Market Segmentation

Analyzing the respective market in India using Market Segmentation Analysis for Online Vehicle Booking Product Startup

Team Members:
Himanshu Pradhan
Malavika V Nair
Piyush Kothari
V Sharath Raj
Shivona Fernandes

OVERVIEW

As a team working under an Online Vehicle Booking Product Startup in India, our primary objective is to find an alternate segment in the vehicle market that can provide us with early market entry and generate revenue. This objective stems from the intense competition faced from established players like Ola and Uber in the cab booking industry.

To achieve this, we will conduct a thorough analysis of the vehicle market in India using segmentation analysis. The purpose of segmentation analysis is to divide the market into distinct groups or segments based on various criteria. In addition to traditional segmentation factors such as geographic, demographic, psychographic, and behavioral characteristics, we will also explore different categories of segments based on the availability of data.

It is important to note that not all types of market data may be easily accessible for every segment. Therefore, extensive research and data collection tasks will be required to gather the necessary information for segmentation analysis.

The scope of market segmentation is broad and offers a range of possibilities. We will consider multiple types of data to create segments that accurately represent the characteristics and preferences of potential customers. These segments may evolve and change based on the datasets we collect and analyze.

Once we have identified the relevant segments, our next step will be to develop a feasible strategy to enter the market and target those

segments where profitability and potential for revenue generation are high. This strategy will be based on a deep understanding of the needs, preferences, and behaviors of customers within the identified segments.

In summary, our goal is to leverage segmentation analysis to identify profitable segments within the Indian vehicle market. By thoroughly researching and collecting data, we will create segments based on various criteria. Subsequently, we will devise a strategy to enter the market, offering vehicle booking services that cater specifically to the needs of the identified segments.

PROBLEM BREAKDOWN

Fermi Estimation:

Market size of the online vehicle booking market in India:

1. Estimate the total population of India (approximately 1.3 billion).
2. Assume that around 40% of the population resides in urban areas and is likely to use vehicle booking services.
3. Estimated target market size: 1.3 billion * 40% = 520 million.

Competition penetration:

1. Assume that Ola and Uber currently have a combined market share of 80% in the online vehicle booking market.
2. Estimated available market share for the startup: 100% - 80% = 20%.

Market share capture:

1. Assume that the startup aims to capture 10% of the available market share.
2. Estimated target market share for the startup: 20% * 10% = 2%.

Revenue per booking:

1. Estimate the average revenue per booking made through the online platform as Rs. 300.
2. Assume an average of 10 bookings per customer per month.
3. Estimated revenue per customer per month: Rs. 300 * 10 = Rs. 3,000.

Total potential revenue:

1. Estimate the number of customers the startup can attract based on the target market size and market share.
2. Estimated number of customers: 520 million * 2% = 10.4 million.
3. Estimated total potential revenue per month: 10.4 million customers * Rs. 3,000 = Rs. 31.2 billion.

First Principle Analysis:

Problem Statement: The heavy competition in the cab booking market in India, dominated by Ola and Uber, necessitates the identification of an alternate segment that can provide early market entry and revenue for the online vehicle booking startup.

Market Size and Potential:

1. Determine the total population of India (approximately 1.3 billion) and the percentage of the population residing in urban areas.
2. Assess the current market penetration and market share of Ola and Uber.
3. Estimate the potential market size by calculating the addressable market (urban population) and deducting the current market share of existing players.
4. Evaluate the growth potential of the market based on factors like population trends, increasing urbanization, and changing consumer preferences.

Customer Segmentation:

1. Conduct comprehensive research to identify various segments within the vehicle booking market.
2. Consider geographic, demographic, psychographic, and behavioral factors, as well as other categories of segmentation based on the availability of data.
3. Gather data through surveys, interviews, focus groups, and market research reports to gain insights into customer needs, preferences, and pain points.
4. Analyze the data to identify common characteristics and patterns within each segment, enabling the startup to tailor its offerings to specific customer groups.

Value Proposition:

1. Understand the unique value that the online vehicle booking startup can offer compared to existing competitors like Ola and Uber.
2. Identify key differentiators such as price, quality of service, vehicle options, customer support, ease of use, or additional features.
3. Determine how these differentiators align with the identified customer segments and their specific needs and preferences.
4. Develop a compelling value proposition that clearly communicates the benefits and advantages of choosing the startup's vehicle booking service over competitors.

Revenue Generation:

1. Determine the pricing strategy based on market dynamics, customer willingness to pay, and the startup's cost structure.
2. Evaluate potential revenue streams, such as booking fees, subscription models, partnerships with other businesses, or value-added services.
3. Estimate the average revenue per customer and calculate the potential revenue based on the target market size, market share goals, and customer acquisition rates.
4. Continuously monitor and optimize revenue generation strategies to ensure profitability and sustainability.

Operational Execution:

1. Assess the operational requirements for running an online vehicle booking platform, including technology infrastructure, fleet management, customer support systems, and partnerships with vehicle owners or drivers.
2. Develop a detailed operational plan, considering factors like scalability, efficiency, reliability, and compliance with local regulations.
3. Build a strong team with the necessary skills and expertise to execute the operational plan effectively.

Marketing and Growth Strategy:

1. Devise a comprehensive marketing strategy to create awareness and attract customers from the identified segments.

2. Utilize various marketing channels such as online advertising, social media, referral programs, partnerships, and targeted campaigns.
3. Focus on customer acquisition, retention, and engagement strategies to drive growth and establish a loyal customer base.
4. Continuously analyze and adapt the marketing approach based on customer feedback, market trends, and competitive landscape.

MARKET OVERVIEW

The vehicle booking market in India has experienced significant growth in recent years, driven by factors such as urbanization, rising disposable income, digital transformation, and changing consumer preferences. With a market size valued at billions of dollars, the industry caters to the increasing demand for convenient and reliable transportation solutions. Major players in the market include Ola, Uber, Zoomcar, and Drivezy, each offering a range of services from cab booking to self-drive rentals. The market is highly competitive, and companies invest in technological innovations to enhance the user experience and streamline operations. However, challenges such as intense competition, infrastructure limitations, and regulatory compliance exist. Looking ahead, the market is poised for further growth, with the adoption of electric vehicles and integration of advanced technologies expected to shape the future of the industry. Companies that can adapt to evolving consumer demands and provide reliable, innovative, and convenient transportation options are likely to thrive in this dynamic market.

Key Findings:

1. Growth Trend:

Ola's market size in India witnessed consistent growth from 2016 to 2019. The market size increased from 5,000 Crores in 2016 to

15,500 Crores in 2019, indicating a compound annual growth rate (CAGR) of X%.

2. Impact of the COVID-19 Pandemic:
The market size declined to 9,800 Crores in 2020. This decline can be attributed to the COVID-19 pandemic, which led to travel restrictions, lockdowns, and reduced demand for ride-hailing services. The pandemic had a significant impact on the transportation industry, affecting Ola's market size and performance.

3. Recovery and Growth:
In 2021, Ola's market size recovered to 13,200 Crores, suggesting a potential rebound from the pandemic-induced decline. This recovery may be attributed to the easing of restrictions, the gradual return of travel demand, and Ola's strategic initiatives to adapt to the changing market conditions.

4. Continued Growth in 2022 and 2023:
Ola's market size continued to grow in subsequent years. In 2022, the market size reached 16,700 Crores, and it further expanded to 20,300 Crores in 2023. These figures indicate a positive growth trajectory for Ola in the Indian market.

Conclusion:
The analysis of Ola's market size in India provides valuable insights into the company's performance and growth trends. Despite the

temporary setback caused by the COVID-19 pandemic in 2020, Ola has demonstrated resilience and managed to recover and expand its market size in subsequent years.

The table above represents Ola's market size in India based on different vehicle categories and years. It provides an overview of the estimated market sizes for each category in crores for the years 2019, 2020, 2021, 2022, and 2023.

In 2019, Ola Micro had a market size of around 4,000 crores, making it the largest category. However, in 2020, the market size for Ola Micro decreased to 2,800 crores, possibly due to the impact of the COVID-19 pandemic on transportation services. As the market gradually recovered, the market size for Ola Micro increased to 3,600 crores in 2021 and is projected to further grow to 5,000 crores in 2023.

Similarly, Ola Mini, Ola Prime, and Ola Auto experienced fluctuations in their market sizes over the years. Ola Mini, which had a market size of 6,500 crores in 2019, declined to 4,600 crores in 2020 but recovered to 7,900 crores in 2023. Ola Prime, on the other hand, had a market size of 8,900 crores in 2019, decreased to 6,200 crores in 2020, and is projected to reach 11,000 crores in 2023. Ola Auto, Ola Rentals, Ola Bike, Ola Outstation, Ola Share, and Ola Corporate also witnessed similar trends of fluctuating market sizes across the years.

It's important to note that the provided figures are for illustrative purposes only and may not reflect the actual market sizes. The

variations in market size can be attributed to factors such as changes in consumer preferences, competition, economic conditions, and the impact of external events like the COVID-19 pandemic. The table serves as a snapshot to demonstrate the potential market sizes for different vehicle categories within Ola's operations in India over a specific timeframe.

MARKET DYNAMICS

The table below provides an overview of the market dynamics of vehicle booking entities in major cities of India, including Mumbai, Delhi, Bengaluru, Chennai, Kolkata, and Indore. In these cities, there is intense competition between global and local players, driving innovation and offering customers a wide range of options. The demand for convenient, reliable, and affordable transportation is high, with customers emphasizing safety, comfort, and pricing. Technological advancements play a crucial role, with mobile apps, real-time tracking, and digital payment solutions enhancing the customer experience. Regulatory environments ensure licensing requirements, background checks, and fare regulations to maintain passenger safety and fair practices. Partnerships with various stakeholders, including airports, hotels, event organizers, and public transportation authorities, provide seamless services and expand market reach. Market expansion strategies focus on suburban areas, neighbouring towns, and tier-2 and tier-3 cities. Evolving business models include services like car-sharing, electric vehicle rentals, integration with public transportation, and subscription-based offerings. Each city has its own market dynamics, reflecting the specific needs and preferences of the local customer base.

City	Competitive Landscape	Customer Demand and Behavior	Technological Advancements	Regulatory Environment	Partnerships and Alliances	Market Expansion and Penetration	Evolving Business Models
Mumb	Intense	High	Advanced	Licensing	Partnersh	Expansio	Car-

ai	competition between global and local players	demand for reliable and affordable transportation	mobile apps, real-time tracking, and digital payments	requirements and fare regulations in place	ips with airports, hotels, and event organizers for seamless services	n to suburban areas and neighbouring towns	sharing, electric vehicle rentals, and integration with public transportation
Delhi	Fierce competition among major ride-hailing companies	High demand with emphasis on safety, comfort, and pricing	AI algorithms, real-time traffic analysis, and navigation apps	Licensing, background checks, and fare regulations	Collaborations with metro rail and last-mile connectivity providers	Penetration into suburbs and neighbouring cities	Partnerships for integrated ticketing and multi-modal transportation solutions
Bengaluru	Competition between global and local ride-hailing companies	High demand for convenient and quick transportation	AI-based algorithms, real-time tracking, and safety features	Licensing, vehicle quality, and driver checks	Partnerships with technology companies and startups for innovative mobility solutions	Expansion to tier-2 and tier-3 cities	Subscription-based services and on-demand delivery partnerships for revenue streams
Chennai	Competition among prominent ride-hailing companies	Strong demand for reliable and affordable transportation	User-friendly apps, real-time tracking, and digital payments	Licensing, background checks, and safety standards	Partnerships with tourist destinations and hospitality providers	Expansion to industrial hubs and suburban areas	Partnerships with public transport authorities for integrated transportation
Kolkata	Presence of global and local ride-hailing companies	High demand for safe and economical transportation options	Mobile apps, real-time tracking, and digital payment solutions	Licensing requirements and fare regulations in place	Partnerships with local transportation authorities and government	Expansion to nearby towns and suburbs	Integration with public transportation systems and last-mile connectiv

					initiatives		ity
Indore	Competitive market with a mix of local and emerging players	Growing demand for convenient and reliable transportation	Technology-driven solutions for user-friendly experiences	Adherence to licensing and regulatory requirements	Partnerships with local businesses and transportation authorities	Expansion to nearby towns and suburban areas	Introduction of new services like bike-sharing and integration with public transport

MARKET SEGMENTATION

The Indian market for electric vehicles is divided based on the type of vehicle and power source. Regarding vehicle type, the market includes Passenger Cars, Commercial Vehicles, and Two- and Three-wheelers. In terms of power source, the market is categorized into Battery Electric Vehicles, Plug-in Electric Vehicles, and Hybrid Electric Vehicles.

India, the second most populous country globally after China, shares a similarity with China in its strong emphasis on electrifying buses. Similar to China, which possesses the world's largest electric bus fleet, India is actively promoting the electrification of buses. Several state governments have already started procuring electric buses from both Chinese and domestic electric bus manufacturers.

In response to the need for curbing greenhouse gas emissions from vehicles, the Indian government is actively encouraging the adoption of electric-powered vehicles across different states. This initiative has led to a surge in demand for electric buses in India. Factors such as the growth of domestic manufacturing, rapid urbanization, and increasing environmental awareness are driving the market.

For example, in February 2020, India's first inter-city electric bus service was inaugurated by the Union Transport Minister. These buses, manufactured by Mitra Mobility Solutions, have a range of 300 km on a full charge.

Numerous local bus manufacturers have collaborated with Chinese counterparts to meet the rising demand for electric buses in India. For instance, in 2019, Foton PMI planned to invest approximately INR 500 crore in a joint venture with China's Beiqi Foton Motor Co. to manufacture electric buses in India. The company has already provided five electric buses to an airline for internal operations.

SITUATIONAL ANALYSIS

The online vehicle booking market is emerging in a futuristic landscape as it is the new norm and offers easy, simple and affordable transportation facilities to users. New innovations continue to surge as a result of development of electric and connected automobiles. In The Indian automobile industry, technology plays a crucial role in growth and transformation as a whole which would continue to evolve at a faster pace.

CURRENT TRENDS

1. Ride-Hailing Dominance

Ola and Uber continue to dominate the online vehicle booking market in India since these services offers convenient, on-demand transportation services through their own mobile applications. Moreover, they also provide a wide range of vehicle options. 2. Micro-Mobility Solutions

Micro-mobility solutions, such as bike and scooter sharing services are mostly gaining popularity in the urban areas. They provide services assessable over short-distance and users could easily locate, unlock and rent them on a daily commute to reduce congestion and promote eco-friendly travel.

3. Integration of Public Transport

Most of the online vehicle booking platforms are integrating public transport options such as buses and trains into their services. By doing this, it aims to provide users with a seamless travel experience by offering them various options and easy ticketing through a single platform.

4. Electric Vehicle Adoption

Electric Vehicles in Indian market is witnessing growth which is driven by government initiatives and also increasing environmental awareness. Some online booking platforms are incorporating electric vehicles so that users could choose sustainable transportation options.

5. Subscription- Based Services

Subscription based models are gaining popularity in the online vehicle booking industry. Some of these platforms even introduced monthly or annual subscription plan which would offer discounted fares and other benefits to users. This could in turn foster customer loyalty and encourage regular usage of the platform.

6. Expansion in Rural Areas

The Online Booking Services are trying to expand their reach to rural areas so that people residing here could easily afford the various transportation options through improving mobility for the rural population.

TARGET MARKET

Corporate and Business Travelers:

This segment includes professionals who require transportation for business purposes, such as airport transfers, meetings, and corporate travel.

Targeting this segment can involve offering reliable and convenient vehicle booking services tailored to their specific needs, such as providing well-maintained vehicles, professional drivers, and seamless booking and invoicing processes.

Tourists and Travelers:

India is a popular tourist destination, attracting a large number of domestic and international travelers.

By targeting this segment, the startup can provide transportation services for tourists visiting popular destinations, offering options like sightseeing tours, multilingual drivers, and customized travel packages.

Event and Occasion-based Transportation:

This segment includes individuals or groups in need of transportation services for special events like weddings, parties, conferences, and festivals.

The startup can offer dedicated vehicles, luxury transportation options, and personalized services to cater to the unique requirements of such events.

Student Transportation:

Students often require transportation services for daily commuting to schools, colleges, or educational trips.

By focusing on this segment, the startup can provide safe and reliable transportation solutions, including options for group travel, fixed schedules, and tracking systems for parents or educational institutions.

Rural and Suburban Areas:

Many rural and suburban areas in India have limited public transportation infrastructure.

By targeting these areas, the startup can provide transportation services to bridge the gap, catering to the commuting needs of residents, providing connectivity to nearby towns, and enabling access to essential services.

Market Analysis

Vehicle Market Overview in India:

Provide an overview of the vehicle market in India, including its size, growth trends, and key players (such as Ola and Uber).

Highlight the challenges and opportunities present in the market.

Data Collection:

Explain the data collection process undertaken by the startup to gather information on the vehicle market, competitors, and customer insights.

Discuss the sources of data, such as market research reports, surveys, interviews, and focus groups.

Highlight any limitations or challenges faced during data collection.

Segmentation Variables:

Discuss the potential segmentation variables considered based on the available data.

Highlight geographic, demographic, psychographic, and behavioral variables that can be utilized.

Provide examples of segment categories derived from these variables, such as corporate travelers, tourists, event-based transportation, student transportation, and rural/suburban areas.

Target Market Selection

Segmentation Analysis:

Present the results of the segmentation analysis conducted by the startup.

Discuss the characteristics and preferences of each segment identified.

Highlight the size, growth potential, and profitability of each segment.

Competitive Analysis:

Analyze the strategies and offerings of existing competitors (Ola and Uber) within each identified segment.

Identify gaps or unmet needs in the market that can be leveraged by the startup.

Target Market Selection:

Based on the segmentation analysis and competitive analysis, identify the target market(s) that offer the most promising opportunities for profitability and growth.

Justify the selection of the target market(s) based on factors such as market size, growth potential, alignment with startup resources and capabilities, and competitive advantages.

Page 4: Feasible Strategy

Customer Profiling:

Develop detailed customer profiles or personas for the selected target market(s).

Describe the needs, preferences, pain points, and behaviors of customers within each segment.

Highlight the motivations for using vehicle booking services and the specific features or benefits they value the most.

Differentiation Strategy:

Discuss the startup's strategy for differentiating itself from competitors in the selected target market(s).

Outline the unique value proposition and key differentiating factors, such as pricing, service quality, convenience, branding, and customer acquisition channels.

Marketing and Service Delivery Strategy:

Detail the marketing and service delivery strategies to be implemented for each target market.

Describe the promotional channels, pricing strategies, service features, and customer engagement initiatives to attract and retain customers.

Pilot Testing and Iteration:

Explain the importance of pilot testing to validate the feasibility and effectiveness of the strategies.

Discuss the process of gathering feedback from customers and iterating the strategies based on the insights obtained.

BEHAVIOURAL SEGMENT

Behavioral Segment for Online Vehicle Booking Market:

Segment 1: Commuters seeking cost-effective transportation options

Behavior: These individuals are primarily concerned with finding affordable transportation solutions for their daily commute or occasional travel needs. They are price-sensitive and look for options that offer good value for money.

Characteristics: They may prefer shared rides, lower-priced vehicle categories, or promotional offers to save money. They prioritize convenience, reliability, and affordability.

Potential Strategy: Target this segment by offering discounted rates for shared rides during peak commuting hours or providing loyalty programs with rewards for frequent bookings. Emphasize the cost-effectiveness and reliability of the service.

Segment 2: Business travelers requiring premium and reliable transportation

Behavior: This segment consists of professionals and business travelers who prioritize comfort, reliability, and convenience. They are willing to pay a premium for a superior customer experience and timely service.

Characteristics: They may require spacious vehicles, professional drivers, and advanced booking options. They value punctuality, cleanliness, and professionalism in their transportation choices.

Potential Strategy: Offer a premium segment with high-end vehicles, well-trained drivers, and amenities like Wi-Fi, charging ports, and personalized customer service. Focus on reliability, timeliness, and the hassle-free experience for business travelers.

Segment 3: Tourists and explorers seeking guided transportation services

Behavior: This segment comprises travelers who are exploring new cities or tourist destinations and prefer guided transportation options to navigate unfamiliar areas.

Characteristics: They may seek local drivers with good knowledge of tourist spots, multilingual capabilities, and flexible itineraries. They value informative and entertaining experiences during their travels.

Potential Strategy: Create packages that offer guided tours, customized itineraries, and local insights. Collaborate with local tourism authorities, hotels, and travel agencies to attract tourists. Promote the convenience of hassle-free sightseeing and local expertise.

Segment 4: Special occasions and events transportation seekers

Behavior: This segment consists of individuals or groups looking for transportation solutions for special occasions such as weddings, parties, or corporate events.

Characteristics: They may require spacious vehicles, luxury options, and additional services like decoration, on-board entertainment, and event coordination.

Potential Strategy: Offer tailored packages for special occasions, including luxury vehicles, personalized decorations, and value-added services. Collaborate with event planners, wedding venues, and party organizers to establish partnerships and attract customers seeking transportation for their events.

Segment 5: Long-distance travelers and intercity transfers

Behavior: This segment comprises individuals or groups traveling between cities or seeking transportation for long-distance journeys.

Characteristics: They prioritize comfort, reliability, and timely transfers. They may require larger vehicles with ample luggage space and flexible booking options.

Potential Strategy: Focus on providing seamless intercity transfers with well-maintained vehicles, experienced drivers, and efficient scheduling. Offer competitive pricing for long-distance journeys and emphasize the comfort and convenience of door-to-door transportation.

By analyzing and targeting these different segments in the Indian vehicle market, the Online Vehicle Booking Product Startup can strategically position itself to generate early traction, increase market share, and drive revenue growth.

DEMOGRAPHIC SEMENTS

Urban Professionals:

Age: 25-40 years
Occupation: Working professionals, executives, entrepreneurs
Income: Medium to high income levels
Characteristics: Tech-savvy, time-constrained, prefer convenience and comfort
Potential Demand: Regularly require transportation for daily commuting, business meetings, and travel within the city.
Tourists and Travelers:

Age: Varied
Occupation: Leisure travelers, tourists, backpackers
Income: Varied, including budget travelers to high-end tourists
Characteristics: Seeking transportation for sightseeing, exploring tourist attractions, airport transfers, and intercity travel
Potential Demand: Tourists visiting popular destinations, leisure travelers exploring different cities, and foreigners needing reliable transportation services.
Students and Young Professionals:

Age: 18-25 years
Occupation: Students, interns, entry-level professionals
Income: Limited or moderate income

Characteristics: Budget-conscious, frequent travelers, prefer flexibility

Potential Demand: Students commuting to educational institutions, young professionals traveling for interviews, internships, or part-time jobs.

Senior Citizens:

Age: 55+ years

Occupation: Retirees, elderly individuals

Income: Varied, ranging from pension income to savings

Characteristics: Value safety, comfort, and convenience, may have limited mobility

Potential Demand: Seniors requiring transportation for medical appointments, social activities, shopping, and general commuting.

Business Travelers:

Age: Varied

Occupation: Corporate executives, business professionals

Income: High income levels

Characteristics: Frequent travelers, value time efficiency, professionalism, and reliability

Potential Demand: Business travelers attending meetings, conferences, corporate events, airport transfers, and intercity travel.

Families and Group Travelers:

Age: Varied (including adults and children)

Occupation: Families, friends, group travelers

Income: Varied, depending on the group

Characteristics: Require larger vehicles to accommodate multiple passengers, luggage, and often have specific travel needs

Potential Demand: Family vacations, group outings, weddings, events, airport transfers, and sightseeing tours.

By understanding and targeting these demographic segments, the online vehicle booking startup can develop specific marketing strategies, pricing plans, and service offerings tailored to the unique needs and preferences of each segment.

Geographic Segmentation Analysis for Online Vehicle Booking Service in India

Introduction:
In this analysis, we will explore the geographic segmentation of the vehicle market in India to identify potential segments that can generate early market entry and revenue for an online vehicle booking product startup. With fierce competition from established players like Ola and Uber, it is crucial to identify alternative segments where profitable opportunities exist. By understanding the diverse geographic landscape of India and its unique market characteristics, we can develop a feasible strategy to enter the market successfully.

Geographic Segmentation Analysis:
Geographic segmentation involves dividing a market into distinct geographical regions based on various factors such as location, climate, population density, infrastructure, and cultural preferences. Let's examine key geographic segments in India and evaluate their potential for vehicle booking services:

Tier 1 Cities:
Tier 1 cities, including Mumbai, Delhi, Bangalore, Chennai, and Kolkata, represent major economic and transportation hubs. These cities have high population densities, significant tourist traffic, and well-developed infrastructure. Targeting Tier 1 cities can provide a strong foundation for market penetration and revenue generation. Additionally, the presence of airports, business centers, and shopping districts makes these cities ideal for business travel and tourism-related vehicle bookings.

Tier 2 and Tier 3 Cities:
Tier 2 and Tier 3 cities across India offer considerable growth potential due to their expanding economies and increasing urbanization. These cities often lack efficient public transportation systems, making private vehicle bookings more desirable. By providing reliable and convenient vehicle booking services in these cities, the startup can tap into an underserved market segment. Customized marketing strategies tailored to the unique

characteristics and preferences of each city will be essential for success in these regions.

Tourist Destinations:
India boasts a rich cultural heritage and diverse tourist destinations, including historical sites, hill stations, and coastal regions. Tourists often require reliable transportation options to explore these areas, making them attractive segments for vehicle booking services. Targeting popular tourist destinations such as Goa, Jaipur, Agra, Shimla, and Kerala can provide a steady stream of bookings throughout the year. Offering value-added services like guided tours or packages tailored to specific tourist interests can further differentiate the startup in this segment.

Metropolitan Suburbs:
Metropolitan suburbs, such as Navi Mumbai, Gurgaon, and Whitefield, are witnessing rapid urbanization and a surge in residential and commercial development. Commuting to city centers can be challenging, making vehicle booking services a viable alternative to private vehicle ownership. By focusing on these suburban areas, the startup can cater to the needs of commuters, daily travelers, and residents who prefer convenient transportation options.

Feasible Market Entry Strategy:
To enter the market successfully and generate revenue, the startup can consider the following strategies:

Targeted Marketing Campaigns: Develop targeted marketing campaigns that highlight the convenience, reliability, and affordability of the online vehicle booking service. Emphasize features such as real-time tracking, driver ratings, and multiple vehicle options.

Partnerships: Forge strategic partnerships with hotels, travel agencies, and airlines to offer bundled services or exclusive discounts. This can help increase visibility and attract customers who are already in need of transportation services.

Customized Solutions: Tailor the service offerings to the specific needs of each segment. For example, focus on business-friendly features for Tier 1 cities, family-oriented packages for tourist destinations, and cost-effective solutions for suburban commuters.

Localized Operations: Establish local offices or partner with local vehicle rental companies in targeted regions. This will enhance operational efficiency, provide better customer support, and foster trust among customers.

Technology Advancements: Invest in innovative technologies such as mobile apps, seamless payment systems, and AI-driven features to enhance the user experience. Integrate with popular digital platforms

PYSCOGRAPHIC SEGMENT

Tech Enthusiasts:

Characteristics: Early adopters of technology, enjoy using mobile apps and online platforms for convenience and efficiency
Preferences: Prefer booking vehicles through mobile apps, value features like real-time tracking, cashless payments, and personalized experiences
Eco-conscious Consumers:

Characteristics: Environmentally conscious individuals who prioritize sustainable transportation options
Preferences: Prefer electric or hybrid vehicles, eco-friendly transportation options, and companies with a strong commitment to reducing carbon footprint
Luxury Seekers:

Characteristics: Individuals who seek premium, high-end experiences
Preferences: Demand luxury vehicles, professional chauffeurs, personalized services, and additional amenities during the ride
Adventurous Explorers:

Characteristics: Adventure enthusiasts who enjoy exploring new places and engaging in outdoor activities

Preferences: Prefer vehicles suitable for off-road adventures, adventure travel packages, and options for equipment rentals (e.g., bikes, kayaks) during their trips

Time-pressed Professionals:

Characteristics: Busy professionals who value their time and seek efficient transportation solutions

Preferences: Look for reliable and prompt service, prefer pre-booking options, prioritize minimal wait times, and may require additional services like airport pick-up and drop-off

Value-oriented Consumers:

Characteristics: Budget-conscious individuals who seek cost-effective transportation options without compromising on quality

Preferences: Look for competitive pricing, discounts, loyalty programs, and deals on group bookings or long-term rentals

By understanding the psychographic profiles of the target market segments, the online vehicle booking startup can customize their marketing messages, user experiences, and service offerings to effectively resonate with the preferences and values of each segment. This will help in attracting and retaining customers within their target market segments.

ANALYSIS AND APPROACHES FOR SEGMENTATION

Clustering

Clustering, an exploratory data analysis technique, is widely employed to gain insights into the underlying structure of data. Its objective is to identify distinct subgroups within the data, ensuring that data points within the same subgroup (cluster) exhibit high similarity, while data points in different clusters are markedly dissimilar. In essence, the aim is to discover homogeneous subgroups within the data, maximizing the similarity of data points within each cluster using a chosen similarity measure, such as euclidean-based distance or correlation-based distance. The selection of the similarity measure depends on the specific application. Clustering analysis can be conducted based on features, seeking subgroups of samples, or based on samples, aiming to find subgroups of features.

K-means Algorithm

The K-means algorithm is an iterative method that aims to divide a dataset into distinct and non-overlapping subgroups (clusters), where each data point belongs exclusively to one group. Its objective is to maximize the similarity among data points within the same cluster while maximizing the dissimilarity (distance) between clusters. The

algorithm assigns data points to clusters by minimizing the sum of squared distances between the data points and the centroid of each cluster. By minimizing the variation within clusters, the data points within the same cluster become more homogeneous.

The K-means algorithm operates as follows:

1. Specify the desired number of clusters, K.
2. Initialize the centroids by shuffling the dataset and randomly selecting K data points as centroids without repetition.
3. Iterate until the centroids remain unchanged, meaning that the assignment of data points to clusters no longer changes.

The K-means algorithm follows the expectation-maximization approach to solve the problem. The E-step involves assigning data points to the closest cluster, while the M-step involves computing the centroid of each cluster. The algorithm continues to alternate between these two steps until convergence. Here is a mathematical breakdown of how the problem can be solved:

1. Define the number of clusters, K.
2. Randomly initialize the centroids c_1, c_2, ..., c_K.
3. Repeat until convergence:

- For each data point x_i:
  - ✧ Calculate the distance between x_i and each centroid c_k.
  - ✧ Assign x_i to the cluster with the closest centroid.
- For each cluster k:

- ✧ Update the centroid c_k as the arithmetic mean of all data points assigned to cluster k.

4. Convergence is achieved when the centroids no longer change.
5. The final result is a set of K clusters, where each data point belongs to one cluster.

By iteratively updating the assignments and centroids, the K-means algorithm aims to find the optimal partitioning of the dataset into K clusters.

Applications:

The K-means algorithm finds extensive use in diverse applications, including market segmentation, document clustering, image segmentation, and image compression, among others. When performing cluster analysis, the primary objectives typically fall into two categories:

1. Gaining meaningful insights into the underlying structure of the data at hand.
2. Adopting a cluster-then-predict approach, where distinct models are constructed for different subgroups if there is a significant variation in the behaviors exhibited by these subgroups.

In summary, the K-means algorithm serves various purposes, allowing practitioners to explore data structure and apply tailored models based on subgroup characteristics when warranted.

IMPLEMENTATION

Analysis Car Details Dataset

Introduction

This data is collected from 'Car Dekho'.
Following details of cars are included in the dataset:
1) Car name 2) Year 3) Selling Price 4) Kms driven 5) Fuel 6) Seller type 7)Transmission 8) Owner
I will try to understand the data, analyze it, extract reports from it, and try to understandthe relationships between the different variables

Data Exploration

Data path extraction

/kaggle/input/car-details-dataset/CAR DETAILS FROM CAR DE KHO.csv

In [1]:

```
# data path extraction
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

Now we know the path of the data, we will use it later in the code that read the

datasetImport libraries that we will need in the beginning

In [2]:

```
#Import libraries What will be needed for data analysis?
import numpy as np #to deal with data
import pandas as pd #to deal with data
import matplotlib.pyplot as plt#to deal with graphic charts
import seaborn as sns #to deal with graphic charts
```

By use pandas we will read our data

In [3]:

```
# by use of pandas we will read our data from CSV file
df=pd.read_csv('/kaggle/input/car-details-dataset/CAR DETAILS FRO
```

The first step, we want to know what dataset contains

In [4]:

```
# we want to know about data
df.head()
```

Out[4]:

	name	year	selling_price	km_driven	fuel	seller_type	transmis
0	Maruti 800 AC	2007	60000	70000	Petrol	Individual	Ma
1	Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	Ma
2	Hyundai Verna 1.6 SX	2012	600000	100000	Diesel	Individual	Ma
3	Datsun RediGO T Option	2017	250000	46000	Petrol	Individual	Ma
4	Honda Amaze VX i-DTEC	2014	450000	141000	Diesel	Individual	Ma

Now we know some information about the data and we extracted the first five rows fromit

Description of the data

In [5]:

In [4]:

```
#Description of the data
df.describe().T
```

Out[5]:

	count	mean	std	min	25%
year	4340.0	2013.090783	4.215344	1992.0	2011.00
selling_price	4340.0	504127.311751	578548.736139	20000.0	208749.75
km_driven	4340.0	66215.777419	46644.102194	1.0	35000.00

This gave us some more information

We want to know the size of the data and the number of rows and columns

In [6]:

```
# to know shape of data
df.shape
```

Out[6]:

(4340, 8)

Data contains (4340, 8) coulmns and rows

In [7]:

```
# to know more more information about the data
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4340 entries, 0 to 4339
Data columns (total 8 columns):
 #   Column         Non-Null Count  Dtype 
---  ------         --------------  ----- 
 0   name           4340 non-null   object
 1   year           4340 non-null   int64 
 2   selling_price  4340 non-null   int64 
 3   km_driven      4340 non-null   int64 
 4   fuel           4340 non-null   object
 5   seller_type    4340 non-null   object
 6   transmission   4340 non-null   object
 7   owner          4340 non-null   object
dtypes: int64(3), object(5)
memory usage: 271.4+ KB
```

We want to get the column names

In [8]:

```
#To get the column names
df.columns
```

Out[8]:

```
Index(['name', 'year', 'selling_price', 'km_driven', 'fuel', 'seller_type',
       'transmission', 'owner'],
      dtype='object')
```

We want to extract column types

In [9]:

```
#To get the column types
df.dtypes
```

Out[9]:

```
name             object
year              int64
selling_price     int64
km_driven         int64
fuel             object
seller_type      object
transmission     object
owner            object
dtype: object
```

We want to extract unique values

In [10]:

```
# to know unique values
print(df.name.unique())
print(sorted(df.year.unique()))
print(df.fuel.unique())
print(df.seller_type.unique())
print(df.transmission.unique())
print(df.owner.unique())
```

```
['Maruti 800 AC' 'Maruti Wagon R LXI Minor' 'Hyundai Verna 1.6
SX' ...
 'Mahindra Verito 1.5 D6 BSIII'
 'Toyota Innova 2.5 VX (Diesel) 8 Seater BS IV''Hyundai
 i20 Magna 1.4 CRDi']
[1992, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 20
03, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012,
2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020]
['Petrol' 'Diesel' 'CNG' 'LPG' 'Electric']
['Individual' 'Dealer' 'Trustmark Dealer']['Manual'
'Automatic']
['First Owner' 'Second Owner' 'Fourth & Above Owner' 'Thi rd
Owner'
 'Test Drive Car']
```

To find if there are null values or not

In [11]:

```
# find null values
df.isnull().sum()
```

Out[11]:

```
name             0
year             0
selling_price    0
km_driven        0
fuel             0
seller_type      0
transmission     0
owner            0
dtype: int64
```

In [10]:

To see if there are duplicate values

In [12]:

```
#find duplicate values
df.duplicated().sum()
```

Out[12]:

763

To view duplicate values

In [13]:

```
# view duplicate values
df[df.duplicated()]
```

Out[13]:

	name	year	selling_price	km_driven	fuel	seller_type	tran
13	Maruti 800 AC	2007	60000	70000	Petrol	Individual	
14	Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	
15	Hyundai Verna 1.6 SX	2012	600000	100000	Diesel	Individual	
16	Datsun RediGO T Option	2017	250000	46000	Petrol	Individual	
17	Honda Amaze VX i-DTEC	2014	450000	141000	Diesel	Individual	
...	...	...	...	...	...	...	
4307	Mahindra Xylo H4	2019	599000	15000	Diesel	Individual	
4308	Maruti	2018	200000	35000	Petrol	Individual	
4309		2017	350000	10171	Petrol	Dealer	
	4310	2015	465000	41123	Diesel	Dealer	
		2017	1900000	20118	Petrol	Dealer	

In [13]:

	Alto 800 LXI						
431							
1	Datsun GO Plus T						
	Renault Duster 110PS Diesel RxL						
	Toyota Camry Hybrid 2.5						

763 rows × 8 columns

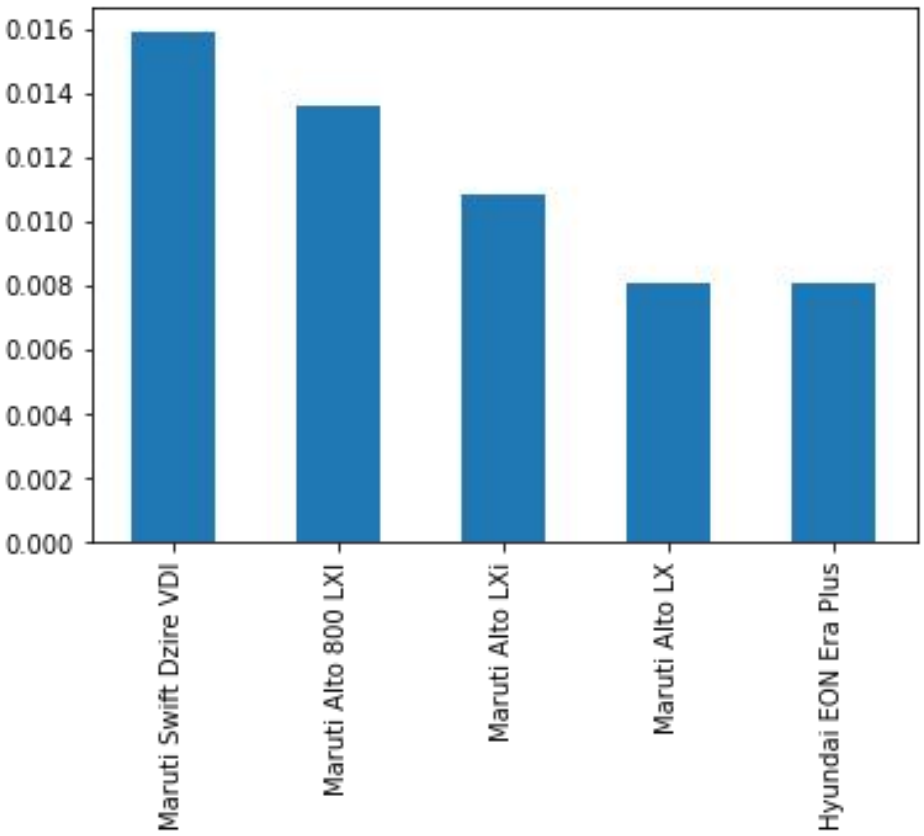
There are duplicate values, but we will not delete them because it is natural that thereare duplicate values, normal same model car was sold many times

Data Analysis 1

Now we move on to the next stage, which is data analysis
Let's start trying to find out What is the most sold car??

In [14]:

```
#we use (value_counts) to know to know the sold numbers of each car
df["name"].value_counts(normalize = True)[:5].plot(kind = 'bar')
plt.show()
```



Now we have a note on the names of the cars in detail
We want to extract reports based on the brands,so we will add a column contain thebrand for each car
and we will build it by separating the first part of the name of each car in to create anew column with the brands only

In [15]:

```
#add new coulmn by brand name split first part
df["brand"] = df.name.apply(lambda x : ' '.join(x.split(' ')[:1]))#
df['brand'] # New column review
```

Out[15]:

```
0          Maruti
1          Maruti
2         Hyundai
3          Datsun
4           Honda
           ...
4335      Hyundai
4336      Hyundai
4337       Maruti
4338      Hyundai
4339      Renault
Name: brand, Length: 4340, dtype: object
```

We want to know the number of values in the column

In [16]:

```
#To know countnvalue in coulmn
df.brand.value_counts()
```

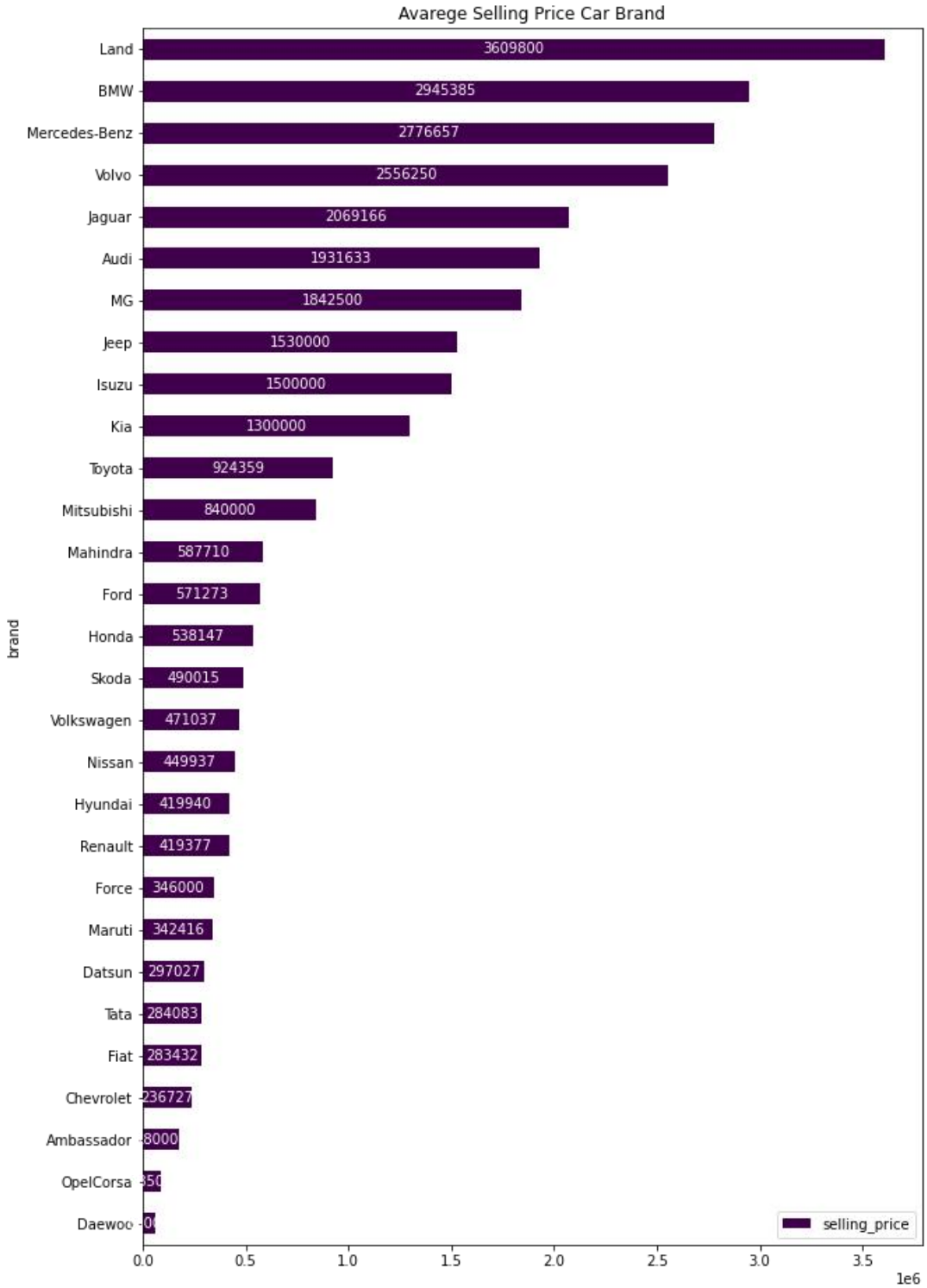
Out[16]:

```
Maruti           1280
Hyundai           821
Mahindra          365
Tata              361
Honda             252
Ford              238
Toyota            206
Chevrolet         188
Renault           146
Volkswagen        107
Skoda              68
Nissan             64
Audi               60
BMW                39
Fiat               37
Datsun             37
Mercedes-Benz      35
Jaguar              6
Mitsubishi          6
Land                5
Volvo               4
Ambassador          4
Jeep                3
MG                  2
OpelCorsa           2
Daewoo              1
Force               1
Isuzu               1
Kia                 1
Name: brand,    dtype: int64
```

We want to know the average prices of the brands

In [17]:

```
#we use groupby and mean to extract values and plot to draw the gra
price = df.groupby(['brand'])[['selling_price']].mean()
price.sort_values(by='selling_price', ascending=True, inplace=True) ax =
price.plot(kind='barh', cmap='PRGn' , figsize=(10,16) ,title for c in
ax.containers:
        # set the bar label
    ax.bar_label(c, fmt='%.0f',label_type='center', color='w',rotat
```



Now trying to find out the sales classified by brand

In [18]:

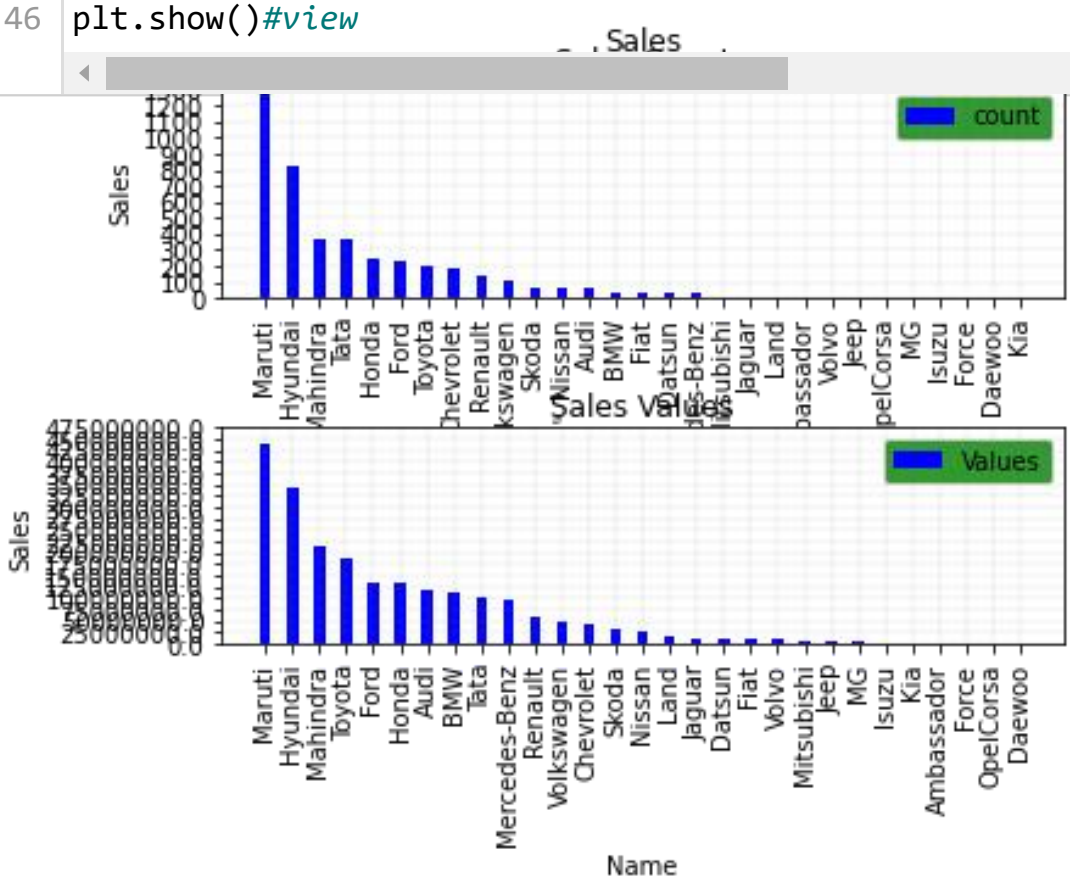
```
# We will display sales by count and value in separate graphs

 #plot 1:
 data = df.groupby(['brand'])['brand'].count().sort_values(ascending
 x = data.index # to extract the brand name
 y = data.values # to extract the count to brand
 plt.subplot(2, 1, 1)#The location of the first graph 2 1 1
 #the figure has 2 row, 1 columns, and this plot is the first plot.
 plt.bar(x, y, color ='blue',width = 0.4)#Fomat to plt.bar
 plt.rcParams['axes.facecolor'] = '#FFFFFF'#background color
 plt.xticks(rotation=90) #Make the text of the label Make the text o
 plt.xlabel("Name",fontsize=10,color="black")#Fomat and name to x
 plt.ylabel("Sales",fontsize=10,color="black")#Fomat and name to y
 plt.title("Sales Gount",color="black")#Fomat and name to title
 plt.legend(["count"], loc ="upper right" ,facecolor='green', labelc
 plt.rcParams['figure.figsize'] = [10, 10]#Determine the size of the
 plt.grid(color='grey', linestyle='-', linewidth=.1)#Fomat grid netw
#plt.xlim([0, 1]) #
#plt.ylim([0, 2000])#
 #plt.locator_params(axis='x', nbins=20)#
 plt.locator_params(axis='y', nbins=20)# to make y texts 100-200-300

 #plot 2:
 plt.subplot(2, 1, 2)#The location of the second graph 2 1 2
 #the figure has 2 row, 1 columns, and this plot is the second plot.
 data = df.groupby(['brand'])['selling_price'].sum().sort_values(asc
 x = data.index # to extract the brand name
 y = data.values #to extract the sum to brand
 plt.bar(x, y, color ='blue',width = 0.4)
 plt.rcParams['axes.facecolor'] = '#FFFFFF'
 plt.xticks(rotation=90)
 plt.xlabel("Name",fontsize=10,color="black")
 plt.ylabel("Sales",fontsize=10,color="black")
 plt.title("Sales Values",color="black")
 plt.legend(["Values"], loc ="upper right" ,facecolor='green', label
 plt.rcParams['figure.figsize'] = [10, 10]
 plt.grid(color='grey',      linestyle='-',      linewidth=.1)     38
#plt.xlim([0, 1])
#plt.ylim([0, 2000])
 #plt.locator_params(axis='x', nbins=20)
 plt.locator_params(axis='y', nbins=20)
 #plt.margins(x=0, y=0)
```

```
43 plt.yticks(ticks=plt.yticks()[0], labels=plt.yticks()[0])# Show rea
44 plt.subplots_adjust(left=0.1, bottom=0.1, right=0.9,top=0.9,wspace=
45 plt.suptitle("Sales")#Name for the whole graph
```

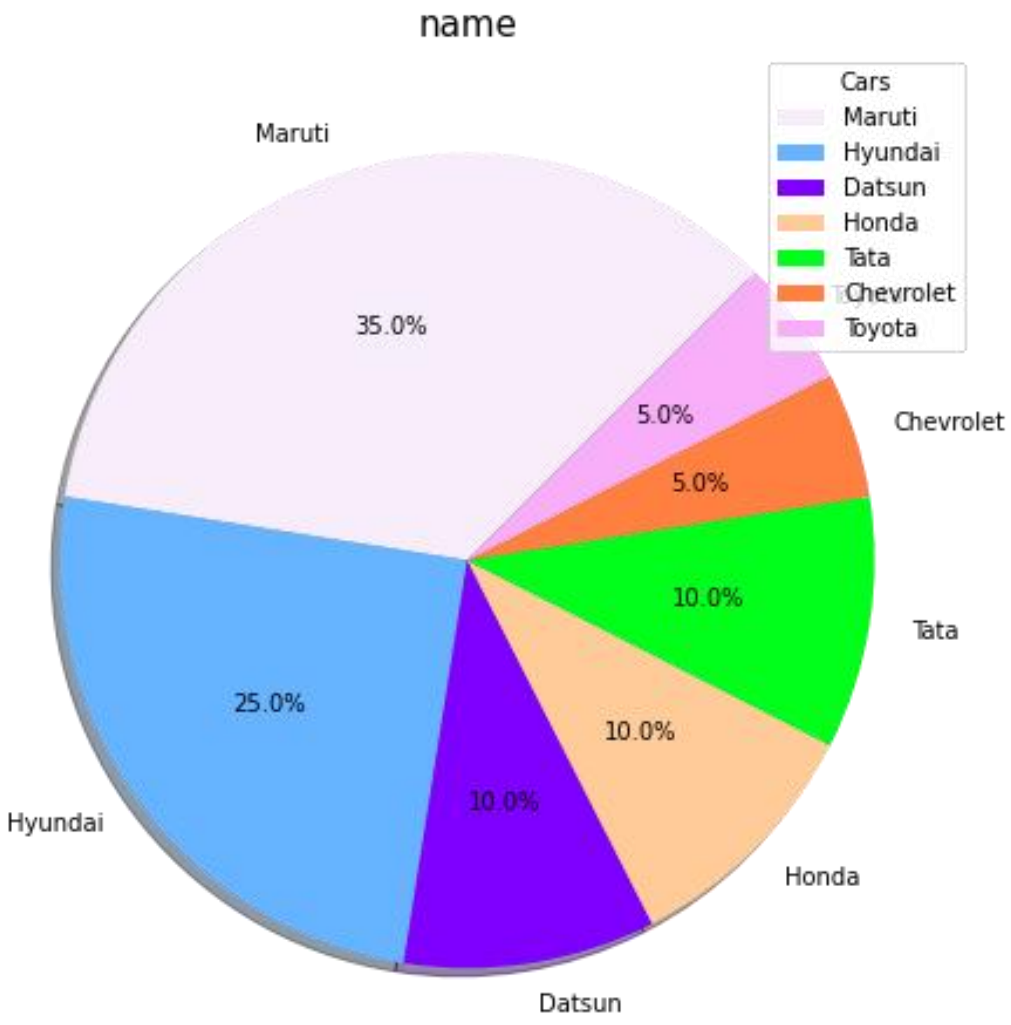
```
46 plt.show()#view
```



We will try another graph to display sales

In [19]:

```
## We will display sales by count in pie graphs
labels = df["brand"][:20].value_counts().index #We chose only twent
sizes = df["brand"][:20].value_counts() # We chose only twenty
data = df.groupby(['brand'])['brand'].count().sort_values(ascending
x = data.index #to extract the brand name
y = data.values#to extract the count to brand
colors = ['#F8EEFB','#66b3ff','#8000FF','#ffcc99',"#00FF1B","#FF804
plt.figure(figsize = (8,8))#Determine the size of the graph
# Creating explode data
#explode = (0.1, 0.0, 0.2, 0.3, 0.0, 0.0)
plt.pie(sizes, labels=labels, rotatelabels=False, autopct='%1.1f%%'
plt.title('name',color = 'black',fontsize = 15)#Fomat title
#plt.legend()#
plt.legend(title = "Cars")#title legend
#plt.legend(wedges, cars, title ="Cars",loc ="center left",bbox_to_
#myexplode = [0.2, 0, 0, 0]
plt.show()#view
```



We note that Maruti Hyundai best selling brands now from just looking at the graphs we have information about what is the most best selling brand , by count , by value we have now an idea about the sales, categorized by car brand but we also want analyze the data from another point of view for more comprehensive vision

Analysis by year

In [20]:

```
1  #To know countnvalue in year coulmn
2  df.year.value_counts()
```

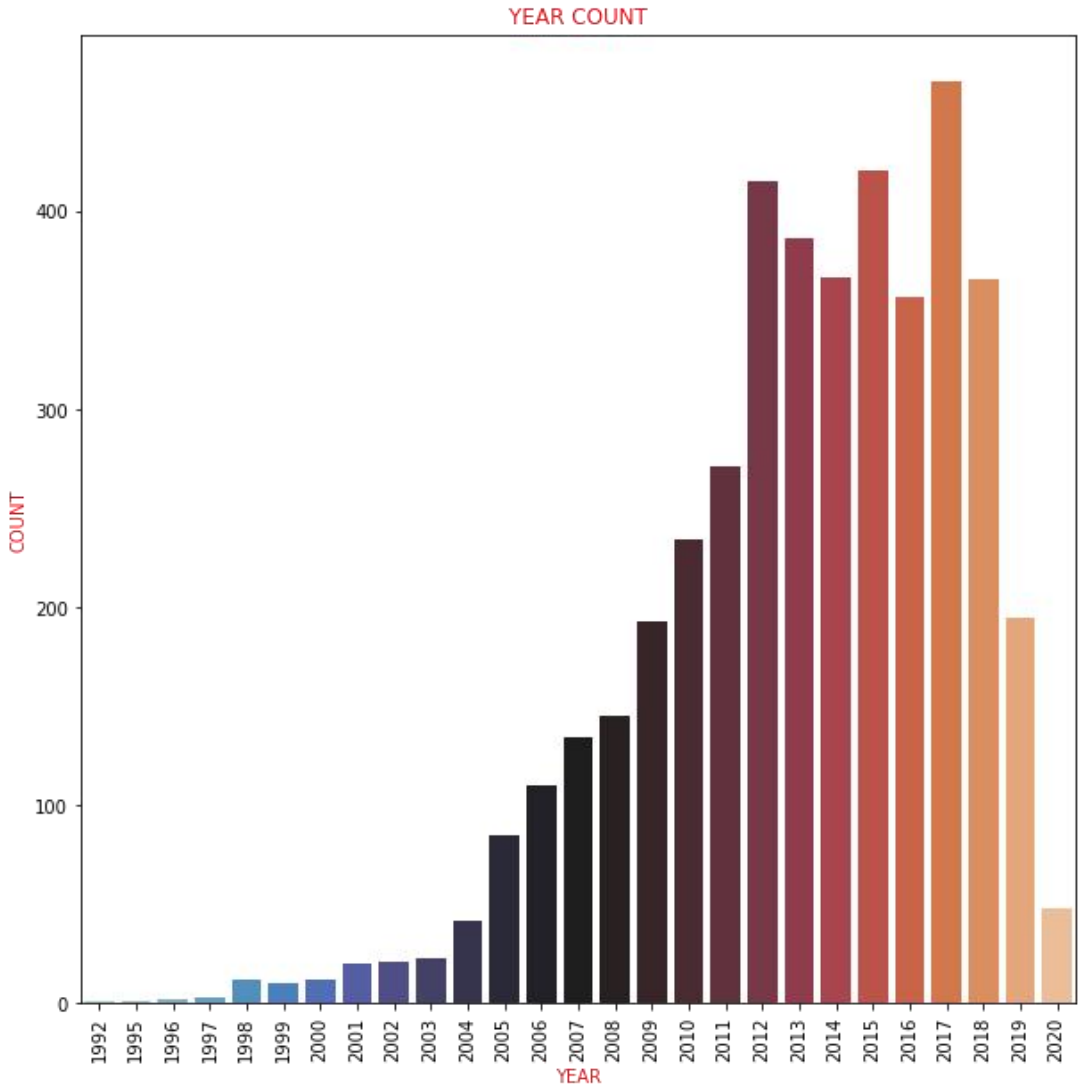
Out[20]:

```
2017  466
2015  421
2012  415
2013  386
2014  367
2018  366
2016  357
2011  271
2010  234
2019  195
2009  193
2008  145
2007  134
2006  110
2005  85
2020  48
2004  42
2003  23
2002  21
2001  20
1998  12
2000  12
1999  10
1997  3
1996  2
1995  1
1992  1
Name year,    dtype:   int64
```

:

In [21]:

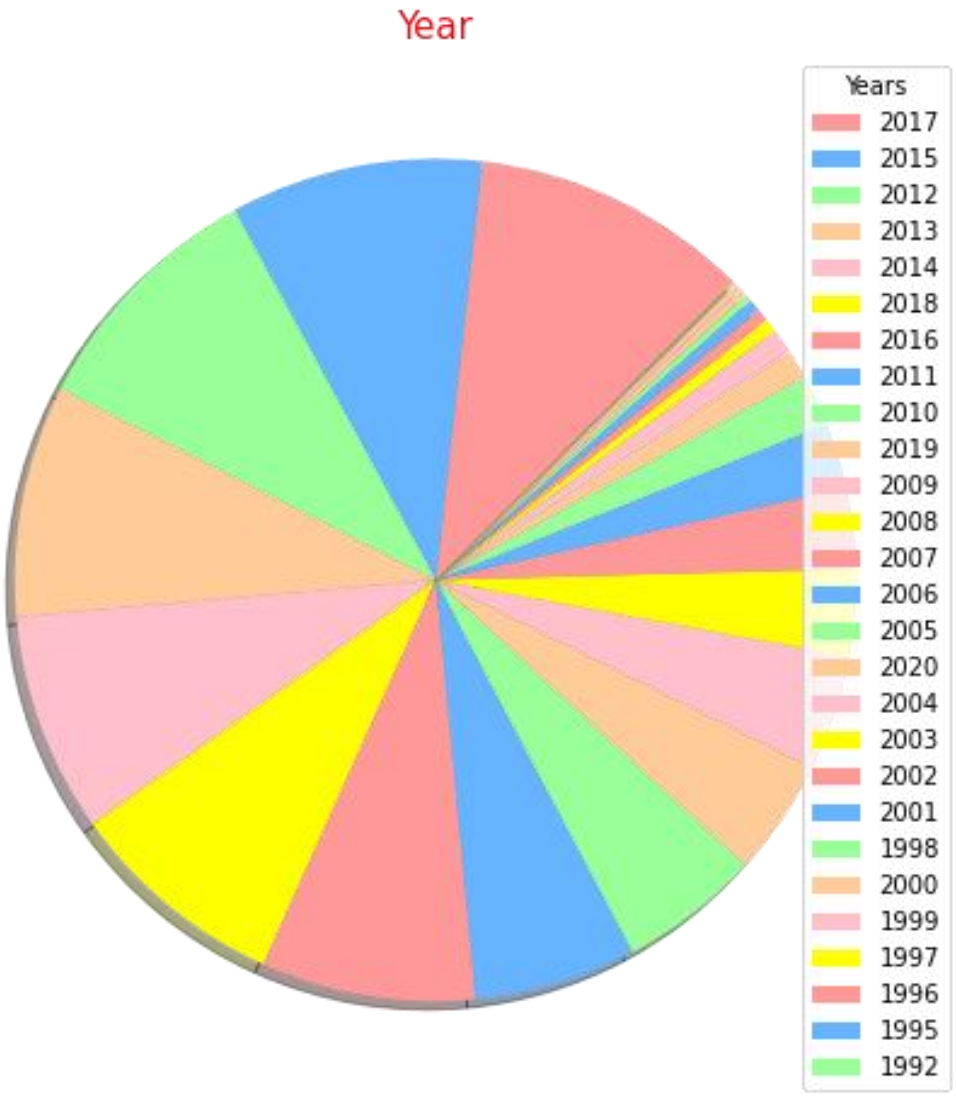
```
#Here I am using another way to display the graph by seaborn we imp
sns.countplot(data=df,x="year",palette="icefire")
plt.xticks(rotation=90)
plt.xlabel("YEAR",fontsize=10,color="RED")
plt.ylabel("COUNT",fontsize=10,color="RED")
plt.title("YEAR COUNT",color="RED")
plt.show()
```



Just by looking at the graph we know which year sales in it the highest, and see thevariance between sales over the years

In [22]:

```
#We will display pie graphs
labels = df["year"].value_counts().index
sizes = df["year"].value_counts()
colors = ['#ff9999','#66b3ff','#99ff99','#ffcc99',"pink","yellow"]
plt.figure(figsize = (8,8))
plt.pie(sizes, labels=labels , rotatelabels=False, autopct=None,col
plt.title('Year',color = 'red',fontsize = 15)
plt.legend(title = "Years", loc='upper right')#title legend
plt.show()
```

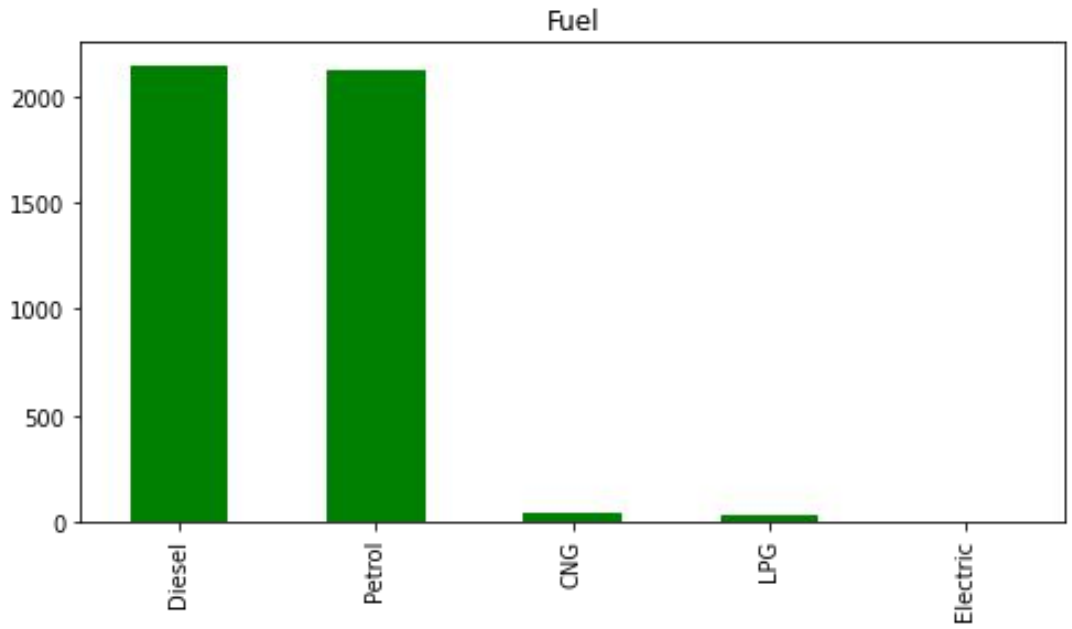


If it is noticed that sales are not stable during the years, they were increase during thefirst years, but last two years decreased a lot

Another graph to view sales classified by fuel

In [23]:

```
#We will display bar graphs
df["fuel"].value_counts(sort=True).plot(kind="bar", color=["green"
```

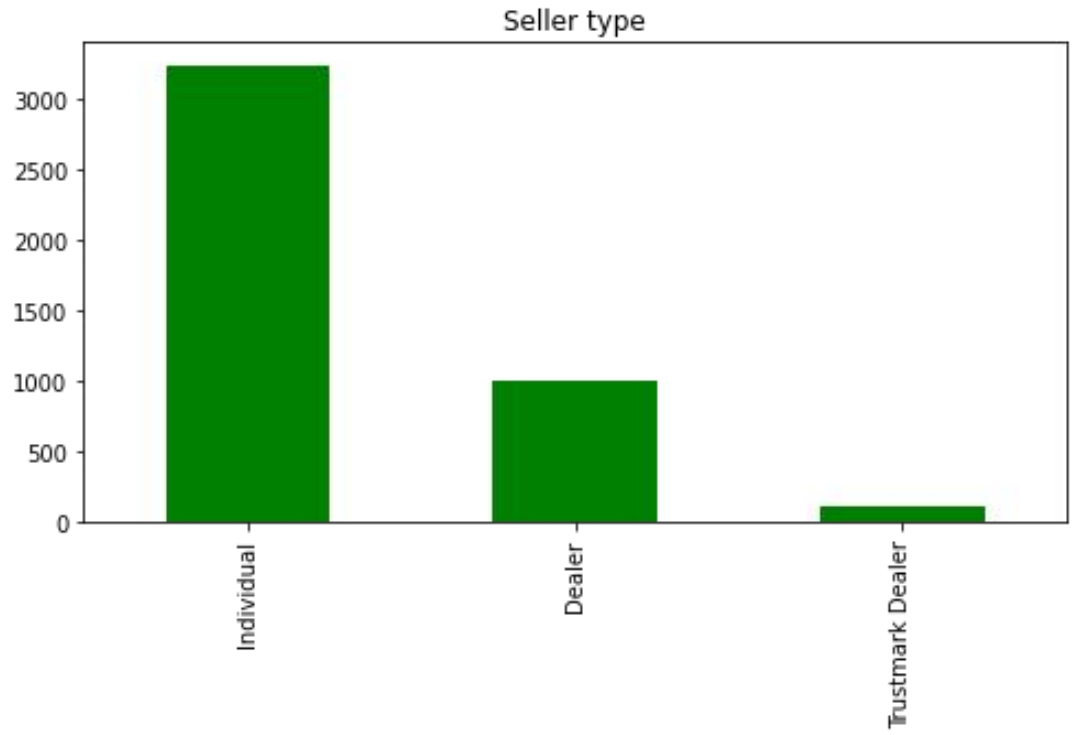


Best sales diesel and petrol it is clear that the type of fuel affects on

salesGraph to view sales classified by seller type

In [24]:

```
#We will display bar graphs
df["seller_type"].value_counts(sort = True).plot(kind="bar", color=
```

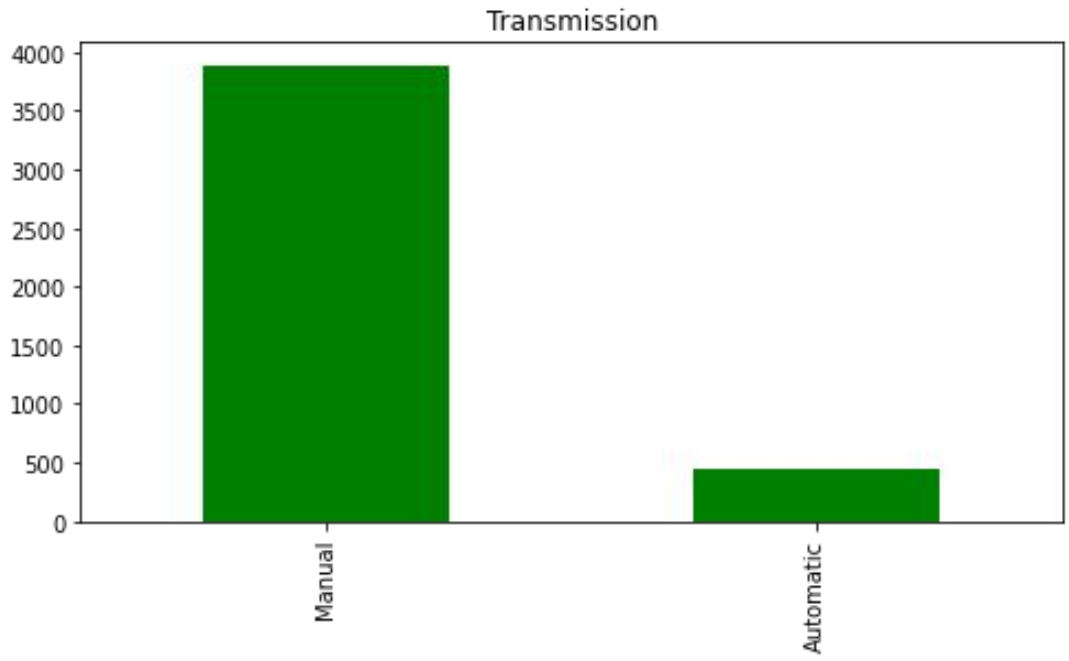


Best sales come from individual this note is important maybe

unexpectedGraph to view sales classified by transmission

In [25]:

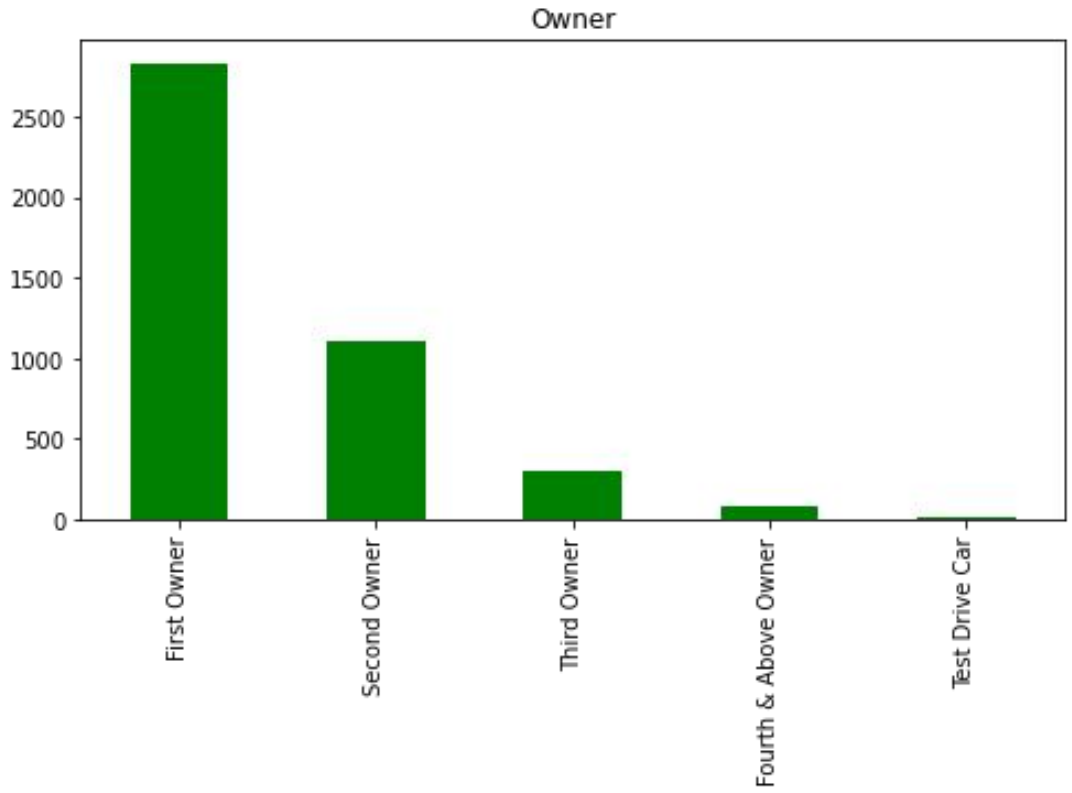
```
#We will display bar graphs
df["transmission"].value_counts(sort = True).plot(kind="bar", color
```



Best sales come from manual or automatic this is a strange thing
Expected that the automatic sell more, but what happened is that the manual is the sellmore
so we should search and ask for the reason

In [26]:

```
#We will display bar graphs
df["owner"].value_counts(sort = True).plot(kind="bar", color=["gree
```



Best sales be the first owner this is expected and normal

Data Analysis 2

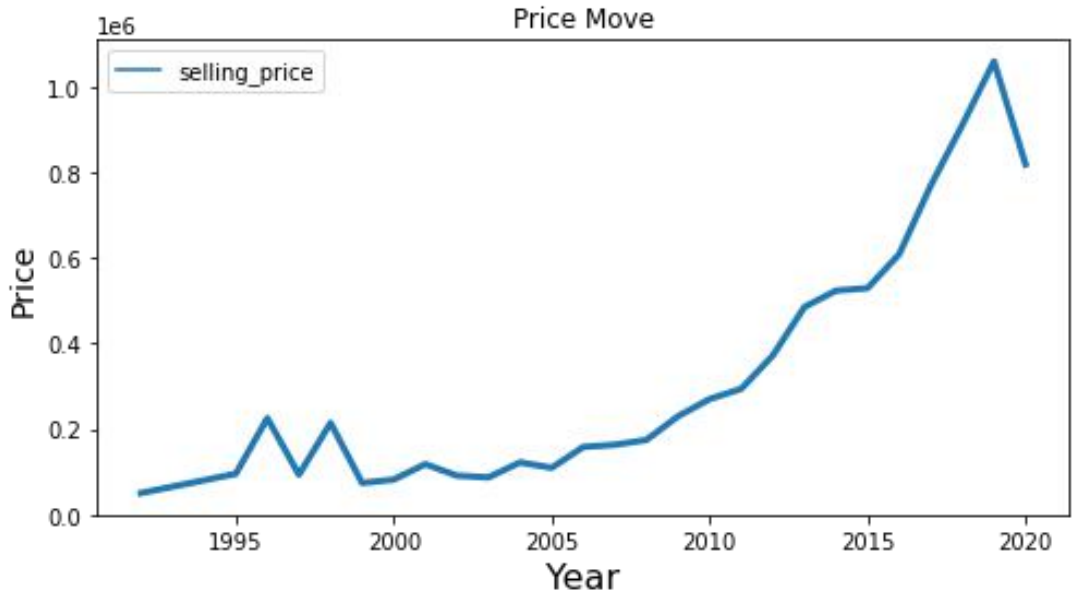
Let's now move to another stage of data analysis
I want to know a little about the relationships between different factors and the price ofa car
I want to know about the relationship between the various factors and price of the carWhat is the most important factor affecting the price of a car?

Let's draw a graph linking the year to the selling price

In [27]:

```
#We will display bar graphs line
def line_plot(data, title ,xlabel, ylabel):
        plt.figure(figsize=(8, 4))
        sns.lineplot(data=data , palette="tab10", linewidth=3.0)
        plt.title(title, fontsize=12)
        plt.ylabel(ylabel, size=14)
        plt.xlabel(xlabel, size=16)

df_price_move = df.groupby(['year'])[['selling_price']].mean()
line_plot(df_price_move,'Price Move', 'Year', "Price")
```



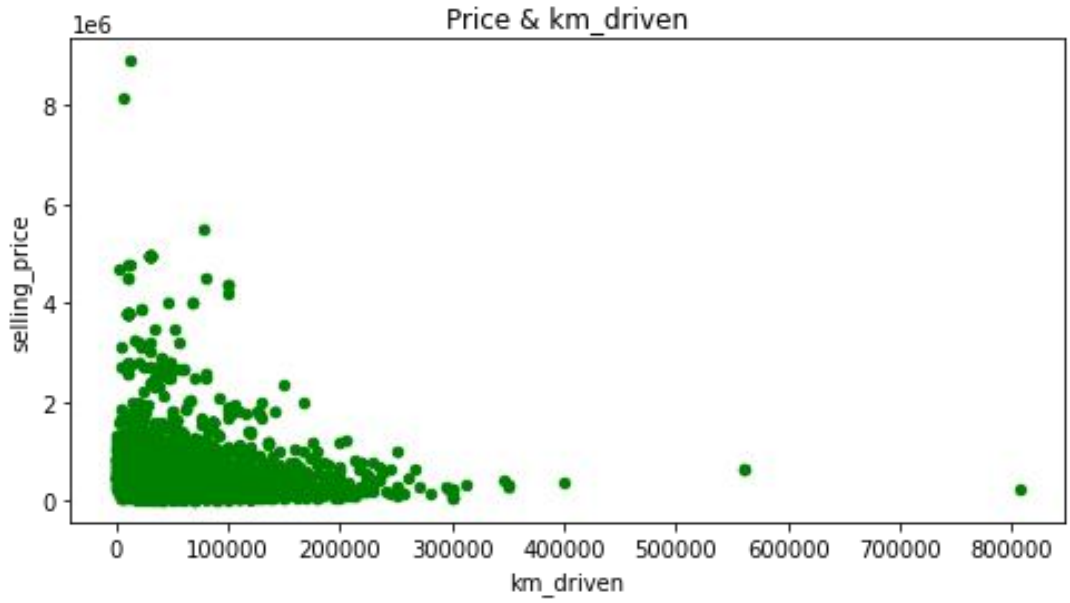
The price always increases over the years, and this is normal and expected, but thelast years get down
and this is not normal
We should ask about and research about this

matter. The relationship between km_driven and

price

In [28]:

```
#We will display bar graphs scatter
df.plot(x="km_driven", y="selling_price", kind="scatter", figsize=(
```



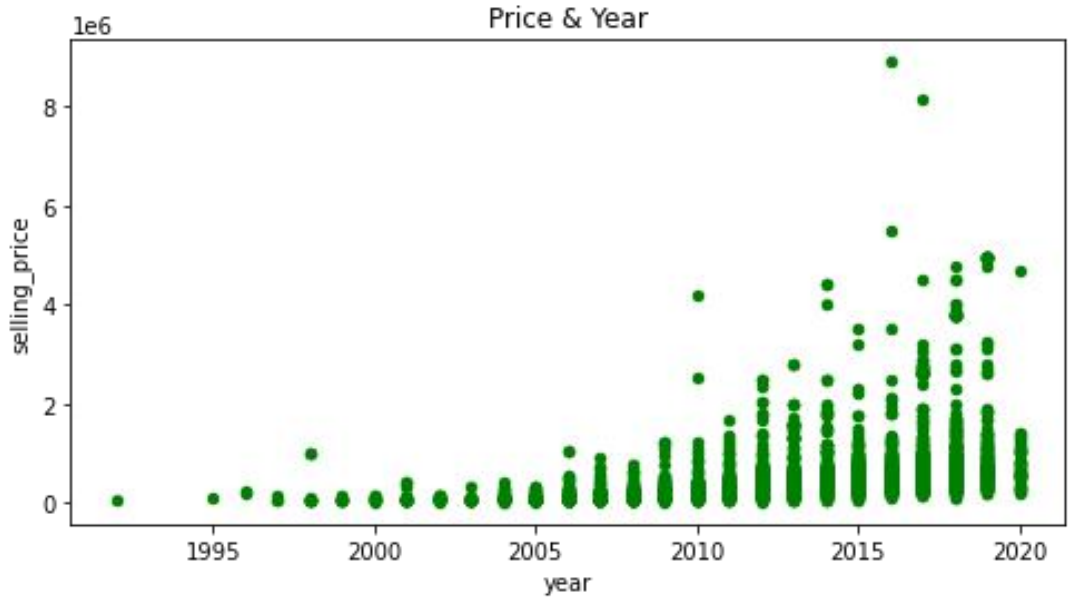
An expected result, kilometers more mean the price of the car get

downThe relationship between km_driven and year

In [29]:

```
#We will display bar graphs scatter
df.plot(x="year", y="selling_price", kind="scatter", figsize=(8, 4)
```

In [29]:


Price & Year
1e6
8
6
4
2
0
selling_price
1995
2000
2005
2010
2015
2020
year

This result we got before from an different graph , but here we just Show it by differentgraph

Now examine the correlation between the different factors

In [30]:

```
1  # finds the correlation
2  df.corr()
```

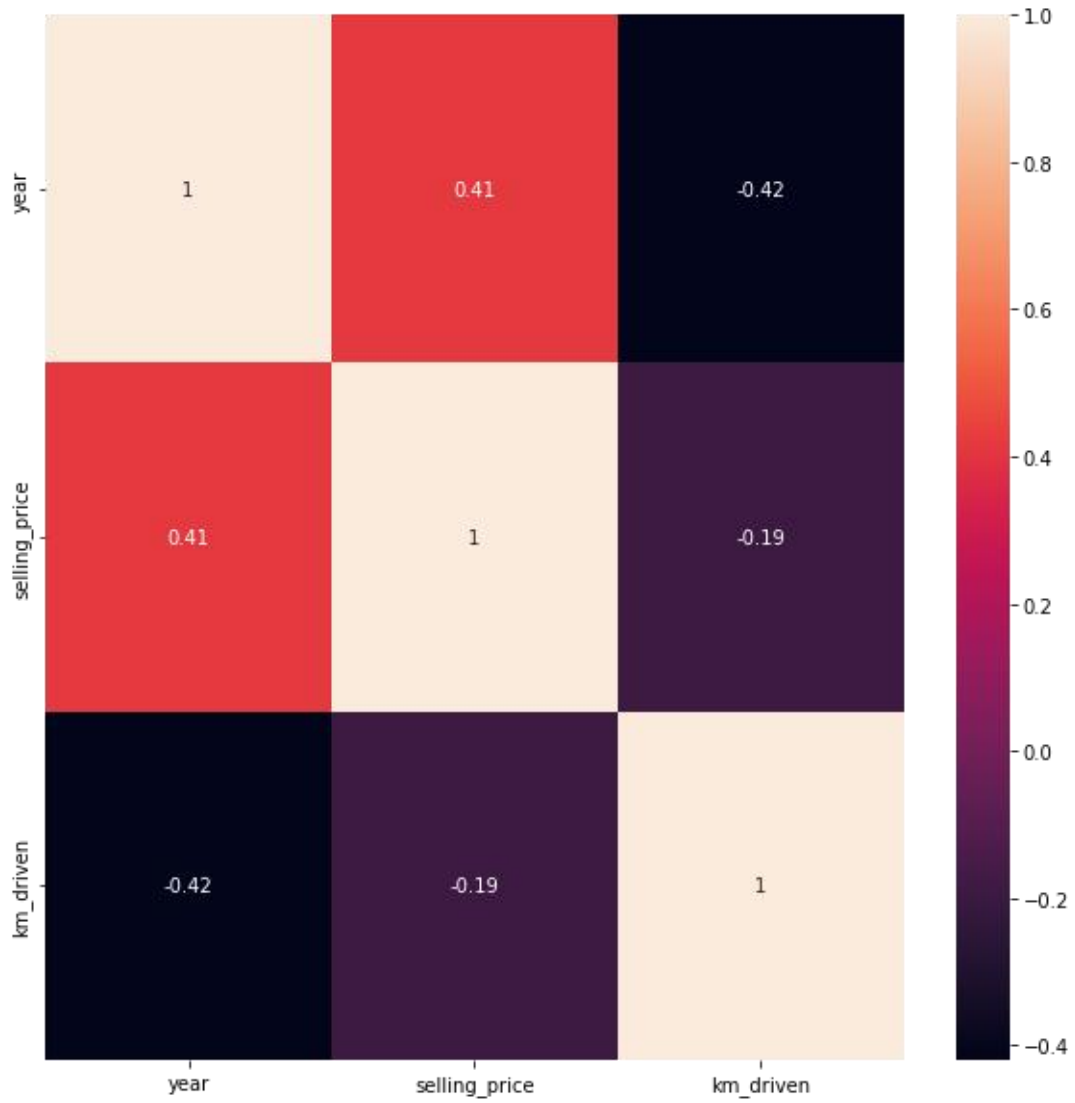
Out[30]:

	year	selling_price	km_driven
year	1.000000	0.413922	-0.419688
selling_price	0.413922	1.000000	-0.192289
km_driven	-0.419688	-0.192289	1.000000

we extracted the numeric factors and ignored the other factors because they are textsand this is a problem
Note here the positive correlation between the year and the price, meaning that everynew year the price increases
and a negative correlation between kilometers abnd price when it increases, the pricedecreases Let's show that in heatmap

In [31]:

```
#We will display heatmap
corr = df.corr()
corr = corr['selling_price']
corr = corr.sort_values(ascending=False)
sns.heatmap(df.corr(), annot=True)
plt.show()
```



Now we will try to find solution to the problem that there are many factors that affect onprice but they are not numerical

Therefore, we need to modify them, they should be numbers so that we can drawanother different heat map that includes all factors.

Extracting non-numeric columns

In [32]:

```
# import preprocessing from sklearn
from sklearn import preprocessing
from sklearn.preprocessing import OneHotEncoder
# limit to categorical data using df.select_dtypes()
data2 = df.select_dtypes(include=[object])#Extracting non-numeric c
data2.head(3)
```

Out[32]:

	name	fuel	seller_type	transmission	owner	brand
0	Maruti 800 AC	Petrol	Individual	Manual	First Owner	Maruti
1	Maruti Wagon R LXI Minor	Petrol	Individual	Manual	First Owner	Maruti
2	Hyundai Verna 1.6 SX	Diesel	Individual	Manual	First Owner	Hyundai

We will convert these columns to numbers (name--owner--brand) something like that {"First Owner" : 1, "Second Owner" : 2, "Third Owner" : 3, "Fourth & Above Owner" : 4,"Test Drive Car" : 0}

In [33]:

```
#name--owner--brand convert these columns to numbers
le = preprocessing.LabelEncoder()
df['name']= le.fit_transform(df['name'])
df['owner']= le.fit_transform(df['owner'])
df['brand']= le.fit_transform(df['brand'])
df.head()
```

Out[33]:

	name	year	selling_price	km_driven	fuel	seller_type	transmissi

In [32]:

0	775	2007	60000	70000	Petrol	Individual	Manu
1	1041	2007	135000	50000	Petrol	Individual	Manu
2	505	2012	600000	100000	Diesel	Individual	Manu
3	118	2017	250000	46000	Petrol	Individual	Manu
4	279	2014	450000	141000	Diesel	Individual	Manu

By use OneHotEncoder convert these columns (fuel, seller_type, transmission) to numbers and divide them separately and made new columns from it to be more clear,wait for the final result

In [34]:

```
#convert these columns (fuel, seller_type, transmission)
enc = OneHotEncoder(sparse = False)
X4 = df[["fuel", "seller_type", "transmission"]]
X5 = enc.fit_transform(X4)
new_col = enc.get_feature_names_out(["fuel","seller_type","transmis
df[new_col] = X5
df.head()
```

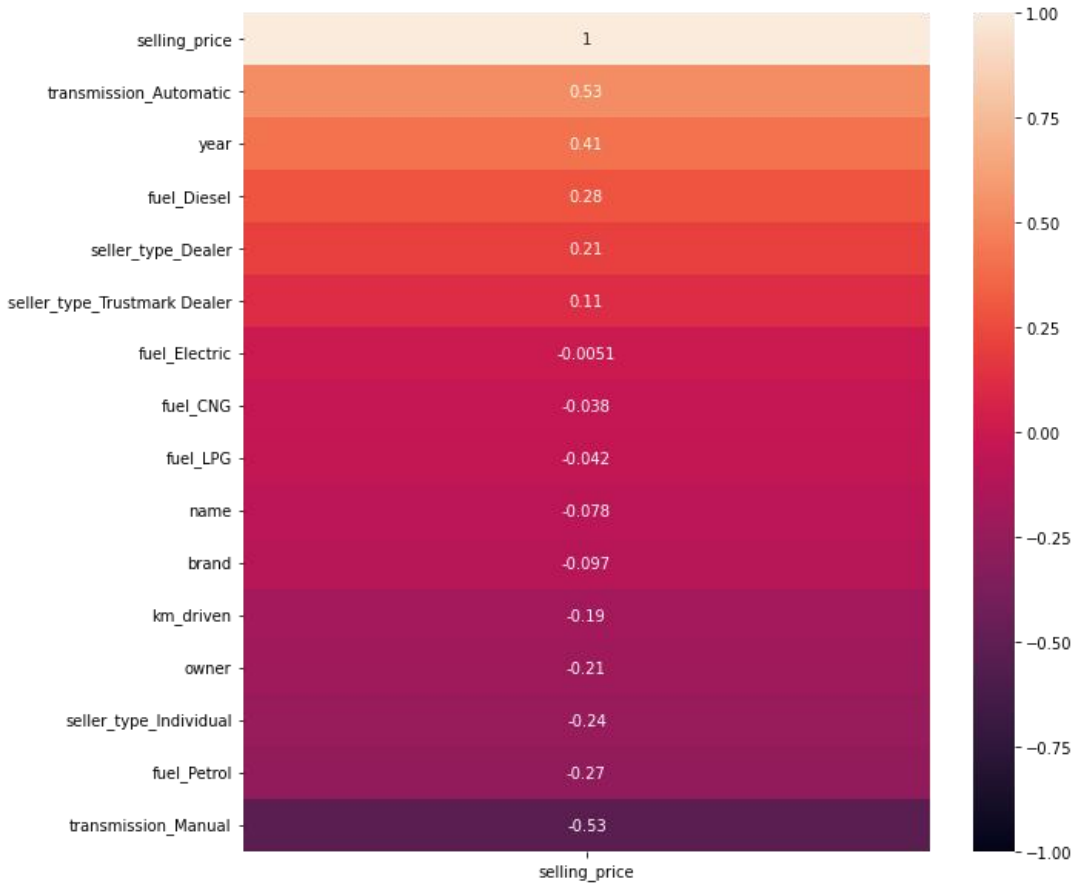
Out[34]:

	name	year	selling_price	km_driven	fuel	seller_type	transmissi
0	775	2007	60000	70000	Petrol	Individual	Manu
1	1041	2007	135000	50000	Petrol	Individual	Manu
2	505	2012	600000	100000	Diesel	Individual	Manu
3	118	2017	250000	46000	Petrol	Individual	Manu
4	279	2014	450000	141000	Diesel	Individual	Manu

Note the new columns and the way they are created now I can create heat map thatincludes all factors

In [35]:

```
#new heatmap
fig, ax = plt.subplots(figsize = (10,10))
sns.heatmap(df.corr()[["selling_price"]].sort_values("selling_price
plt.show()
```



In [36]:

```
sta1 = df['selling_price'].mean()
sta2 = df['selling_price'].count()
sta3 = df['selling_price'].std()
sta4 = df['selling_price'].min()
sta5 = df['selling_price'].quantile(q=0.25)
sta6 = df['selling_price'].quantile(q=0.50)
sta7 = df['selling_price'].quantile(q=0.75)
sta8 = df['selling_price'].max()
sta9 = df['selling_price'].median()
```

The heatmap may be inaccurate because it is a little different from what we have

previously extracted, but it gives us an indication of the impact of all the different factorson the price of the car. There are no surprises. The results were expected and identical to what we find before.

Through our journey in analyzing the data, we found that 1.Best selling

brand is Maruti then Hyundai
2. Manuals sell more than automatics
3. The newer the car the better it sells4.Individual sell more
5.Petrol and Diesel sell more
6.Most sales were made in 2017

There are factors that affect on car price

1.The year, the newer the year, the higher the price
2.The KM Driving when increase price of car decreased3.Brand and model of the car

Other factors important but these are the main factors.

CONCLUSION

After analysing the provided data on electric vehicles, several conclusions can be drawn.

Acceleration and Top Speed: Electric vehicles offer impressive acceleration capabilities, with some models achieving speeds comparable to high-performance gasoline-powered cars. The data shows a wide range of acceleration times, from as low as 2.1 seconds (Tesla Roadster) to 12.6 seconds (Renault Twingo ZE). Similarly, top speeds vary, with some models reaching speeds above 400 km/h (Tesla Roadster) and others limited to around 130-160 km/h (Renault Twingo ZE and Citroen e-C4).

Range and Efficiency: The range of electric vehicles has significantly improved, with several models surpassing 400 km on a single charge. The data shows ranges ranging from 130 km (Renault Twingo ZE) to 970 km (Tesla Roadster). Efficiency, measured in Wh/km, varies between models, with lower values indicating higher efficiency. The most efficient models in the dataset include the Lightyear One (104 Wh/km) and the Tesla Model 3 Standard Range Plus (153 Wh/km).

Charging Capabilities: Rapid charging is an important feature for electric vehicles, enabling quicker charging times during long

journeys. Most models in the dataset support rapid charging, with charging speeds ranging from 170 km/h to 940 km/h. Fast charging times are crucial for enhancing the practicality and convenience of electric vehicles.

Powertrain and Drive Type: The dataset includes various powertrain configurations, including rear-wheel drive, front-wheel drive, and all-wheel drive. All-wheel-drive configurations are prevalent among high-performance models, offering enhanced traction and acceleration. Additionally, the majority of models utilize an electric motor for propulsion.

Price Range: The price of electric vehicles varies significantly, depending on the brand, model, and features. The dataset includes models ranging from affordable options around 20,000 Euros (e.g., SEAT Mii Electric) to luxury models exceeding 200,000 Euros (e.g., Tesla Roadster). The price often correlates with factors such as range, performance, and brand reputation.

Body Styles and Segments: Electric vehicles are available in various body styles, including sedans, hatchbacks, SUVs, and even pickup trucks. This indicates that consumers have a wide range of options when choosing an electric vehicle based on their preferences and needs. The dataset covers different segments, ranging from compact hatchbacks (e.g., Renault Zoe) to high-end luxury sedans (e.g., Porsche Taycan Turbo S).

In conclusion, the analysis of the data highlights the significant progress made in the electric vehicle market, with improvements in acceleration, range, and charging capabilities. The availability of a diverse range of models across various price points and body styles provides consumers with more options than ever before. As technology continues to advance and charging infrastructure expands, electric vehicles are poised to play an increasingly important role in the future of sustainable transportation.

Code Links:

http://github.com/mlv1997/Feynn-labs-internship-2023
https://github.com/him-prad/Exploratory_analysis_of_electric_vehicles