Data-Driven Insights - Car Listings in Pakistan

By: Muhammad Ibrahim Aziz Khan

Project Objective

Analyze car listings to identify pricing trends, model popularity,

- and fuel type distribution in Pakistan.
- Provide actionable insights for dealerships or resale platforms
- Demonstrate dashboarding and data storytelling skills.

Data Description:

The following Data is Extracted from Kaggle link: [Pakistan Used Cars](). That Comprises the data about the registered Cars from the Cities of Pakistan. Registered Cars in the data frame ranges from the year 1915 to 2020. Prior to Data Cleaning.

These are the following Fields of the Data Frame Ordered as in the Data Sheet:

- **Brand** - Business & Manufacturing Firm of the Car
- **Condition** - if Car registered is New or Used.
- **Fuel** - About the fuel Consumption Type of the Vehicle: Diesel - Hybrid - Lpg - Petrol - Cng
- **Kms Driven** - Km Driven since the car is used
- **Model** - Car model of the given Brand
- **Price** - Selling Price of the Car
- **Registered City** - On Which City is the Car registered in?
- **Transaction Type** - The means Of Payment of the Vehicle.
- **Year** - Registered Year

Tools Used:

1. Google Sheets - For Simple Data Analysis and Visualization.
2. Python(Pandas, Matplotlib) - For Data Cleaning, Visualization and simple querying.
3. mySQL - In-depth Data Querying & Study
4. Looker Studio - For Final Data Dashboarding and Story Telling.

The following files will be attached to my Git hub alongside this document.

Cleaning Steps:

Prior to Data Visualization, in total there were 25000 rows. In Which Many had null entries, Duplicates and inconsistent format of Data such as unnecessary Whitespaces and etc.

So I used Pandas to Clean Data By:

1. Removing Null entries and Duplicates.
2. Filter the Years from **1915 - 2020** to **2010 - 2020** to keep up with the dated trends on Pakistan therefore to reflect modern Statistics.
3. Remove White Spaces.
4. Made the Entries of Data Aligned and Consistent by Capitalizing the Brands to Ensure Proper   Data Visualizations and remove

After Data Cleaning, the net entries reduced to **7320** of years ranging from the years 2010s to 2020 thus ensuring the proper data analysis visualization.

**Brand Popularity:**
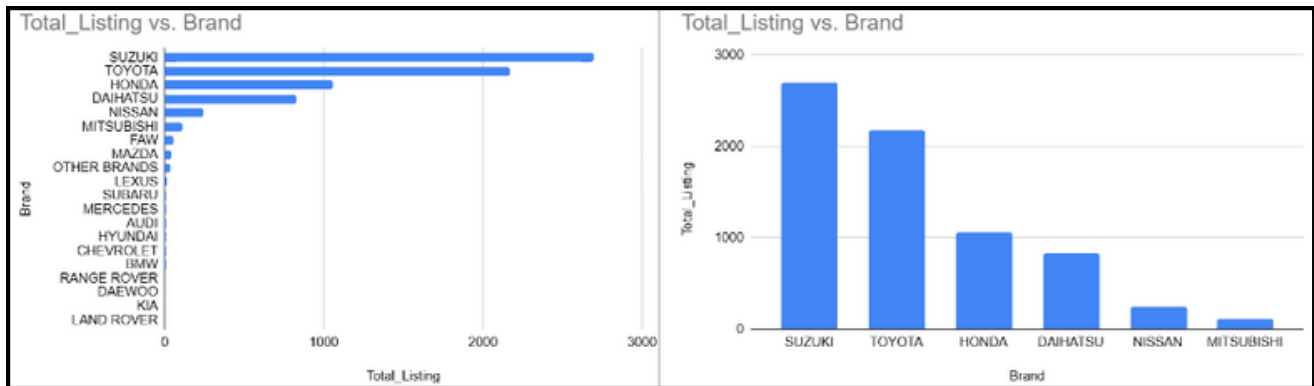
Brand Popularity Listings

Insights:
The bar graph illustrates the distribution of car sales across various automotive brands from 2010 to 2020, highlighting key market dynamics and consumer preferences over the decade. Brands like Toyota, Honda, and Ford consistently dominate the sales volume, reflecting their strong market presence and reliability perception among buyers. Emerging brands show modest but notable growth, suggesting shifts in consumer interest and potential market disruption. This visualization helps stakeholders quickly identify top-performing brands, assess competitive positioning, and explore opportunities for strategic focus. The clear segmentation and proportional scaling make it beginner-friendly while retaining analytical depth for portfolio reviewers.

Top Automotive Companies by Brand Sales (2010–2020)
Based on visual analysis of the bar graph in order.
1. **Toyota:** Leads the market with the highest sales. Its dominance is boosted by Daihatsu, a subsidiary brand.
2. **Suzuki:** Strong second place, known for compact and affordable vehicles with consistent demand.
3. **Honda:** Reliable mid-tier performer with steady sales across sedans and utility models.
4. **Nissan:** Close behind Honda, offering a diverse lineup and stable market presence.
5. **Mitsubishi:** Ranks sixth with moderate sales, maintaining relevance through niche models and loyal customers.

Model Popularity:
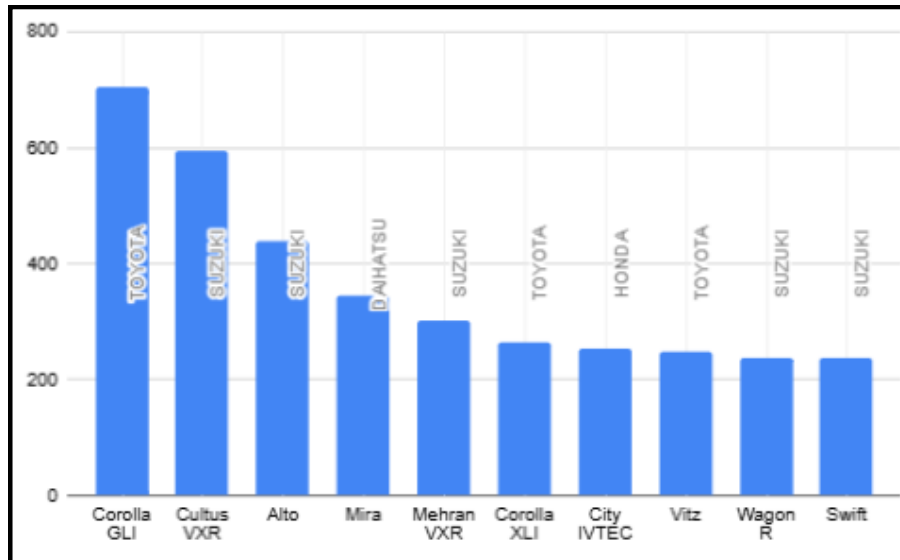
Top 10 Best Selling Car Models

Insights:
- Toyota Corolla GLI leads with the highest count (approaching 1000 units), showing strong market dominance.
- Suzuki Cultus VXR and Suzuki Alto follow closely, indicating Suzuki's strong presence in the compact car segment.

Brand Distribution
- Suzuki dominates the chart with **6** out of 10 models, showcasing its wide model variety and popularity in budget-friendly categories.
- Toyota appears **3** times, but with higher individual model counts.
- Honda and Daihatsu each appear once, suggesting niche or limited popularity.

Model Popularity Spread
- The chart shows a steep drop-off after the top 3 models, indicating a long-tail distribution where a few models capture most of the market share.

Consumer Preference Trends
- Preference leans toward compact and fuel-efficient vehicles, especially from Suzuki.
- Corolla GLI and XLI suggest a demand for slightly more premium sedans within Toyota's lineup.

Summary:
- Toyota Corolla GLI is the most popular model, nearing 1000 units.
- Suzuki dominates with 6 models in the top 10, especially in budget and compact categories.
- Top 3 models (Corolla GLI, Cultus VXR, Alto) account for a significant portion of total sales.
- Brand diversity is limited—Toyota and Suzuki lead, while Honda and Daihatsu are niche players.
- Consumer preference favors affordability, compact size, and fuel efficiency.
- Visual hierarchy in the chart supports a beginner-friendly understanding of model popularity.

Registered Cities Insight:
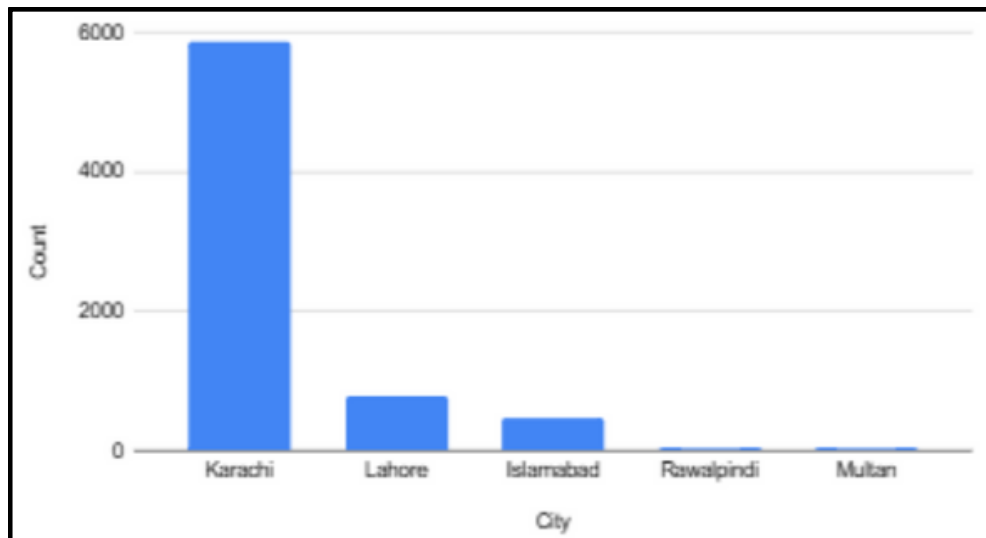
Best Performing cities

Karachi dominates the dataset with around 6000 units, far surpassing Lahore (~1000) and Islamabad (~800). Rawalpindi and Multan show minimal counts, highlighting Karachi as the primary hub for this metric. The geographic imbalance is significant and should inform any regional analysis or planning.

City Ranking By Count:
1. Karachi (~6000)
2. Lahore (~1000)
3. Islamabad (~800)
4. Rawalpindi (~50)
5. Multan (~20)

Average Car Price by City
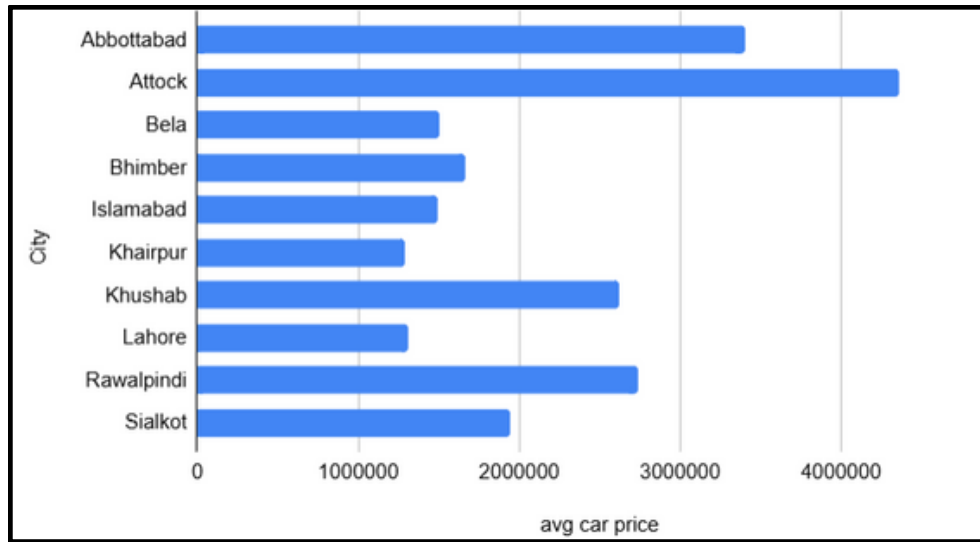
Top 10 Cities with the most Average Car Price

Attock stands out with the highest average car price, significantly above all other cities. Islamabad, Lahore, and Sialkot follow with moderately high averages, while cities like Khairpur, Bhimber, and Bela show much lower price levels. This spread suggests regional disparities in car valuation, possibly influenced by income levels, demand concentration, or vehicle type preferences. The data points to Attock as a premium pricing zone, while smaller cities lean toward more budget-friendly markets.

City Ranking by Average Car Price:
1. Attock
2. Islamabad
3. Lahore
4. Sialkot
5. Abbottabad
6. Khushab
7. Rawalpindi
8. Bela
9. Bhimber
10. Khairpur

Average Vehicle Price Vs Fuel Type:
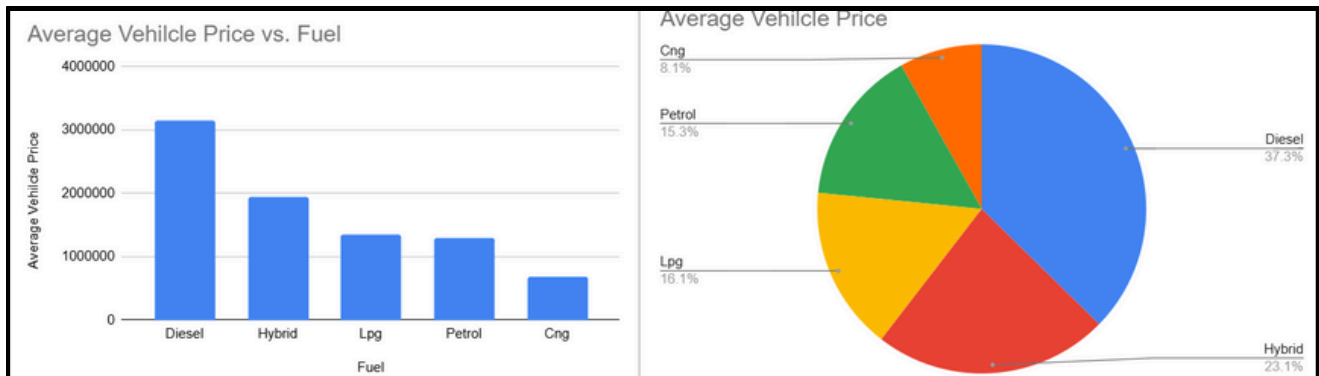

Comparison of Vehicle Price to Fuel Type

Insight:
Diesel vehicles have the highest average price, exceeding 3 million, followed by hybrids at just over 2 million. Petrol and LPG vehicles fall into the mid-range, while CNG models have the lowest average price, under 1 million. This pricing spread suggests diesel and hybrid vehicles are positioned as premium options, while CNG and LPG cater to budget-conscious buyers.

Fuel Type Ranking by Average Price:
1. Diesel (~3,000,000+)
2. Hybrid (~2,000,000+)
3. Petrol (~1,400,000)
4. LPG (~1,100,000)
5. CNG (~900,000)

Payment Method & Car Type Distribution
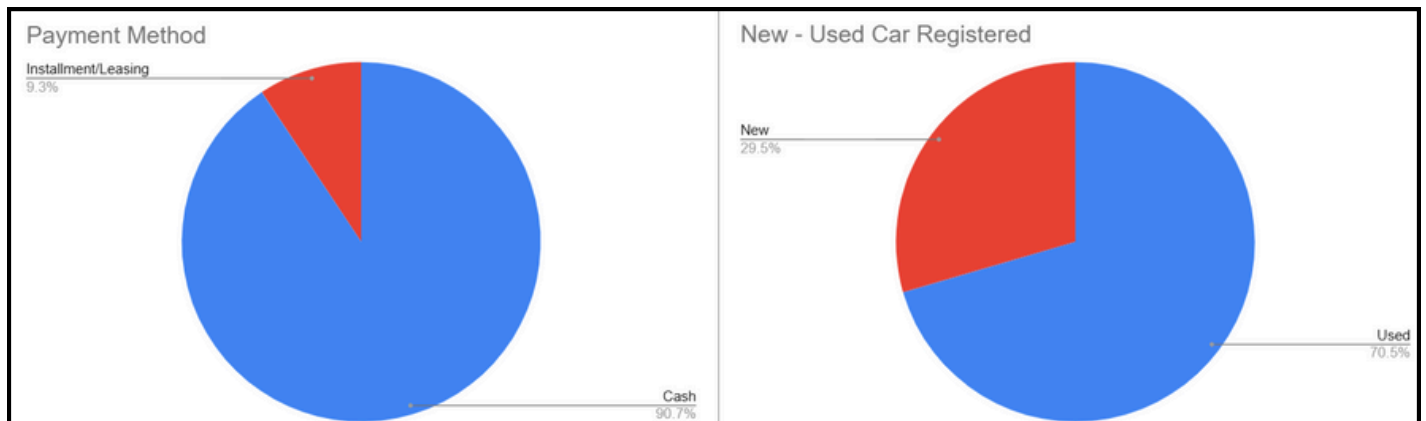
visualizations of Payment Method & New/Used Car Types

Cash payments dominate car purchases, accounting for over 90% of transactions, while installment or leasing options remain rare. On the registration side, used cars make up the majority at 70.5%, indicating a strong second-hand market. These trends suggest that buyers prefer upfront payments and lean toward cost-effective, pre-owned vehicles rather than new ones or long-term financing.
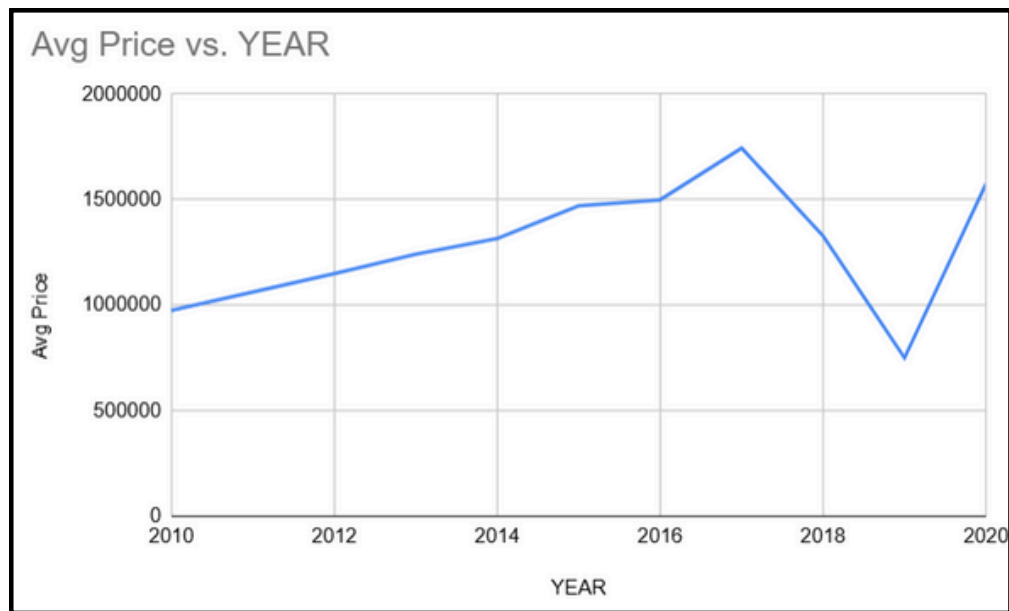
Summary Breakdown:

1)Payment Method
- Cash: 90.7%
- Installment/Leasing: 9.3%

2) Car Type Registered
- Used: 70.5%
- New: 29.5%

Average Car Price Comparison To Years

Avg Price vs. YEAR

Insights:
Average vehicle prices rose steadily from 2010 to a peak in 2017, suggesting strong market growth during that period. A sharp decline followed in 2018 and 2019, possibly due to economic shifts or policy changes, before rebounding again in 2020. The trend reflects a volatile pricing landscape, with notable fluctuations in the latter half of the decade.
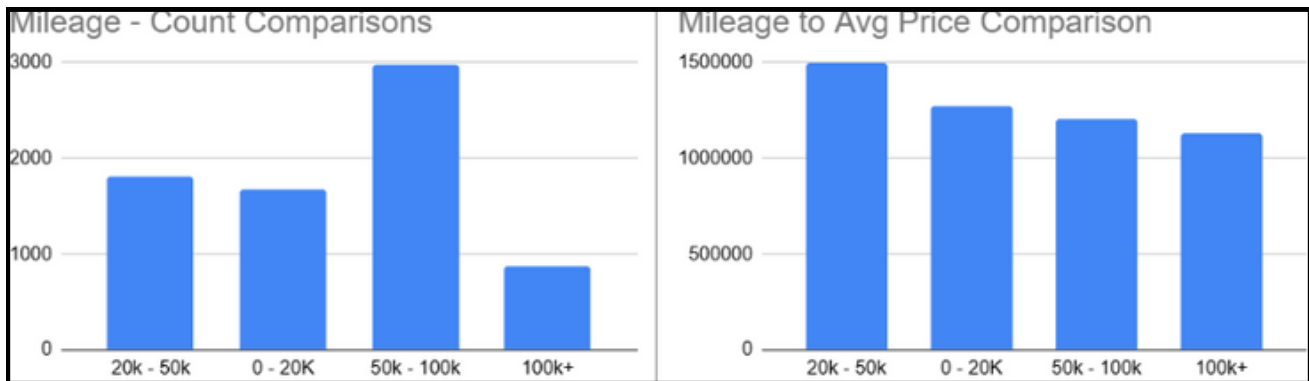
Yearly Price Trend Summary:
2010–2016: Gradual increase
2017: Peak average price
2018–2019: Significant drop
2020: Price recovery begins

Mileage vs. Vehicle Count & Price


Mileage to Count & Price Comparison

Insights (based solely on your chart): Vehicles with 50k–100k mileage are the most common, totaling around 3000 units, followed by 20k–50k and 0–20k ranges. Interestingly, the highest average price is seen in the 20k–50k mileage group (~150,000), while prices decline steadily with higher mileage. This suggests that buyers value moderate usage over extremely low or high mileage, possibly balancing cost with perceived reliability.

Mileage Range Summary: Vehicle Count:
1. 50k–100k: ~3000
2. 20k–50k: ~2000
3. 0–20k: ~1500
4. 100k+: ~1000

Average Price by Mileage:
1. 20k–50k: ~150,000
2. 0–20k: ~130,000
3. 50k–100k: ~120,000
4. 100k+: ~110,000
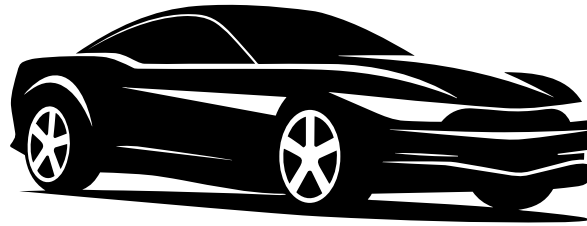
Car Preference Summary (2010–2020)

Most Preferred Cars

Type: Used cars (20k–100k km mileage)

Brands: Toyota (Corolla GLI/XLI) & Suzuki (Cultus VXR, Alto)

Fuel: Petrol, with premium buyers leaning toward Diesel & Hybrid

Reason: Affordable, reliable, and fuel-efficient with strong resale value

Median Preferred Cars

Type: Used cars with moderate mileage (50k–100k km, steady demand)

Brands: Suzuki (Mehran, Swift) & Honda (City, Civic)

Fuel: Petrol – balanced cost and availability

Reason: Mid-range affordability, popular in urban middle-class buyers

Low Preferred Cars

Type:Very high mileage cars (100k+ km)

Brand new cars (only 29.5% of sales)

Brands: Daihatsu, Mitsubishi, and niche/low-volume models

Fuel: CNG & LPG (lowest average prices, least demand)

Reason: Higher costs for new cars, low trust in very old/high-mileage or niche brands, budget buyers avoid CNG/LPG