**وارد کردن کتابخانه‌ها:**

|  |
| --- |
| import pandas as pd  import matplotlib  matplotlib.use('TkAgg')  import matplotlib.pyplot as plt  import seaborn as sns |

pandas یک کتابخانه برای مدیریت و تحلیل داده است.  
Matplotlib یک کتابخانه برای رسم نمودارهای دوبعدی جهت ایجاد تصاویر است.  
matplotlib.use('TkAgg') بک‌اند Matplotlib را به 'TkAgg' تنظیم می‌کند.  
Seaborn یک کتابخانه تصویرسازی داده‌های آماری بر پایه Matplotlib است که رابط سطح بالا برای رسم گرافیک‌های آماری جذاب و اطلاعاتی ارائه می‌دهد.

**تعریف یک تابع برای خواندن فایل CSV:**

|  |
| --- |
| def read\_file(path, encoding="utf-16"):  df = pd.read\_csv(path, encoding=encoding)  return df |

این تابع (read\_file) یک فایل CSV از مسیر مشخص شده با استفاده از pandas می‌خواند و یک DataFrame را برمی‌گرداند.

**خواندن فایل‌های داده:**

|  |
| --- |
| questions = read\_file("Data/Questions.txt")  answers = read\_file("Data/Answer.csv")  answers = answers.drop(0)  q\_a = read\_file("Data/Q\_A.csv") |

فایل‌های "Questions.txt"، "Answer.csv" و "Q\_A.csv" با استفاده از تابع read\_file به کمک pandas به DataFrame تبدیل می‌شوند.

**ادغام DataFrame‌ها:**

|  |
| --- |
| merged\_Q\_A = pd.merge(questions, q\_a, left\_on='Id', right\_on='QId', how='inner') |

DataFrame‌های questions و q\_a بر اساس ستون‌های 'Id' و 'QId' به ترتیب ادغام می‌شوند.

**ایجاد یک نشانگر دودویی برای پاسخ‌های پذیرفته شده:**

|  |
| --- |
| accepted\_flag = merged\_Q\_A.groupby('Id')['accepted'].any().astype(int)  questions['accepted\_flag'] = accepted\_flag |

یک ستون جدید به نام 'accepted\_flag' در DataFrame questions ایجاد می‌شود که نشان‌دهنده پذیرفته شدن (1) یا عدم پذیرفته شدن (0) است.

**محاسبه همبستگی:**

|  |
| --- |
| correlation = questions['accepted\_flag'].corr(questions['ViewCount']) |

همبستگی بین ستون‌های 'accepted\_flag' و 'ViewCount' محاسبه می‌شود.

**محاسبه میانه تعداد بازدید:**

|  |
| --- |
| median\_view\_count = questions.groupby('accepted\_flag')['ViewCount'].median() |

میانه تعداد بازدید برای پاسخ‌های پذیرفته شده و عدم پذیرفته شدن محاسبه می‌شود.

**ایجاد یک نمودار میله‌ای:**

|  |
| --- |
| sns.barplot(x=median\_view\_count.index.map({1: 'Accepted', 0: 'Not Accepted'}), y=median\_view\_count.values, color='blue') |

یک نمودار میله‌ای با استفاده از Seaborn ایجاد می‌شود، با برچسب‌های محور x نگاشت شده به 'پذیرفته شده' و 'عدم پذیرفته شدن' و مقادیر محور y به میانه تعداد بازدید محاسبه شده.

**سفارشی‌سازی نمودار و ذخیره به عنوان تصویر:**

|  |
| --- |
| plt.ylabel('Median View Count')  plt.title('Median Of View Count for Accepted vs Not Accepted Answers')  plt.xticks(rotation=90)  plt.yticks(rotation=0)  plt.savefig('Result question 1.png', dpi=1200, bbox\_inches='tight') |

سفارشی‌سازی‌های مختلف به نمودار اعمال می‌شود و نمودار نتیجه به عنوان یک تصویر با نام 'Result question 1.png' با رزولوشن بالا (dpi=1200) و باکس محدوده فشرده ذخیره می‌شود.