**GOVERNMENT COLLEGE OF TECHNOLOGY COIMBATORE**

**(An Autonomous Institution Affiliated to Anna University)**

**COLLEGE CODE-7177**

BIG DATA ANALYTICS USING IBM CLOUD DATABASES

TEAM MEMBERS SIVAGOVIND G U – 71772114141

SOWMIYA P –71772114143

VISHVA SAKTHI M – 71772114150

SURITHIKA T - 71772114313

PHASE 4

**BIG DATA ANALYSIS**

**PROBLEM STATEMENT:**

* Continue building the big data analysis solution by applying advanced Analysis techniques and visualizing the results.
* Apply more complex analysis techniques, such as machine learning Algorithms, time series analysis, or sentiment analysis, depending on the Dataset and objectives.
* Create visualizations to showcase the analysis results. Use tools like

Matplotlib, Plotly, or IBM Watson Studio for creating graphs and charts.

**SOLUTION:**

Certainly, building a big data analysis solution that incorporates advanced Techniques and visualizations is essential for deriving meaningful insights from Your data. Let’s continue with the process:

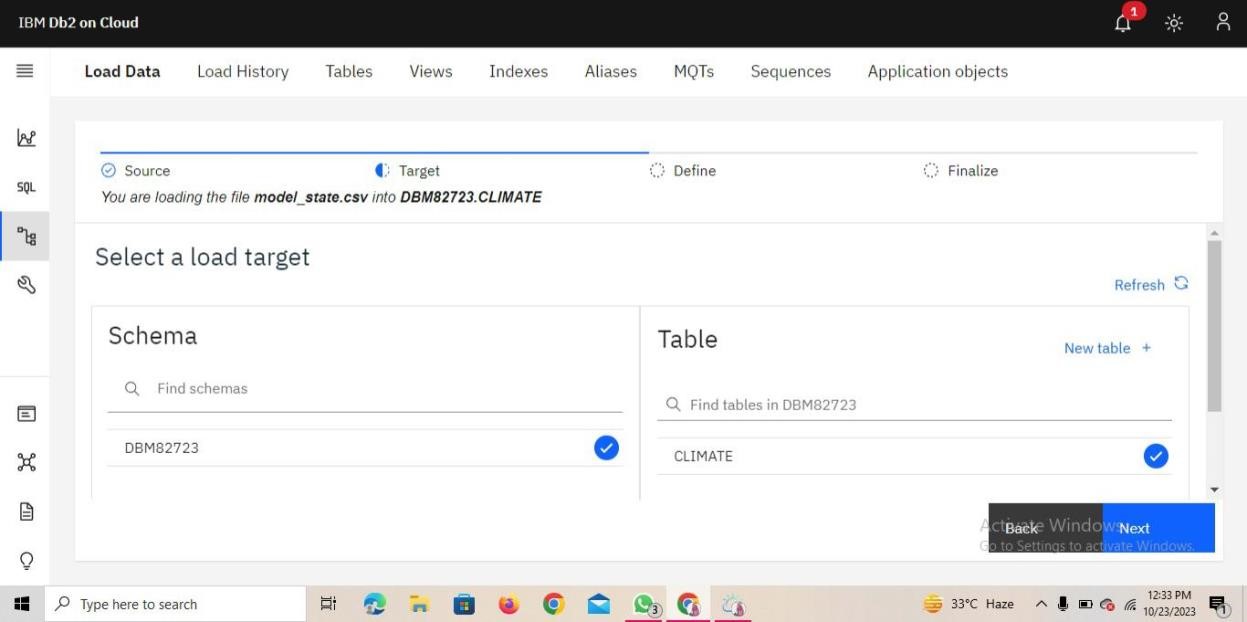
**Step 1:**

Download a CSV or xlsx file for upload in the DB2 database.

Example: open the wwb browser.

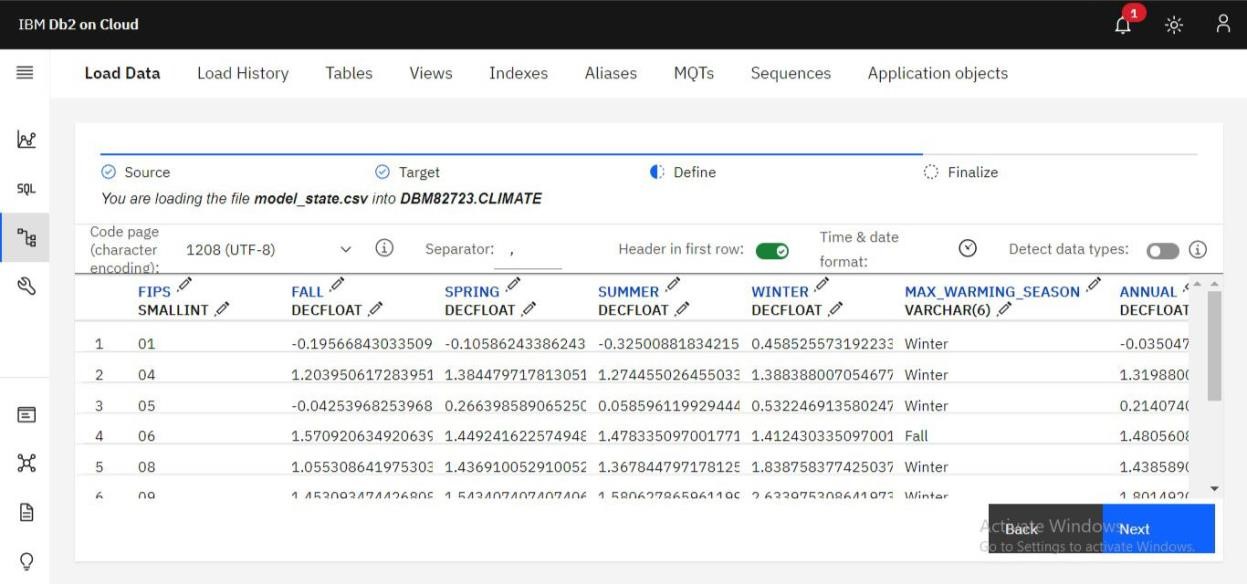
Search for the convenient topic to download database.(eg:kaggle,Data.world..) **Step 2:**

Create a data table in IBM Cloud DB2 Database.



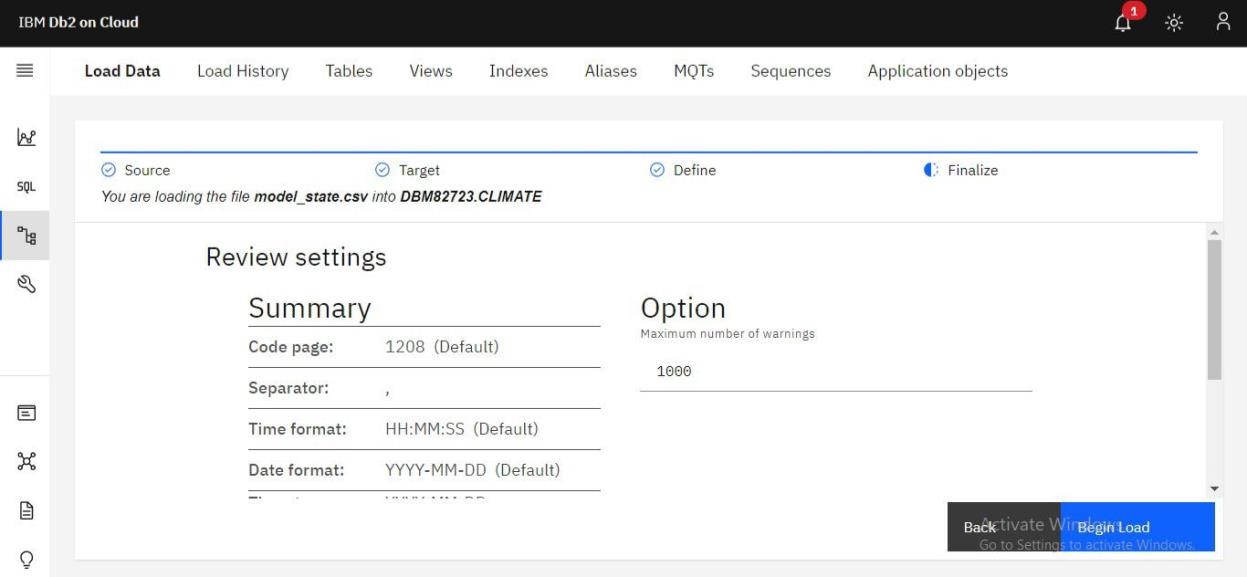
**Step 3:**

Upload the downloaded CSV. File in the database.



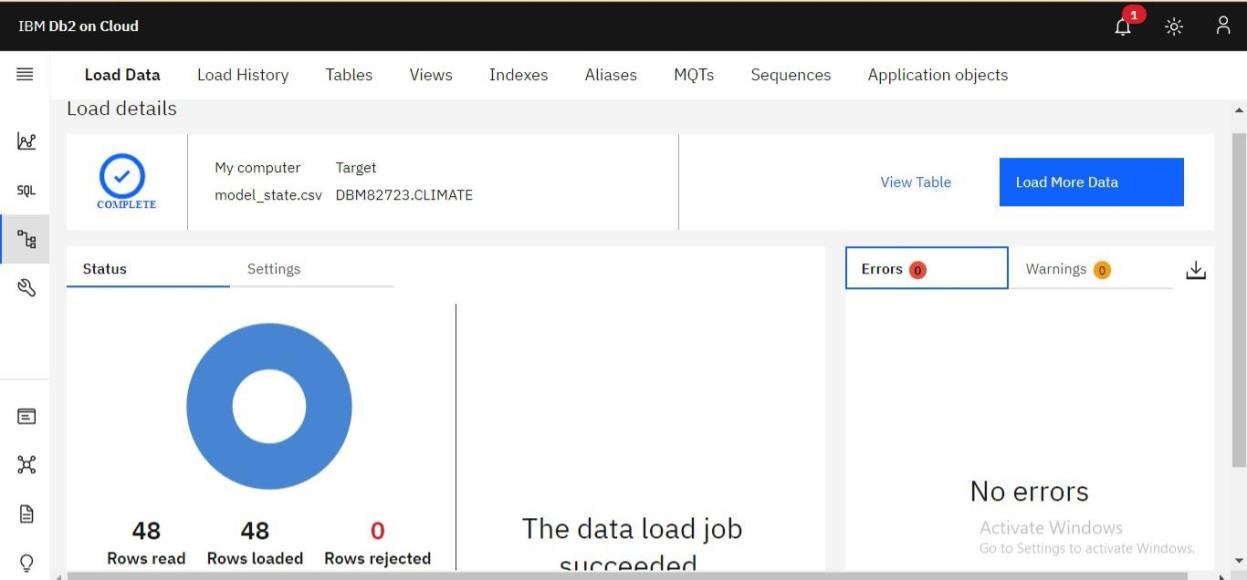
**Step 4:**

Finalize the uploading settings.



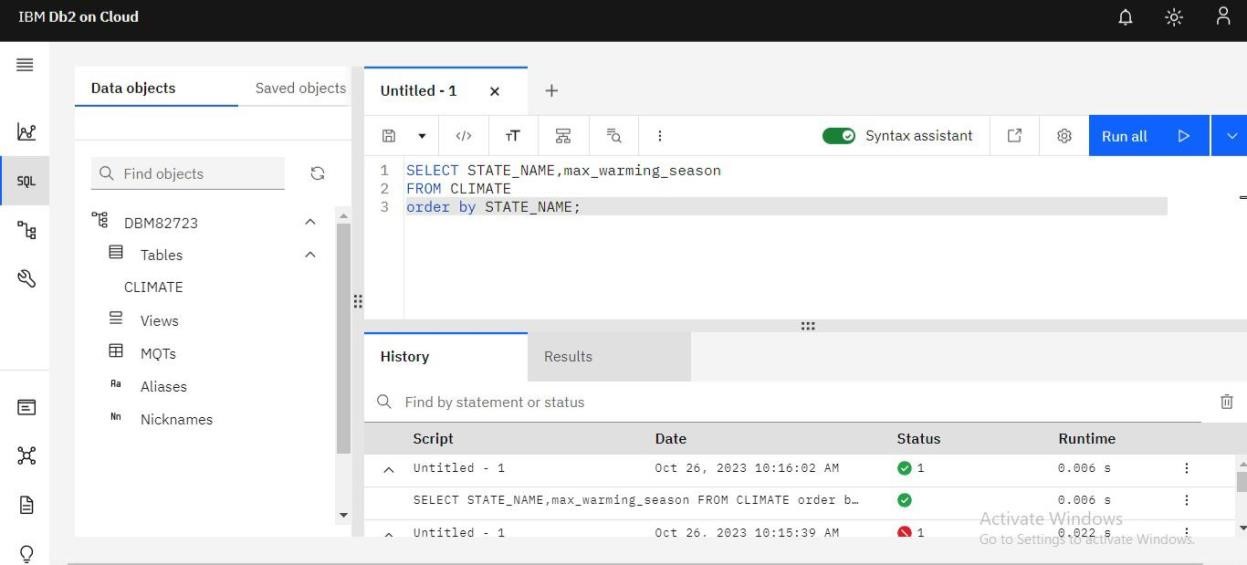
**Step 5:**

Run the loaded data to check it is contain error or not.



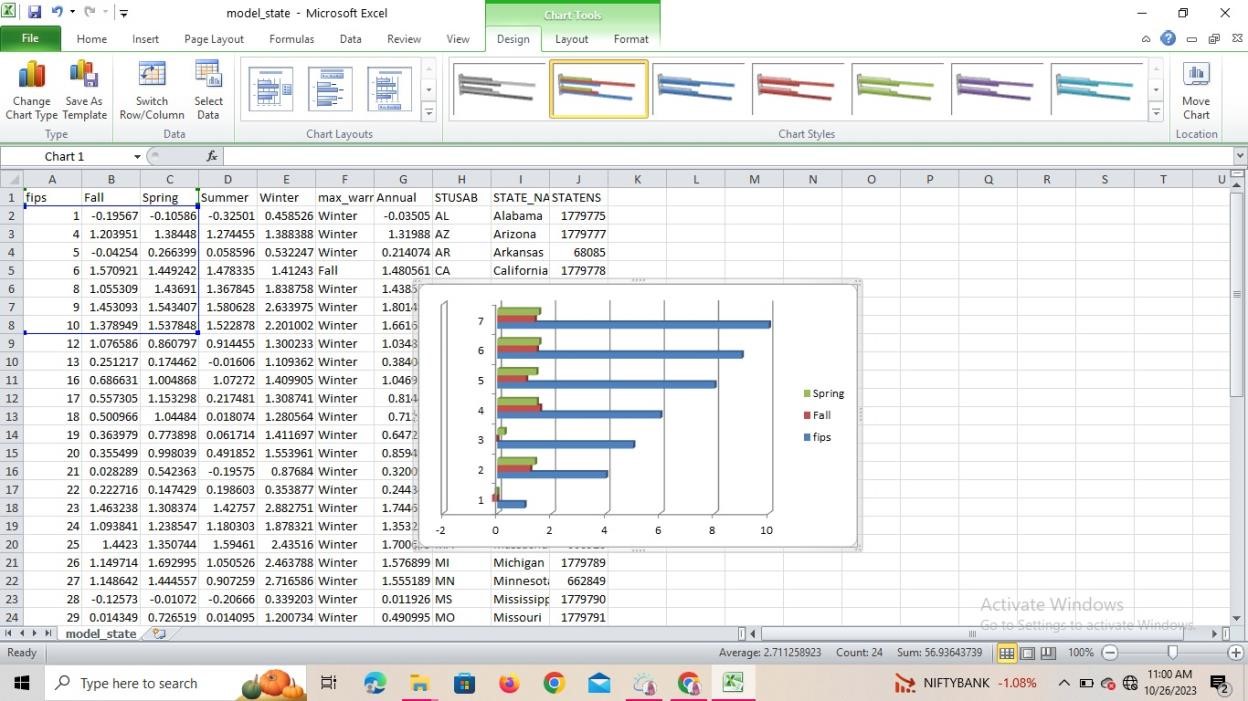
**Step 6:**

Create SQL queries to run the database table.

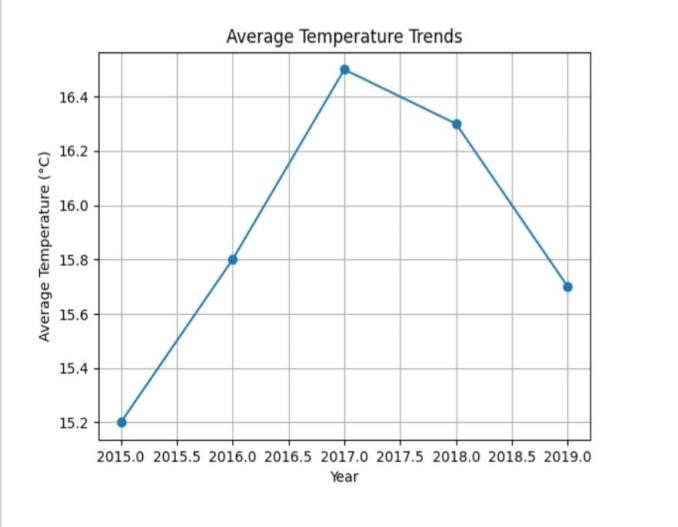
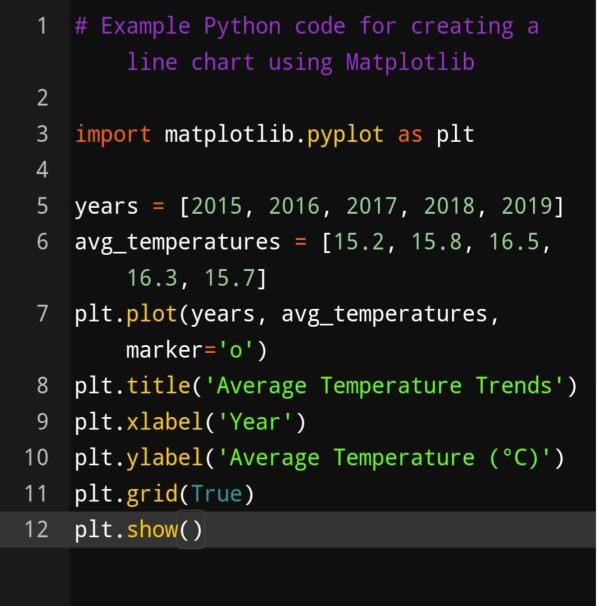


**Step 7:**

For development the analysis data we need to use the virtualization techniques in the datasets.



**Step 8: Using python.**



**Step 9:**

**Using Machine Learning techniques.**

**Select Appropriate Analysis Techniques:**

Depending on the nature of your dataset and specific objectives, consider various  **Advanced analysis techniques:**

**Machine Learning Algorithms:** Use supervised or unsupervised machine learning

Algorithms like decision trees, random forests, support vector machines, or Clustering algorithms for predictive modeling or pattern recognition. **Time Series Analysis:** If your data involves time-based data points, use time

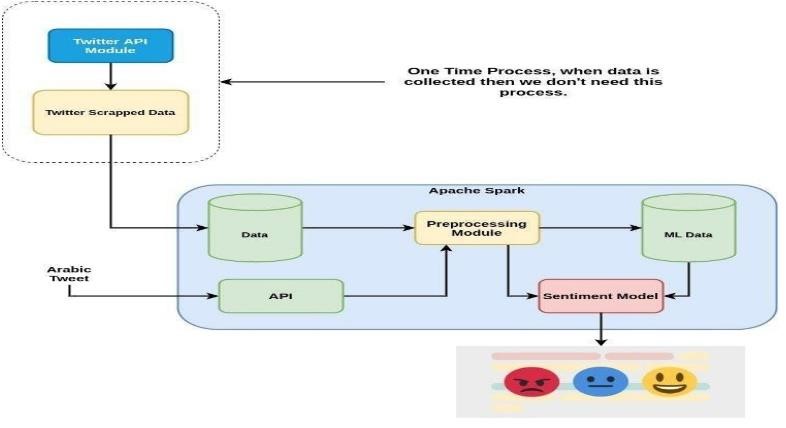
Series analysis techniques to identify trends, seasonality, and forecast future Values.

**Sentiment Analysis**: Apply natural language processing techniques to extract Sentiment from text data, useful for social media or customer reviews analysis.

**Example:**

# Example Python code for sentiment analysis using NLTK import nltk from nltk.sentiment import SentimentIntensityAnalyzer nltk.download(‘vader\_lexicon’) sia = SentimentIntensityAnalyzer()

text = “The weather is wonderful and the scenery is breathtaking.” sentiment\_score = sia.polarity\_scores(text) print(sentiment\_score)



**Conclusion:**

Thus the ,Continue building the big data analysis solution by applying advanced analysis techniques

And visualizing the results has been completed.