

# **Maximizing Profits and Market Differentiation: Unveiling Insights from Used Car Sales Data for Dealership Success**

Prepared By

**Joseph Macharia**

Data Analyst



# Table of Contents

- 
- 1 Executive Summary Pg. 3
  - 2 Project Description Pg. 4
  - 3 Sidebar and Data Sources Pg. 5
  - 4 Data Cleaning Pg. 6
  - 5 Insights Exploration Pg. 7
  - 6 Key Questions and Answers Pg. 23
  - 7 Summary of Key Findings Pg. 34
  - 8 Recommendations & Conclusion Pg. 35

# Executive Summary

This analysis of used car sales data across the United States aimed to provide actionable insights to optimize inventory selection, pricing strategies, and market positioning for a car dealership.

Key findings and recommendations are as follows:

## Price Dynamics

- Average used car price nationally is \$17,002, with significant state variations (e.g. \$10,633 in DC vs \$27,368 in Alaska)
- Rural areas command higher prices than urban centers
- Top price factors are location, condition, body type, size, and year of manufacture

## Market Trends

- Ford, Chevrolet, RAM, Toyota, and GMC dominate inventory with models like F-150, Silverado 1500 driving sales
- Emerging models gaining traction include Toyota Tundra, Chevy Trailblazer, Rivian SUV
- Trucks most popular in states like Montana, Arkansas; SUVs favored in Rhode Island, New Hampshire

## Optimal Inventory Mix

- Prioritize inventory segmentation by price, condition, top brands/models to meet varied buyer needs
- Offer affordable options like trucks <\$10K and variety of sedans including luxury/sport
- Stock per location preferences (e.g. more trucks in Alabama, SUVs in Massachusetts)
- Diversify body types focusing on full-size trucks/SUVs and selection of sedans

## Website & Customer Experience

- Invest in SEO-optimized website with location filters to capitalize on online buyer journey
- Provide transparent pricing, rich media and details to elevate browsing experience
- Continuously monitor trends, sales and feedback to align inventory and marketing

By strategically managing inventory mix, pricing, online presence and customer experience based on these data-driven insights, the dealership can maximize profitability and solidify market differentiation.



# Project Description

To better understand the underlying trends and patterns in the sale of used cars across the United States. This analysis aims to provide valuable insights, enabling data-driven decision making regarding inventory selection, pricing strategies, and market positioning.

This will help us to:

- Identify market approaches that could help the dealership maximize profits.
- Ascertain the most desirable makes, models, and characteristics of used cars for different regions to guide the dealership in optimizing its inventory selection to effectively meet customer needs.



# Sidebar and Data Sources

I conducted an in-depth analysis of Craigslist's used car for sale dataset for different states across the United States for the year 2021. The dataset includes various attributes such as manufacturer, model, year of manufacture, odometer reading, condition, location, paint color, car type and size, price, and more.

The objective of this analysis was to derive actionable insights to inform strategic decision-making and optimization efforts within the used cars for sale market for a U.S. car dealership.

Please note that certain details within this data analysis report, like the name of the car dealership, have been intentionally altered or omitted to maintain confidentiality and safeguard proprietary information of the company. These modifications have been made in accordance with confidentiality protocols to ensure the protection of sensitive data and maintain the privacy of our business operations.

Unless otherwise stated, the data sources used for this analysis are:

- Used Cars Dataset by Austin Reese: <https://www.kaggle.com/datasets/austinreese/craigslist-carstrucks-data>
- Google Trends: <https://trends.google.com/trends/explore>



# Data Cleaning

After sourcing the data from Kaggle, I used Python for data cleaning, including:

## Removing duplicates:

```
Notebook ▾ Python 3 (ipykernel) C  
[21]: cars = cars.drop_duplicates()  
cars  
  
[21]:  
      id          url   region      region_url  price  year  manu  
0  7222695916  https://prescott.craigslist.org/cto/d/prescott...  prescott  https://prescott.craigslist.org  6000  NaN  
1  7218891961  https://fayar.craigslist.org/ctd/d/bentonville... fayetteville  https://fayar.craigslist.org  11900  NaN  
2  7221797935  https://keys.craigslist.org/cto/d/summerland-k... florida keys  https://keys.craigslist.org  21000  NaN
```

## Removing null values

```
[8]: cars.isnull().sum()  
  
Unnamed: 0          0  
id                0  
region             0  
price              0  
year               1178  
manufacturer       17619  
model              5250  
condition           174077  
cylinders           177651  
fuel                2986  
odometer            4373  
title_status        8215  
transmission         2529  
VIN                 161015  
drive               130540  
size                306334  
type                92831  
paint_color          130176  
description          43  
state                  0  
lat                  6522  
long                  6522  
posting_date          41  
dtype: int64  
  
[8]: cars.dropna(subset=["year", "manufacturer", "model", "condition"], inplace=True)  
[8]: cars.dropna(subset=["year", "manufacturer", "model", "condition"], inplace=True)  
[10]: cars.dropna(subset=["cylinders", "fuel", "odometer", "title_status"], inplace=True)  
[12]: cars.dropna(subset=["transmission", "VIN", "drive", "size", "type"], inplace=True)  
[14]: cars.dropna(subset=["paint_color", "lat", "long"], inplace=True)  
[15]: cars.isnull().sum()  
  
Unnamed: 0          0  
id                0  
region             0  
price              0  
year               0  
manufacturer       0  
model              0  
condition           0  
cylinders           0  
fuel                0  
odometer            0  
title_status         0  
transmission         0  
VIN                 0  
drive               0  
size                0  
type                0  
paint_color          0  
description          0  
state                  0  
lat                  0  
long                  0  
posting_date          0  
dtype: int64
```

## Removing unnecessary columns like the URL columns, image columns, etc.

```
Notebook ▾ Python 3 (ipykernel) C  
[23]: cars = cars.drop(columns = 'url')  
cars  
  
[23]:  
      id   region      region_url  price  year  manufacturer    model  condition  cylinders  
0  7222695916  prescott  https://prescott.craigslist.org  6000  NaN     NaN     NaN     NaN     NaN  
1  7218891961  fayetteville  https://fayar.craigslist.org  11900  NaN     NaN     NaN     NaN     NaN  
2  7221797935  florida keys  https://keys.craigslist.org  21000  NaN     NaN     NaN     NaN     NaN  
3  7222270760  worcester / central MA  https://worchester.craigslist.org  1500  NaN     NaN     NaN     NaN     NaN  
4  7210384030  greensboro  https://greensboro.craigslist.org  4900  NaN     NaN     NaN     NaN     NaN  
...  ...  ...  ...  ...  ...  ...  ...  ...  ...
```





# Insights Exploration

# Used cars are cheapest in DC and Pennsylvania and most expensive in Alaska and Tennessee

The average price of used cars across the United States as recorded in the dataset is **\$17,002**

The price of used cars does vary from one state to another:

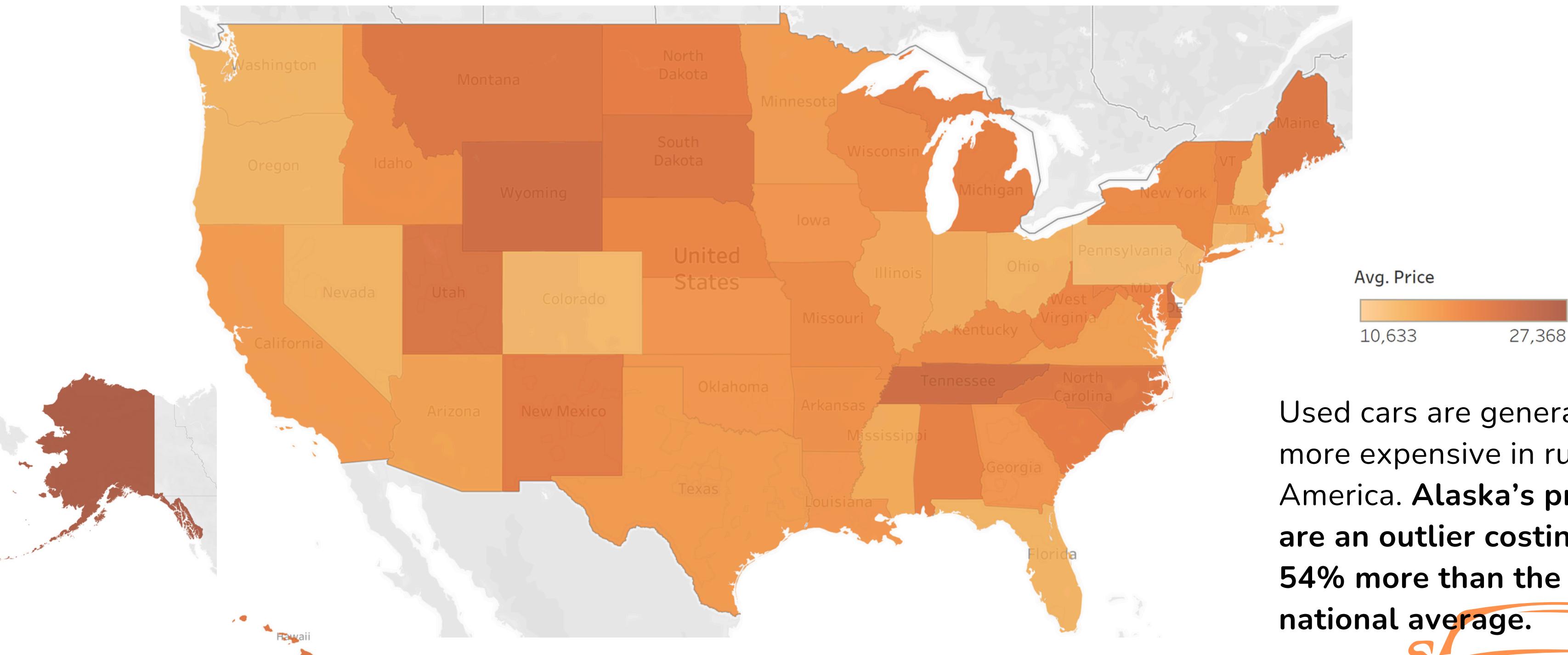
- Used cars are cheapest in the District of Columbia at an average price of \$10,633. Pennsylvania has the second lowest used car prices at \$13,148
- Used cars are most expensive in Alaska at an average price of \$27,368 followed by Tennessee at an average price of \$23,185.
- Alaska's prices are an outlier costing 54% more than the national average.

Most Expensive		Cheapest	
State	Average Price	State	Average Price
ak	\$27,368	dc	\$10,633
tn	\$23,185	pa	\$13,148
wy	\$23,178	nj	\$13,444
de	\$22,597	co	\$13,550
ut	\$21,755	ct	\$13,559



# Used cars are generally more expensive in rural America

Map showing the average cost of buying a used car in different states across the U.S.

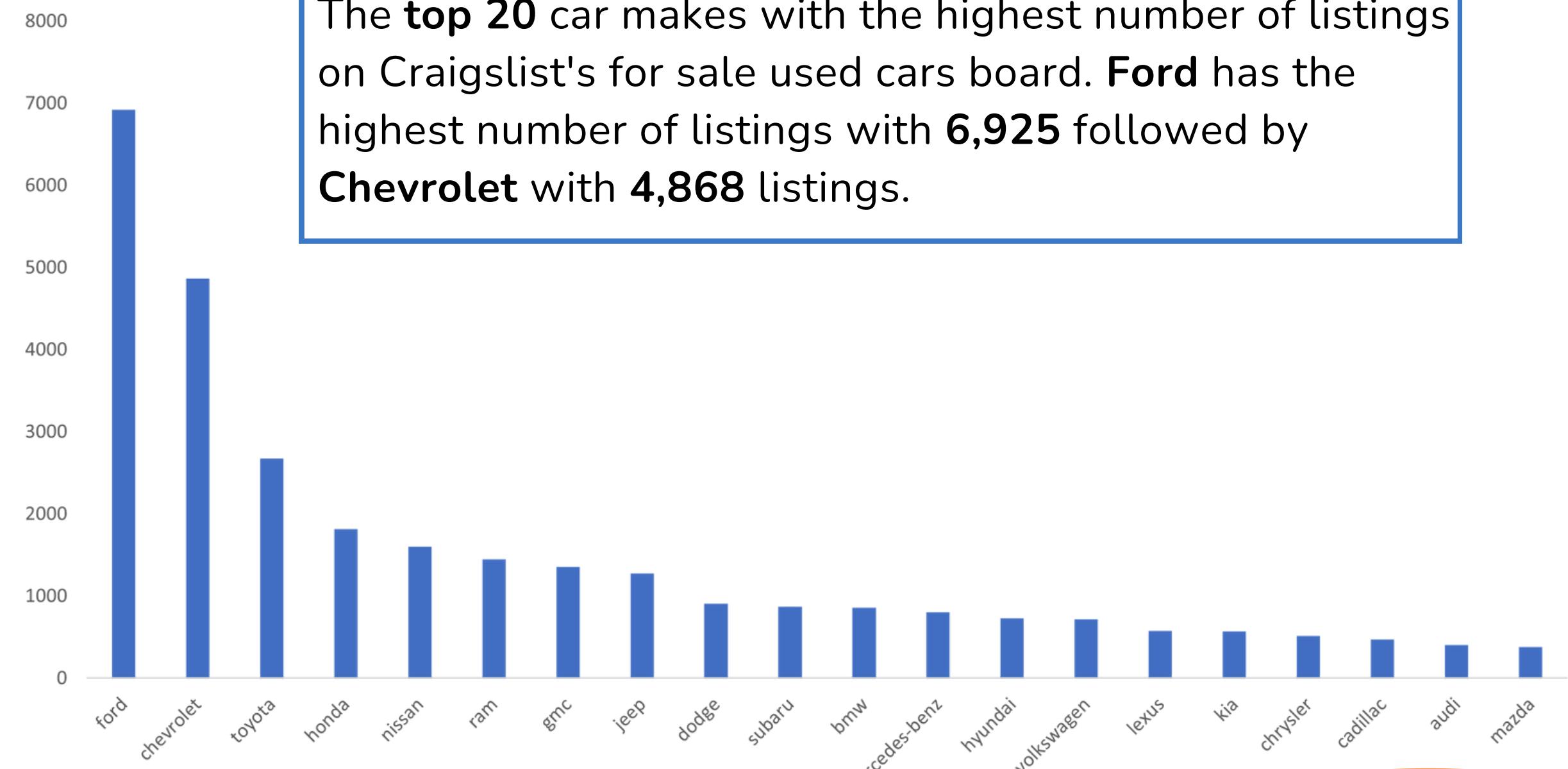


Used cars are generally more expensive in rural America. Alaska's prices are an outlier costing 54% more than the national average.



# Ford and Chevrolet dominate the inventory with close to 12,000 listings

Manufacturer	Number of Listings
ford	6925
chevrolet	4868
toyota	2676
honda	1817
nissan	1600
ram	1449
gmc	1354
jeep	1278
dodge	909
subaru	872
bmw	860
mercedes-benz	803
hyundai	727
volkswagen	717
lexus	579
kia	572
chrysler	513
cadillac	469
audi	407
mazda	381



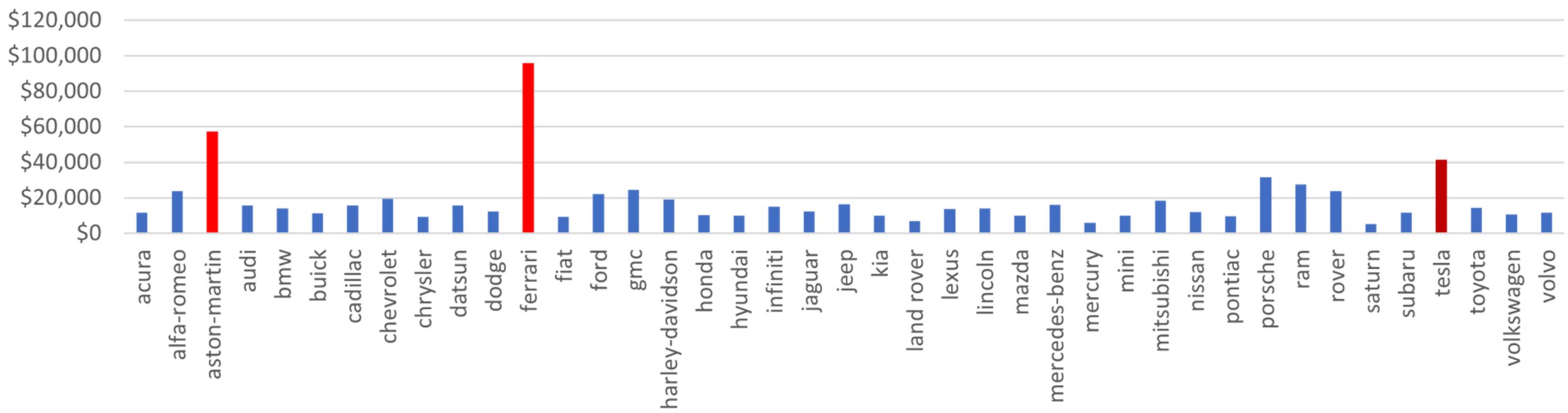
The **top 20** car makes with the highest number of listings on Craigslist's for sale used cars board. **Ford** has the highest number of listings with **6,925** followed by **Chevrolet** with **4,868** listings.



# The price of used cars is within the same range for different manufacturers except for luxury car brands like Ferrari and Aston Martin, and Tesla.

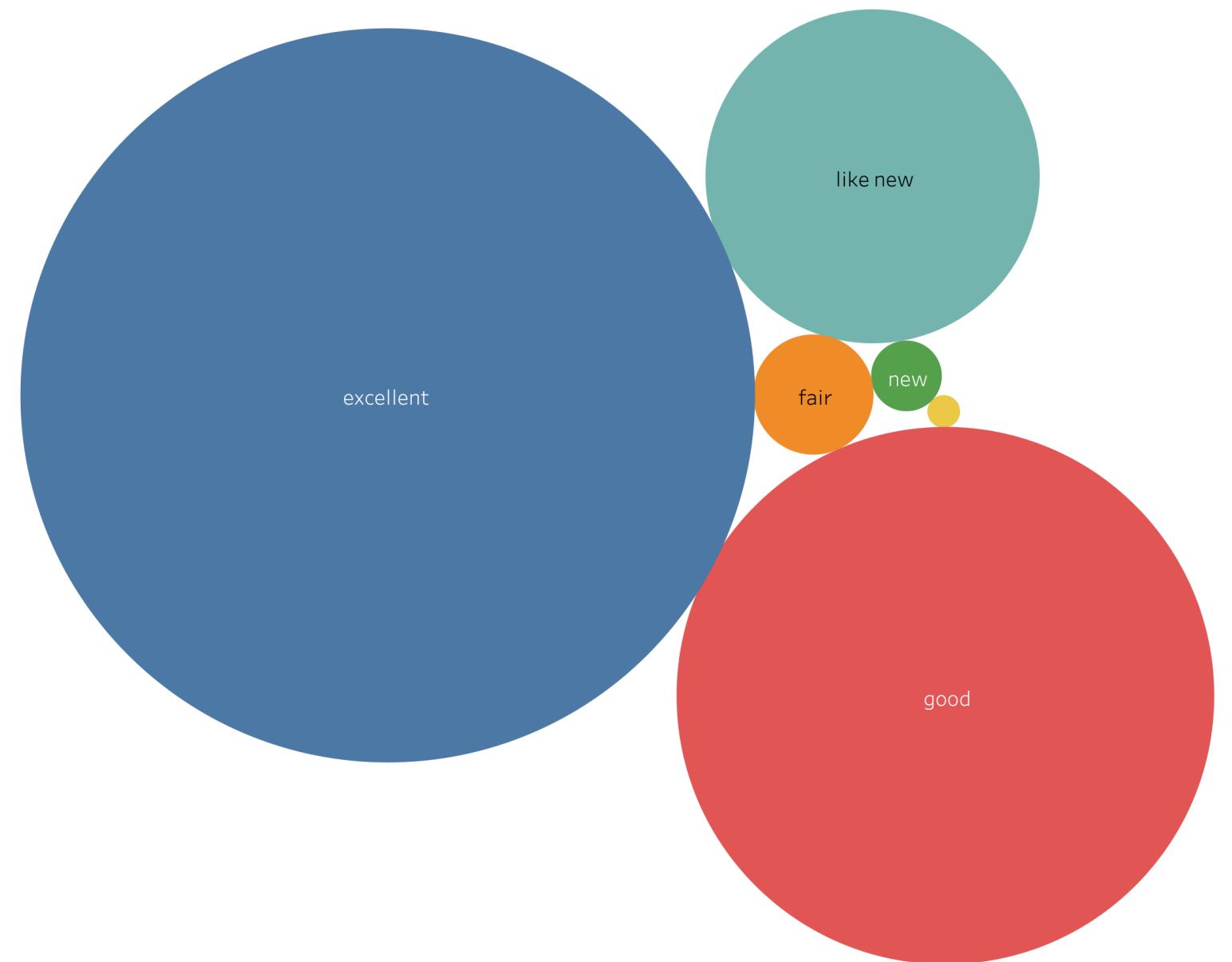
Make	Price
ferrari	\$95,829
aston-martin	\$57,280
tesla	\$41,424
porsche	\$31,622
ram	\$27,398
gmc	\$24,446
alfa-romeo	\$23,893
rover	\$23,731
ford	\$21,918
chevrolet	\$19,282
harley-davidson	\$19,023
mitsubishi	\$18,434
jeep	\$16,251
mercedes-benz	\$16,098
audi	\$15,758
cadillac	\$15,543
datsun	\$15,500
infiniti	\$15,059
toyota	\$14,362
bmw	\$14,036

As expected, luxury car manufacturers Ferrari and Aston Martin are outliers in their cost followed by Tesla. For the rest of the manufacturers, there's no significant price difference between them.



# Used cars listed as in excellent, good condition or like new accounting for 98% of all the cars on sale.

Car Condition
excellent
fair
good
like new
new
salvage



Car Condition	Number of Cars
excellent	18211
good	9757
like new	3769
fair	488
new	169
salvage	36

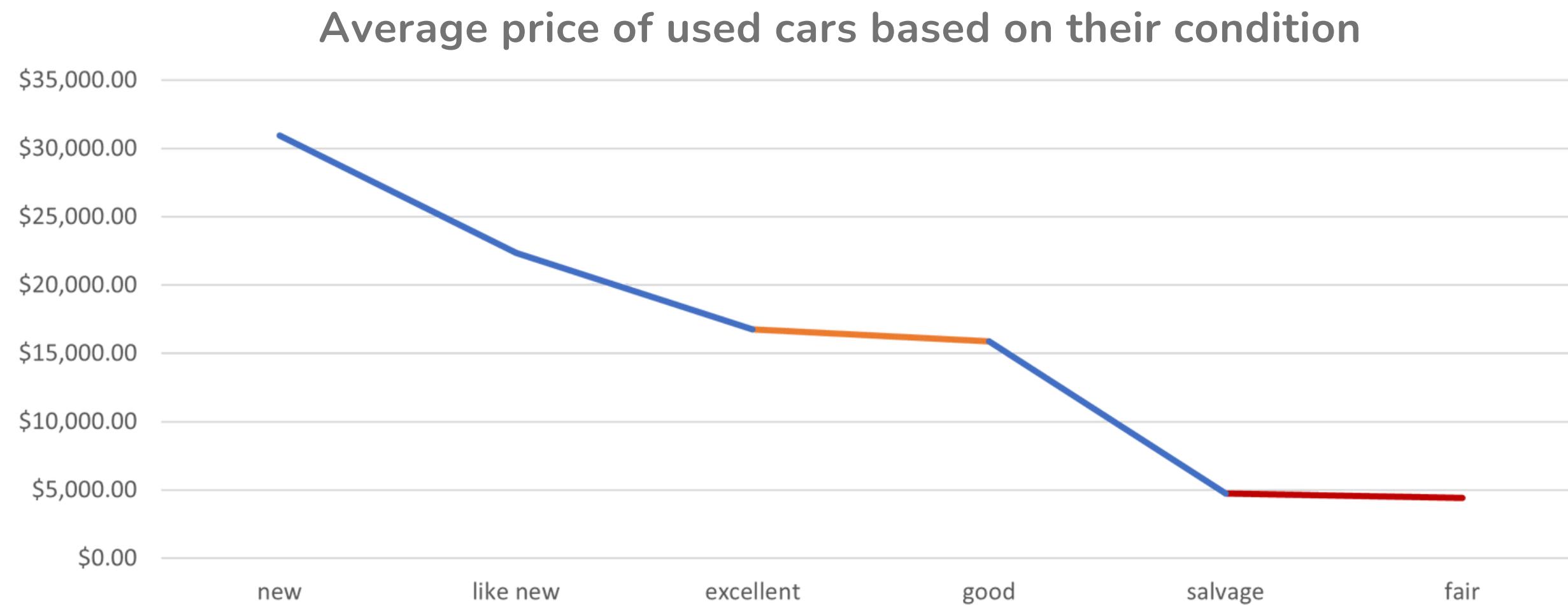
Most of the used cars on sale are either in excellent condition, in good condition or like new, accounting for **97.86%** of the total cars on sale.



## Salvage cars are slightly more expensive than those labelled as being in a fair condition.

There's no significant difference between the cost of a car labelled as in excellent condition compared to that said to be in good condition. Used cars said to be in the poorest condition (salvage) are slightly more expensive than those labelled as being in a fair condition.

Car Condition	Number of Cars
excellent	18211
good	9757
like new	3769
fair	488
new	169
salvage	36

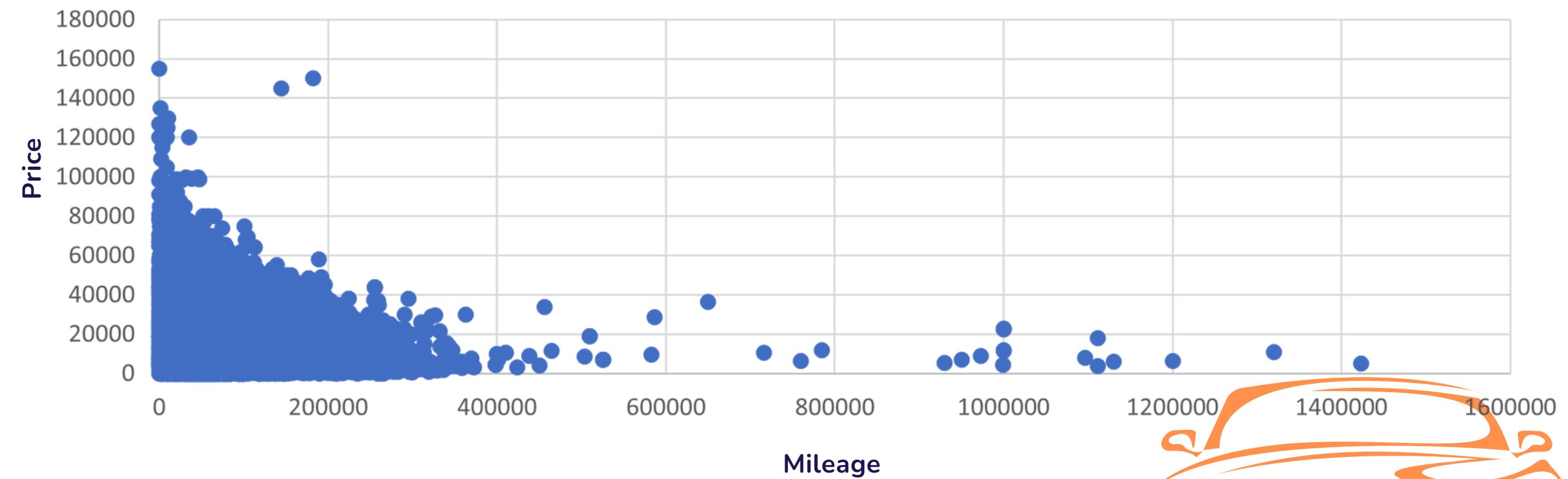


# Mileage does not seem to be a good predictor of the price of a used car

The average mileage of the cars in the dataset is **108,933.75 miles**.

manufacturer	odometer	price
mercedes-benz	250	\$155,000
ford	182415	\$150,000
volvo	144832	\$144,832
dodge	1588	\$135,000
porsche	10374	\$129,928
cadillac	50	\$126,995
porsche	10100	\$125,000
rover	8916	\$124,900
rover	8916	\$124,900
ferrari	35498	\$120,000
ferrari	9350	\$120,000
ford	1	\$119,999
chevrolet	601	\$119,900
rover	5000	\$118,300
bmw	3800	\$115,000
chevrolet	2107	\$109,000
bmw	9000	\$104,900
nissan	1234	\$99,990
ford	46113	\$99,888
ford	46113	\$99,888

The mileage of a used car does not seem to be a good indicator of the price of the car. There are many cars with low mileage that are also cheap.



# The majority of used cars on sale across the country use gas and have automatic transmission

The majority of cars have automatic transmission at just shy of 93%.

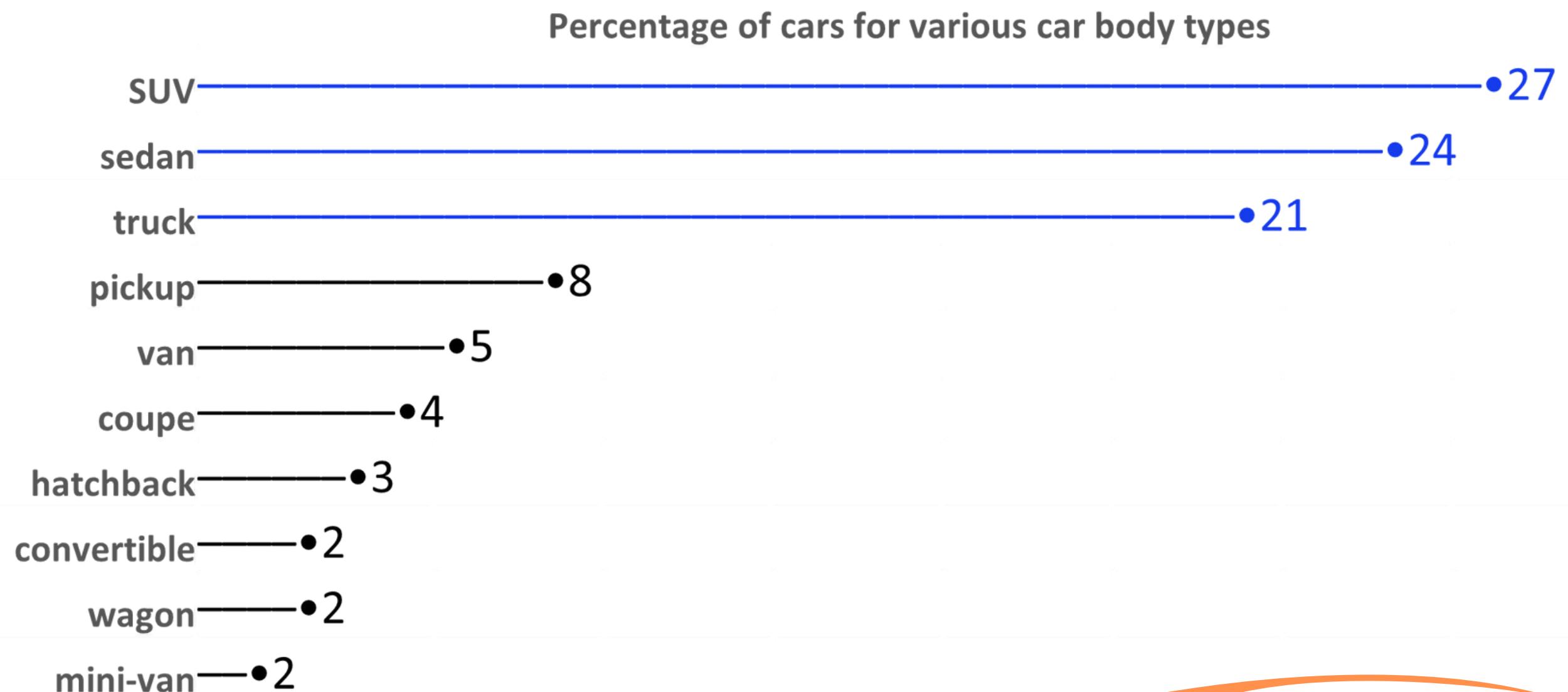
Transmission	Percentage
automatic	92.79%
manual	5.07%
other	2.14%

The majority of cars (89%) use gas.

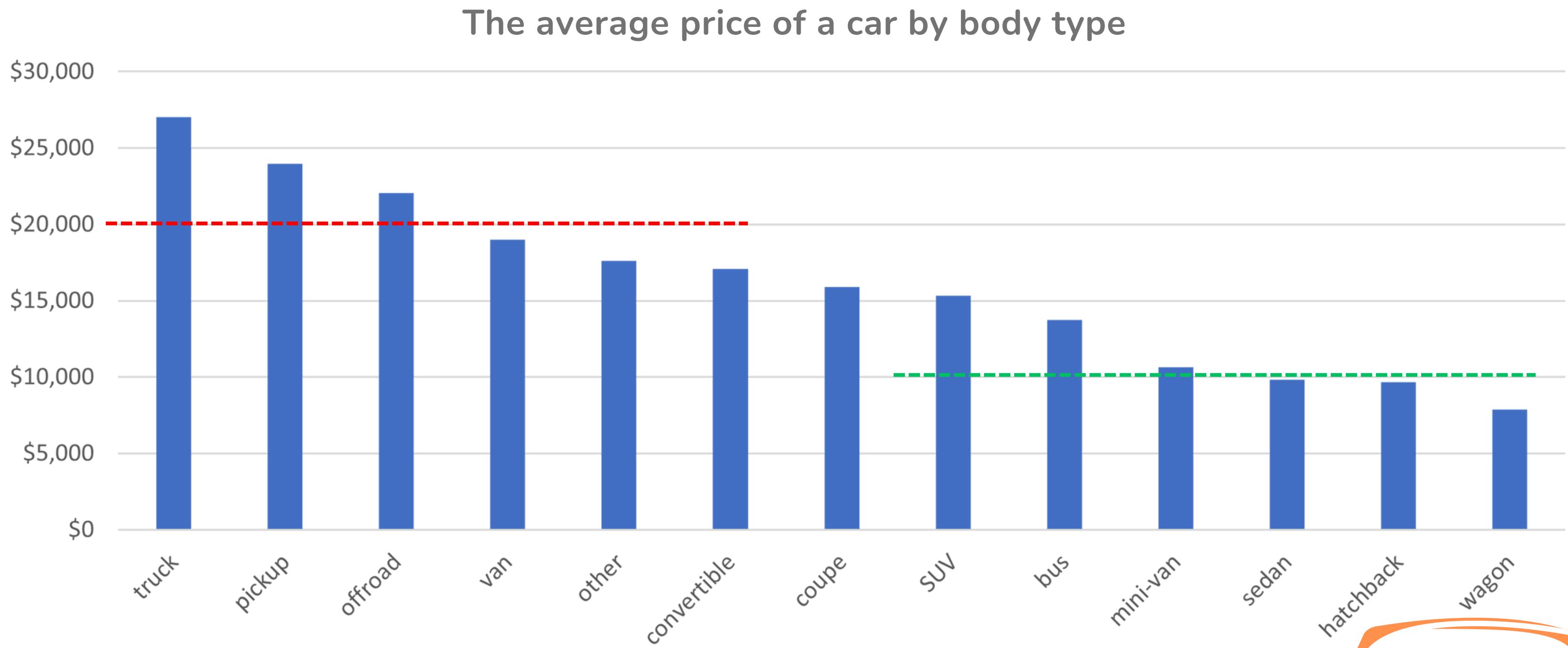
Fuel Type	Percentage
gas	89.77%
diesel	8.70%
hybrid	1.20%
electric	0.23%
other	0.10%

# SUVs, sedans, and trucks account for 72% of used car body types

Body Type	Number of Cars	Percentage
SUV	8608	26.54%
sedan	7803	24.06%
truck	6949	21.43%
pickup	2563	7.90%
van	1632	5.03%
coupe	1409	4.34%
hatchback	1091	3.36%
convertible	796	2.45%
wagon	772	2.38%
mini-van	590	1.82%
other	93	0.29%
offroad	81	0.25%
bus	43	0.13%



Truck, pickup, and off-road cars are the most expensive at **over \$20,000** while sedan, hatchback, and wagon cars are the cheapest at **under \$10,000**.



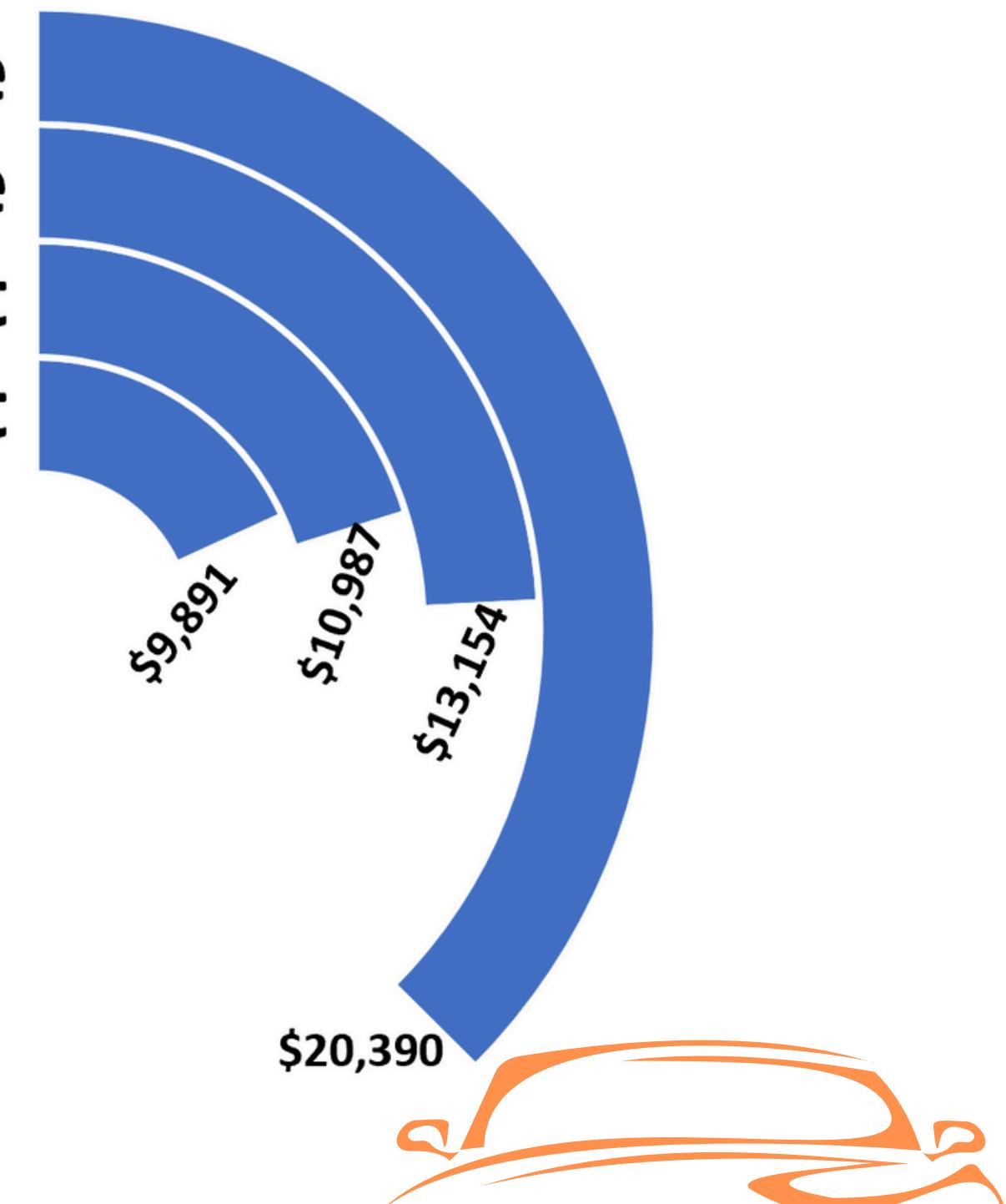
## The price of a full-size car is almost 1.8 times that of a mid-size car.

Full-size cars are the most expensive on average while compact cars are the cheapest.

Body Size	Average Price
full-size	\$20,390
mid-size	\$13,154
sub-compact	\$10,987
compact	\$9,891

The average price of a car based on its body size.

**full-size  
mid-size  
sub-compact  
compact**



# Truck is the most preferred body type across the U.S. followed by SUV. Sedan takes third place.

Rhode Island is the only state where mid-size cars are preferred. In all other states, full size is the most popular.

Body Type	Number of States
truck	18
SUV	17
sedan	16

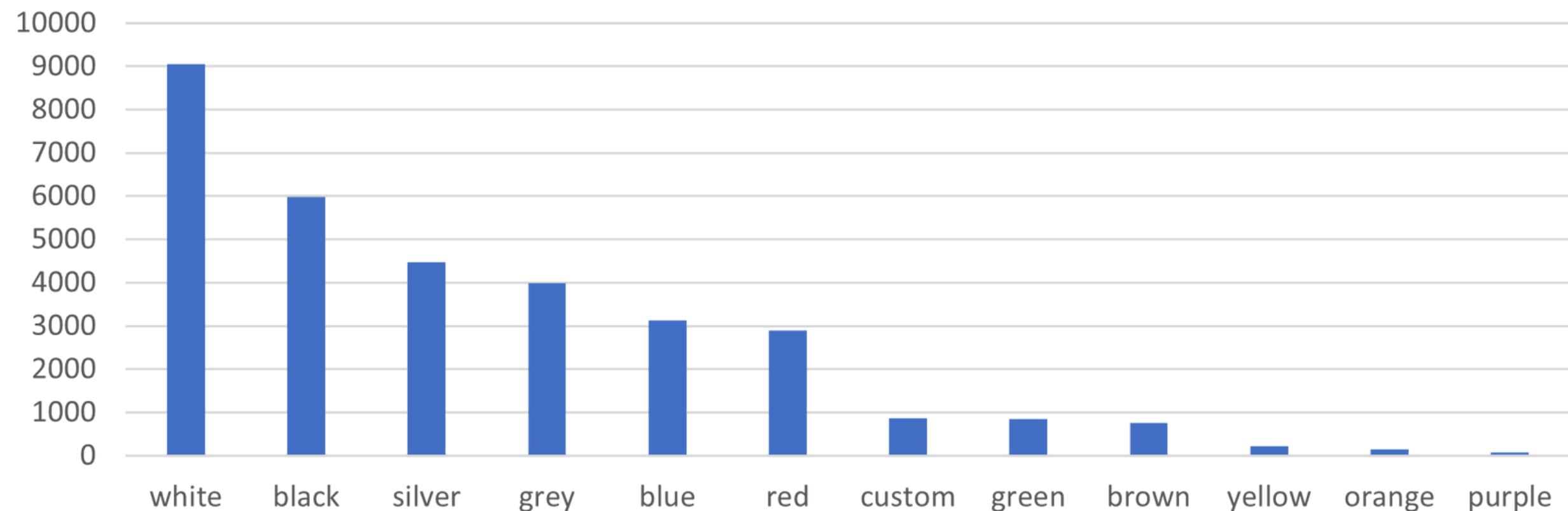


SUV	Truck	Sedan
Alaska	Alabama	Arizona
Arkansas	Delaware	California
Colorado	Georgia	Connecticut
Idaho	Iowa	District of Columbia
Illinois	Louisiana	Florida
Kansas	Maine	Hawaii
Kentucky	Mississippi	Indiana
Michigan	Montana	Massachusetts
Minnesota	North Carolina	Maryland
Missouri	New Mexico	New Jersey
North Dakota	Oklahoma	Nevada
Nebraska	South Carolina	Ohio
New Hampshire	South Dakota	Oregon
New York	Tennessee	Pennsylvania
Rhode Island	Utah	Texas
Vermont	Wisconsin	Virginia
Washington	West Virginia	
		Wyoming

# Black and white cars make up almost half of all used cars sold across the country

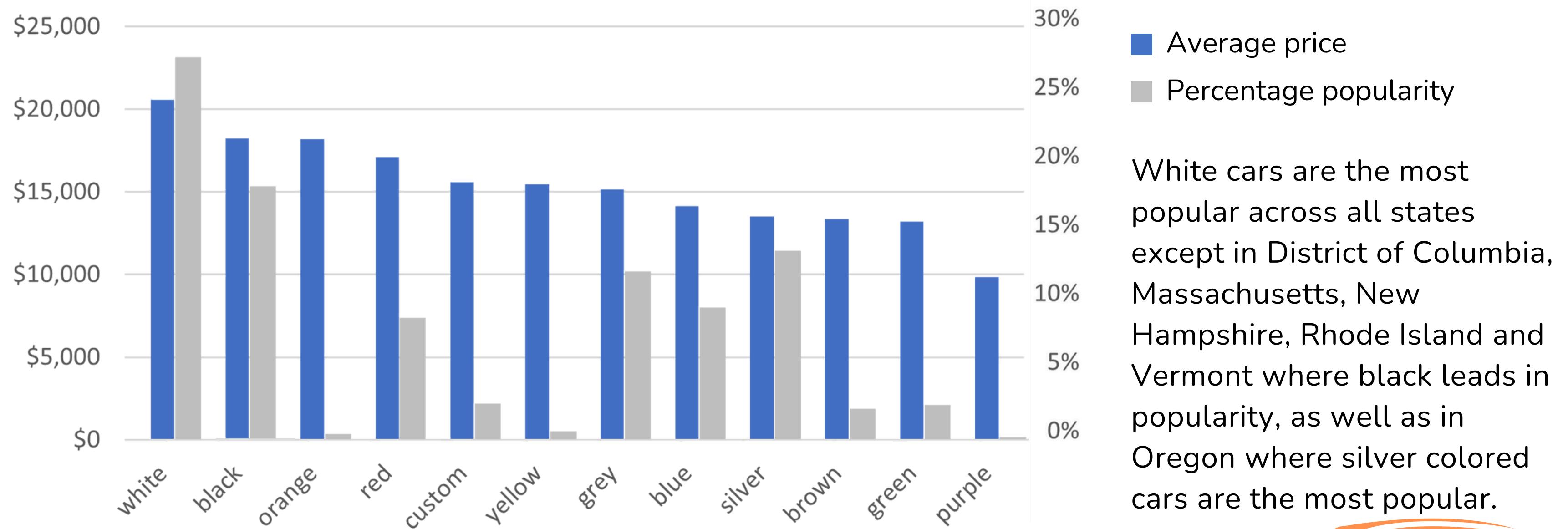
Paint Color	Number of Cars	Percentage
white	9059	27.93%
black	5985	18.46%
silver	4477	13.81%
grey	3986	12.29%
blue	3131	9.65%
red	2897	8.93%
custom	867	2.67%
green	842	2.60%
brown	748	2.31%
yellow	214	0.66%
orange	152	0.47%
purple	72	0.22%

Popularity of used cars by their color. **White** is the most popular at **27.93%**. Black and white cars make up almost half (46.39%) of all used cars sold across the country.



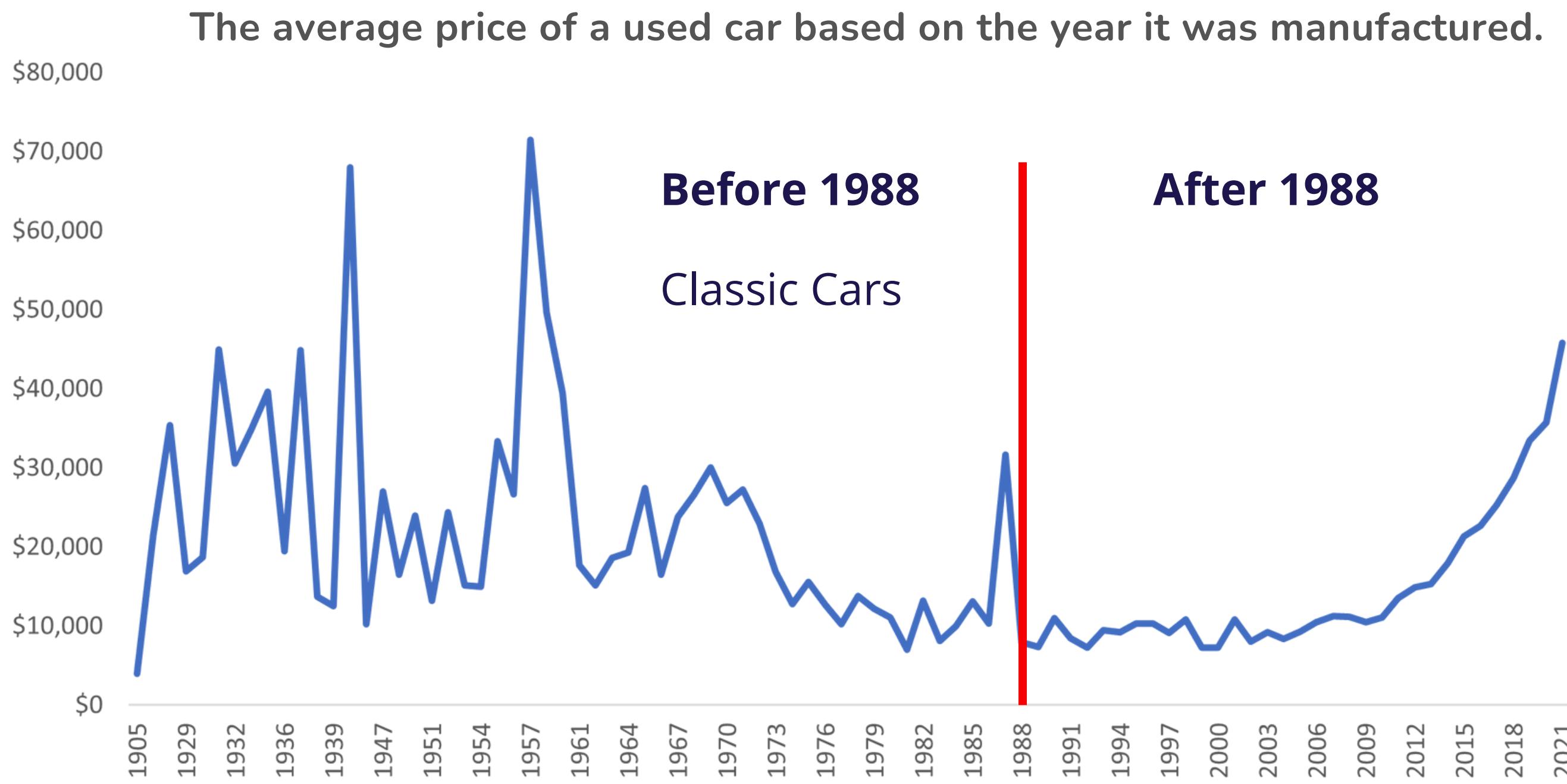
# Impact of paint color on the price of a used car in different states across the U.S.

Paint color seems to impact the price of some cars while having little effect on the price of others. White used cars are both the most expensive and the most popular while purple are the least preferred and the cheapest.



**For cars manufactured after 1987, the older the car is, the cheaper its price, but classic cars tend to be more expensive the older they are.**

For cars manufactured before 1988, other factors have a greater impact on their price.



The data suggests that the market considers a classic car to be a car manufactured before 1988. This is **10 years more** than the common definition of a classic car: a vehicle 25 years or older. Classic cars tend to become more expensive the older they are.





# Key Questions and Answers

# Key Questions

- 1 What are the factors that have the greatest effect on the price of a used car?
- 2 Are there any specific models or manufacturers that are driving sales growth, and how can we capitalize on this trend?
- 3 What is the optimal mix of inventory (e.g., popular models, price ranges) to maximize sales?
- 4 What other noteworthy insights should the leadership consider?



# Q.1 Answer:

**What are the factors that have the greatest effect on the price of a used car?**

The factors that seem to have the greatest effect on the price of a used car are:

- **Location** - used cars in rural America are more expensive than used cars in urban areas.
- **Condition** - The newer the car seems, the more expensive it is.
- **Body type** - Trucks, pickups, and off-roads are the most expensive at over \$20,000. Sedans, hatchbacks, and wagons are the cheapest at under \$10,000.
- **Body size** - full-size used cars are the most expensive
- **Year of manufacture** - The older the car, the cheaper it is, except for classic cars (cars manufactured before 1988) where the older it is, the more expensive it likely will be.



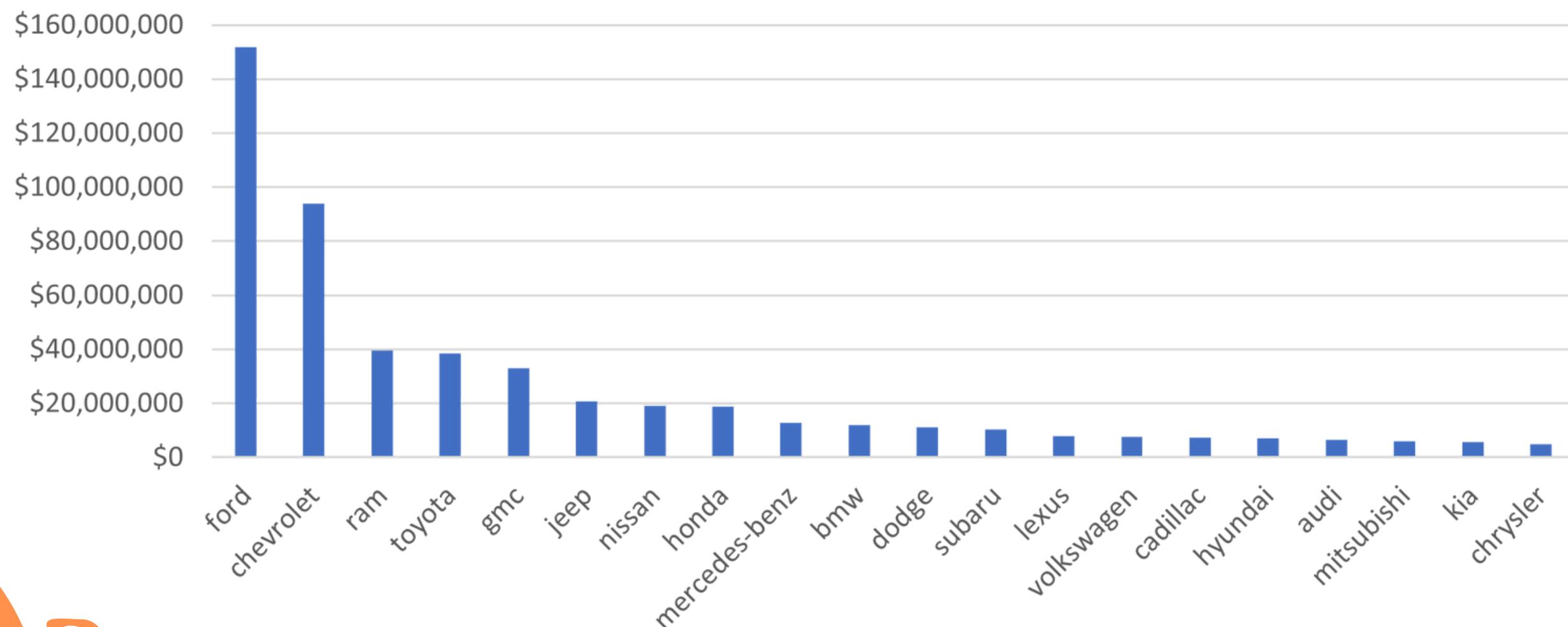
# Question #2

**Are there any specific models or manufacturers that are driving sales growth, and how can we capitalize on this trend?**

The top 20 potentially most profitable used cars by manufacturer based on available inventory.

The best performing manufacturers include:

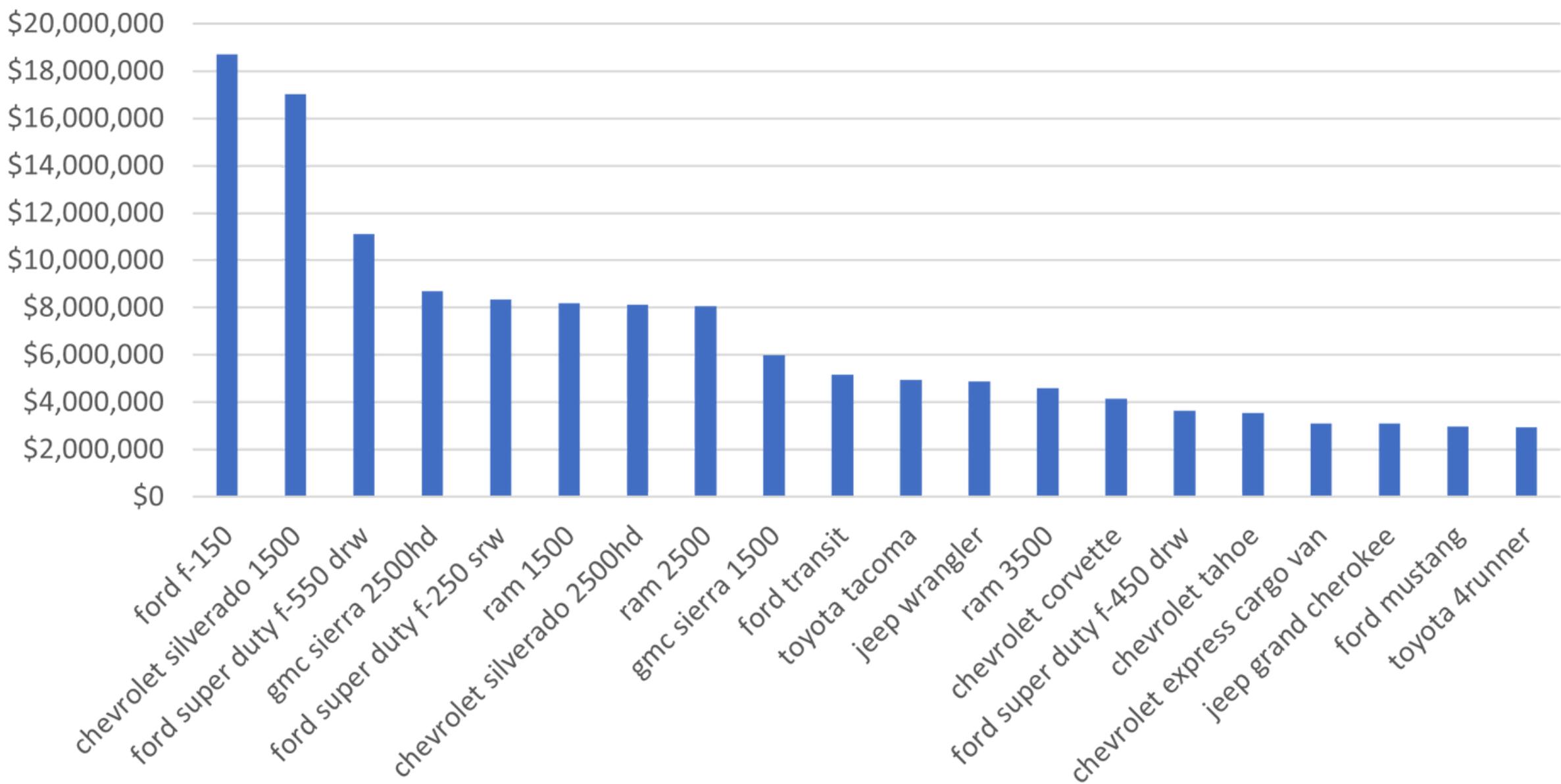
- Ford
- Chevrolet
- RAM
- Toyota
- GMC



## Top 20 most profitable used car models based on possible sales of available inventory.

The best performing models include:

- Ford F-150
- Chevy Silverado 1500
- Ford Super Duty F-550 DRW
- GMC Sierra 2500 HD
- Ford Super Duty F-250 SRW



**In the last 12 months**, the following trucks, SUVs, and sedans have been gaining in popularity amongst potential used car buyers:

**Trucks:**

- Used tow trucks under \$10,000
- Used 3/4 ton trucks
- Used 2500 trucks under \$30,000
- Used 4x4 trucks under \$10,000
- Used Toyota Tundra

**SUVs:**

- Chevy Trailblazer
- Used Rivian SUV
- Used Toyota Sequoia
- 3 row SUV

**Sedans:**

- Used Acura TLX
- Used Honda Civic
- Used sport sedans
- Used luxury sedans
- Used Honda Accord
- Used Hyundai Sonata

**In the last 12 months interest for:**

**Used cars** was highest in:

- District of Columbia
- Illinois
- California
- New Jersey
- Hawaii

**Trucks** was highest in:

- Montana
- Arkansas
- Alabama
- Vermont
- West Virginia

**SUVs** was highest in:

- Rhode Island
- New Hampshire
- Massachusetts
- Minnesota
- Pennsylvania

**Sedans** was highest in:

- Rhode Island
- Minnesota
- Nebraska
- Kansas
- Mississippi

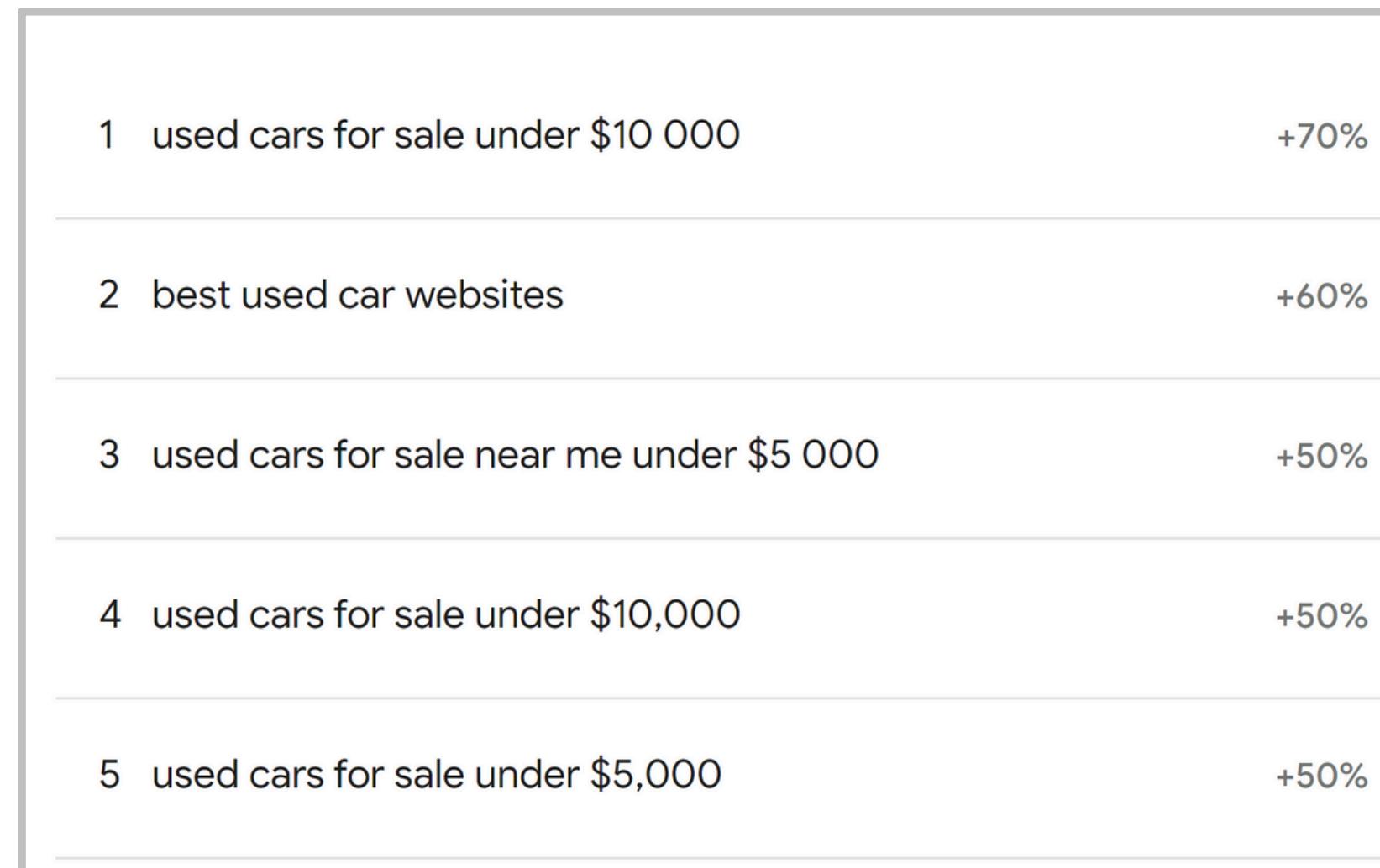
Source: Google Trends: <https://trends.google.com/trends/explore>



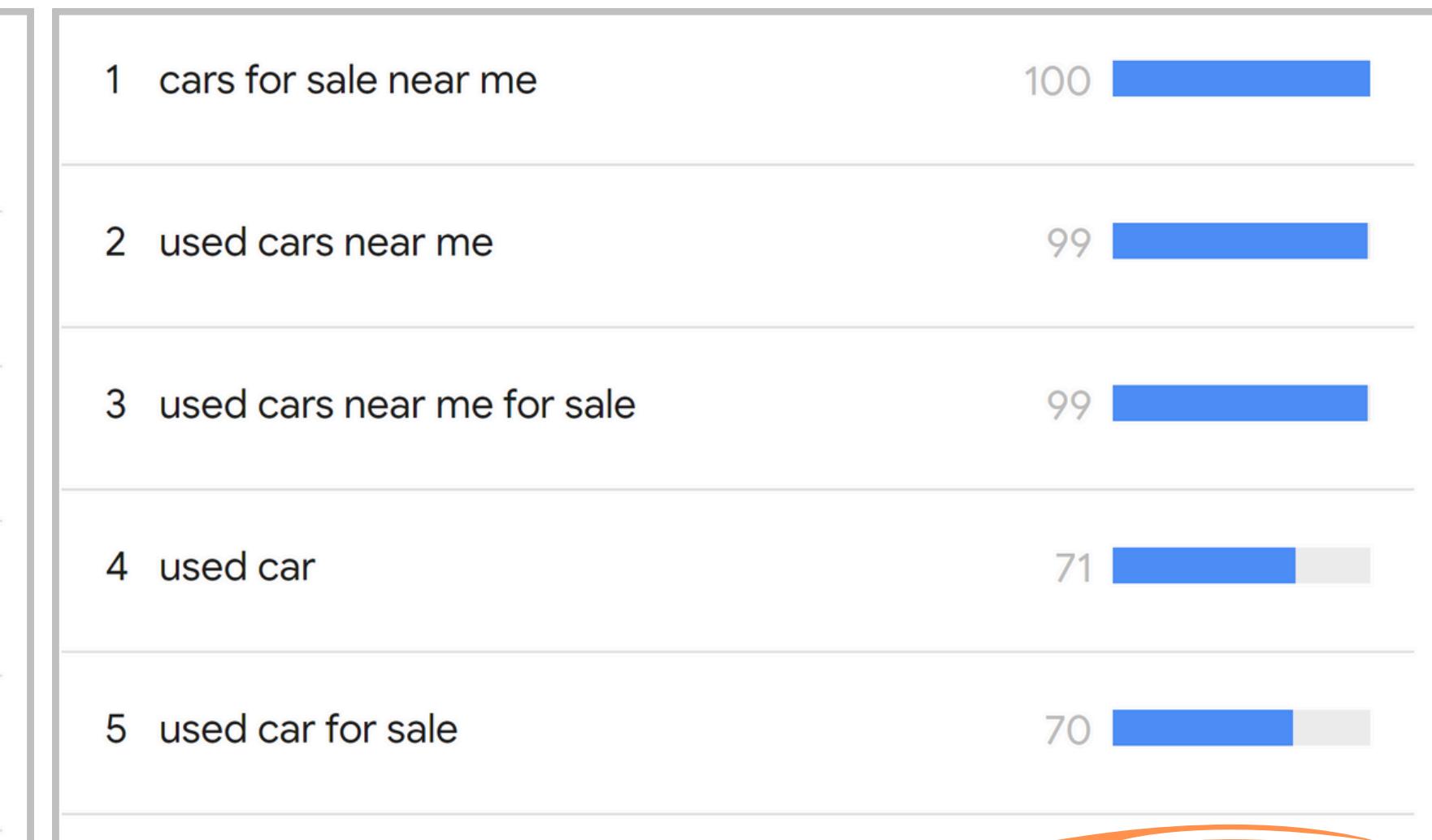
The majority of Google searches for used cars for sale are price-based and location-centric, evidenced by the use of phrases like “near me” and “under \$10,000”.

Customers in the last 12 months also seem interested in finding reliable websites for used cars, evidenced by a 60% rise in search volume.

### Rising search keywords



### Top search keywords



Source: Google Trends: <https://trends.google.com/trends/explore>

# Q.2 Answer:

The brands dominating the used cars market are **Ford, Chevrolet, RAM, Toyota, and GMC**. So it would be prudent to ensure the dealership's inventory includes a good number of these brands.

**Ford and Chevrolet alone account for 44.55%** of all for sale used cars in the inventory.

Models like the **Toyota Tundra, Chevy Trailblazer**, Rivian SUV, Toyota Sequoia, Acura TLX, and Honda Civic have gained traction among used car buyers. When creating the dealership's inventory, it would be prudent to take this into account.

Given the growing interest in reliable websites from which to buy used cars, the dealership should **revamp its website** to take advantage of this shift in customer behavior. I would recommend investing in SEO-optimized content created to target states where used cars are gaining traction and based on the type of used car that would-be car buyers are interested in.

The data suggests that **a buyer's buying process begins online but most buyers prefer to buy used cars that are near their location**. The dealership could include a location category in its website where users can browse through the used cars available in the dealership's branch nearest to them.



# Q.3 Answer:

What is the optimal mix of inventory (e.g., popular models, price ranges) to maximize sales?

Here are my recommendations towards an optimal mix of inventory to maximize sales:

## 1. Price Segmentation

- We should segment the inventory based on price ranges and condition categories to cater to price-sensitive buyers and those seeking newer, well-maintained vehicles, keeping in mind the average price of used cars in the dataset (\$17,002).
- We should also consider offering a range of options, including trucks under \$10,000, 3/4 ton trucks, and SUVs under \$30,000 to attract a broader audience.
- Focus on stocking a variety of affordable options, particularly in regions where prices are relatively lower, such as Pennsylvania and the District of Columbia.
- Consider stocking a few higher-priced vehicles, especially in regions with higher average prices, like Alaska and Alabama, to cater to customers with higher budgets

## 2. Focus on Top-Performing Brands and Models

- Given that Ford, Chevrolet, RAM, Toyota, and GMC dominate the used car market, we should prioritize stocking inventory from these brands.
- We should pay special attention to popular models such as Ford F-150, Chevy Silverado 1500, Toyota Tundra, etc., which have gained traction among used car buyers.
- Monitor market trends and adjust inventory accordingly to capitalize on emerging preferences and interests of potential buyers.



### **3. Consider Location-Specific Preferences**

- We should tailor the dealership's inventory mix based on location-specific preferences. For example, trucks are popular in states like Montana, Arkansas, and Alabama, while SUVs are in demand in Rhode Island, New Hampshire, and Massachusetts. We should adjust our inventory to meet the preferences of buyers in these regions.

### **4. Body Type and Size Consideration**

- We should take into account the preferences for specific body types and sizes. Given that trucks, pickups, and off-roads are more expensive, we can ensure a selection of full-size trucks in the inventory.
- Additionally, let's consider offering a variety of sedan options, including luxury sedans and sport sedans, to appeal to different buyer preferences.

### **5. Condition and Mileage Proposition**

- Focus on stocking a mix of used cars in excellent, good, and like-new condition, as they account for the majority of cars on sale.
- Highlight the value proposition of cars labeled as in excellent or good condition, as they tend to attract higher prices.

### **6. Paint Color Preferences**

- Include a significant number of white and black cars in the inventory, as they are the most popular colors nationwide.
- Consider regional preferences for paint colors, such as offering more black cars in states like District of Columbia, Massachusetts, and Oregon.



# Q.4 Answer:

What other noteworthy insights should the leadership consider?

## Online Presence and SEO Optimization

- Invest in revamping the dealership's website to align with the shift in customer behavior towards online research and purchasing.
- Optimize the website with SEO-friendly content targeting states with growing interest in used cars and specific models.
- Incorporate location-based browsing features to facilitate easy access to nearby inventory.

## Customer Engagement and Experience

- Enhance customer engagement and experience by providing detailed information, high-quality images, and transparent pricing for all inventory listings on the website.
- Offer tools such as location-based browsing and price filters to facilitate an efficient and personalized shopping experience.

- Actively engage with customers through surveys, feedback forms, and social media channels to gather insights into their preferences and satisfaction levels.
- Use customer feedback to refine the inventory mix, improve service offerings, and tailor the dealership's approach to meet the needs of its target audience effectively.

## Regular Inventory Updates and Monitoring

- Continuously monitor market trends, sales performance, and customer feedback to make informed decisions about inventory management.
- Regularly update the inventory mix based on changing preferences, seasonal demand fluctuations, and emerging market opportunities.



# Summary of Key Findings

## Key Finding #1

Used car prices vary significantly across states, with the District of Columbia offering the lowest average prices (\$10,633) and Alaska being the most expensive (\$27,368).

## Key Finding #2

Analysis reveals that location, condition, body type, and year of manufacture are primary determinants of used car prices, with rural areas commanding higher prices and full-size trucks being the most expensive body type.

## Key Finding #3

Certain models, such as the Toyota Tundra and Chevy Trailblazer, have gained traction among buyers in the past 12 months, suggesting opportunities for inventory optimization.

## Key Finding #4

Insights into customer preferences, including transmission type, fuel type, body size, and paint color, provide valuable guidance for inventory selection and marketing strategies.

## Key Finding #5

The analysis highlights the importance of enhancing the dealership's online presence, leveraging SEO-optimized content and location-based browsing features to cater to evolving customer behaviors.





# Recommendations & Conclusion

# Recommended Actions

## Recommendation #1

Prioritize stocking popular models from top-performing manufacturers, such as Ford and Chevrolet, while considering emerging market trends and regional preferences.

## Recommendation #2

Tailor pricing strategies based on location, condition, and body type, leveraging insights into price drivers to maximize profitability and market competitiveness.

## Recommendation #3

Revamp the dealership's website to improve user experience and incorporate SEO-optimized content targeting states with high search volume and specific vehicle preferences.

## Recommendation #4

Implement location-based browsing features to facilitate convenient browsing of inventory and enhance the overall customer experience.





# Conclusion

This data analysis project has yielded invaluable insights that hold significant potential to influence strategic decision-making and optimization efforts within our U.S. car dealership's used car market. Through a meticulous examination of various attributes, including price, location, manufacturer, condition, and customer preferences, we have acquired a deeper comprehension of market dynamics and emerging trends.

The findings emphasize the necessity of adapting our inventory selection, pricing strategies, and marketing endeavors to harmonize with customer preferences and market conditions. Implementing these recommendations and adopting a proactive approach to inventory management and customer engagement will help us become a leader in the used car industry. This positions us for sales growth and long-term success in the highly competitive used car market.



Thank you!