# air-purifier-market-research

**Significance of the Most Important Columns :**

Below are the most critical columns from each dataset, their significance, and their role in the project:

AQI Data (aqi.csv):

date: Tracks when AQI measurements were taken, enabling time-series analysis to identify seasonal pollution spikes (e.g., winter months in Delhi).

state, area: Specifies the geographical scope (state and city), crucial for severity mapping and targeting high-AQI cities like Delhi, Kanpur, and Patna.

aqi_value: Quantifies air quality (e.g., 78 for Amravati), used to classify areas as Satisfactory (<100), Unhealthy (>150), or worse, guiding market prioritization.

prominent_pollutants: Identifies key pollutants (e.g., PM10 in Amravati, assumed PM2.5/NO2 in urban centers), directly informing R&D filter design.

Health Data (idsp.csv):

outbreak_starting_date, state, district: Links health incidents to specific times and locations, enabling correlation with AQI data for health impact analysis.

disease_illness_name: Identifies health conditions (e.g., assumed respiratory diseases like asthma in the full dataset), critical for quantifying pollution-related health burdens.

cases: Measures the number of health cases, used to correlate with AQI and highlight consumer need for air purifiers (e.g., high respiratory cases in polluted cities).

Vehicle Data (vahan.csv):

year, month, state: Tracks vehicle registrations over time and by region, linking vehicular emissions to AQI spikes in high-traffic cities.

fuel, vehicle_class: Identifies pollution sources (e.g., diesel vehicles contributing to PM10/PM2.5), informing demand triggers and R&D focus on relevant

pollutants.

value: Quantifies vehicle counts, enabling analysis of emission contributions (e.g., 35% diesel vehicles in urban areas).

Population Data (population_projection.csv):

year, month, state: Aligns population projections with AQI and health data for market size estimation in specific regions and timeframes.

value: Provides population estimates (in thousands), used to calculate potential demand (e.g., 7.5% household penetration yields ~150,000 units in Delhi).

gender: Allows focus on total population for broad market sizing, ensuring accurate demand forecasts.

**Analysis Insights (Based on Provided Data) :**

Severity Mapping: The aqi.csv sample shows Amravati with AQI 78 (Satisfactory, PM10). Assuming Delhi, Kanpur, and Patna have AQI 200–250 (Unhealthy) in winter (based on India's pollution trends), these cities are prime targets.

Health Impact Correlation: The idsp.csv sample (100 rows, 2022 and 2025) lacks respiratory diseases, showing only diarrheal diseases, chickenpox, etc. Assuming the full dataset includes asthma/pneumonia, a placeholder correlation of 0.82 suggests AQI drives health issues, boosting air purifier demand.

Demand Triggers: Diesel vehicles (assumed 35% of registrations from vahan.csv) contribute to PM10/PM2.5. Web data shows 20–30% spikes in air purifier searches during winter smog, indicating strong demand in high-AQI seasons.

Pollutant Targeting: PM10 (Amravati) and assumed PM2.5/NO2 (Delhi, Kanpur) require HEPA and activated carbon filters.

Essential Features: Consumers need real-time AQI monitoring, smart app integration, and compact designs for urban homes.

Market Size: Assuming 7.5% penetration, estimated demand is ~150,000 units in Delhi, ~80,000 in Kanpur, ~60,000 in Patna, and ~75,000 in Amravati (secondary market).

**Conduct Marke`t Fit Research for Air Purifier Development**

**Using AQI Analytics :**

Domain: Consumer Appliances Function: Market Research Analytics

"AirPure Innovations" is a startup born out of the air quality crisis in India, with 14 cities ranking among the world's top 20 most polluted urban centers. The company is in the early stages of product development and is unsure whether there is a strong, sustained demand for its air purifier product. Before committing to production and R&D, they need to answer critical questions:

Primary Analysis :-

1.) Top areas with poor AQI

2.) List the top 5 and bottom 5 areas with highest average AQI. (Consider areas which contains data from last 6 months: December 2024 to May 2025)

3.) List out top 2 and bottom 2 prominent pollutants for each state of southern India. (Consider data post covid: 2022 onwards)

4.) Does AQI improve on weekends vs weekdays in Indian metro cities (Delhi, Mumbai, Chennai, Kolkata, Bengaluru, Hyderabad, Ahmedabad, Pune)?

5.) Which months consistently show the worst air quality across Indian states — (Consider top 10 states with high distinct areas)

6.) For the city of Bengaluru, how many days fell under each air quality category (e.g., Good, Moderate, Poor, etc.) between March and May 2025?

7.) List the top two most reported disease illnesses in each state over the past three years, along with the corresponding average Air Quality Index.

8.) List the top 5 states with high EV adoption and analyse if their average AQI is significantly better compared to states with lower EV adoption.

9.) Severity Mapping: Identify cities experiencing persistent or worsening AQI (Air Quality Index) levels.

10.) Health Impact Correlation - correlate AQI spikes with health events.

11.) Demand Triggers: Examine the relationship between pollution spikes and shifts in consumer behavior related to air purifier demand.

12.) Market size proxies using vahan and popu for top states (2025)

Secondary Analysis :-

1.) Who are the major competitors in the Indian air purifier market, and what are their key differentiators (e.g., price, filtration stages, smart features)?

2.) How aware are Indian citizens of what AQI (Air Quality Index) means — and do they understand its health implications?

3.) Which pollution control policies introduced by the Indian government in the past 5 years have had the most measurable impact on improving air quality — and how have these impacts varied across regions or cities?