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

from collections import Counter

import re

df = pd.read\_csv(r"C:\Users\jampa\OneDrive\文档\customer\_reviews.csv")

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

tokens = re.findall(r'\b\w+\b', text.lower())

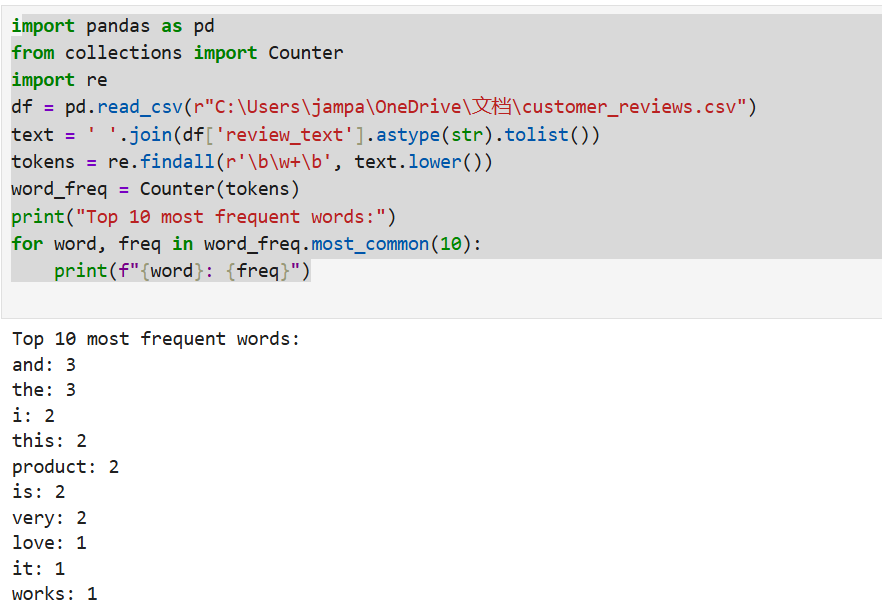
word\_freq = Counter(tokens)

print("Top 10 most frequent words:")

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

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

output:



Dataset:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| review\_id | review\_text | |  |  |  |  |  |
| 1 | I love this product! It works great and is very affordable. | | | | | |  |
| 2 | Terrible experience. The product broke after one use. Very disappointed. | | | | | | |
| 3 | Excellent quality and amazing customer service. Highly recommend! | | | | | | |
| 4 | Not worth the money. Poor build and unreliable. | | | | |  |  |
| 5 | This is the best purchase I've made. Extremely satisfied! | | | | | |  |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |