Overview

The Approach to the problem is on the understanding that Weather is the primary driver of AC usage which is directly correlated to AC service.

1. Weather impacts AC service demand.
2. Marketing spend would inturn capture a percentage of the AC service demand

PART A : Descriptive Analytics Objective:

To Summarize historical performance and key trends M-O-M and W-O-W level.

Orders

Annual orders in 2023 was 4.8 Cr , 2024 was 5.6 cr and 2.3 Cr till now in 2025

M-O-M trend for Orders

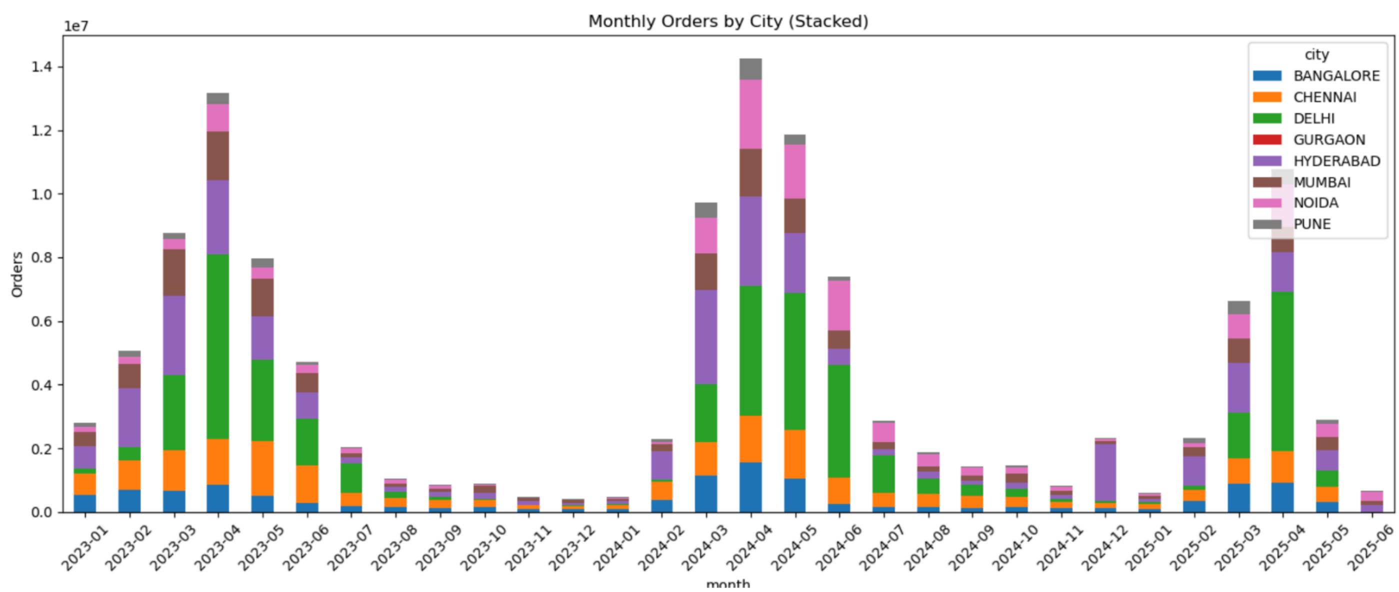


Fig (a) : M-o-M trend for orders

- Month on Month trend shows a seasonal pattern
- Top Months are March, April and May for each year as seen by Table 1
- The top 2 city for 2023 and 2024 are constantly Delhi followed by Hyderabad
- Top City by Month are Delhi and Gurgaon and their top months are April and March which is shown in Fig (b).

Table 1 : Top Orders by Month in each year

Years	**Month**	Orders
2023	Apr-23	13175514.0
2023	Mar-23	8773974.0
2024	Apr-24	14260059.0
2024	May-24	11857725.0
2025	Apr-25	10771992.0
2025	Mar-25	6641910.0

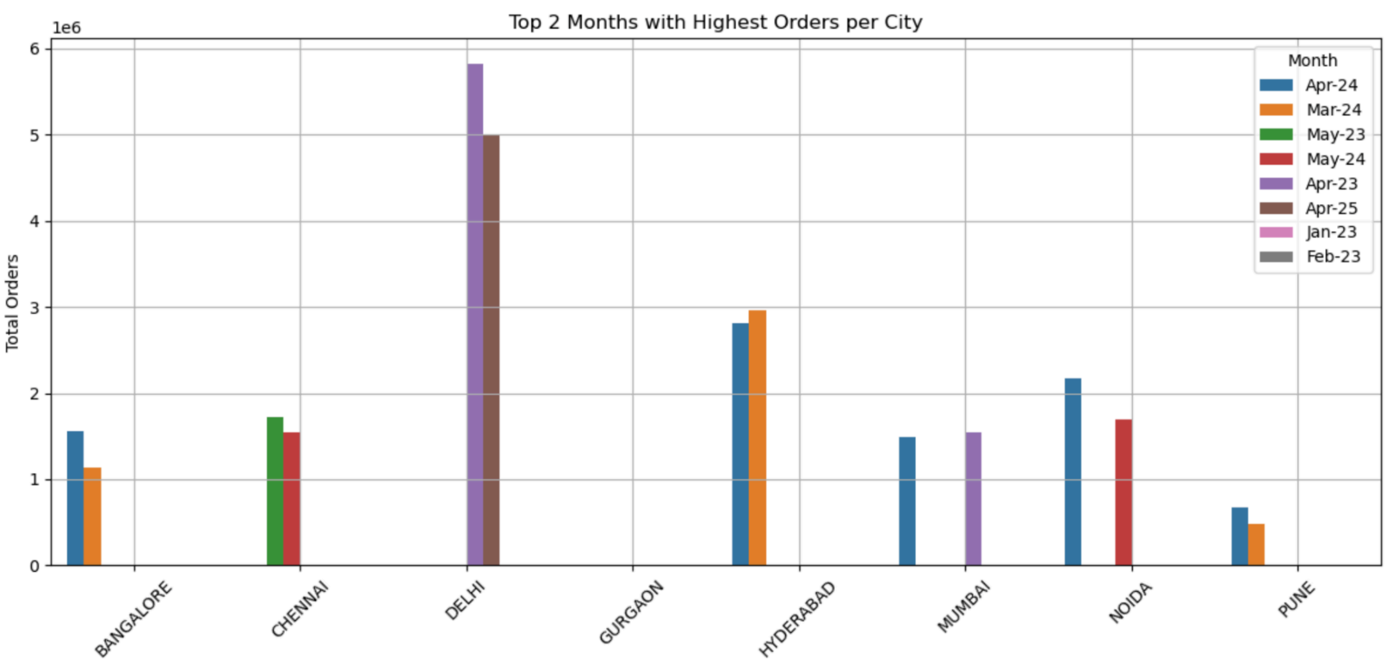


Fig (b) : Top Months Orders by City

W-O-W trend for Orders

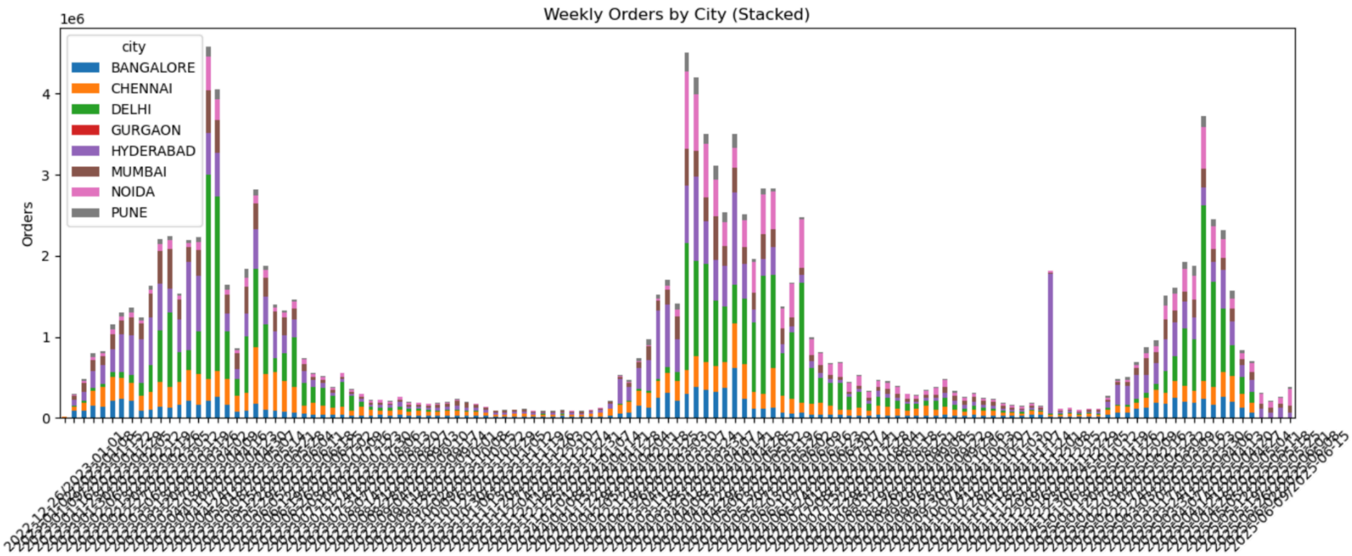


Fig (c) : W-o-W trend for Orders

- Week on Week pattern performs similarly to M-o-M , and shows seasonality and repeatability each year
- Most Ordered weeks are present in March, April and May as shown by Table 2
- Weekly streaks of 3-4 weeks where the growth is increasing followed by decreasing growth , and it follows cyclically as shown in Fig (d)

Table 2 : Showing top weeks by Orders

No.	Weeks	Orders
1	2023-04-10/2023-04-16	4579146.0
2	2023-04-17/2023-04-23	4049694.0
3	2023-05-15/2023-05-21	2821500.0
4	2024-03-25/2024-03-31	4506876.0
5	2024-04-01/2024-04-07	4194630.0

6	2024-04-08/2024-04-14	3508560.0
7	2025-04-07/2025-04-13	3723885.0
8	2025-04-14/2025-04-20	2448864.0
9	2025-04-21/2025-04-27	2314422.0

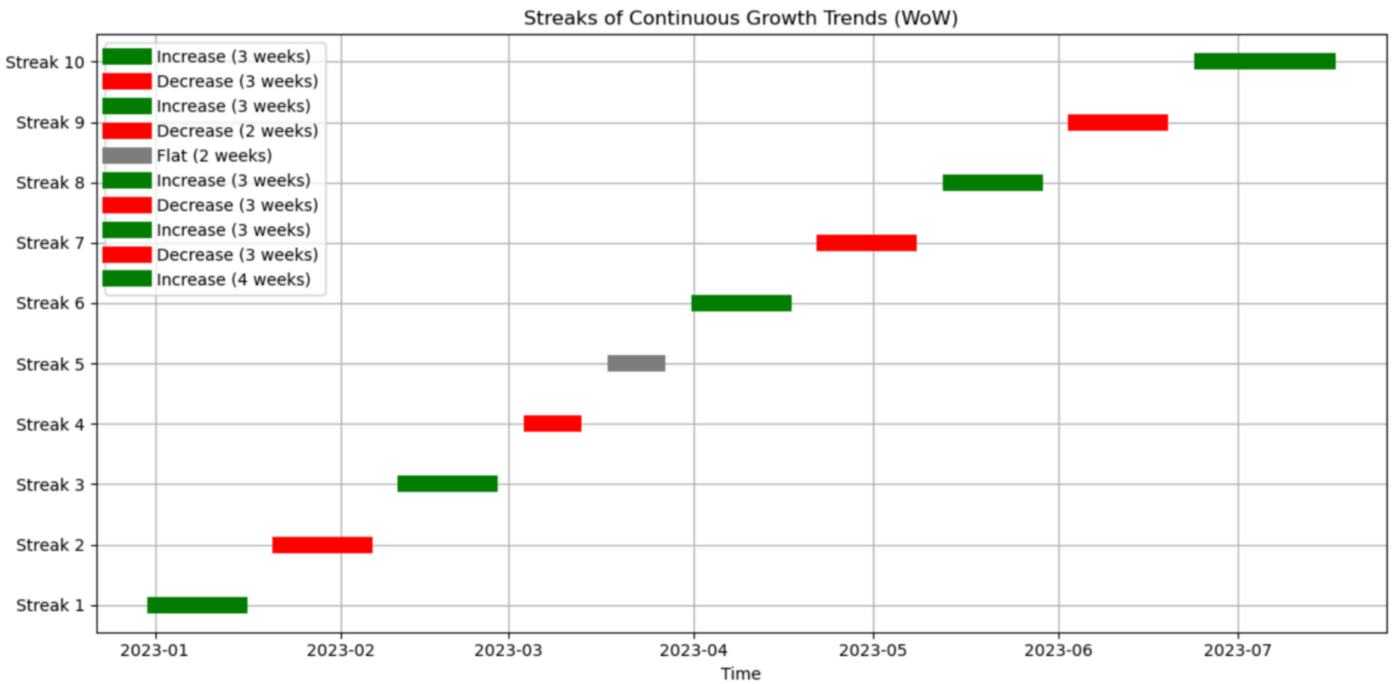


Fig (d) : Weekly streaks where growth is increasing consecutively or decreasing consecutively

Weather

Temp, Tempmax, Tempmin, feelslike, Humidity, Precip are the main features that could drive the Market.

After analyzing tempmax, tempmin, feelslike have more than 75% correlation with temp and humidity. So having these features would be redundant as its characteristics are already present in temp and humidity.

temp, humidity, precip are the main factors to be considered.

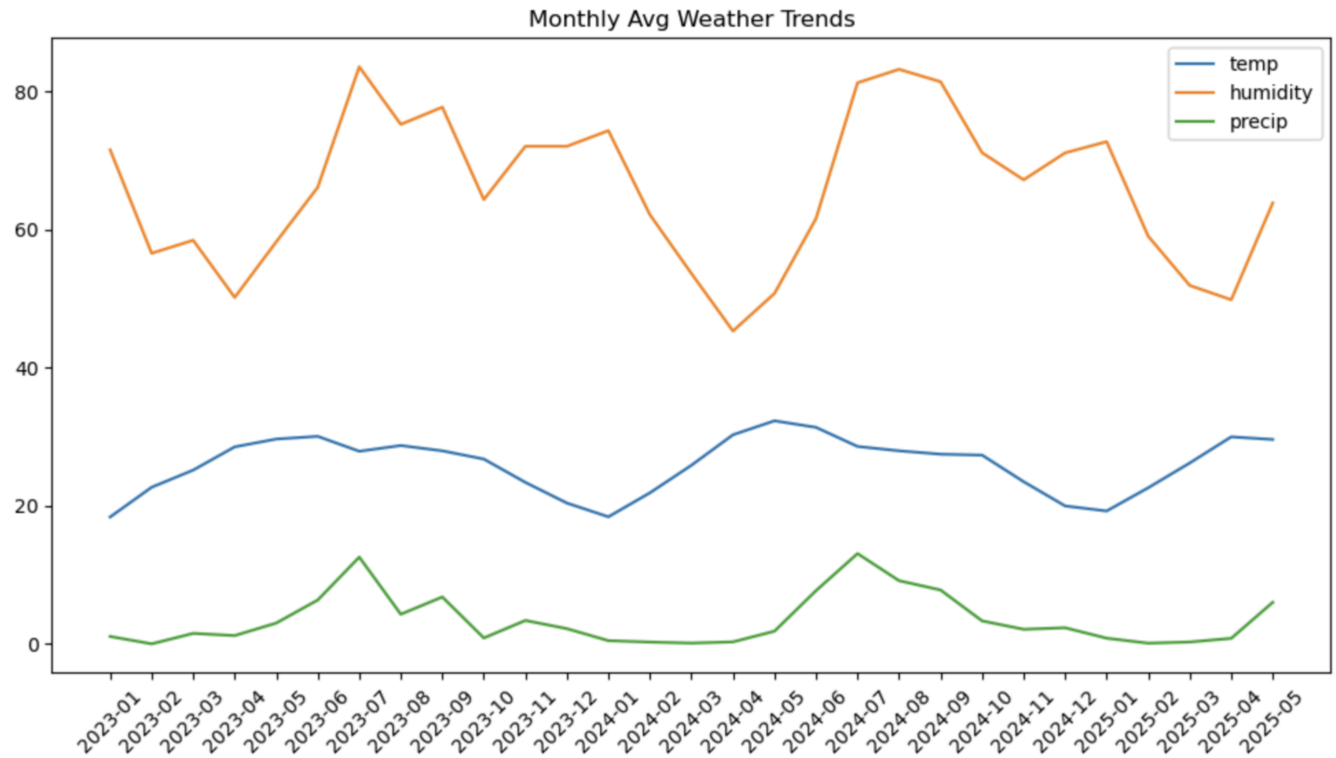


Fig (e) : Monthly Weather Data

- Temperature, humidity and precip shows cyclical trends

Marketing

Annual Marketing spend is 335 Cr in 2023, 280 Cr in 2024 and 47 Cr till May in 2025

M-O-M Trends for Marketing

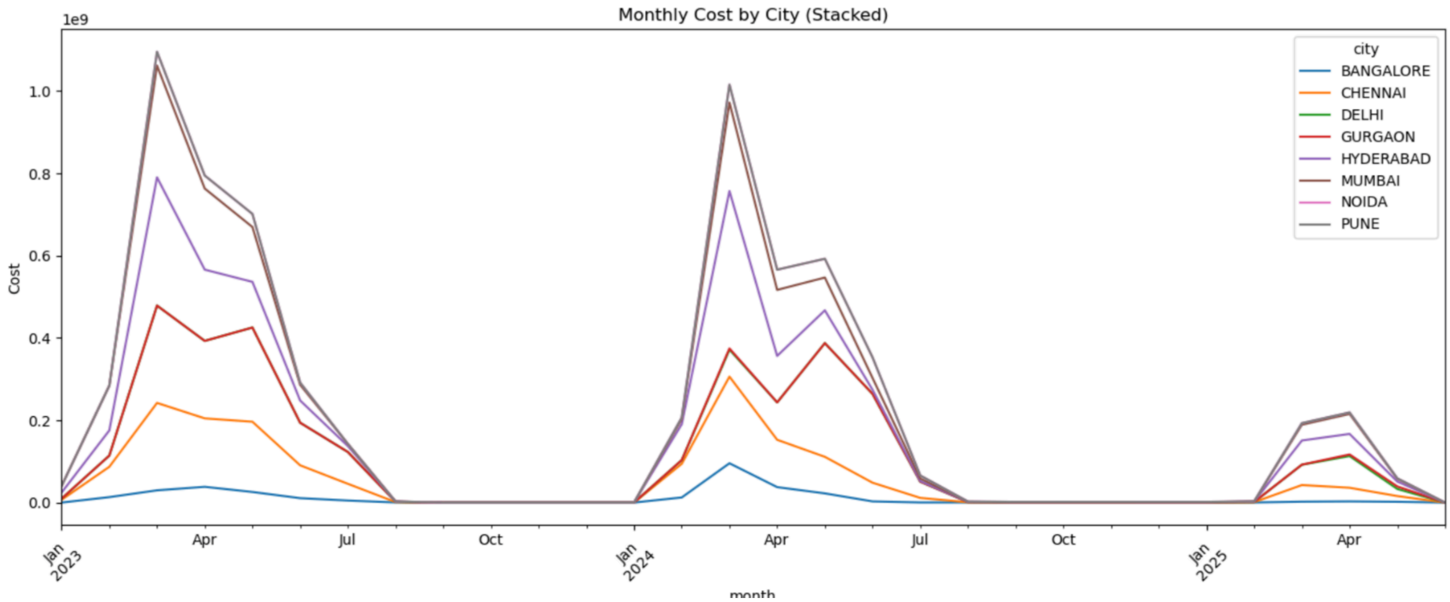
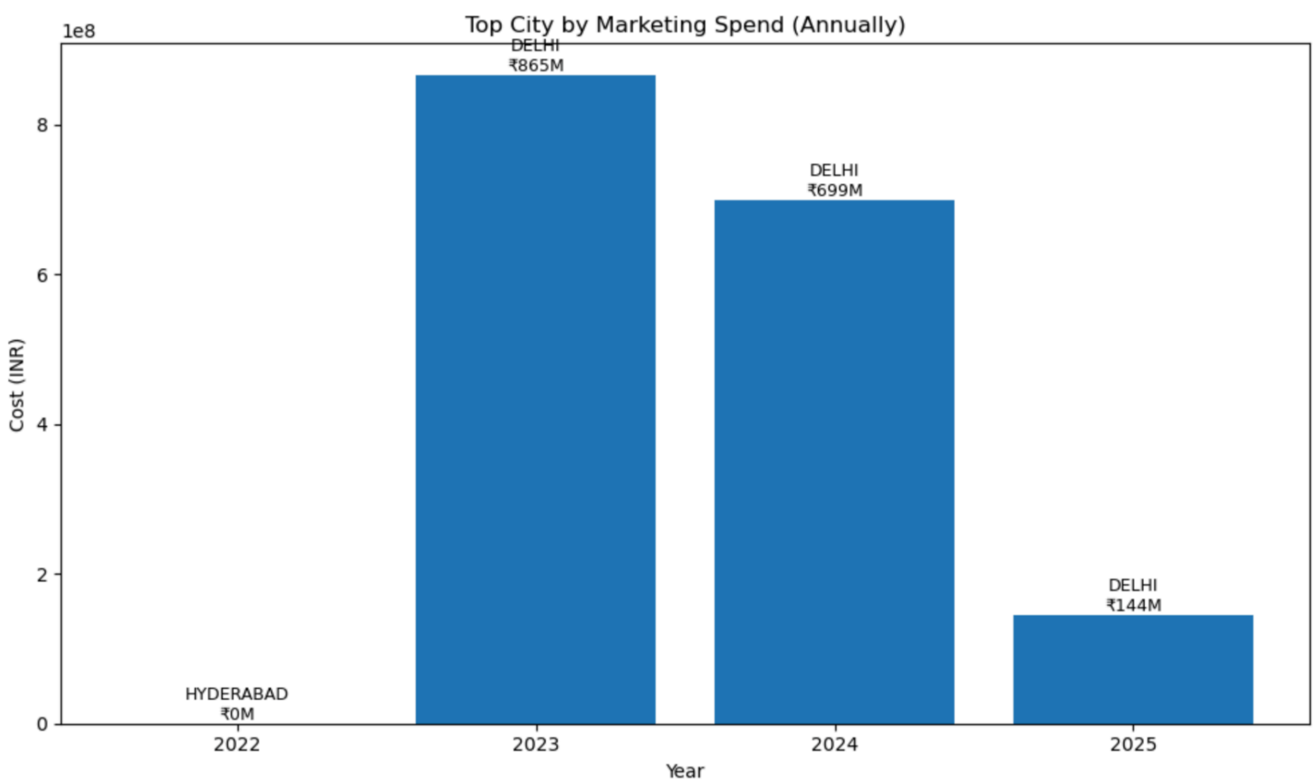


Fig (f) : Monthly Marketing Spend

- Monthly marketing spend shows sharp peaks and a pattern which is repeated in subsequent years
- Top Marketing Spend Annually by City is **Delhi**



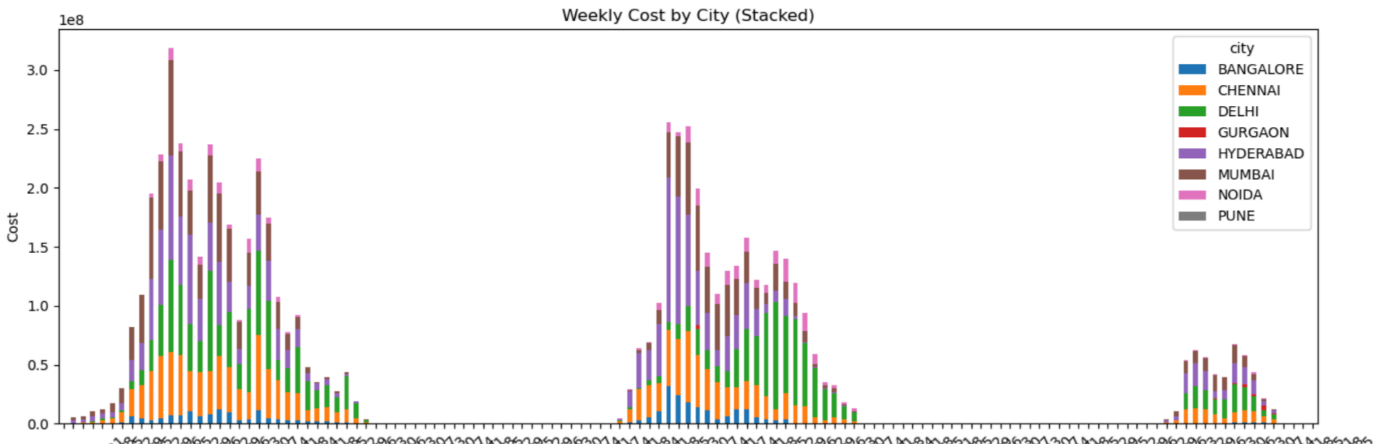
Fig(g) : Top Marketing Spend by City Anually

- **Mumbai, Hyderabad, Gurgaon, Pune** show consistently higher marketing spends.
- **Bangalore, Chennai** have **lower or flatter spend curves**, possibly indicating less aggressive campaigns or lesser market focus.
- Month March, April, May are the top Marketing Spend months as shown by Table 3

Table 3 : Top 10 Marketing Spend by Month

Marketing Spend	**Month**
1.095438e+09	Mar-23
1.015562e+09	Mar-24
7.948044e+08	Apr-23
7.014837e+08	May-23
5.924961e+08	May-24
5.660023e+08	Apr-24
3.527270e+08	Jun-24
2.914296e+08	Jun-23
2.841590e+08	Feb-23
2.189398e+08	Apr-25

W-o-W Analysis for Marketing Spend



Fig(h) : W-o-W trend for Marketing spend

- The results are similar to monthly results.
- Top weeks are in March, April which clearly reflects the seasonality of Marketing spend related to summer and increased market.

Table 4: Top 10 weeks Marketing Spend

No	Year	Week Start	Spend
0	2022	2022-12-26	2.639629e+05
1	2023	2023-03-13	3.185481e+08
2	2023	2023-03-20	2.380691e+08
3	2023	2023-04-10	2.372676e+08
4	2024	2024-03-04	2.553441e+08
5	2024	2024-03-18	2.518962e+08
6	2024	2024-03-11	2.467319e+08

7	2025	2025-04-14	6.738512e+07
8	2025	2025-03-17	6.249076e+07
9	2025	2025-04-21	5.851970e+07

Part B – Diagnostic Analytics Objective:

Uncover the underlying drivers behind observed performance fluctuations.

1. Overall Demand Growth Analysis

- There is a **clear MoM and WoW increase in order volumes** from January through April/May every year.
- **Peak weeks and months** for orders coincide with **summer onset**, confirming that temperature is a strong driver.
- The year-wise growth in orders is there as 4.8 Cr in 2023 to 5.6 Cr in 2024 and is 2.3 Cr till Mid-May which indicates **growth**

2. Seasonal Trends

- Demand for AC servicing **peaks consistently in March, April, and May** each year.
- This trend **repeats weekly** with clear spikes seen in early to mid-April.
- **Post-May decline** suggests demand is **highly temperature-driven**, tapering off as monsoon arrives.

3. Marketing Spend vs. Orders

- Marketing spend also shows **sharp peaks in March–May**, aligning with peak demand.
- This indicates a **conscious strategy** to concentrate spend during high-return months.

- However, **2025 shows lower marketing spend** (~₹47 Cr till May) compared to previous years (₹335 Cr in 2023, ₹280 Cr in 2024).

Insight:

- Despite reduced spend in 2025, **orders are still strong**, indicating potential **spend efficiency** or better organic demand capture.

4. City-wise Marketing Budget Trends

- Cities like **Delhi, Mumbai, Hyderabad, Pune, and Gurgaon** consistently receive higher spend.
- **Bangalore and Chennai** have lower spends, suggesting:
  - Lower demand,
  - More organic traffic,
  - Or strategic deprioritization.

Insight:

- Top-spending cities are **also the top in orders**, implying a **positive ROI** and strategic targeting.

5. Budget Efficiency & Strategic Implications

- The fact that spend is **heavily skewed to Q1/Q2** and **drops to near-zero post-May** implies:
  - Campaigns are **time-bound** and **ROI-driven**.
  - Marketing is likely **trigger-based**, i.e., aligned with weather signals.

6. Investigating 2025 Spend Drop

Despite comparable temperature and demand trends, **2025 spend is significantly lower**.

Possible Explanations:

- **Budget cuts**
- **Operational constraints** (vendor issues, supply limitations)
- **City saturation or market maturity**
- **Efficiency strategy** — perhaps testing **lower spend with same demand**

7. Weather Impact on Demand

- Clear correlation between **rising temperatures and increase in orders**.
- Humidity shows **inverse or weak correlation**.
- Precipitation appears to **suppress demand** slightly, especially post-June.

Insight:

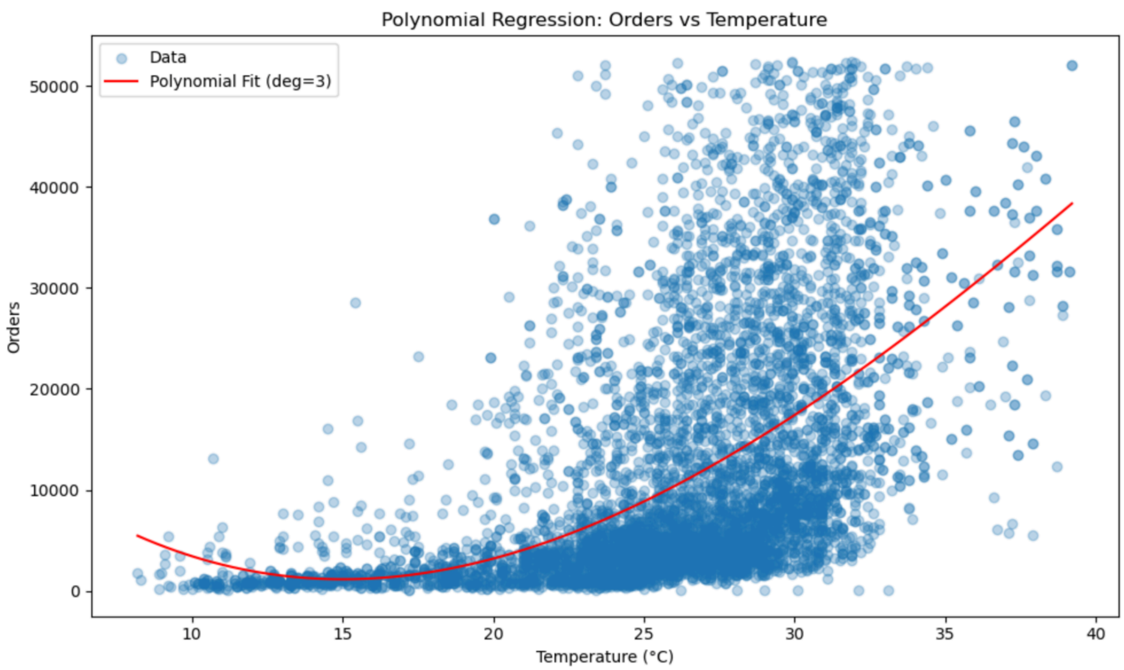
- **Temperature** can serve as a **leading indicator** to trigger marketing campaigns.

---

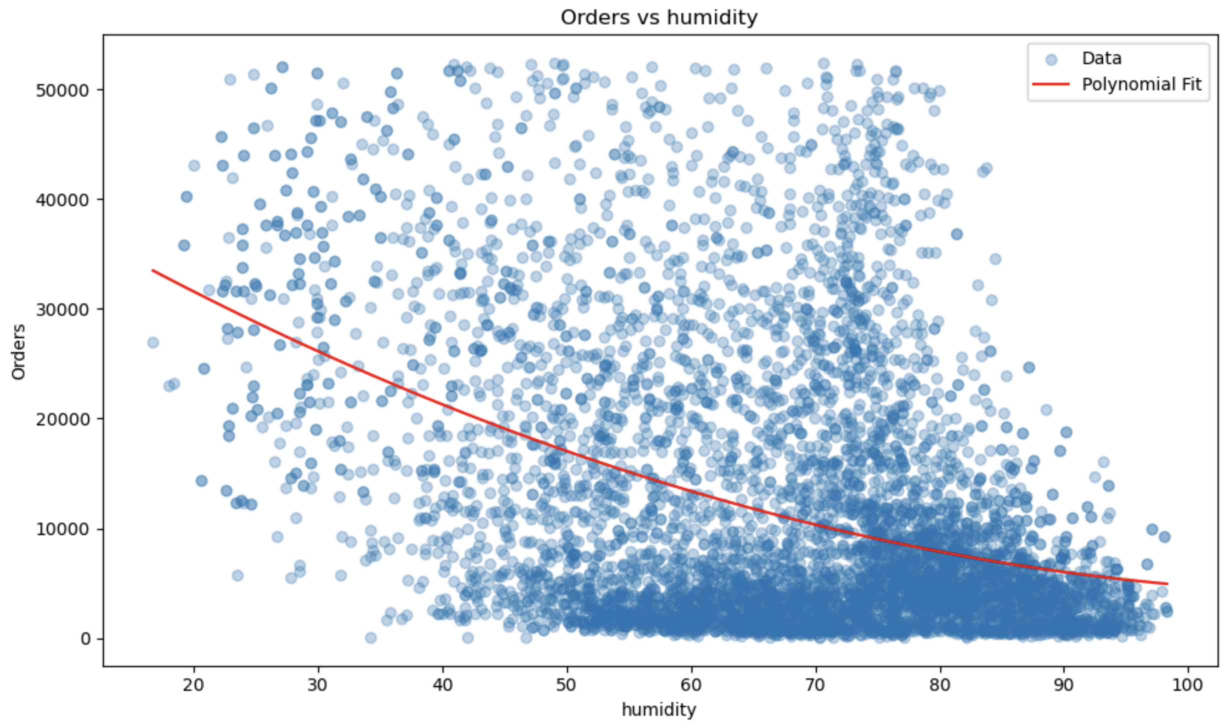
8. Charts and Mathematical Evidence for Correlation

- Monthly and Weekly Line plots of temperature and orders confirm seasonal behavior.
- Weekly stacked plots show **city-wise marketing alignment with order trends**

A. T **temperature and Orders** have a positive correlation

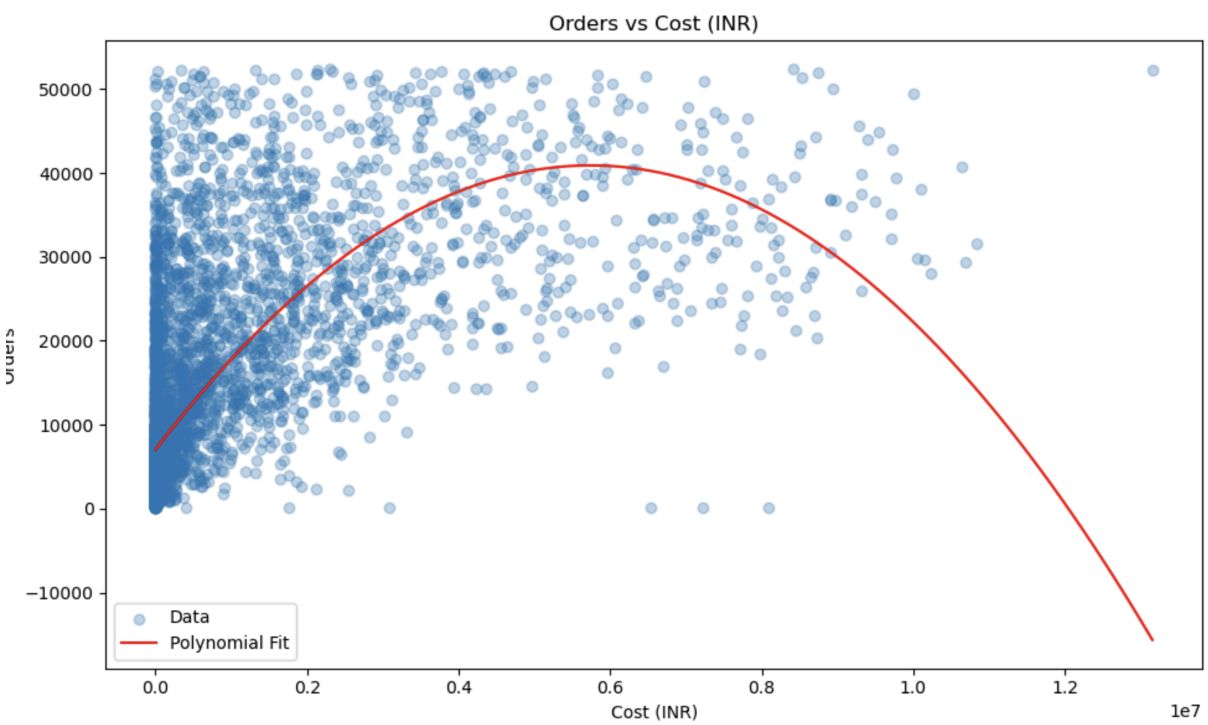


B. **Order and Humidity** have a **inverse correlation of -0.498**



📈 Correlation between humidity and orders: −0.408

C. Order and Marketing spend have positive correlation of 0.498



📈 Correlation between Cost (INR) and orders: 0.620

Part C – Predictive Analytics Objective

Overview(Jupyter Notebook)

Modeling Begins: `Cell [667]`

Exploratory Data Analysis (EDA): `Cell [1] to [666]`

This section documents the predictive modeling efforts for AC servicing demand and marketing cost estimation, using weather and temporal data.

Modeling Workflow

Step 1: Predict Marketing Spend

Model: **Random Forest Regressor**

- Predict marketing spend using weather + date-based features.
- Null marketing values imputed with **0**, assuming **no campaign activity**.

Step 2: Predict Future Orders

Model: **XGBoost Regressor**

- Used **predicted marketing spend + weather features** to forecast **order volume**.

Data Preparation

Null Value Handling

- `Cost (INR)` (Marketing Spend): Replaced nulls with `0`.
- Minor nulls in other columns handled with drop or forward-fill as appropriate.

Features Used

Category	Features
City	`city`
Weather	`temp`, `humidity`, `precip`

Date-Derived	`month`, `year`, `day`, `weekday`, `weekofyear`, `dayofyear`, `is_weekend`

- Instead of using raw date, engineered features like `month`, `weekday`, etc. capture **seasonality**, **trends**, and **cyclicality** in a more model-friendly way.

Feature Scaling

- All features were **standard scaled** (zero mean, unit variance).
- This ensures:
  - Equal contribution of all features
  - Faster convergence for gradient-based models
  - Eliminates magnitude bias in feature influence

Model Comparison & Selection

Models Tried:

- Linear Regression
- KMeans + Regression
- AdaBoost Regressor
- Random Forest Regressor
- XGBoost Regressor
- Decision Tree
- Gradient Boost

Best Models:

Task	Best Model
Marketing Spend	Random Forest
Order Prediction	XGBoost

Model 1: XGBoost Regressor (Predicting Orders)

Performance Metrics

Metric	Training Set	Test Set
R^2 Score	0.9388	**0.8961**
RMSE	7,059.26	8,666.99
MAE	774.96	3,165.14

Comment:

The XGBoost model performs excellently, explaining nearly 90% of the variance in the test set.
The low RMSE and MAE confirm strong predictive accuracy with minimal overfitting.
This model is robust and reliable for forecasting AC servicing demand and can be effectively used in capacity planning and business decision-making.

Model 2: Random Forest Regressor (Predicting Marketing Spend)

Performance Metrics

Metric	Training Set	Test Set
R^2 Score	0.8511	**0.7666**
RMSE	₹743,712.32	₹929,088.75
MAE	₹266,365.81	₹350,661.20

Comment:

The Random Forest model performs moderately well.
It captures ~77% of marketing spend variation, but with higher errors, especially on unseen data.
This suggests that external factors (like budget decisions or campaign-specific strategies) are not fully represented in the available features.

Added features, better data sourcing, tuning can be some of the ways to improve the model performance

Nonetheless, it serves as a **reasonable first-step estimator** for planning spend distribution.

Result

- **XGBoost is highly accurate for order forecasting**, demonstrating both precision and generalization.
- **Random Forest works for estimating marketing cost**, but has room for improvement, possibly through additional features like channel-wise spend or campaign tags.
- Both models align with **seasonal and weather-driven behavior** and can be deployed to guide **marketing strategy and operational planning**.