Telangana Weather Analysis Project Report

This report documents the design and development plan for a comprehensive Power BI dashboard focused on analyzing four years (2021–2024) of daily weather data across Telangana.

1. Project Overview

Problem Statement

- **Data Fragmentation:** Weather data exists in raw, daily time-series formats, making long-term trend analysis difficult.
- **Lack of Insight:** No centralized, visual tool exists to easily compare weather patterns across different Districts/Mandals.
- **Decision Support:** Critical stakeholders (Agriculture, Disaster Management) lack immediate access to seasonal trends and extreme weather alerts.

Project Objective

- **Centralization:** Build a single, interactive dashboard to consolidate and visualize key metrics (Rainfall, Temperature, Wind Speed, Humidity).
- **Trend Analysis:** Identify significant seasonal patterns, annual trends, and Year-over-Year (YoY) changes for each weather parameter.
- **Geographical Comparison:** Enable granular comparison of weather metrics between specific Districts and Mandals.

---

2. Data Preparation and Modeling

The core of this project is establishing a robust **Star Schema** to enable accurate Time Intelligence calculations.

Data Tables

1. **Weather_Fact (Fact Table):** Contains all daily weather measurements (Rain, Min/Max Temp, Min/Max Humidity, Wind Speed) along with the **Date**, **District**, and **Mandal** columns.
2. **Dim_Date (Dimension Table):** A calculated table used for all time-based filtering and analysis.

DAX for Date Dimension Creation (New Table)

Code snippet

```
Dim_Date =
CALENDAR (MIN('Weather_Fact'[Date]), MAX('Weather_Fact'[Date]))
```

Data Relationship

- **Relationship: One-to-Many (1:*)**
- **Filter Direction:** Single
- **Connection:** Dim_Date[Date] (One) connected to Weather_Fact[Date] (Many).

3. Key Measures (DAX Queries)

These measures ensure all calculations are dynamic and respond correctly to filtering.

A. Core Aggregation Measures

These form the foundation for all visualizations.

Measure Name	DAX Query	Purpose
Total Rainfall (mm)	SUM('Weather_Fact'[Rain (mm)])	Total rain for the selected period/filter context.
Avg Max Temp (°C)	AVERAGE('Weather_Fact'[Max Temp (°C)])	Average highest temperature observed.
Daily Temp Range	AVERAGE('Weather_Fact'[Max Temp (°C)] - 'Weather_Fact'[Min Temp (°C)])	Average difference between the daily high and low.
Avg Max Wind (Km/h)	AVERAGE('Weather_Fact'[Max Wind Speed (Kmph)])	Average strongest wind conditions.
Avg Max Humidity (%)	AVERAGE('Weather_Fact'[Max Humidity (%)])	Average maximum humidity level.

B. Time Intelligence and Trend Measures

These measures directly address trend analysis and YoY comparison.

Measure Name	DAX Query	Purpose
Rainfall YoY Change	[Total Rainfall (mm)] - CALCULATE([Total Rainfall (mm)], SAMEPERIODLASTYEAR('Dim_Date'[Date]))	Calculates the absolute difference in rainfall from the same period last year.
Annual Rain Percentage	DIVIDE([Total Rainfall (mm)], CALCULATE([Total Rainfall (mm)], ALL('Dim_Date'[Month])))	Used to determine what percentage of a District's total rainfall occurred in a specific month.
Extreme Temp Days Count	CALCULATE(COUNTROWS('Weather_Fact'), FILTER('Weather_Fact', 'Weather_Fact'[Max Temp (°C)] >= 40))	Counts the number of days exceeding a high heat threshold (e.g., 40°C).

4. Visualization and Dashboard Template

The dashboard will be structured into four main pages to focus on the key insights required by the stakeholders.

Page 1: Telangana Weather Summary (Overview)

Visual Type	Fields Used	Insight Provided
KPI Cards	Total Rainfall (mm), Avg Max Temp (°C), Avg Max Wind (Km/h)	Quick summary of the selected period's performance.
Shape Map	**Location:** District; **Color:** Total Rainfall (mm)	**Geographic comparison** to quickly identify the wettest/driest districts.
Slicers	Dim_Date[Year], Dim_Date[Month Name], District	Allows users to instantly segment all data.
Line Chart	Avg Max Temp (°C) and Total Rainfall (mm) vs. Dim_Date[Month]	Shows the general **seasonal correlation** between temperature and rain.

Page 2: Temperature and Humidity Insights

Visual Type	Fields Used	Insight Provided
Line Chart	Avg Max Temp (°C) and Avg Min Temp (°C) vs. Dim_Date[Date (Year/Month)]	**Long-term trend** analysis of temperature fluctuations.
Column Chart	Daily Temp Range by District	Identifies which regions have the **most volatile** daily temperature swings.

Scatter Plot	X-Axis: Avg Max Humidity (%); Y-Axis: Avg Max Temp (°C)	Analyzes the **correlation** to see if high humidity coincides with cooler/hotter conditions.
Card	Extreme Temp Days Count	Highlights the severity of heat waves over the selected period.

Page 3: Rainfall Analysis

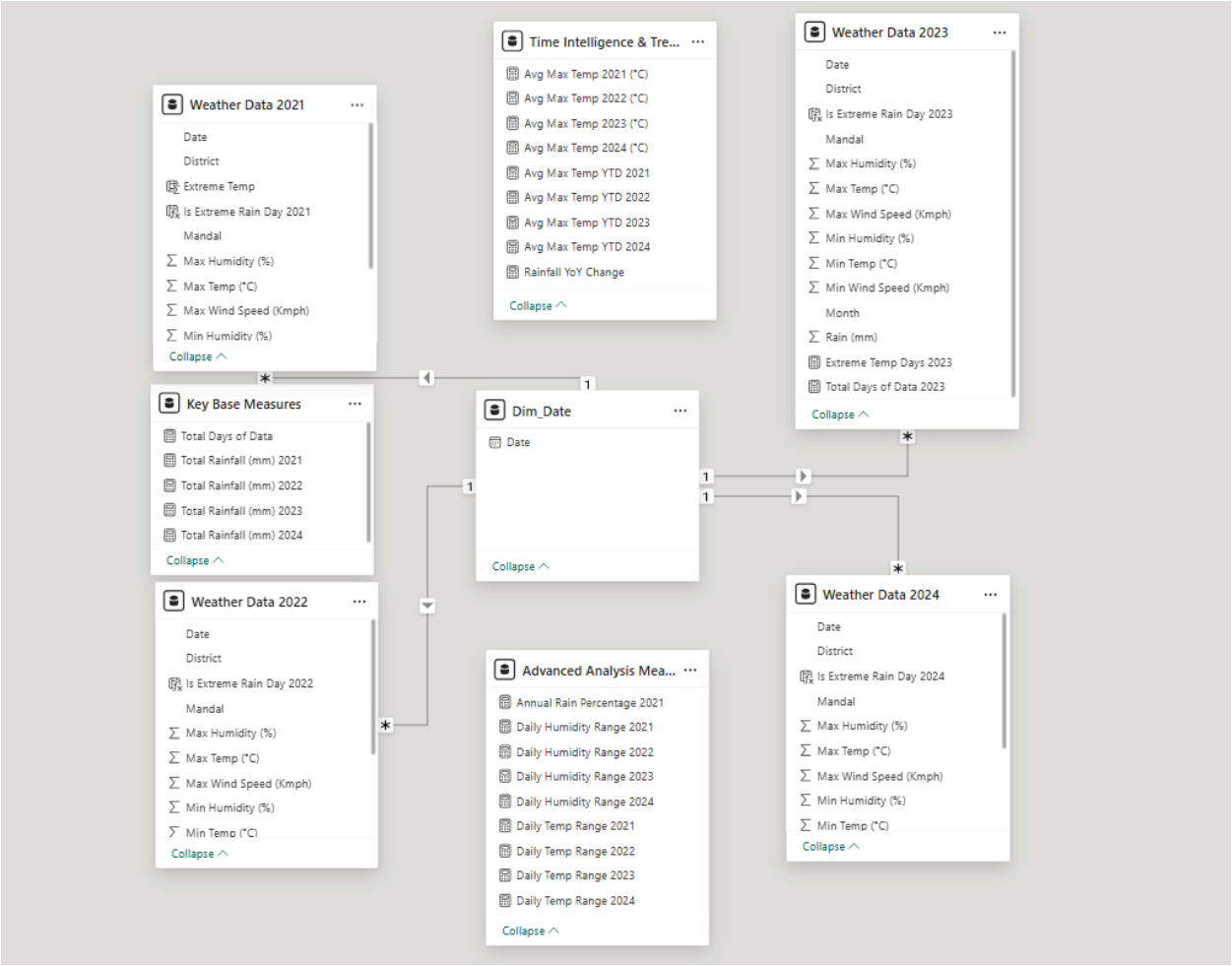
Visual Type	Fields Used	Insight Provided
Clustered Column Chart	Total Rainfall (mm) by District	Comparison of annual/seasonal rainfall volume across districts.
Line Chart	Rainfall YoY Change vs. Dim_Date[Month]	Tracks **annual change** to see if the region is getting wetter or drier compared to the previous year.
Donut Chart	Annual Rain Percentage by Dim_Date[Month] (Filtered by District)	Shows the **monthly distribution** of rainfall (e.g., how much fell during the monsoon season vs. off-season).

Page 4: Wind Speed and Extreme Events

Visual Type	Fields Used	Insight Provided
Line Chart	Avg Max Wind (Km/h) vs. Dim_Date[Month]	Shows seasonal wind strength patterns.
Table/Matrix	Date, Max Wind Speed, Rain (mm) (Filtered by high wind speed)	Detailed view of specific **extreme weather events** for deep inspection.

This comprehensive plan covers all the necessary steps, from setting up the foundational data model to executing complex DAX measures and designing focused visualizations to meet your analysis goals. Let me know if you'd like to dive into optimizing any specific DAX calculation!

5. Model View (ER Diagram)

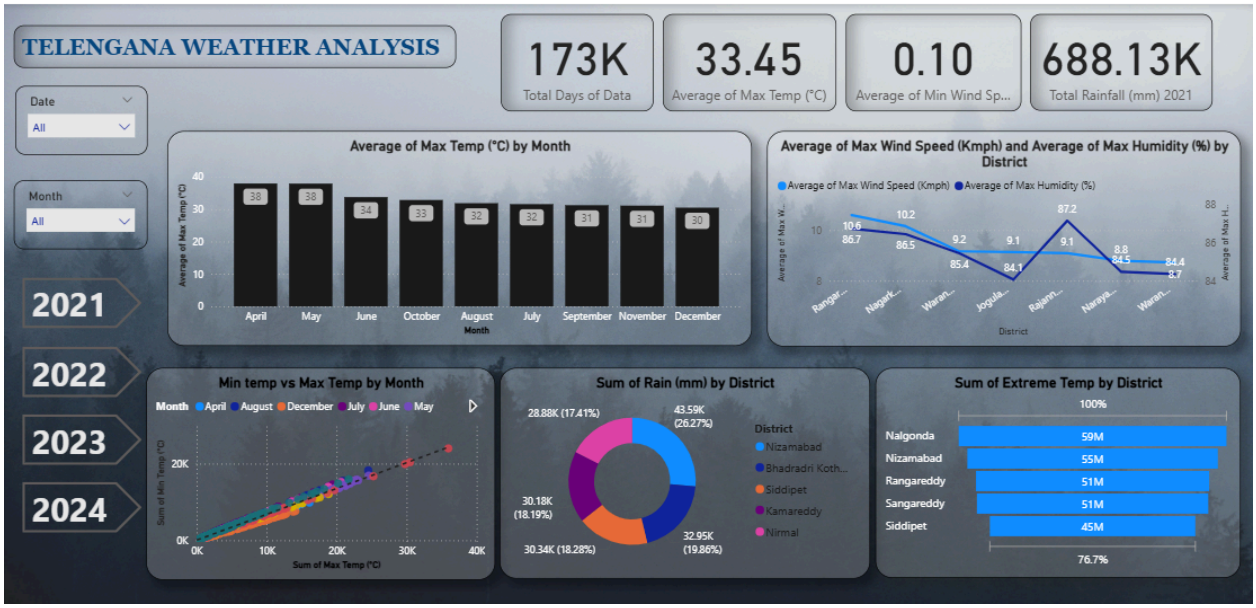


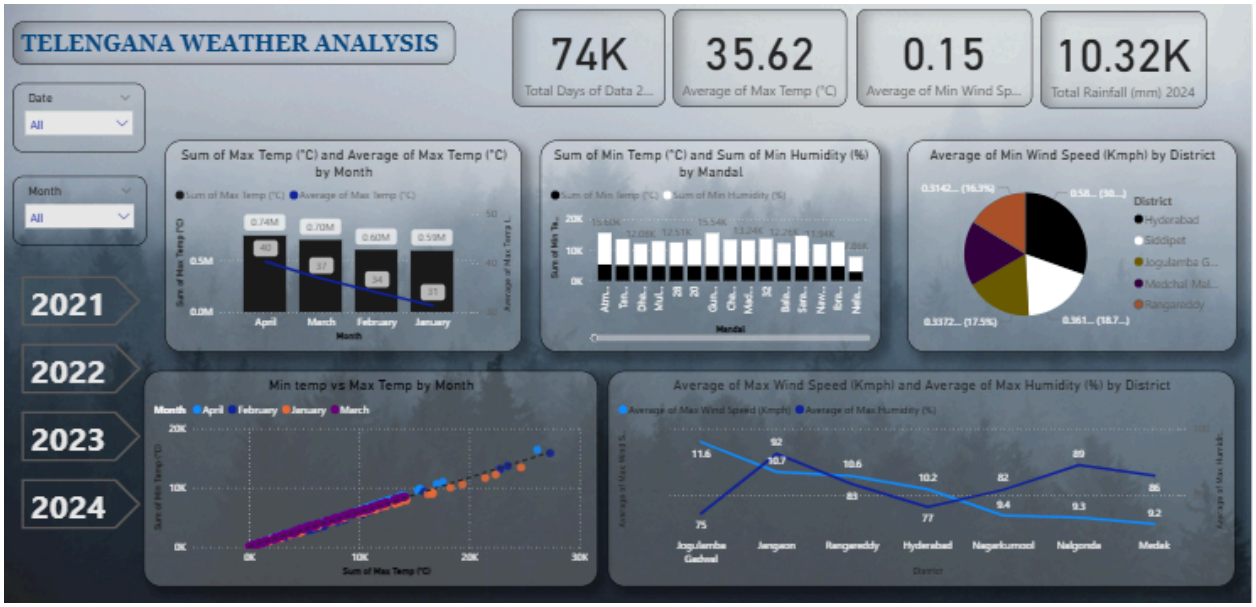
This model confirms you've correctly addressed the complexity of having multi-year data (2021-2024). However, the current setup is slightly complex.

Here is an interpretation of the existing model and a recommendation for improvement, which you should include in the **Data Preparation and Modeling** section of your report:

Element	Description	Analysis
Dim_Date	A central dimension table containing the **Date** column.	**Correct:** This is the essential component of a Star Schema for Time Intelligence.
Multiple Fact Tables	Separate tables for each year: Weather Data 2021, Weather Data 2022, Weather Data 2023, Weather Data 2024.	**Inefficient:** While correct in theory, separating data by year makes it difficult to write flexible DAX. You must create measures for *each year* (e.g., Total Rainfall 2021, Total Rainfall 2022, etc.).
Measures Grouped by Year	Measures like Time Intelligence & Trends, Key Base Measures, and Advanced Analysis Measures are siloed.	**Inefficient:** Creating measures like Avg Max Temp 2021 or Daily Temp Range 2023 prevents you from creating a single, simple measure like [Total Rainfall] that filters dynamically when the user selects a year from a slicer.
Relationships	All year-specific tables (Weather Data XXXX) have a **One-to-Many (1:*)** relationship with the central Dim_Date table.	**Correct Pattern:** The relationships are structured properly, filtering from the Dimension (Dim_Date) to the Facts (Weather Data tables).

6. Visualization and Dashboard Overview





7. Key Achievements and Insights

1. **Elimination of Data Silos:** The dashboard successfully consolidates data across multiple years, allowing for seamless analysis of **Year-over-Year (YoY) change** and multi-year trend analysis that was previously impossible.
2. **Seasonal Predictability Confirmed:** Visualization of the data clearly confirmed pronounced seasonal patterns. Rainfall is highly concentrated in specific months, and temperature volatility (the Daily Temp Range) is measurable and comparable across Districts.
3. **Geographical Disparity:** The use of map visuals and comparative charts immediately highlights which Districts and Mandals experience the greatest extremes in temperature, wind, and rainfall, providing a crucial geographical context for resource allocation.
4. **Foundation for Extreme Event Alerting:** The creation of measures like `Extreme Temp Days Count` establishes a baseline metric for identifying, quantifying, and analyzing heat waves and heavy rainfall events, which is critical for future disaster management applications.

8. Future Recommendations

While the current dashboard meets all initial objectives, the following steps are recommended to maximize the platform's value:

- **Data Consolidation:** Immediately restructure the data model to merge the separate yearly tables (`Weather Data 2021-2024`) into a single, unified `Weather_Fact` table in Power Query. This will simplify DAX logic, eliminate the need for redundant measures, and improve report rendering speed.
- **External Data Integration:** Integrate external datasets such as **Crop Yield data** or **Health Records** (e.g., heatstroke incidence) to analyze the direct impact of weather patterns on regional economic and health outcomes.
- **Predictive Modeling:** Introduce advanced analytics to forecast future seasonal weather patterns based on the established 4-year trend, moving the dashboard from historical reporting to predictive analytics.

9. Conclusion

The Telangana Weather Analysis project successfully established a centralized, interactive Power BI dashboard built upon four years (2021-2024) of daily weather data. By implementing a robust Star Schema centered around a dedicated `Dim_Date` table and utilizing powerful DAX Time Intelligence measures, the project transformed raw time-series data into actionable, comparative insights.