**Problem Statement :**

**Using the Data Sheets as below, predict the chickpea yield for a State in India named “XYZ” given the   
weather data for the years 2021-22, 2022-23, 2023-24.**

1. Worksheet named Yield: This has the annual chickpea yield data for XYZ State for the period

2013-14 to 2020-21

2. Worksheet named Weather Data: This contains district wise historical weather data for

major chickpea growing districts in XYZ state for 2013-2023

**Deliverables in the Problem Statement:**

**1.** Clean and reorganize the weather data to find monthly rainfall, and minimum and maximum

temperatures for the state XYZ (Steps taken to clean the data are left to your discretion)

**2.** Upload the cleaned weather data to any SQL server (local/cloud)

**3.** Write a query to find which month got the highest rainfall in the State

**4.** Based on the weather data and the yield data given, predict the yield of chickpea for the

years (2021-22, 2022-23, 2023-24) using python/R

**5.** Create a dashboard using a visualization tool showing exploratory data analysis on the given

data. You can use community version of any tool like Tableau, powerBI

**6.** Create a consolidated report which includes the data cleaning approach, exploratory data

analysis, modeling approach