Weather Report Forecasting Analysis & Graphical Speculation on Dashboard

Weather prediction is one of the most required information all over the regions. It involves collecting global meteorological surface and upper-air observations and preparing a worldwide surface and upper-air pressure, temperature, moisture, and wind analyses at frequent time intervals based upon these observations we predict some data for upcoming days' weather conditions.

Project Description

This data contains day-wise weather attributes from 2022 to July 2033 (predicted data).

Columns:

Date	Average temperature (°F)
Average humidity (%)	Average dewpoint (°F)
Average barometer (in)	Average windspeed (mph)
Average gust speed (mph)	Average direction (°degree)
Rainfall for month (in)	Rainfall for year (in)
Maximum rain per minute	Maximum temperature (°F)
Minimum temperature (°F)	Maximum humidity (%)
Minimum humidity (%)	Maximum pressure
Minimum pressure	Maximum wind speed (mph)
Maximum gust speed (mph)	Maximum heat index (°F)

The Project will consist of 2 modules:

Module 1: Pre-processing, Analysing data using Python and SQL.

Module 2: Visualizing data using Power BI/Tableau.

Module 1: Pre-processing, Analysing data using SQL

In this module, you will query the dataset using SQL to gain insights from the database. The problem statements to be solved will be provided to you and you need to provide the solution for the same using your log

Task 1: Pre-processing the data

Data pre-processing is one of the important steps in data analytics because data that is not processed can lead to different unwanted results when the data is used for further applications. This task includes sub-tasks such as handling null values, deletion or transformation of irrelevant values, data type transformation, removing duplicates, etc.

Data set:

https://drive.google.com/drive/folders/1gcogu8wm0KOQ1biyVxf_dHtqLpFoV2OV

- **Subtask 1**: Correct years for given data set
- **Subtask 2**: removal of duplicate rows and duplicate Columns
- **Subtask 3**: fix a few labels in the given data set
- **Subtask 4**: Encoding data into suitable format

The dataset we used in this project was populated using a scraping technique, the encoding method may be different. We prefer UTF-8 encoding as it is majorly used in all the database servers. To do so, perform the below tasks:

1. In the menu bar of excel, select option data. Then select from text/CSV and choose the file in which you made changes using excel.

2. The default file origin will be UTF -8. Keep all the field as it is and select load.

3. Save the file as a CSV file.

4. Import the file into your SQL database (Analyze the dataset with respect to your database system and then change it, if necessary).

Task 2: Analysing data

Since you have analyzed the complete data set accurately, Let's answer the following requirements to proceed further for a better understanding

- Give the count of the minimum number of days for the time when temperature reduced
- Find the temperature as Cold/hot by using the case and avg of values of the given data set
- Can you check for all 4 consecutive days when the temperature was below 30 Fahrenheit
- Can you find the maximum number of days for which temperature dropped
- Can you find the average humidity average from the dataset
  (NOTE: should contain the following clauses: group by, order by, date)
- Use the GROUP BY clause on the Date column and make a query to fetch details for average windspeed (which is now windspeed done in task 3)
- Please add the data in the dataset for 2034 and 2035 as well as forecast predictions for these years
  (NOTE: data consistency and uniformity should be maintained)

- If the maximum gust speed increases from 55mph, fetch the details for the next 4 days
- Find the number of days when the temperature went below 0 degrees Celsius
- Create another table with a "Foreign key" relation with the existing given data set.

Module 2: Visualizing data using Power BI

Here comes the next module. You will use the weather datasets and the data table you created to visualize what the numbers in and behind the data want to convey. You have to create a dashboard for the same using different statistical graphs and diagrams for a visual understanding and analysis. Given below are the requirements, create a dashboard consisting of mentioned visuals.

- Design a dashboard of your choice and imagination with attractive wallpapers and design visuals
- The dashboard must consist of basic power bi visuals like stacked bar charts, cards, line charts, pie charts, etc.
- You can use any visual capable of representing the given dataset.
- Columns having numeric data types are to be majorly used in visuals
- Include the Weather dataset and the data table you created in the above task (question 10).