NAME: ARPAN CHATTERJEE

Roll: 2312RES162

Adm. No: IITP000646

4th SEM BSc. CSDA

GUVI + HCL SUMMER INDUSTRY INTERNSHIP

Email-id: arpan_2312res162@iitp.ac.in

GitHub Link: https://github.com/ArpanC03/Guvi-Project-1-Data-Science

AQI Data

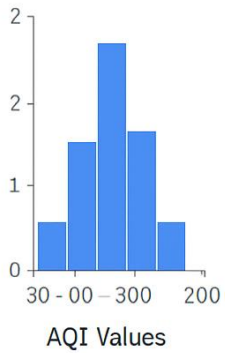
City	State	AQI
Delhi	Delhi	252
Mumbai	Maharashtra	191
Bengaluru	Kamafaka	132
Kalkata	West Bengal	146
Chennai	Tamil Nadu	125
Hyderabad	Telangana	153
Anmedabad	Gujarat	129

1 / 10 >

Dashboard

AQI Scatter Map

AQI Histogram



AQI Distribution

DATA SCIENCE INTERN

Problem Statement 5:

- AQI Mapping of Indian Cities
- Mapping of AQI readings using geospatial visualizations.
- Choropleth map by AQI
- Tooltip: City name + AQI
- AQI filter: Good, Moderate, Poor

DESCRIPTION

Air pollution in Indian cities is a major environmental and health concern, but city-level AQI data is complex, scattered, and difficult for citizens or policymakers to interpret quickly.

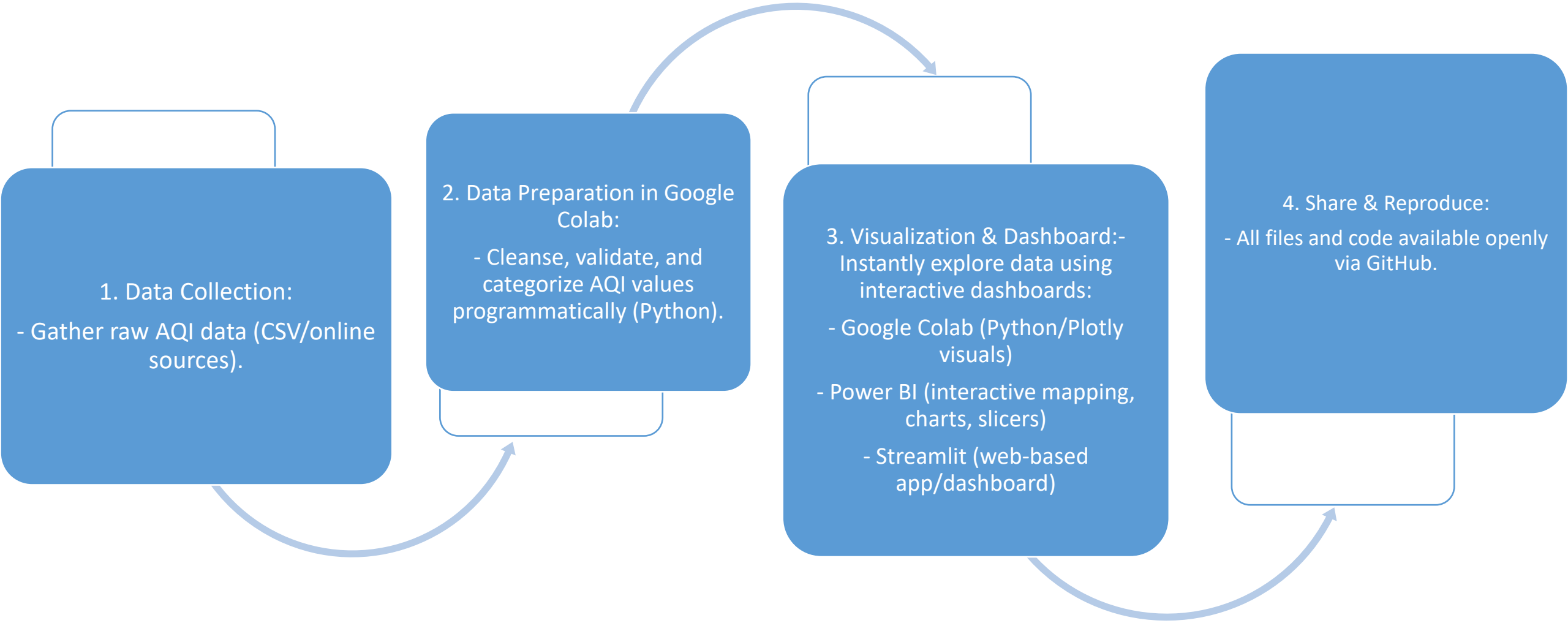
Our goal: Build an interactive dashboard that streamlines exploration, visualization, and analysis of real AQI data for all major Indian cities, to empower evidence-driven action and awareness.

OBJECTIVES

- Aggregate, clean, and categorize AQI data for Indian cities, ensuring consistency and accuracy.
- Enable easy visual exploration of air quality by city, state, and AQI category.
- Provide actionable insights—top/bottom ranked cities, category-wise breakdowns, and distribution analysis—for diverse stakeholders.
- Share the project openly as a reproducible workflow and dashboard.

DATA PIPELINE AND WORKFLOW

1. Data Collection:
   - Gather raw AQI data (CSV/online sources).
2. Data Preparation in Google Colab:
   - Cleanse, validate, and categorize AQI values programmatically (Python).
3. Visualization & Dashboard:
   - Instantly explore data using interactive dashboards:
     - Google Colab (Python/Plotly visuals)
     - Power BI (interactive mapping, charts, slicers)
     - Streamlit (web-based app/dashboard)
4. Share & Reproduce:
   - All files and code available openly via GitHub.


1. Data Collection:
- Gather raw AQI data (CSV/online sources).
2. Data Preparation in Google Colab:
- Cleanse, validate, and categorize AQI values programmatically (Python).
3. Visualization & Dashboard:- Instantly explore data using interactive dashboards:
- Google Colab (Python/Plotly visuals)
- Power BI (interactive mapping, charts, slicers)
- Streamlit (web-based app/dashboard)
4. Share & Reproduce:
- All files and code available openly via GitHub.

KEY FEATURES

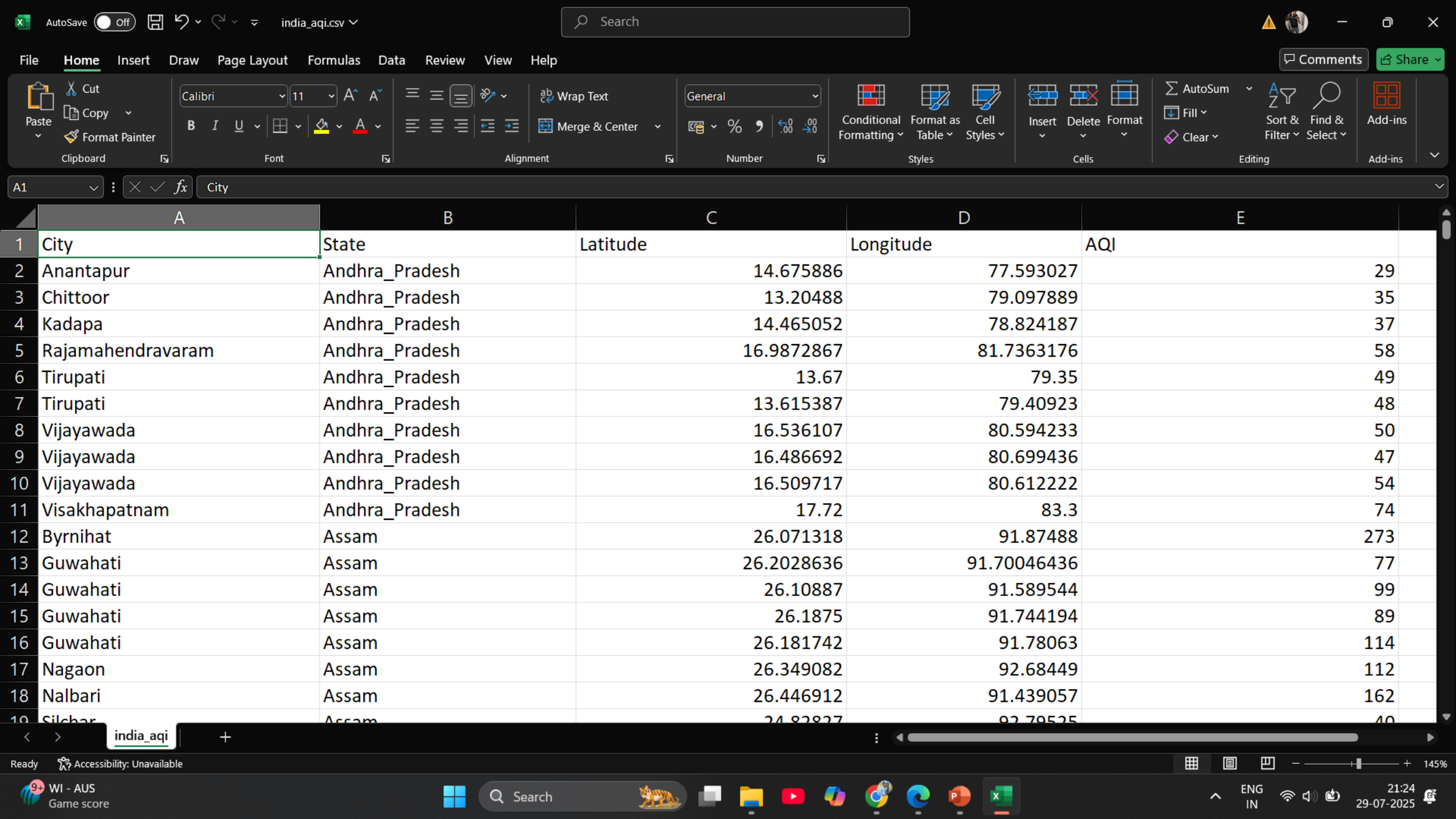
- Scrollable, filterable city-state AQI table for instant lookup.
- Dynamic KPIs: total cities, average AQI, most/least polluted city at a glance.
- Interactive map of India with color-coded AQI markers.
- Distribution plots, city ranking charts, and categorical breakdowns.
- One-click reproducibility: all dashboards work from one dataset.

data_aqi_cpcb.xml

File C:/Users/hp/Desktop/ARPAN/PROFESSIONAL/Intern%20Works/guvi%20HCL/data_aqi_cpcb.xml

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
<AqIndex>
  <Country id="India">
    <State id="Andhra_Pradesh">
      <City id="Amaravati">
        <Station id="Secretariat, Amaravati - APPCB" lastupdate="20-11-2023 12:00:00" latitude="16.5150833" longitude="80.5181667">
          <Pollutant_Index id="PM2.5" Min="NA" Max="NA" Avg="NA"/>
          <Pollutant_Index id="PM10" Min="NA" Max="NA" Avg="NA"/>
          <Pollutant_Index id="SO2" Min="NA" Max="NA" Avg="NA"/>
          <Pollutant_Index id="CO" Min="NA" Max="NA" Avg="NA"/>
          <Pollutant_Index id="OZONE" Min="NA" Max="NA" Avg="NA"/>
        </Station>
      </City>
      <City id="Anantapur">
        <Station id="Gulzarpet, Anantapur - APPCB" lastupdate="20-11-2023 12:00:00" latitude="14.675886" longitude="77.593027">
          <Pollutant_Index id="PM2.5" Min="15" Max="34" Avg="24"/>
          <Pollutant_Index id="PM10" Min="16" Max="39" Avg="24"/>
          <Pollutant_Index id="NO2" Min="17" Max="39" Avg="23"/>
          <Pollutant_Index id="NH3" Min="2" Max="3" Avg="3"/>
          <Pollutant_Index id="SO2" Min="4" Max="17" Avg="7"/>
          <Pollutant_Index id="CO" Min="20" Max="40" Avg="29"/>
          <Pollutant_Index id="OZONE" Min="16" Max="52" Avg="25"/>
          <Air_Quality_Index Value="29" Predominant_Parameter="CO"/>
        </Station>
      </City>
      <City id="Chittoor">
        <Station id="Gangineni Cheruvu, Chittoor - APPCB" lastupdate="20-11-2023 12:00:00" latitude="13.204880" longitude="79.097889">
          <Pollutant_Index id="PM2.5" Min="11" Max="30" Avg="19"/>
          <Pollutant_Index id="PM10" Min="13" Max="35" Avg="22"/>
          <Pollutant_Index id="NO2" Min="8" Max="21" Avg="14"/>
          <Pollutant_Index id="NH3" Min="1" Max="2" Avg="2"/>
          <Pollutant_Index id="SO2" Min="4" Max="18" Avg="9"/>
          <Pollutant_Index id="CO" Min="2" Max="5" Avg="4"/>
          <Pollutant_Index id="OZONE" Min="23" Max="92" Avg="35"/>
          <Air_Quality_Index Value="35" Predominant_Parameter="OZONE"/>
        </Station>
      </City>
      <City id="Kadapa">
        <Station id="Yerramukkapalli, Kadapa - APPCB" lastupdate="20-11-2023 12:00:00" latitude="14.465052" longitude="78.824187">
          <Pollutant_Index id="PM2.5" Min="19" Max="57" Avg="37"/>
          <Pollutant_Index id="PM10" Min="18" Max="48" Avg="32"/>
          <Pollutant_Index id="NO2" Min="21" Max="54" Avg="30"/>
          <Pollutant_Index id="NH3" Min="1" Max="3" Avg="1"/>
          <Pollutant_Index id="SO2" Min="1" Max="2" Avg="2"/>
          <Pollutant_Index id="CO" Min="12" Max="30" Avg="21"/>
          <Pollutant_Index id="OZONE" Min="6" Max="48" Avg="21"/>
```



AutoSave Off india_aqi.csv Search

File Home Insert Draw Page Layout Formulas Data Review View Help Comments Share

Paste Cut Copy Format Painter Clipboard | Calibri 11 Font | Wrap Text Merge & Center Alignment | General Number | Conditional Formatting Format as Table Cell Styles Styles | Insert Delete Format Cells | AutoSum Fill Clear Sort & Filter Find & Select Editing | Add-ins Add-ins

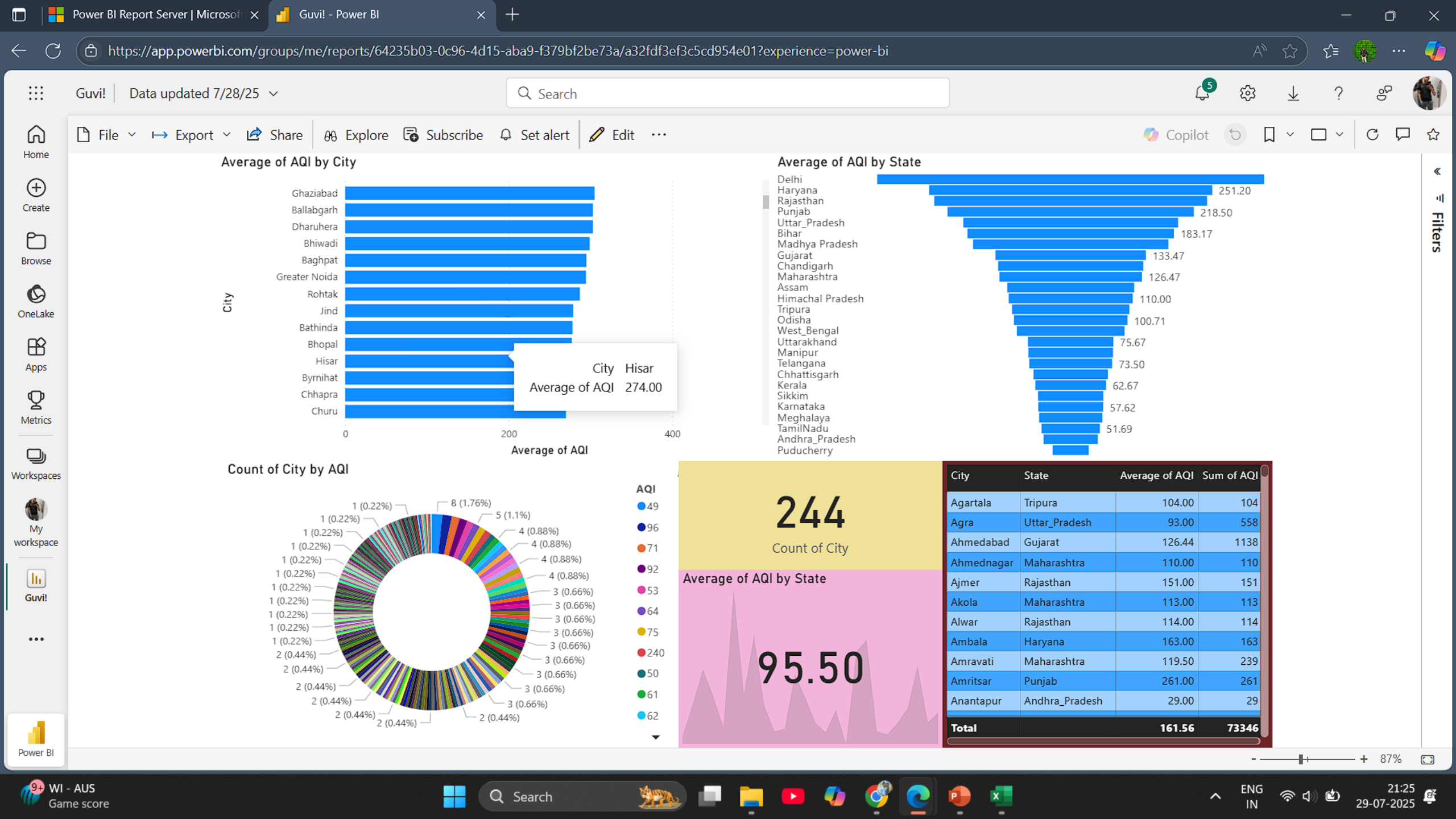
A1 City

	A	B	C	D	E
1	City	State	Latitude	Longitude	AQI
2	Anantapur	Andhra_Pradesh	14.675886	77.593027	29
3	Chittoor	Andhra_Pradesh	13.20488	79.097889	35
4	Kadapa	Andhra_Pradesh	14.465052	78.824187	37
5	Rajamahendravaram	Andhra_Pradesh	16.9872867	81.7363176	58
6	Tirupati	Andhra_Pradesh	13.67	79.35	49
7	Tirupati	Andhra_Pradesh	13.615387	79.40923	48
8	Vijayawada	Andhra_Pradesh	16.536107	80.594233	50
9	Vijayawada	Andhra_Pradesh	16.486692	80.699436	47
10	Vijayawada	Andhra_Pradesh	16.509717	80.612222	54
11	Visakhapatnam	Andhra_Pradesh	17.72	83.3	74
12	Byrnihat	Assam	26.071318	91.87488	273
13	Guwahati	Assam	26.2028636	91.70046436	77
14	Guwahati	Assam	26.10887	91.589544	99
15	Guwahati	Assam	26.1875	91.744194	89
16	Guwahati	Assam	26.181742	91.78063	114
17	Nagaon	Assam	26.349082	92.68449	112
18	Nalbari	Assam	26.446912	91.439057	162
19	Silchar	Assam	24.82827	92.79525	40

india_aqi

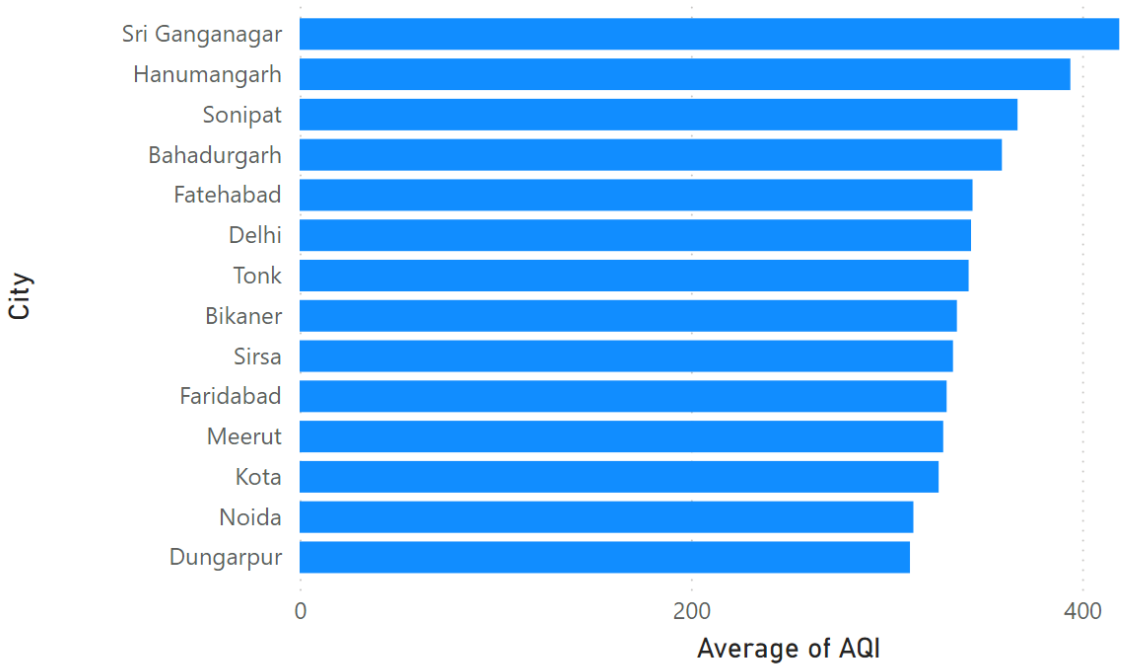
Ready Accessibility: Unavailable 145%

WI - AUS Game score Search ENG IN 21:24 29-07-2025

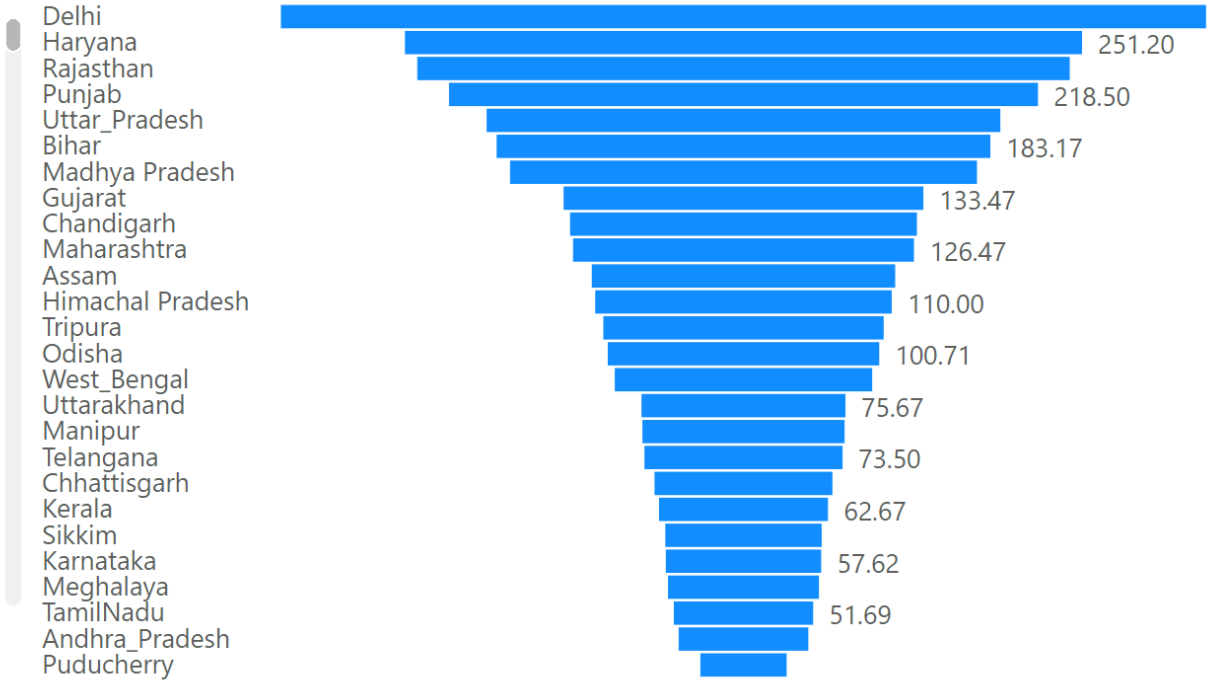


City	State	Average of AQI	Sum of AQI
Agartala	Tripura	104.00	104
Agra	Uttar_Pradesh	93.00	558
Ahmedabad	Gujarat	126.44	1138
Ahmednagar	Maharashtra	110.00	110
Ajmer	Rajasthan	151.00	151
Akola	Maharashtra	113.00	113
Alwar	Rajasthan	114.00	114
Ambala	Haryana	163.00	163
Amravati	Maharashtra	119.50	239
Amritsar	Punjab	261.00	261
Anantapur	Andhra_Pradesh	29.00	29
Total		**161.56**	**73346**

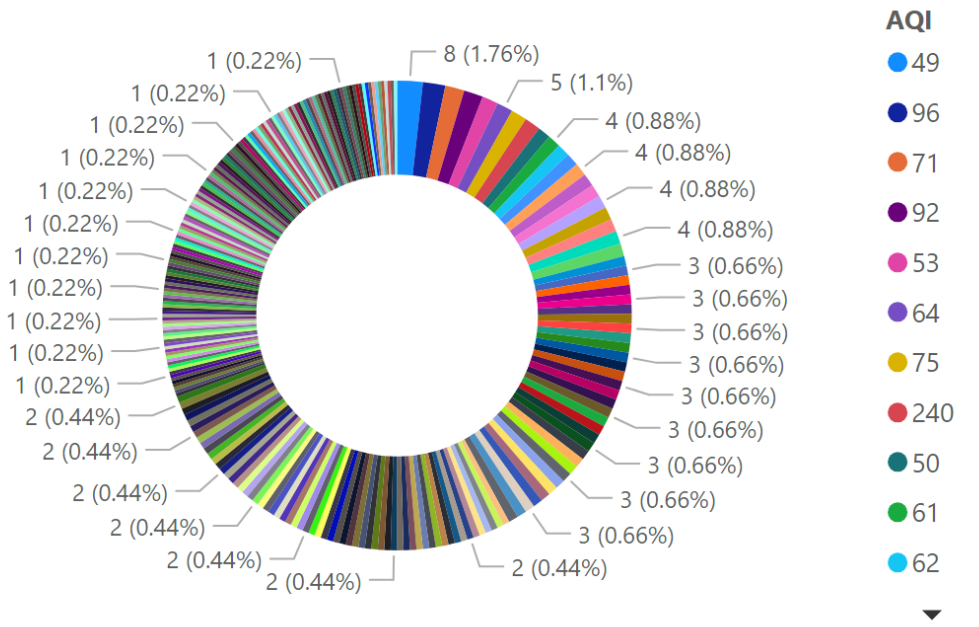
Average of AQI by City



Average of AQI by State



Count of City by AQI



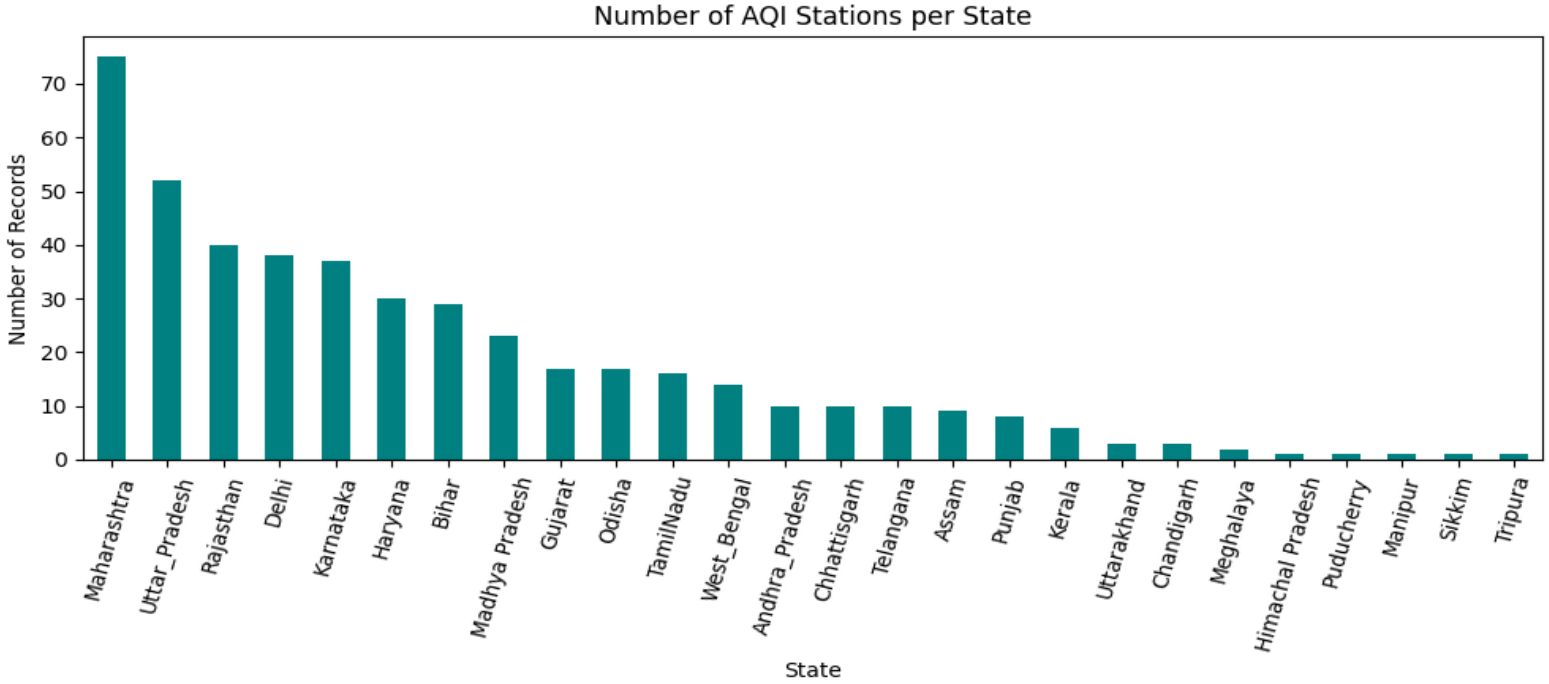
244

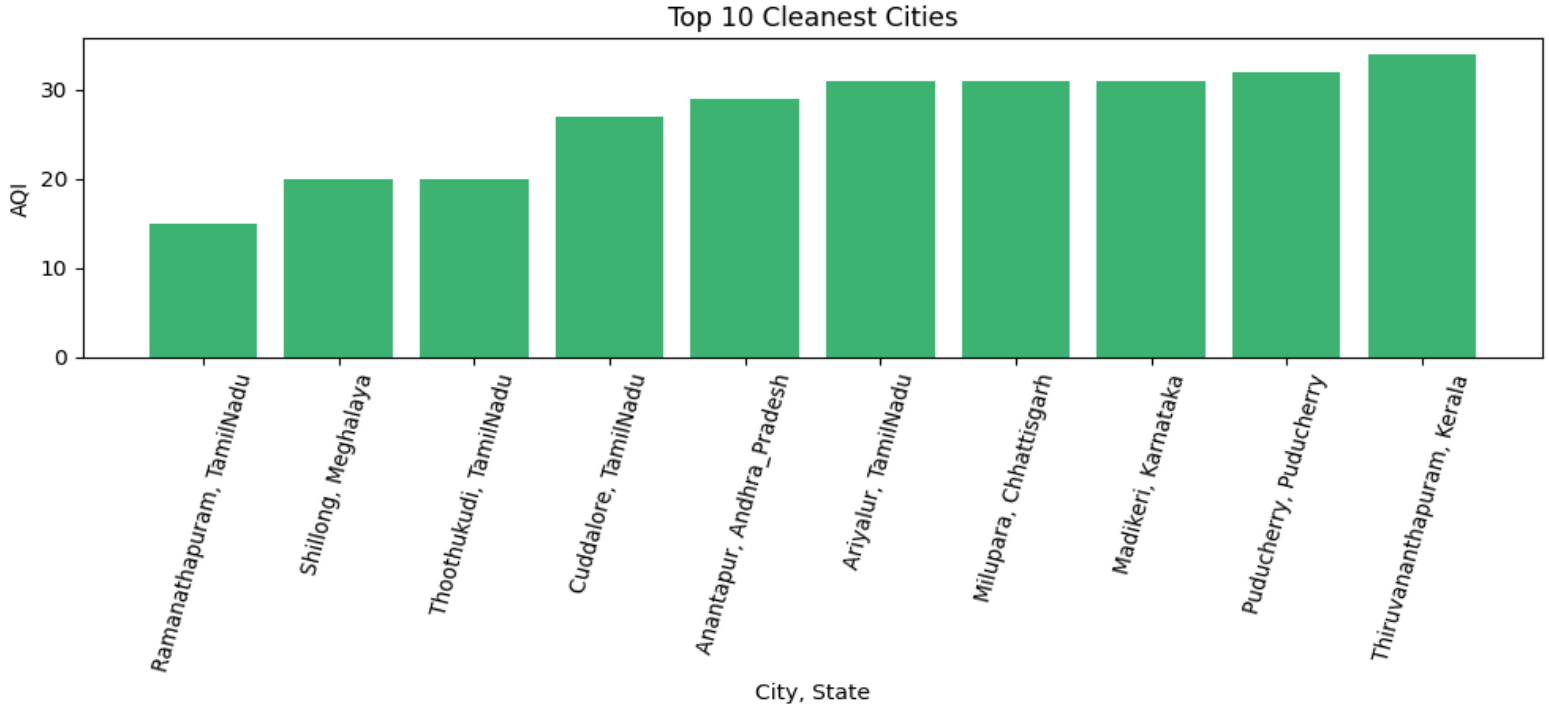
Count of City

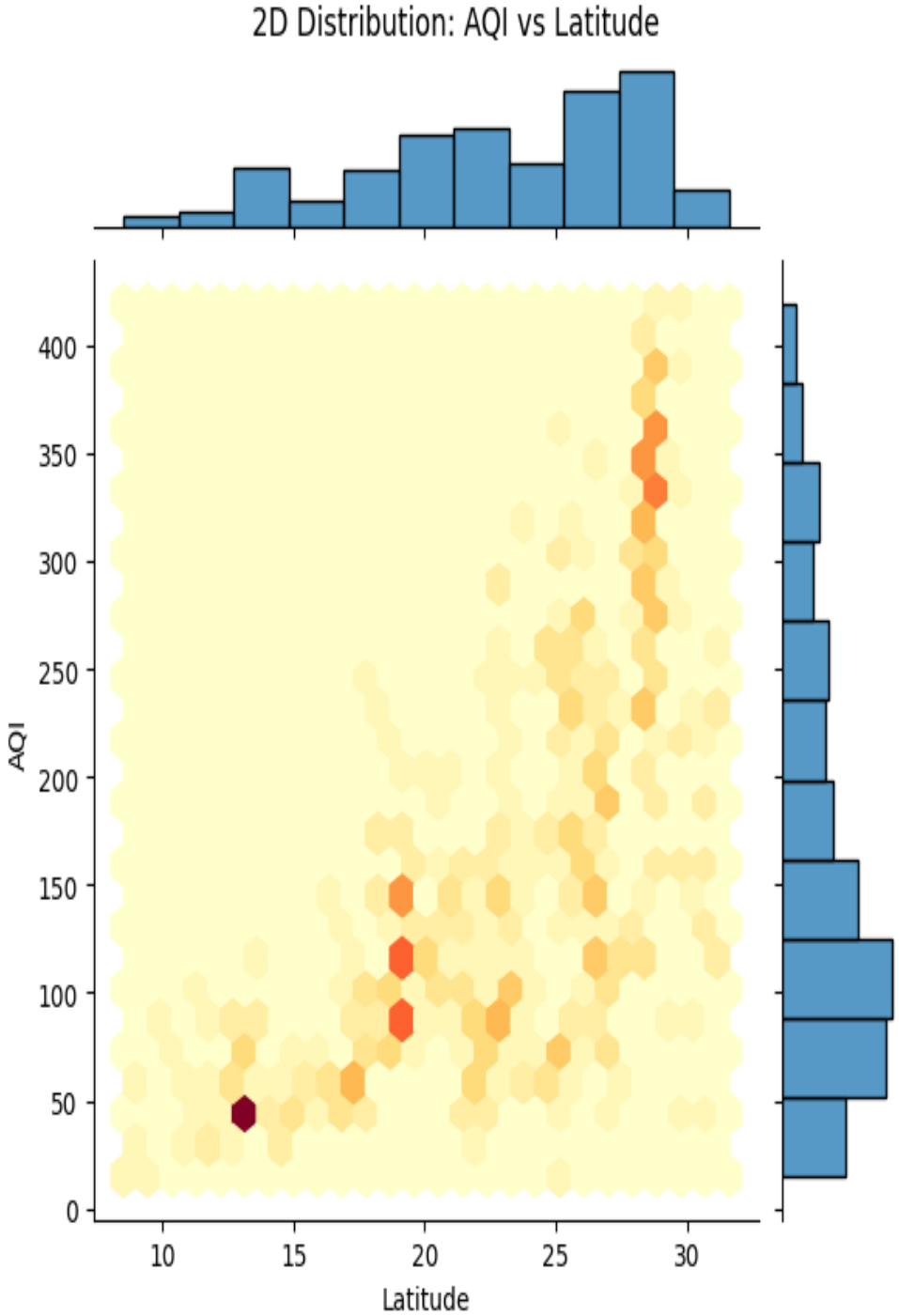
Average of AQI by State

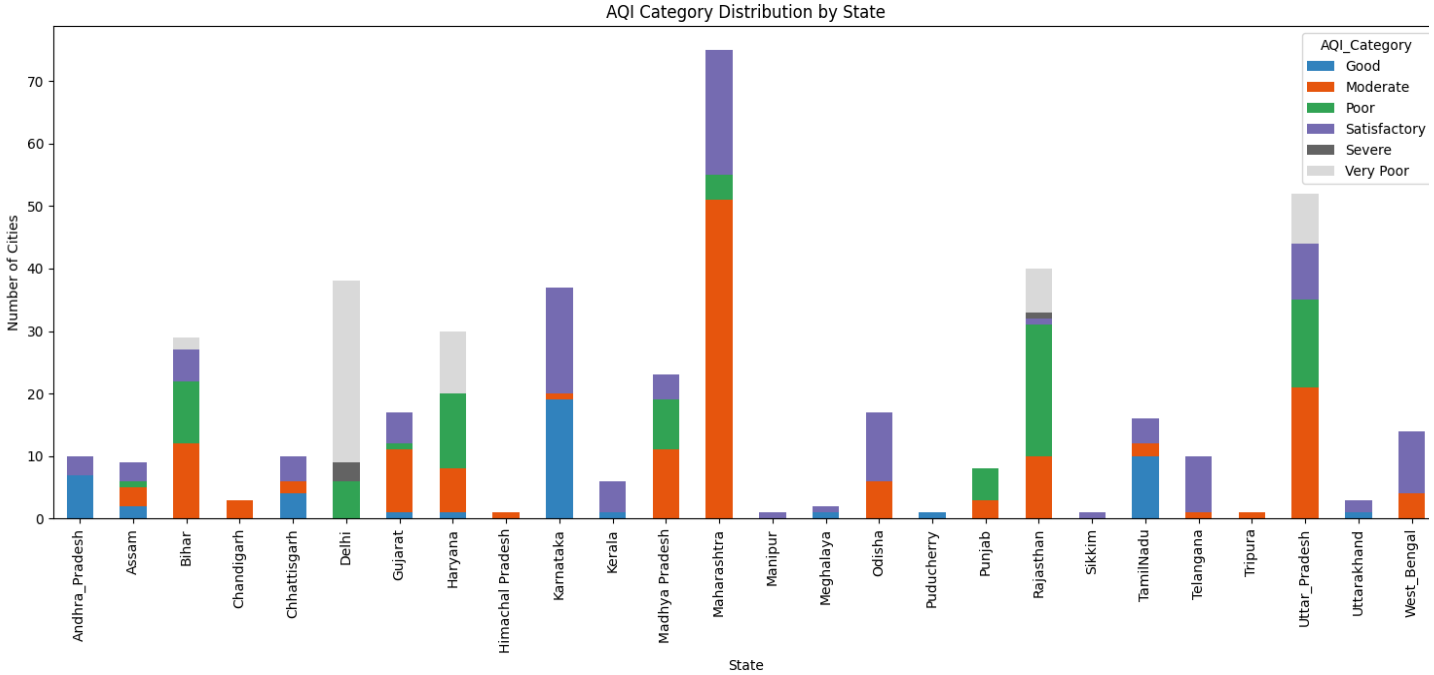
95.50

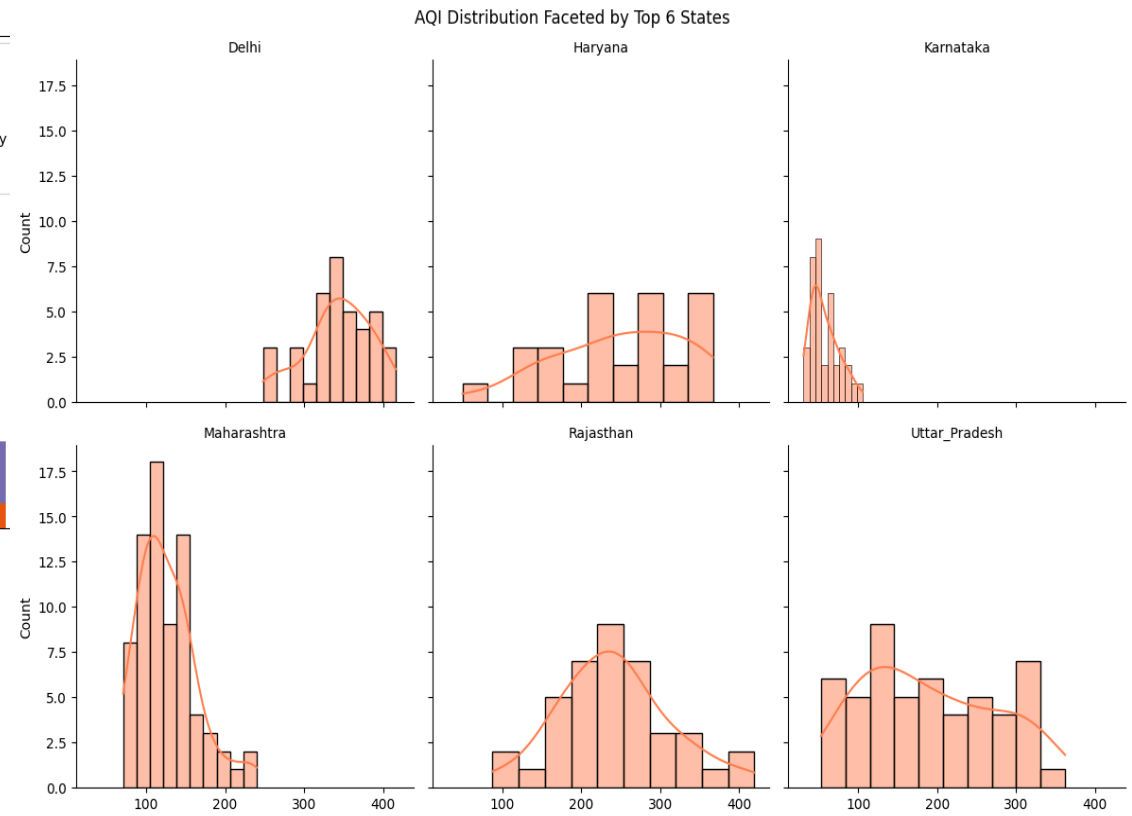
City	State	Average of AQI	Sum of AQI
Agartala	Tripura	104.00	104
Agra	Uttar_Pradesh	93.00	558
Ahmedabad	Gujarat	126.44	1138
Ahmednagar	Maharashtra	110.00	110
Ajmer	Rajasthan	151.00	151
Akola	Maharashtra	113.00	113
Alwar	Rajasthan	114.00	114
Ambala	Haryana	163.00	163
Amravati	Maharashtra	119.50	239
Amritsar	Punjab	261.00	261
Anantapur	Andhra_Pradesh	29.00	29
Total		**161.56**	**73346**

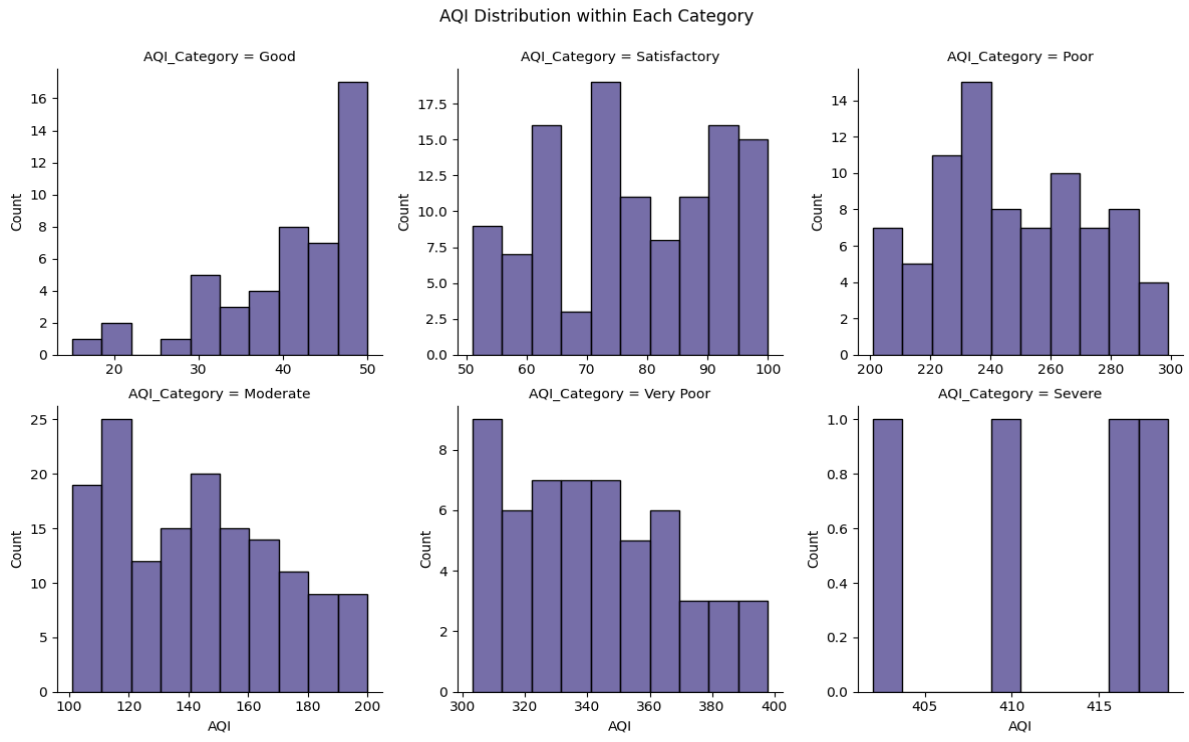

Number of AQI Stations per State
Number of Records
0
10
20
30
40
50
60
70
Maharashtra
Uttar_Pradesh
Rajasthan
Delhi
Karnataka
Haryana
Bihar
Madhya Pradesh
Gujarat
Odisha
TamilNadu
West_Bengal
Andhra_Pradesh
Chhattisgarh
Telangana
Assam
Punjab
Kerala
Uttarakhand
Chandigarh
Meghalaya
Himachal Pradesh
Puducherry
Manipur
Sikkim
Tripura
State

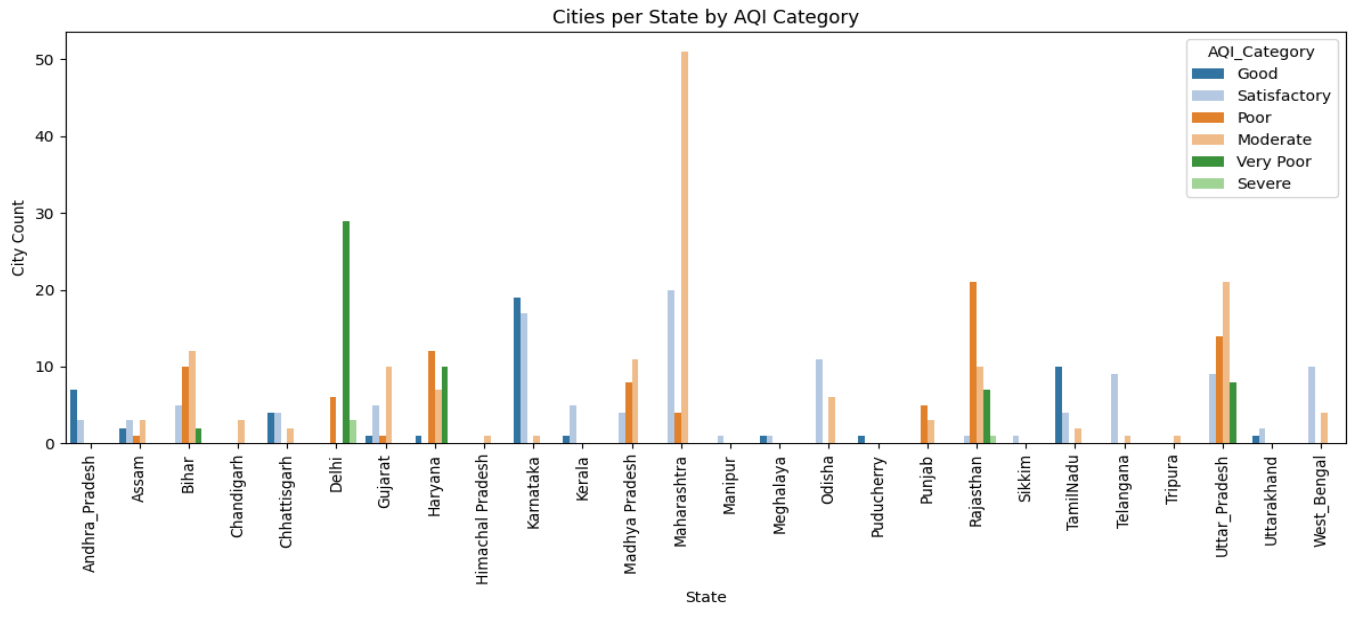

Top 10 Cleanest Cities
AQI
0
10
20
30
Ramanathapuram, TamilNadu
Shillong, Meghalaya
Thoothukudi, TamilNadu
Cuddalore, TamilNadu
Anantapur, Andhra_Pradesh
Ariyalur, TamilNadu
Milupara, Chhattisgarh
Madikeri, Karnataka
Puducherry, Puducherry
Thiruvananthapuram, Kerala
City, State

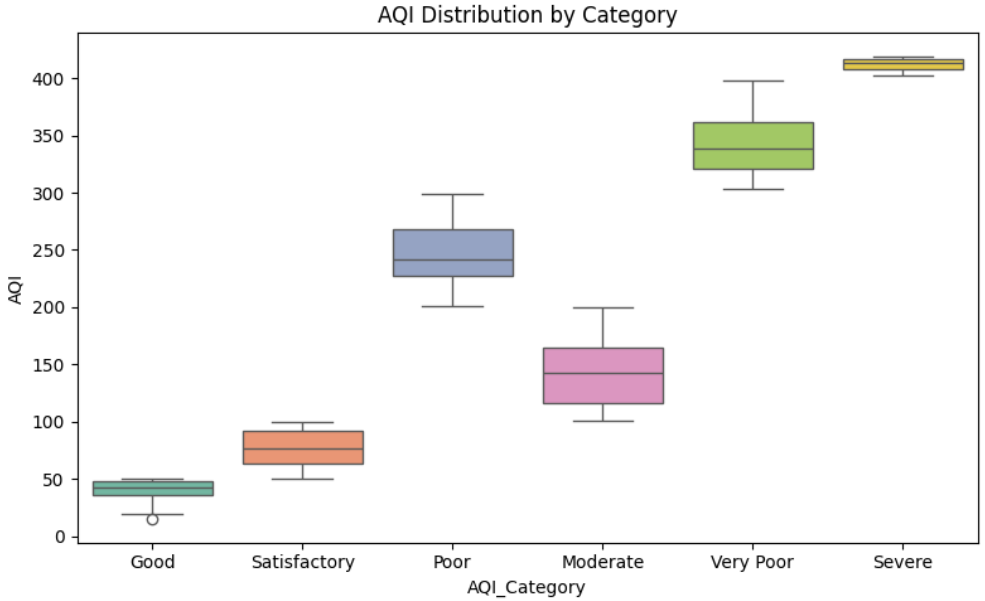

2D Distribution: AQI vs Latitude
AQI
0
50
100
150
200
250
300
350
400
10
15
20
25
30
Latitude

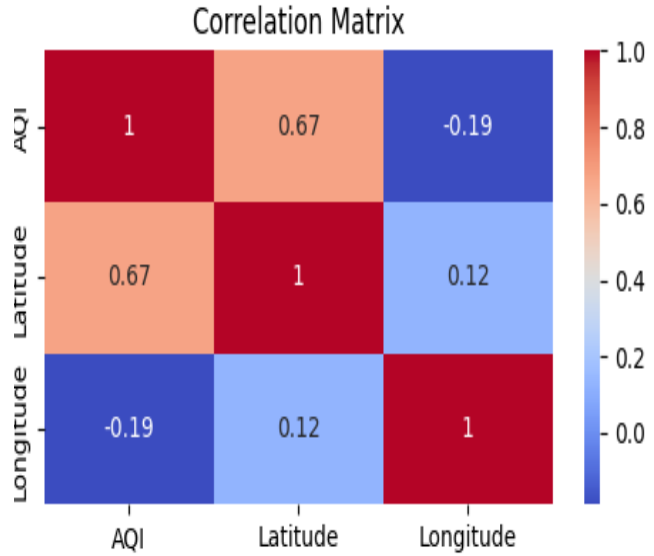

AQI Category Distribution by State
AQI_Category
Good
Moderate
Poor
Satisfactory
Severe
Very Poor
Number of Cities
State
Andhra_Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Delhi
Gujarat
Haryana
Himachal Pradesh
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
TamilNadu
Telangana
Tripura
Uttar_Pradesh
Uttarakhand
West_Bengal

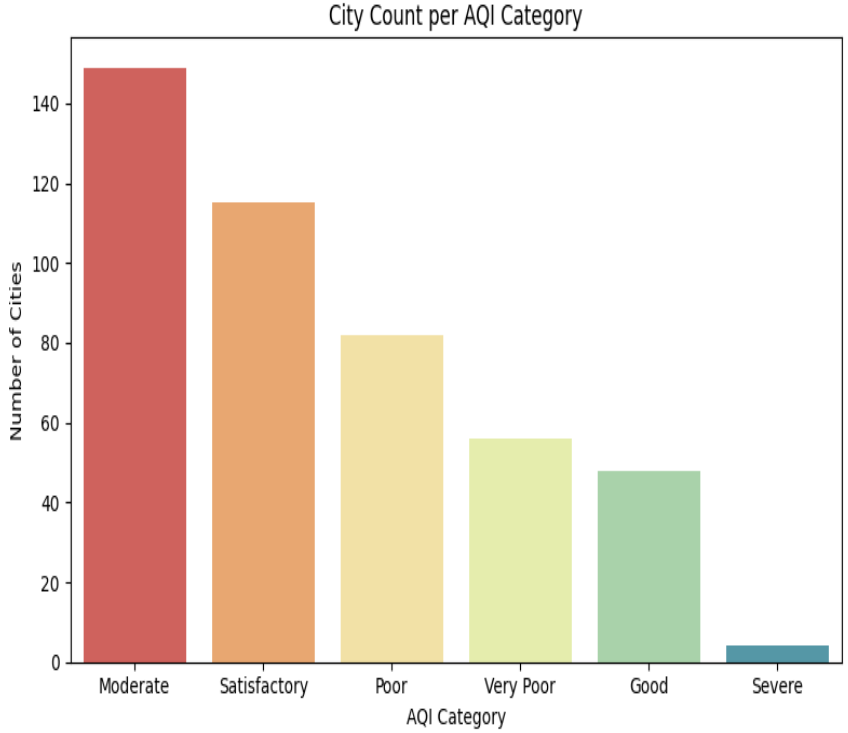

AQI Distribution Faceted by Top 6 States
Delhi
Haryana
Karnataka
Maharashtra
Rajasthan
Uttar_Pradesh
Count

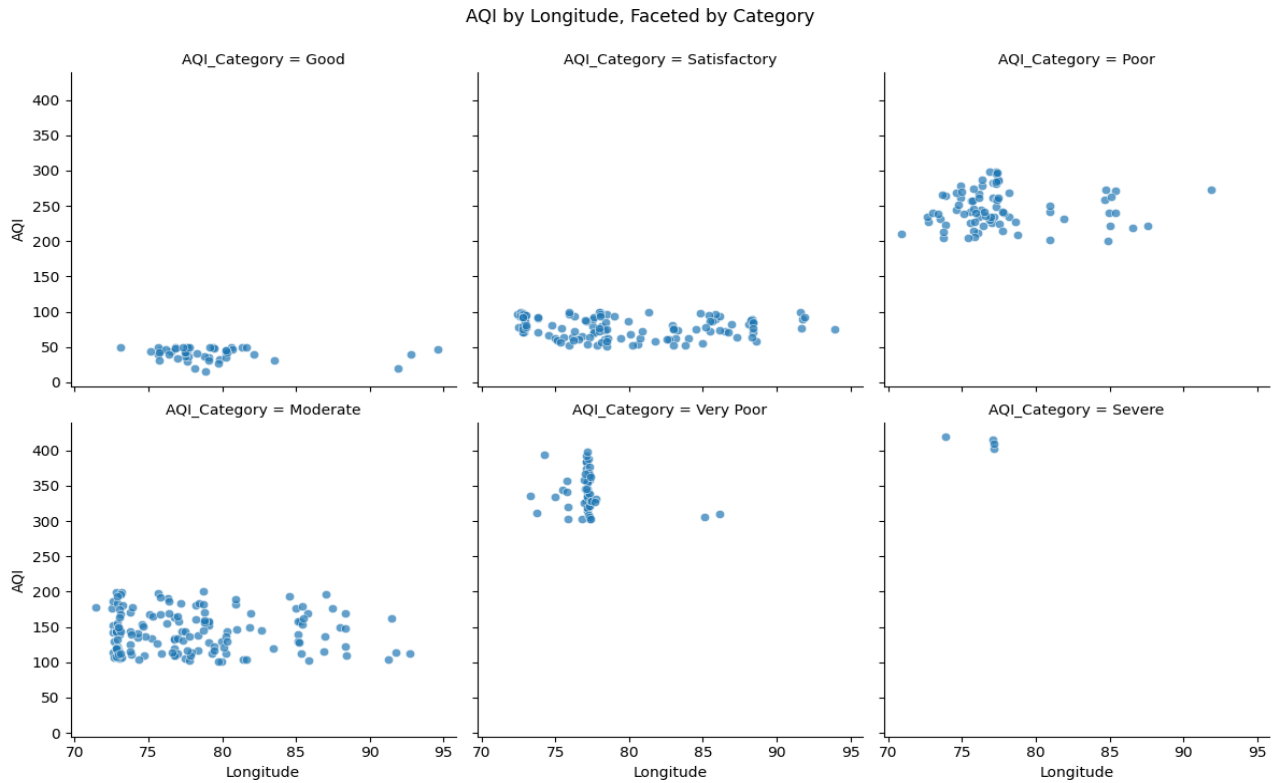

AQI Distribution within Each Category
AQI_Category = Good
AQI_Category = Satisfactory
AQI_Category = Poor
AQI_Category = Moderate
AQI_Category = Very Poor
AQI_Category = Severe
Count
AQI

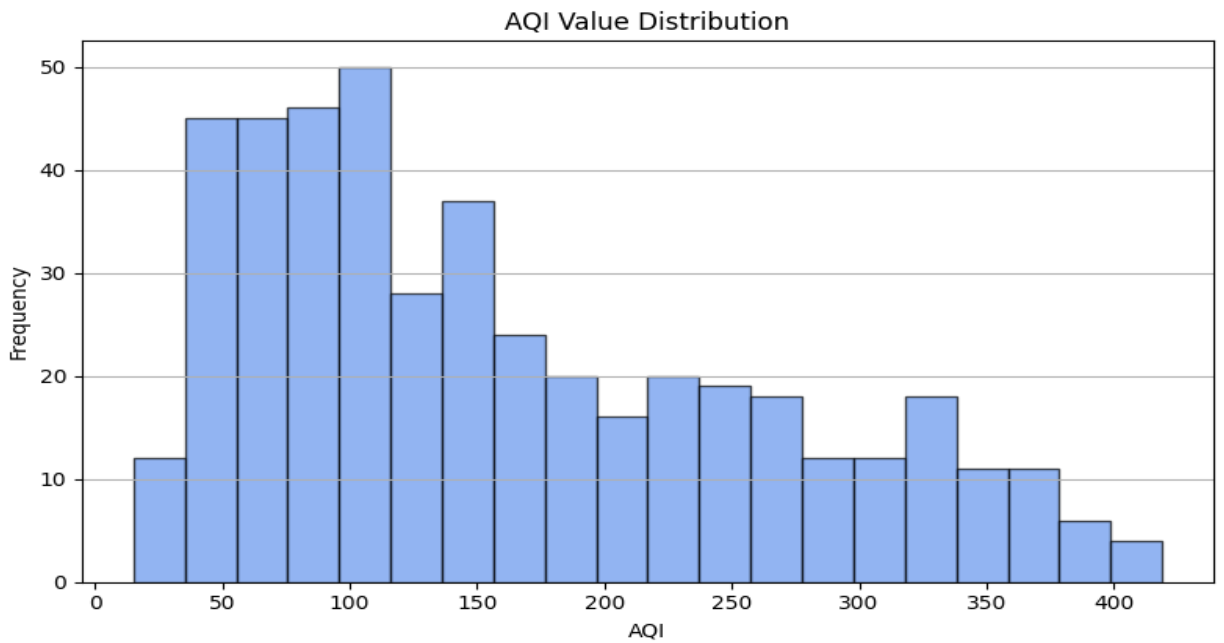

Cities per State by AQI Category
AQI_Category
Good
Satisfactory
Poor
Moderate
Very Poor
Severe
City Count
State
Andhra_Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Delhi
Gujarat
Haryana
Himachal Pradesh
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
TamilNadu
Telangana
Tripura
Uttar_Pradesh
Uttarakhand
West_Bengal

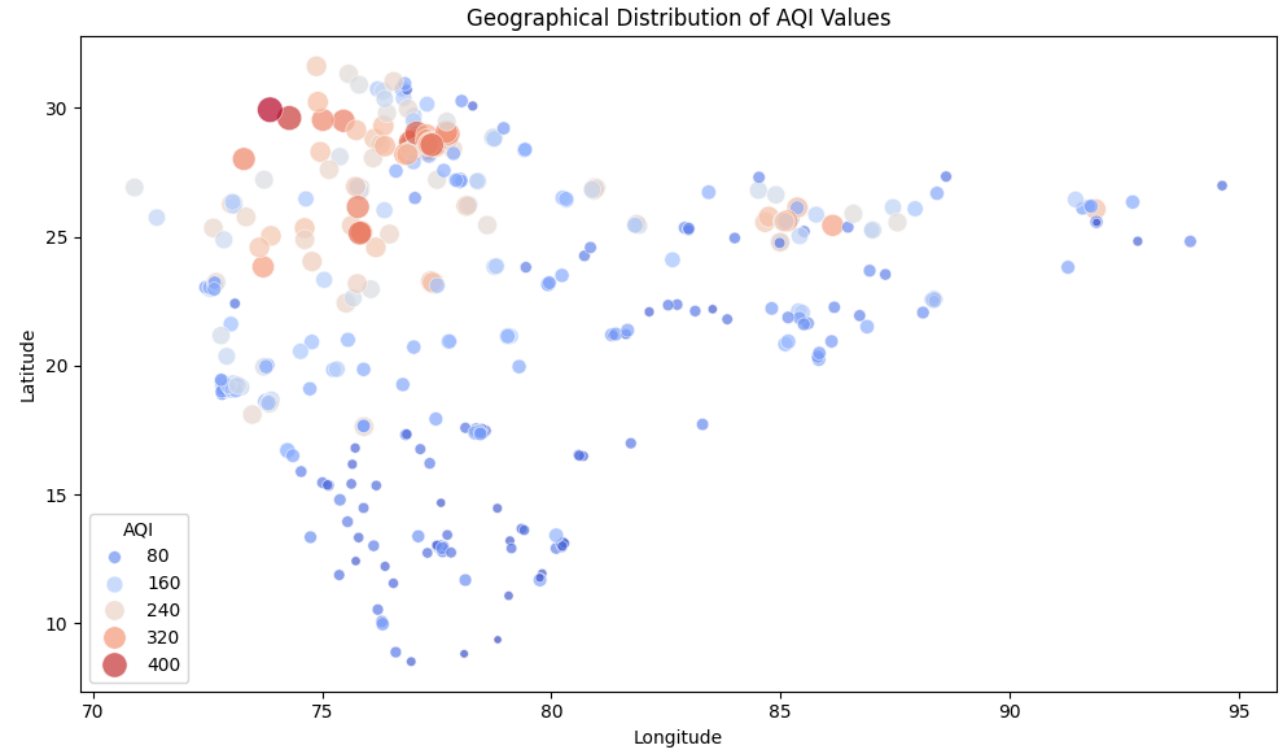

AQI Distribution by Category
AQI
AQI_Category
Good
Satisfactory
Poor
Moderate
Very Poor
Severe

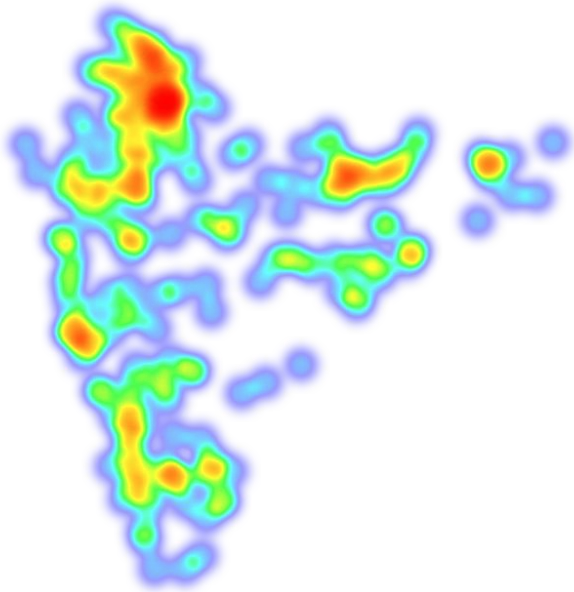

Correlation Matrix
AQI
Latitude
Longitude
1
0.67
-0.19
0.67
1
0.12
-0.19
0.12
1

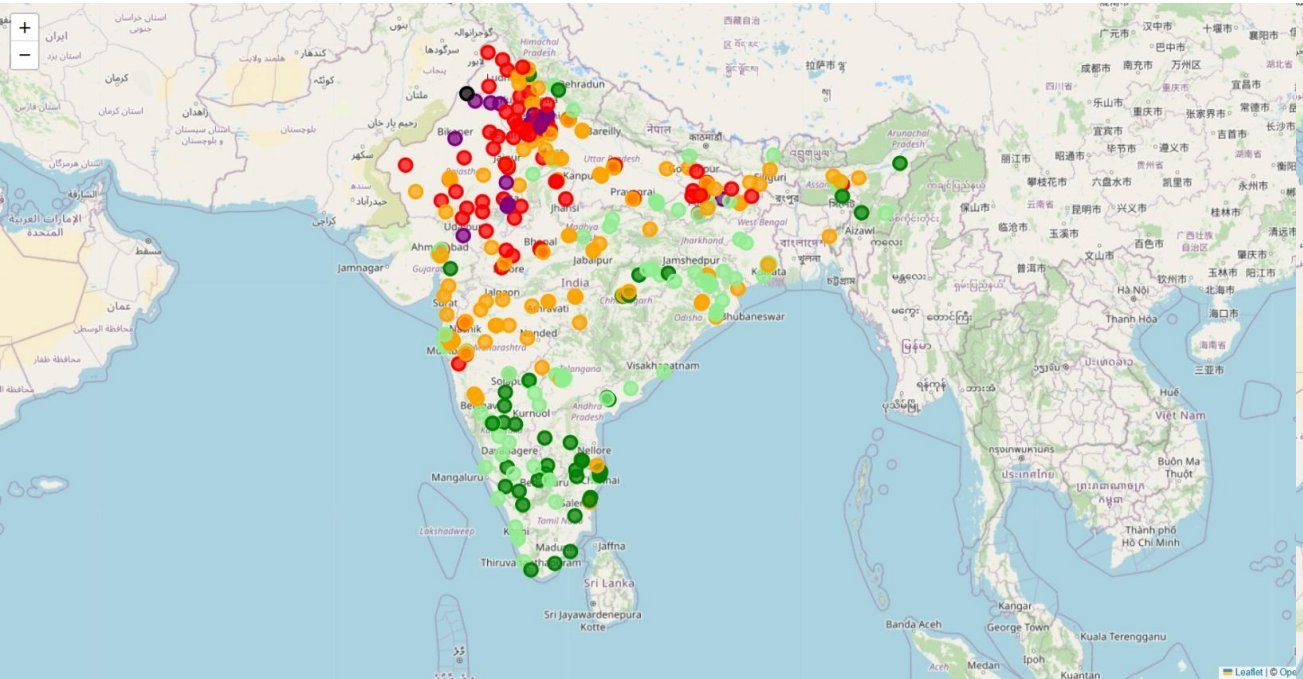

City Count per AQI Category
Number of Cities
AQI Category
Moderate
Satisfactory
Poor
Very Poor
Good
Severe

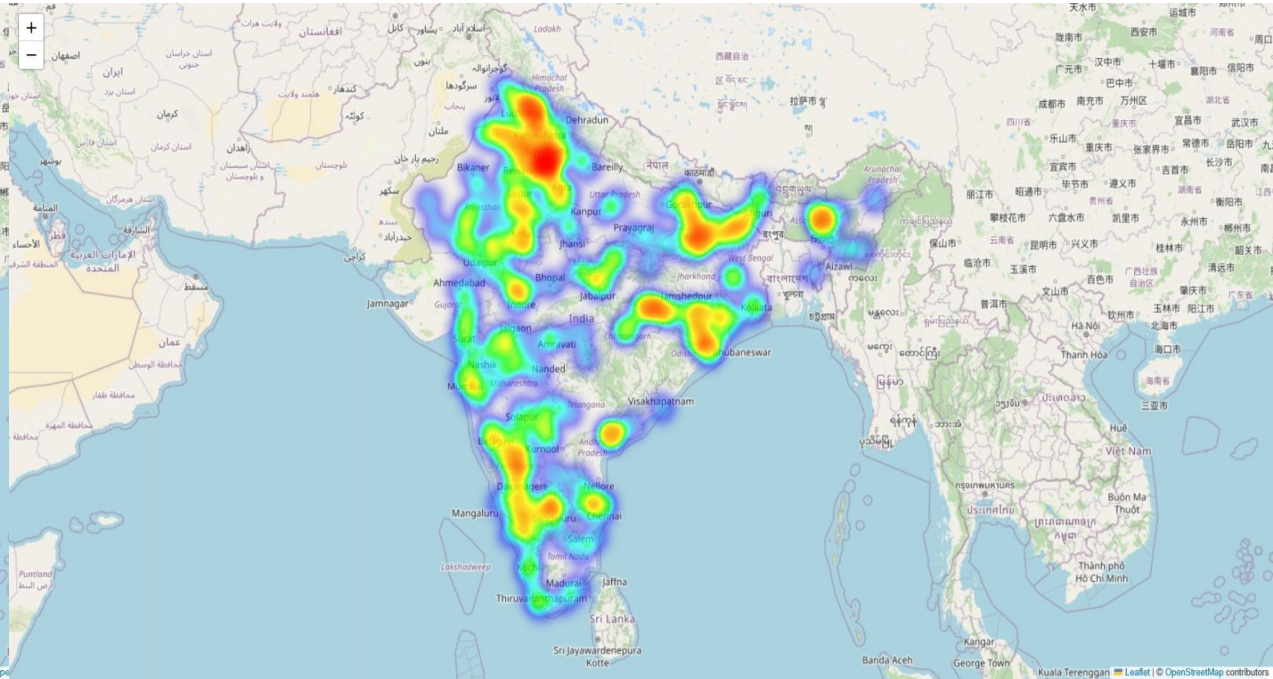

AQI by Longitude, Faceted by Category
AQI_Category = Good
AQI_Category = Satisfactory
AQI_Category = Poor
AQI_Category = Moderate
AQI_Category = Very Poor
AQI_Category = Severe
AQI
Longitude


AQI Value Distribution
Frequency
AQI


Geographical Distribution of AQI Values
Latitude
Longitude
AQI
80
160
240
320
400
30
25
20
15
10
70
75
80
85
90
95




Leaflet | © Ope


Leaflet | © OpenStreetMap contributors

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.p...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distri...
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png
- state_vs_aqi_cat.png
- top_cleanest_cities.png

Disk 69.48 GB available

8s

```
# Install necessary packages; repeat only if you get ImportError later
!pip install pandas folium xmltodict matplotlib seaborn missingno

import pandas as pd
import folium
from folium.plugins import MarkerCluster, HeatMap
import matplotlib.pyplot as plt
import xmltodict
from google.colab import files
import seaborn as sns
import missingno as msno
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (2.2.2)
Requirement already satisfied: folium in /usr/local/lib/python3.11/dist-packages (0.20.0)
Requirement already satisfied: xmltodict in /usr/local/lib/python3.11/dist-packages (0.14.2)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.11/dist-packages (3.10.0)
Requirement already satisfied: seaborn in /usr/local/lib/python3.11/dist-packages (0.13.2)
Requirement already satisfied: missingno in /usr/local/lib/python3.11/dist-packages (0.5.2)
Requirement already satisfied: numpy>=1.23.2 in /usr/local/lib/python3.11/dist-packages (from pandas) (2.0.2)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
Requirement already satisfied: branca>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from folium) (0.8.1)
Requirement already satisfied: jinja2>=2.9 in /usr/local/lib/python3.11/dist-packages (from folium) (3.1.6)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from folium) (2.32.3)
Requirement already satisfied: xyzservices in /usr/local/lib/python3.11/dist-packages (from folium) (2025.4.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (1.3.2)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (4.59.0)
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (1.4.8)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (25.0)
Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (11.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib) (3.2.3)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from missingno) (1.16.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2>=2.9->folium) (3.0.2)
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:10 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

21s

```
# Upload your XML AQI dataset
uploaded = files.upload()
xml_file = list(uploaded.keys())[0]

# Read and parse the XML content
with open(xml_file, 'r', encoding='utf-8') as f:
    xml_content = f.read()
data_dict = xmltodict.parse(xml_content)

def extract_aqi_entries(xml_dict):
    records = []
    for state in xml_dict['AqIndex']['Country']['State']:
        state_name = state['@id']
        if 'City' not in state:
            continue
        cities = state['City']
        if isinstance(cities, dict):
            cities = [cities]
        for city in cities:
            city_name = city['@id']
            if 'Station' not in city:
                continue
            stations = city['Station']
            if isinstance(stations, dict):
                stations = [stations]
            for station in stations:
                latitude = station.get('@latitude')
                longitude = station.get('@longitude')
                aqi_value = None
                if 'Air_Quality_Index' in station:
                    aqi_info = station['Air_Quality_Index']
                    if isinstance(aqi_info, list):
                        aqi_info = aqi_info[0]
                    aqi_value = aqi_info.get('@Value')
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:10 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

[38]

	City	State	Latitude	Longitude	AQI
0	Anantapur	Andhra_Pradesh	14.675886	77.593027	29
1	Chittoor	Andhra_Pradesh	13.204880	79.097889	35
2	Kadapa	Andhra_Pradesh	14.465052	78.824187	37
3	Rajamahendravaram	Andhra_Pradesh	16.987287	81.736318	58
4	Tirupati	Andhra_Pradesh	13.670000	79.350000	49

Next steps: Generate code with df View recommended plots New interactive sheet

[39]
```
# Dataset shape, structure, summary
print(f"Rows: {df.shape[0]}, Columns: {df.shape[1]}")
df.info()
df.describe(include='all')
```

```
Rows: 454, Columns: 5
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 454 entries, 0 to 453
Data columns (total 5 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   City       454 non-null    object
 1   State      454 non-null    object
 2   Latitude   454 non-null    float64
 3   Longitude  454 non-null    float64
 4   AQI        454 non-null    int64
dtypes: float64(2), int64(1), object(2)
memory usage: 17.9+ KB
```

| | City | State | Latitude | Longitude | AQI |

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

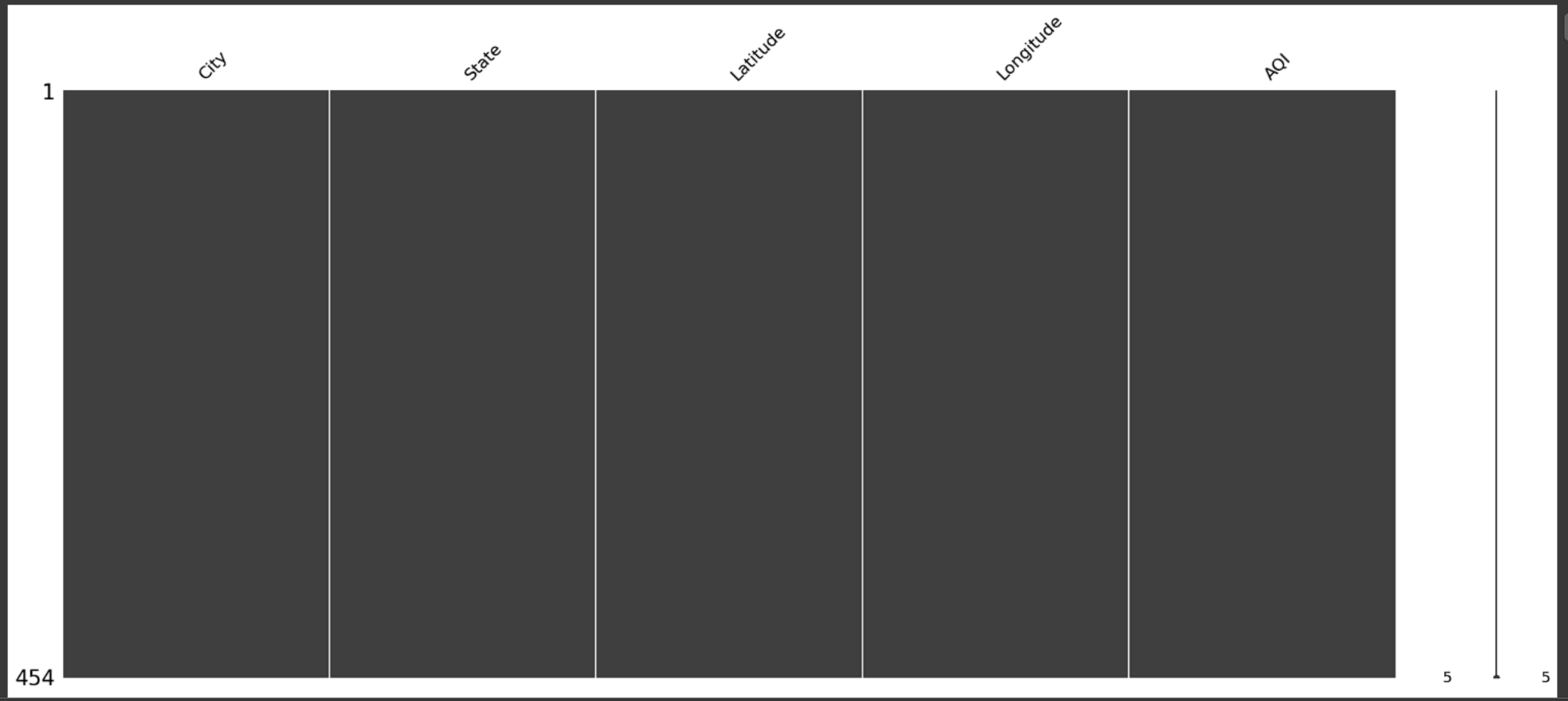
Files

- ..
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.p...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distri...
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png
- state_vs_aqi_cat.png
- top_cleanest_cities.png

Disk 69.48 GB available

```
msno.matrix(df)
plt.show()
```

0s



Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

```
def aqi_category(aqi):
    if aqi <= 50:
        return 'Good'
    elif aqi <= 100:
        return 'Satisfactory'
    elif aqi <= 200:
        return 'Moderate'
    elif aqi <= 300:
        return 'Poor'
    elif aqi <= 400:
        return 'Very Poor'
    else:
        return 'Severe'
df['AQI_Category'] = df['AQI'].apply(aqi_category)
df.head()
```

	City	State	Latitude	Longitude	AQI	AQI_Category
0	Anantapur	Andhra_Pradesh	14.675886	77.593027	29	Good
1	Chittoor	Andhra_Pradesh	13.204880	79.097889	35	Good
2	Kadapa	Andhra_Pradesh	14.465052	78.824187	37	Good
3	Rajamahendravaram	Andhra_Pradesh	16.987287	81.736318	58	Satisfactory
4	Tirupati	Andhra_Pradesh	13.670000	79.350000	49	Good

Next steps: Generate code with df View recommended plots New interactive sheet

```
[43] # Data coverage across states
     state_counts = df['State'].value_counts()
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN 21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

0s

```
# Create a map figure using Plotly
fig_map = px.scatter_mapbox(
    df, lat="Latitude", lon="Longitude", color="AQI_Category", size="AQI",
    hover_name="City", hover_data=["State","AQI"],
    color_discrete_map=color_dict,
    mapbox_style="carto-positron", zoom=4, height=420,
    title="Indian Cities AQI Map"
)

# Histogram
fig_hist = px.histogram(df, x="AQI", nbins=20, color="AQI_Category", title="AQI Value Distribution")

# Pie chart
fig_pie = px.pie(df, names="AQI_Category", title="City Count by AQI Category", color="AQI_Category", color_discrete_map=color_dict)

# Dash Layout
app = dash.Dash(__name__)
app.layout = html.Div([
    html.H2("Indian Cities AQI Dashboard (via Dash & Plotly)"),
    dcc.Tabs([
        dcc.Tab(label='Map', children=[dcc.Graph(figure=fig_map)]),
        dcc.Tab(label='Histogram', children=[dcc.Graph(figure=fig_hist)]),
        dcc.Tab(label='Pie Chart', children=[dcc.Graph(figure=fig_pie)]),
        dcc.Tab(label='Table', children=[
            dash_table.DataTable(
                data=df.sort_values(by='AQI', ascending=False).to_dict('records'),
                columns=[{"name": i, "id": i} for i in df.columns],
                page_size=15,
                style_table={'overflowX': 'auto'},
                style_cell={'textAlign': 'left', 'padding': '4px'},
                style_header={'backgroundColor': 'lightgray', 'fontWeight': 'bold'}
            )
        ])
    ])
])
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

31s

```
# Required imports
!pip install folium --quiet
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import folium
from folium.plugins import HeatMap

# Load your dataset (ensure india_aqi.csv exists from previous steps!)
df = pd.read_csv('india_aqi.csv')

# Ensure AQI_Category column exists
def aqi_category(aqi):
    if aqi <= 50: return 'Good'
    elif aqi <= 100: return 'Satisfactory'
    elif aqi <= 200: return 'Moderate'
    elif aqi <= 300: return 'Poor'
    elif aqi <= 400: return 'Very Poor'
    else: return 'Severe'
if 'AQI_Category' not in df.columns:
    df['AQI_Category'] = df['AQI'].apply(aqi_category)

color_dict = {'Good': 'green','Satisfactory': 'lightgreen','Moderate': 'orange',
              'Poor': 'red','Very Poor': 'purple','Severe': 'black'}

# --- Dataset summary info (prints, not image) ---
print("DATASET SUMMARY:")
print(df.describe(include='all'))
print(f"\nRows: {df.shape[0]}, Columns: {df.shape[1]}")
print(f"Columns: {list(df.columns)}\n")

# --- Visual 1: Records Per State (Bar Plot)
plt.figure(figsize=(10,5))
df['State'].value_counts().plot(kind='bar', color='teal')
plt.title('Number of AQI Stations per State')
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

31s

```
print(f"Columns: {list(df.columns)}\n")

# --- Visual 1: Records Per State (Bar Plot)
plt.figure(figsize=(10,5))
df['State'].value_counts().plot(kind='bar', color='teal')
plt.title('Number of AQI Stations per State')
plt.xlabel('State')
plt.ylabel('Number of Records')
plt.xticks(rotation=75)
plt.tight_layout()
plt.savefig('records_per_state.png')
plt.show()

# --- Visual 2: Top 10 Cleanest Cities
top_clean = df.sort_values(by='AQI').head(10)
plt.figure(figsize=(10,5))
plt.bar(top_clean['City'] + ", " + top_clean['State'], top_clean['AQI'], color='mediumseagreen')
plt.title('Top 10 Cleanest Cities')
plt.xlabel('City, State')
plt.ylabel('AQI')
plt.xticks(rotation=75)
plt.tight_layout()
plt.savefig('top_cleanest_cities.png')
plt.show()

# --- Visual 3: AQI Category Distribution by State
plt.figure(figsize=(15,7))
pd.crosstab(df['State'], df['AQI_Category']).plot(kind='bar', stacked=True, colormap="tab20c", ax=plt.gca())
plt.title('AQI Category Distribution by State')
plt.xlabel('State')
plt.ylabel('Number of Cities')
plt.tight_layout()
plt.savefig('aqi_category_distribution_state.png')
plt.show()
```

Variables Terminal

8:43PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

31s

```
# --- Visual 4: AQI Distribution Histogram
plt.figure(figsize=(8,5))
plt.hist(df['AQI'], bins=20, color='cornflowerblue', edgecolor='black', alpha=0.7)
plt.title('AQI Value Distribution')
plt.xlabel('AQI')
plt.ylabel('Frequency')
plt.grid(axis='y')
plt.tight_layout()
plt.savefig('aqi_histogram.png')
plt.show()

# --- Visual 5: City Count per AQI Category (Countplot)
plt.figure(figsize=(8,5))
sns.countplot(x='AQI_Category', data=df, order=df['AQI_Category'].value_counts().index, palette='Spectral')
plt.title('City Count per AQI Category')
plt.xlabel('AQI Category')
plt.ylabel('Number of Cities')
plt.tight_layout()
plt.savefig('city_count_aqi_category.png')
plt.show()

# --- Visual 6: Geo scatter (Longitude vs Latitude, hue=AQI)
plt.figure(figsize=(10,6))
sns.scatterplot(x='Longitude', y='Latitude', hue='AQI', data=df, palette='coolwarm', size='AQI', sizes=(20,200), alpha=0.7, legend='brief')
plt.title('Geographical Distribution of AQI Values')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.tight_layout()
plt.savefig('geo_scatter_aqi.png')
plt.show()

# --- Visual 7: State vs AQI Category (2D Categorical)
plt.figure(figsize=(12,6))
sns.countplot(x='State', hue='AQI_Category', data=df, palette='tab20')
plt.title('Cities per State by AQI Category')
```

Variables Terminal

8:43 PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

31s

```
plt.show()

# --- Visual 7: State vs AQI Category (2D Categorical)
plt.figure(figsize=(12,6))
sns.countplot(x='State', hue='AQI_Category', data=df, palette='tab20')
plt.title('Cities per State by AQI Category')
plt.xlabel('State')
plt.ylabel('City Count')
plt.xticks(rotation=90)
plt.tight_layout()
plt.savefig('state_vs_aqi_cat.png')
plt.show()

# --- Visual 8: Faceted AQI Histogram by AQI Category
g = sns.FacetGrid(df, col="AQI_Category", col_wrap=3, height=4, sharex=False, sharey=False)
g.map(sns.histplot, "AQI", bins=10, color='darkslateblue')
g.fig.subplots_adjust(top=0.9)
g.fig.suptitle('AQI Distribution within Each Category')
g.savefig('facet_aqi_category_distribution.png')
plt.show()

# --- Visual 9: Jointplot (AQI vs Latitude)
jp = sns.jointplot(x='Latitude', y='AQI', data=df, kind='hex', gridsize=25, cmap='YlOrRd')
jp.fig.suptitle('2D Distribution: AQI vs Latitude', y=1.02)
jp.savefig('joint_lat_aqi.png')
plt.show()

# --- Visual 10: Boxplot AQI by Category
plt.figure(figsize=(8,5))
sns.boxplot(x='AQI_Category', y='AQI', data=df, palette='Set2')
plt.title('AQI Distribution by Category')
plt.tight_layout()
plt.savefig('boxplot_aqi_category.png')
plt.show()

# --- Visual 11: Faceted Scatterplot AQI by Longitude and Category
```

Variables Terminal

8:43PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available

31s

```
sns.boxplot(x='AQI_Category', y='AQI', data=df, palette='Set2')
plt.title('AQI Distribution by Category')
plt.tight_layout()
plt.savefig('boxplot_aqi_category.png')
plt.show()

# --- Visual 11: Faceted Scatterplot AQI by Longitude and Category
fg = sns.FacetGrid(df, col='AQI_Category', col_wrap=3, height=4)
fg.map_dataframe(sns.scatterplot, x='Longitude', y='AQI', alpha=0.7)
fg.set_axis_labels("Longitude", "AQI")
fg.fig.subplots_adjust(top=0.9)
fg.fig.suptitle('AQI by Longitude, Faceted by Category')
fg.savefig('facet_scatter_aqi_long.png')
plt.show()

# --- Visual 12: Statewise Distributions (FacetGrid top 6 states)
top_states = df['State'].value_counts().head(6).index.tolist()
df_facet = df[df['State'].isin(top_states)]
g = sns.FacetGrid(df_facet, col="State", col_wrap=3, height=4)
g.map_dataframe(sns.histplot, x='AQI', bins=10, kde=True, color='coral')
g.set_titles(col_template="{col_name}")
g.fig.suptitle('AQI Distribution Faceted by Top 6 States', y=1.02)
g.savefig('facet_aqi_by_state.png')
plt.show()

# --- Visual 13: Correlation Matrix Heatmap
plt.figure(figsize=(5,3))
sns.heatmap(df[['AQI', 'Latitude', 'Longitude']].corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.tight_layout()
plt.savefig('correlation_matrix.png')
plt.show()

# --- Visual 14: Folium AQI Marker Map (Saves to HTML)
india_map = folium.Map(location=[22.97, 78.65], zoom_start=5)
for _, row in df.iterrows():
```

Variables Terminal

8:43PM Python 3

27°C Haze

Search

ENG IN

21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.p...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distri...
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png
- state_vs_aqi_cat.png
- top_cleanest_cities.png

Disk 69.48 GB available

31s

```
# --- Visual 13: Correlation Matrix Heatmap
plt.figure(figsize=(5,3))
sns.heatmap(df[['AQI', 'Latitude', 'Longitude']].corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.tight_layout()
plt.savefig('correlation_matrix.png')
plt.show()

# --- Visual 14: Folium AQI Marker Map (Saves to HTML)
india_map = folium.Map(location=[22.97, 78.65], zoom_start=5)
for _, row in df.iterrows():
    folium.CircleMarker(
        location=[row['Latitude'], row['Longitude']],
        radius=7,
        popup=f"{row['City']}, {row['State']}<br>AQI: {row['AQI']}",
        tooltip=f"{row['City']}: AQI {row['AQI']}",
        color=color_dict[row['AQI_Category']],
        fill=True,
        fill_opacity=0.7
    ).add_to(india_map)
india_map.save('india_aqi_map.html')
print("Folium AQI map saved as 'india_aqi_map.html'.")

# --- Visual 15: Folium Heatmap (Saves to HTML)
heat_data = df[['Latitude', 'Longitude', 'AQI']].values.tolist()
india_heatmap = folium.Map(location=[22.97, 78.65], zoom_start=5)
HeatMap(heat_data, radius=18, blur=12, max_zoom=1).add_to(india_heatmap)
india_heatmap.save('india_aqi_heatmap.html')
print("Folium AQI heatmap saved as 'india_aqi_heatmap.html'.")

print("\\nAll charts saved as PNG images. Both folium maps saved as HTML and can be downloaded for sharing or embedding.")
```

Variables Terminal

8:43PM Python 3

27°C Haze

Search

ENG IN 21:11 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=lqh-OcmvvfqH

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distribution_...
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available

```
from plotly.subplots import make_subplots
import plotly.graph_objs as go

# ... Assume df, cat_counts, color_dict already defined above ...

fig = make_subplots(
    rows=1, cols=2,
    specs=[[{"type": "xy"}, {"type": "domain"}]],  # <-- THIS LINE IS KEY
    subplot_titles=("AQI Value Distribution", "Cities by AQI Category")
)

# Histogram in the first subplot
fig.add_trace(
    go.Histogram(x=df['AQI'], nbinsx=20, name='AQI', marker=dict(color='skyblue'), opacity=0.8),
    row=1, col=1
)

# Pie chart in the second subplot
cat_counts = df['AQI_Category'].value_counts()
fig.add_trace(
    go.Pie(
        labels=cat_counts.index,
        values=cat_counts.values,
        marker=dict(colors=[color_dict[c] for c in cat_counts.index]),
        hole=0.3
    ),
    row=1, col=2
)

fig.update_layout(height=400, title_text="AQI Distribution & Category Share")
fig.show()
```

Variables Terminal

9:20 PM Python 3

News for you
'Avatar: Fire and...

Search

ENG IN 21:21 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distribution_...
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available

5s

```
# 1. Install Plotly (skip if already installed)
!pip install plotly --quiet

# 2. Imports
import pandas as pd
import plotly.graph_objs as go
from plotly.subplots import make_subplots
import plotly.express as px
from IPython.display import display, Markdown

# 3. Load Data
df = pd.read_csv('india_aqi.csv')

# 4. Add AQI_Category if needed
def aqi_category(aqi):
    if aqi <= 50: return 'Good'
    elif aqi <= 100: return 'Satisfactory'
    elif aqi <= 200: return 'Moderate'
    elif aqi <= 300: return 'Poor'
    elif aqi <= 400: return 'Very Poor'
    else: return 'Severe'
if 'AQI_Category' not in df.columns:
    df['AQI_Category'] = df['AQI'].apply(aqi_category)

color_dict = {
    'Good': 'green', 'Satisfactory': 'lightgreen', 'Moderate': 'orange',
    'Poor': 'red', 'Very Poor': 'purple', 'Severe': 'black'
}

# --- TOP KPIs ---
total_cities = df['City'].nunique()
avg_aqi = df['AQI'].mean()
worst_row = df.loc[df['AQI'].idxmax()]
best_row = df.loc[df['AQI'].idxmin()]
```

Variables Terminal

9:20 PM Python 3

News for you
'Avatar: Fire and...

Search

ENG IN 21:21 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- aqi_category_distribution_...
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.p...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distri...
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png

Disk 69.48 GB available

5s

```
# --- TOP KPIs ---
total_cities = df['City'].nunique()
avg_aqi = df['AQI'].mean()
worst_row = df.loc[df['AQI'].idxmax()]
best_row = df.loc[df['AQI'].idxmin()]

display(Markdown(f"""
## <span style="color:#008080;">Project AQI Dashboard</span>
**Cities:** {total_cities}
**States:** {df['State'].nunique()}
**Average AQI:** <span style='color:orange;font-weight:bold'>{avg_aqi:.1f}</span>
**Most Polluted:** <span style='color:red'>{worst_row['City']} ({worst_row['State']}): AQI {worst_row['AQI']}</span>
**Cleanest:** <span style='color:green'>{best_row['City']} ({best_row['State']}): AQI {best_row['AQI']}</span>
"""))

# --- 1. INTERACTIVE AQI MAP (Plotly) ---
fig_map = px.scatter_mapbox(
    df, lat="Latitude", lon="Longitude", color="AQI_Category", size="AQI",
    hover_name="City", hover_data=["State","AQI"],
    color_discrete_map=color_dict,
    mapbox_style="carto-positron", zoom=4, height=400,
    title="Indian Cities: AQI Map by Category"
)
fig_map.show()

# --- 2. AQI HISTOGRAM & CATEGORY PIE CHART ---
fig = make_subplots(
    rows=1, cols=2,
    specs=[[{"type": "xy"}, {"type": "domain"}]],  # histogram + pie in grid
    subplot_titles=("AQI Value Distribution", "Cities by AQI Category")
)

fig.add_trace(
    go.Histogram(x=df['AQI'], nbinsx=20, name='AQI', marker=dict(color='skyblue'), opacity=0.8),
    row=1, col=1
```

Variables Terminal

9:20 PM Python 3

News for you 'Avatar: Fire and...

Search

ENG IN 21:21 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

sample_data

aqi_category_distribution_...

aqi_histogram.png

boxplot_aqi_category.png

city_count_aqi_category.p...

correlation_matrix.png

data_aqi_cpcb.xml

facet_aqi_by_state.png

facet_aqi_category_distri...

facet_scatter_aqi_long.png

geo_scatter_aqi.png

india_aqi.csv

india_aqi_heatmap.html

india_aqi_map.html

joint_lat_aqi.png

records_per_state.png

Disk 69.48 GB available

5s

```
fig_map.show()

# --- 2. AQI HISTOGRAM & CATEGORY PIE CHART ---
fig = make_subplots(
    rows=1, cols=2,
    specs=[[{"type": "xy"}, {"type": "domain"}]],  # histogram + pie in grid
    subplot_titles=("AQI Value Distribution", "Cities by AQI Category")
)

fig.add_trace(
    go.Histogram(x=df['AQI'], nbinsx=20, name='AQI', marker=dict(color='skyblue'), opacity=0.8),
    row=1, col=1
)

cat_counts = df['AQI_Category'].value_counts()
fig.add_trace(
    go.Pie(
        labels=cat_counts.index,
        values=cat_counts.values,
        marker=dict(colors=[color_dict[c] for c in cat_counts.index]),
        hole=0.3
    ),
    row=1, col=2
)
fig.update_layout(height=400, title_text="AQI Distribution & Category Share")
fig.show()

# --- 3. TOP/BOTTOM CITIES BAR CHARTS ---
top10 = df.nlargest(10, 'AQI')
bot10 = df.nsmallest(10, 'AQI')

fig2 = make_subplots(
    rows=1, cols=2,
    subplot_titles=("Top 10 Most Polluted Cities", "Top 10 Cleanest Cities")
)
```

Variables Terminal

9:20 PM Python 3

News for you
'Avatar: Fire and...

Search

ENG IN

21:21
29-07-2025

Files
..
sample_data
aqi_category_distribution_...
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.p...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
Disk 69.48 GB available

```
)
fig.update_layout(height=400, title_text="AQI Distribution & Category Share")
fig.show()

# --- 3. TOP/BOTTOM CITIES BAR CHARTS ---
top10 = df.nlargest(10, 'AQI')
bot10 = df.nsmallest(10, 'AQI')

fig2 = make_subplots(
    rows=1, cols=2,
    subplot_titles=("Top 10 Most Polluted Cities", "Top 10 Cleanest Cities")
)
fig2.add_trace(go.Bar(
    x=top10['City'] + '-' + top10['State'],
    y=top10['AQI'],
    marker=dict(color='crimson')
), row=1, col=1)
fig2.add_trace(go.Bar(
    x=bot10['City'] + '-' + bot10['State'],
    y=bot10['AQI'],
    marker=dict(color='forestgreen')
), row=1, col=2)
fig2.update_layout(height=400, title_text="City Rankings by AQI")
fig2.update_xaxes(tickangle=60)
fig2.show()

# --- 4. STATEWISE AQI CATEGORY DISTRIBUTION (STACKED BAR) ---
catby = pd.crosstab(df['State'], df['AQI_Category'])
catby = catby.loc[catby.sum(axis=1).sort_values(ascending=False).index]  # Sort by state size

fig3 = go.Figure()
for category in catby.columns:
    fig3.add_trace(go.Bar(
        name=category,
        x=catby.index,
        y=catby[category],
```

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- aqi_category_distribution_...
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.p...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distri...
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png

Disk 69.48 GB available

5s

```
fig2.show()

# --- 4. STATEWISE AQI CATEGORY DISTRIBUTION (STACKED BAR) ---
catby = pd.crosstab(df['State'], df['AQI_Category'])
catby = catby.loc[catby.sum(axis=1).sort_values(ascending=False).index]  # Sort by state size

fig3 = go.Figure()
for category in catby.columns:
    fig3.add_trace(go.Bar(
        name=category,
        x=catby.index,
        y=catby[category],
        marker_color=color_dict[category]
    ))
fig3.update_layout(
    barmode='stack',
    title='AQI Category Distribution by State',
    xaxis_title='State',
    yaxis_title='City Count',
    height=500
)
fig3.update_xaxes(tickangle=60)
fig3.show()

# --- Optional: Dashboard note
display(Markdown("""
*Dashboard sections:
- KPIs
- Interactive AQI map
- Distribution & share
- Top/bottom 10 cities
- Statewise category breakdown*

*Interact with plots: Zoom, hover, or export as PNG for slides.*
"""))
```

Variables Terminal

9:20 PM Python 3

News for you
'Avatar: Fire and...

Search

ENG IN 21:21 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=CTVepwPkArg8

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
app.py
aqi_category_distribution_state.png
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.png
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distribution.png
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png

Disk 69.18 GB available

```
# main.ipynb
app_code = '''
import streamlit as st
import pandas as pd
import plotly.express as px

@st.cache_data
def load_data():
    df = pd.read_csv('india_aqi.csv')
    def aqi_category(aqi):
        if aqi <= 50: return 'Good'
        elif aqi <= 100: return 'Satisfactory'
        elif aqi <= 200: return 'Moderate'
        elif aqi <= 300: return 'Poor'
        elif aqi <= 400: return 'Very Poor'
        else: return 'Severe'
    if 'AQI_Category' not in df.columns:
        df['AQI_Category'] = df['AQI'].apply(aqi_category)
    return df

df = load_data()

st.title("Air Quality Index (AQI) Dashboard - Indian Cities")

# Scrollable data table with pagination
st.subheader("AQI Data Table")
rows_per_page = 10
total_rows = len(df)
pages = (total_rows // rows_per_page) + 1
page_number = st.number_input("Page Number", min_value=1, max_value=pages, value=1, step=1)
start_idx = (page_number - 1) * rows_per_page
end_idx = start_idx + rows_per_page
st.dataframe(df.iloc[start_idx:end_idx, :], height=300)
```

Variables Terminal

10:30 PM Python 3

27°C Haze

Search

ENG IN

22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=CTVepwPkArg8

HCL+GUVI_Project-1.ipynb Saving...

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- app.py
- aqi_category_distribution_state.png
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.png
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distribution.png
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png

Disk 69.18 GB available

[121]
```
# AQI histogram
fig_hist = px.histogram(
    df, x="AQI", nbins=30, color="AQI_Category", color_discrete_map=color_dict
)
st.plotly_chart(fig_hist, use_container_width=True)

# Top 10 most polluted cities bar
top10 = df.nlargest(10, "AQI")
fig_bar_top = px.bar(
    top10, x="City", y="AQI", color="AQI_Category", color_discrete_map=color_dict, title="Top 10 Most Polluted Cities"
)
st.plotly_chart(fig_bar_top, use_container_width=True)

# Top 10 cleanest cities bar
bot10 = df.nsmallest(10, "AQI")
fig_bar_bot = px.bar(
    bot10, x="City", y="AQI", color="AQI_Category", color_discrete_map=color_dict, title="Top 10 Cleanest Cities"
)
st.plotly_chart(fig_bar_bot, use_container_width=True)

# Cities by AQI Category pie chart
cat_counts = df["AQI_Category"].value_counts()
fig_pie = px.pie(
    names=cat_counts.index,
    values=cat_counts.values,
    color=cat_counts.index,
    color_discrete_map=color_dict,
    hole=0.3,
)
st.plotly_chart(fig_pie, use_container_width=True)
'''
with open("app.py", "w") as f:
    f.write(app_code)
```

Variables Terminal

10:30 PM Python 3

27°C Haze

Search

ENG IN

22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=xikot8VBArcE

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
app.py
aqi_category_distribution_state.png
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category.png
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distribution.png
facet_scatter_aqi_long.png
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png

Disk 69.18 GB available

```
      --exclude-domains=LIST      comma-separated list of rejected domains
      --follow-ftp                follow FTP links from HTML documents
      --follow-tags=LIST          comma-separated list of followed HTML tags
      --ignore-tags=LIST          comma-separated list of ignored HTML tags
  -H, --span-hosts                go to foreign hosts when recursive
  -L, --relative                  follow relative links only
  -I, --include-directories=LIST  list of allowed directories
      --trust-server-names        use the name specified by the redirection
                                    URL's last component
  -X, --exclude-directories=LIST  list of excluded directories
  -np, --no-parent                don't ascend to the parent directory

Email bug reports, questions, discussions to <bug-wget@gnu.org>
and/or open issues at https://savannah.gnu.org/bugs/?func=additem&group=wget.
```

0s
```
[128] !wget -q -O - ipv4.icanhazip.com
```

```
34.23.201.254
```

4m
```
!streamlit run app.py & npx localtunnel --port 8501
```

```
Collecting usage statistics. To deactivate, set browser.gatherUsageStats to false.


  You can now view your Streamlit app in your browser.

  Local URL: http://localhost:8501
  Network URL: http://172.28.0.12:8501
  External URL: http://34.23.201.254:8501

your url is: https://full-crews-smell.loca.lt
  Stopping...
^C
```

Variables Terminal

10:30 PM Python 3

27°C Haze

Search

ENG IN

22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1k9a5ybzIHdWJiIEIJFbzclDsPCgbLwkq#scrollTo=-HTREFfUA92F

app.py

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

sample_data

india_aqi.csv

Disk 69.43 GB available

```
# Install streamlit and dependencies
!pip install streamlit pandas plotly
!pip install pyngrok
```

11s

```
Collecting streamlit
  Downloading streamlit-1.47.1-py3-none-any.whl.metadata (9.0 kB)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (2.2.2)
Requirement already satisfied: plotly in /usr/local/lib/python3.11/dist-packages (5.24.1)
Requirement already satisfied: altair<6,>=4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.5.0)
Requirement already satisfied: blinker<2,>=1.5.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (1.9.0)
Requirement already satisfied: cachetools<7,>=4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.5.2)
Requirement already satisfied: click<9,>=7.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (8.2.1)
Requirement already satisfied: numpy<3,>=1.23 in /usr/local/lib/python3.11/dist-packages (from streamlit) (2.0.2)
Requirement already satisfied: packaging<26,>=20 in /usr/local/lib/python3.11/dist-packages (from streamlit) (25.0)
Requirement already satisfied: pillow<12,>=7.1.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (11.3.0)
Requirement already satisfied: protobuf<7,>=3.20 in /usr/local/lib/python3.11/dist-packages (from streamlit) (5.29.5)
Requirement already satisfied: pyarrow>=7.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (18.1.0)
Requirement already satisfied: requests<3,>=2.27 in /usr/local/lib/python3.11/dist-packages (from streamlit) (2.32.3)
Requirement already satisfied: tenacity<10,>=8.1.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (8.5.0)
Requirement already satisfied: toml<2,>=0.10.1 in /usr/local/lib/python3.11/dist-packages (from streamlit) (0.10.2)
Requirement already satisfied: typing-extensions<5,>=4.4.0 in /usr/local/lib/python3.11/dist-packages (from streamlit) (4.14.1)
Collecting watchdog<7,>=2.1.5 (from streamlit)
  Downloading watchdog-6.0.0-py3-none-manylinux2014_x86_64.whl.metadata (44 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 44.3/44.3 kB 3.1 MB/s eta 0:00:00
Requirement already satisfied: gitpython!=3.1.19,<4,>=3.0.7 in /usr/local/lib/python3.11/dist-packages (from streamlit) (3.1.45)
Collecting pydeck<1,>=0.8.0b4 (from streamlit)
  Downloading pydeck-0.9.1-py2.py3-none-any.whl.metadata (4.1 kB)
Requirement already satisfied: tornado!=6.5.0,<7,>=6.0.3 in /usr/local/lib/python3.11/dist-packages (from streamlit) (6.4.2)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas) (2025.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from altair<6,>=4.0->streamlit) (3.1.6)
Requirement already sat                                                    om altair<6,>=4.0->streamlit) (4.25.0)
Requirement already sa                                                     om altair<6,>=4.0->streamlit) (1.48.1)
Requirement already sat                                                    om gitpython!=3.1.19,<4,>=3.0.7->streamlit) (4.0.12)
```

What can I help you build?

Variables Terminal

9:59 PM Python 3

27°C Haze

Search

ENG IN 22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1k9a5ybzIHdWJiIEIJFbzclDsPCgbLwkq#scrollTo=-HTREFfUA92F

app.py

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

sample_data

india_aqi.csv

Disk 69.43 GB available

```
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3,>=2.27->streamlit) (2025.7.14)
Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.11/dist-packages (from gitdb<5,>=4.0.1->gitpython!=3.1.19,<4,>=3.0.7->streamlit) (5.0.2)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->altair<6,>=4.0->streamlit) (3.0.2)
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit) (25.3.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit) (2025.
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit) (0.36.2)
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=3.0->altair<6,>=4.0->streamlit) (0.26.0)
```

0s

```python
import streamlit as st
import pandas as pd
import plotly.express as px

@st.cache_data
def load_data():
    df = pd.read_csv('india_aqi.csv')
    def aqi_category(aqi):
        if aqi <= 50: return 'Good'
        elif aqi <= 100: return 'Satisfactory'
        elif aqi <= 200: return 'Moderate'
        elif aqi <= 300: return 'Poor'
        elif aqi <= 400: return 'Very Poor'
        else: return 'Severe'
    if 'AQI_Category' not in df.columns:
        df['AQI_Category'] = df['AQI'].apply(aqi_category)
    return df

df = load_data()

st.title("Air Quality Index (AQI) Dashboard - Indian Cities")

# Scrollable data table with pagination
st.subheader("AQI Data Table")
rows_per_page = 10
total_rows = len(df)
pages = (total_rows // rows_per_page) + 1
```

Variables Terminal

9:59 PM Python 3

27°C Haze

Search

ENG IN

22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1k9a5ybzIHdWJiIEIJFbzclDsPCgbLwkq#scrollTo=-HTREFfUA92F

app.py

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

sample_data

india_aqi.csv

Disk 69.43 GB available

```
total_cities = df['City'].nunique()
avg_aqi = df['AQI'].mean()
worst_city = df.loc[df['AQI'].idxmax()]
cleanest_city = df.loc[df['AQI'].idxmin()]

col1, col2, col3 = st.columns(3)
col1.metric("Total Cities", total_cities)
col2.metric("Average AQI", f"{avg_aqi:.1f}")
col3.metric("Most Polluted City", f"{worst_city['City']} ({worst_city['AQI']})")
st.write(f"**Cleanest City:** {cleanest_city['City']} ({cleanest_city['AQI']})")

# Interactive AQI map
fig_map = px.scatter_mapbox(
    df, lat="Latitude", lon="Longitude", color="AQI_Category", size="AQI",
    hover_name="City", hover_data=["State", "AQI"],
    color_discrete_map=color_dict,
    mapbox_style="carto-positron",
    zoom=4,
    height=500
)
st.plotly_chart(fig_map, use_container_width=True)

# AQI histogram
fig_hist = px.histogram(
    df, x="AQI", nbins=30, color="AQI_Category", color_discrete_map=color_dict
)
st.plotly_chart(fig_hist, use_container_width=True)

# Top 10 most polluted cities bar
top10 = df.nlargest(10, "AQI")
fig_bar_top = px.bar(
    top10, x="City", y="AQI", color="AQI_Category", color_discrete_map=color_dict, title="Top 10 Most Polluted Cities"
)
st.plotly_chart(fig_bar_top, use_container_width=True)

# Top 10 cleanest cities bar
```

Variables Terminal

9:59 PM Python 3

27°C Haze

Search

ENG IN

22:31 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

colab.research.google.com/drive/1k9a5ybzIHdWJiIEIJFbzclDsPCgbLwkq#scrollTo=-HTREFfUA92F

app.py

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

..

sample_data

india_aqi.csv

Disk 69.43 GB available

```
# Top 10 cleanest cities bar
bot10 = df.nsmallest(10, "AQI")
fig_bar_bot = px.bar(
    bot10, x="City", y="AQI", color="AQI_Category", color_discrete_map=color_dict, title="Top 10 Cleanest Cities"
)
st.plotly_chart(fig_bar_bot, use_container_width=True)

# Cities by AQI Category pie chart
cat_counts = df["AQI_Category"].value_counts()
fig_pie = px.pie(
    names=cat_counts.index,
    values=cat_counts.values,
    color=cat_counts.index,
    color_discrete_map=color_dict,
    hole=0.3,
)
st.plotly_chart(fig_pie, use_container_width=True)
```

```
2025-07-29 16:29:07.144 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.146 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.147 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.148 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.149 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.150 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.150 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.151 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.152 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.152 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.155 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.159 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.159 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.160 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.161 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.161 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.163 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.166 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
2025-07-29 16:29:07.167 Thread 'MainThread': missing ScriptRunContext! This warning can be ignored when running in bare mode.
```

Variables Terminal

9:59 PM Python 3

27°C Haze

Search

ENG IN

22:32 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

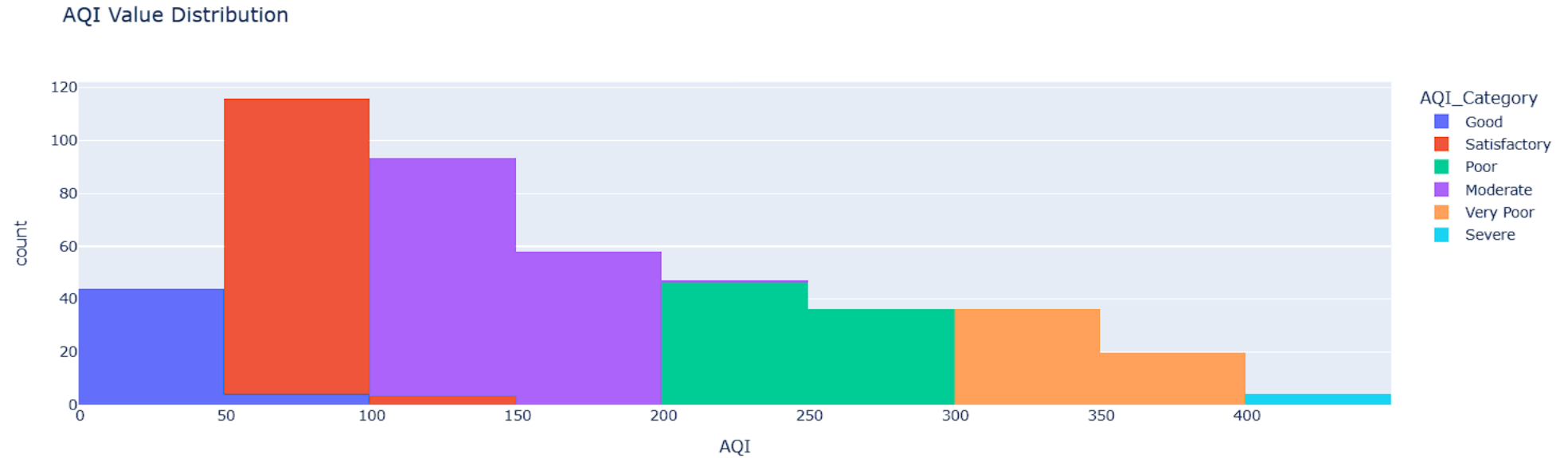
Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available

Indian Cities AQI Dashboard (via Dash & Plotly)

Map | Histogram | Pie Chart | Table



Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:44 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- aqi_category_distributi...
- aqi_histogram.png
- boxplot_aqi_category.p...
- city_count_aqi_categor...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_dis...
- facet_scatter_aqi_long....
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.html
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png

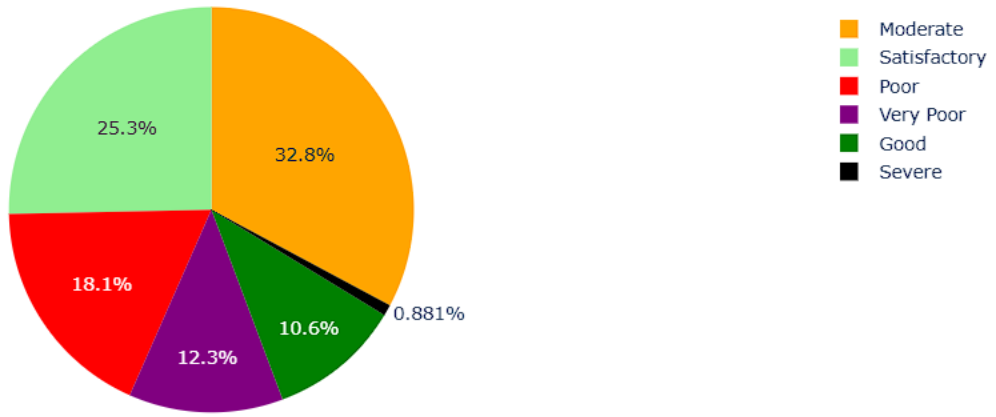
Disk 69.48 GB available

31s

Indian Cities AQI Dashboard (via Dash & Plotly)

Map | Histogram | Pie Chart | Table

City Count by AQI Category



Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:45 29-07-2025

Indian Cities AQI Dashboard (via Dash & Plotly)

Map	Histogram	Pie Chart	Table

City	State	Latitude	Longitude	AQI	AQI_Category
Sri Ganganagar	Rajasthan	29.931624	73.864511	419	Severe
Delhi	Delhi	28.7762	77.051074	416	Severe
Delhi	Delhi	28.674045	77.131023	410	Severe
Delhi	Delhi	28.73282	77.170633	402	Severe
Delhi	Delhi	28.699793	77.165453	398	Very Poor
Hanumangarh	Rajasthan	29.61075	74.283608	394	Very Poor
Delhi	Delhi	28.732528	77.11992	392	Very Poor
Delhi	Delhi	28.56789	77.250515	388	Very Poor
Delhi	Delhi	28.684678	77.076574	384	Very Poor
Delhi	Delhi	28.5710274	77.0719006	384	Very Poor
Delhi	Delhi	28.623763	77.287209	377	Very Poor
Delhi	Delhi	28.822836	77.101981	374	Very Poor
Delhi	Delhi	28.530785	77.271255	370	Very Poor
Delhi	Delhi	28.5504249	77.2159377	368	Very Poor
Sonipat	Haryana	29.0272	77.0621	367	Very Poor

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

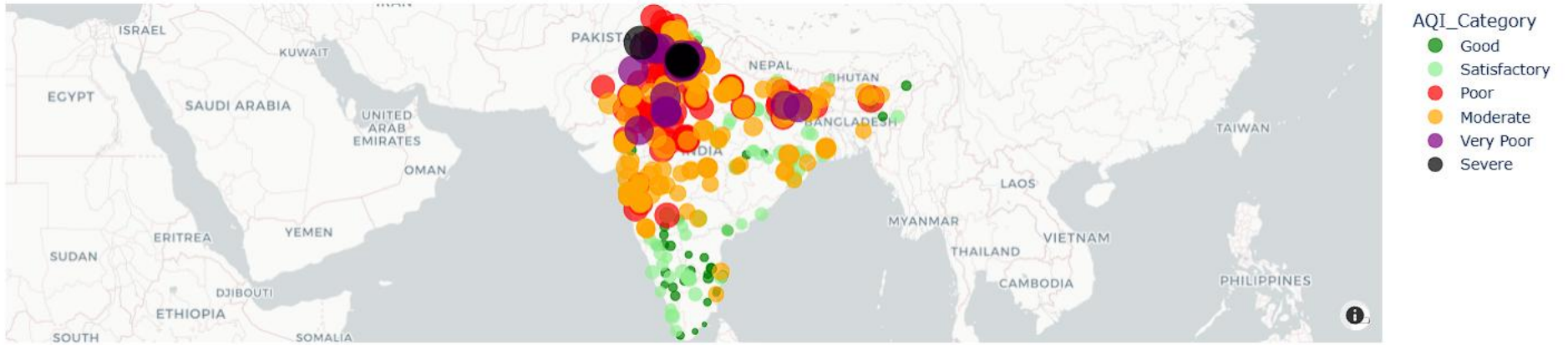
..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available

Indian Cities AQI Dashboard (via Dash & Plotly)

Map | Histogram | Pie Chart | Table

Indian Cities AQI Map



Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:45 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

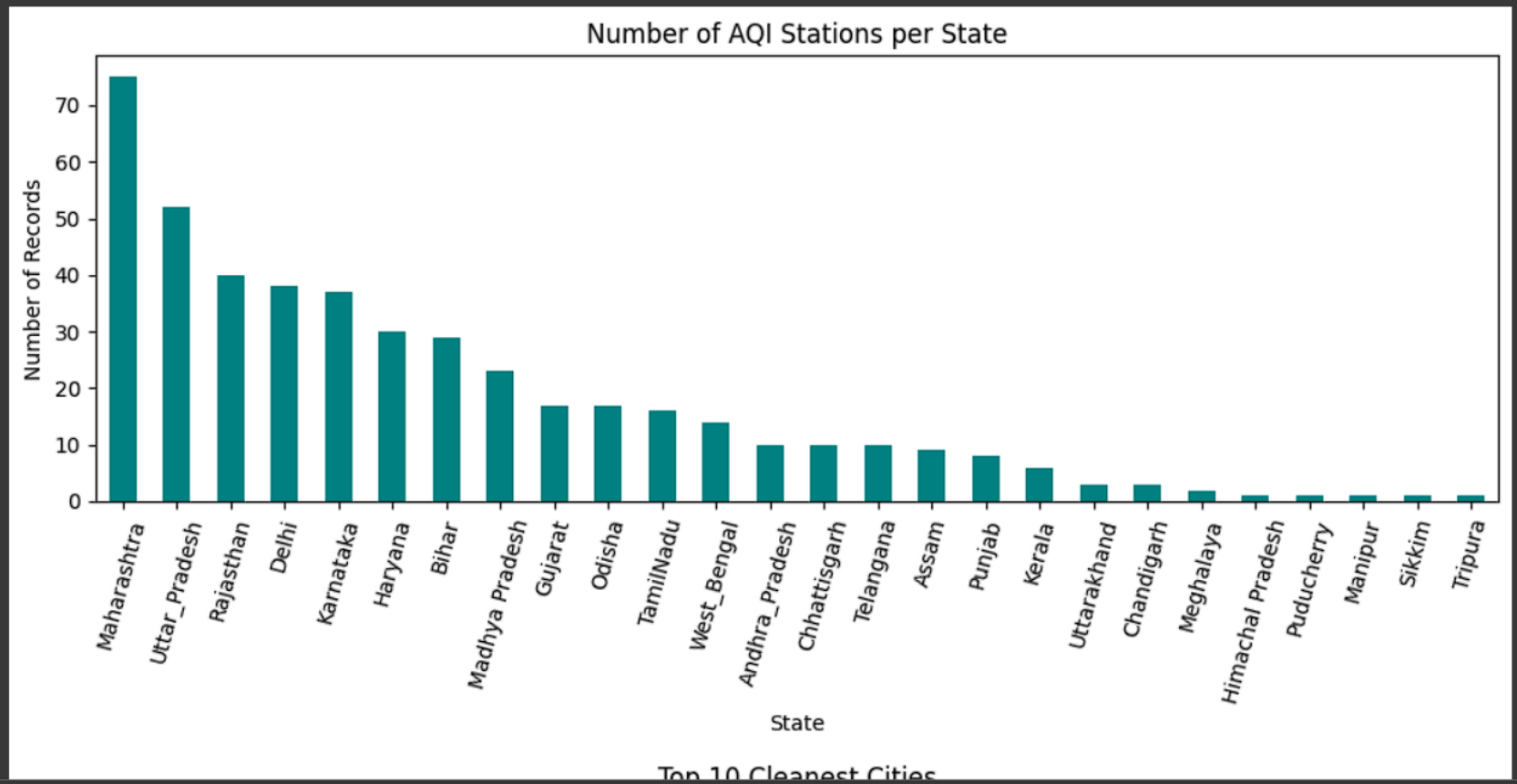
Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html

Disk 69.48 GB available

31s

```
Rows: 454, Columns: 6
Columns: ['City', 'State', 'Latitude', 'Longitude', 'AQI', 'AQI_Category']
```



Variables Terminal

8:43PM Python 3

28°C Haze

Search

ENG IN

20:46 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

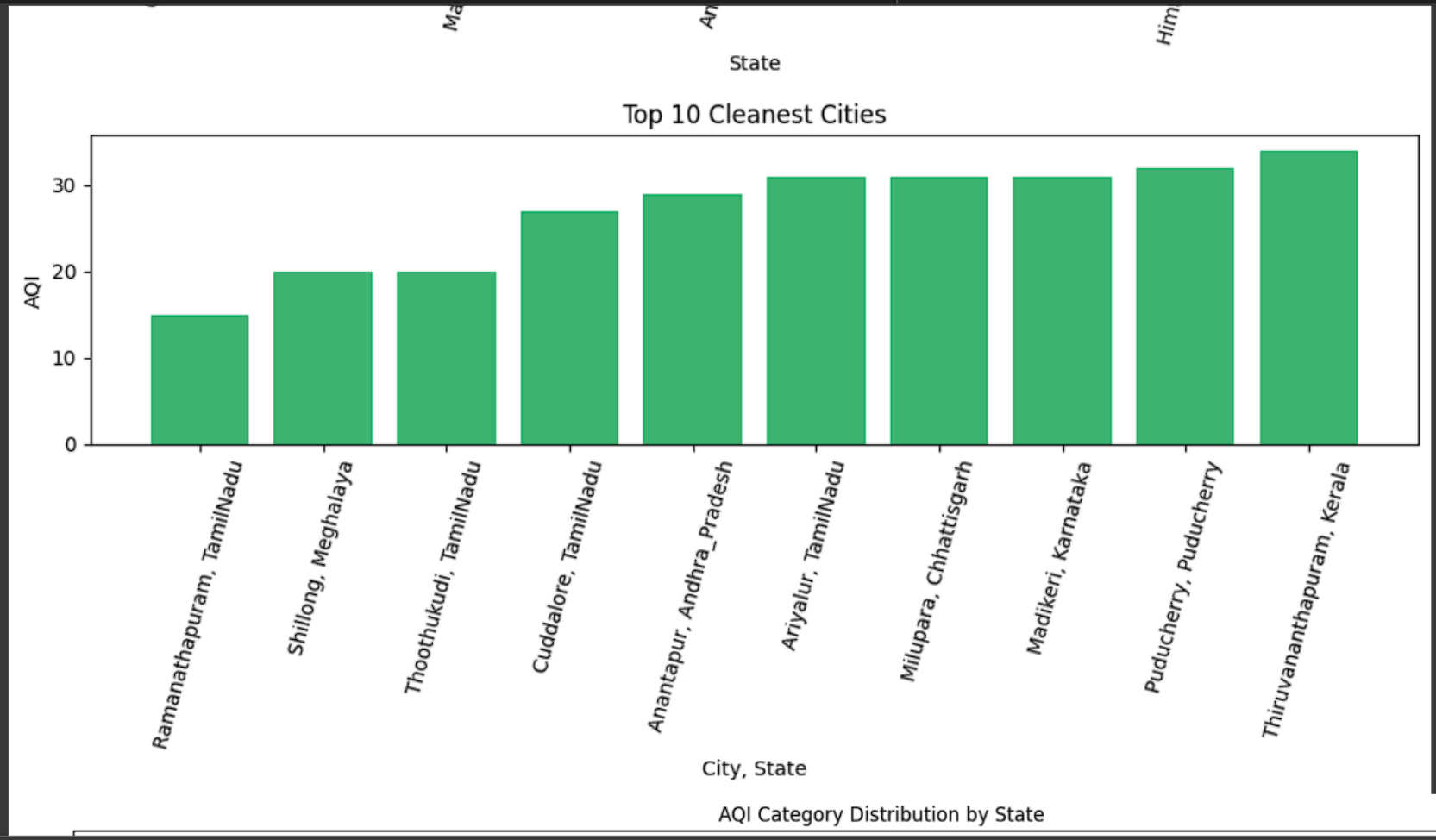
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html

Disk 69.48 GB available



Variables Terminal

8:43PM Python 3

28°C Haze

Search

ENG IN

20:46 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

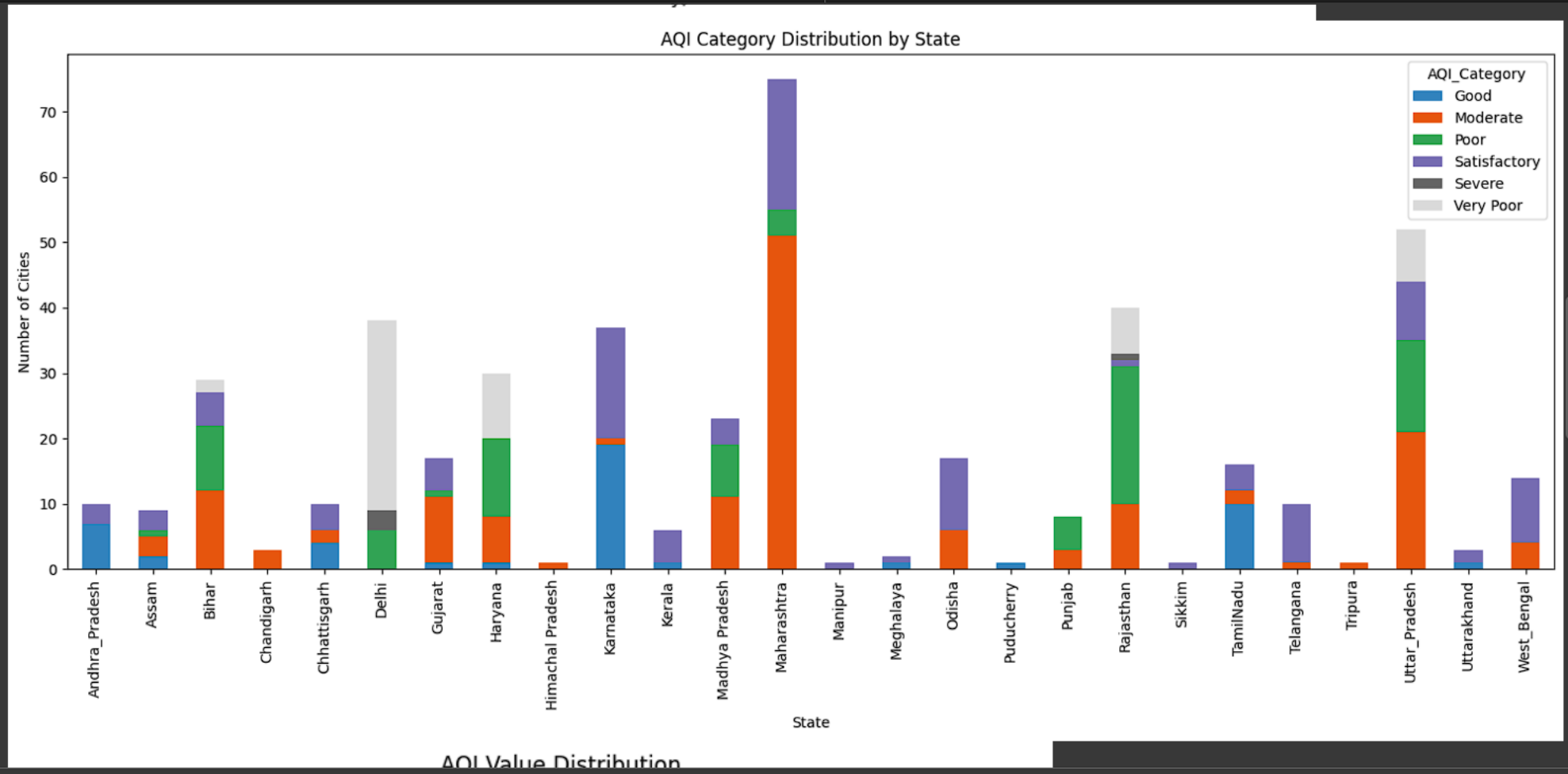
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html

Disk 69.48 GB available



AQI Value Distribution

Variables Terminal

8:43PM Python 3

28°C Haze

Search

ENG IN

20:47 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

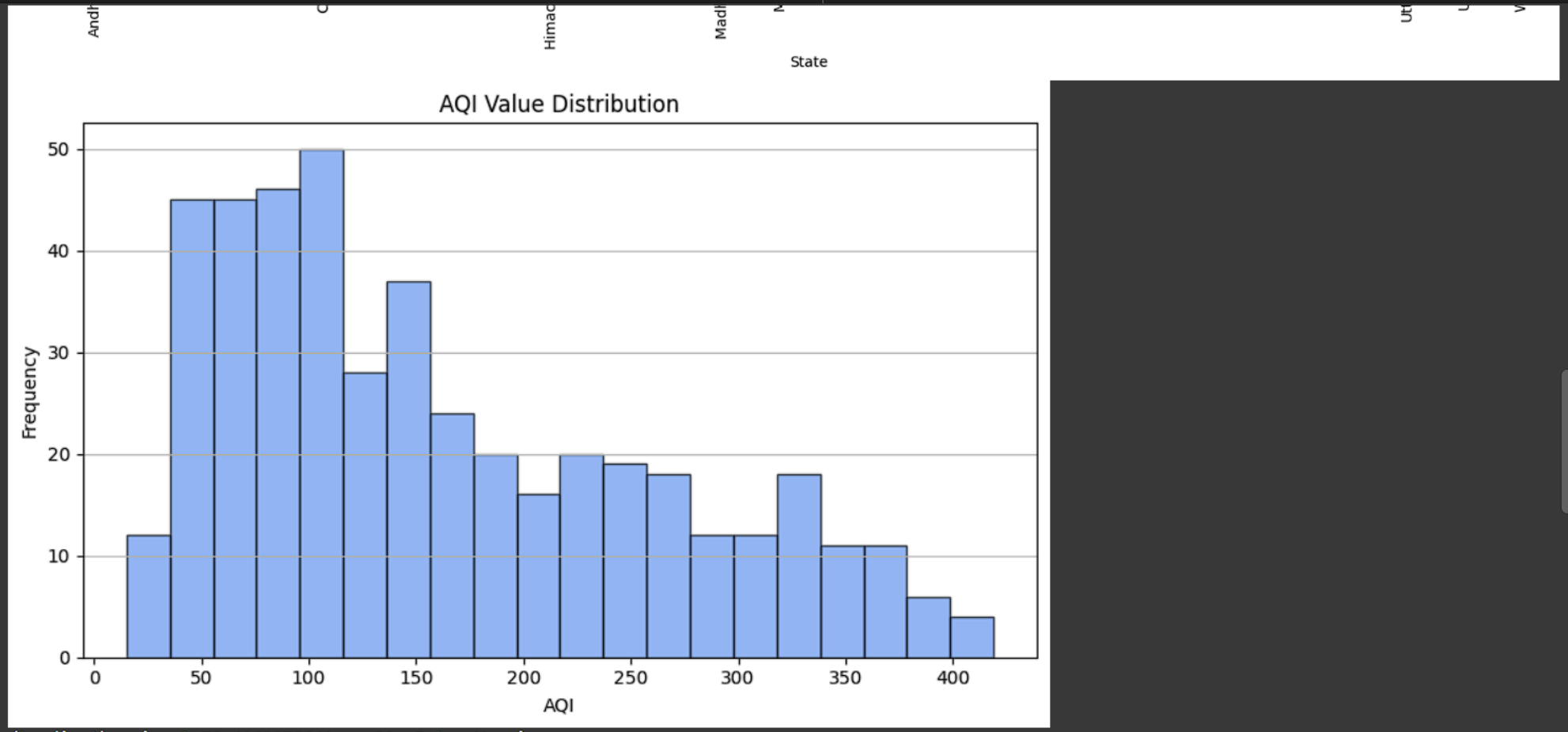
RAM Disk

31s

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html

Disk 69.48 GB available



/tmp/ipython-input-70-4169078641.py:80: FutureWarning:

Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:47 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

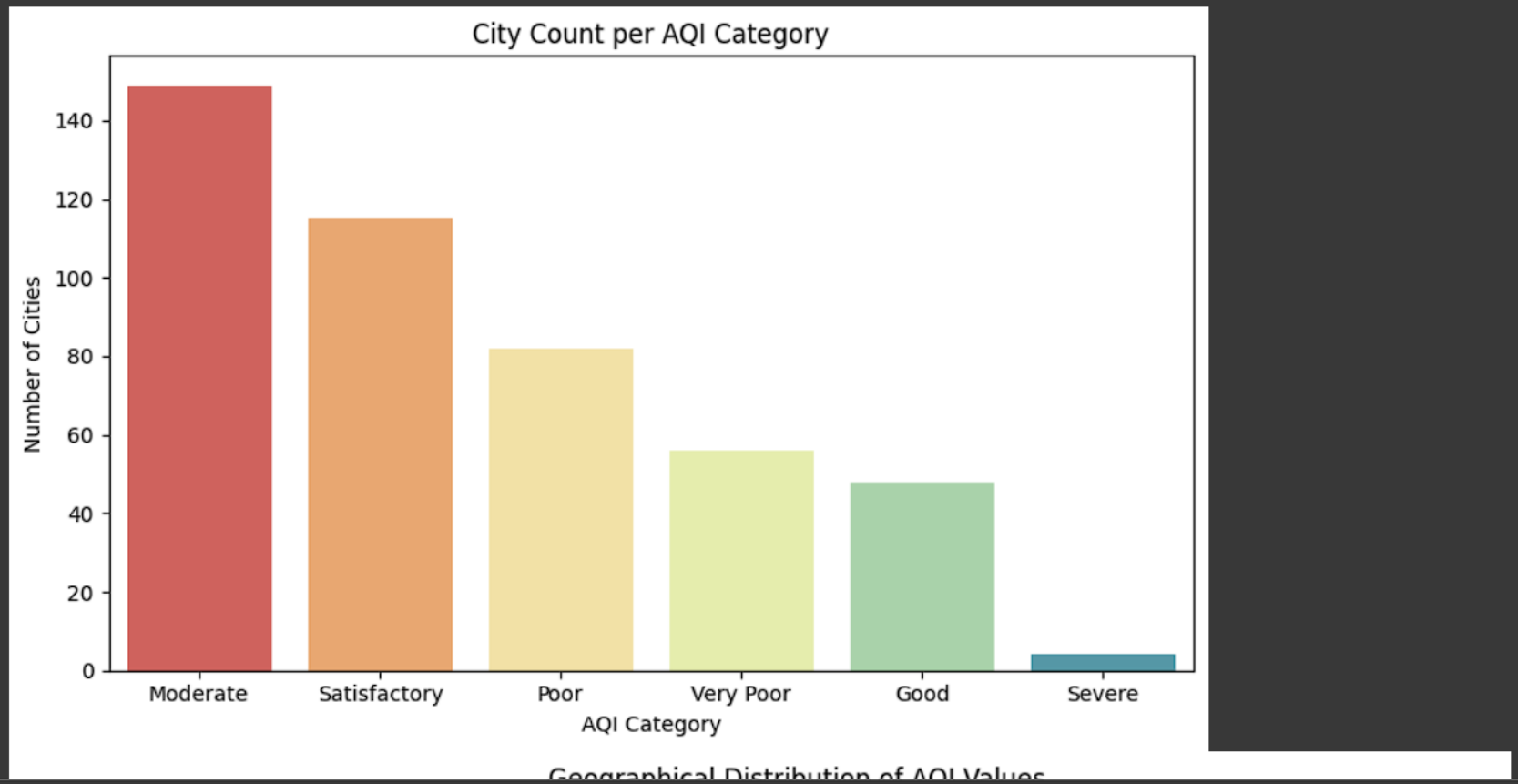
31s

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html

Disk 69.48 GB available

```
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the s
```



Variables Terminal

8:43PM Python 3

28°C Haze

Search

ENG IN

20:47 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

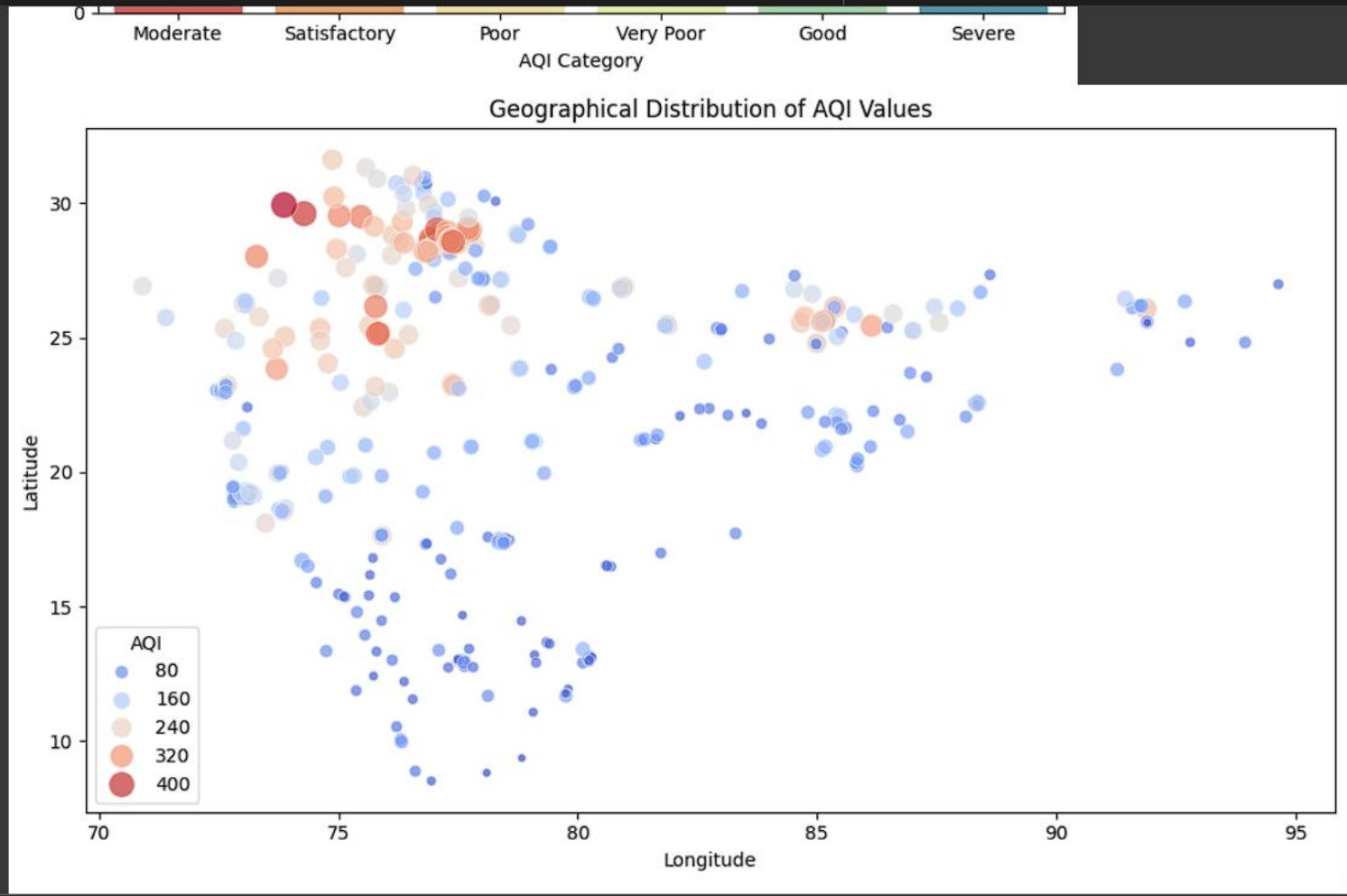
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available



Variables Terminal

8:43 PM Python 3

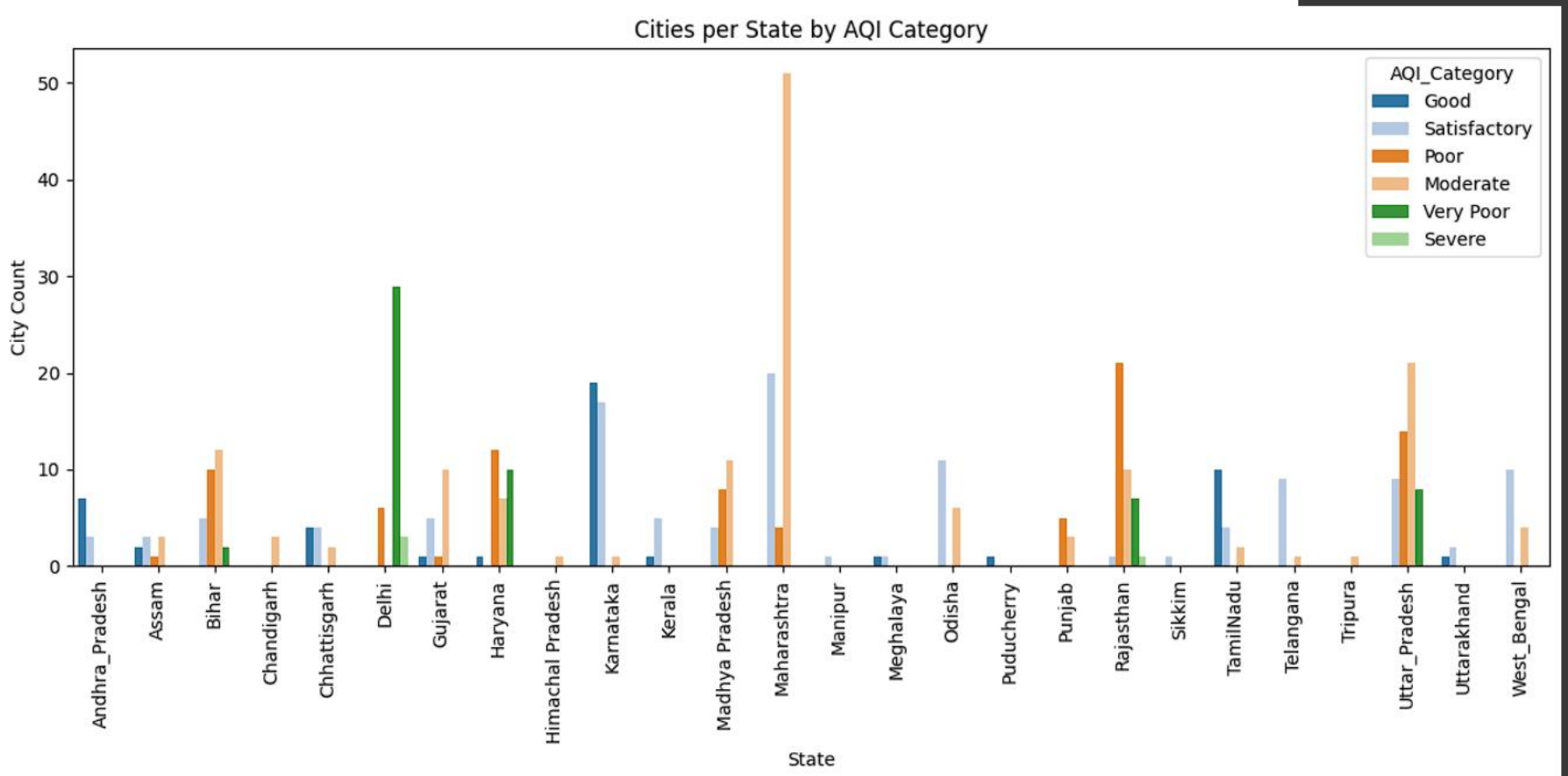
28°C Haze

Search

ENG IN

20:47 29-07-2025

Longitude



HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

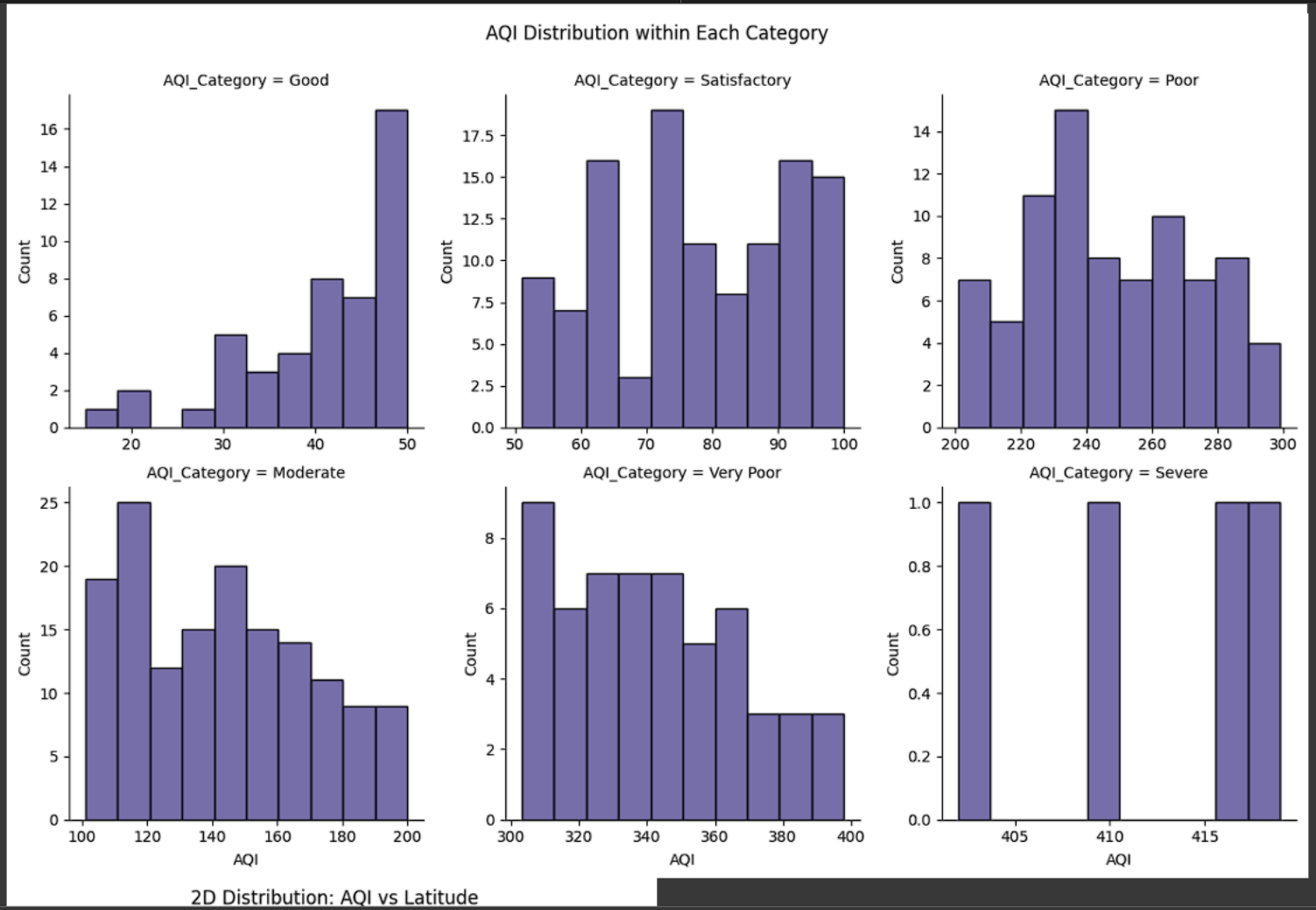
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distribution...
aqi_histogram.png
boxplot_aqi_category.png
city_count_aqi_category....
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_distri...
facet_scatter_aqi_long.p...
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png

Disk 69.48 GB available



Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:48 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=cFbEsHxCNDC-

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

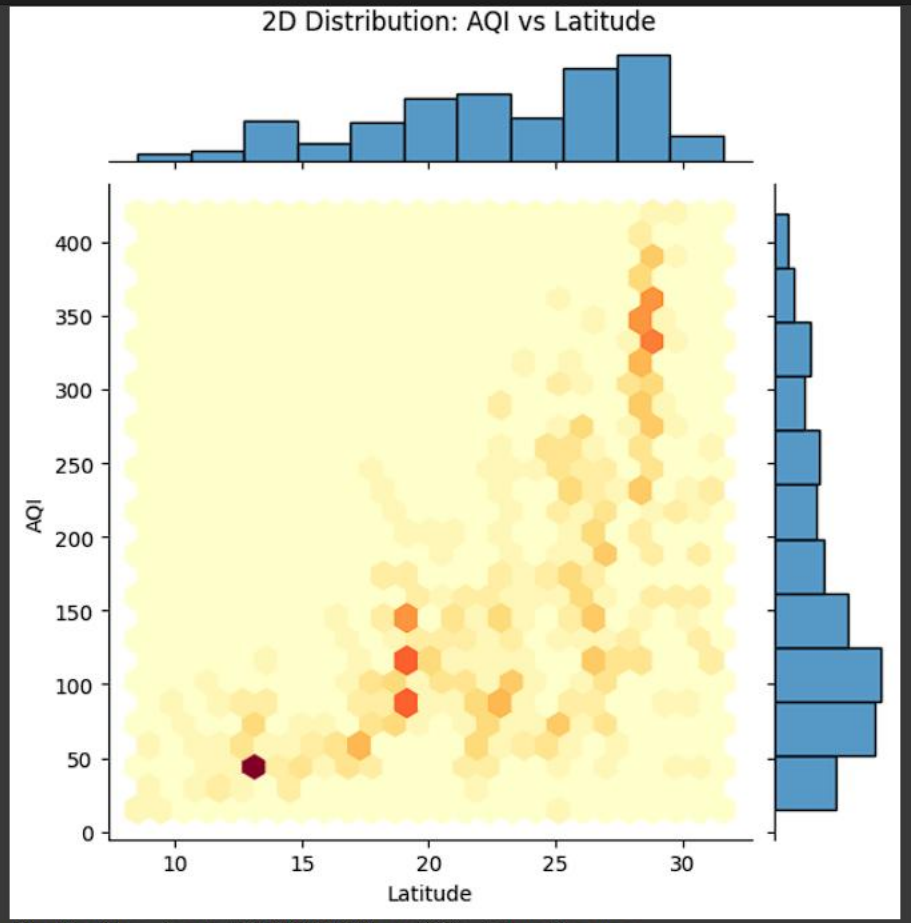
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available



/tmp/ipython-input-70-4169078641.py:125: FutureWarning:

Variables Terminal

8:43 PM Python 3

28°C Haze

Search

ENG IN

20:48 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=S7VfRS3M1Wcm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

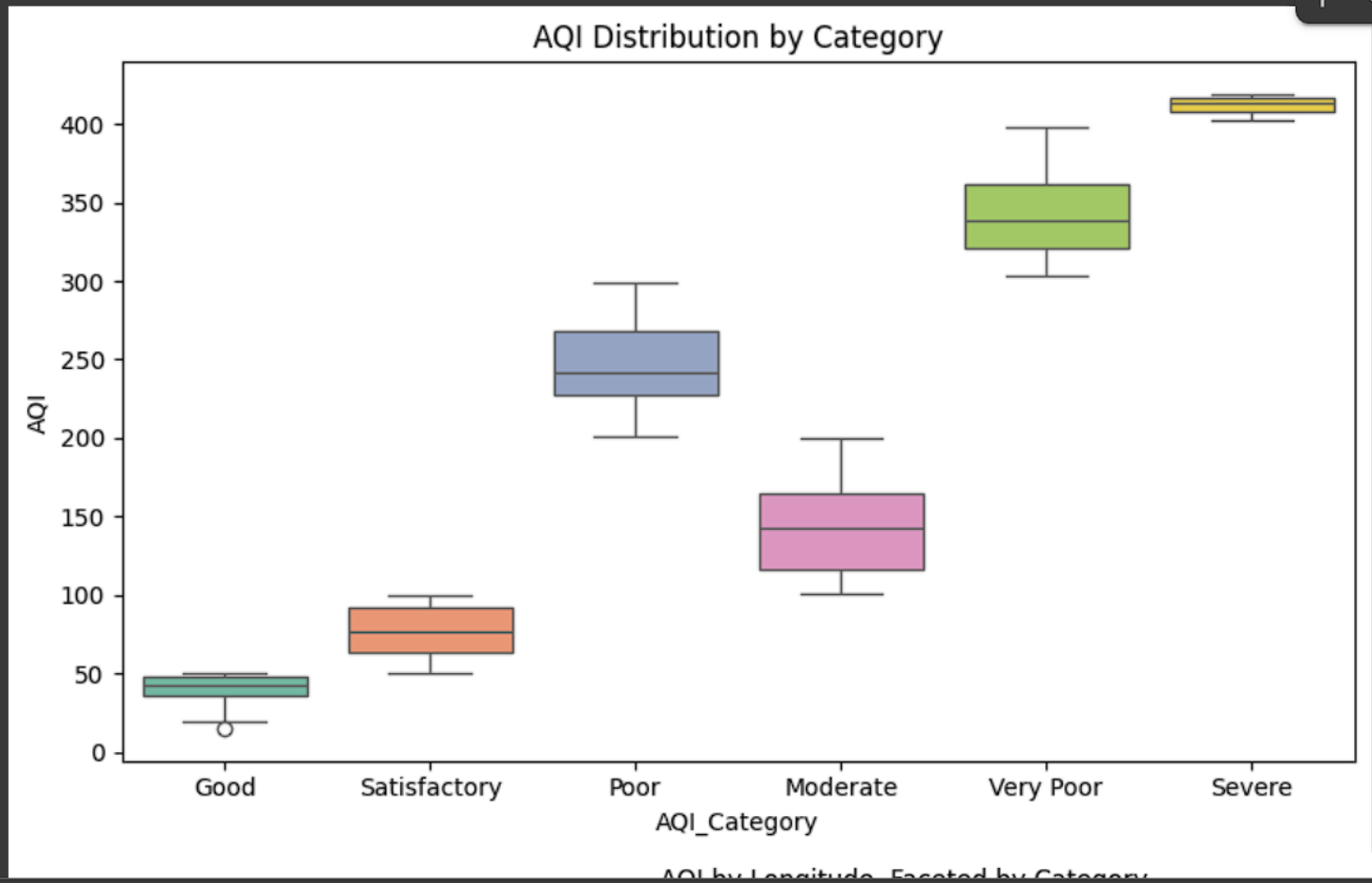
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long....
geo_scatter_aqi.png
india_aqi.csv

Disk 69.48 GB available



AQI by Longitude, Faceted by Category

Variables Terminal

8:43 PM Python 3

28°C Haze

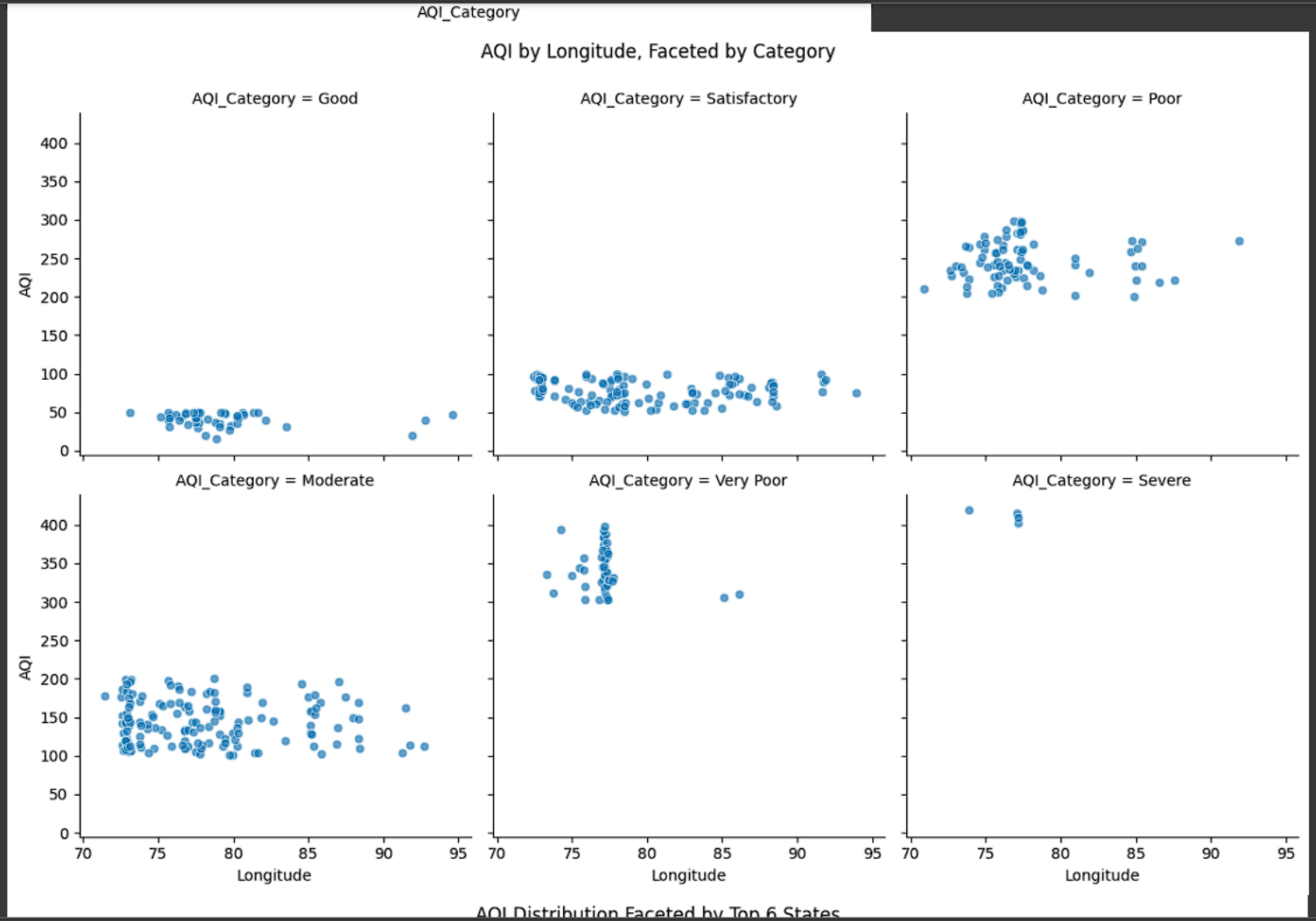
Search

ENG IN

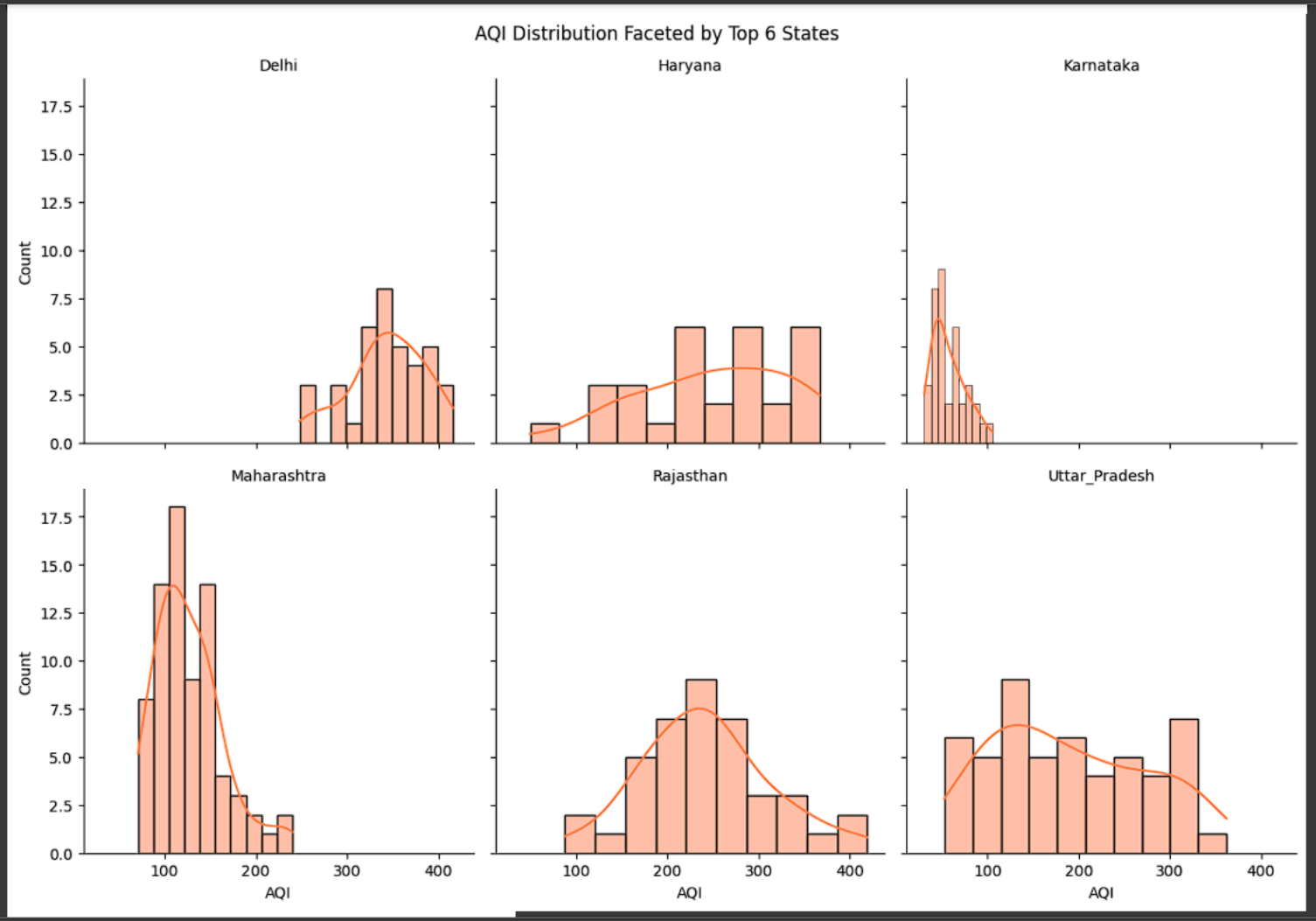
20:48 29-07-2025

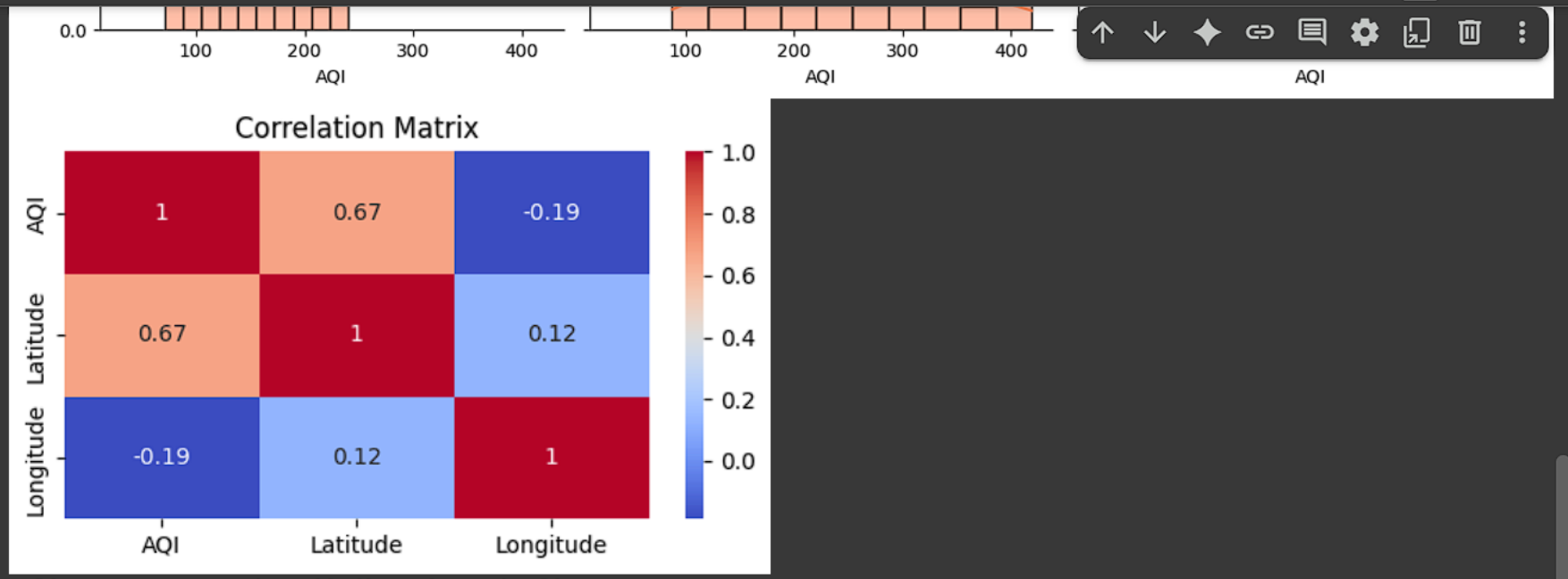
Files
..
sample_data
aqi_category_distributi...
aqi_histogram.png
boxplot_aqi_category.p...
city_count_aqi_categor...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_dis...
facet_scatter_aqi_long...
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.html
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png
state_vs_aqi_cat.png
top_cleanest_cities.png
Disk 69.48 GB available

AQI_Category



AQI Distribution Faceted by Top 6 States






```
Folium AQI map saved as 'india_aqi_map.html'.
Folium AQI heatmap saved as 'india_aqi_heatmap.html'.
\nAll charts saved as PNG images. Both folium maps saved as HTML and can be downloaded for sharing or embedding.
```

Files

- ..
- sample_data
- aqi_category_distribution_state.png
- aqi_histogram.png
- boxplot_aqi_category.png
- city_count_aqi_category.png
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_distribution.png
- facet_scatter_aqi_long.png
- geo_scatter_aqi.png
- india_aqi.csv

Disk 69.48 GB available

```
[70] Start coding or generate with AI.
```

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

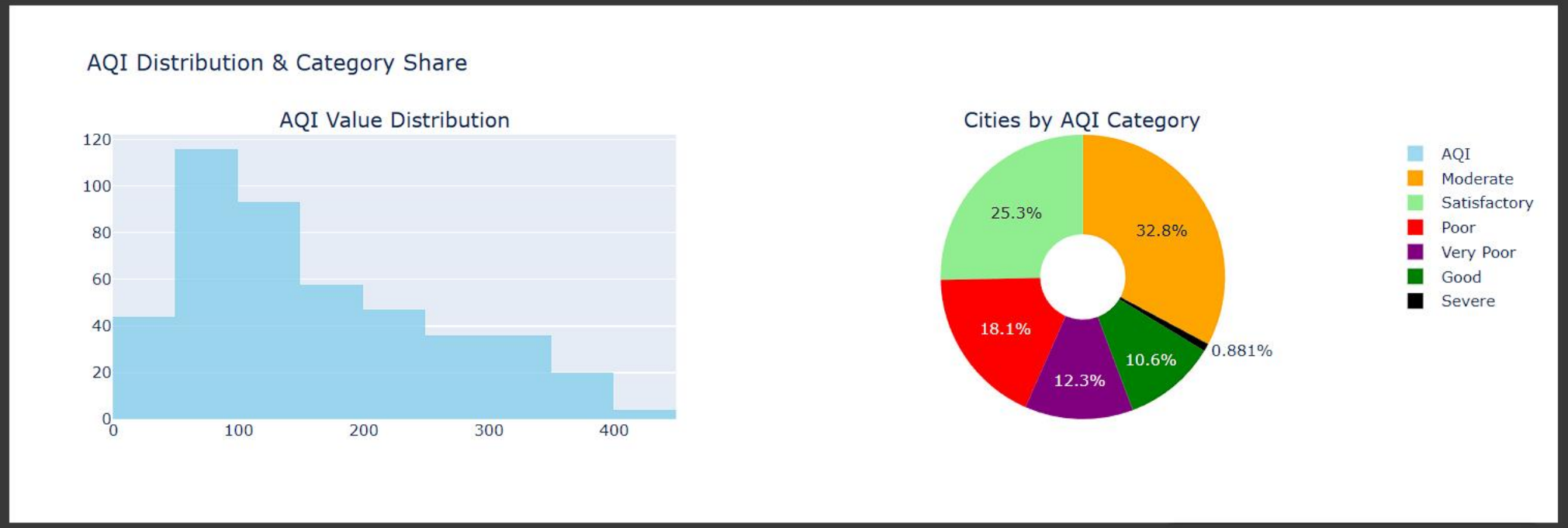
RAM Disk

Files

..
sample_data
aqi_category_distribu...
aqi_histogram.png
boxplot_aqi_category....
city_count_aqi_categ...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_di...
facet_scatter_aqi_lon...
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.ht...
india_aqi_map.html

Disk 69.48 GB available

```
fig.update_layout(height=400, title_text="AQI Distribution & Category Share")
fig.show()
```



5s

```
# ---FINAL DASHBOARD CELL FOR AQI PROJECT IN COLAB---

# 1. Install Plotly (skip if already installed)
```

Variables Terminal

9:20 PM Python 3

1 cm of rain Thursday

Search

ENG IN 21:22 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- aqi_category_distribu...
- aqi_histogram.png
- boxplot_aqi_category....
- city_count_aqi_categ...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_di...
- facet_scatter_aqi_lon...
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.ht...
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png

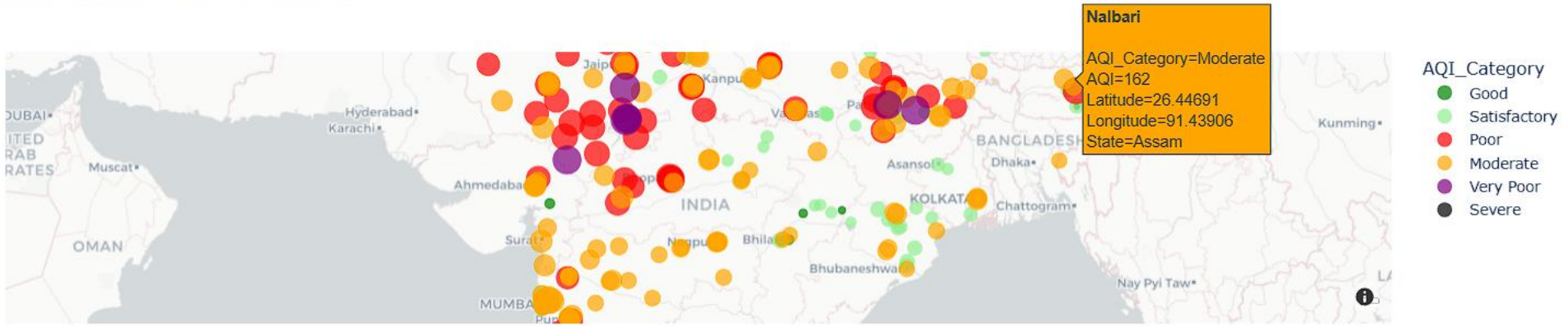
Disk 69.48 GB available

"""))

Project AQI Dashboard

Cities: 244
States: 26
Average AQI: 161.6
Most Polluted: Sri Ganganagar (Rajasthan): AQI 419
Cleanest: Ramanathapuram (TamilNadu): AQI 15

Indian Cities: AQI Map by Category



AQI Distribution & Category Share

Variables Terminal

9:20 PM Python 3

1 cm of rain Thursday

Search

ENG IN 21:22 29-07-2025

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

Commands + Code + Text Run all

RAM Disk

Files

- ..
- sample_data
- aqi_category_distribu...
- aqi_histogram.png
- boxplot_aqi_category....
- city_count_aqi_categ...
- correlation_matrix.png
- data_aqi_cpcb.xml
- facet_aqi_by_state.png
- facet_aqi_category_di...
- facet_scatter_aqi_lon...
- geo_scatter_aqi.png
- india_aqi.csv
- india_aqi_heatmap.ht...
- india_aqi_map.html
- joint_lat_aqi.png
- records_per_state.png

Disk 69.48 GB available

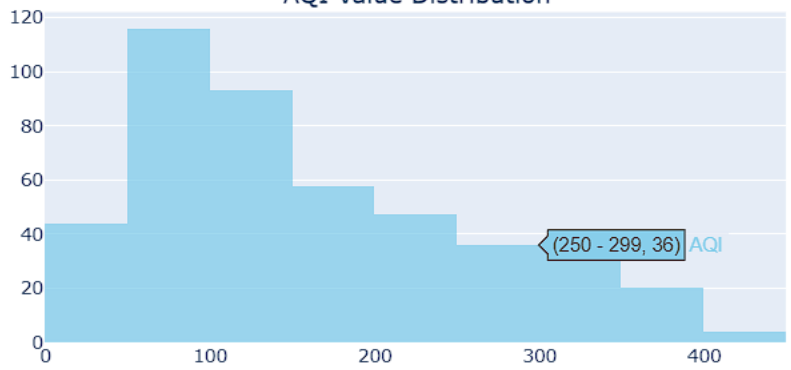
"""))

MUMBA

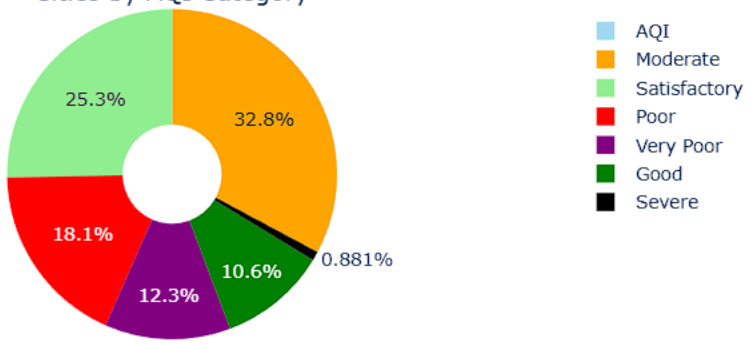
Longitude=85.83678
State=Odisha

AQI Distribution & Category Share

AQI Value Distribution



Cities by AQI Category



City Rankings by AQI

Top 10 Most Polluted Cities

Top 10 Cleanest Cities

Variables Terminal

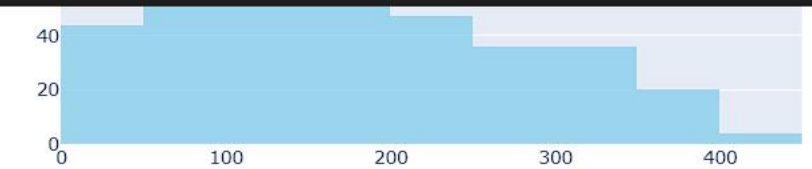
9:20 PM Python 3

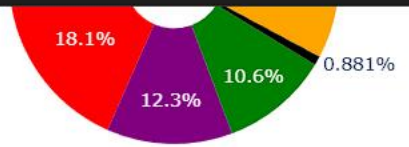
1 cm of rain Thursday

Search

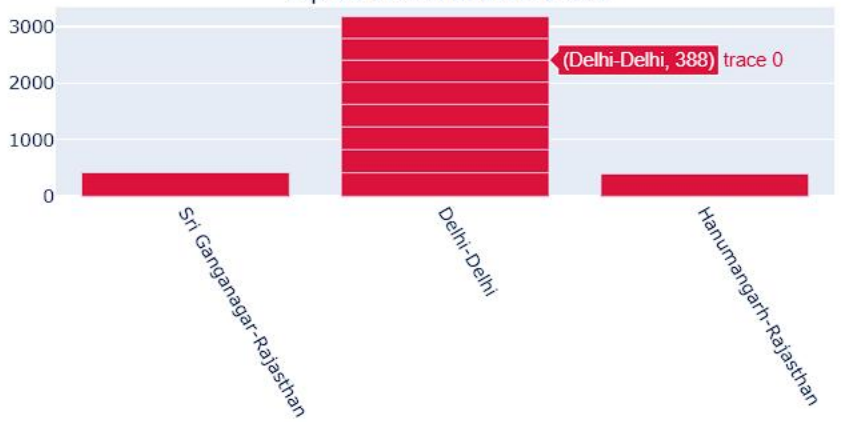
ENG IN

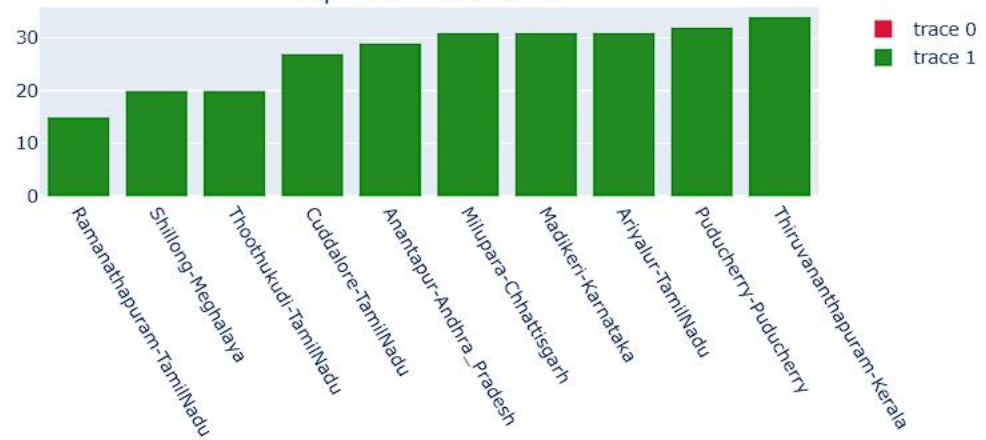
21:23 29-07-2025





City Rankings by AQI





AQI Category Distribution by State

HCL+GUVI_Project-1.ipynb - Co

colab.research.google.com/drive/1pJcC1wpmAxnKKzuPl23ovWL1liotzhpn#scrollTo=C7uoM2om2ycm

HCL+GUVI_Project-1.ipynb

File Edit View Insert Runtime Tools Help

Share Gemini

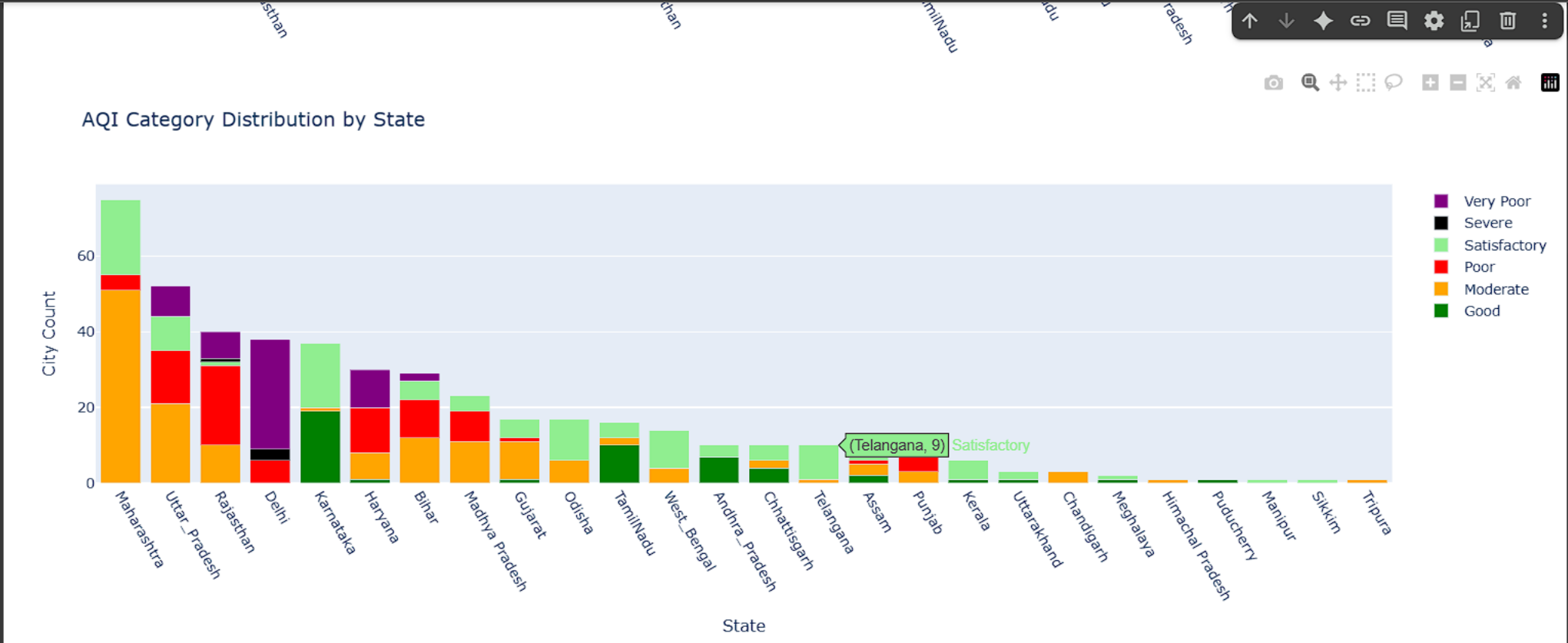
Commands + Code + Text Run all

RAM Disk

Files

..
sample_data
aqi_category_distribu...
aqi_histogram.png
boxplot_aqi_category....
city_count_aqi_categ...
correlation_matrix.png
data_aqi_cpcb.xml
facet_aqi_by_state.png
facet_aqi_category_di...
facet_scatter_aqi_lon...
geo_scatter_aqi.png
india_aqi.csv
india_aqi_heatmap.ht...
india_aqi_map.html
joint_lat_aqi.png
records_per_state.png

Disk 69.48 GB available



*Dashboard sections:

- KPIs
- Interactive AQI map
- Distribution & share
- Top/bottom 10 cities
- Statewise category breakdown*

Variables Terminal

9:20 PM Python 3

1 cm of rain Thursday

Search

ENG IN 21:23 29-07-2025

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

full-crews-smell.loca.lt

Air Quality Index (AQI) Dashboard - Indian Cities

AQI Data Table

Page Number

1

	City	State	Latitude	Longitude	AQI	AQI_Category
0	Anantapur	Andhra_Pradesh	14.6759	77.593	29	Good
1	Chittoor	Andhra_Pradesh	13.2049	79.0979	35	Good
2	Kadapa	Andhra_Pradesh	14.4651	78.8242	37	Good
3	Rajamahendravaram	Andhra_Pradesh	16.9873	81.7363	58	Satisfactory
4	Tirupati	Andhra_Pradesh	13.67	79.35	49	Good
5	Tirupati	Andhra_Pradesh	13.6154	79.4092	48	Good
6	Vijayawada	Andhra_Pradesh	16.5361	80.5942	50	Good
7	Vijayawada	Andhra_Pradesh	16.4867	80.6994	47	Good

Dashboard Visualizations

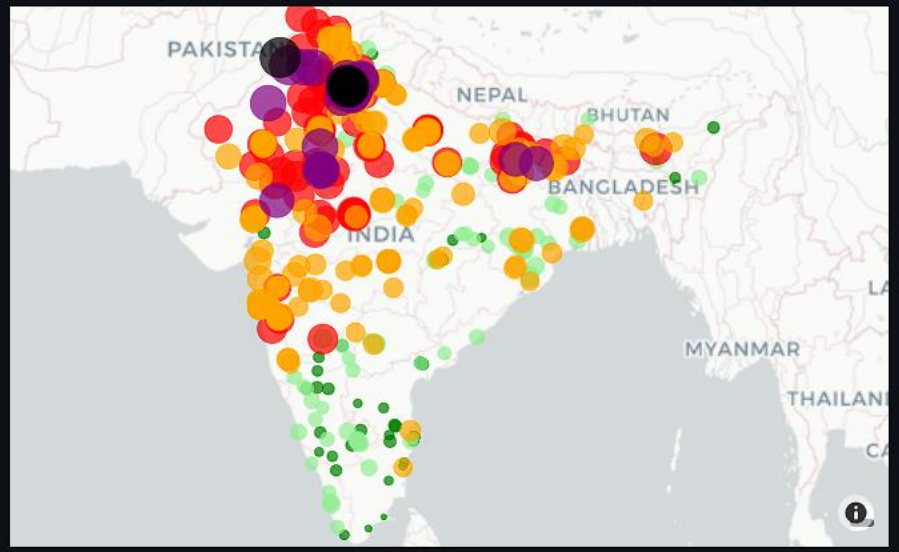
HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

full-crews-smell.loca.lt

Dashboard Visualizations

Total Cities	Average AQI	Most Polluted City
244	161.6	Sri Gangana...

Cleanest City: Ramanathapuram (15)



27°C Haze

Search

ENG IN

22:30 29-07-2025

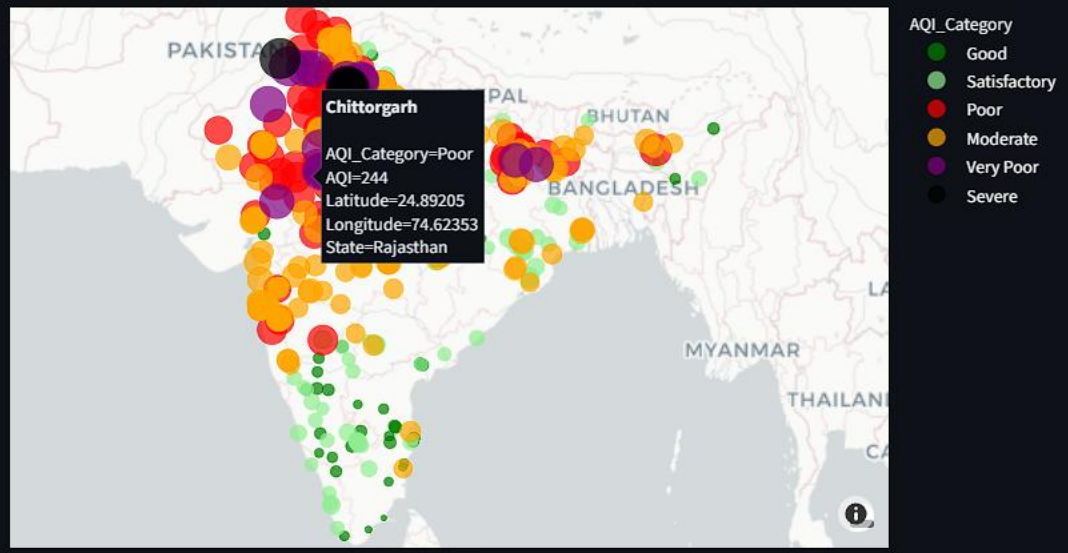
HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

full-crews-smell.loca.lt

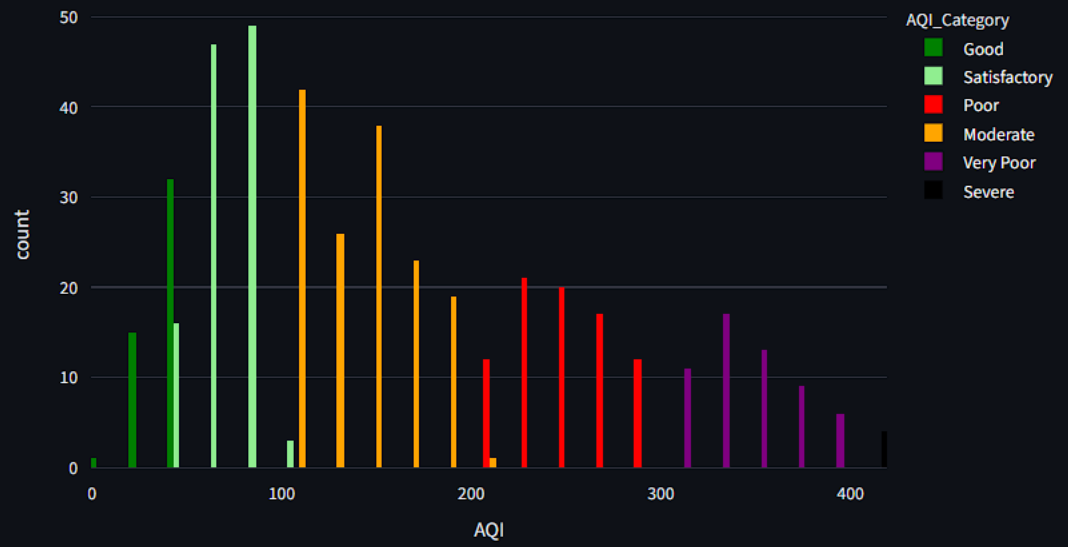
Dashboard Visualizations

Total Cities	Average AQI	Most Polluted City
244	161.6	Sri Gangana...

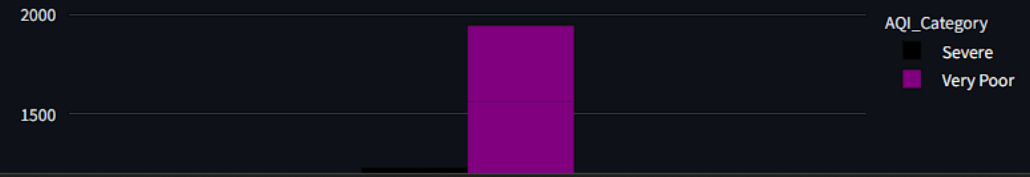
Cleanest City: Ramanathapuram (15)



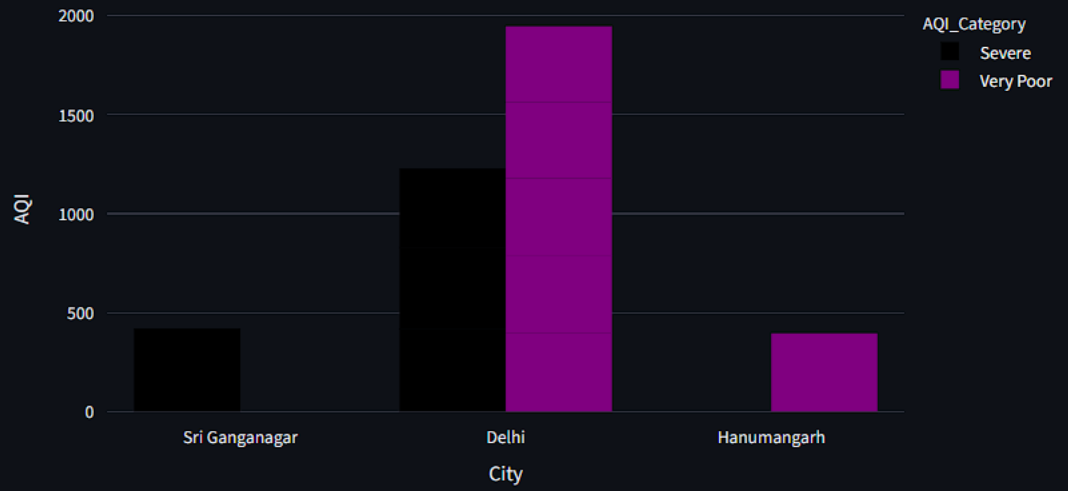
27°C Haze | Search | ENG IN | 22:30 29-07-2025



Top 10 Most Polluted Cities



Top 10 Most Polluted Cities

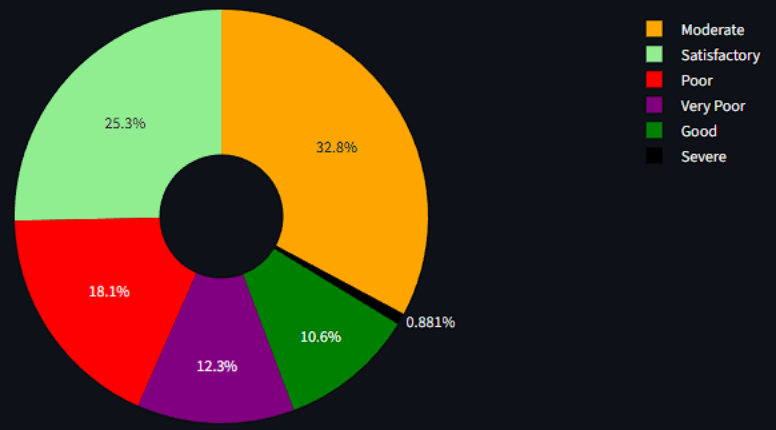


Top 10 Cleanest Cities

HCL+GUVI_Project-1.ipynb - Co | Streamlit | app.py - Colab

full-crews-smell.loca.lt

Top 10 Cleanest Cities



27°C Haze | Search | ENG IN | 22:30 29-07-2025


Thank You...