

AATISH AROTE

◇ +91 862-307-1089 ◇ aatisharote07@gmail.com ◇ linkedin.com/in/aatisharote/ ◇ aatisharote.com ◇ Mumbai, India

OBJECTIVE

Aspiring **Data Analyst** skilled in Python, SQL, ML pipelines, and Power BI dashboards. Seeking an entry-level role to drive insights and support data-driven decisions.

EXPERIENCE

Data Visualization Intern
Infosys Springboard

Sep 2025 – Dec 2025
Remote

- **Engineered** an advanced **Power BI** dashboard to analyze Global Tourism data (ETL, DAX); selected as the **top design** in the initial phase, for superior data modeling and visualization of seasonal trends.
- **Developed** a scalable **Flask** web application to embed the dashboard, integrating **Python** analytics; the solution was chosen to represent the **entire batch of 20 interns** over 5 competing teams.
- **Led** the final stakeholder presentation and technical walkthrough, orchestrating the live demo of AI-driven insights and validating the application flow for the official batch showcase.

SKILL SET

- **Programming Languages:** Python, SQL
- **Libraries:** NumPy, Pandas, Scikit-Learn
- **Data Visualization:** Power BI, Tableau, Matplotlib, Seaborn
- **Tools:** Jupyter Notebook, Git/GitHub, Microsoft Excel
- **Databases:** MySQL
- **Cloud:** AWS (EC2, RDS), Azure
- **Certifications:** AZ-900 (Microsoft), AWS PartyRock Hackathon, Infosys Springboard Internship

PROJECTS

Real Estate Analytics

Flask, Scikit-Learn, AWS, Python ([Link](#))

- **Built an end-to-end data pipeline** processing **4,125+ raw listings** across 150+ sectors; engineered **11 bespoke features** (including "Luxury Score" and KMeans-clustered furnishing) to optimize the dataset for high-dimensional modeling.
- **Benchmarked 11 regression algorithms** (inc. XGBoost, Lasso) using 10-fold CV; tuned a **Random Forest Regressor** via GridSearchCV to achieve an **R squared of 0.9012** (a **22.4% improvement** over baseline).
- **Deployed** a production-grade web application on **AWS**, integrating the serialized model pipeline with three custom recommendation engines and interactive geospatial maps for real-time decision support.

Tourism AI

Flask, Power BI, Python, SQL ([Link](#))

- **Architected** a refined preprocessing pipeline with normalization, aggregation, and feature engineering to optimize complex global datasets for Power BI.
- **Developed** three interactive dashboards for 232 countries using heat maps, scatter plots, and time-series charts to visualize tourism metrics.
- **Implemented** a unified analytics platform by embedding interactive Power BI dashboards into a Flask web application for real-time exploration of global tourism trends.

National Air Quality Analysis

Python, Pandas, Seaborn, Matplotlib ([Link](#))

- **Conducted** a comprehensive analysis of real-time Air Quality Index (AQI) data sourced from data.gov.in, focusing on **7** key pollutants including PM2.5, NO2, and Ozone across various Indian states.
- **Executed** Exploratory Data Analysis (EDA) to uncover regional pollution trends, utilizing Seaborn and Matplotlib to generate boxplots and histograms that visualized the distribution and spread of pollutant levels.

EDUCATION

B.E. in Computer Engineering, University of Mumbai
Current SGPI: **8.43/10**

Aug 2022 – May 2026

HSC (12th) - 67.67%

Mar 2022

SSC (10th) - 92.20%

Mar 2020