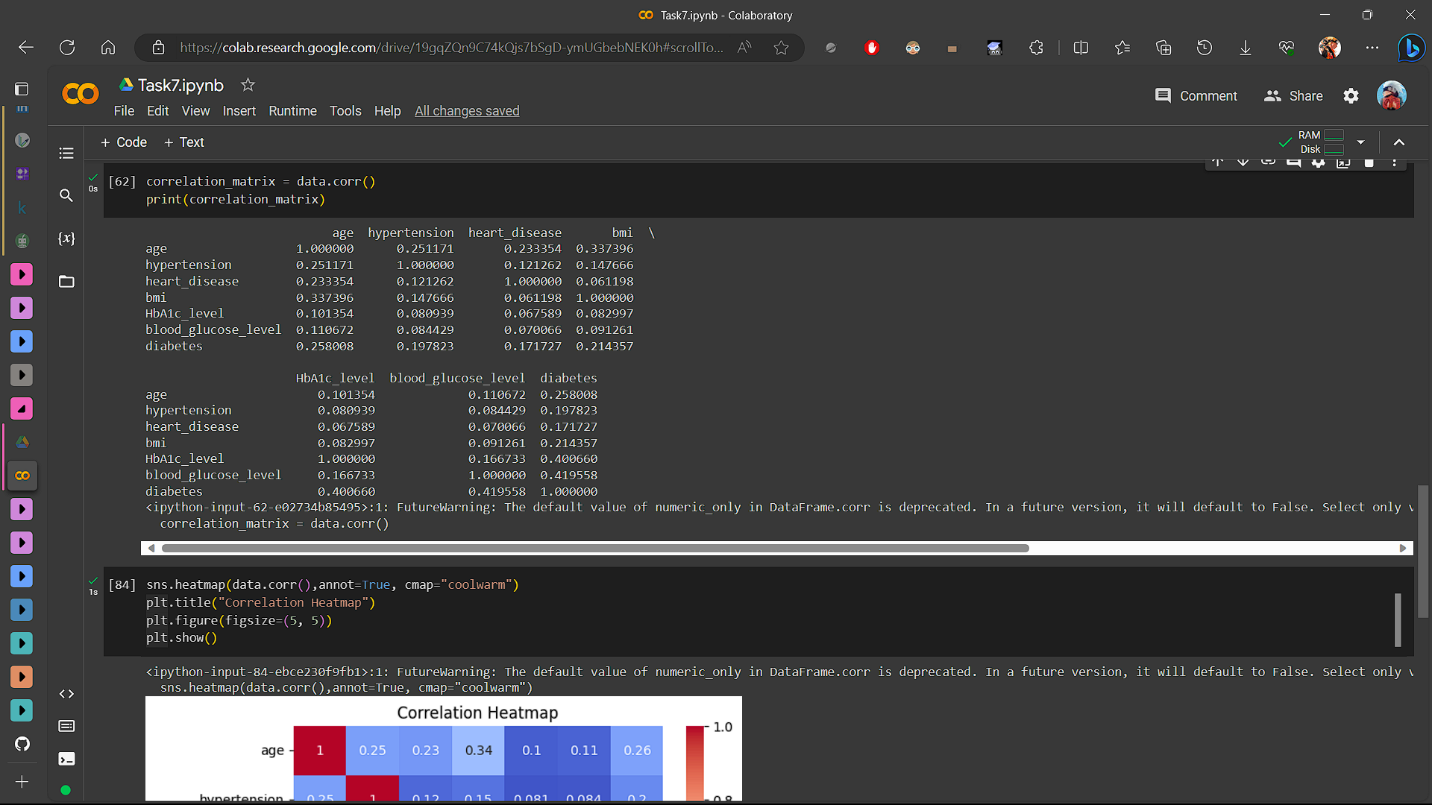
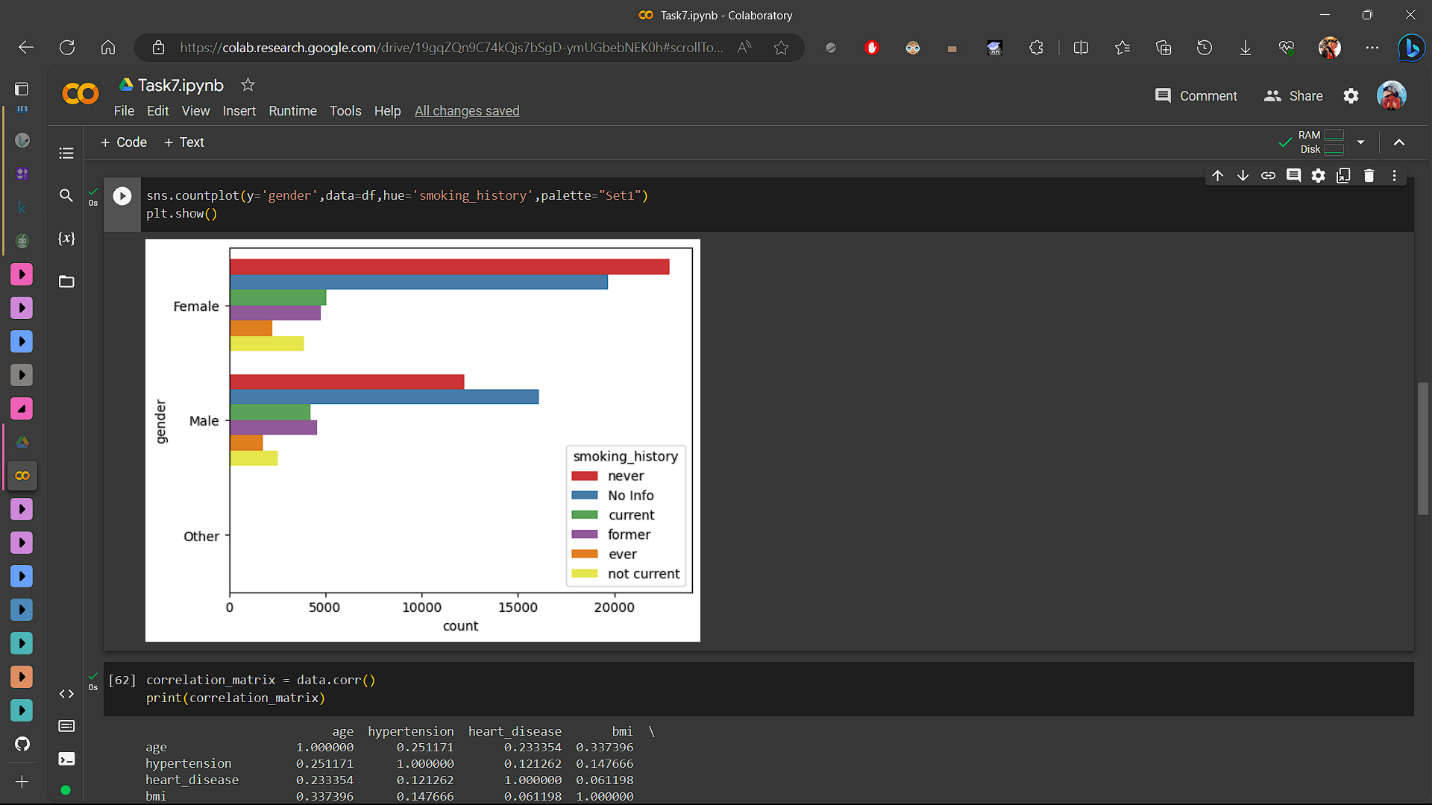
plt.title("Correlation Heatmap")

plt.figure(figsize=(5, 5))

plt.show()







**Data Analysis Report**

In this Analysis we have gone through the data and understand the nature of the data. And made a Count Plot and Heatmap. The following analysis is given below:

**Count Plot Analysis:**

I have made Two Count plots using seaborn as sns and matplotlib as plt which are based on gender which gives who have diabetes in the data using Hue function. And the other one is on gender and who have a smoking\_history in the data.

**Observation:**

* In this Count Plot we can observe that the number people which have a diabetes in genders in which more than 50000 females in males its less than 40000 and there are none in the others.
* We can see that the no of females tend to have diabetes than male’s
* In the next Count plot we can observe that the number people which have a somking\_history in genders in which female rank higher then male in which they never smoked and had no info on the smoking.
* We can also observe that the currently female’s are smoking more than male but not currently It may be because of there population is higher than males.
* We can also see that the female’s are more active smoker’s than male’s

**HEATMAP Analysis:**

I have a heatmap of the dataset by using seaborn as sns and matplotlib as plt which are based on the correlation of the data.

**Observation:**

* In the heatmap darker the color the stronger the correlation.
* Age is the strong correlation to the every other categories by 1.
* Moderate positive correlations exist between HbA1c levels, blood glucose levels, and diabetes presence.
* These correlations suggest that these variables are interrelated and can be valuable for diabetes prediction and diagnosis.
* Caution is needed as correlation does not prove causation.
* Further research and clinical analysis are essential to establish causative relationships and draw definitive conclusions about diabetes factors.