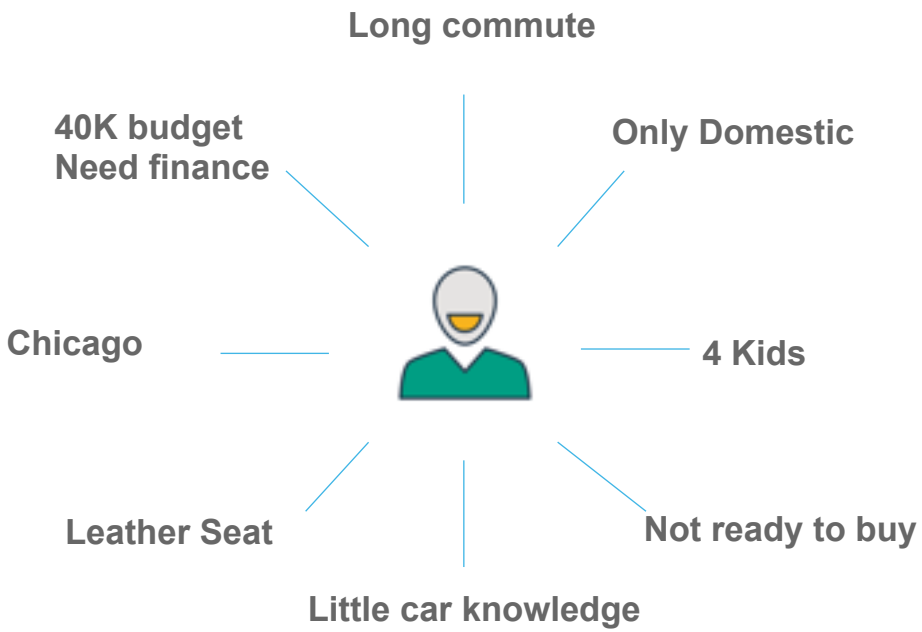
Using Data Science to Solve Platform Problems at TrueCar

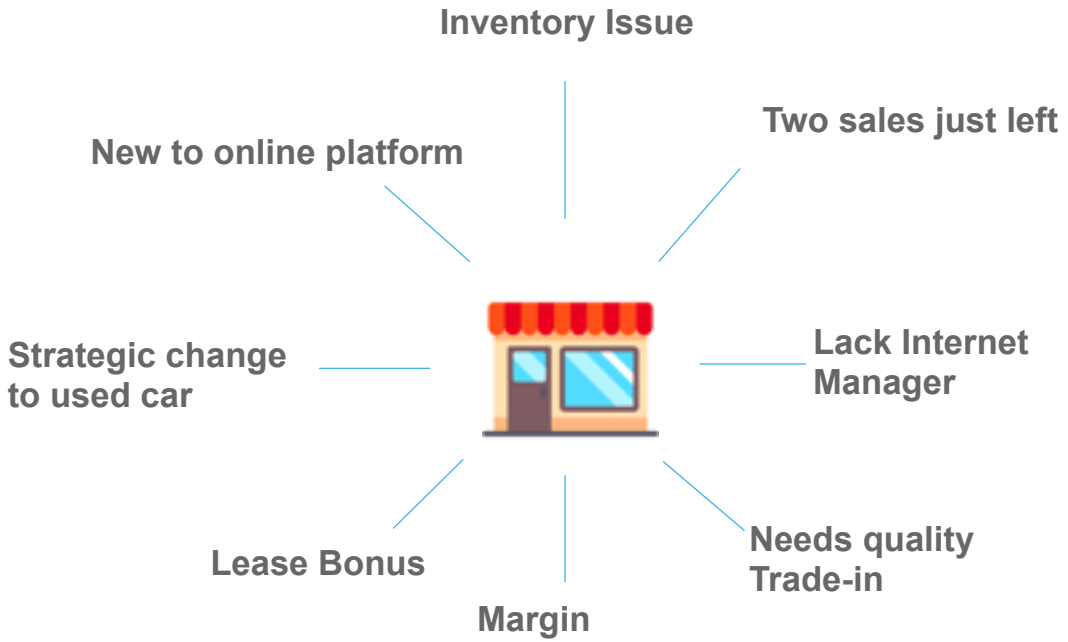
TRUECar

Rachel Wang

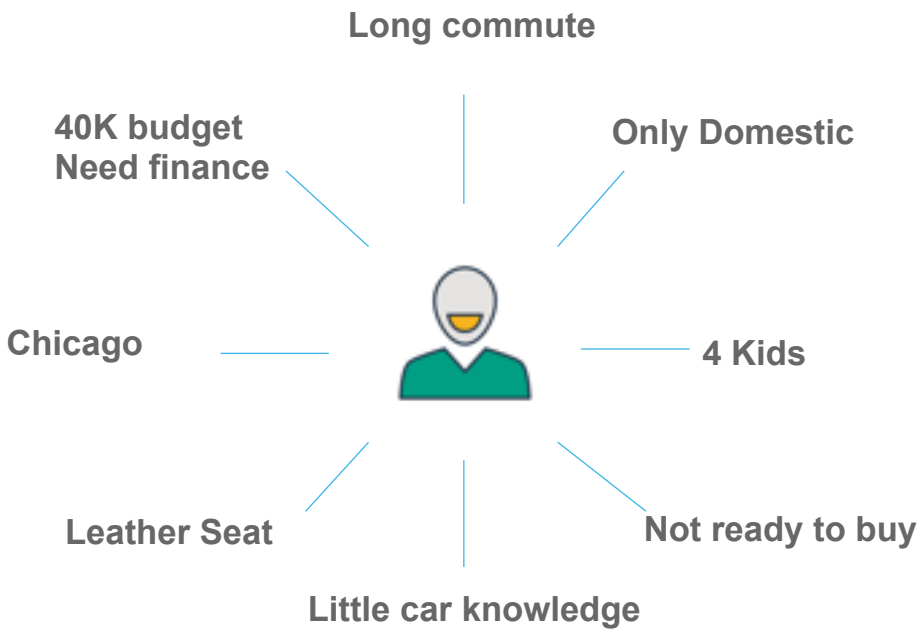
Platform Problems in the Auto Industry

Various Needs of Each Individual Adds Tremendous Complexity

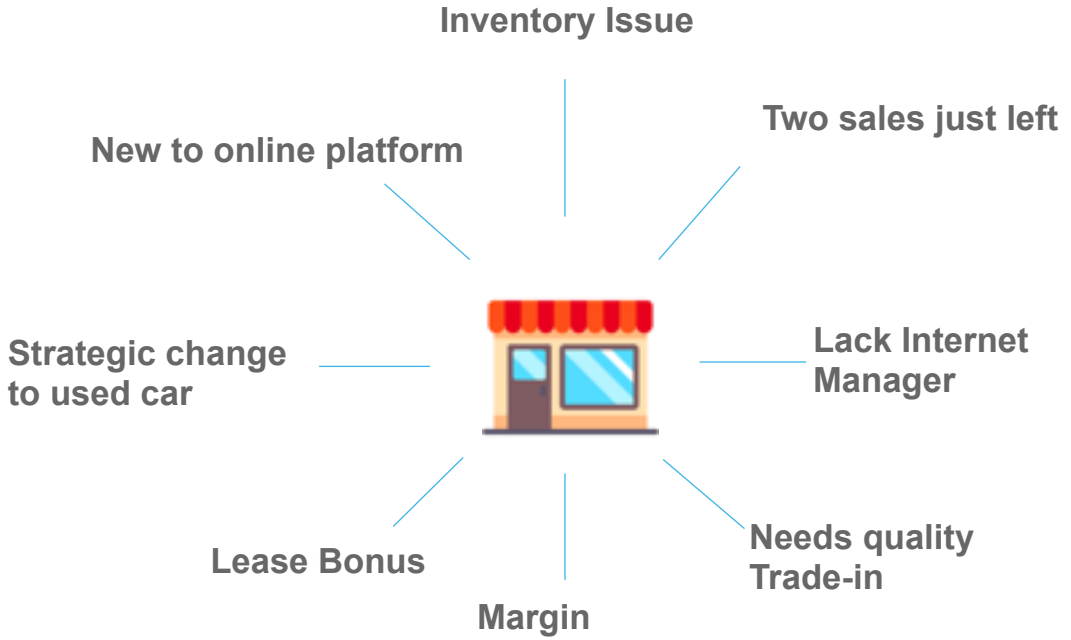




Various Needs of Each Individual Adds Tremendous Complexity

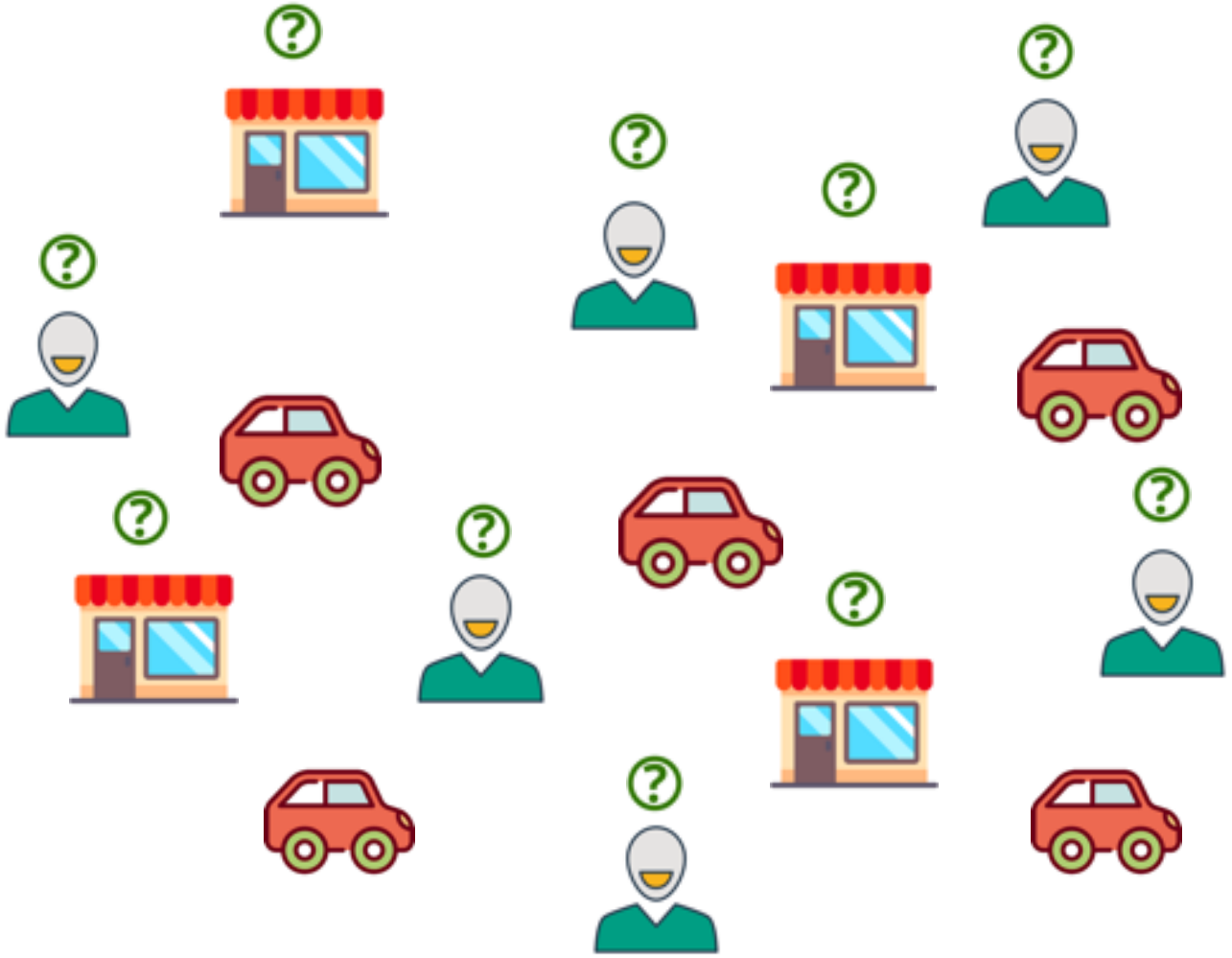


X 17M



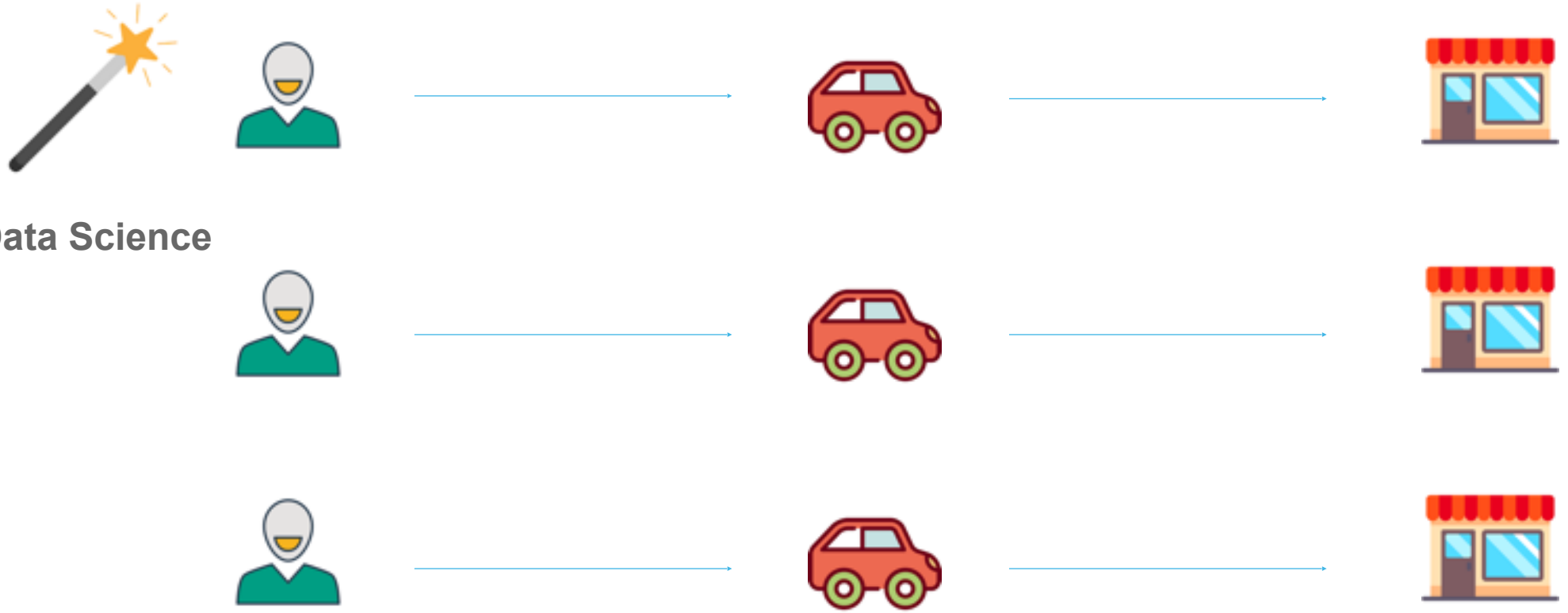
X 35K

A Marketplace Puzzle Only Can be Solved Through Data



Craigslist

Platform problem is the perfect use case for modern Data Science



Data Science

Platform Problems are generally…

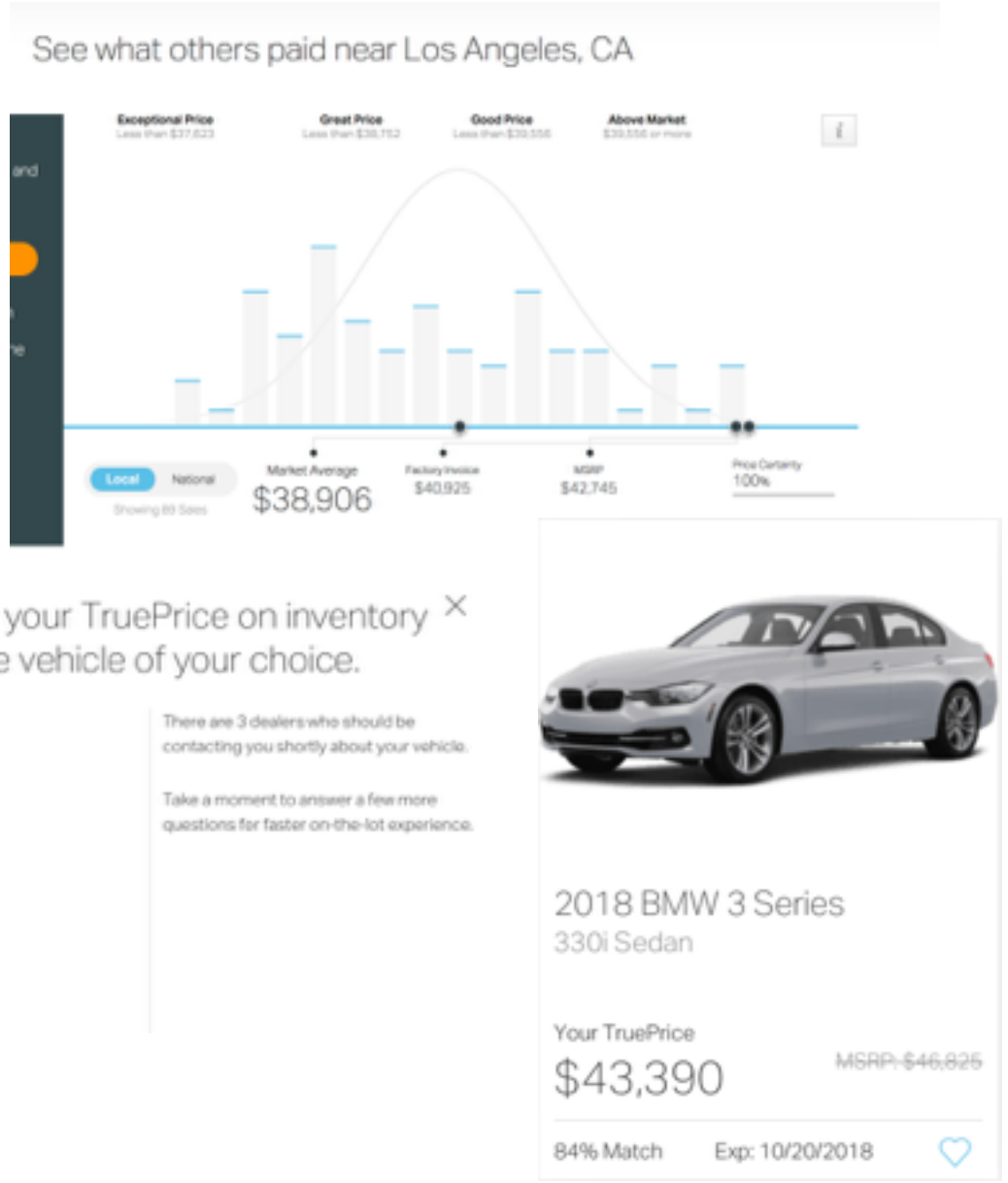
Classification
Customer Segmentation

Recommendation
Recommending Trim/Dealers

Optimization
Help dealer reduce time cost

Allocation
Inventory Allocation

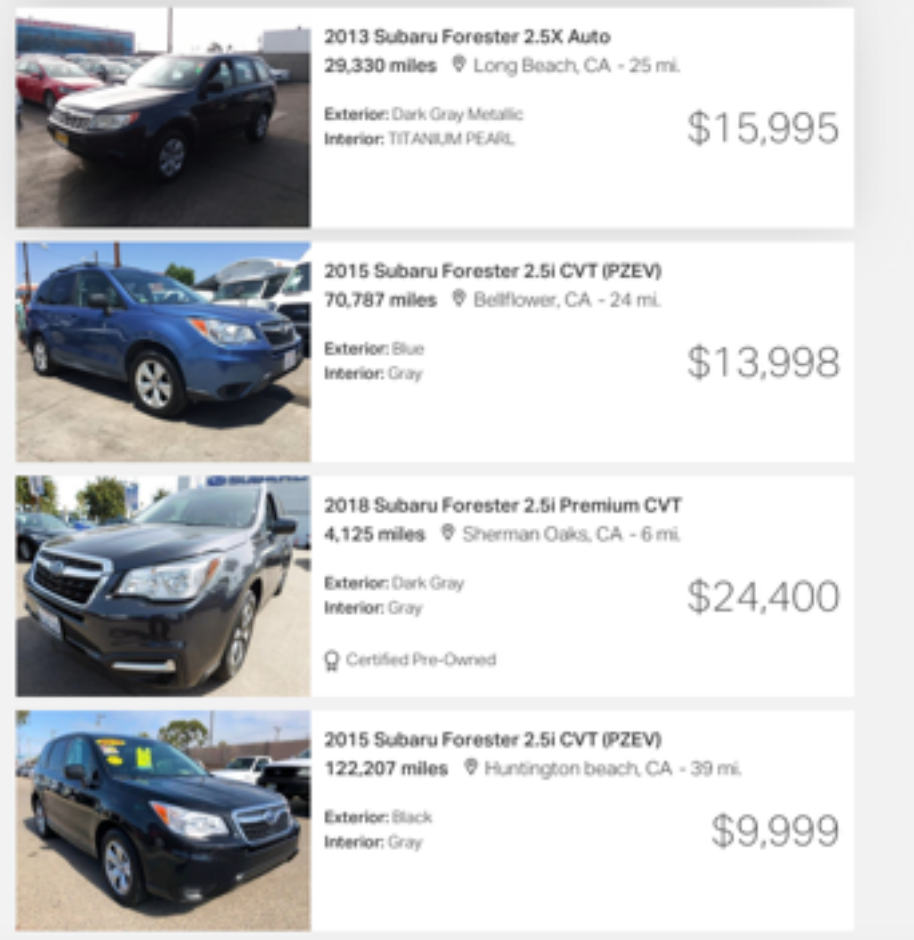
Sorting & Ranking
Used Car Sorting

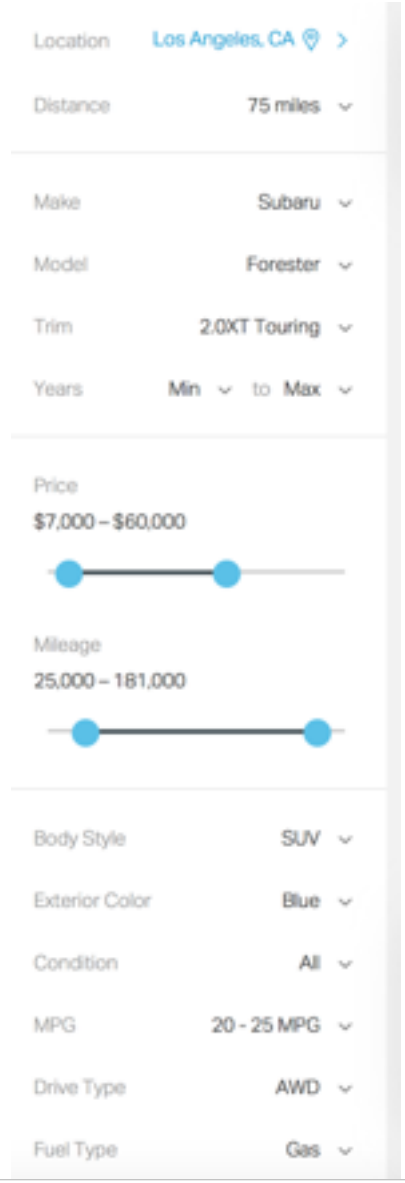


Platform Problem #1: Sorting

Used Car Sort at TrueCar

How to rank the used vehicles in our inventory?

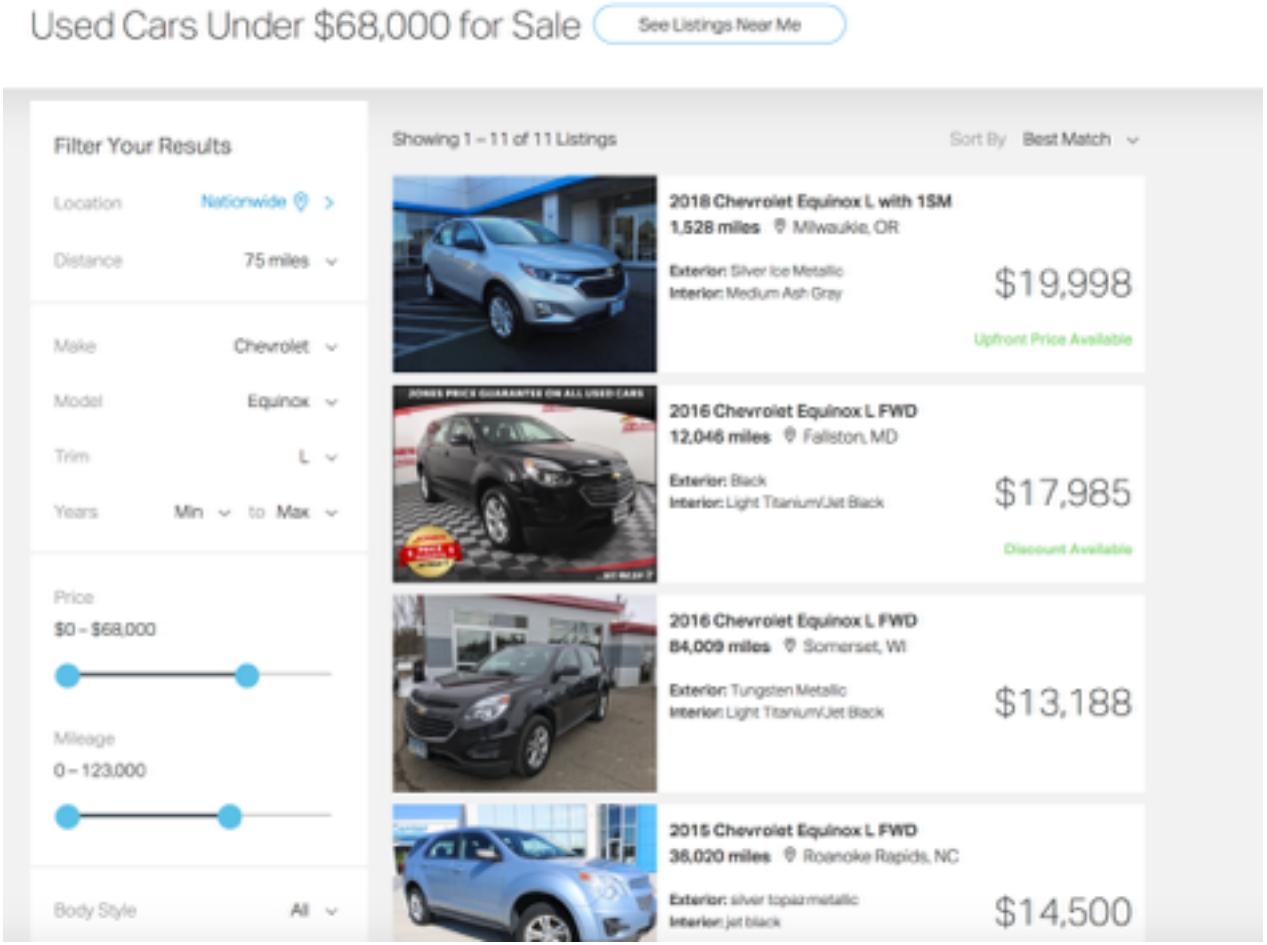




Used Car Sort

How to rank the used vehicles in our inventory?

Goal: show the most relevant vehicles to the top



Used Car Sort

How to rank the used vehicles in our inventory?

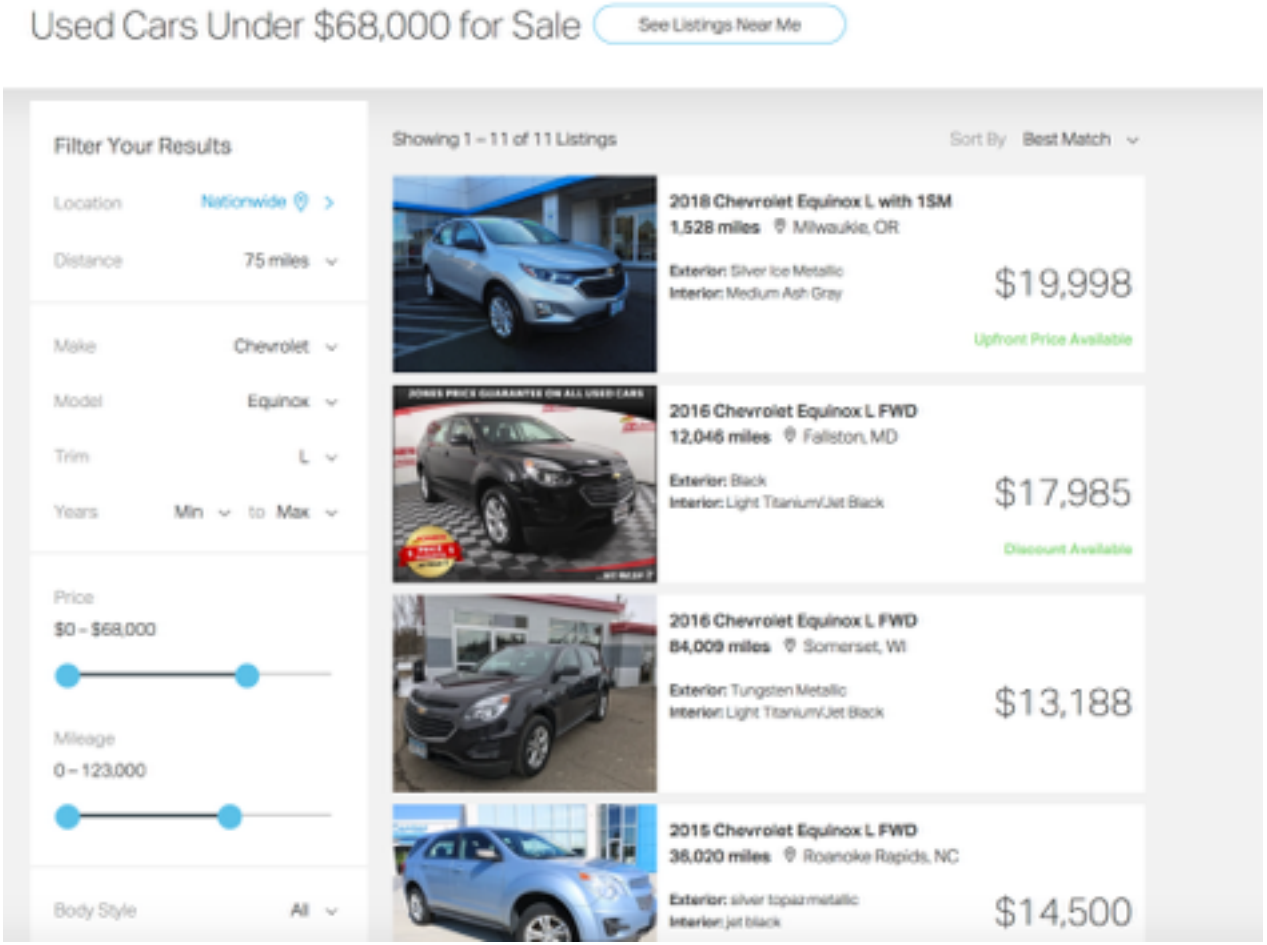
Goal: show the most relevant vehicles to the top

Clickstream Data

Filters

Vehicle

Clicks

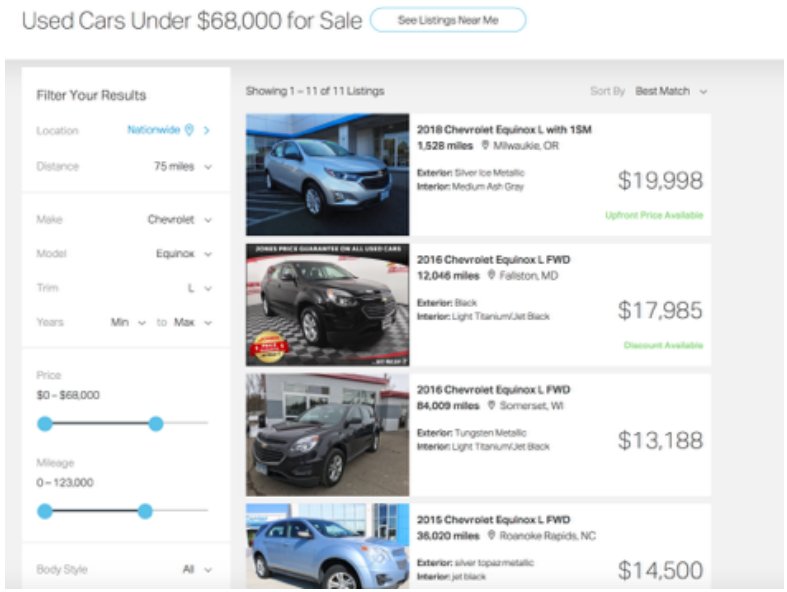


Used Car Sort

How to rank the used vehicles in our inventory?

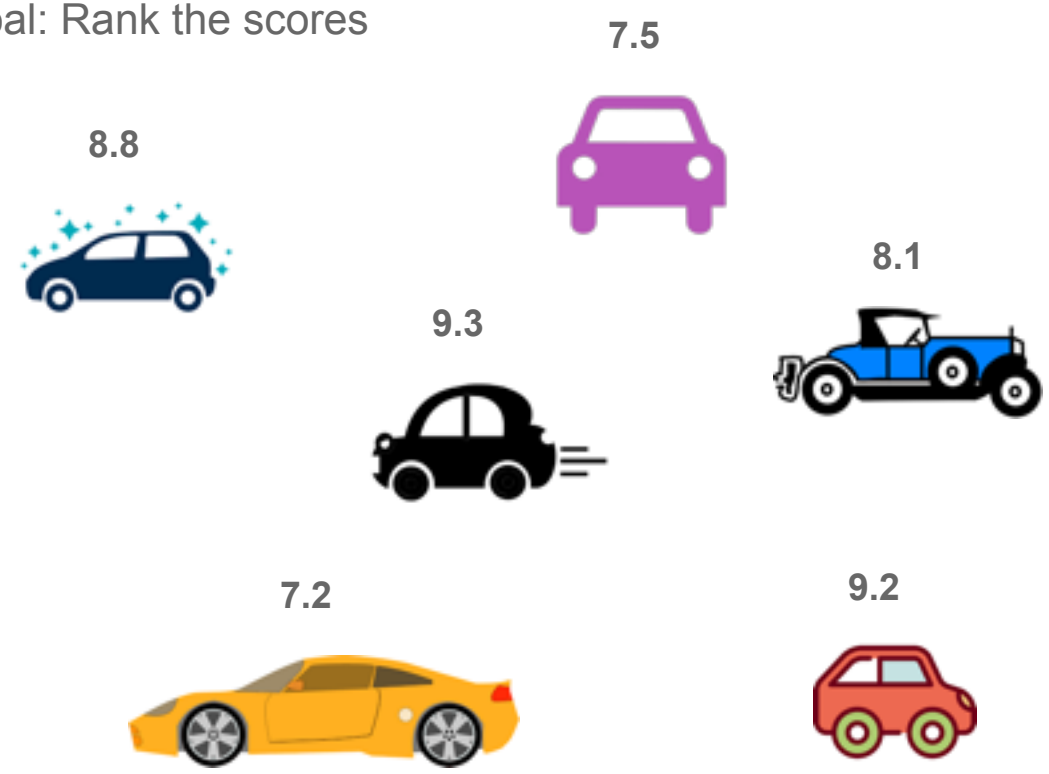
Goal: show the most relevant vehicles to the top

Clickstream Data: Filters, Vehicle, Clicks



Pointwise approach

Condition: Filters + Vehicle
Training Set: Click
Goal: Rank the scores

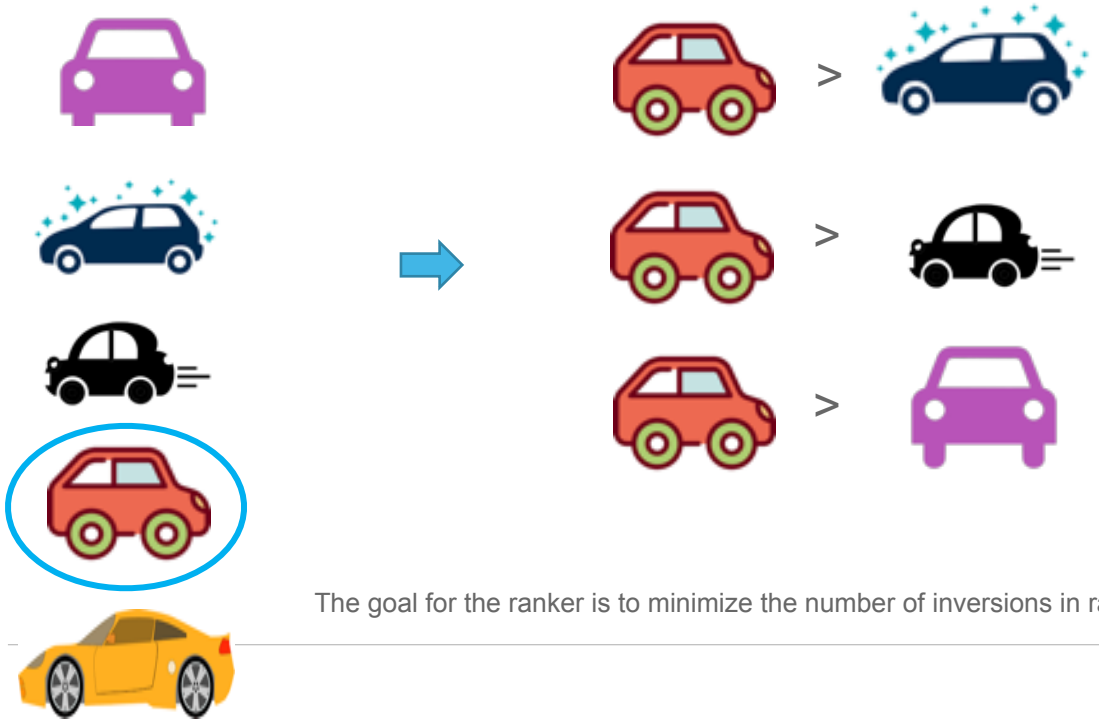


They essentially take a single car and train a classifier / regressor on it to predict how relevant it is for the current query. The final ranking is achieved by simply sorting the result list by these car scores. For pointwise approaches, the score for each car is independent of the other cars that are in the result list for the query

Used Car Sort

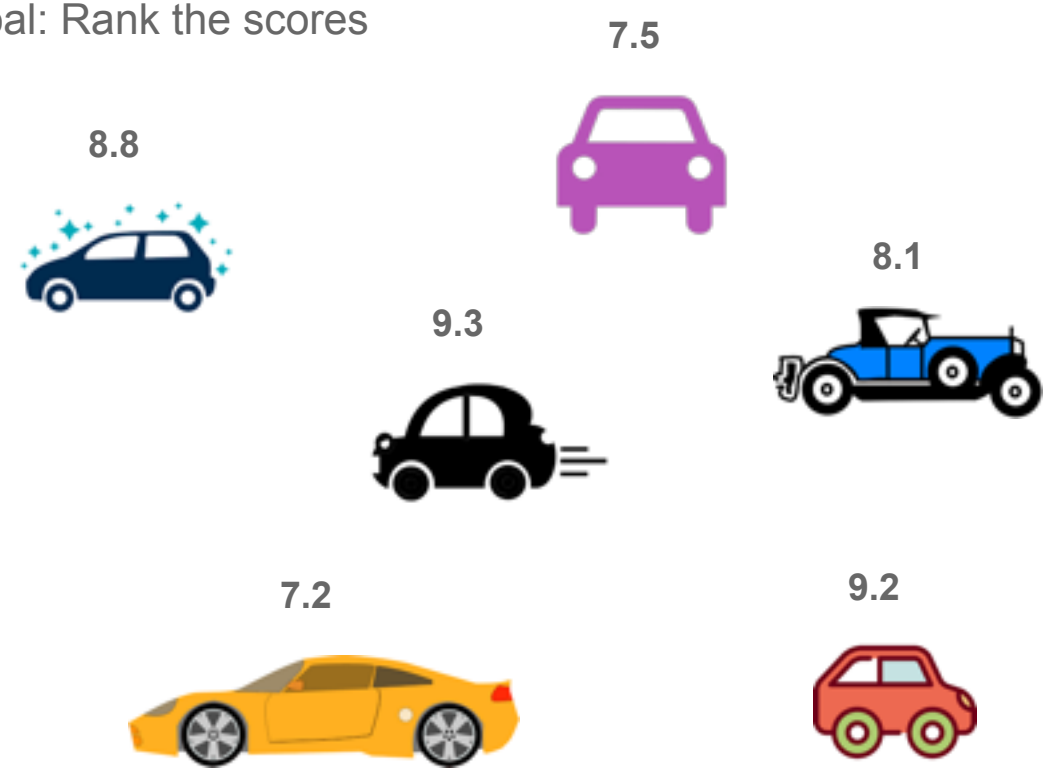
Pairwise approach

Condition: Filters + Vehicle
Training Set: Relative Position of A Pair of Cars
Goal: Minimize the number of inversions



Pointwise approach

Condition: Filters + Vehicle
Training Set: Click
Goal: Rank the scores



The goal for the ranker is to minimize the number of inversions in ranking i.e. cases where the pair of results are in the wrong order relative to the ground truth.

Used Car Sort

Pairwise approach

Condition: Filters + Vehicle
Training Set: Relative Position of A Pair of Cars
Goal: Minimize the number of inversions

Pointwise approach

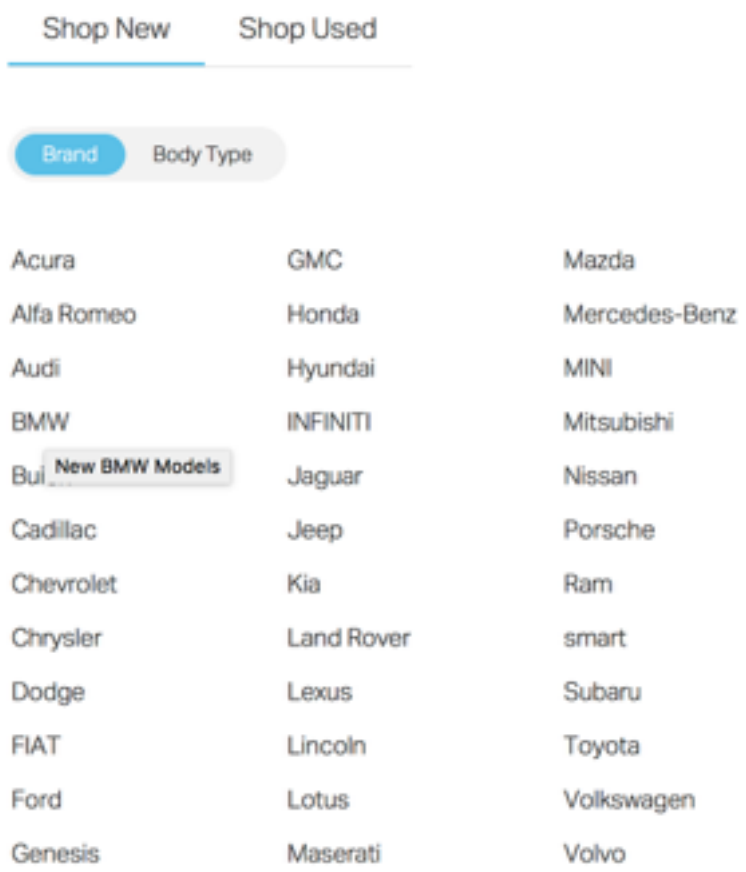
Condition: Filters + Vehicle
Training Set: Click
Goal: Rank the scores

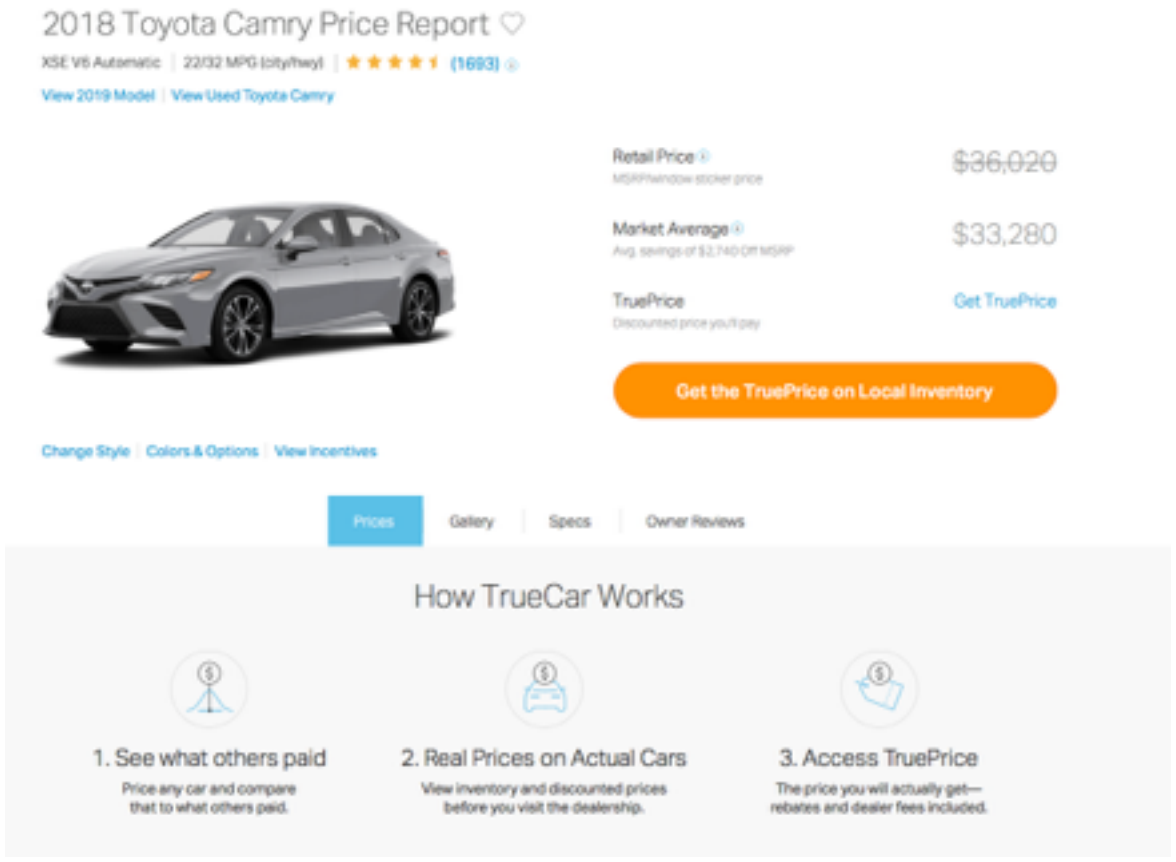
Improving?

Personalization, Real-time Adjustment

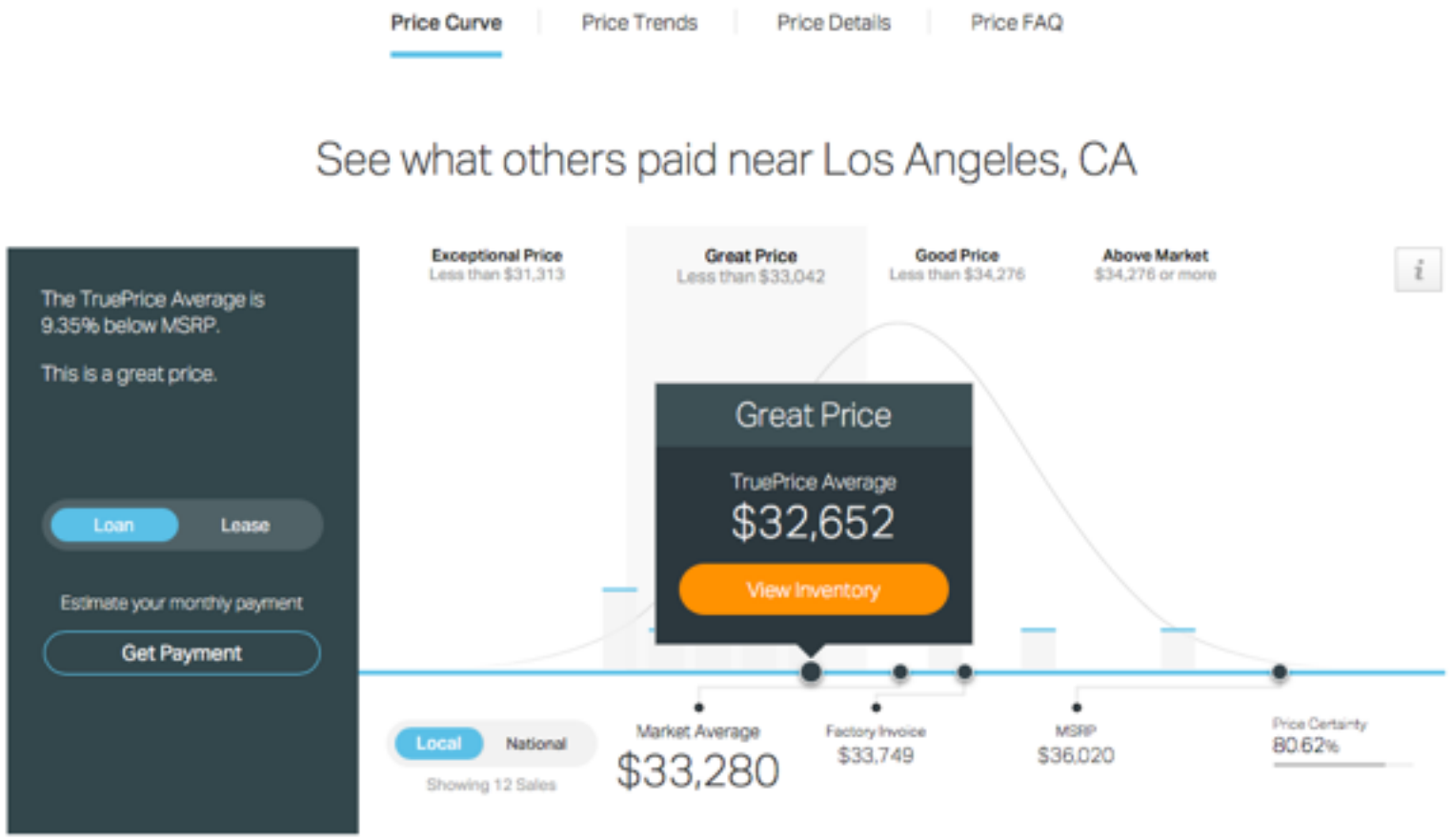
Platform Problem #2: Market Simulation

New Car Price Report Page





New Car Price Report Page



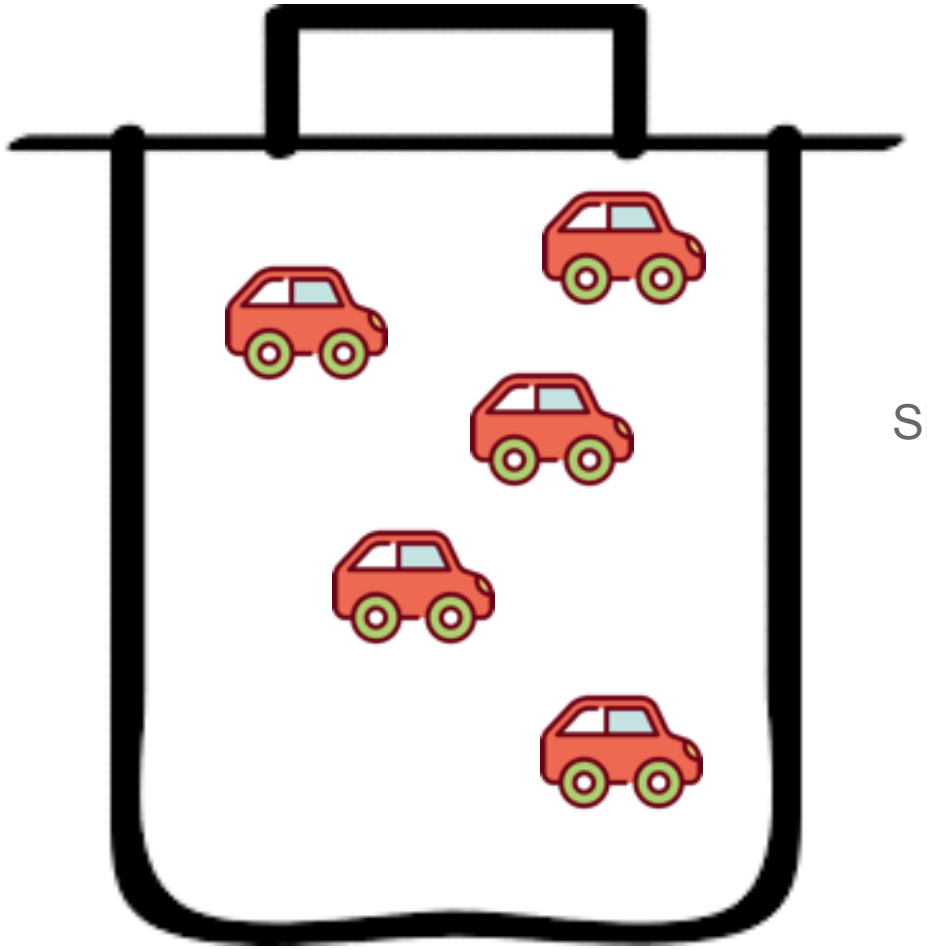
Using transaction data to mimic the auto market

for all makes, models, regions for all time

New Car Price Model Problem for Data Science to solve

A Binning Approach

New Car Price Model Problem for Data Science to solve



Simple Average

A Binning Approach

Statistical Model vs. Machine Learning

Random Forest?

Clustering?

Neutral Network?

Easy to explain

Cost of Implementation

Visualization

Implementing Data Science to Platform Problems

Classification

Recommendation

Optimization

Allocation

Sorting & Ranking

X

Understand the Business Need

Dance with Engineers

Know Your Audience

Automation Is Cool

Create Value

Questions?