Vehicle Registration Analysis in India (2009-2020)

---

1. Project Title

Vehicle Registration Analysis in India (2009-2020)

2. Dataset

The dataset used is the Indian Vehicle Registration Dataset from Kaggle. It contains vehicle registration data for different states and vehicle types in India from 2009 to 2015.

3. Project Objective

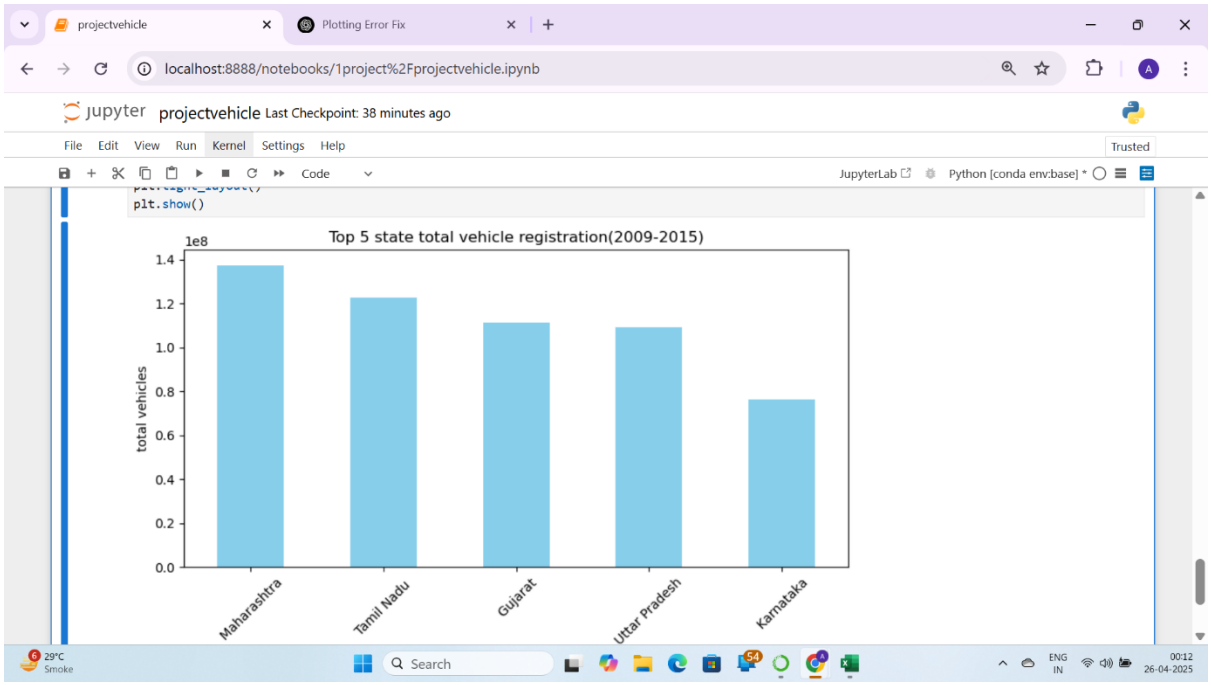
- **Primary Goal**: Identify the five states with the highest number of registered vehicles and determine which vehicle type (2-wheelers, 4-wheelers, etc.) had the highest growth between 2009 and 2015.
- **Secondary Goal**: Predict the number of vehicle registrations in Maharashtra for the years 2016 to 2020, focusing on two-wheelers.

4. Data Preprocessing

- **Handling Missing Data**: Checked for missing values it is filled by fi.
- **Data Cleaning**: Removed irrelevant columns and ensured consistent formats.
- **Feature Engineering**: Created a dictionary for two-wheeler counts across years.

5. Exploratory Data Analysis (EDA)

Top 5 States with Most Registered Vehicles (2015):

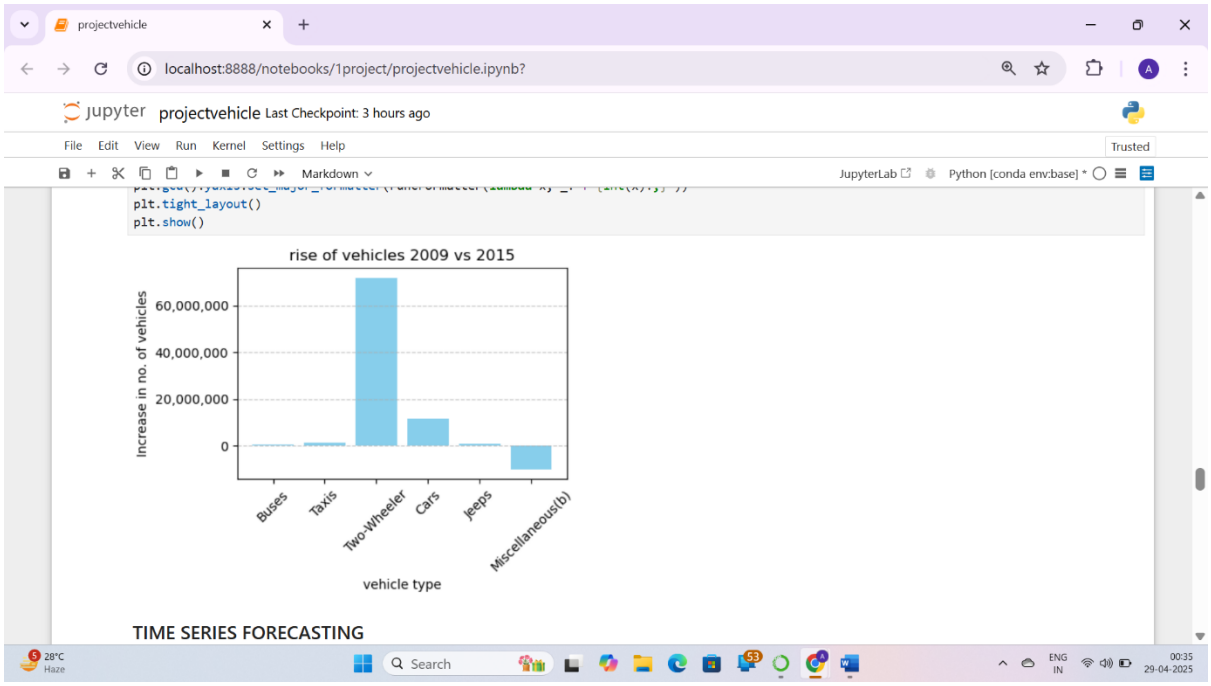


1. Maharashtra: 137,529,842
2. Tamil Nadu: 122,617,415

3. Gujarat: 111,198,964
4. Uttar Pradesh: 109,298,346
5. Karnataka: 76,307,517

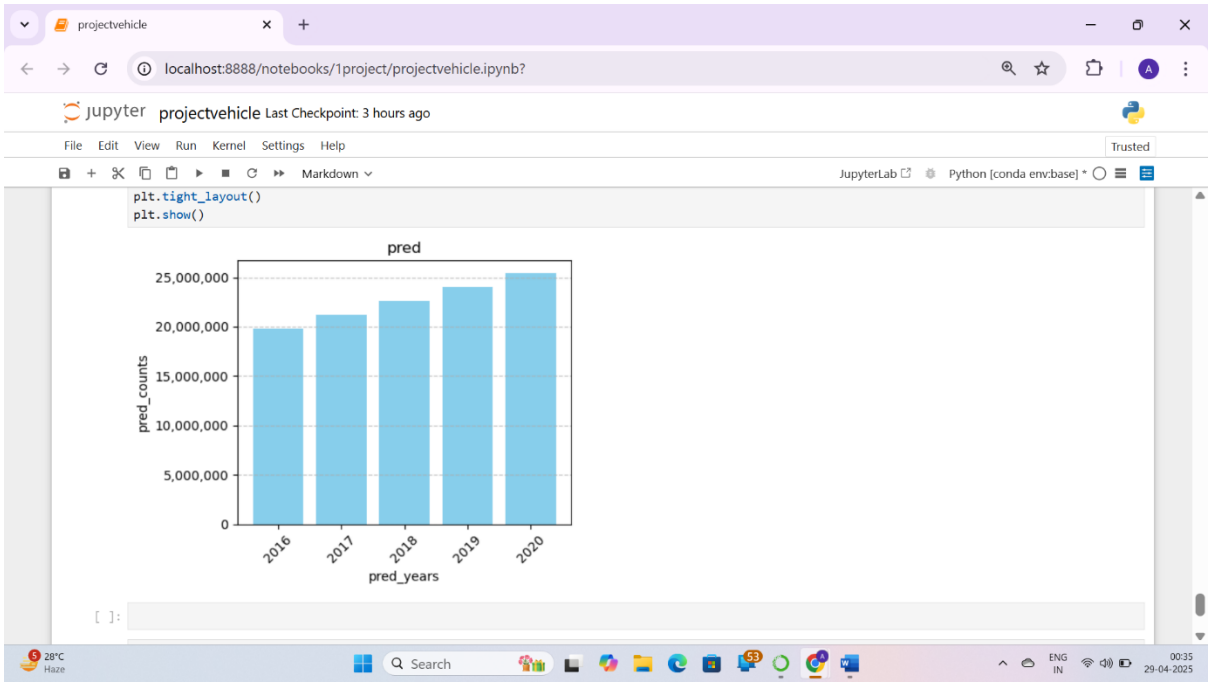
Growth Analysis (2009-2015)

- **Two-wheelers** had the most significant growth.
- **Method**: Calculated absolute growth between 2009 and 2015 using a for-loop across vehicle types.



6. Modeling and Analysis

- **Model Used**: Linear Regression (using scikit-learn)
- **Data**: Two-wheeler counts from 2009 to 2015
- **Prediction**: Number of two-wheeler registrations in Maharashtra from 2016 to 2020



Year	Predicted Two-Wheeler Registrations
2016	19,768,007
2017	21,184,429
2018	22,600,850
2019	24,017,272
2020	25,433,694

7. Conclusion

- Maharashtra dominates the vehicle registration market in India.
- Two-wheelers show strong and consistent growth.
- Projections suggest continued growth in two-wheeler registrations in Maharashtra.

8. Tools and Libraries Used

- **Python Libraries**:
  - Pandas (data handling)
  - Matplotlib (visualizations)
  - Scikit-learn (linear modeling)

9. Limitations and Future Work

- **Limitations**:
  - Dataset ends at 2015.

    - External factors like economy, government policies not considered.

- **Future Work**:
    - Update model with recent data.
    - Analyze the impact of external factors.
    - Expand to other vehicle types and states.

Project By: Aditi Singh

Dataset Source: Kaggle - Indian Vehicle Registration Dataset