

Used Car Recommendation System

Team 10

Cid Hernandez, Manuel Sanchez, Minting Chen, Sunny Chu, Jiawen Zeng

Motivations for our System

- Buying a car is a common task
- Used car sales are the most popular
- Can be a difficult task
- Remove emotion



Our Goals

- Find the best priced used car
- Provide useful insights about the cars
- Allow for user preference selection



Data Overview

- Data source: Craigslist
- Web scraped in January 2020
- 1.34 GB with 26 columns
- Removed NaN entries, selected useful rows

RangeIndex: 458213 entries, 0 to 458212
Data columns (total 26 columns):

#	Column	Non-Null Count	Dtype
0	Unnamed: 0	458213 non-null	int64
1	id	458213 non-null	int64
2	url	458213 non-null	object
3	region	458213 non-null	object
4	region_url	458213 non-null	object
5	price	458213 non-null	int64
6	year	457163 non-null	float64
7	manufacturer	439993 non-null	object
8	model	453367 non-null	object
9	condition	265273 non-null	object
10	cylinders	287073 non-null	object
11	fuel	454976 non-null	object
12	odometer	402910 non-null	float64
13	title_status	455636 non-null	object
14	transmission	455771 non-null	object
15	VIN	270664 non-null	object
16	drive	324025 non-null	object
17	size	136865 non-null	object
18	type	345475 non-null	object
19	paint_color	317370 non-null	object
20	image_url	458185 non-null	object
21	description	458143 non-null	object
22	state	458213 non-null	object
23	lat	450765 non-null	float64
24	long	450765 non-null	float64
25	posting_date	458185 non-null	object

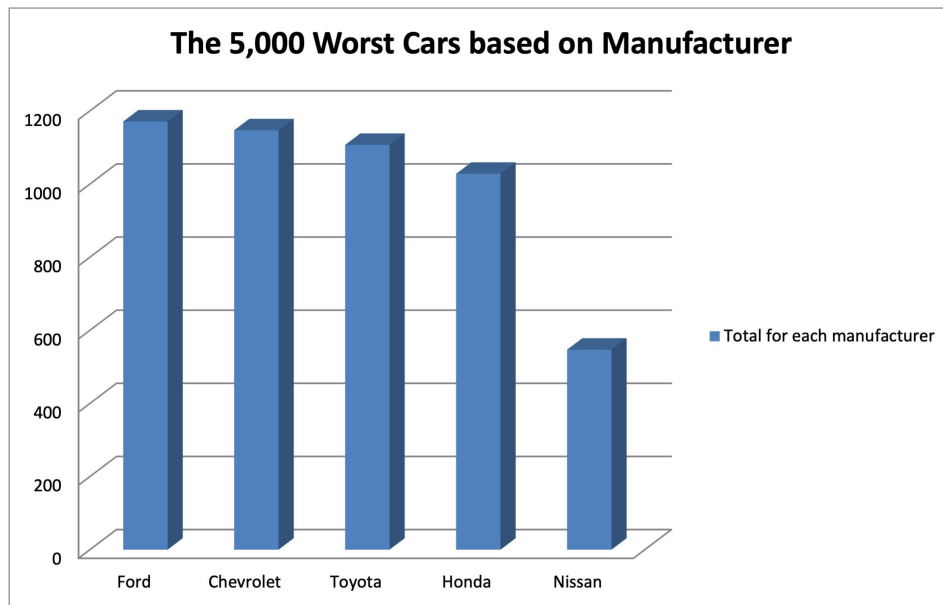


RangeIndex: 458213 entries, 0 to 458212
Data columns (total 7 columns):

#	Column	Non-Null Count	Dtype
0	condition	265273 non-null	object
1	odometer	402910 non-null	float64
2	title_status	455636 non-null	object
3	manufacturer	439993 non-null	object
4	model	453367 non-null	object
5	price	458213 non-null	int64
6	year	457163 non-null	float64

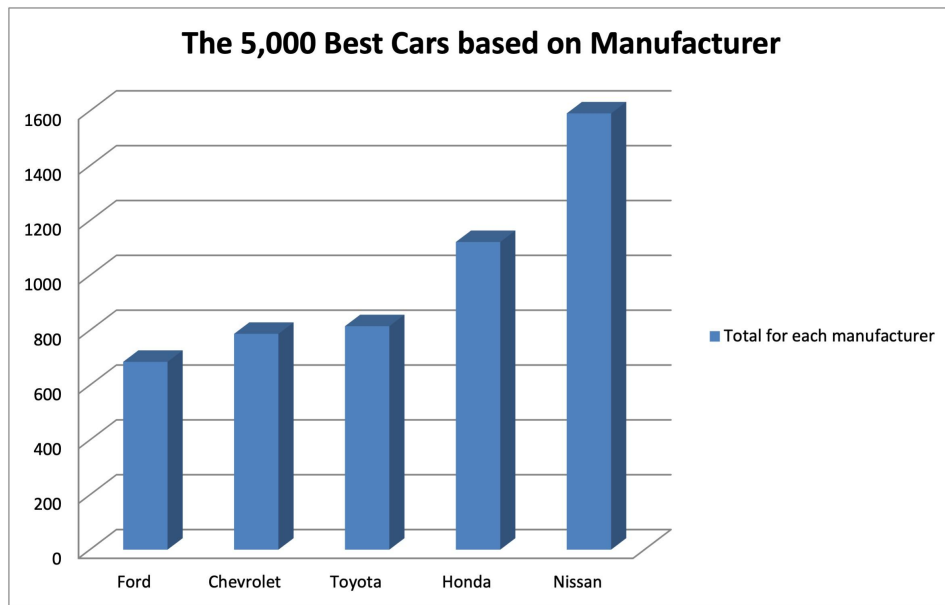
Choosing Wisely

- Ford, Chevrolet, Toyota, Honda, and Nissan are the most popular used manufacturers listed for sale
- Among these manufacturers, here's how they rank against each other
- Be Careful choosing your all time favorite manufacturer!
- This is important to keep in mind as you move forward with buying a used car



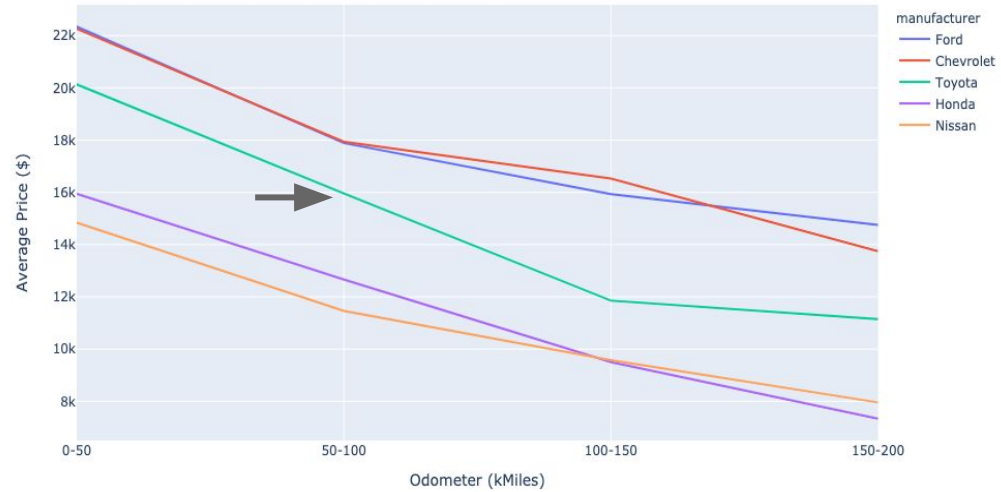
Choosing Wisely

- So what tends to be the least problematic car based on popular manufacturers?
- How does this help us?



Odometer vs Prices

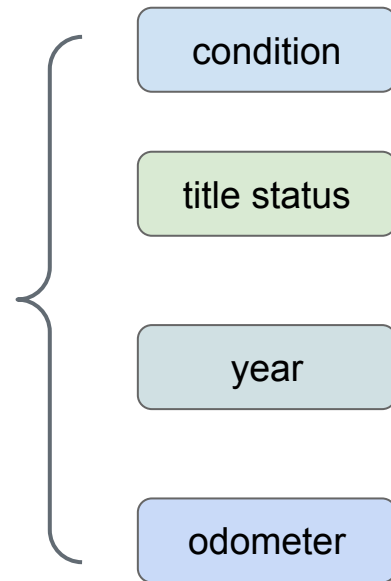
- Relationship between odometer and prices
- Ford and Chevrolet have higher prices
- Honda and Nissan have lower prices



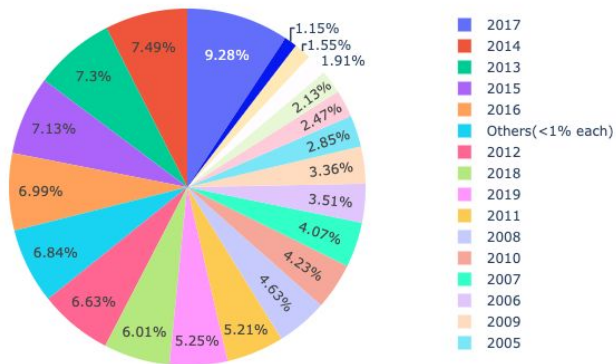
Factors to take into Consideration



Factors

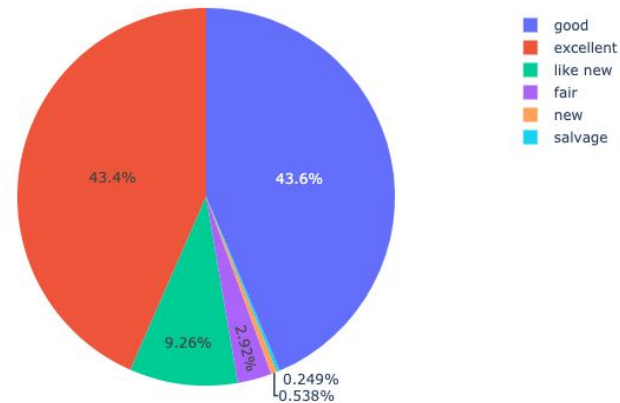


year statistics

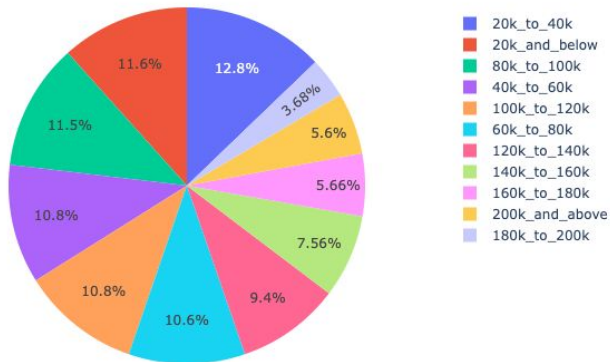


condition statistics

- Sectors on the left is balanced
- Sectors on the right is imbalanced

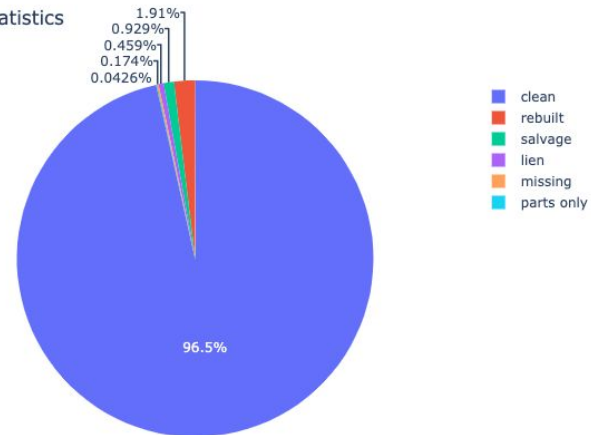


odometer_range statistics



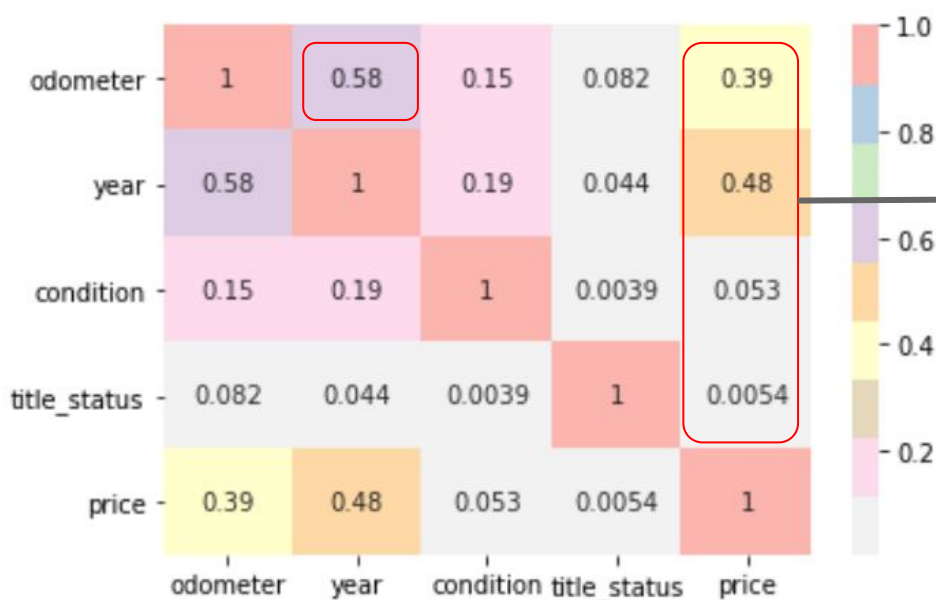
title_status statistics

- Balanced - > each sector is important - > weight more
- Imbalanced - > one or two sector(s) dominates - > weight less

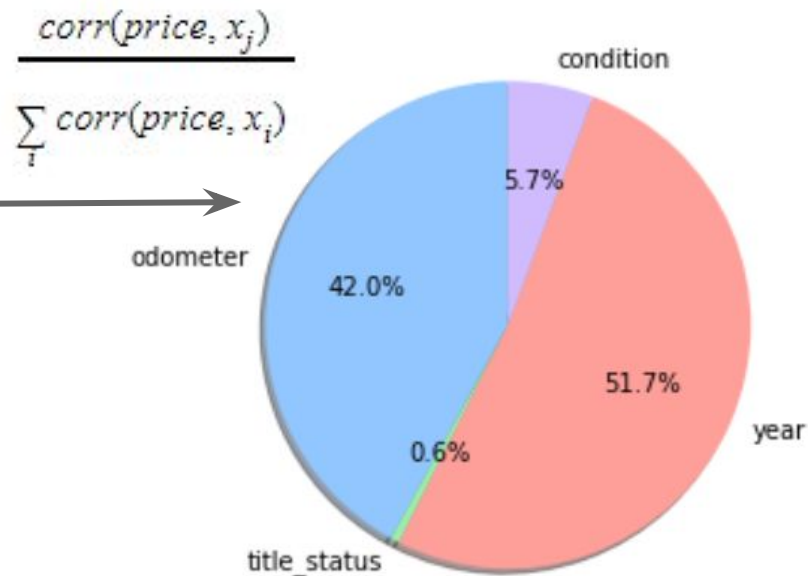


Correlation and Weight

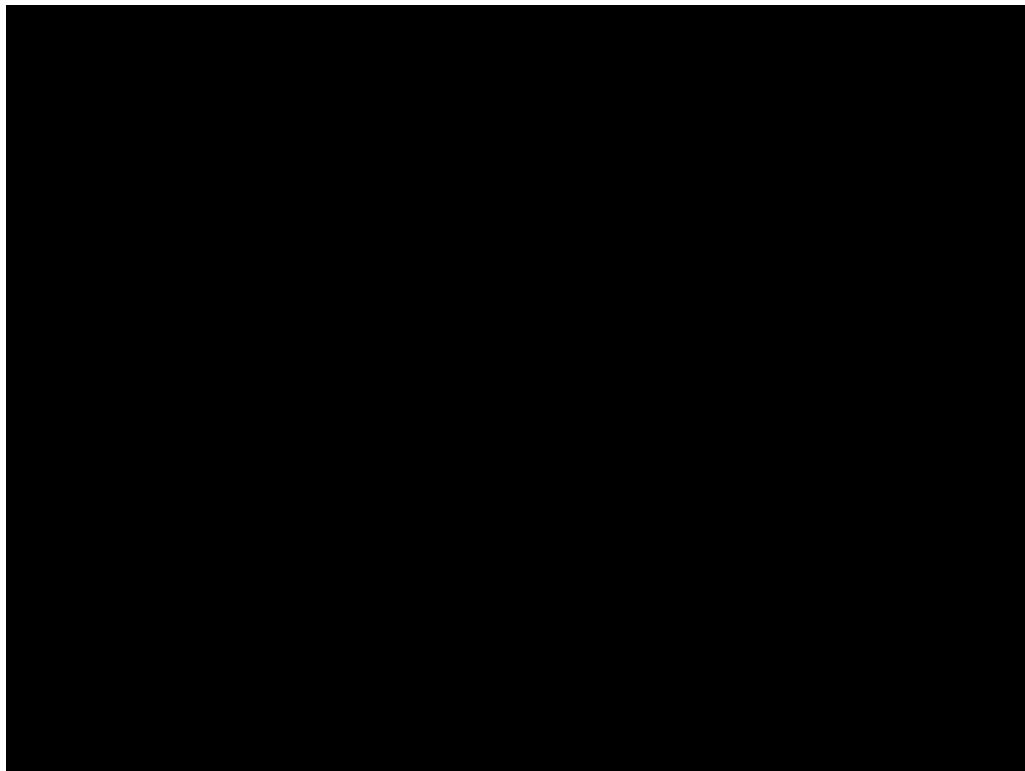
Correlation between Different Factors



Correlation Weight between Different Factors and Price



Demo



Conclusion

- Car buying can be easy with our system
- Car valuations backed with logic
- Possible depreciation system for future work

