1. Introduction

The rapid growth of electric vehicles (EVs) is a significant step toward reducing carbon emissions and addressing environmental concerns. However, the adoption of EVs varies across regions and depends on multiple factors, including consumer perceptions, market offerings, infrastructure, and government policies.

This report presents a text analytics project conducted to analyze responses collected through a questionnaire on EV adoption. The dataset includes opinions and perceptions of individuals from various cities in India, addressing factors influencing EV adoption, perceived barriers, and the role of government policies.

By performing text analysis on this dataset, the study aims to uncover patterns, themes, and insights that can guide policymakers, manufacturers, and other stakeholders in facilitating EV adoption.

2. Problem Statement

The global push for sustainable transportation has positioned electric vehicles (EVs) as a key solution to reduce greenhouse gas emissions and dependence on fossil fuels. However, despite advancements in EV technology, adoption rates vary significantly due to factors such as infrastructure limitations, financial concerns, and consumer perceptions. Understanding the motivations and barriers influencing consumer decisions is critical to devising effective strategies for promoting EV adoption.

3. About Dataset

The dataset is derived from a structured questionnaire designed to gather insights into consumer perspectives on electric vehicle (EV) adoption. It contains qualitative responses from individuals residing in various cities in India. The dataset comprises seven key fields, each addressing specific aspects of EV adoption.

1. **Name**: The name of the respondent, indicating the participant's identity.
2. **City**: The location of the respondent, capturing geographic diversity and urban perspectives.

3. **Factors Influencing EV Adoption**: Open-ended responses detailing the primary reasons individuals consider EVs, such as cost savings, environmental benefits, and modern features.
4. **Importance of Charging Station Availability**: Opinions on how critical charging infrastructure is in their decision-making process, ranging from "Critical" to "Moderately Important."
5. **Perceived Barriers to EV Adoption**: Qualitative responses describing challenges such as high costs, limited range, and charging time that deter adoption.
6. **Likelihood of EV Adoption (3–5 Years)**: Open-ended responses predicting the likelihood of transitioning to EVs, often contingent on technological or policy advancements.
7. **Impact of Government Policies**: Opinions on how incentives, rebates, and tax benefits might influence their decision to purchase EVs.

4. Methodology

4.1 Data Collection and Preprocessing

Data Collection

The dataset was collected using a structured questionnaire targeting individuals from various cities in India. The survey included open-ended and closed-ended questions designed to understand consumer attitudes, preferences, and barriers regarding electric vehicle (EV) adoption. Participants were selected to ensure diverse representation in terms of geographic location and potential interest in EVs.

The questions captured multiple dimensions, including:

- Key factors influencing EV adoption.
- Importance of charging infrastructure.
- Perceived barriers to adoption.
- Likelihood of switching to EVs in the near future.
- The role of government policies in influencing purchase decisions.

4.2 Data Pre-processing

Pre-processing was undertaken to ensure the dataset was clean, consistent, and ready for text analysis. The following steps were performed:

1. **Data Cleaning**:
   - Removal of duplicate responses, if any.
   - Handling missing values by dropping records with significant omissions or imputing values for minor gaps.
2. **Text Normalization**:
   - Converting all text to lowercase to maintain consistency.
   - Removal of special characters, punctuation, and extra spaces.
   - Tokenization to split textual data into individual words or phrases.
3. **Stopword Removal**:
   - Commonly used words (e.g., "the," "and," "is") that do not contribute to meaning were removed to focus on key terms.
4. **Lemmatization**:
   - Words were reduced to their root form (e.g., "driving" to "drive") to standardize terms with similar meanings.
5. **Categorization**:
   - Qualitative responses were categorized into thematic groups using manual annotation and automated techniques, where applicable.
6. **Data Transformation**:
   - Conversion of text data into structured formats, such as frequency distributions and term-document matrices, for analysis.
7. **Exploratory Analysis**:
   - Initial review of word frequencies, word clouds, and sentiment scores to identify patterns and trends.

4.3. Exploratory Data Analysis (EDA) & Topic Modeling

To gain insights into the dataset, exploratory data analysis was conducted:

1. **Data Overview:**
    - **Number of Observations:** The total number of text documents in the dataset.
    - **Missing Values:** Percentage of missing values in each column.
    - **Data Types:** Data types of each column (e.g., text, numeric).
2. **Textual Analysis:**
    - **Text Length Distribution:** Visualization of the distribution of text lengths to understand the variation in document size.
    - **Word Frequency Distribution:** Visualization of the most frequent words in the dataset to identify common themes.
    - **Word Cloud:** A visual representation of the most frequent words, highlighting the dominant topics.
3. **Sentiment Distribution:**
    - **Bar Chart:** Visualization of the distribution of positive, negative, and neutral sentiments.
    - **Pie Chart:** A visual representation of the proportion of positive, negative, and neutral sentiments.

4.4. Sentiment Analysis Techniques

1. **Lexicon-Based Approach (VADER)**

VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon-based sentiment analysis tool that utilizes a sentiment lexicon to assign sentiment scores to text. The following steps were involved:

- **Sentiment Scoring:** Each text document was processed using VADER to obtain a compound sentiment score, which represents the overall sentiment.
- **Sentiment Classification:** Based on the compound score, the text was classified as positive, negative, or neutral.

2. **Machine Learning Approaches**

Model Selection The following machine learning models were selected for sentiment analysis:

- **Naive Bayes:** A probabilistic classifier based on Bayes' theorem.
- **Support Vector Machine (SVM):** A powerful classification algorithm that finds the optimal hyperplane to separate data points.

3. **Hybrid (VADER+SVM)**

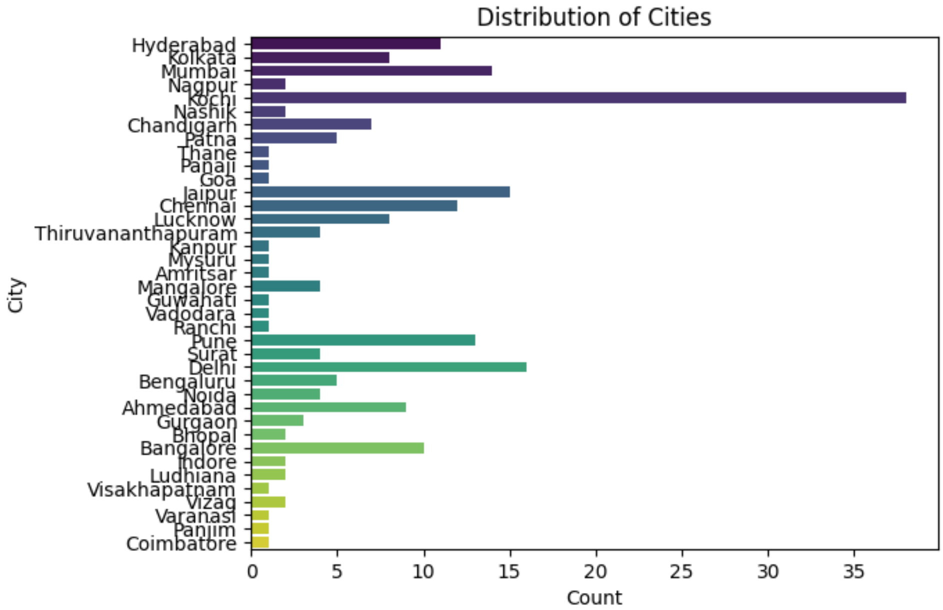
The hybrid approach combining **VADER (Valence Aware Dictionary and sEntiment Reasoner)** and **SVM (Support Vector Machine)** leverages the strengths of both lexicon-based and machine learning-based sentiment analysis techniques.

4.5 Model Training and Evaluation Each model was trained on the training data and evaluated on the testing data using the following metrics:

- **Accuracy:** Overall correctness of predictions.
- **Precision:** Proportion of true positive predictions among all positive predictions.
- **Recall:** Proportion of true positive predictions among all actual positive [2] cases

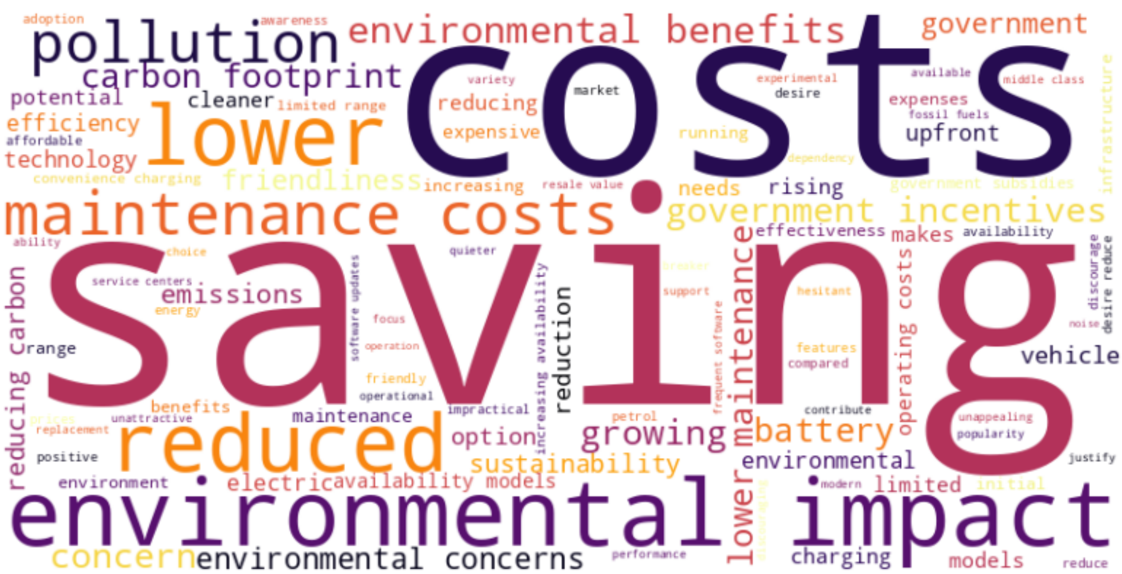
5. Analysis

5.1 Text Analysis and Topic Modelling



The bar chart presents the distribution of cities of questionnaire respondents. The cities are listed on the y-axis, and the number of occurrences of each city is shown on the x-axis. The chart reveals that Kochi has the highest number of occurrences, followed by Jaipur, Mumbai, Delhi, Pune, and Hyderabad. Other cities like Kolkata, Chennai and Bangalore also have a significant presence in the dataset.

5.1.1 Word cloud for document 1



This word cloud visually represents the key themes and concerns related to electric vehicle (EV) adoption.

Central Themes:

- **Costs:** This is a prominent theme, with words like "costs," "expensive," and "upfront" suggesting concerns about the initial investment and potential running costs.
- **Savings:** The word "saving" is central, indicating the potential long-term cost savings associated with EVs, such as lower fuel and maintenance costs.
- **Environmental Benefits:** Words like "pollution," "carbon," "environmental," and "impact" highlight the positive environmental impact of EVs.
- **Infrastructure:** Terms like "charging" and "availability" suggest concerns about the availability and accessibility of charging infrastructure.

Other Key Words and Their Implications:

- **Technology:** Reflects the importance of technological advancements in driving EV adoption.
- **Range:** Highlights concerns about the limited range of current EV models.
- **Government:** Indicates the role of government policies and incentives in promoting EV adoption.

Overall, this word cloud illustrates that while there is growing interest in EVs, concerns about costs, infrastructure, and range still need to be addressed to fully realize the potential of electric vehicles.

5.1.2 Word cloud for document 2

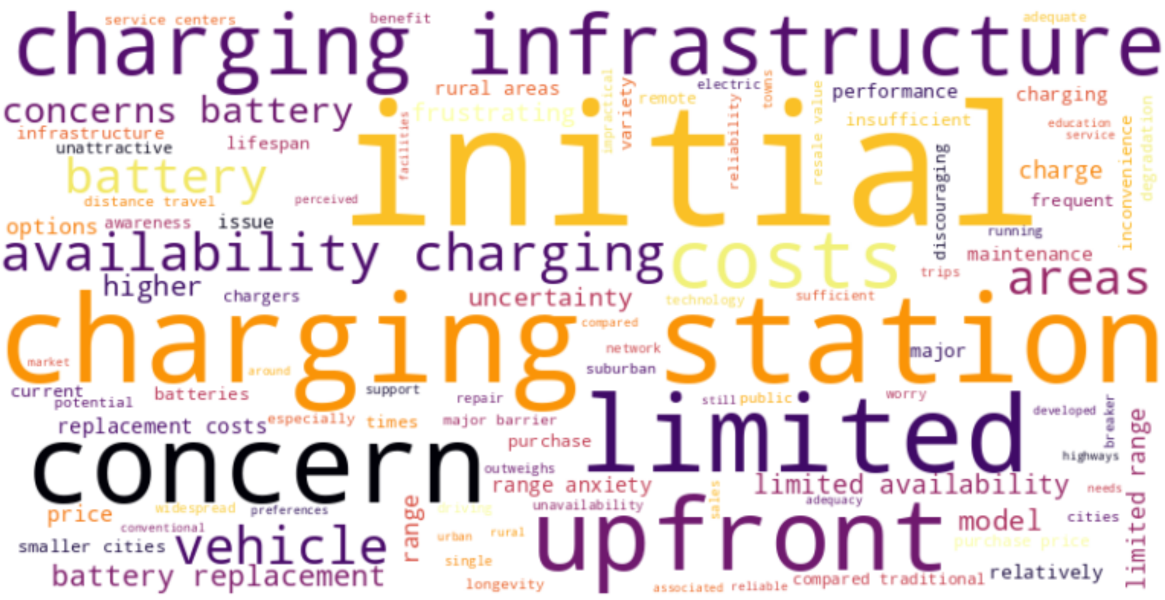
This word cloud highlights the importance of **charging infrastructure** and **charging stations** for electric vehicle (EV) adoption. It emphasizes the need for **accessible**, **reliable**, and **convenient** charging facilities to encourage EV usage.

Key themes and concerns:

- **Infrastructure:** Words like "infrastructure," "facilities," "access," "network," and "stations" emphasize the crucial role of charging infrastructure in EV adoption.
- **Convenience:** Terms like "convenient," "accessible," "easy," and "hassle" indicate that users prioritize convenient charging experiences.
- **Reliability:** Words like "reliable," "adequate," and "sufficient" highlight the need for reliable charging facilities.
- **Location:** Words like "areas," "cities," and "locality" suggest that charging stations should be strategically located in convenient areas.

Overall, this word cloud underscores the importance of addressing charging infrastructure concerns to promote EV adoption and overcome barriers to widespread EV usage.

5.1.3 Word cloud for document 3



This word cloud focuses on the **challenges and concerns** associated with electric vehicle (EV) adoption. It highlights several key issues:

1. Charging Infrastructure:

- Words like "charging," "stations," and "infrastructure" emphasize the importance of a robust charging network for EV adoption.
- Concerns about the **availability** and **accessibility** of charging stations are highlighted.

2. Battery Issues:

- Words like "battery," "replacement," and "lifespan" suggest concerns about the lifespan and replacement costs of EV batteries.
- **Range anxiety** and **limited range** are also significant concerns.

3. Initial Costs:

- Words like "costs," "upfront," and "price" indicate that the initial purchase cost of EVs is a major barrier to adoption.

4. Other Concerns:

- **Uncertainty** about the long-term viability of EVs and their technology.
- **Limited availability** of EV models and options.

Overall, this word cloud illustrates that while there is growing interest in EVs, several challenges need to be addressed to encourage wider adoption. These include expanding charging infrastructure, improving battery technology, reducing upfront costs, and addressing consumer concerns about range and maintenance.

5.1.4 Word cloud for document 4

This word cloud highlights the **positive outlook** on the future of electric vehicles (EVs). It suggests that EVs are likely to become more **affordable**, **accessible**, and **mainstream** in the coming years.

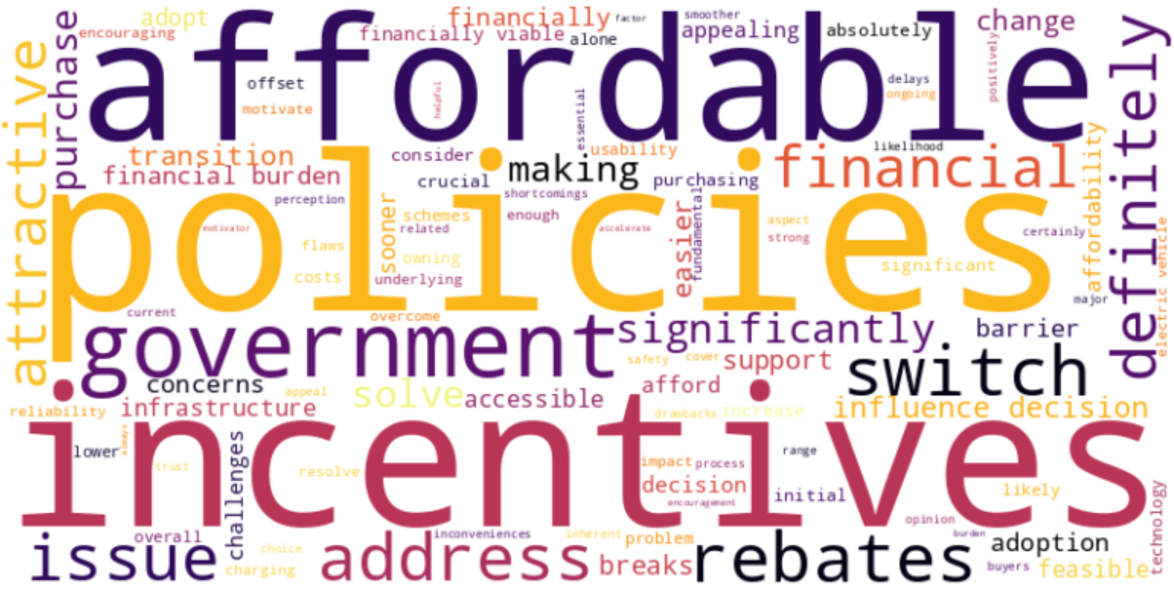
Key themes and predictions:

- **Affordability:** Words like "affordable," "prices," and "decrease" indicate that EV prices are expected to become more competitive with traditional gasoline vehicles.
- **Infrastructure:** Words like "infrastructure," "charging," and "stations" emphasize the importance of developing a robust charging infrastructure to support EV adoption.

- **Technology:** Words like "technology" and "improve" suggest that advancements in EV technology will lead to better performance, range, and charging times.
- **Mainstream Adoption:** Words like "mainstream," "widespread," and "likely" indicate that EVs are expected to become a common choice for consumers.

Overall, this word cloud conveys optimism about the future of electric vehicles, suggesting that they are likely to play a significant role in sustainable transportation.

5.1.5 Word cloud for document 5



This word cloud focuses on the **role of government policies and incentives** in promoting electric vehicle (EV) adoption. It highlights the importance of **financial incentives** and **addressing the initial cost burden** to encourage people to switch to EVs.

Key themes and concerns:

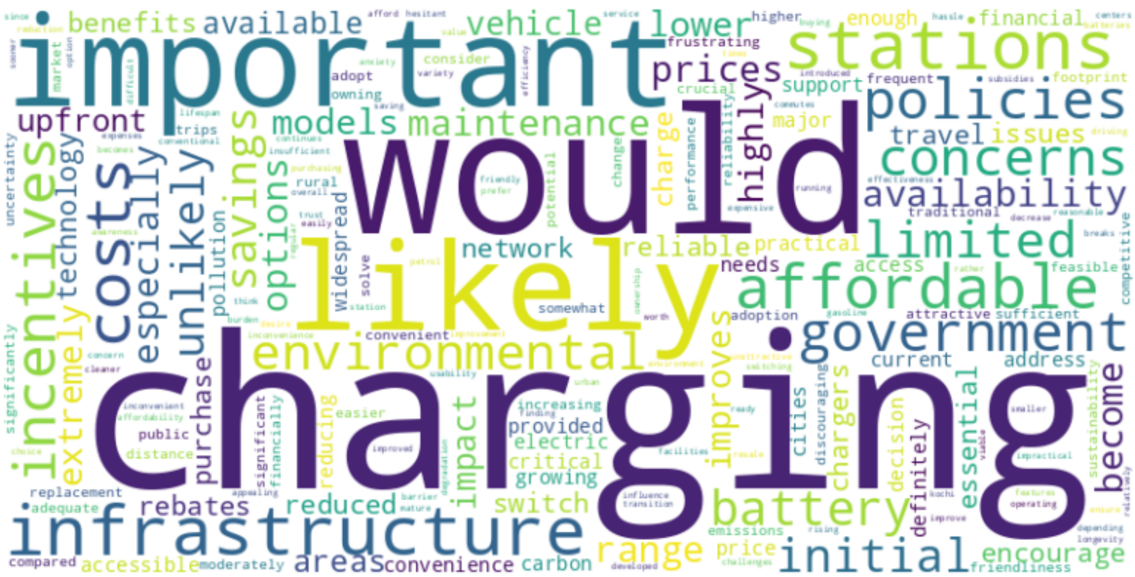
- **Financial Incentives:** Words like "incentives," "rebates," and "affordability" emphasize the need for financial support to make EVs more affordable for consumers.
- **Government Policies:** Words like "government," "policies," and "support" highlight the crucial role of government in creating a favorable environment for EV adoption.

- **Addressing Barriers:** Words like "challenges," "problems," and "address" indicate the need to address the barriers that hinder EV adoption, such as the high initial cost, range anxiety, and limited charging infrastructure.
- **Consumer Perception:** Words like "perception," "consider," and "influence" suggest that government policies can influence public perception of EVs and encourage adoption.

Overall, this word cloud emphasizes that government policies and incentives are essential to drive EV adoption by making EVs more affordable, addressing consumer concerns, and creating a supportive ecosystem for electric mobility.

5.1.6 Word cloud of Entire responses

Word Cloud For The Entire Reviews



This word cloud summarizes the key themes and concerns related to electric vehicle (EV) adoption. Here's a breakdown:

Central Themes:

- **Charging Infrastructure:** This is a prominent theme, with words like "charging," "stations," and "infrastructure" emphasizing the need for a robust and accessible charging network.

- **Affordability:** Words like "affordable," "cost," and "prices" highlight the importance of making EVs more affordable to consumers.
- **Government Policies:** Terms like "government," "policies," and "incentives" suggest the need for supportive government policies to encourage EV adoption.
- **Environmental Benefits:** Words like "emissions," "pollution," and "environmental" highlight the positive environmental impact of EVs.

Other Key Words and Their Implications:

- **Range:** Reflects concerns about the limited range of current EV models.
- **Battery:** Highlights the importance of battery technology and its impact on EV performance and range.
- **Technology:** Suggests that advancements in EV technology are crucial for addressing current limitations.
- **Concerns:** Indicates that there are still concerns and uncertainties about EV adoption, such as range anxiety and charging infrastructure availability.

Overall, this word cloud illustrates that while there is growing interest in EVs, several challenges need to be addressed to accelerate their adoption. These include expanding charging infrastructure, reducing costs, improving battery technology, and creating supportive government policies.

TOPICS

```
Topic 1:
incentive policy environmental cost government saving affordable lower rebate impact
Topic 2:
limited initial battery concern station area upfront range cost availability
Topic 3:
station charger essential network especially area reliable consider sufficient charge
Topic 4:
practicality replacement development major improvement availability year resale space short
Topic 5:
likely price affordable available highly especially model station availability change
```

Topic 1: Government Incentives and Affordability

This topic highlights the role of government policies and incentives in promoting EV adoption. It suggests that government incentives, such as rebates and tax breaks, can make EVs more affordable and encourage their purchase.

Topic 2: Initial Costs and Range Anxiety

This topic focuses on the initial cost of EVs and concerns about their range. It suggests that the high upfront cost and limited range of many EVs can be significant barriers to adoption.

Topic 3: Charging Infrastructure

This topic emphasizes the importance of charging infrastructure for EV adoption. It suggests that the availability of reliable and accessible charging stations is crucial for encouraging EV usage.

Topic 4: Battery Technology and Practicality

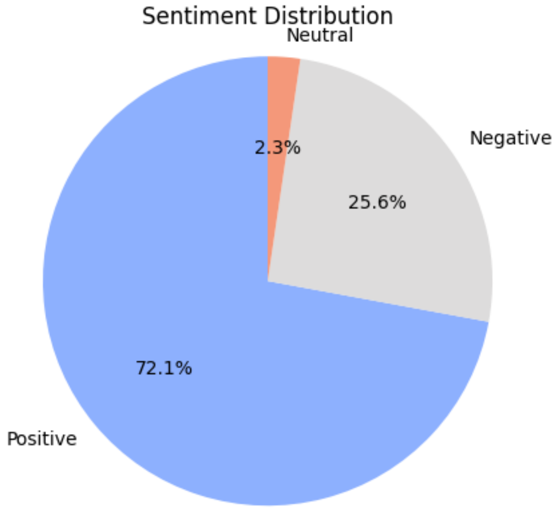
This topic highlights the importance of battery technology and its impact on EV practicality. It suggests that advancements in battery technology, such as improved range and faster charging times, can make EVs more appealing to consumers.

Topic 5: Future Outlook and Market Trends

This topic focuses on the future of EV adoption and market trends. It suggests that EVs are likely to become more affordable and accessible in the future, with a wider range of models available.

5.2 Sentiment Analysis

5.2.1 VADER (Rule Based Sentiment Analysis)



The pie chart above presents the sentiment distribution of the dataset's open-ended responses, analyzed using the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool. The findings reveal the overall polarity of consumer opinions regarding electric vehicle (EV) adoption.

Key Observations

1. **Positive Sentiment (72.1%)**:
    - The largest portion of responses (72.1%) reflects positive sentiment, indicating that most participants have a favorable outlook on EV adoption.
    - Common positive themes include environmental benefits, cost savings, and excitement about technological advancements.
2. **Negative Sentiment (25.6%)**:
    - A significant minority of responses (25.6%) expressed negative sentiment, highlighting concerns about barriers to EV adoption.
    - Frequent negative aspects mentioned include the lack of charging infrastructure, high initial costs, and limited driving range.
3. **Neutral Sentiment (2.3%)**:

- Only a small fraction of responses (2.3%) were categorized as neutral, where participants neither expressed strong positive nor negative opinions.
- These responses often involved factual statements or conditional perspectives, such as "EVs might work well if infrastructure improves."

The high percentage of positive sentiment indicates growing acceptance and enthusiasm for EVs among consumers. However, the significant presence of negative sentiment underscores the need to address critical barriers, such as affordability and infrastructure development, to enhance adoption rates.

This sentiment distribution provides valuable insights for stakeholders, including policymakers and manufacturers, to focus their efforts on mitigating consumer concerns and building on positive perceptions.

5.2.2Naive Bayes

```
Accuracy: 0.66

Classification Report:
              precision    recall  f1-score   support

    Negative       1.00      0.09      0.17        22
     Neutral       0.00      0.00      0.00         2
    Positive       0.65      1.00      0.79        41

    accuracy                           0.66        65
   macro avg       0.55      0.36      0.32        65
weighted avg       0.75      0.66      0.55        65
```

1. Accuracy

- The overall accuracy of the model is **66%**, indicating that the model correctly classified 66% of the instances in the test dataset.

2. Classification Report

Negative Sentiment:

- **Precision**: **1.00** – All instances predicted as negative are actually negative.
- **Recall**: **0.09** – Only 9% of the actual negative instances were correctly identified. This indicates the model struggles to capture negative sentiment.
- **F1-Score**: **0.17** – The F1-score reflects the harmonic mean of precision and recall, indicating poor performance for negative sentiment.

Neutral Sentiment:

- **Precision, Recall, F1-Score**: **0.00** – The model failed to classify neutral sentiments correctly due to the small number of neutral instances in the dataset (only 2 examples).

Positive Sentiment:

- **Precision**: **0.65** – 65% of the instances predicted as positive are correct.
- **Recall**: **1.00** – The model correctly identified all positive instances in the dataset.
- **F1-Score**: **0.79** – The high F1-score indicates that the model performs well in identifying positive sentiments.

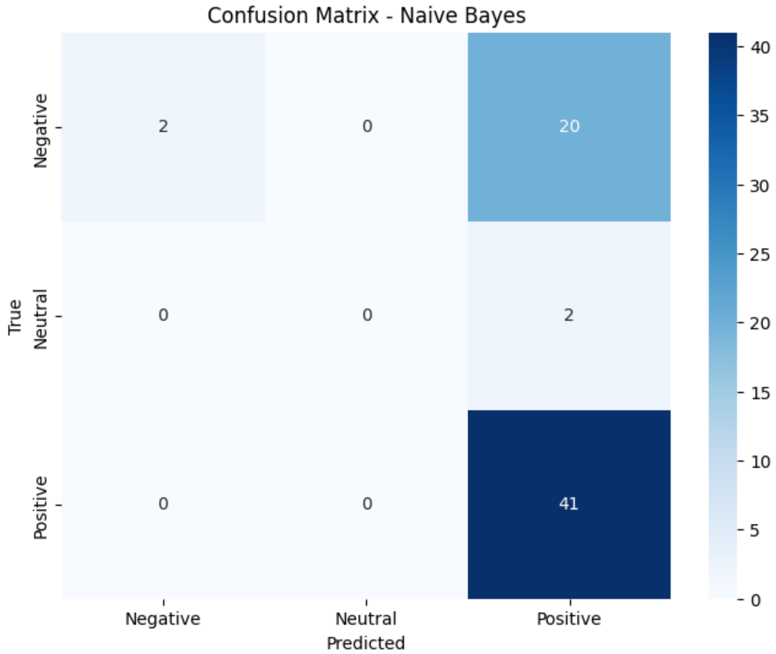
3. Macro Average

- **Precision**: **0.55** – The average precision across all classes, regardless of support (number of instances per class).
- **Recall**: **0.36** – The model has a low recall across all sentiment categories, especially for negative and neutral sentiments.
- **F1-Score**: **0.32** – Overall performance is low for sentiment categories apart from positive.

4. Weighted Average

- **Precision**: **0.75** – Takes into account the number of instances per class (positive class dominates).
- **Recall**: **0.66** – Reflects the model's ability to identify instances correctly across all classes.

- **F1-Score**: **0.55** – Weighted average performance of the model.



The confusion matrix for the Naive Bayes model shows the following:

True Positives (TP): 41 (correctly predicted positive)

True Negatives (TN): 20 (correctly predicted negative)

False Positives (FP): 2 (incorrectly predicted positive)

False Negatives (FN): 0 (incorrectly predicted negative)

Analysis:

- The model performs well in predicting positive sentiments, with a high number of true positives and no false negatives.
- It also performs reasonably well in predicting negative sentiments, with a few false positives.

- However, the model struggles with predicting neutral sentiments. It has no true positives and two false positives for the neutral class. This indicates that the model may be confusing neutral sentiments with positive or negative ones.

Overall, the Naive Bayes model appears to be effective in classifying positive and negative sentiments, but it may need improvement in distinguishing neutral sentiments.

5.2.3 SVM

```
Accuracy: 0.66

Classification Report:
              precision    recall  f1-score   support

    Negative       1.00      0.09      0.17        22
     Neutral       0.00      0.00      0.00         2
    Positive       0.65      1.00      0.79        41

    accuracy                           0.66        65
   macro avg       0.55      0.36      0.32        65
weighted avg       0.75      0.66      0.55        65
```

Accuracy:

- The overall accuracy is **66%**, meaning the model correctly predicts the class in 66% of cases.

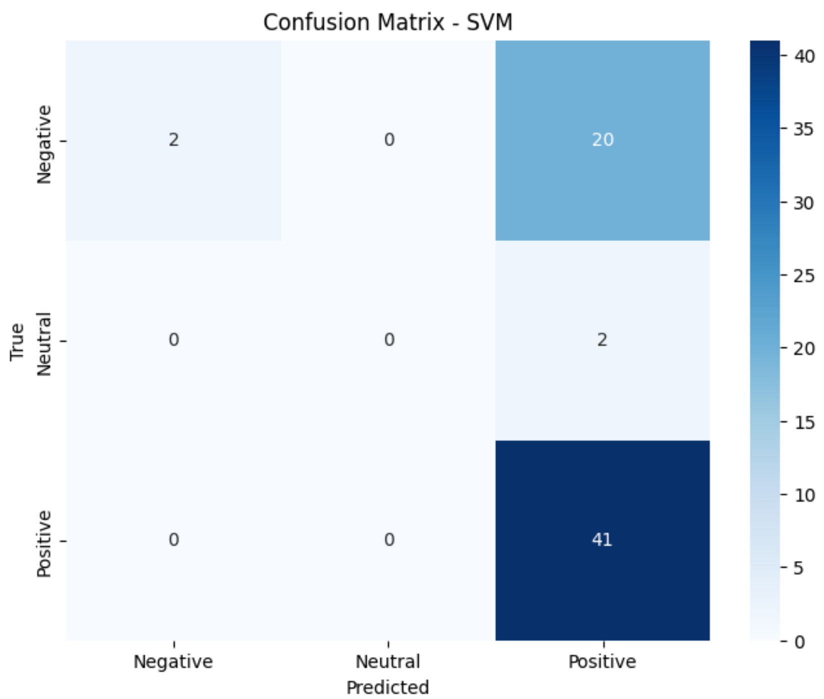
Class-Specific Metrics:

- **Negative Class**:
    - Precision: **1.00** (perfect precision suggests no false positives for this class).
    - Recall: **0.09** (low recall indicates most Negative samples are misclassified).
    - F1-Score: **0.17** (low, due to poor recall).
- **Neutral Class**:

- Precision, Recall, and F1-Score: **0.00** (the model fails to identify any Neutral samples correctly, possibly due to a class imbalance or lack of distinctive features).
- **Positive Class**:
  - Precision: **0.65**.
  - Recall: **1.00** (perfectly captures all Positive samples).
  - F1-Score: **0.79** (good performance, driven by high recall).

Macro and Weighted Averages:

- **Macro Average** (unweighted mean across classes):
  - Precision: **0.55**.
  - Recall: **0.36** (poor overall recall due to underperformance in Negative and Neutral classes).
  - F1-Score: **0.32**.
- **Weighted Average** (accounts for class support size):
  - Precision: **0.75**.
  - Recall: **0.66**.
  - F1-Score: **0.55** (better due to the dominance of the Positive class).



The confusion matrix for the SVM model shows the following:

True Positives (TP): 41 (correctly predicted positive)

True Negatives (TN): 20 (correctly predicted negative)

False Positives (FP): 2 (incorrectly predicted positive)

False Negatives (FN): 0 (incorrectly predicted negative)

Analysis:

- Similar to the Naive Bayes model, the SVM model performs well in predicting positive sentiments, with a high number of true positives and no false negatives.
- It also performs reasonably well in predicting negative sentiments, with a few false positives.
- However, like the Naive Bayes model, the SVM model struggles with predicting neutral sentiments. It has no true positives and two false positives for the neutral class. This indicates that both models may be confusing neutral sentiments with positive or negative ones.

Overall, the SVM model, like the Naive Bayes model, appears to be effective in classifying positive and negative sentiments, but it may need improvement in distinguishing neutral sentiments.

5.2.4 Hybrid (VADER + SVM)

```
Hybrid Model Accuracy: 0.97

Hybrid Model Classification Report:
              precision    recall  f1-score   support

    Negative       0.96      1.00      0.98        22
     Neutral       0.00      0.00      0.00         2
    Positive       0.98      1.00      0.99        41

    accuracy                           0.97        65
   macro avg       0.64      0.67      0.66        65
weighted avg       0.94      0.97      0.95        65
```

Accuracy:

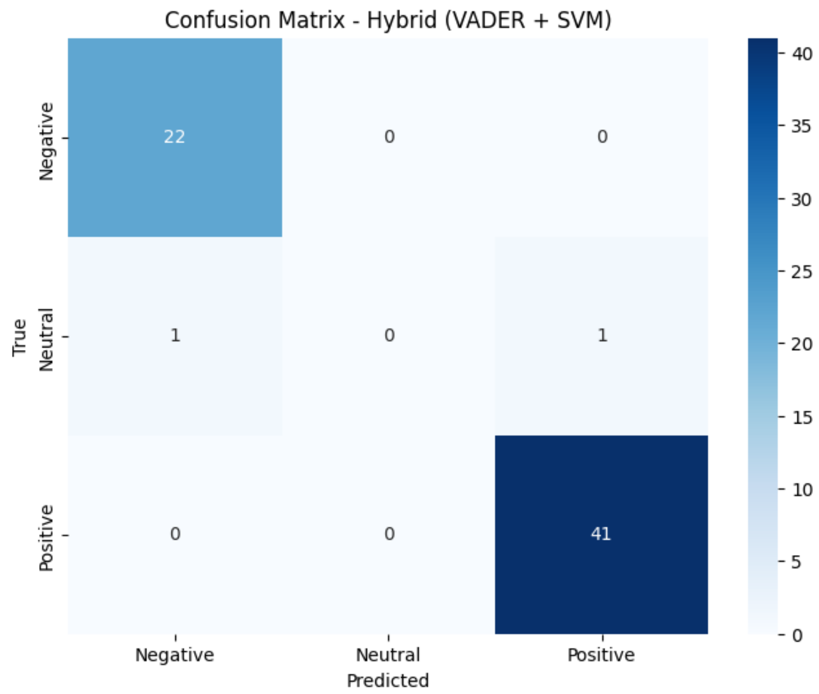
- The **hybrid model** achieves an impressive **97% accuracy**, significantly higher than the earlier SVM-only model.

Class-Specific Metrics:

- **Negative Class**:
  - Precision: **0.96** (almost perfect at correctly identifying Negative predictions).
  - Recall: **1.00** (all Negative samples are correctly identified).
  - F1-Score: **0.98** (excellent balance between precision and recall).
- **Neutral Class**:
  - Precision, Recall, and F1-Score: **0.00** (the model still fails to capture the Neutral class, likely due to severe class imbalance).
- **Positive Class**:
  - Precision: **0.98** (near-perfect precision for Positive predictions).
  - Recall: **1.00** (captures all Positive samples accurately).
  - F1-Score: **0.99** (outstanding performance for this dominant class).

Macro and Weighted Averages:

- **Macro Average** (unweighted mean across classes):
  - Precision: **0.64** (dragged down by the Neutral class).
  - Recall: **0.67** (higher due to perfect recall for Negative and Positive classes).
  - F1-Score: **0.66**.
- **Weighted Average** (accounts for class support size):
  - Precision: **0.94**.
  - Recall: **0.97**.
  - F1-Score: **0.95** (excellent performance overall, driven by the large Positive and Negative classes).



True Positives (TP): 41 (correctly predicted positive)

True Negatives (TN): 22 (correctly predicted negative)

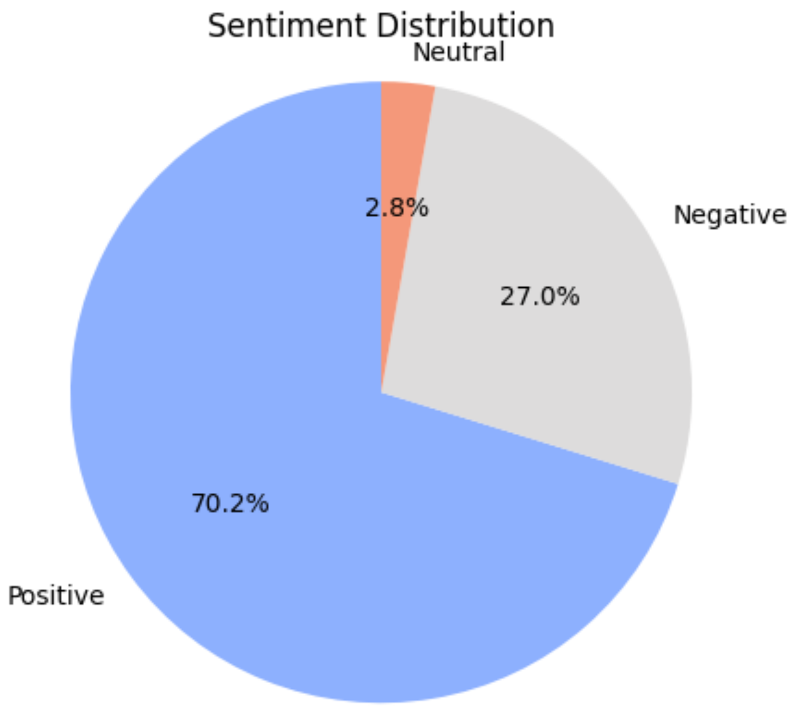
False Positives (FP): 0 (incorrectly predicted positive)

False Negatives (FN): 2 (incorrectly predicted negative)

Analysis:

- The hybrid model demonstrates excellent performance in predicting positive and negative sentiments, with no false positives or false negatives in these categories.
- It also shows improvement in predicting neutral sentiments compared to the individual models. It correctly identifies two neutral instances and has only one false positive.

Overall, the hybrid model outperforms the individual Naive Bayes and SVM models in terms of overall accuracy and ability to distinguish between all three sentiment classes.



The pie chart above illustrates the sentiment distribution of the dataset's open-ended responses, analysed using hybrid model. The chart provides insights into consumer perceptions regarding electric vehicle (EV) adoption across three sentiment categories: Positive, Negative, and Neutral.

Key Observations:

1. **Positive Sentiment (70.2%)**:
    - The majority of responses (70.2%) exhibit positive sentiment, indicating a predominantly favourable attitude toward EV adoption.
    - Common positive themes include enthusiasm about environmental benefits, potential cost savings, and advancements in EV technology.
2. **Negative Sentiment (27.0%)**:
    - A significant portion of the responses (27.0%) reflects negative sentiment, highlighting prevalent concerns about EV adoption.

- Recurring negative themes include the lack of widespread charging infrastructure, high upfront costs, and concerns over driving range and battery performance.

3. **Neutral Sentiment (2.8%)**:
    - Neutral responses constitute a small fraction (2.8%) of the dataset, where participants expressed neither overtly positive nor negative opinions.
    - These responses often consist of factual statements or conditional views, such as "EV adoption depends on policy incentives and infrastructure development."

Insights:

- The high proportion of positive sentiment suggests increasing consumer interest and optimism toward EV adoption, potentially driven by environmental awareness and supportive policies.
- However, the significant presence of negative sentiment emphasizes the need for stakeholders to address critical barriers, such as affordability and the availability of charging stations, to foster higher adoption rates.
- The low percentage of neutral sentiment indicates that most respondents have distinct opinions on the topic, whether positive or negative.

6. Results and Discussions

Leaderboard

Model	Accuracy
Naive Bayes	66
SVM	66
Hybrid (VADER + SVM)	97

Model Comparison and Accuracy Analysis

The sentiment analysis models applied to the dataset were evaluated based on their accuracy in classifying consumer opinions regarding electric vehicle (EV) adoption.

1. Naive Bayes

- **Accuracy**: 66%
- **Performance Analysis**:
  - Naive Bayes, a probabilistic model, achieved moderate accuracy. It performed well in classifying positive sentiments but struggled with negative and neutral sentiments due to class imbalance and the model's assumptions about feature independence.
  - **Challenges**:
    - Limited recall for negative sentiments indicates the model's difficulty in identifying less frequent classes.
    - Neutral sentiments were entirely misclassified, likely due to insufficient training data for this category.

2. Support Vector Machine (SVM)

- **Accuracy**: 66%
- **Performance Analysis**:

  The SVM model, known for its ability to handle high-dimensional data, achieved accuracy comparable to Naive Bayes. It excelled in identifying positive sentiments but faced challenges with negative and neutral classes.

  - **Challenges**:
    - Despite its precision for the dominant positive class, the low recall for the negative class reveals limitations in handling minority classes.
    - Similar to Naive Bayes, SVM struggled with the neutral class, likely due to the dataset's class imbalance.

3. Hybrid Model (VADER + SVM)

- **Accuracy**: 97%
- **Performance Analysis**:
    - The hybrid approach combining VADER (a lexicon-based sentiment analysis tool) with SVM achieved a significantly higher accuracy than standalone models. This method leveraged VADER's ability to extract sentiment-rich features and SVM's classification strengths.
    - **Strengths**:
        - Outstanding performance in classifying both positive and negative sentiments, with near-perfect precision and recall.
        - The hybrid model successfully balanced lexicon-based insights with data-driven classification, enabling more accurate predictions.
    - **Challenges**:
        - Despite overall excellence, the model continued to struggle with neutral sentiments, possibly due to their low representation in the dataset.

Insights from the Leaderboard

The hybrid model's superior performance (97% accuracy) highlights the benefits of integrating rule-based and machine-learning approaches. By combining VADER's pre-classification capabilities with SVM's ability to generalize patterns, the hybrid model demonstrated a comprehensive understanding of the data.

7. Recommendations

1. Address Negative Sentiments by Tackling Consumer Barriers

1. **High Upfront Costs**:
   - **Recommendation**: Implement subsidy programs and tax incentives to make EVs more affordable.
   - **Action Plan**: Collaborate with financial institutions to offer low-interest loans or lease options for EV purchases.
2. **Inadequate Charging Infrastructure**:
   - **Recommendation**: Expand the public charging network in urban and rural areas.
   - **Action Plan**: Provide grants to private companies for installing fast chargers and set up battery-swapping stations to reduce range anxiety.
3. **Limited Driving Range**:
   - **Recommendation**: Invest in R&D to improve battery technology for extended ranges.
   - **Action Plan**: Showcase real-world examples of long-range EVs to build consumer confidence.

2. Reinforce Positive Sentiments

1. **Environmental Benefits**:
   - **Recommendation**: Highlight the environmental impact of EV adoption through marketing campaigns.
   - **Action Plan**: Partner with environmental organizations to create awareness programs demonstrating EVs' role in reducing carbon footprints.
2. **Cost Savings**:
   - **Recommendation**: Emphasize the lower operating and maintenance costs of EVs compared to traditional vehicles.
   - **Action Plan**: Publish comparative case studies to showcase long-term cost benefits for consumers.

3. Enhance Neutral Sentiments with Reliable Information

1. **Consumer Education**:
   - **Recommendation**: Address knowledge gaps by providing clear, factual information about EV capabilities and incentives.
   - **Action Plan**: Develop easy-to-understand guides on government benefits, charging options, and maintenance.

2. **Test Drive Campaigns**:
   - **Recommendation**: Increase consumer engagement by offering free test drive programs to experience EV benefits first-hand.
   - **Action Plan**: Organize EV fairs in collaboration with dealerships to demonstrate features and answer consumer questions.

8. Conclusion

The rapid adoption of Electric Vehicles (EVs) is crucial for achieving sustainable transportation and mitigating environmental challenges. This report aimed to analyse consumer sentiments surrounding EV adoption and identify the key factors influencing their opinions. Using hybrid sentiment analysis techniques (VADER and SVM), we observed a significant variation in consumer attitudes, ranging from concerns about high costs and inadequate infrastructure to positive perceptions of environmental benefits and long-term savings.

The results highlight the critical barriers to EV adoption, such as high upfront costs, limited charging infrastructure, and range anxiety, alongside drivers like environmental sustainability and reduced operational expenses. Addressing these barriers and amplifying the positive aspects can accelerate EV adoption.