

AutoBridge Analytics: Data Methodology

Data Methodology (English)

1. Philosophy and Transparency

Accuracy and absolute transparency are at the core of AutoBridge Analytics. We are convinced that the Georgian automotive market, being the largest re-export hub in the region, requires a reliable digital infrastructure. Our statistics are not built on seller surveys or manual expert estimates; they represent a direct, continuous cross-section of the actual digital automotive market.

2. Data Collection

Our ETL (Extract, Transform, Load) algorithms daily aggregate data from the largest automotive marketplaces and classifieds in Georgia.

- **Volume:** Our database is replenished with tens of thousands of new unique offers every month.
- **Frequency:** Data is tracked in real time, allowing us to record price drops or hikes, mileage changes, and the final publication status (sold/removed).

3. Cleaning and Verification (ETL Pipeline)

Raw data is always full of "noise," duplicates, and spam. We apply strict engineering filters:

1. **Deduplication:** Mega-dealers often publish the exact same vehicle across multiple websites or create duplicate listings to refresh the date. Our scripts merge these publications: we count **unique vehicles**, not the number of their advertising placements.
2. **Anomaly & Spam Filtering:** Listings with prices like \$1 , \$999,999 , or a listed mileage of 1 km for older cars are automatically excluded from analytical samples.
3. **Segment Isolation:** Within passenger car metrics, we algorithmically exclude trucks, heavy equipment, totaled cars meant for parts (dismantling), and daily rentals.

4. Key Metrics

- **Median Price:** To evaluate a model's cost, we **never** use the arithmetic mean (Average). A single erroneously expensive or severely damaged car can heavily distort the average price. The Median shows the actual price at which the "middle of the market" operates.
- **Liquidity Index:** This unique indicator reflects the median lifespan (in days) of a listing on the platform. The lower the index (fewer days), the faster the vehicle is resold, signaling hot demand.
- **Market Share:** Determined proportionally by the volume of valid, deduplicated publications of a specific brand within a chosen timeframe.

5. Data Limitations

Like any analytical platform, we have metric boundaries. We analyze the market's "digital footprint." Transactions conducted by dealers entirely "offline" (directly at the physical auto market without internet

publication) do not enter our statistics until digitized. However, the sheer scale of our digital sample ensures trend accuracy exceeding 95%.
