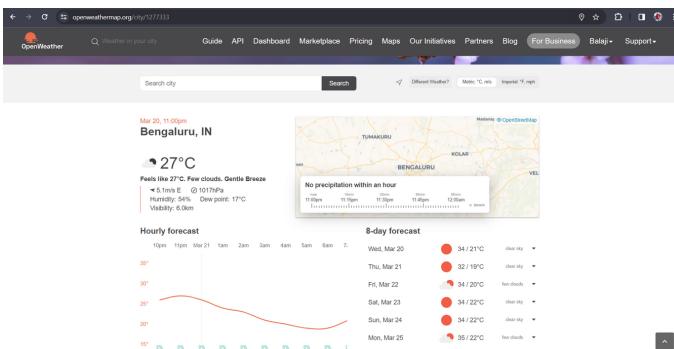


Automated Weather Data Pipeline

Developed a comprehensive automated data pipeline for real-time weather data acquisition and analysis. Leveraged Python (BeautifulSoup, pandas) for web scraping and data manipulation. Integrated PySpark for efficient processing of large datasets and Snowflake for scalable storage. Implemented dynamic schema creation, reducing manual efforts in data structuring. Provided structured, query-ready datasets facilitating advanced weather analytics for informed decision-making. Enabled seamless workflow from data collection to analysis, supporting various weather-dependent applications and industries.

Data Source & Data Stored:



The screenshot shows a table of weather information for Bengaluru. The table includes columns for rank, temperature, humidity, wind speed, general weather, current date and time, sunrise, and sunset.

Rank	Temperature in bangalore	Humidity in bangalore	Wind Speed in bangalore	General Weather in bangalore	Current date and time in bangalore	Sunrise in bangalore	Sunset in bangalore
1	31.77°C or 89.19°F						
2	Temperature in bangalore feels like: 30.45°C or 86.81°F						
3	Humidity in bangalore: 29%						
4	Wind Speed in bangalore: 2.57 m/s						
5	General Weather in bangalore: clear sky						
6	Current date and time in bangalore: 2024-03-19 15:52:09						
7	Sunrise in bangalore: 2024-03-19 06:24:21 local time.						
8	Sunset in bangalore: 2024-03-19 18:30:18 local time.						

ETL Process Design Diagram

