CAR SALES ANALYSIS

Humza Khan

Agenda

- About Me
- Motivation
- Data Collection
- Features
- Exploratory Data Analysis
- Data Cleaning
- Multi-class Classification
- Customer Cluster Analysis
- PCA Visualization

About Me

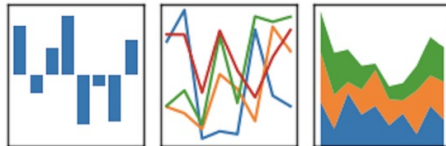- Hospital analytics

- Management consulting

- Why Metis?

Motivation

- Green Family Stores

- Which car brand will sell?

- What types of sales do I have?

Data Collection

Features

MSRP	Sale Price	Stock Number
VIN	Model Year	**Make** (Dodge, Kia, Subaru, RAM, Other)
Model Type	Ext. Warranty Miles	Lender
Purchase Date	**Sales Type** (New, Used)	**Purchase Type** (Cash, Finance, Lease)

Features

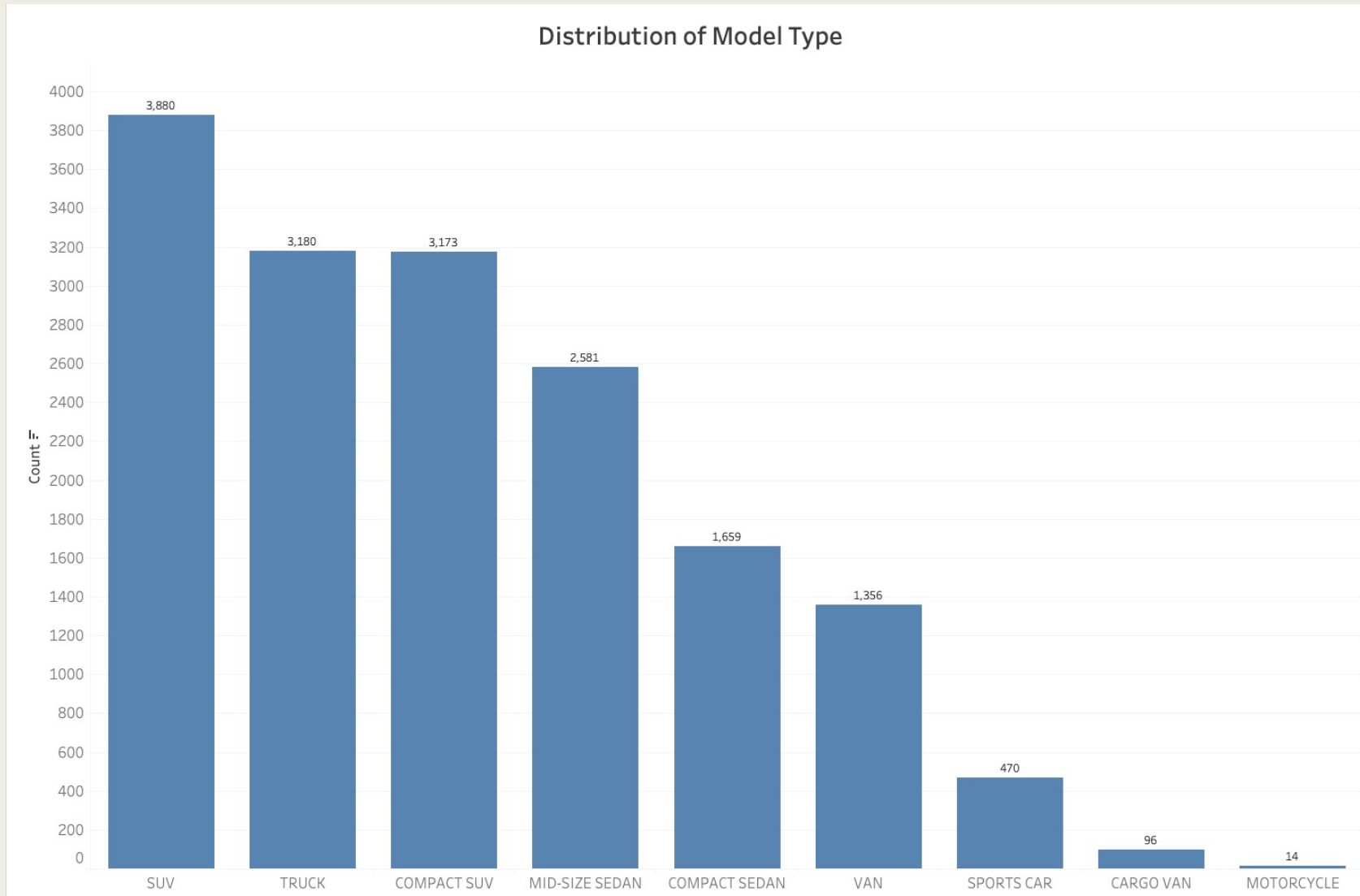|		
MSRP	Sale Price	~~Stock Number~~
~~VIN~~	Model Year	Make (Dodge, Kia, Subaru, RAM, Other)
Model Type	Ext. Warranty Miles	Lender
Purchase Date	Sales Type (New, Used)	Purchase Type (Cash, Finance, Lease)

Exploratory Data Analysis

Exploratory Data Analysis
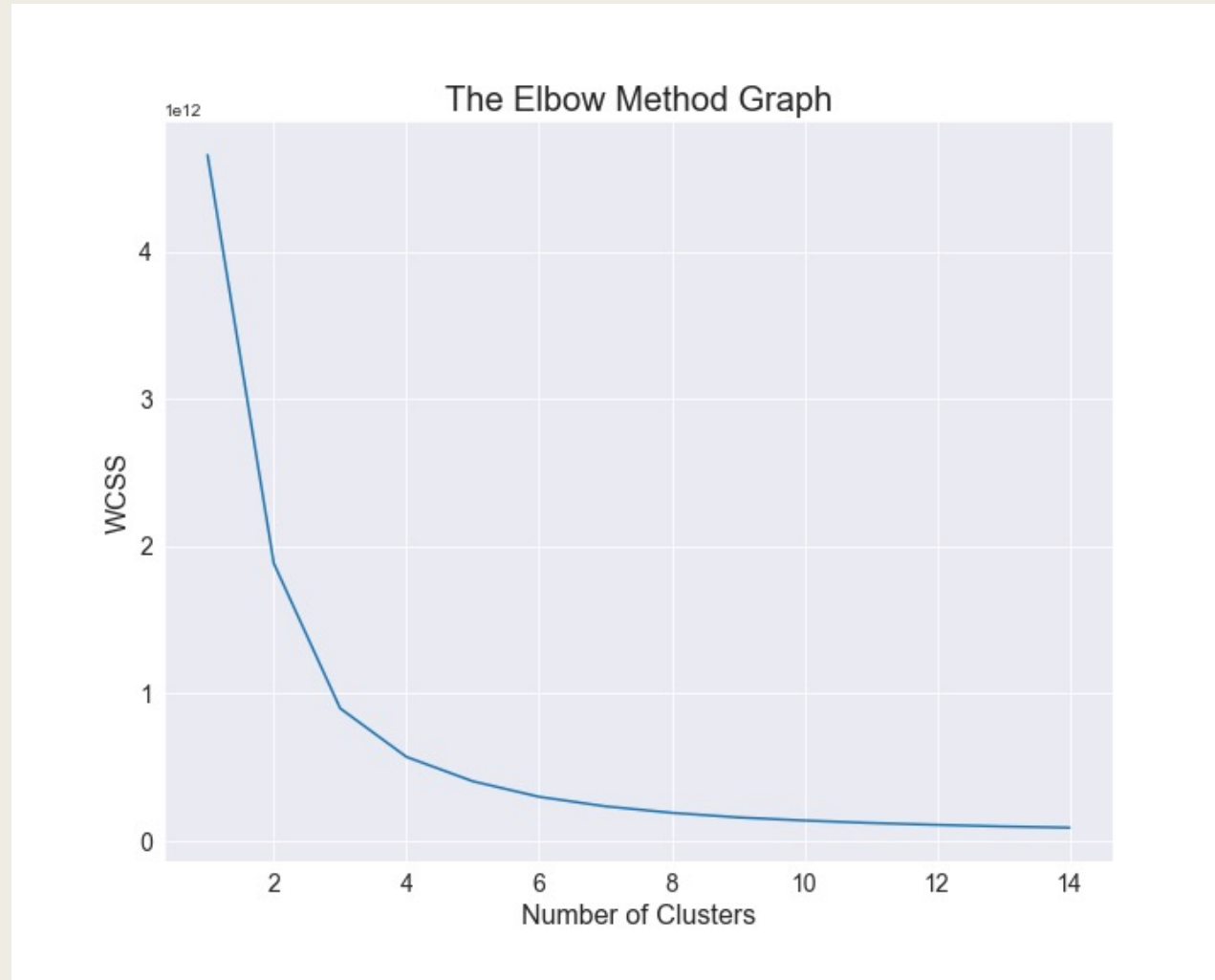

Distribution of Model Type

Data Cleaning

- Extended Warranty Miles

- Dodge, Kia, Subaru, RAM, Other

- 9 Model Types

- Purchase and Sales Type

Multi-Class Classification

ALGORITHM	COHEN KAPPA SCORE
Decision Tree	.36
KNN	.36
Naïve Bayes	.47

Customer Segmentation Analysis

The Elbow Method Graph

Types of Customers

Cluster	1	2	3	4
Sale Price	$35,175	$18,824	$15558	$27205
Cash	No	No	Yes	Split
Finance	Yes	Yes	No	Split
Lease	No	No	No	No
New	Yes	No	No	Yes
Used	No	Yes	Yes	No

PCA Visualization

Visualizing Clusters in Three Dimensions Using PCA
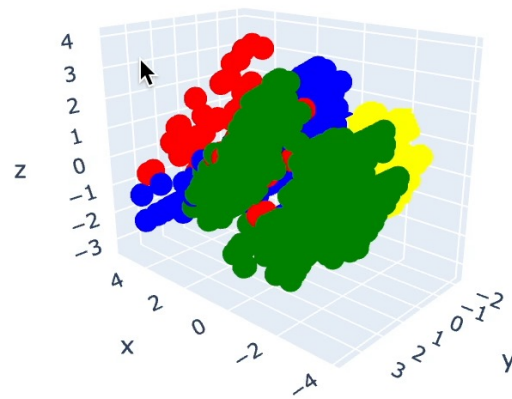
Next Steps

- Refine model types

- Add more car brands

- Add gas or hybrid

Contact

- [LinkedIn](#)

- [GitHub](#)

- [hkhan94@gmail.com](#)