**🚗 Final Answer to:**

**What is the most popular car listing on this platform?**

**1. Definition of Popular Car Listing (for our analysis)**

👉 **Popular** means a car that:

* Has a **high average ConsumerRating** (buyers love the car)
* Has a **high number of ConsumerReviews** (many buyers engaged with the car)
* Was **sold by a seller with high SellerRating and SellerReviews** (trustworthy seller experience)

**We created a Total Popularity Score** based on:

Total Popularity Score=(ConsumerRating×ConsumerReviews)+(SellerRating×SellerReviews)\text{Total Popularity Score} = (\text{ConsumerRating} \times \text{ConsumerReviews}) + (\text{SellerRating} \times \text{SellerReviews})Total Popularity Score=(ConsumerRating×ConsumerReviews)+(SellerRating×SellerReviews)

✅ So **Popularity = Quality × Quantity**, for both car and seller.

**2. Why not just number of listings?**

* **Number of listings** alone would only tell you what cars are **frequently posted** — not if buyers **actually love them**.
* A car might be posted often but have **low ratings** or **bad reviews** (i.e., not truly popular with buyers).

✅ That's why we didn’t define popularity just by how many times it was listed.

**3. Why not just rating?**

* A car with **one perfect review** (5.0) is not as trustworthy as a car with **500 good reviews** (4.7).
* We needed to consider **both** **average rating** **and** **volume of reviews**.

✅ That’s why we multiplied **Rating × Reviews** — to balance quality and quantity.

**4. Our Final Method:**

| **Step** | **What We Did** |
| --- | --- |
| Grouped by Year, Make, Model | Car type definition |
| Averaged ConsumerRating, SellerRating | Measure satisfaction |
| Summed ConsumerReviews, SellerReviews | Measure engagement |
| Created Consumer and Seller Popularity Scores | Separate popularity angles |
| Combined them into Total Popularity Score | Overall buyer love + seller trust |

**🚗 Why do we define unique car type using Year, Make, and Model?**

| **Reason** | **Explanation** |
| --- | --- |
| **Year matters** | Cars change every year — new designs, features, tech. A 2018 Camry is not the same as a 2020 Camry. |
| **Make matters** | Different brands build cars differently — Honda ≠ Toyota ≠ Ford. Buyer loyalty and reliability perceptions vary by brand. |
| **Model matters** | Even within a brand, different models serve different purposes — Civic vs Accord, Corolla vs Camry — different sizes, audiences, prices. |
| **Trim optional** | Trim (like EX, LX) adds fine details (sunroof, leather), but basic car type is already defined by Year + Make + Model. |
| **Available data** | Our dataset (c\_dataf) consistently has Year, Make, and Model for every car, but other details (like Trim) may be missing or inconsistent. |
| **Grouping logic** | In our analysis code (groupby(['Year', 'Make', 'Model'])), this grouping cleanly aggregates ratings and reviews without mixing different car types. |

**✏️ In plain words:**

*Buyers don’t just shop for a "Camry" — they shop for a "2019 Toyota Camry."  
Year, Make, and Model together describe the specific version of the car that matters for pricing, popularity, and buyer decision-making.*

**📌 Key Summary:**

* **What defines a car type?**  
  → Year + Make + Model
* **Why these?**  
  → Because they together capture the true identity of the car: its design generation, brand reputation, and functional category.
* **Why not just Make and Model?**  
  → Because Year changes make huge differences in features, safety, performance, and buyer preferences.
* **Why not include VIN?**  
  → VIN identifies an individual *car* (one car). We want to define *car types* (categories) for analysis.
* **Why is Trim optional?**  
  → Not all listings specify Trim cleanly, and core buyer decisions are still heavily based on Year + Make + Model.

**🎯 Final Professional Statement (polished for report or slide):**

*We define a car type using Year, Make, and Model because these three fields together accurately capture the identity of a car as buyers experience it.  
This grouping reflects critical differences in design, brand reputation, and intended market, allowing for accurate popularity and rating analysis without unnecessary noise.*

Fixed 'Price' column formatting, created 'PopularityScore' (ConsumerRating × ConsumerReviews), grouped by Year+Make+Model, identified and removed inconsistent duplicates in grouping by dropping rows with fewer reviews, and re-identified most popular car listing.

What is the most popular car listing on this platform? - Define what is popular car listing. Is it number of listings of certain type of car? Rating? Or something else? - How to uniquely define a type of car? By country made, brand, year, etc?

Selected and filtered listings for the top car model.Explored variation across InteriorColor, Engine, and Transmission. Created new popularity metrics: **Rating × Reviews** and **Rating × Reviews × Listings**. Ranked variations by total impact score. Visualized the Top 5 variations with a bar chart.

“Create new popularity score based on ConsumerReviews, ConsumerRating, and sqrt(Listings) so that low quality high listing cars don’t overwhelm the rankings; rank and plot Top 10 cars. Created new Jupyter Notebook (Untitled1.ipynb) to double-check and validate results. Started fresh to avoid confusion from earlier steps.”

Create new popularity score based on ConsumerReviews, ConsumerRating, and sqrt based on VIN listings so that low quality high listing cars don’t overwhelm the rankings; rank and plot Top 10 cars. Created new Jupyter Notebook title Untitled1.ipynb to double-check and validate results; did this to avoid confusion from earlier steps in Untitled.ipynb file