16. Scenario: You are working on a project that involves analyzing customer reviews for a product.

You have a dataset containing customer reviews, and your task is to develop a Python program that

calculates the frequency distribution of words in the reviews.

Question: Develop a Python program to calculate the frequency distribution of words in the

customer reviews dataset?

**Code:**

import pandas as pd

import re

from collections import Counter

df = pd.read\_csv(r"C:\Users\vara prasad\Downloads\reviews.csv")

text = ' '.join(df['review\_text'].astype(str))

cleaned\_text = re.sub(r'[^a-z\s]', '', text.lower())

words = cleaned\_text.split()

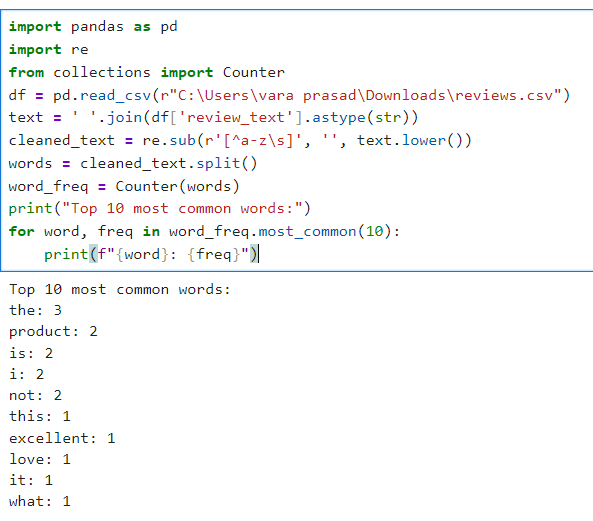
word\_freq = Counter(words)

print("Top 10 most common words:")

for word, freq in word\_freq.most\_common(10):

print(f"{word}: {freq}")

**output:**

****

**Dataset:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| review\_id | review\_text |  |  |  |
| 1 | This product is excellent! I love it. | | |  |
| 2 | Not what I expected. The quality could be better. | | | |
| 3 | Amazing value for the price. Highly recommend! | | | |
| 4 | The product is okay, but shipping was delayed. | | | |
| 5 | Terrible experience. Will not buy again! | | |  |