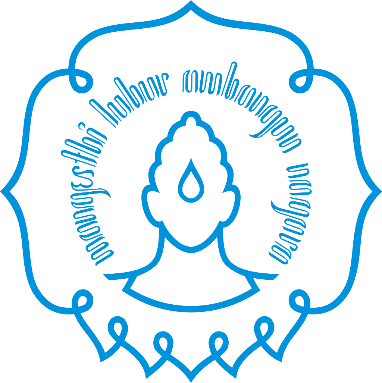
UJIAN AKHIR SEMESTER

PRAKTIK PEMROGRAMMAN

PYHTON



## Disusun oleh :

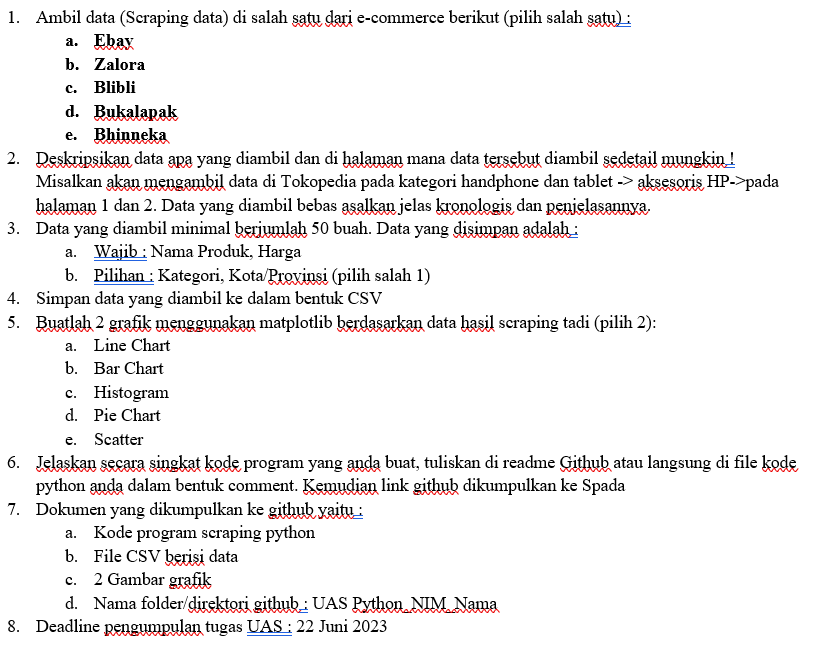
Wahyu Bagas Dwi Prasetyo

V3922045

## Dosen

Yusuf Fadila Rachman, S. Kom., M. Kom

# PS D-III TEKNIK INFORMATIKA SEKOLAH VOKASI UNIVERSITAS SEBELAS MARET 2023

 SOAL

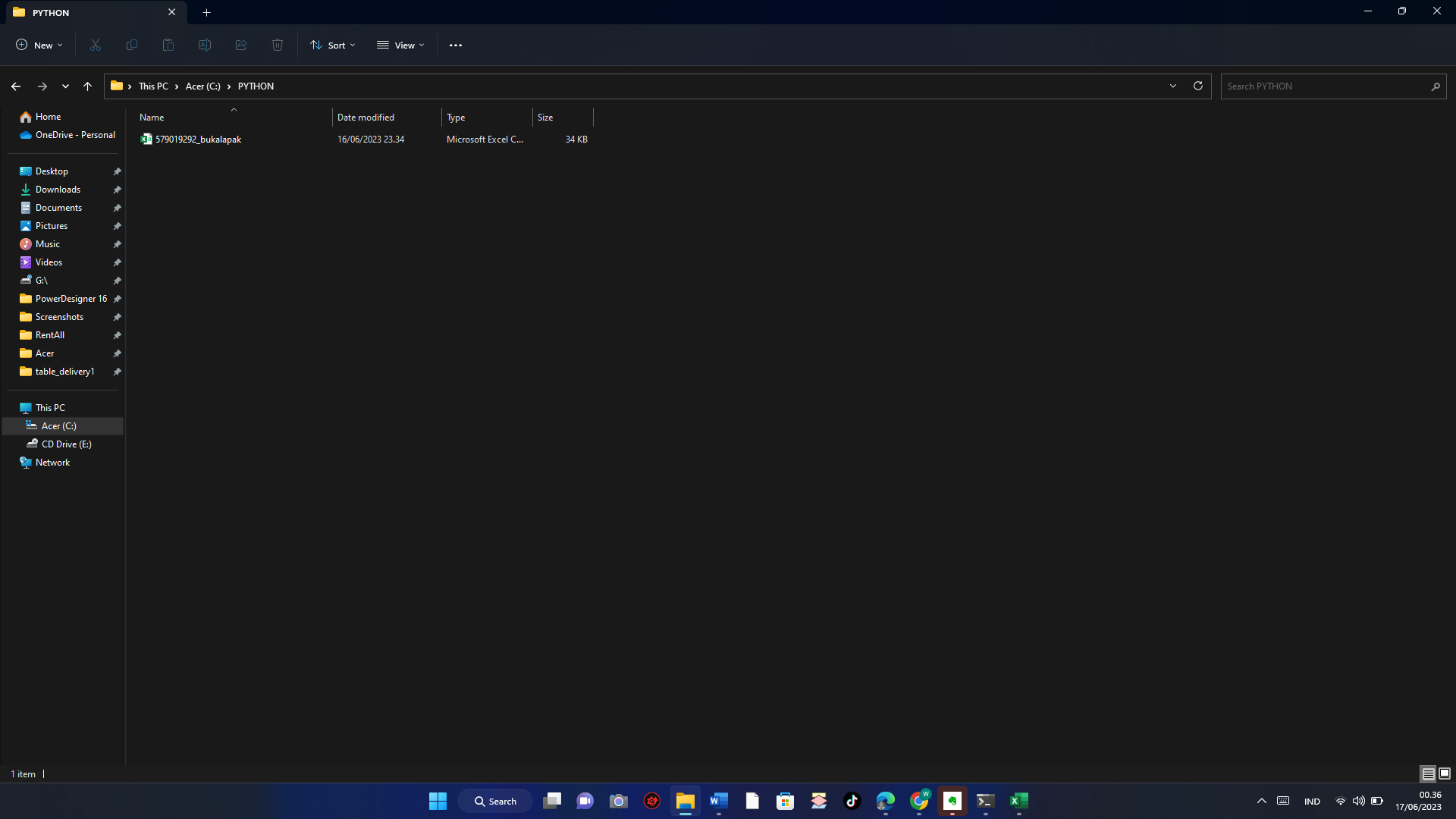
Jawab:

1.Disini saya mengambil Scrapping Data di Website Blibli

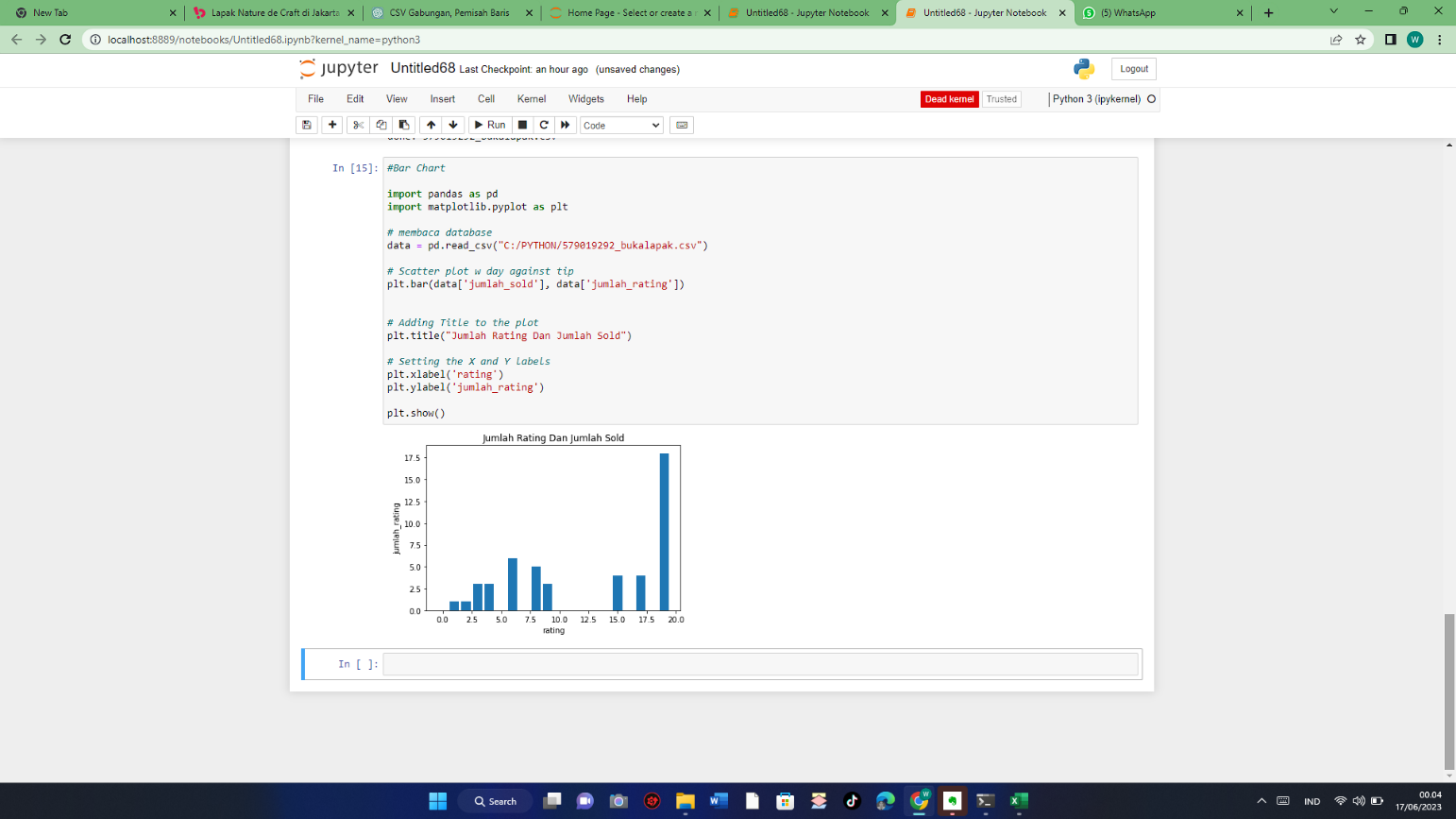
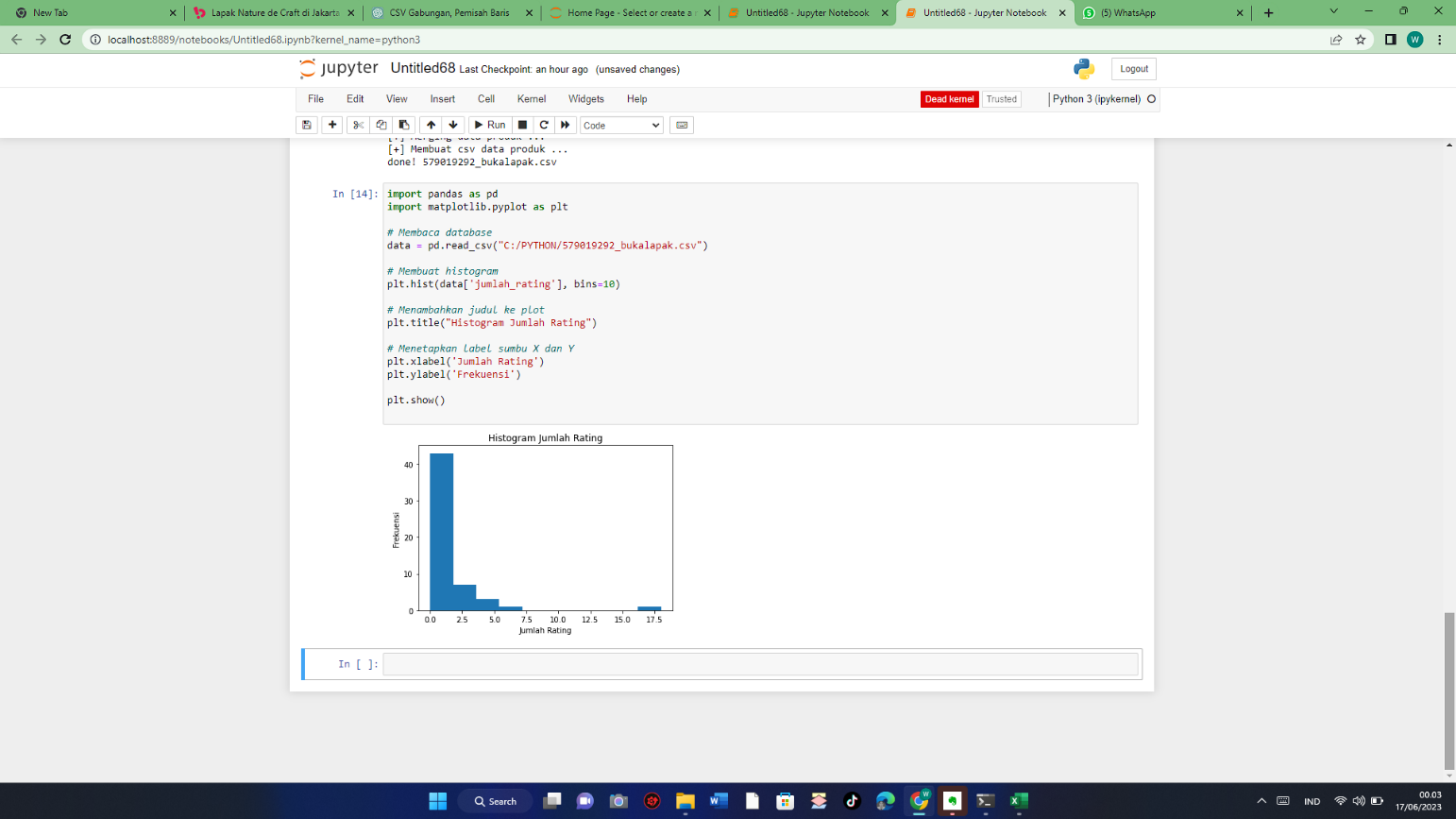
2.Saya mengambil data diblibli pada toko:https://www.bukalapak.com/u/nature\_de\_craft\_46327 dan untuk halamannya saya mengambil di halaman 1-4

3.Data yang saya ambil berjumlah 56 buah untuk data waib seperti: Nama Produk,Harga sudah tersimpan. Lalu Untuk data Pilihan saya Memilih Kategori

4.Disini sudah saya simpan dan saya buat file baru (Untuk file csv nanti saya sertakan di github)



5.Disini saya mengambil barchart dan histogram untuk memvisualisasikan/menampilkan data yang sudah saya simpan sebagai contoh disini saya ingin menampilkan rating dan jumlah sold nya



6.Penjelasan Kode Disini saya cantumkan comment dan file readme untuk readme saya upload digithub:

**A.Berikut kode dengan menggunakan penjelasan comment pada program untuk menyimpan file csv**

#!/usr/bin/env python

# coding: utf-8

# In[ ]:

#!/usr/bin/env python

# coding: utf-8

# Import library yang diperlukan

from bs4 import BeautifulSoup as bs

import requests as req

import glob

import json

import os

import csv

import time

# URL dasar

baseurl = 'https://www.bukalapak.com/u/'

apiurl = 'https://api.bukalapak.com/stores/'

# Definisikan kelas Bukalapak

class Bukalapak:

def \_\_init\_\_(self, username):

self.username = username

self.urltoko = baseurl + username

self.headerbrowser = {

'User-Agent': 'Mozilla/5.0 (X11; Linux x86\_64; rv:74.0) Gecko/20100101 Firefox/74.0'

}

pg\_data = req.get(self.urltoko, headers=self.headerbrowser, timeout=3000)

try:

if pg\_data.status\_code == 200:

soup = bs(pg\_data.text, 'html.parser')

# Mengambil id penjual

print("[+] Mengambil id penjual ...")

self.idseller = soup.find('a', attrs={'data-user-id': True}).get('data-user-id')

print("[+] Id penjual: " + self.idseller)

# Mengambil access token untuk API

print("[+] Mengambil access token ...")

stoken = soup.find\_all('script')

gtoken = str(stoken[4]).replace("<script>localStorage.setItem('bl\_token', '", "").replace("');</script>", "")

gtoken = json.loads(str(gtoken))

self.token = gtoken['access\_token']

print("[+] Access token: " + self.token)

self.grab\_produk()

else:

print("[!] Username tidak ditemukan!")

except Exception as e:

print(e)

print("[!] Website down!")

exit()

def grab\_produk(self):

print("=== GRABBING PRODUK ===")

print("[+] Memulai download halaman produk ...")

print("\x1B[3m" + "(delay 3 detik untuk menghindari anti-spam!)" + "\x1B[0m")

api = apiurl + str(self.idseller) + '/products?offset=0&limit=50&access\_token=' + self.token

total\_product = req.get(api, headers=self.headerbrowser, timeout=3000).json()

print("[+] Total produk: " + str(total\_product['meta']['total']))

print("[+] Hapus file lama ...")

if not os.path.exists("data"):

os.makedirs("data")

for filename in glob.glob("data/" + str(self.idseller) + "bukalapak\*.json"):

os.remove(filename)

for filename in glob.glob(str(self.idseller) + "\_bukalapak.csv"):

os.remove(filename)

# Mengambil halaman produk

a = 0

b = 0

while True:

print("-> download halaman ke-" + str(b + 1))

api = apiurl + str(self.idseller) + '/products?offset=' + str(a) + '&limit=50&access\_token=' + self.token

pg\_data = req.get(api, headers=self.headerbrowser, timeout=3000).json()

if len(pg\_data['data']) == 0:

break

with open("data/" + str(self.idseller) + "bukalapak" + str(b) + '.json', 'w') as json\_file:

json.dump(pg\_data['data'], json\_file)

a += 50

b += 1

time.sleep(3)

# Menggabungkan JSON

print("[+] Menggabungkan data produk ...")

data = []

for f in glob.glob("data/" + str(self.idseller) + "bukalapak\*.json"):

with open(f) as infile:

data.extend(json.load(infile))

with open("data/" + str(self.idseller) + "\_bukalapak\_all.json", 'w') as outfile:

json.dump(data, outfile)

# Membuat CSV

print("[+] Membuat csv data produk ...")

f\_data = []

with open("data/" + str(self.idseller) + "\_bukalapak\_all.json") as f:

f\_read = json.load(f)

for i in f\_read:

f\_data.append([

self.idseller,

self.username,

i['store']['name'],

i['store']['level']['name'],

i['store']['premium\_level'],

i['id'],

i['name'],

i['category']['name'],

i['condition'],

i['rating']['average\_rate'],

i['rating']['user\_count'],

i['stock'],

i['stats']['interest\_count'],

i['stats']['sold\_count'],

i['stats']['view\_count'],

i['price'],

i['original\_price'],

i['discount\_percentage'],

i['description']

])

f\_header = ['id\_seller', 'username', 'nama\_toko', 'level\_toko', 'premium\_toko', 'id\_produk', 'nama\_produk',

'kategori', 'kondisi', 'rating', 'jumlah\_rating', 'stok', 'jumlah\_interest', 'jumlah\_sold',

'jumlah\_view', 'harga', 'harga\_asli', 'diskon', 'deskripsi']

with open(str(self.idseller) + '\_bukalapak.csv', 'w', newline='', encoding='utf-8') as file:

writer = csv.writer(file)

writer.writerow(f\_header)

writer.writerows(f\_data)

print('done! ' + str(self.idseller) + '\_bukalapak.csv')

f.close()

# Main program

print("[+] https://github.com/heryandp/bukalapak-product-scrap")

sname = input("[+] Masukkan username seller: https://www.bukalapak.com/u/")

act = Bukalapak(sname)

# In[ ]:

**B.Berikut adalah kode dengan penjelasan menggunakan comment untuk memvisualisasikan dalam bentuk barchart**

#!/usr/bin/env python

# coding: utf-8

# In[28]:

#!/usr/bin/env python

# coding: utf-8

# Import library yang diperlukan

import pandas as pd

import matplotlib.pyplot as plt

# Membaca database

data = pd.read\_csv("C:/PYTHON/579019292\_bukalapak.csv")

# Membuat bar chart dengan jumlah\_sold sebagai sumbu X dan jumlah\_rating sebagai sumbu Y

plt.bar(data['jumlah\_sold'], data['jumlah\_rating'])

# Menambahkan judul ke plot

plt.title("Jumlah Rating dan Jumlah Sold")

# Menetapkan label sumbu X dan Y

plt.xlabel('Jumlah Sold')

plt.ylabel('Jumlah Rating')

# Menampilkan plot

plt.show()

# In[ ]:

**C.Berikut adalah kode dengan penjelasan menggunakan comment untuk memvisualisasikan dalam bentuk Histogram**

#!/usr/bin/env python

# coding: utf-8

# In[1]:

#!/usr/bin/env python

# coding: utf-8

# Import library yang diperlukan

import pandas as pd

import matplotlib.pyplot as plt

# Membaca database

data = pd.read\_csv("C:/PYTHON/579019292\_bukalapak.csv")

# Membuat histogram dengan menggunakan data jumlah\_rating dan membaginya ke dalam 10 bins

plt.hist(data['jumlah\_rating'], bins=10)

# Menambahkan judul ke plot

plt.title("Histogram Jumlah Rating")

# Menetapkan label sumbu X dan Y

plt.xlabel('Jumlah Rating')

plt.ylabel('Frekuensi')

# Menampilkan plot

plt.show()

# In[ ]: