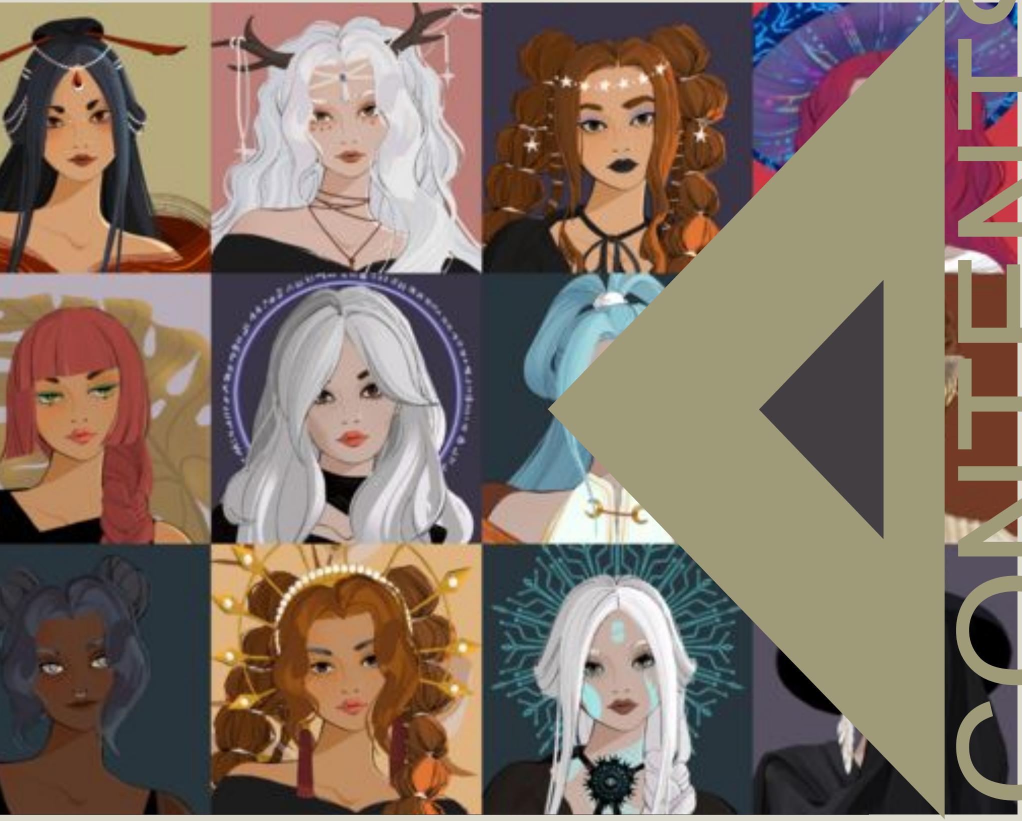
IDS 702 Team Project

Crypto Coven Analysis

Team Green

Dany Jabban, Scott Lai, Zhonglin Wang

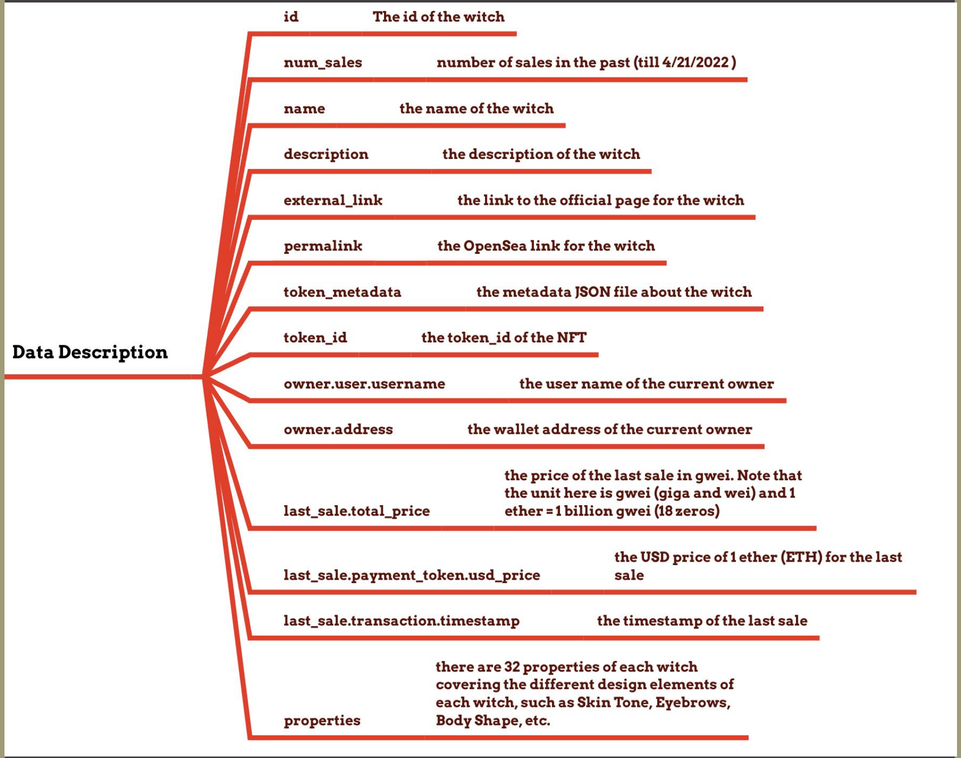


CONTENTS

Background

“NFT stands for Non-fungible token. “Non-fungible” more or less means that it’s unique and can’t be replaced with something else. Most NFTs are made of digital art, and that art contains many pieces with different rarity, which will decide the selling price of the NFT. Usually, the rarer, the more expensive, but we still find people collect those NFTs in their interest. Although 100% accurate technical analysis for NFT cryptocurrency is hardly possible, we are trying to find what factors could affect the NFT price and sell status.”

Data Description



About Dataset

Crypto Coven

We get this dataset from Kaggle. The dataset contains information about the 9761 witches picture from the Crypto Coven NFT project collected using OpenSea API. The full data returned by the API is provided in witches_full.csv and a subset of the data is chosen by team and shared in witches.csv. We show all the variables in the left.

Our research mainly focuses on properties, which is the decide variable for the NFT price. Each properties represent one of the feature of the NFT. The properties include hair color, hat, necklace, etc.

Research Questions

1. Does rarity have an affect on sale likelihood

We are particularly interested in what kind of NFT people are willing to buy. There are 9761 Coven NFTs in the dataset, and most didn't make the sale. We will focus on the research of the selling group and non-sell group to see if particular properties made them sell more than others.

2. Does the rarity of the properties increasing the sell price.

With the sell price data, we can find out the different listing price of Coven NFTs, find out if there is any connection between the rarity, sell action, and NFT prices.

Logistic Regression

Linear Regression

Variable Structure

Base Structure Factors

- These are features every Crypto Coven NFT has
- Examples: skin tone, eyebrows, hair color, etc.
- These are treated as separate categorical predictors

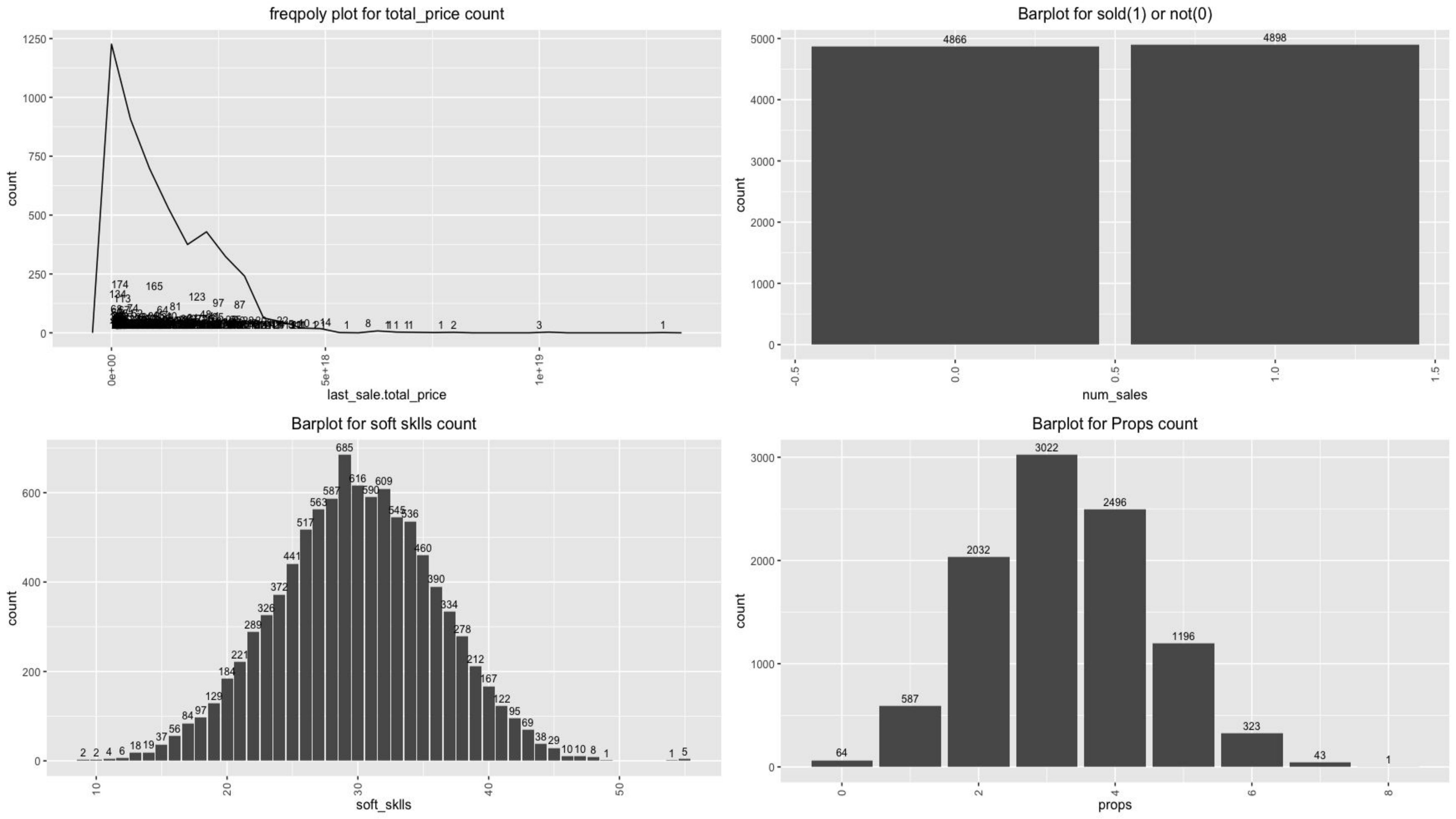
Soft Skills

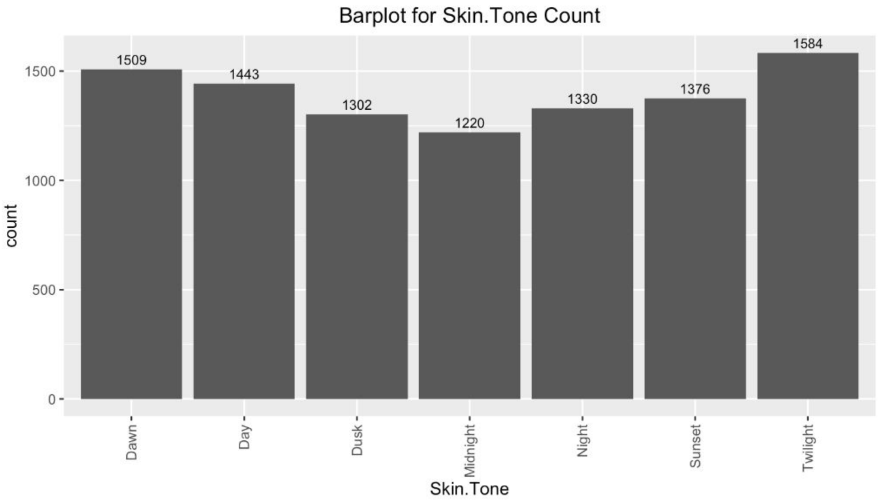
- These are ordinal categorical variables ranging from 0-10
- examples : wonder, wisdom, wit, etc
- Summing soft skills together into one predictor - new scale from 0-60

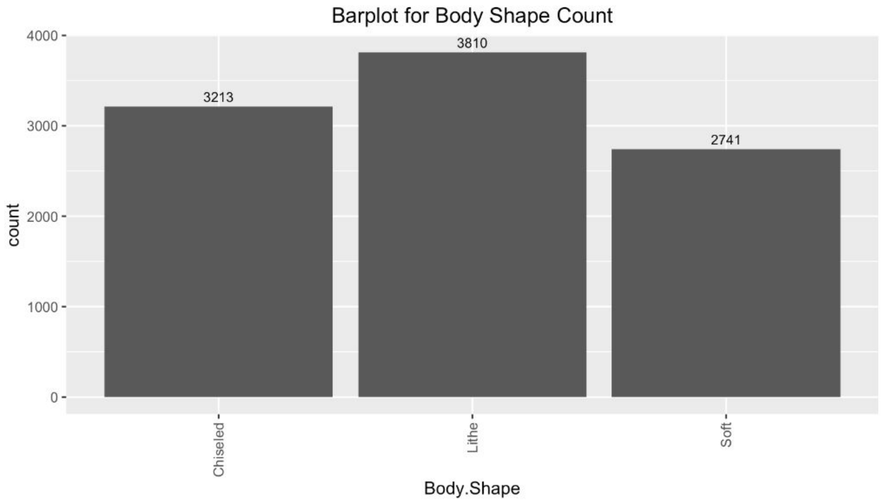
Props Existence Level

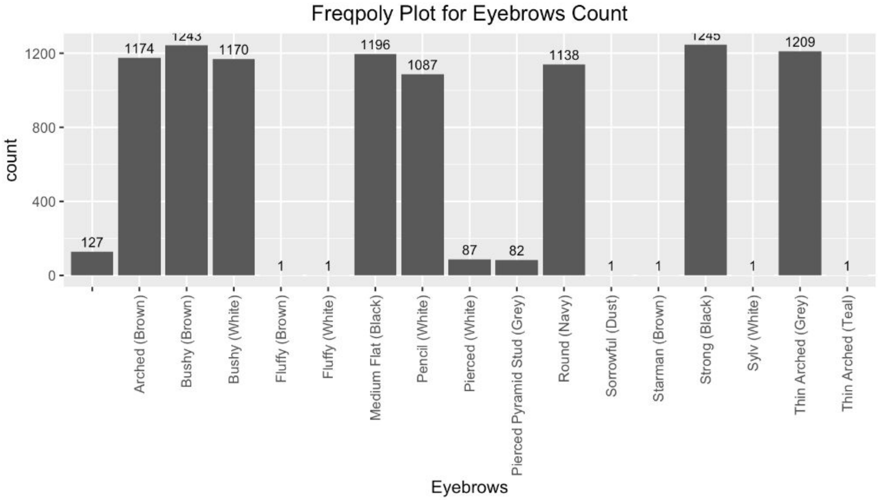
- These are binary features some NFTs have them some do not
- Examples: faceware, earrings, necklace, etc.
- Summing props together into one predictor - new scale from 0-12

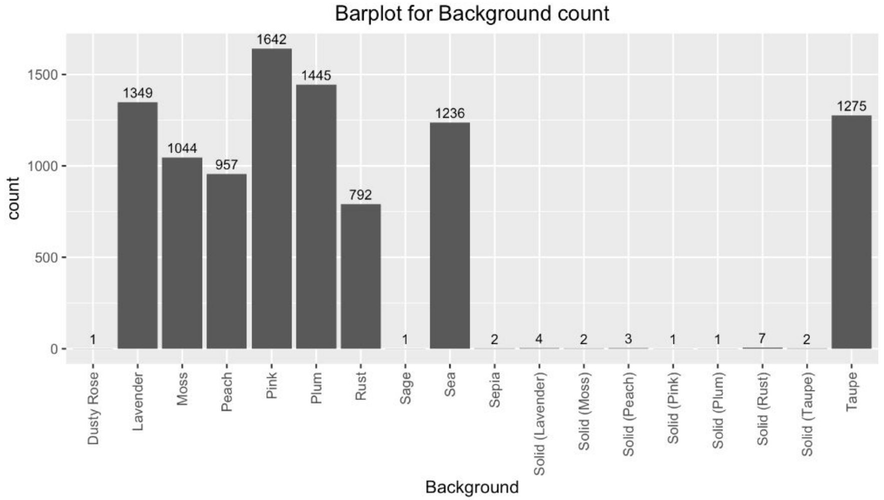
Exploratory Data Analysis


freqpoly plot for total_price count
count
last_sale.total_price
0e+00
5e+18
1e+19
Barplot for sold(1) or not(0)
4866
4898
count
num_sales
Barplot for soft sklls count
count
soft_sklls
685
616
609
590
587
563
545
536
517
460
441
390
372
334
326
289
278
221
212
184
167
129
122
97
95
84
69
56
38
37
29
19
18
10
10
8
6
4
2
2
1
1
5
Barplot for Props count
count
props
3022
2496
2032
1196
587
323
64
43
1


Barplot for Skin.Tone Count
count
1509
1443
1302
1220
1330
1376
1584
Dawn
Day
Dusk
Midnight
Night
Sunset
Twilight
Skin.Tone


Barplot for Body Shape Count
count
3213
3810
2741
Chiseled
Lithe
Soft
Body.Shape


Freqpoly Plot for Eyebrows Count
count
127
1174
1243
1170
1
1
1196
1087
87
82
1138
1
1
1245
1
1209
1
Arched (Brown)
Bushy (Brown)
Bushy (White)
Fluffy (Brown)
Fluffy (White)
Medium Flat (Black)
Pencil (White)
Pierced (White)
Pierced Pyramid Stud (Grey)
Round (Navy)
Sorrowful (Dust)
Starman (Brown)
Strong (Black)
Sylv (White)
Thin Arched (Grey)
Thin Arched (Teal)
Eyebrows


Barplot for Background count
count
1
1349
1044
957
1642
1445
792
1
1236
2
4
2
3
1
1
7
2
1275
Dusty Rose
Lavender
Moss
Peach
Pink
Plum
Rust
Sage
Sea
Sepia
Solid (Lavender)
Solid (Moss)
Solid (Peach)
Solid (Pink)
Solid (Plum)
Solid (Rust)
Solid (Taupe)
Taupe
Background

Question 1

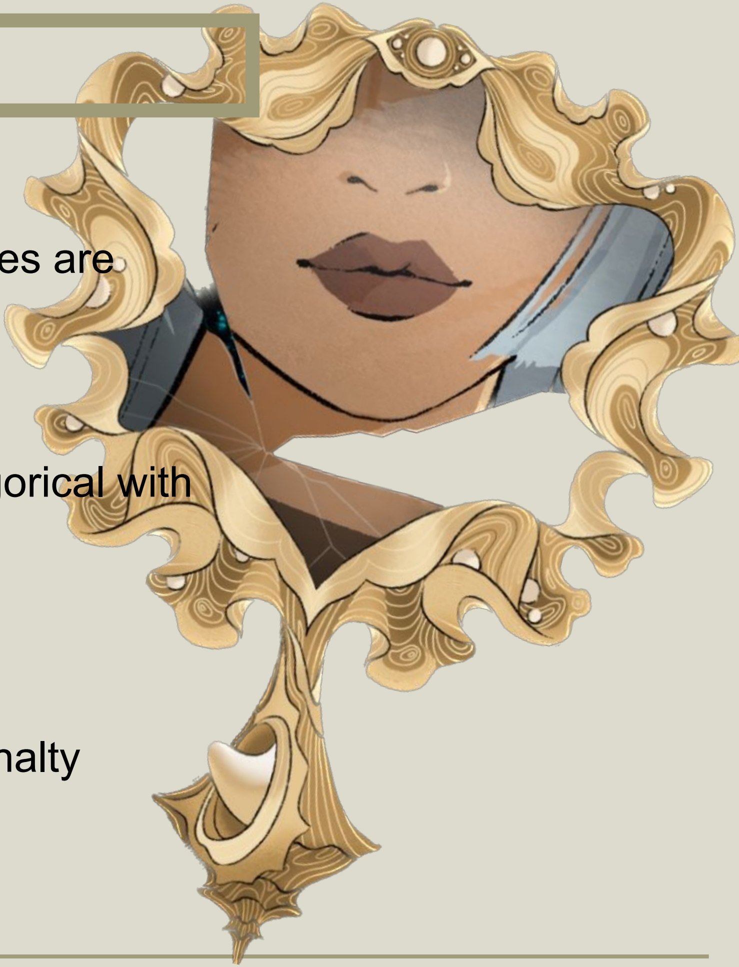
What features are associated with the increasing of odds of Sale

Odds of a Sale

- Logistic regression model to determine what features are associated with increase odds of a sale

- Still had 16 predictors with the majority being categorical with over 10 subcategories
  - Some had up to 40 subcategories

- Used stepwise selection with AIC criterion
  - Want to use a medium strength parameter penalty

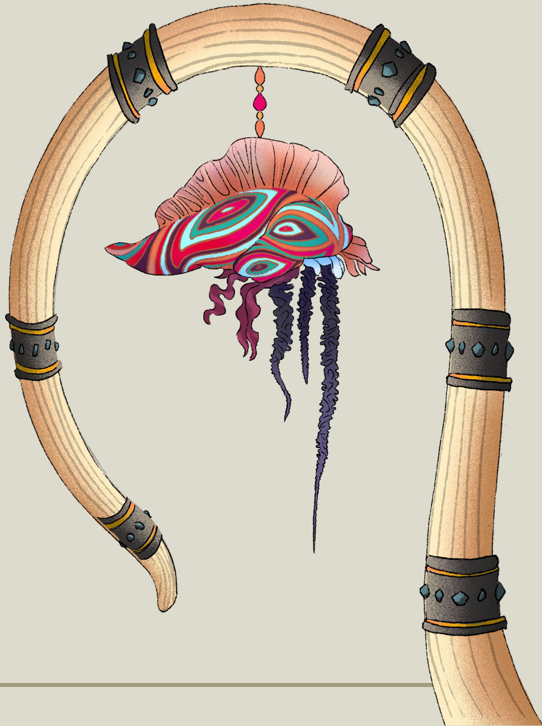


Odds of a Sale

- Our final model had 2 features with only a single subcategory being statistically significant

Logistic Regression for Sale Odds

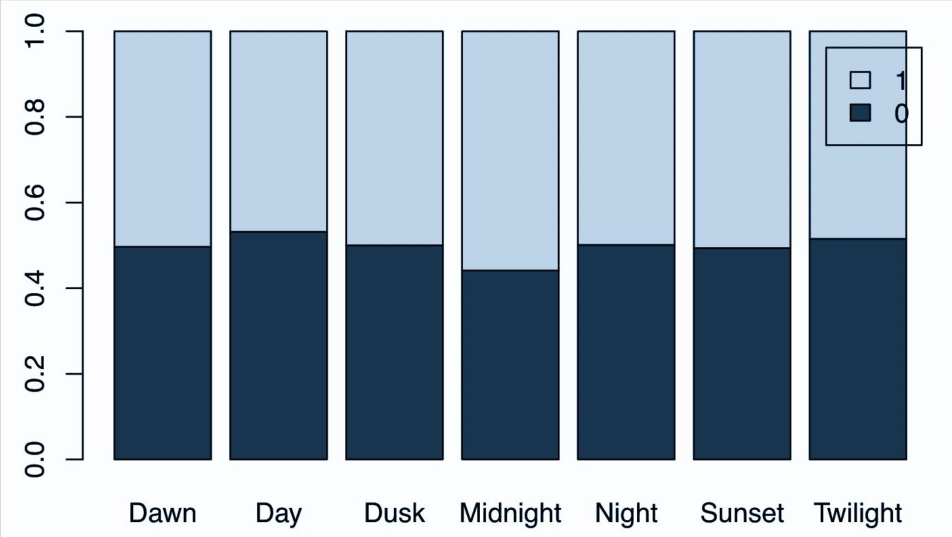
	Estimate	SE	t	p-value
Intercept	0.022	0.0589	0.3732	0.710
Skin Tone: Dawn	-	-	-	-
Skin Tone: Day	-0.1431	0.0738	-1.9402	0.053*
Skin Tone: Dusk	-0.021	0.0757	-0.2774	0.782
Skin Tone: Midnight	0.2128	0.0774	2.7493	<0.001***
Skin Tone: Night	-0.0268	0.0753	-0.3565	0.722
Skin Tone: Sunset	0.0035	0.0746	0.0463	0.964
Skin Tone: Twilight	-0.0758	0.072	-1.0529	0.293
Body Shape: Chiseled	-	-	-	-
Body Shape: Lithe	-0.0732	0.048	-1.5253	0.128
Body Shape: Soft	0.0929	0.0521	1.7819	0.075*



Odds of a Sale

Body Shape vs Sale Proportion

Skin Tone vs Sale Proportion

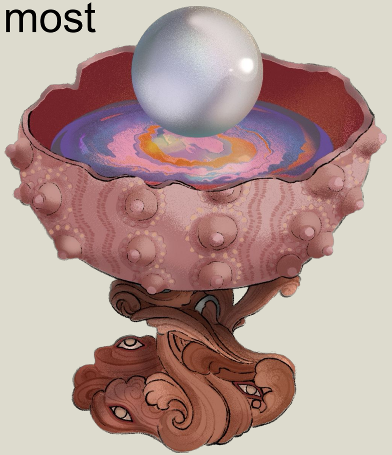


Odds of a Sale: Conclusion

- Midnight skin tone had biggest increase in odds of sale
  Day skin tone had biggest decrease in odds of sale

- Most informative predictors are the facial features that stand out the most
    - Not the smaller intricate detailed differences

- This was actually counter to what Scott originally hypothesized



Question 2

What features associated with the increase in NFT’s last sale total price.

Increase in NFT’s Price

- Linear regression model to determine what features are associated with the increase in NFT’s last sale total price.

- Used stepwise selection with AIC criterion
  - Want to use a medium strength parameter penalty

- Question - **High value NFTs features?**

Increase in Price

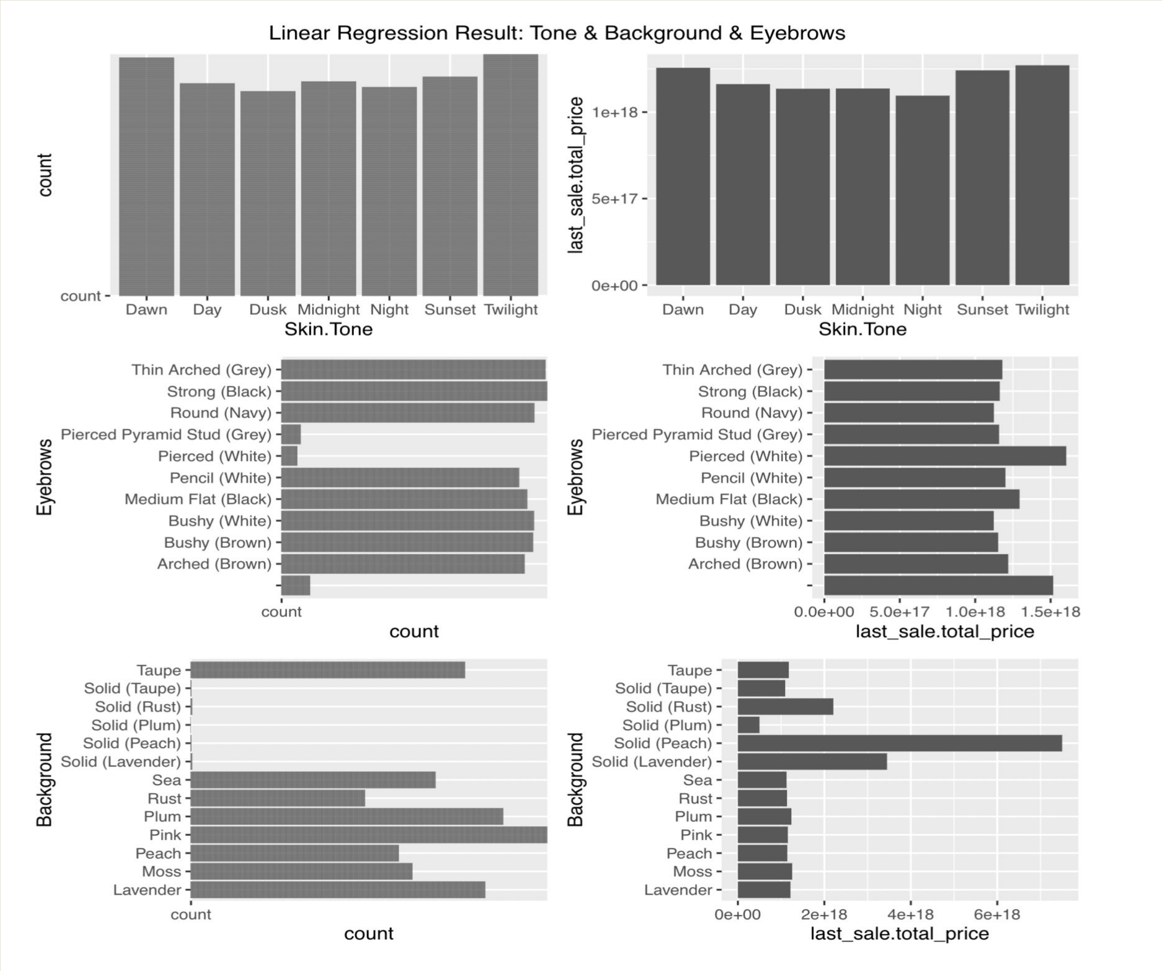
- The final model (with step-AIC selection) had 3 features with subcategories being statistically significant

Table 1: Summary Table of coefficients for the Regression Model

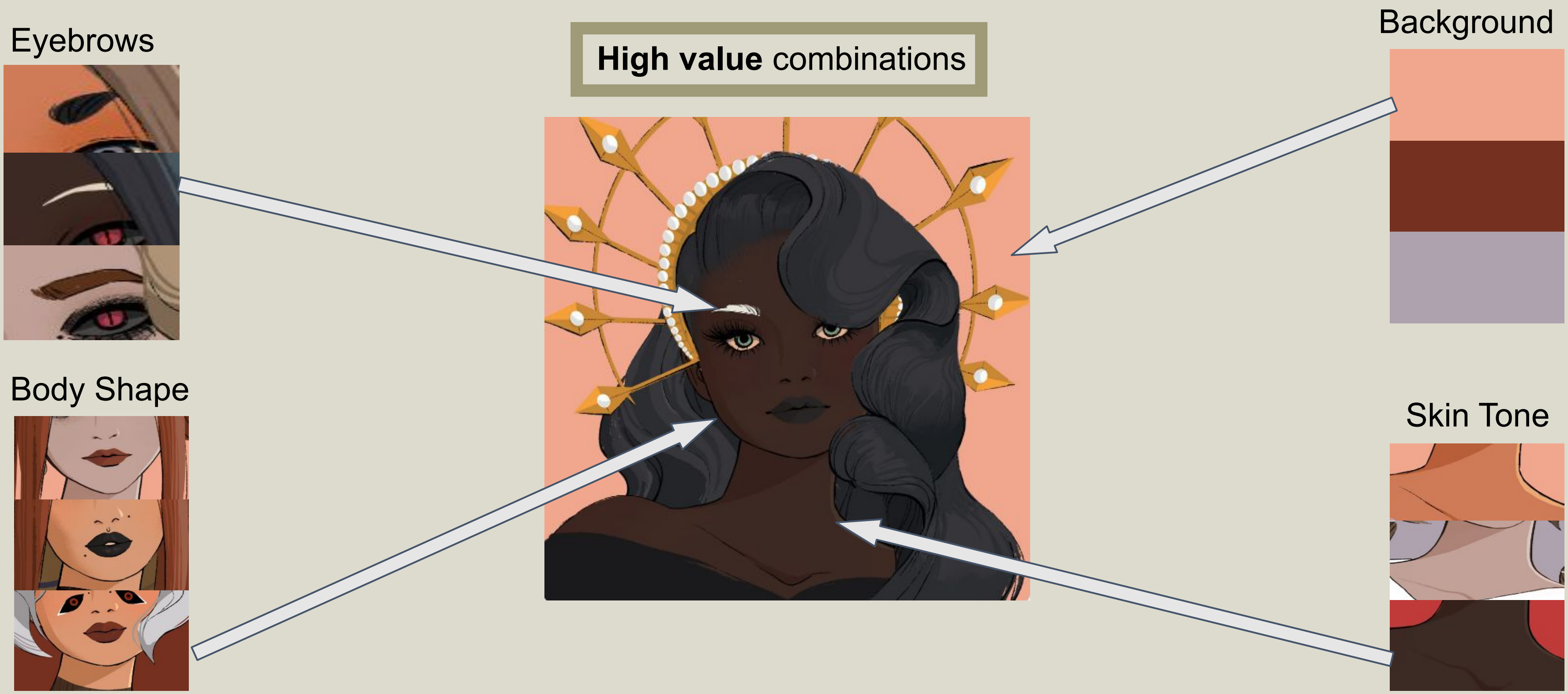
	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.593e+18	1.435e+17	11.105	< 2e-16	***
Skin.ToneDay	-7.382e+16	5.848e+16	-1.262	0.20693	
Skin.ToneDusk	-9.998e+16	5.913e+16	-1.691	0.09092	.
Skin.ToneMidnight	-9.725e+16	5.836e+16	-1.666	0.09570	.
Skin.ToneNight	-1.471e+17	5.873e+16	-2.504	0.01232	*
Skin.ToneSunset	4.707e+15	5.798e+16	0.081	0.93530	
Skin.ToneTwilight	2.394e+16	5.660e+16	0.423	0.67228	
EyebrowsArched (Brown)	-2.952e+17	1.418e+17	-2.081	0.03747	*
EyebrowsBushy (Brown)	-3.645e+17	1.416e+17	-2.574	0.01008	*
EyebrowsBushy (White)	-3.967e+17	1.416e+17	-2.802	0.00510	**
EyebrowsMedium Flat (Black)	-2.409e+17	1.417e+17	-1.699	0.08930	.
EyebrowsPencil (White)	-3.167e+17	1.420e+17	-2.230	0.02581	*
EyebrowsPierced (White)	1.099e+17	2.236e+17	0.491	0.62310	
EyebrowsPierced Pyramid Stud (Grey)	-3.440e+17	2.108e+17	-1.632	0.10271	
EyebrowsRound (Navy)	-3.924e+17	1.416e+17	-2.772	0.00559	**
EyebrowsStrong (Black)	-3.553e+17	1.413e+17	-2.515	0.01193	*
EyebrowsThin Arched (Grey)	-3.347e+17	1.411e+17	-2.371	0.01777	*
BackgroundMoss	5.585e+16	6.429e+16	0.869	0.38505	
BackgroundPeach	-7.066e+16	6.536e+16	-1.081	0.27972	
BackgroundPink	-4.772e+16	5.687e+16	-0.839	0.40144	
BackgroundPlum	2.122e+16	5.859e+16	0.362	0.71729	
BackgroundRust	-8.114e+16	6.906e+16	-1.175	0.24012	
BackgroundSea	-8.616e+16	6.246e+16	-1.379	0.16782	
BackgroundSolid (Lavender)	2.237e+18	5.533e+17	4.043	5.36e-05	***
BackgroundSolid (Peach)	6.214e+18	7.814e+17	7.952	2.26e-15	***
BackgroundSolid (Plum)	-6.508e+17	1.105e+18	-0.589	0.55576	
BackgroundSolid (Rust)	1.003e+18	5.531e+17	1.814	0.06970	.
BackgroundSolid (Taupe)	-1.242e+17	7.820e+17	-0.159	0.87380	
BackgroundTaupe	-3.051e+16	6.064e+16	-0.503	0.61494	

Increase in Price

Quantity vs Sale Price

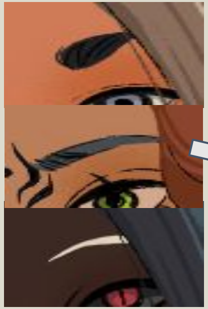


High value combinations



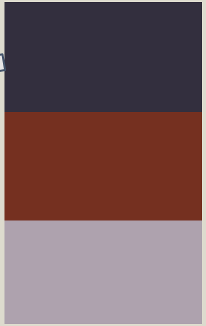
Low value combinations

Eyebrows

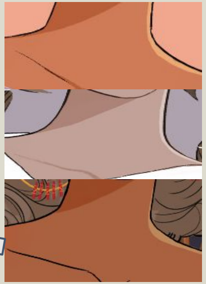


Body Shape

Background



Skin Tone



Increase in Price: Conclusion

- Dawn skin tone corresponds to relatively high prices.
  Solid (Lavender) and Solid (Peach) backgrounds correspond to high prices.

- Night skin tone corresponds to relatively low prices.
  Thin Arched (Grey), Pencil (white), Round (Navy), Strong(Black), Bushy (white), Bushy(Brown), Arched (Brown) correspond to low prices.

Conclusion

- From the result above, we can find out that Skin tone and Body Shape are the most significant variables for NFT sale.

- After people make the buy decision, the factor of rarity is joined to affect the NFT price change. The rarer the elements, the higher the price. For example the eyebrow and background we seen from before.

The evidence show that most people buy NFTs with their first impression of preference for the NFTs. The appearance features are associated with buying or not buying, and the properties' rarity affects the buying price.

Shortage

- We might lost some of the detailed affections by add up variables.
- Data not up to date.

Challenge

- NFT Diversities.
- Buying behavior might affect buy the auction price.

Future Plan

- Continue researching the buying behavior affections.
- Partially predicting the NFTs' price with the model.

Thank You

Questions?