Project Semester January–April 2025

DATA SCIENCE MINOR PROJECT REPORT ON

A Dashboard on Electric Vehicle population Data

INRODUCTION TO DATA MANAGEMENT

COURSE CODE: INT217

B. TECH COMPUTER SCIENCE AND ENGINEERING

LOVELY PROFESSIONAL UNIVERSITY

PHAGWARA, PUNJAB

PROJECT SUBMITTED BY:

Prashanth Hosamani(12303950)

Section: K23GW

Roll No.:57

PROJECT SUBMITTED TO:

Baljinder Kaur

CERTIFICATE

This is to certify that Prashanth bearing Registration no. 12303950 has complete INT217 project titled, **"Electric Vehicle Population"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor Designation of the Supervisor

School of Computer Science and Engineering
Lovely Professional University Phagwara, Punjab.

Date: 13/04/2025

DECLARATION

I, Prashanth, student of BTech CSE under CSE/IT Discipline at, Lovely Professional University, Punjab, here by declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 13/04/2025 Signature: - Prashanth

Registration No:- 12303950

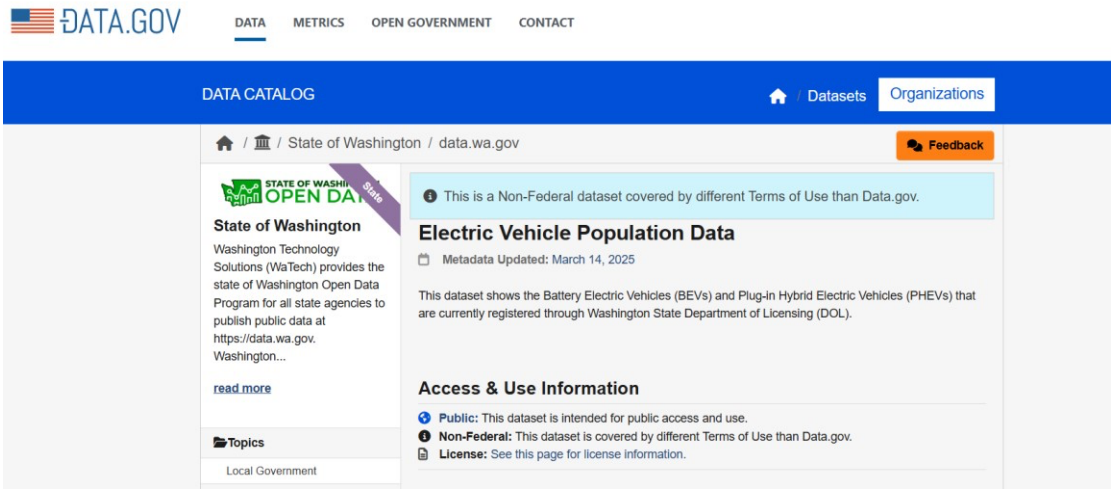
Name of the Student:- Prashanth

1. INTRODUCTION:

2. • **Total EVs on the Road**: By the end of 2023, approximately **40 million electric cars** were in use worldwide.
3. • **Annual Sales**: In 2023, nearly **14 million new electric cars** were registered globally, marking a 35% increase from the previous year.
4. • **Regional Distribution**: China, Europe, and the United States accounted for about **95% of global EV sales** in 2023.

5. SOURCE OF DATASET:

https://catalog.data.gov/dataset/electric-vehicle-population-data



6. DATASET PRE-PROCESSING:

- **Inspect the Data**: Examine the dataset's shape, data types, and summary statistics to understand its structure and contents.

- **Identify Missing Values**: Detect columns with missing or null values. GitHub
- **Check for Duplicates**: Identify and address any duplicate records that may skew analysis.

🧹 2. Data Cleaning

- **Handle Missing Values**: Depending on the context, fill missing values using appropriate strategies such as mean, median, or mode imputation. GitHub
- **Remove Duplicates**: Eliminate any duplicate rows to maintain data integrity.
- **Correct Inconsistencies**: Standardize categorical entries to ensure uniformity (e.g., converting all entries to lowercase).
- **Convert Data Types**: Ensure that each column has the appropriate data type (e.g., dates in datetime format, numerical columns as integers or floats).

⚙️ 3. Feature Engineering

- **Extract New Features**: Derive new columns from existing data, such as extracting the year from a date or calculating the age of a vehicle.
- **Encode Categorical Variables**: Transform categorical variables into numerical format using techniques like one-hot encoding or label encoding.
- **Normalize Numerical Features**: Scale numerical features to a standard range using normalization or standardization techniques.

📊 4. Exploratory Data Analysis (EDA)

- **Visualize Distributions**: Use histograms, box plots, or density plots to understand the distribution of numerical features.
- **Analyze Relationships**: Employ scatter plots or correlation matrices to explore relationships between variables.
- **Geospatial Analysis**: If location data is available, create maps to visualize geographic distributions of EV registrations.

4. ANALYSIS ON DATASET (FOR EACH ANALYSIS)

Objective 1: School Demographics and Location Analysis

General Description

This objective focuses on identifying countries that have experienced the highest number of terrorist attacks. It helps highlight global hotspots where terrorism is most prevalent and persistent, offering insights into the geopolitical conditions that may contribute to frequent attacks. This forms the basis for comparative global studies and risk assessments.

Specific Requirements

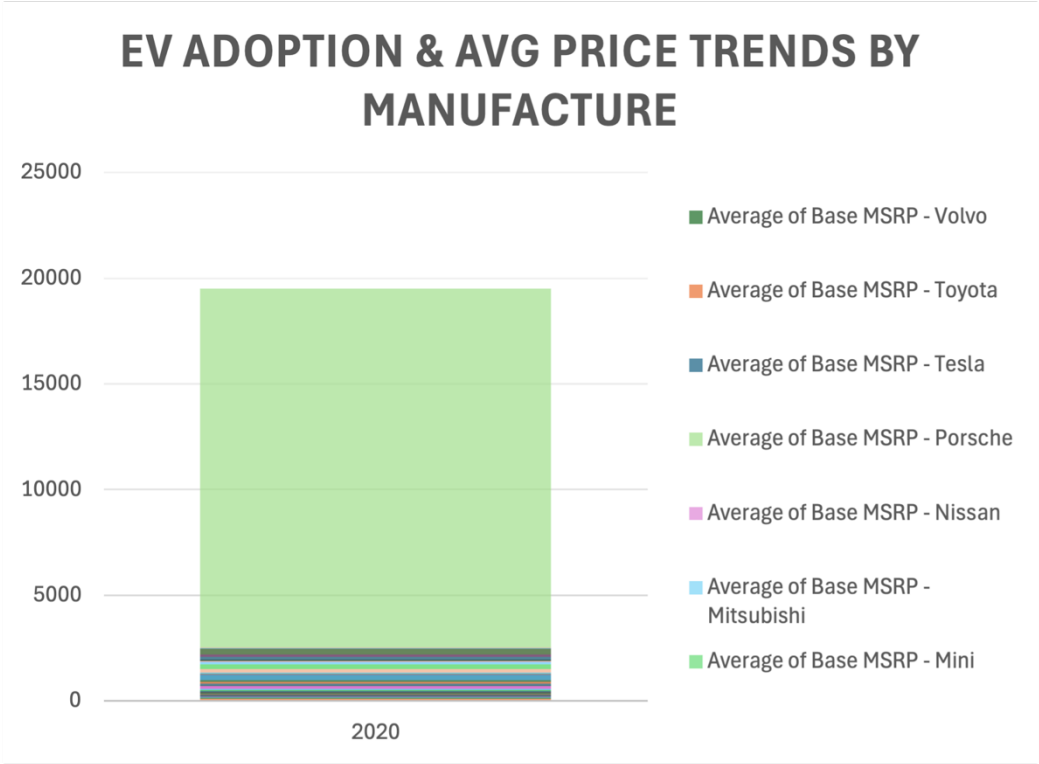
To meet this objective, the dashboard uses a clean dataset that includes school names along with their geographic classification (ULOCALE). The analysis is performed by counting the number of schools in each location type. Filtering options are provided to narrow down by county or district, allowing deeper regional exploration

Analysis Results

The results show that a significant proportion of schools are concentrated in suburban and rural areas, with urban schools also playing a critical role in densely populated regions. Some areas lack adequate schooling facilities, especially in isolated rural zones, which highlights potential gaps in infrastructure planning.

Visualization

A **Line** is used to visually represent the proportion of schools in each geographic category (Urban, Suburban, Town, Rural). This chart gives an immediate understanding of how schools are distributed across different locality types. An interactive slicer allows users to filter by district or county to analyze specific regions.



Objective 2: EV Adoption & AVG Price

General Description

This objective aims to analyze enrollment trends across different schools and explore diversity through the distribution of students across grade levels. By examining student counts in early grades such as Grade 1 and Grade 2, the dashboard provides a snapshot of population flow into the education system and highlights differences across districts or regions.

Specific Requirements

To fulfill this objective, the dashboard requires a dataset with grade-wise enrollment figures—specifically for G01 (Grade 1) and G02 (Grade 2)—alongside district-level identifiers like LEA_NAME. Pivot tables are structured to sum enrollments by district, and slicers allow filtering by district or school status for more granular analysis.

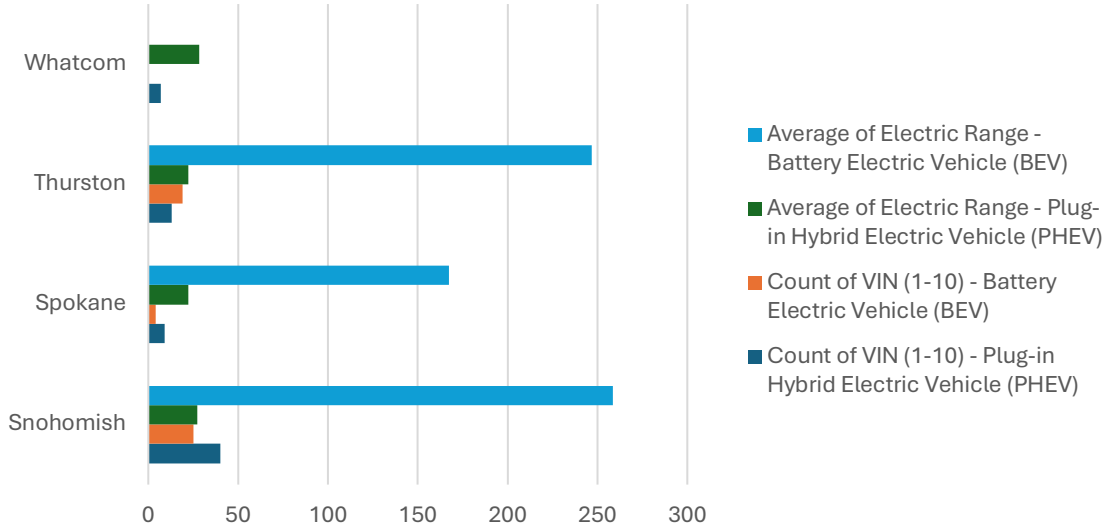
Analysis Results

The data shows that certain districts have higher enrollments in early grades, indicating strong population growth or better school participation rates. Other districts display lower numbers, possibly pointing to migration,

low birth rates, or access issues. Comparing Grade 1 and Grade

Visualization

A **Clustered Column Chart** is used to compare Grade 1 and Grade 2 enrollments side by side for each district. This makes it easy to observe patterns and anomalies across regions. Slicers for LEA_NAME and ULOCALE further enhance interactivity, enabling deeper exploration by users.



General Description

This objective focuses on understanding the current operational status of schools, such as whether they are active, closed, merged, or reopened. It provides a quick yet meaningful snapshot of how the school landscape is evolving over time, helping decision-makers assess trends in school closures and new openings across regions.

Specific Requirements

The dashboard uses the SY_STATUS_TEXT field to classify each school based on its current operational status. A pivot table counts the number of schools under each status category, providing a clear comparison. Filtering options are included to allow breakdowns by district, county, or location type.

Analysis Results

The analysis shows that the majority of schools are active, but a noticeable percentage fall under categories such as merged or closed—often in rural or low-population areas. This may reflect changing demographics, policy decisions, or consolidations to improve resource use.

Visualization

A **Line Chart** presents the school status distribution, visually distinguishing active, closed, reopened, and merged schools. This format provides a clear visual segmentation while keeping the focus on proportions. Slicers for location (ULOCALE) or district (LEA_NAME) allow users to explore trends across different regions.

General Description

This objective focuses on evaluating the level of resources and support provided to students at each school. One key indicator is the number of students directly certified for assistance programs, which reflects both the economic background of the student population and the school's outreach capabilities. This analysis helps in understanding how effectively schools are supporting students in need.

Specific Requirements

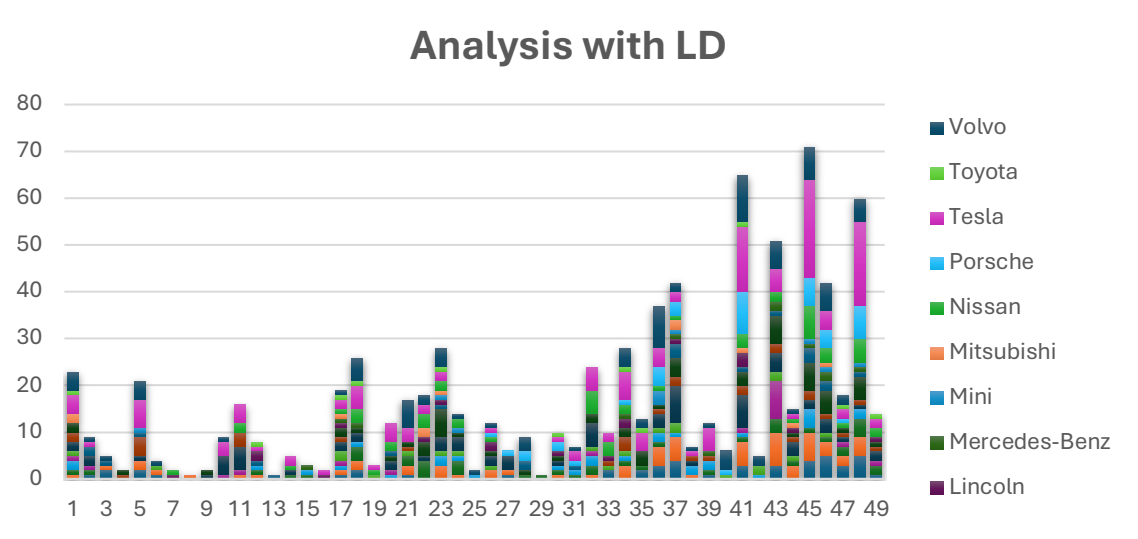
The dashboard uses the DIRECTCERT field to analyze the number of students who are directly certified for support programs. The data is aggregated using school names (SCH_NAME) to show support levels per school. Filters are provided to isolate specific districts, counties, or school types for deeper evaluation.

Analysis Results

The results reveal considerable variation in student support across schools. Some schools show high numbers of direct certifications, indicating higher concentrations of economically disadvantaged students. Others reflect lower counts, either due to demographic differences or underutilization of available programs.

Visualization

A **Bar Chart** is used to display the total number of directly certified students per school. This clear, straightforward visualization highlights which schools are providing significant levels of support. Filters like LEA_NAME or SY_STATUS_TEXT help users explore how certification levels vary across school districts or statuses.



Objective 5:Utility area

General Description

This objective analyzes the extent to which students across various regions benefit from free and reduced lunch programs. These programs are vital indicators of student economic need and overall school support systems.

The goal is to understand how these benefits are distributed across counties and identify areas with the highest student dependency on meal assistance.

Specific Requirements

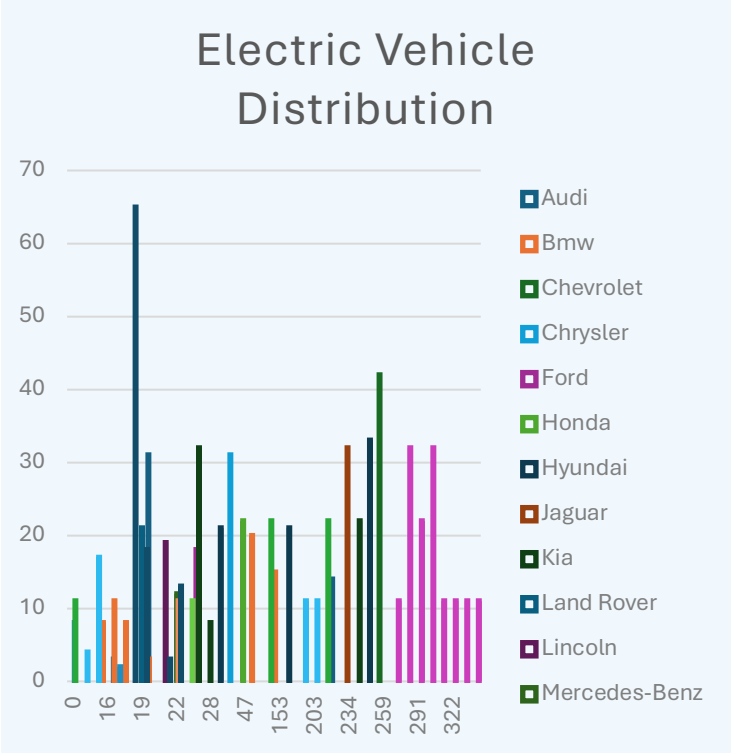
The dashboard utilizes the fields FRELCH (Free Lunch) and REDLCH (Reduced Lunch) to calculate the total number of students receiving each type of meal support. Data is grouped by county (NMCNTY) to allow for regional comparison. Filtering options enable exploration by district, school type, or operational status.

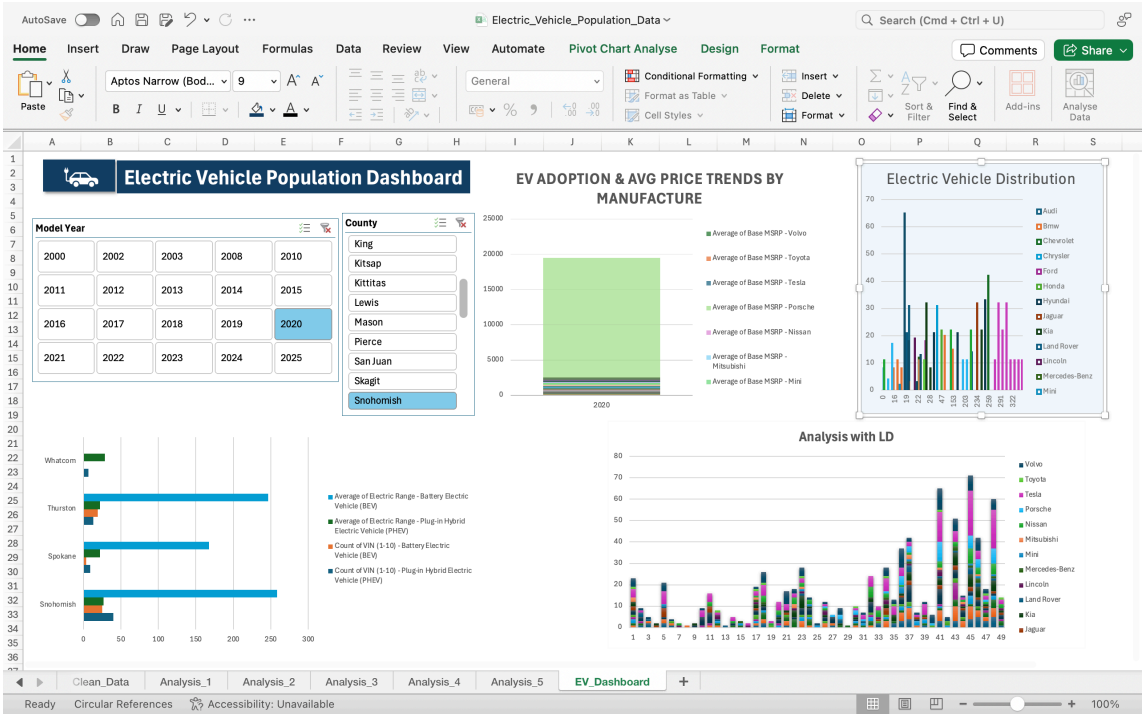
Analysis Results

The analysis shows that several counties have high numbers of students relying on lunch programs, with some showing more dependency on free lunch over reduced lunch. This points to economic challenges faced by students and highlights areas where additional funding or support services may be required.

Visualization

A **Stacked Column Chart** displays the sum of free and reduced lunch recipients per county. This format clearly shows both the total volume and the proportion of each lunch type within each region. Slicers enhance the user's ability to interact with the data and focus on specific areas of interest.





DASHBOARD

5.CONCLUSION:

In short, this project helped me:

- Improve my Excel and data analysis skills,
- Understand how to use visuals for storytelling,
- And gain real experience working with meaningful data.

5. FUTURE SCOPE:

This Excel project has the potential to grow even further. Here are some ideas for future improvement and expansion:

1. **Add More Years of Data** ₀ Including data from multiple years would help identify trends over time, such as changes in enrollment or school support.
2. **Include Performance Metrics** ₀ Adding academic results (like test scores or graduation rates) can help measure the quality of education, not just the support.

3. **Deeper Geographic Analysis** ₀ With mapping tools or Power BI integration, the data can be shown on interactive maps for better geographic understanding.
4. **Automated Data Updates** ₀ Linking Excel to a live database or regularly updated file can make the dashboard dynamic and always up-to-date.
5. **User Role-Based Filters** ₀ Create versions of the dashboard for different stakeholders (teachers, administrators, government) with filters that matter most to them.
6. **Exportable Reports** ₀ Add buttons or macros to automatically generate reports based on selected filters—great for school meetings or decision-making.
7. **Advanced Visuals with Power BI** ₀ Move the project to Power BI or Tableau for more powerful visuals, drill-down features, and online sharing.

7. REFERENCES:

- **Dataset Source**
  - The data used in this project is taken from official school-level statistics provided in .csv format. It includes information such as school names, enrollment numbers, lunch program participation, and school statuses.
- **Microsoft Excel Documentation**
  - Microsoft Support. *Pivot Tables in Excel*. https://support.microsoft.com/excel/pivottable
- **Data Visualization Concepts**
  - Evergreen Data. *Best Practices for Chart Design*. https://stephanieevergreen.com
- **Dashboard Design Tips**
  - Excel Campus. *How to Create a Dashboard in Excel*. https://www.excelcampus.com/charts/create-dashboard/
- **Educational Insights**
  - U.S. Department of Education – National Center for Education

Statistics (NCES) https://nces.ed.gov
(if applicable to your dataset)

Excel Slicers and Interactivity

- Excel Easy. *Using Slicers in Excel*.
  https://www.excel-easy.com/examples/slicer.html

8. **LINKEDIN**

9.Google Drive Excel Link:-