## Rajiv Gandhi University of knowledge and Technologies,

## IIIT-Srikakulam,Andhra Pradesh.

## A

## Project

## Report

## On

## MOBILE PRICE RANGE CLASSIFICATION

## 

## 

## Submitted to : Submitted By :

## Mr.CH.Sateesh sir T.Gopichand,

## ID:S180136,

## Class:CSE-2A.

## Department Of Computer Science Engineering

## RGUKT-IIIT,SKLM

## AP.

## MOBILE PRICE RANGE PREDICTION

## STATEMENT:

 Mobile phones are the best selling electronic devices as people keep updating their cell phones whenever they find new features in a new device. Thousands of mobiles are sold daily, in such a situation it is a very difficult task for someone who is planning to set up their own mobile phone business to decide what the price of the mobile should be.

**OBJECTIVE:**

To estimate the price of a mobile that can cover both marketing and manufacturing costs. So in this task, I don’t have to predict the actual prices of the mobiles but I have to predict the price range of the mobiles.

**APPROACH:**

I have collected the dataset from Kaggle.com . The dataset contains information (2000rows and 21columns) regarding mobile phone features, specifications etc and their price range. For Dataset and code link:

**The data features are**

1. **battery\_Power :** Total energy a battery can store in one time measured in mAh
2. **blue :** Has bluetooth or not
3. **clock\_speed :** speed at which microprocessor executes instructions
4. **dual \_sim :** Has dual sim support or not
5. **fc :** Front Camera megapixels
6. **four\_g :** Has 4G or not
7. **int\_memory :** Internal Memory in Gigabytes
8. **m\_dep :** Mobile Depth in cm
9. **mobile\_wt :** Weight of mobile phone
10. **n\_cores :** Number of cores of processor
11. **pc :** Primary Camera megapixels
12. **px\_height :** Pixel Resolution Height
13. **px\_width :** Pixel Resolution Width
14. **ram :** Random Access Memory in Megabytes
15. **sc\_h :** Screen Height of mobile in cm
16. **sc\_w :** Screen Width of mobile in cm
17. **talk\_time** : longest time that a single battery charge will last when you are
18. **three\_g :** Has 3G or not
19. **touch\_screen :** Has touch screen or not
20. **wifi :** Has wifi or not
21. **price\_range :** This is the target variable with values of 0(low cost), 1(medium cost), 2(high cost) and 3(very high cost).

**The price range of mobile phones as:**

1. 0 (Low Price)
2. 1 (Moderate Price)
3. 2 (Normal Price)
4. 4 (Expensive)

I have used **Ordinal** **Logistic Regression** Technique to predict the required output.

* **Ordinal logistic regression**: The response variable can belong to one of three or more categories and there *is* a natural ordering among the categories.

**Methodology**

After collecting dataset

Import the require libraries and applied various data preprocessing methods.

Data preprocessing is a technique in which we can remove missing values in the data. Because of these missing values, it is not possible to apply machine learning algorithms.

After removal of missing values, we need to apply label encoding, one hot encoding data to the categorical features.

Categorical features are the features whose values are labels instead of values.

After that, apply standardization or normalization techniques to our data. This method is used when all the attribute values are not in the same scale.

Import require Libraries

**import numpy as np**

**import pandas as pd**

**import matplotlib.pyplot as plt**

**import seaborn as sns**

**from sklearn.model\_selection import train\_test\_split**

**from sklearn.preprocessing import StandardScaler**

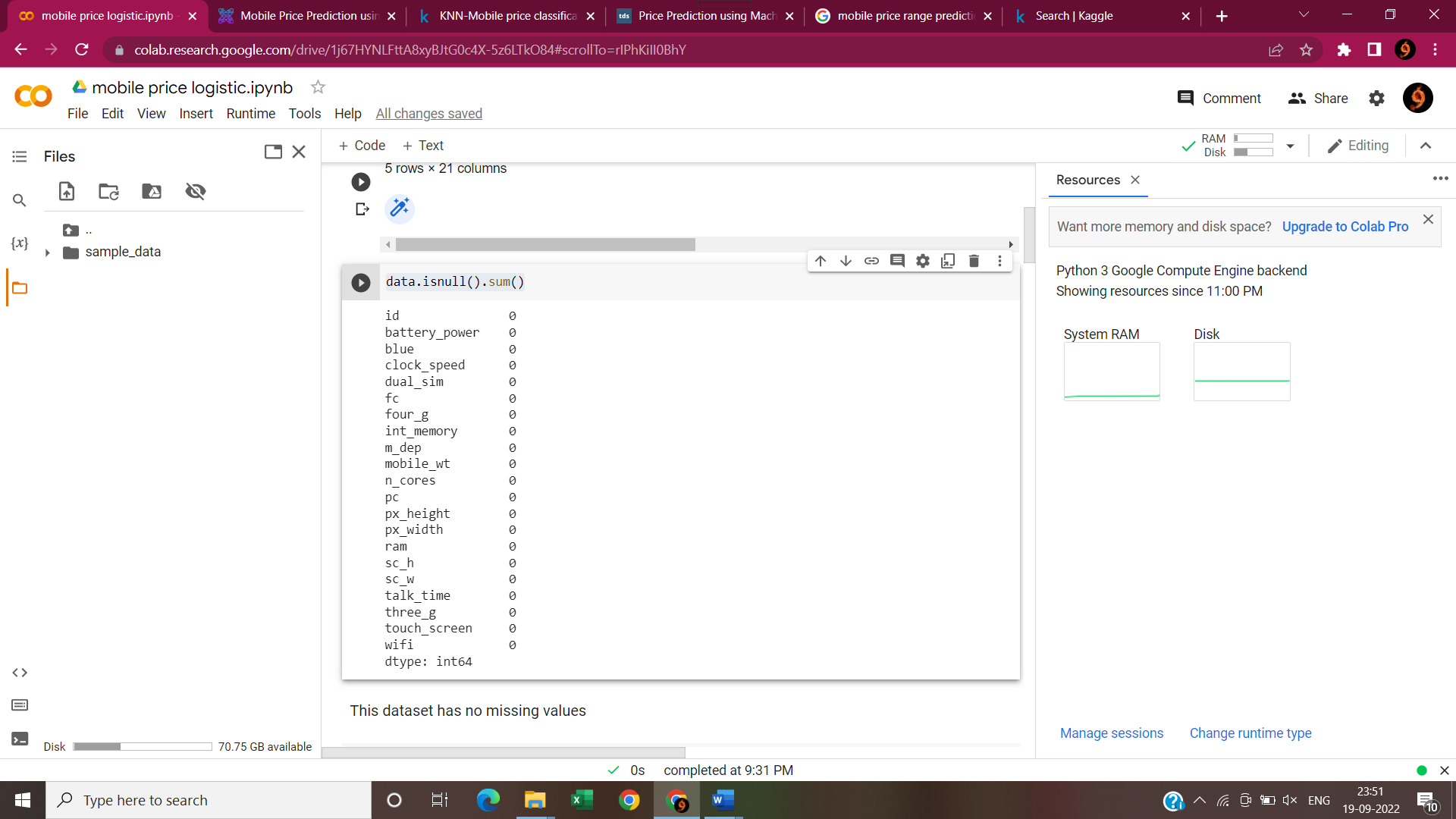
**from sklearn.linear\_model import LogisticRegression**

**from sklearn.metrics import accuracy\_score**

**sns.set()**

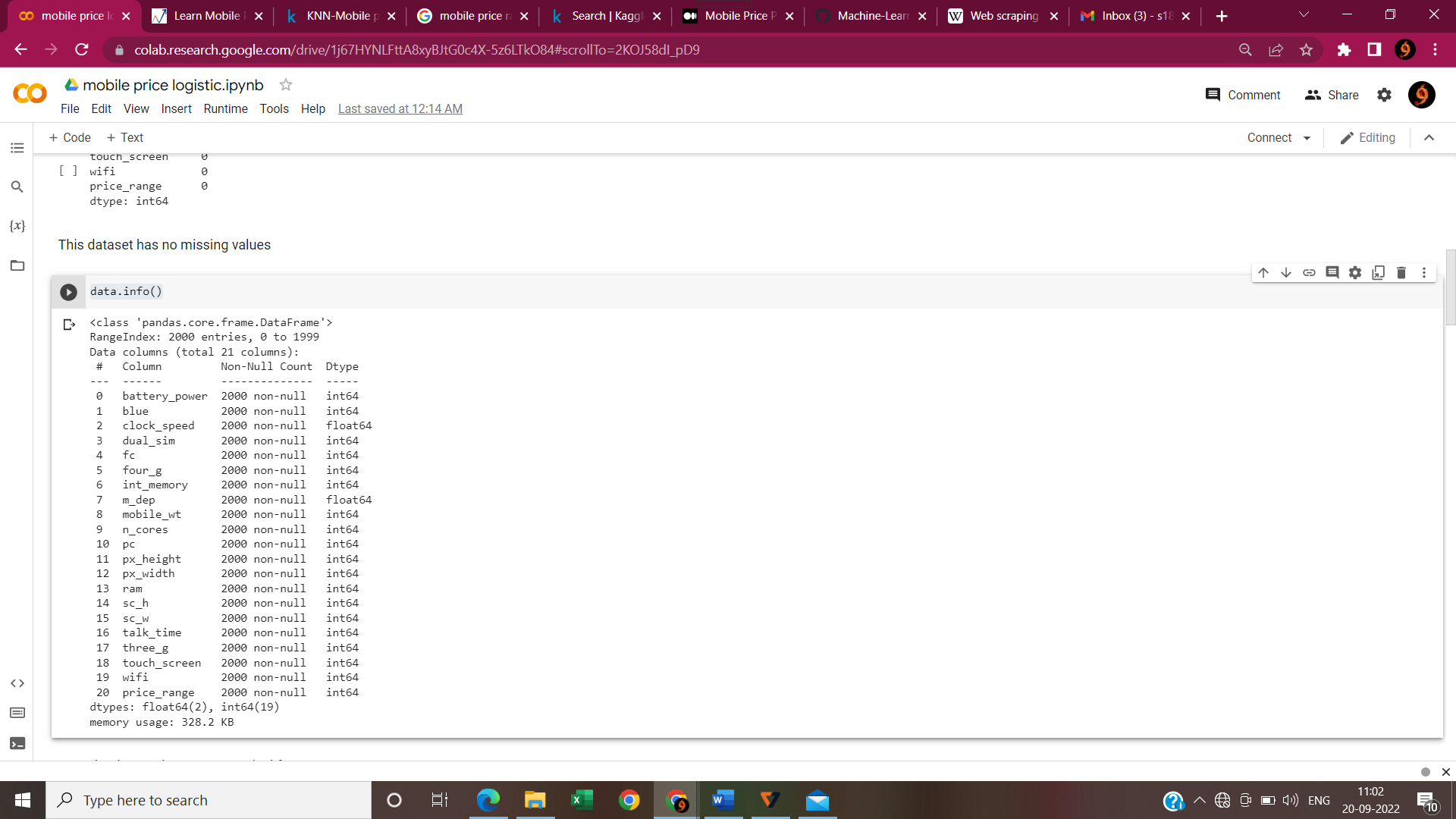
**Data preprocessing:**

Checking null values



Here,The dataset contains no NULL values .

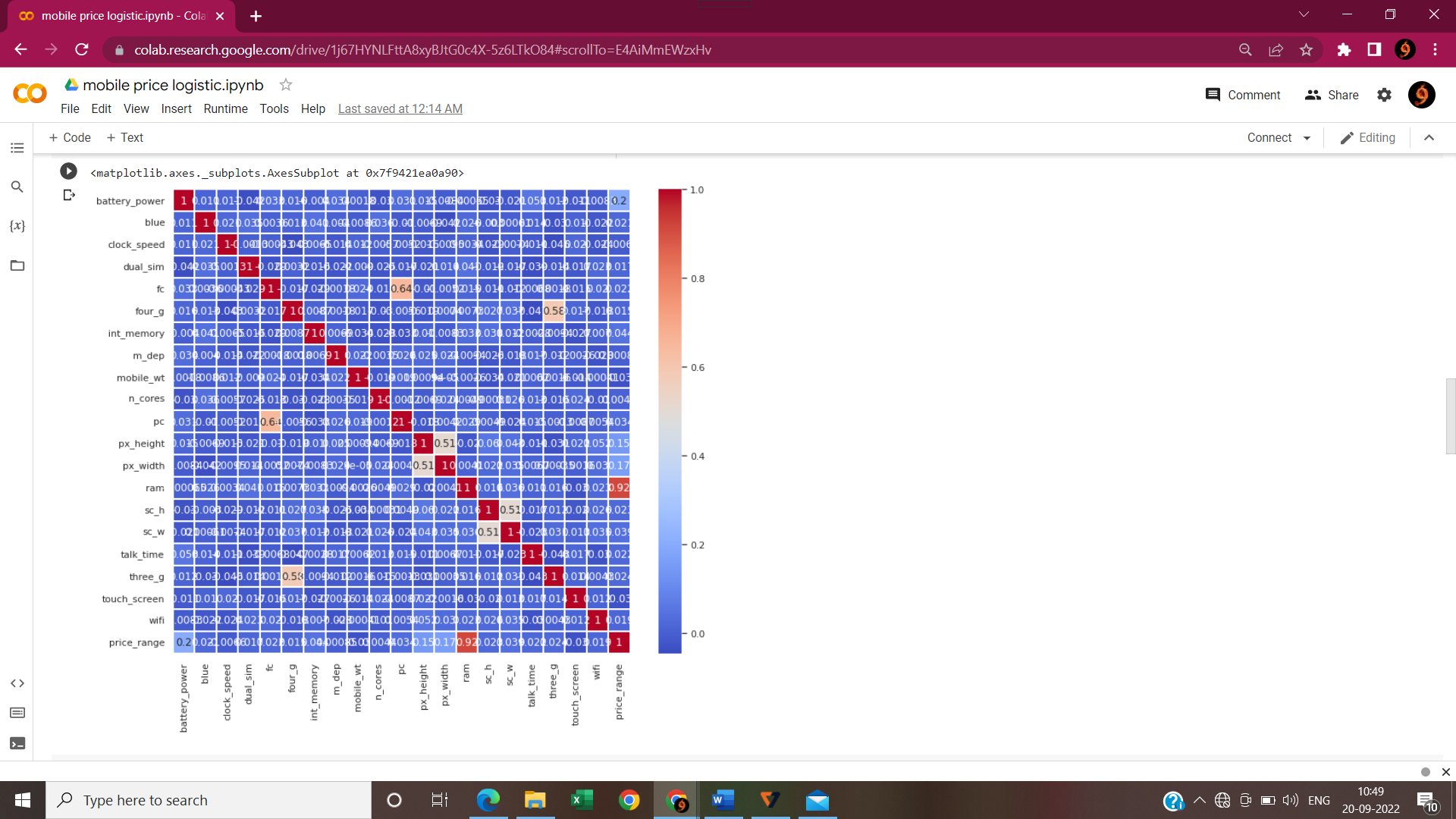
**Checking for categorical values:**



**NOW,using heat mapchecking Correlation between features in the dataset:**

**plt.figure(figsize=(12, 10))**

**sns.heatmap(data.corr(), annot=True, cmap="coolwarm", linecolor='white', linewidths=1)**

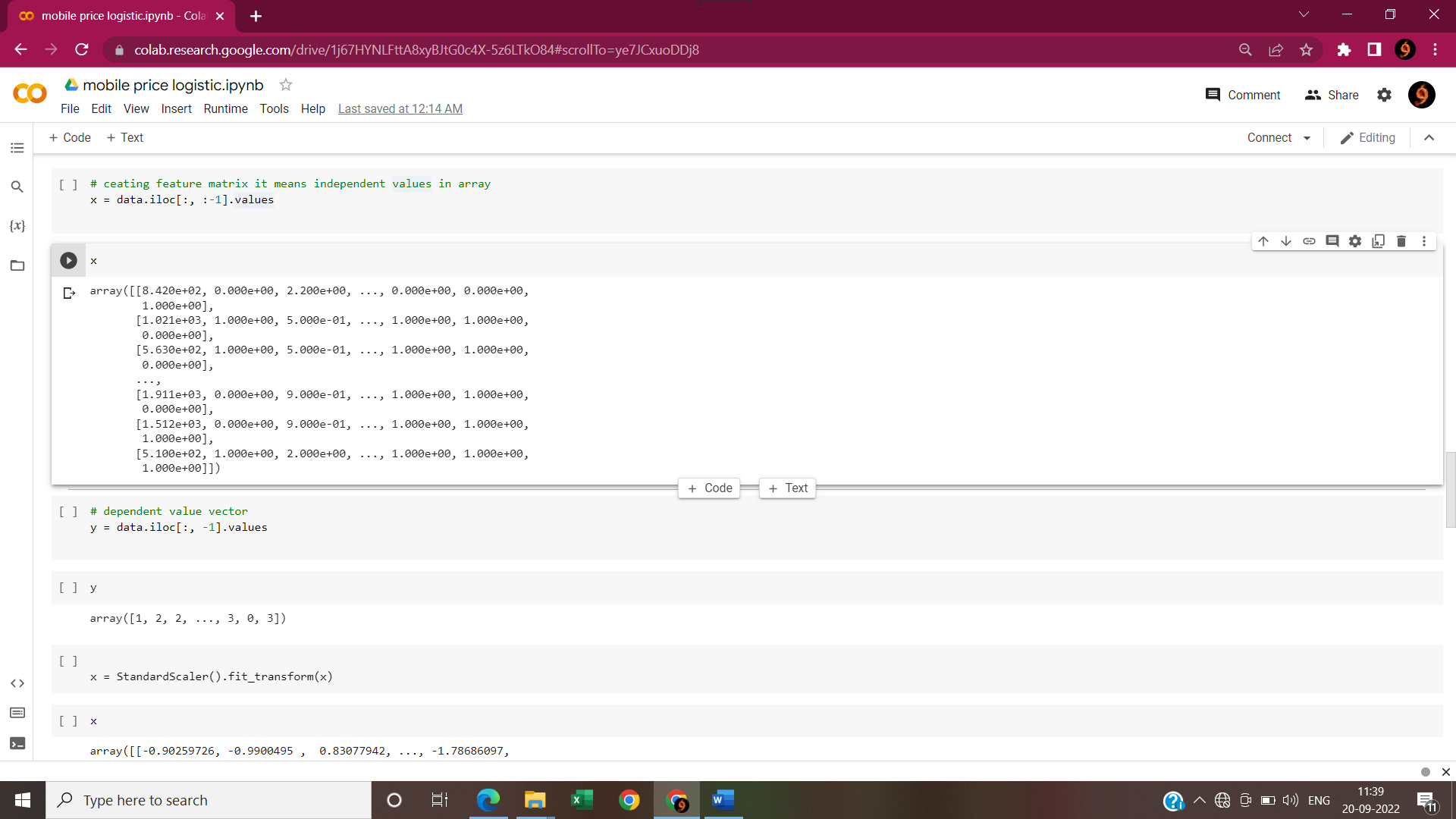
****

In this heatmap the co related features are

Ram is highly co-related 0.920.

**# ceating feature matrix it means independent values in array**

**x = data.iloc[:, :-1].values**

****

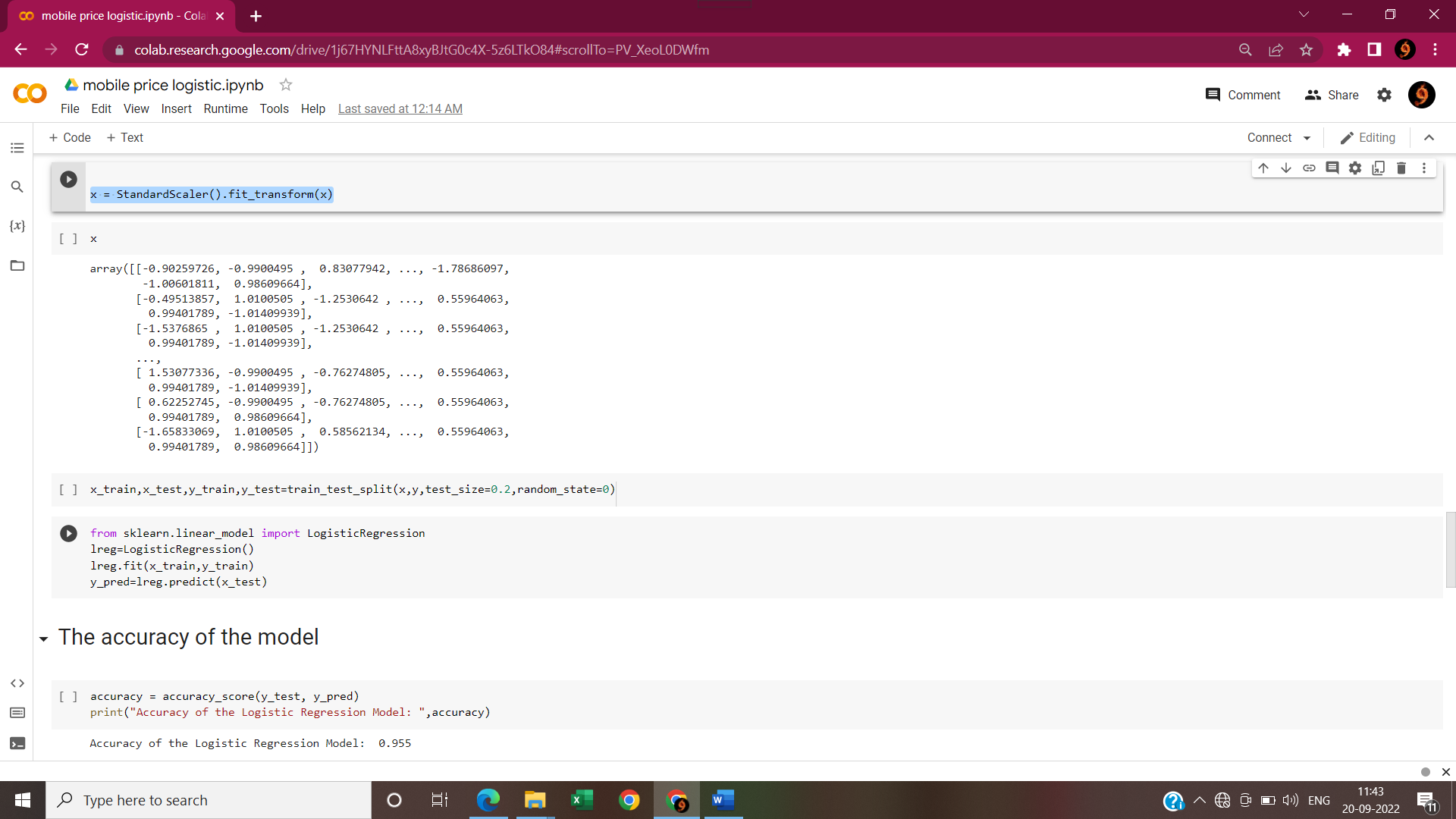
**# dependent value vector**

**y = data.iloc[:, -1].values**

**y**

array([1, 2, 2, ..., 3, 0, 3])

**x = StandardScaler().fit\_transform(x)**



**x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=0)**

**from sklearn.linear\_model import LogisticRegression**

**lreg=LogisticRegression()**

**lreg.fit(x\_train,y\_train)**

**y\_pred=lreg.predict(x\_test)**

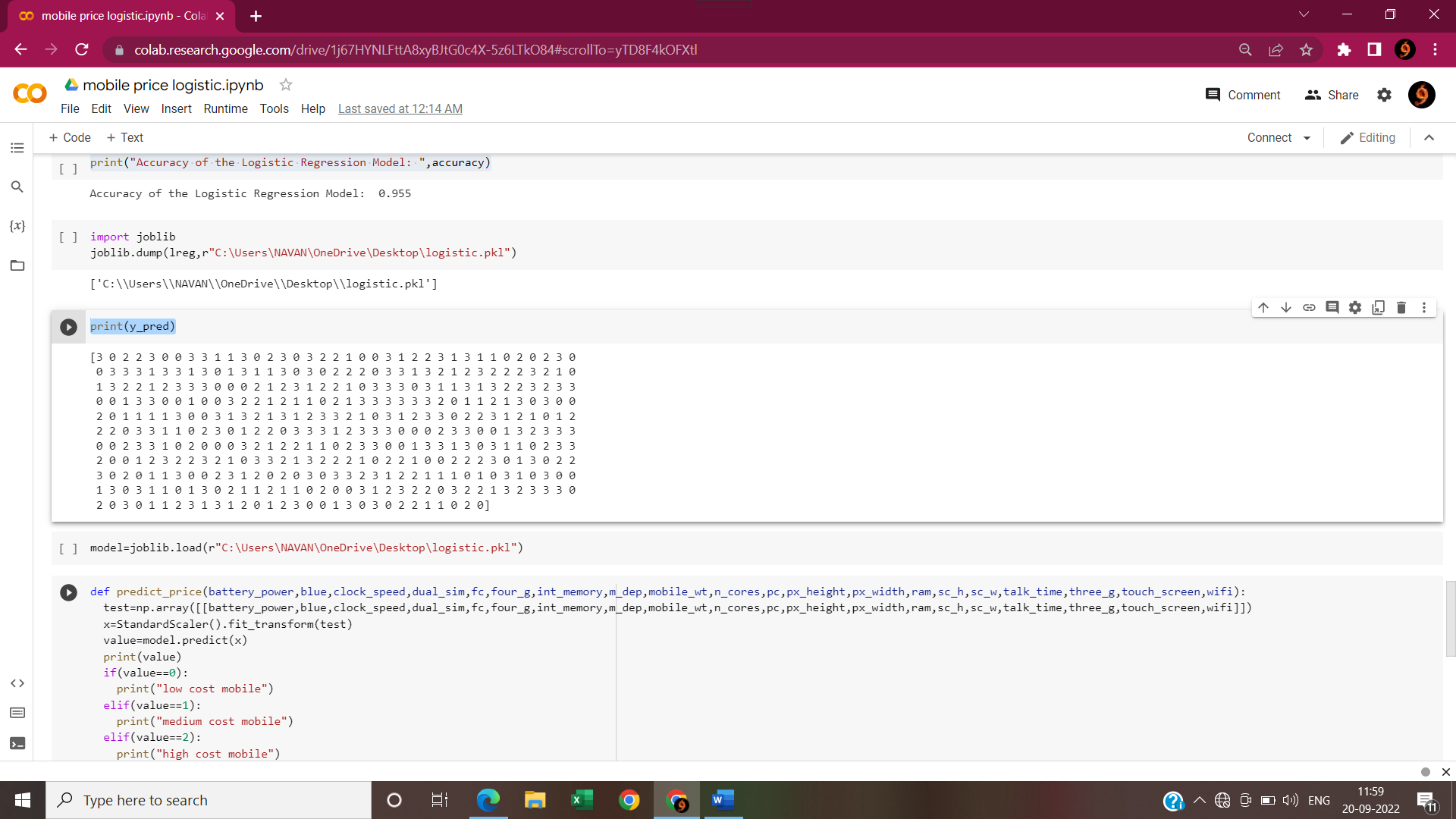
# **The accuracy of the model:**

**accuracy = accuracy\_score(y\_test, y\_pred)**

**print("Accuracy of the Logistic Regression Model: ",accuracy)**

A**ccuracy of the Logistic Regression Model: 0.955**

**print(y\_pred)**



**CONCLUSION:**

As cost is very Important to the marketing analysis, the model Will give people a sharp knowledge about various products price.

From this strategy a people can find their required products with minimum cost and maximum specifications.

It determines the price range so a good product can be suggested to a customer.

It will be so helpful to the businessman ,Who is going to build a mobile .

By using this machine learning project the business man will be able to know the budget or capital to build the mobile phones with features he wants to build.