Laptop Price Modeling

TIMOTHY LU

SPRINGBOARD CAPSTONE

JULY 2022

Capstone Goal:

Consumer demand for laptops is increasing

Laptop companies need to understand what features to focus on in order to optimize their pricing strategy

Having inappropriately price products can either cause decreased sales or missed profit

What features are consumers looking for in laptops?

What is an appropriate price point for laptops with certain technical specifications?

We will use a dataset of scraped laptop pricing data to answer these questions
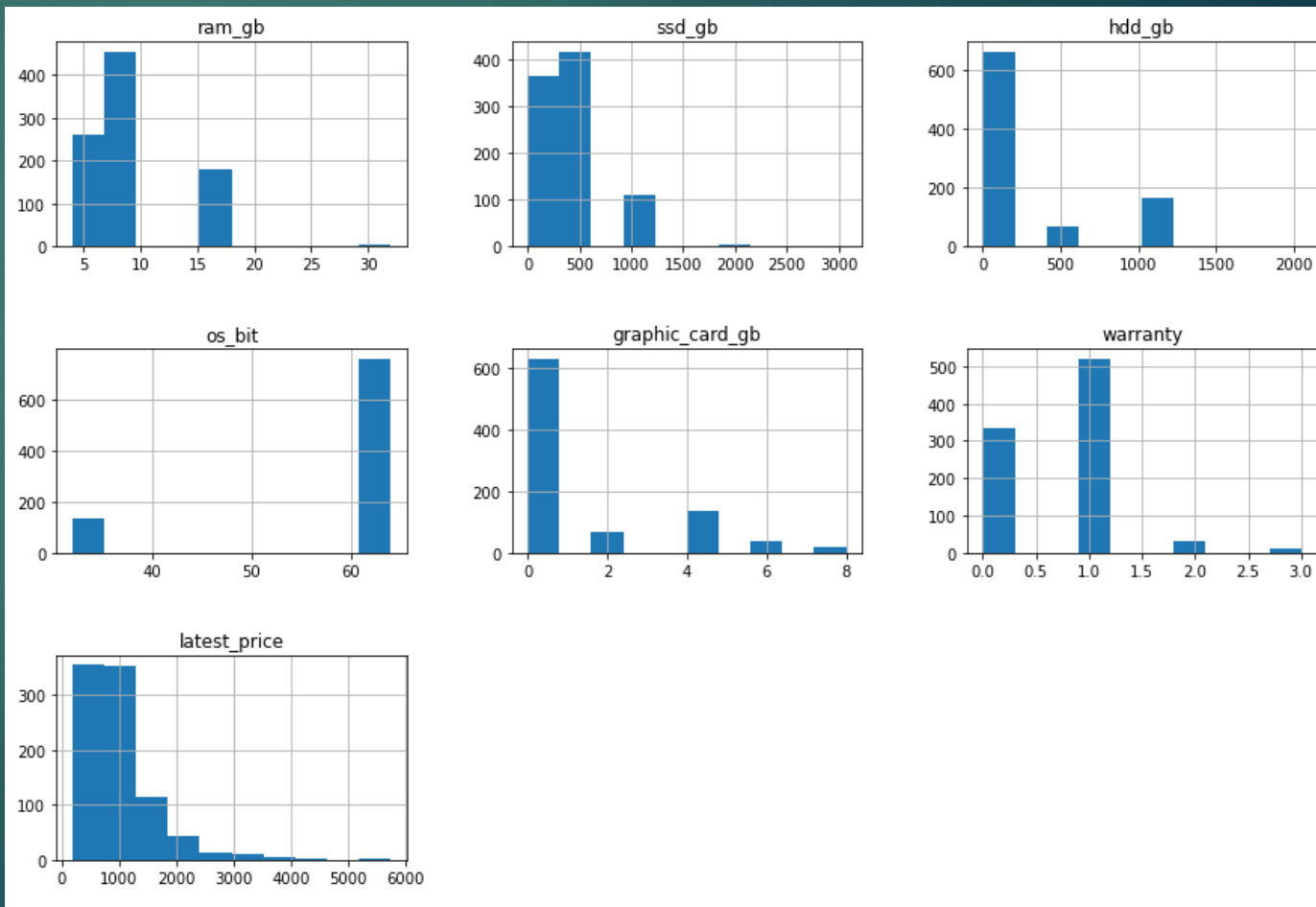
Interested Parties

Any company that sells laptops

Data Wrangling

▶ GOALS: Remove unnecessary features and tidy up some data

▶ Following columns unnecessary and dropped:

  ▶ old_price

  ▶ discount

  ▶ reviews

  ▶ star_rating

  ▶ ratings

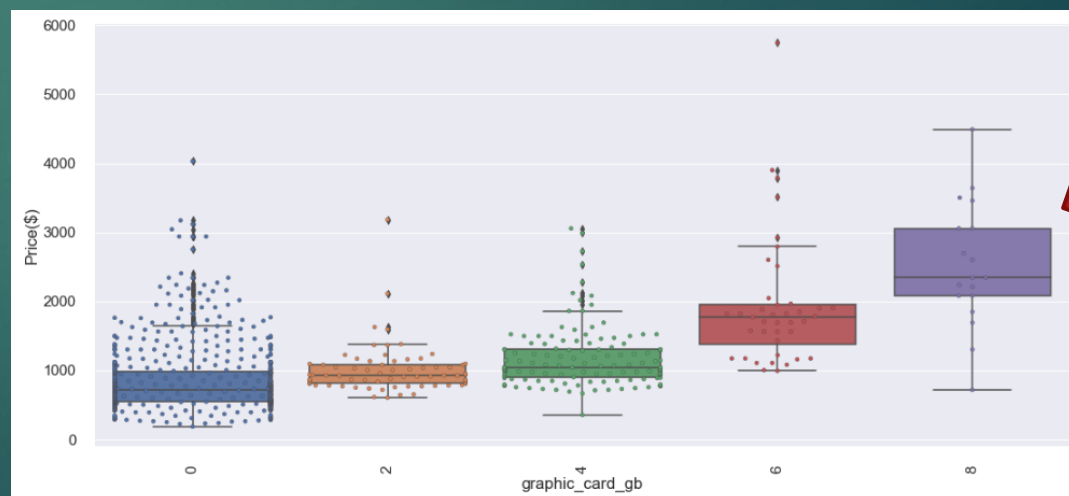Understanding data distributions

Exploratory Data Analysis
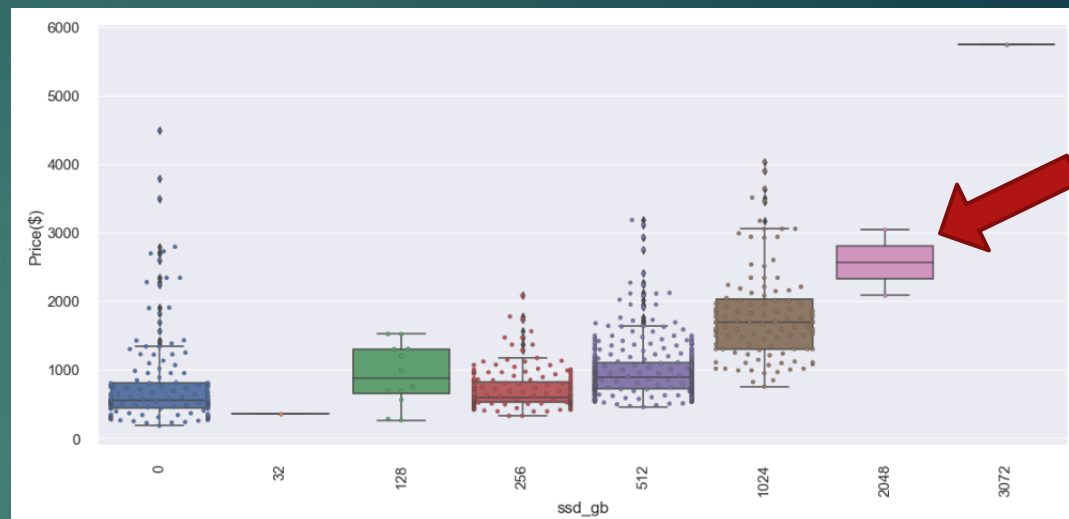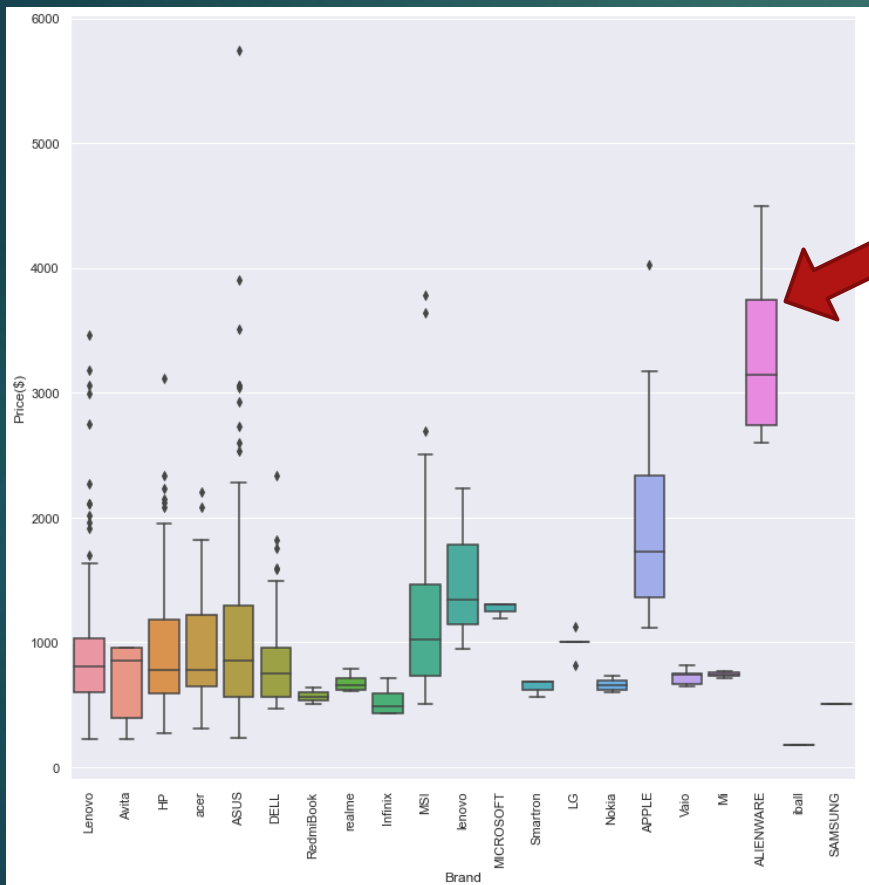
▶ GOALS: Understand our data and building our model hypothesis

▶ Primary Exploration - Median and Correlation
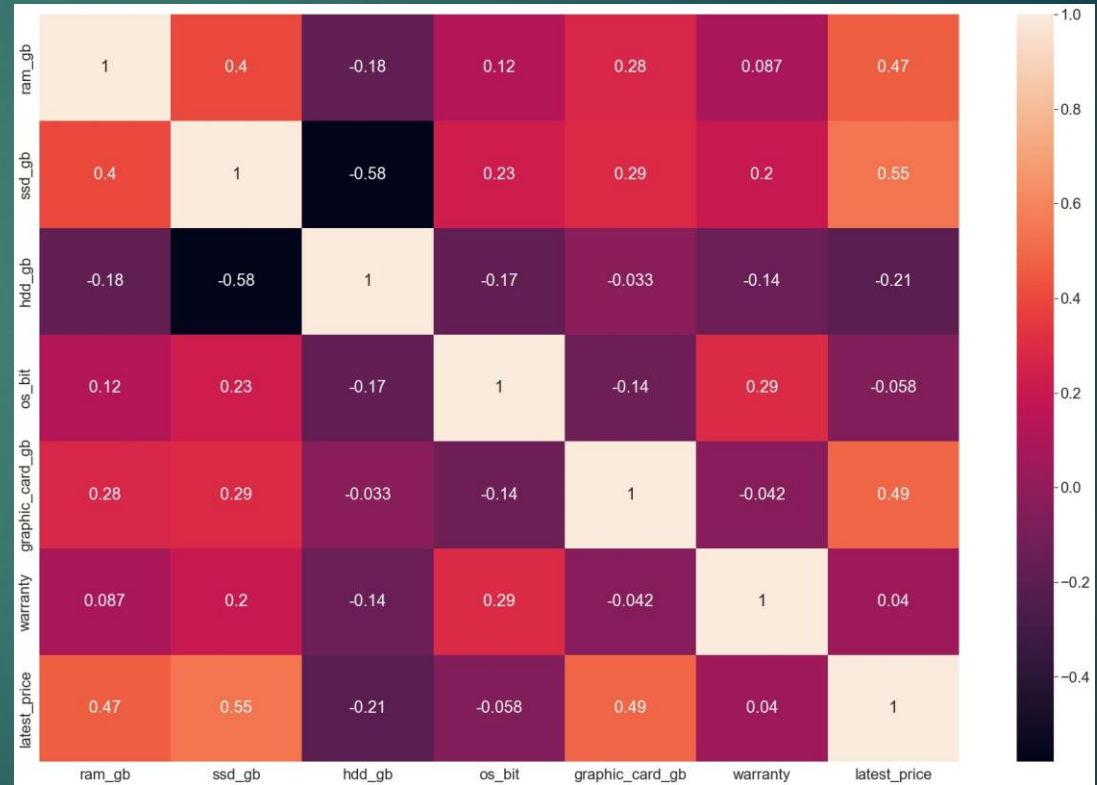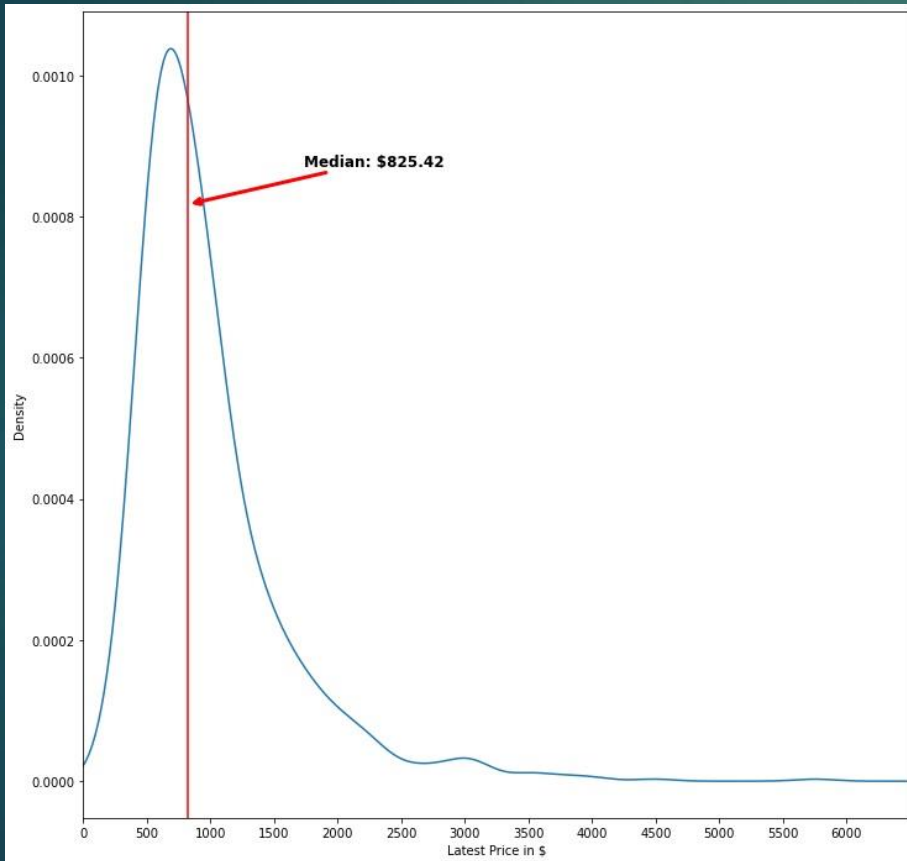
Features with correlations

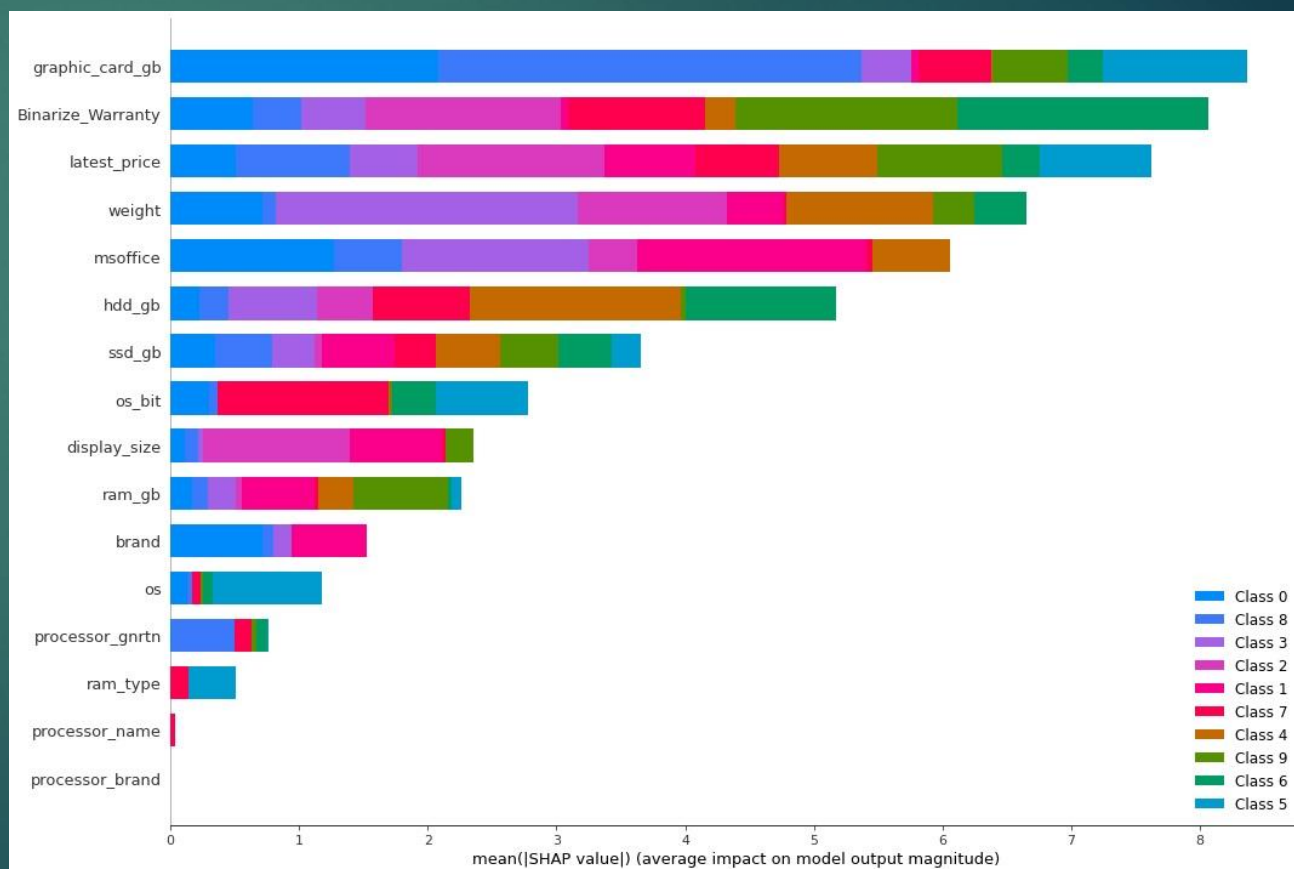# Exploratory Data Analysis

Box Plots show similarities and differences

Data Pre-Processing

- ▶ GOAL: Make data usable for models

- ▶ Remove -> "Model": Too sparse

- ▶ Impute -> display size and processor gen

- ▶ MinMaxScaler with our numerical data

RAM GB	SSD GB	HDD GB	Graphics Card GB	Display Size	Latest Price

- ▶ Dummy encode categorical data

| Brand	Processor Brand	Processor Name	Processor Generation	RAM Type	OS	OS BIT	Weight	MS Office	Binarize Warrant

Model 1 – K-Prototypes

- 12 Clusters seemed ideal
- SHAP Analysis helped me understand the clustering breakdown
- Cost of: 191.30
- F1 Score of: 0.79

Model 1 – K-Prototypes

- Assessing the clusters:

| latest_price	ram_gb	ssd_gb	hdd_gb	graphic_card_gb	display_size	brand	processor_name	processor_gnrtn	os_bit	weight	msoffice	Binarize_Warranty
489.0	5.0	180.0	414.0	0.0	15.0	ASUS	Core i3	11th	64	Casual	No	0
579.0	6.0	121.0	1051.0	0.0	15.0	DELL	Core i3	11th	64	ThinNlight	Yes	1
728.0	7.0	387.0	21.0	0.0	15.0	Lenovo	Core i3	10th	64	ThinNlight	Yes	1
830.0	9.0	446.0	0.0	0.0	15.0	Avita	Core i5	11th	64	ThinNlight	No	0
865.0	8.0	485.0	12.0	0.0	16.0	HP	Core i5	11th	64	ThinNlight	Yes	1
884.0	9.0	291.0	588.0	1.0	15.0	ASUS	Core i5	11th	64	Casual	No	1
1106.0	9.0	484.0	194.0	2.0	16.0	ASUS	Core i5	10th	64	Casual	No	0
1263.0	14.0	627.0	0.0	1.0	15.0	HP	Core i7	11th	64	Casual	Yes	1
1475.0	7.0	26.0	486.0	4.0	16.0	acer	Core i5	10th	32	Casual	No	0
1597.0	15.0	745.0	84.0	5.0	16.0	MSI	Core i7	11th	64	Casual	No	1

Model 2 – Isomapping

▶ Ultimately, difficult to discern

▶ Overlapping components without a way to understand the eigenvalues
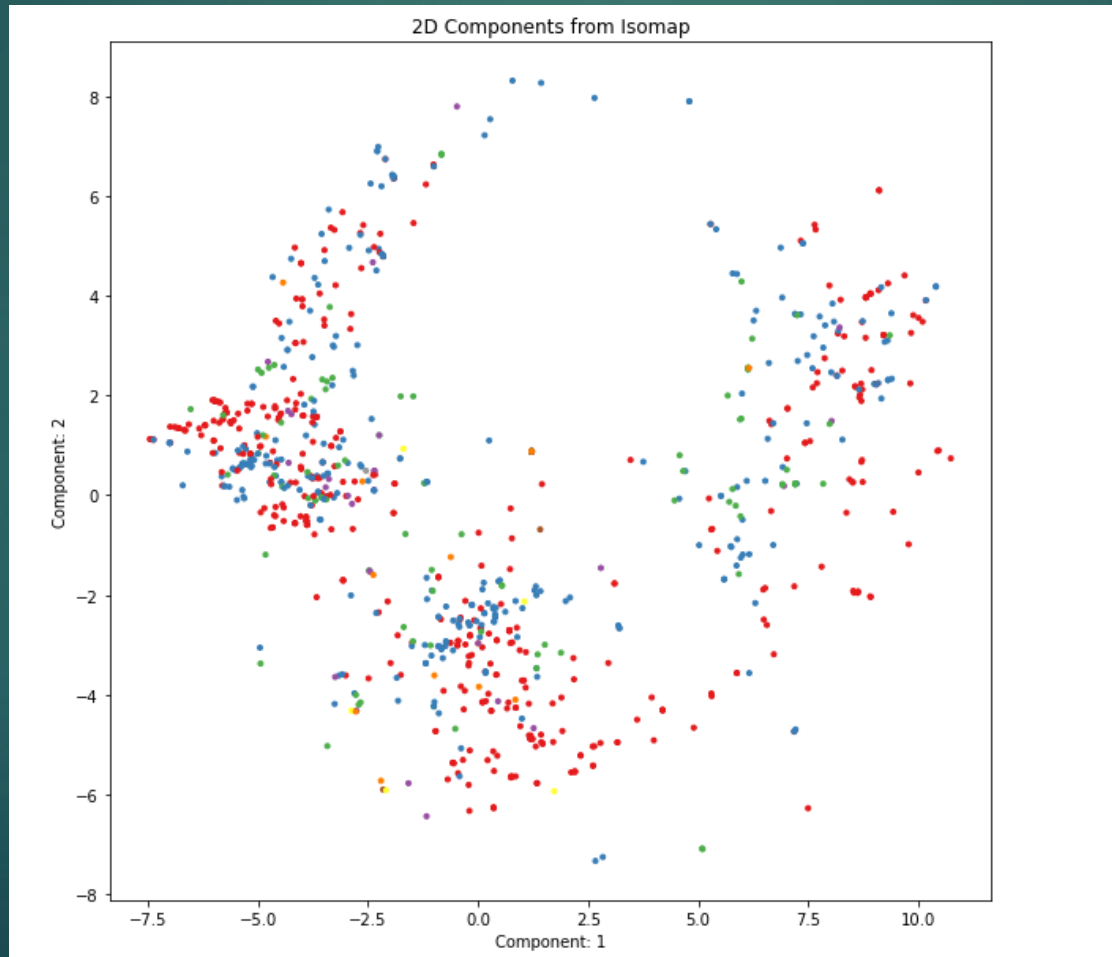
Model 3 – XGBoost

▶ Great for understanding classification

▶ Gave good understanding of features

▶ Helped show important features for classification model
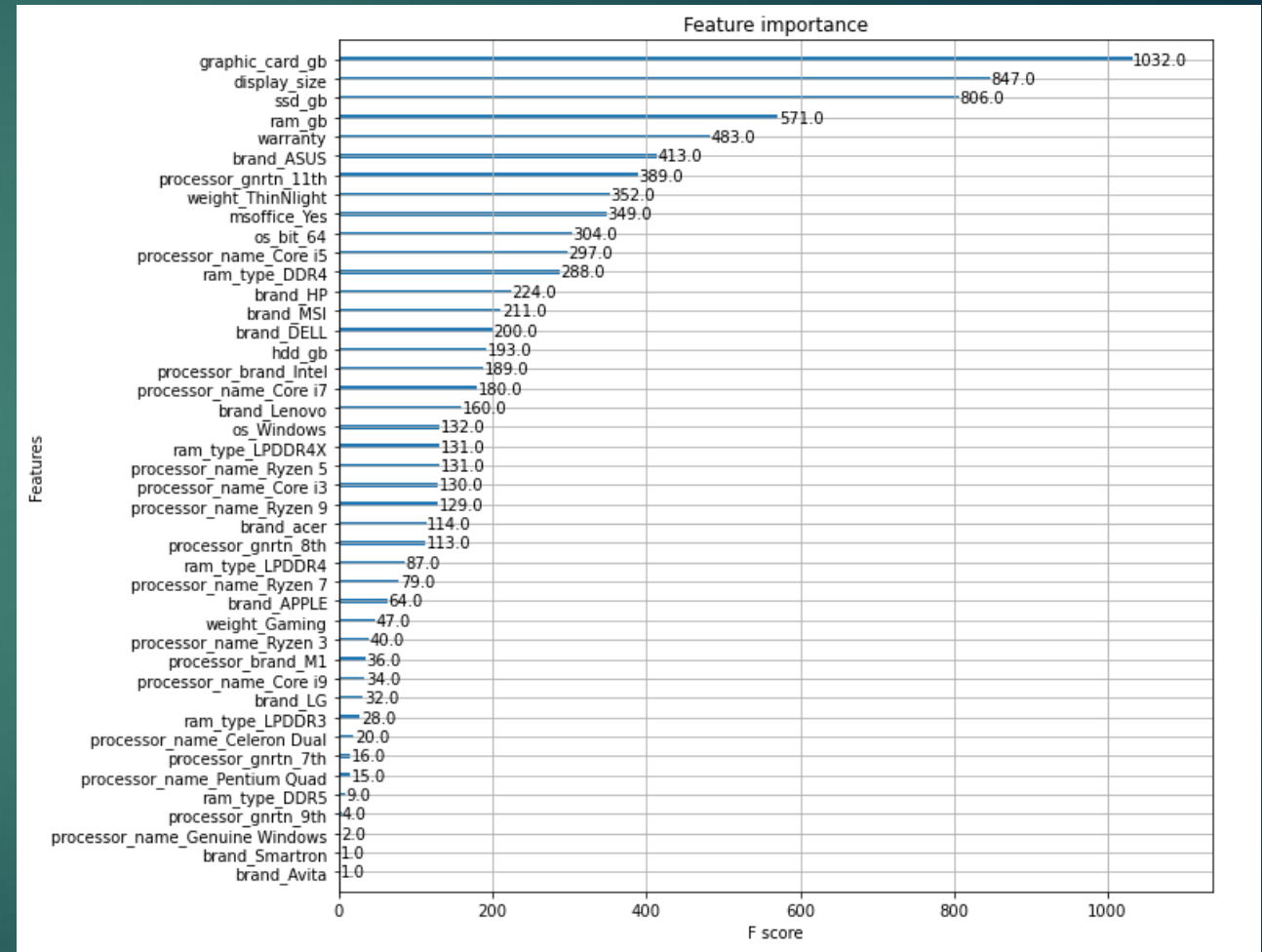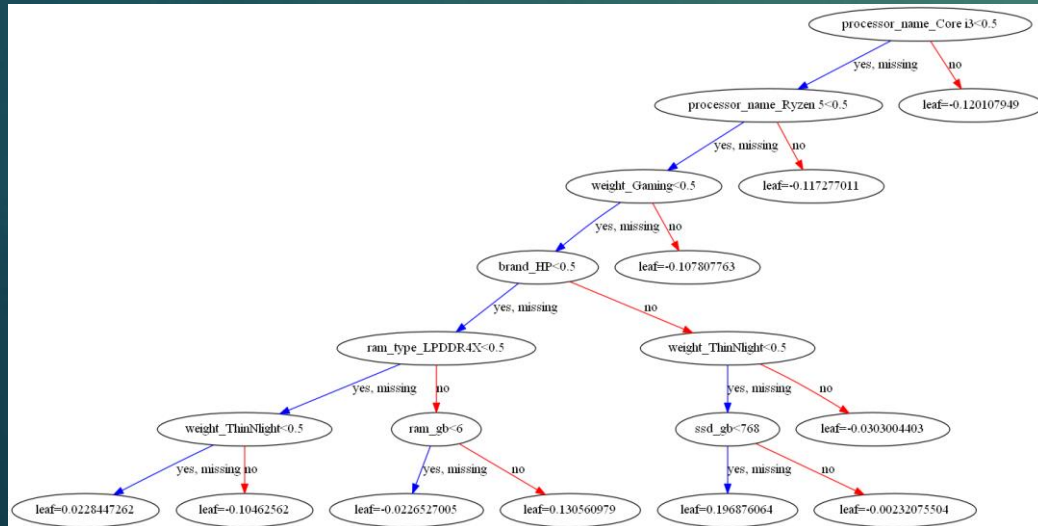
Model Decision – K-prototypes w/ XGBoost Knowledge

- K-Prototypes gives easy to understand clusters
- XGBoost gives great delineation between certain features
  - e.g if SSD above 770GB we can be confident the price should be in a higher classification group
- No one model was immediately the best

Recommended Action

▶ Use the K-protoypes clusters as a starting point for pricing

▶ Use XGBoost tree for understanding exactly which price range it should fall in -> can label a laptop based on inputted specifications

▶ Focus on developing graphics card memory size, SSD size, display size, and providing warranties

▶ These are the main ways to optimize our pricing strategy

Constraints and Limitations

▶ Initially missing a lot of data (almost 33%)

▶ Assuming that all features are weighed the same

▶ Relatively small dataset compared to the volume of laptops available

▶ Potentially could become outdated without updated laptop information

Further Study

- Gather more laptop data

- Improve the decision tree modeling to increase accuracy

- Get fewer missing data points with less imputing to minimize bias

- Gather consumer data and apply weighting to our data points

- Combine with NLP to understand what consumers are saying about laptops and what they feel are important features

Thank you

TIMOTHY LU

PROJECT REPORT: HTTPS://GITHUB.COM/NATURESBLESS/LAPTOPPRICEMODELING/BLOB/MAIN/REPORTS%20AND%20METRICS/LAPTOP_PRICE_MODELING_CAPSTONE_REPORT.PDF

PROJECT GITHUB: HTTPS://GITHUB.COM/NATURESBLESS/LAPTOPPRICEMODELING