13. Scenario: You are working on a text analysis project and need to determine the frequency

distribution of words in a given text document. You have a text document named "sample\_text.txt"

containing a paragraph of text. Your task is to develop a Python program that reads the text

document, processes the text, and generates a frequency distribution of the words.

Question: How would you develop a Python program to calculate the frequency distribution of

words in a text document?

Code:

#13

import string

from collections import Counter

with open(r"C:\Users\hares\Downloads\data\_text.txt") as file:

text = file.read()

text = text.lower()

text = text.translate(str.maketrans('', '', string.punctuation))

words = text.split()

word\_freq = Counter(words)

print("Top 10 Most Common Words:\n")

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

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

output:

Top 10 Most Common Words:

data: 4

and: 4

science: 3

is: 2

field: 2

that: 2

to: 2

an: 2

of: 2

a: 1

dataset:

Data science is a field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data. It is an interdisciplinary field that combines statistics, computer science, and domain expertise to solve real-world problems. With the increasing availability of data, data science has become an essential part of business decision-making.