**Lab Manual 5 – From Keggle**

A logo of a computer science

Description automatically generated

Session: 2022 – 2026

**Submitted by:**

Abdullah Zahid

2022-CS-03

**Submitted To:**

Sir Nauman Shafi

Department of Computer Science

**University of Engineering and Technology**

**Lahore Pakistan**

# Dataset Link:

[Laptop Price](https://www.kaggle.com/datasets/muhammetvarl/laptop-price/)

<https://www.kaggle.com/datasets/muhammetvarl/laptop-price/>

# Queries:

1. To Create Data frame-extract data from CSV.
2. I have Converted the Price Column into Array and Calculated the Average, Minimum and Maximum Price.
3. Filter the Laptops which have Ram more than 8GB and Price below the average.
4. Group the Dataset by Laptop Company and calculate the Average Price of Each Company.
5. Plot a line graph showing the price variation over the companies
6. Scatter Plot shows how prices change with RAM sizes

# Code:

# -\*- coding: utf-8 -\*-

"""

Created on Fri Oct 4 14:13:08 2024

@author: zabdu

"""

import datetime

import pandas as pd

import numpy as np

import re

import matplotlib.pyplot as plt

def createDataFrame():

df=pd.read\_csv(f'laptop\_price.csv',encoding='latin1')

return df

def priceAMM(df):

price=df['Price\_euros'].to\_numpy()

print(f"Average Price: {np.mean(price):.2f} euros")

print(f"Minimum Price: {np.min(price):.2f} euros")

print(f"Max Price: {np.max(price):.2f} euros")

def laptopFilter(df):

count = 0

avg\_salary = np.mean(df['Price\_euros'].to\_numpy())

dt = pd.DataFrame(columns=['laptop\_ID', 'Company', 'Product', 'TypeName', 'Inches', 'ScreenResolution', 'Cpu', 'Ram', 'Memory', 'Gpu', 'OpSys', 'Weight', 'Price\_euros'])

for i in range(len(df)):

ram = int(re.findall(r"\d+", df.loc[i, 'Ram'])[0])

if ram > 8 and df.loc[i, 'Price\_euros'] < avg\_salary:

dt.loc[count] = df.loc[i]

count += 1

return dt

def groupedData(df):

grouped=df.groupby('Company')['Price\_euros'].mean()

for index in grouped.index:

print(f"{index} : {grouped[index]} euros")

return grouped

def lineGraph(df):

grouped=df.groupby('Company')['Price\_euros'].mean()

companies=grouped.index

prices=grouped.values

plt.figure(figsize=(12, 6))

plt.plot(companies,prices, marker='o')

plt.xticks(rotation=45)

plt.xlabel('Brand Names')

plt.ylabel('Price (euros)')

plt.title('Price Variation of Laptops Across Companies')

plt.grid()

plt.show()

def ramGraph(dt):

dt['Ram\_numbers'] = dt['Ram'].str.extract('(\d+)').astype(int)

plt.figure(figsize=(10, 5))

plt.scatter(dt['Ram\_numbers'], dt['Price\_euros'], marker='o')

plt.xlabel('RAM (GB)')

plt.ylabel('Price (euros)')

plt.title('Price Variation of Laptops by RAM Size')

plt.xticks(dt['Ram\_numbers'])

plt.show()

df=pd.DataFrame()

#Task 1

df=createDataFrame()

#Task 2

# I have Converted the Price Column into Array and Calcuated the Average, Minimum and Maximum Price

#priceAMM(df)

#Task 3

#Filter the Laptops which have Ram more than 8GB and Price below the average.

#laptopFilter(df)

#Task 4

#Group the Dataset by Laptop Compoany and calculate the Average Price of Each Company

#groupedData(df)

#Task 5

# Plot a line graph showing the price variation over the companies

#lineGraph(df)

#Task 6

#Scatter Plot shows how prices change with RAM sizes

#ramGraph(df)