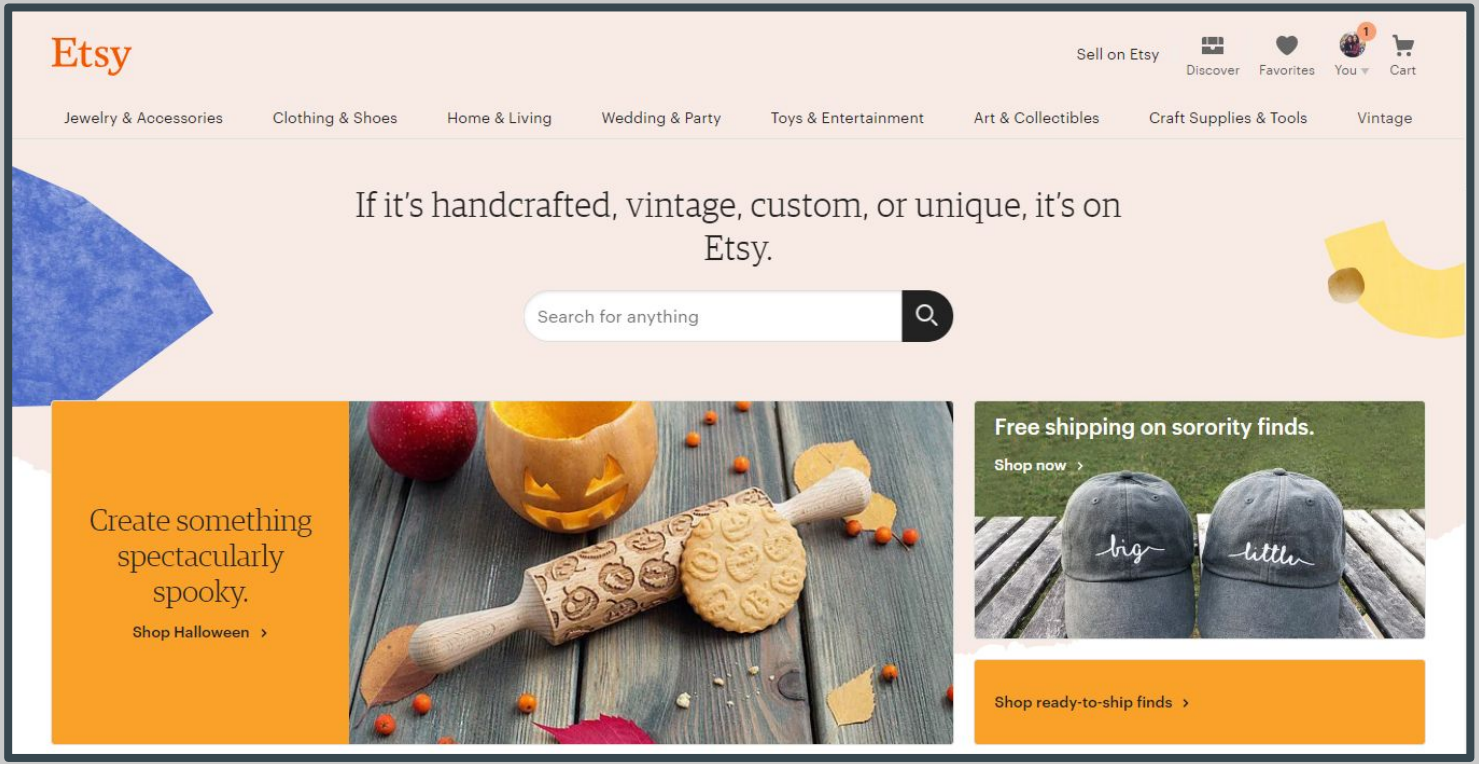
Thinkful Final Capstone:
“Etsy Epiphany”

•••

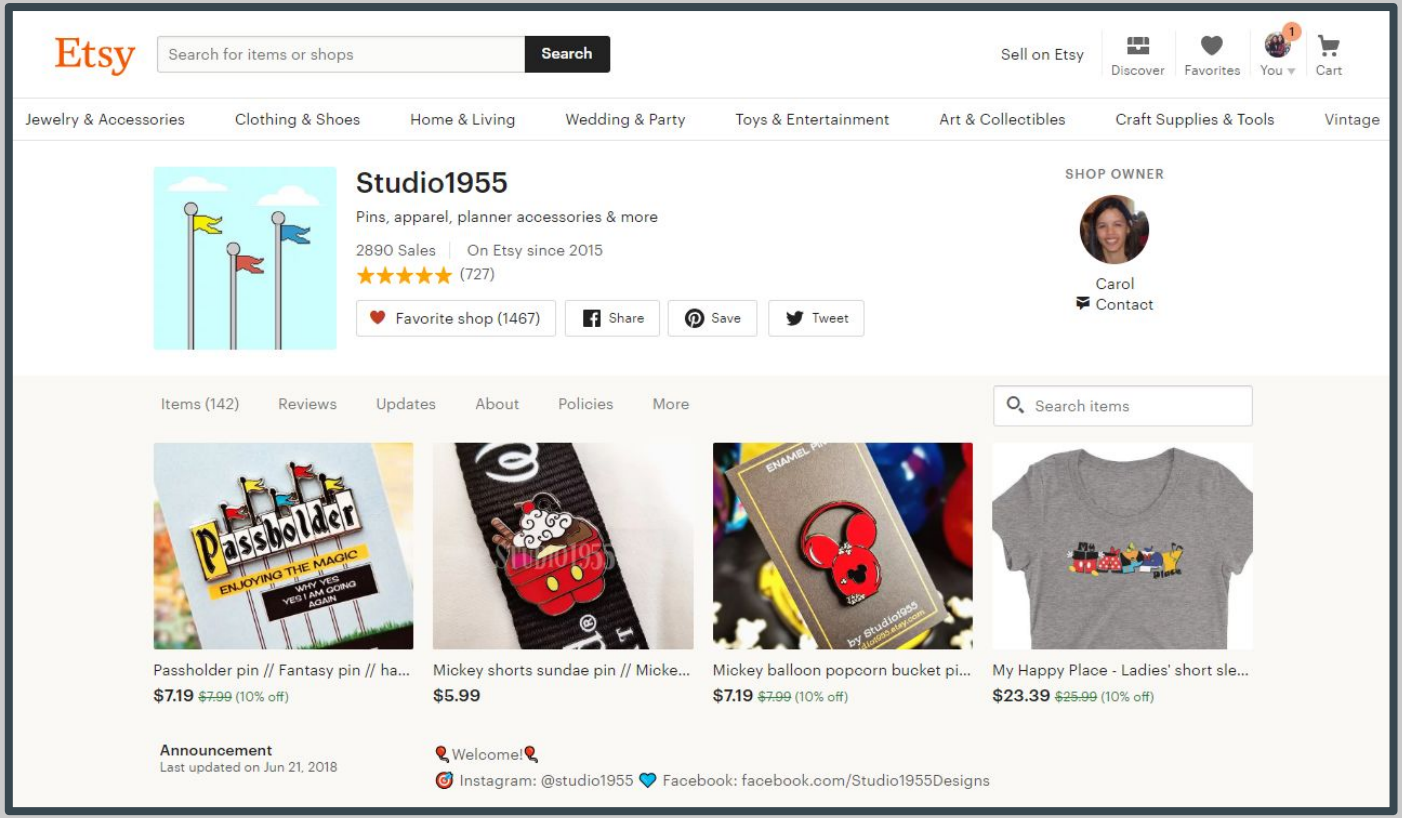
Tiani Calip

October 2018

Overview - What is Etsy?



Overview - My Personal Interest



Project Objective:

See how her products are doing compared to similar items on Etsy, and be able to respond to the demand for certain types of popular items and create and sell her own.

Providing review mapping and sentiment analysis will allow her and other users to identify what qualities of their products are favorable or not, be able to address those during production, and thereby possibly increase their chances of getting favorited.

Target Audience

- Etsy Store Owners
  - Don’t know what to produce next

The Competition

- “Whatify”- analyzes pictures
- “Marmalead”- analyzes keywords
- **None** look at data containing reviews, sentiments, and favorites

Understanding the problem

- Reviews are skewed heavily toward 5 stars
- Not all reviews have accompanying messages

Reviews ★★★★★ (728)

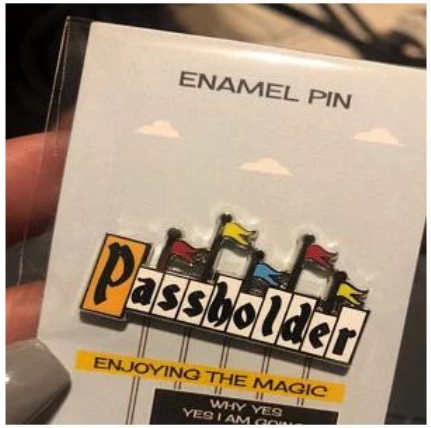
41 people left 5-star reviews this month

Reviewed by Cindy Meyers

★★★★★ Sep 24, 2018

Passholder pin // Fantasy pin // hard enamel pin // annual passholder // Disney pin // Fantasy pins

Reviewed by Grace Martinez



★★★★★ Sep 24, 2018

Absolutely love! Got this for my boyfriend and I as our 1 year Passholder anniversary. We get a new pin every trip we take and he loved this one!!

Mickey popcorn pin // p...

$7.19 ~~$7.99~~ (10% off)

Monorail tote pin // Disn...

$5.39 ~~$5.99~~ (10% off)

Mickey suitcase pin // Mi...

$3.99 ~~$9.98~~ (60% off)

Mickey toaster pin // Mic...

$8.99 ~~$9.99~~ (10% off)

Reviewed by Chris

★★★★★ Aug 4, 2018

Nice passholder pin and had it in an instant!

Passholder pin // Fantasy pin // hard enamel pin // annual passholder // Disney pin // Fantasy pins

Reviewed by Mallory Clark

★★★★★ Jul 23, 2018

Passholder pin // Fantasy pin // hard enamel pin // annual passholder // Disney pin // Fantasy pins

Reviewed by Brie

★★★★★ Jul 20, 2018

Passholder pin // Fantasy pin // hard enamel pin // annual passholder // Disney pin // Fantasy pins

Reviewed by MFAP8278

★★★★★ Jul 19, 2018

Passholder pin // Fantasy pin // hard enamel pin // annual passholder // Disney pin // Fantasy pins

Hypothesis:

Using NLP and machine learning techniques, I can identify the top 10 important qualities and price range needed for any given Etsy user to use to improve their chances of getting favorited.

Methodology

Etsy API

- Grab info:
  - Get the shop info for every shop that can be found with the keyword: "Disney"
  - Use shop IDs to get the listing info
- Model:
  - Use the shop and listing features to train a regression model to identify the amount a listing will be favorited

Analyze Outcomes:

- Use vectorization to identify the most important words in a shop's collective texts that lead to a positive sentiment and being favorited
- Check correlations of the number of favorites with other shop features
- Compare the number of favorites of the current listings with altered descriptions that have those positively associated words the model suggests

```
import lxml
```

In [ ]:
```
# Find shops w/ keyword "Disney", limit to first 200 shops
# Tells me which shops sell Disney themed items
disney_shops_results = []
disney_shops_results.append(json.loads(requests.get(
    'https://openapi.etsy.com/v2/shops?shop_name=Disney&limit=50&offset=0&api_key=j6ws81elytduo5sqhdm5egqv').text)['results'])
disney_shops_results.append(json.loads(requests.get(
    'https://openapi.etsy.com/v2/shops?shop_name=Disney&limit=50&offset=50&api_key=j6ws81elytduo5sqhdm5egqv').text)['results'])
disney_shops_results.append(json.loads(requests.get(
    'https://openapi.etsy.com/v2/shops?shop_name=Disney&limit=50&offset=100&api_key=j6ws81elytduo5sqhdm5egqv').text)['results'])
disney_shops_results.append(json.loads(requests.get(
    'https://openapi.etsy.com/v2/shops?shop_name=Disney&limit=50&offset=150&api_key=j6ws81elytduo5sqhdm5egqv').text)['results'])
```

In [13]:
```
# Display results of API request
# There's more info than we need
disney_shops_results
```

Out[13]:
```
[{'shop_id': 5001407,
  'shop_name': 'DisneyObscura',
  'user_id': 20413,
  'creation_tsz': 1139103736,
  'title': 'Obscure, rare, odd, vintage Disney memories on canvas.',
  'announcement': 'Spreading the love of Disney movies and Disneyland attractions through art.  Stills from your favorite movies, recreations of event and attraction posters and more to come.  *Custom Requests are Welcome!*  \nRight now these are all one of a kind, painted canvas pictures.  In the future, I may add prints and posters and I&#39;m working on wood cutouts and more.  &quot;Favorite&quot; my store and get updates as items are listed.  Thanks!',
  'currency_code': 'USD',
  'is_vacation': False,
  'vacation_message': None,
  'sale_message': 'Thank you for your purchase!  Your item will be wrapped and custom boxed and sent via USPS priority mail with tracking and insurance in the U.S. and first class international everywhere else.  Since this is artwork, it is non-refundable/no returns.  If it arrives damaged, please keep the packaging and notify me.  ~R~',
  'digital_sale_message': None,
  'last_updated_tsz': 1537570840,
  'listing_active_count': 18,
  'digital_listing_count': 0,
```

Examples of API Request Results

	user_id	login_name	shop_name	shop_id	listing_active_count	url
0	20413	shamanicvisions	DisneyObscura	5001407	18	https://www.etsy.com/shop/DisneyObscura?utm_so...
1	30837077	marywinkler5	DisneyforFUN	7867596	711	https://www.etsy.com/shop/DisneyforFUN?utm_sou...
2	24112050	charliemarieriri	LittleWallop	8159392	23	https://www.etsy.com/shop/LittleWallop?utm_sou...
3	38103717	disneynerdclique	DisneyNerdClique	8632402	141	https://www.etsy.com/shop/DisneyNerdClique?utm...
4	20850951	sharkattack2878	MakingCruiseMemories	8635548	184	https://www.etsy.com/shop/MakingCruiseMemories...
5	44053694	disneyfreakboutique	DisneyFreakBoutique	9140037	122	https://www.etsy.com/shop/DisneyFreakBoutique?...
6	44647121	kdorlamarquis	Disneykenscollection	9188796	1132	https://www.etsy.com/shop/Disneykenscollection...
7	42828296	nuhertz	RetroCot	9524912	207	https://www.etsy.com/shop/RetroCot?utm_source=...
8	6112479	jenstar519	DisneyBottleCaps	9544253	226	https://www.etsy.com/shop/DisneyBottleCaps?utm...
9	33819301	jenndukes06	DisneypassionJennie	9881742	115	https://www.etsy.com/shop/DisneypassionJennie?...
10	27992879	elliemcguire1	DisneyCollectorPlate	10850759	53	https://www.etsy.com/shop/DisneyCollectorPlate...
11	75643457	sallyahmadi1	SallysEnchantingEars	11973795	52	https://www.etsy.com/shop/SallysEnchantingEars...

	feedback_id	user_id	message	value
0	321233074.0	20413	if i could give this 1,000 stars i would!! i l...	1
1	319990968.0	20413	You CANNOT go wrong with ordering from this sh...	1
2	314893601.0	20413	Rhonda! You are truly a magical woman! You bri...	1
3	314226292.0	20413	Absolutely in love! They're exactly what I...	1
4	308755577.0	20413	I love it! I like that she made it her own by ...	1
5	302109457.0	20413	Came back to commission this next painting for...	1
6	301569979.0	20413	Rhonda is a great find on Etsu	1
7	301473094.0	20413	Second time coming to Rhonda for a purchase. ...	1
8	295495440.0	20413	Absolutely a masterpiece. The best painting I ...	1
9	292736232.0	20413	This is a masterpiece that belongs at Disneyla...	1
10	288215630.0	20413	Could not have asked for a better product. Mat...	1
11	285651217.0	20413	Looks great!	1

Final Data Frame with Etsy Shop Info

	user_id	shop_name	num_favorites	title	description	tags	price	message	value	listing_active_count	creation_yr
0	20413	DisneyObscura	355.0	Moana toddler baby with Stitch crossover mashu...	Baby Moana making friends with Stitch on the b...	moana, baby moana, toddler moana, stitch, beac...	57.277778	if i could give this 1,000 stars i would!! i l...	1.000000	18	2006
1	6112479	DisneyBottleCaps	1113.0	Disney Princess POCAHONTAS 3D Finished Bottle ...	♥ Finished Bottle Cap ♥Finished Bottle Caps fo...	disney, disney princess, princess pocahontas, ...	7.759690	Perfect party favor for my daughter's birthday...	0.992278	226	2015
2	7543374	DisneyGirl84	12.0	Princess Leia custom disney ears	These cute ears are inspired by Star Wars and ...	gifts, disney, vacation, disneyland, star wars...	20.000000	Great buyer and easy to communicate with Grea...	1.000000	1	2011
3	8082338	disneyfarmhouse	0.0	Disney autograph signing board Disney autograp...	Disneyland, Walt Disney World, or Disney cruis...		15.000000	I was expecting a canvas, and got a night piec...	1.000000	5	2018

Final Data Frame with Listing Info

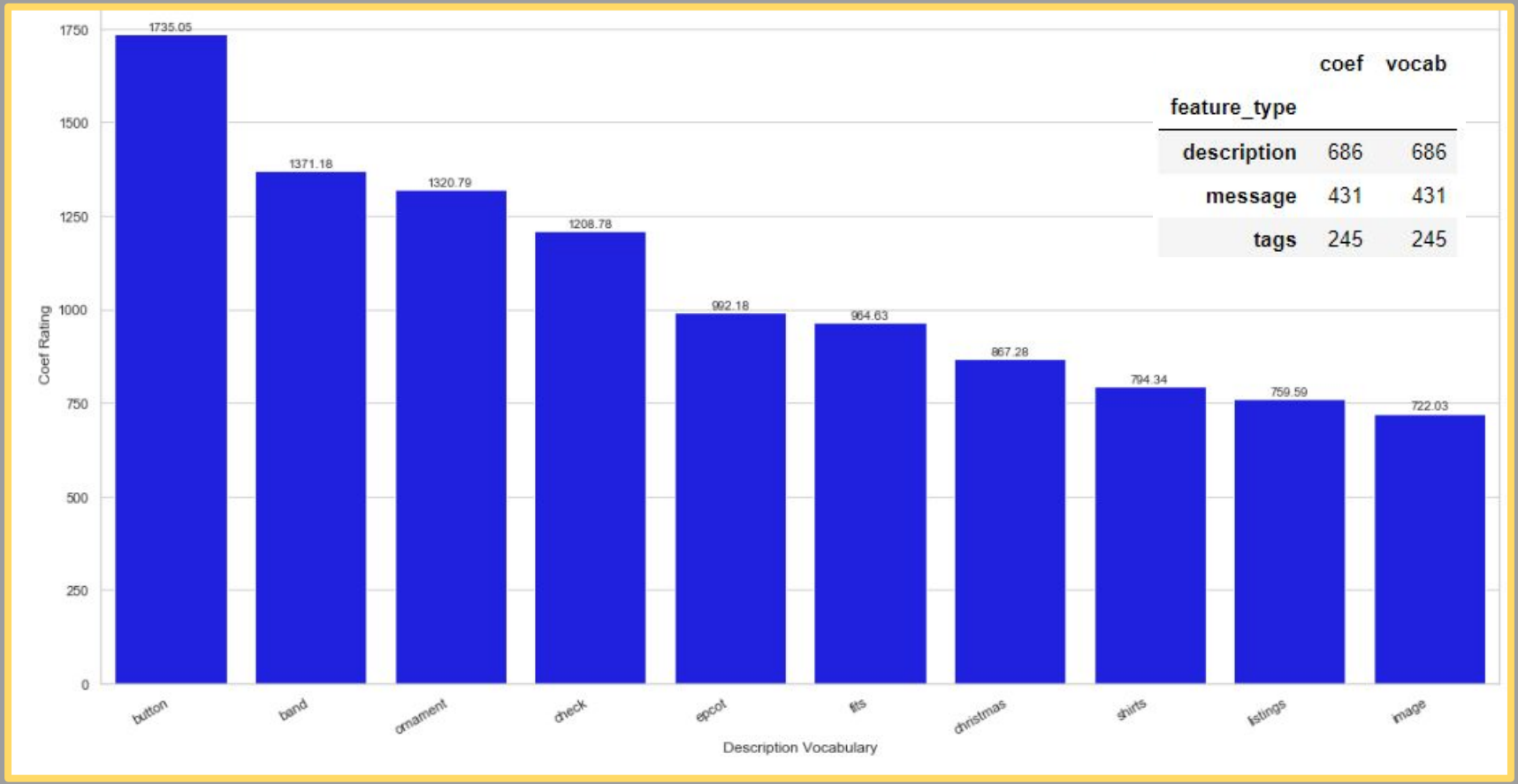
	listing_id	num_favorites	user_id	title	description	tags	style	is_digital	is_customizable	price	creation_yr
0	262981963	100	38103717	Up inspired our love is sn adventure 3 button	Our LE Up inspired our love is an adventure 3 ...	['balloons', 'Ellie', 'Carl', 'up', 'disney mo...		0	0	5.00	2018
1	591514611	100	103655661	Flower and garden Mickey Mouse ears| flower an...	Gorgeous rose gold ears with peach, pink and ...	['Mickey ears', 'Mickey Mouse ears', 'Minnie ...		0	1	35.00	2018
2	528942794	100	30837077	Dcl FE gift (2) LED Lanterns great party favor...	You get 2Stand about 4 inches tall very cute d...	['Disney shower', 'hidden Mickey', 'DISNEYCRUI...		0	1	20.00	2018
3	549566703	100	44053694	Black and White Disney Family Shirts - Disney ...	Halloween Special !! **********PL...	['disney family shirts', 'halloween disney', '...		0	0	14.00	2018

Machine Learning Regression Model

Great! It seems like the error for our lasso regression model is not as high as when we take the averages and predict that way Next I will try and compare the results with ridge regression.

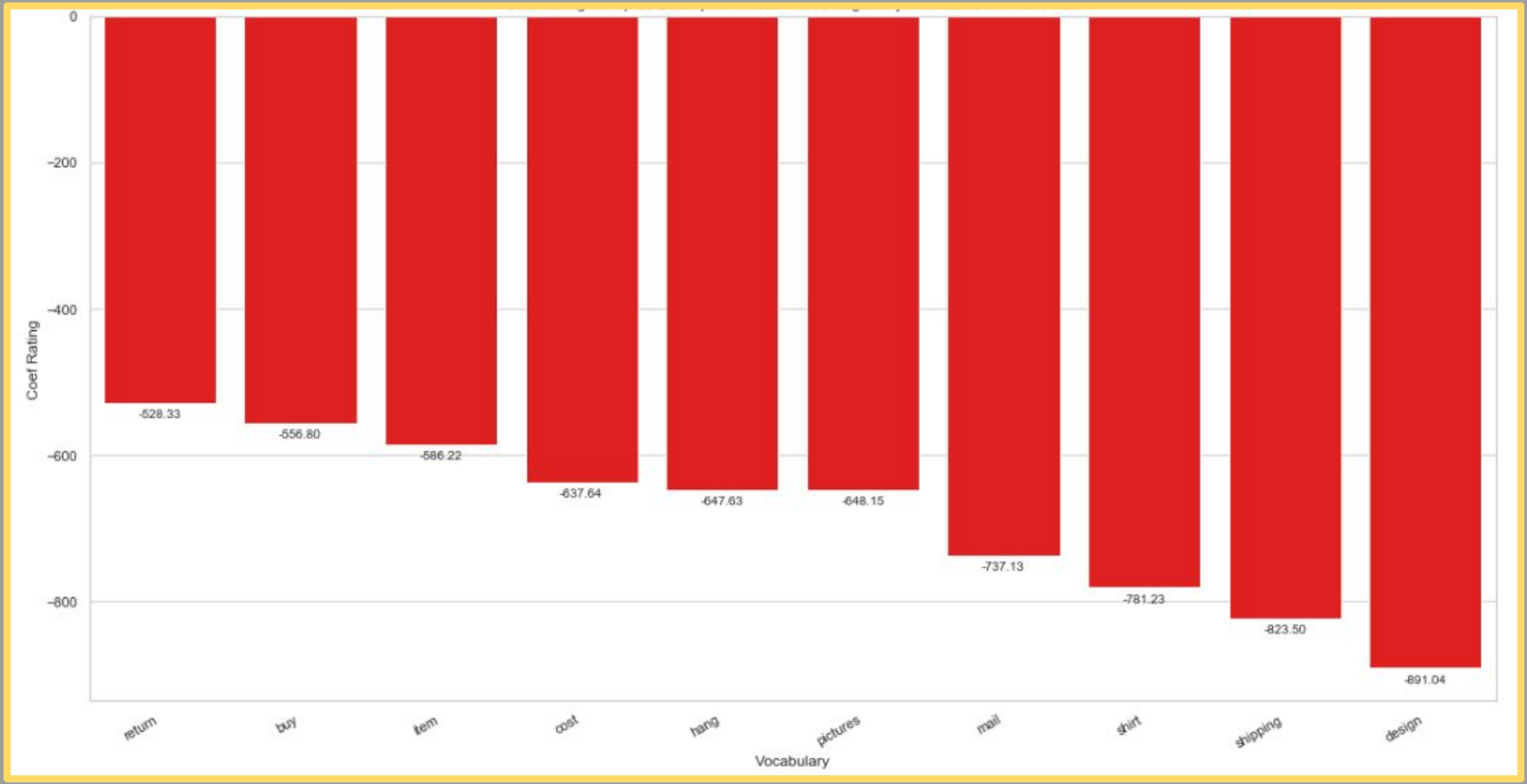
In [87]:
```
# Play around with the alpha parameter and see if we can improve it
ridgeregr = linear_model.Ridge(alpha=0.35)
err = []
err_dumb = []
for train, test in kf.split(X_matrix):
    print(" ")
    X_train = X_matrix[train]
    y_train = y[train]
    X_test = X_matrix[test]
    ridgeregr.fit(X_train, y_train)
    y_test = y[test]
    y_pred = ridgeregr.predict(X_test)
    err.append(np.mean(np.abs(y_pred - y_test)))
    y_pred_dumb = np.array([np.mean(y_train)]*len(y_test))
    err_dumb.append(np.mean(np.abs(y_pred_dumb - y_test)))
```

Coef Rating of Top 25 Description Vocab Words Associated with Favorites



feature_type	coef	vocab
description	686	686
message	431	431
tags	245	245

Coef Rating of Top 25 Description Vocab Words Negatively Associated with Favorites



Outcomes - Shop Level

Top 10 vocabulary:

Descriptions: 'button', 'band', 'ornament', 'check', 'epcot', 'fits', 'christmas', 'shirts', 'listings', 'image'

Messages: 'magic', 'shipping', 'trip', 'christmas', 'ordered', 'shirts', 'family', 'disney', 'husband', 'pin'

Tags: 'epcot', 'band', 'vinyl', 'magic', 'ornament', 'custom', 'silver', 'movie', 'cruise', 'tank'

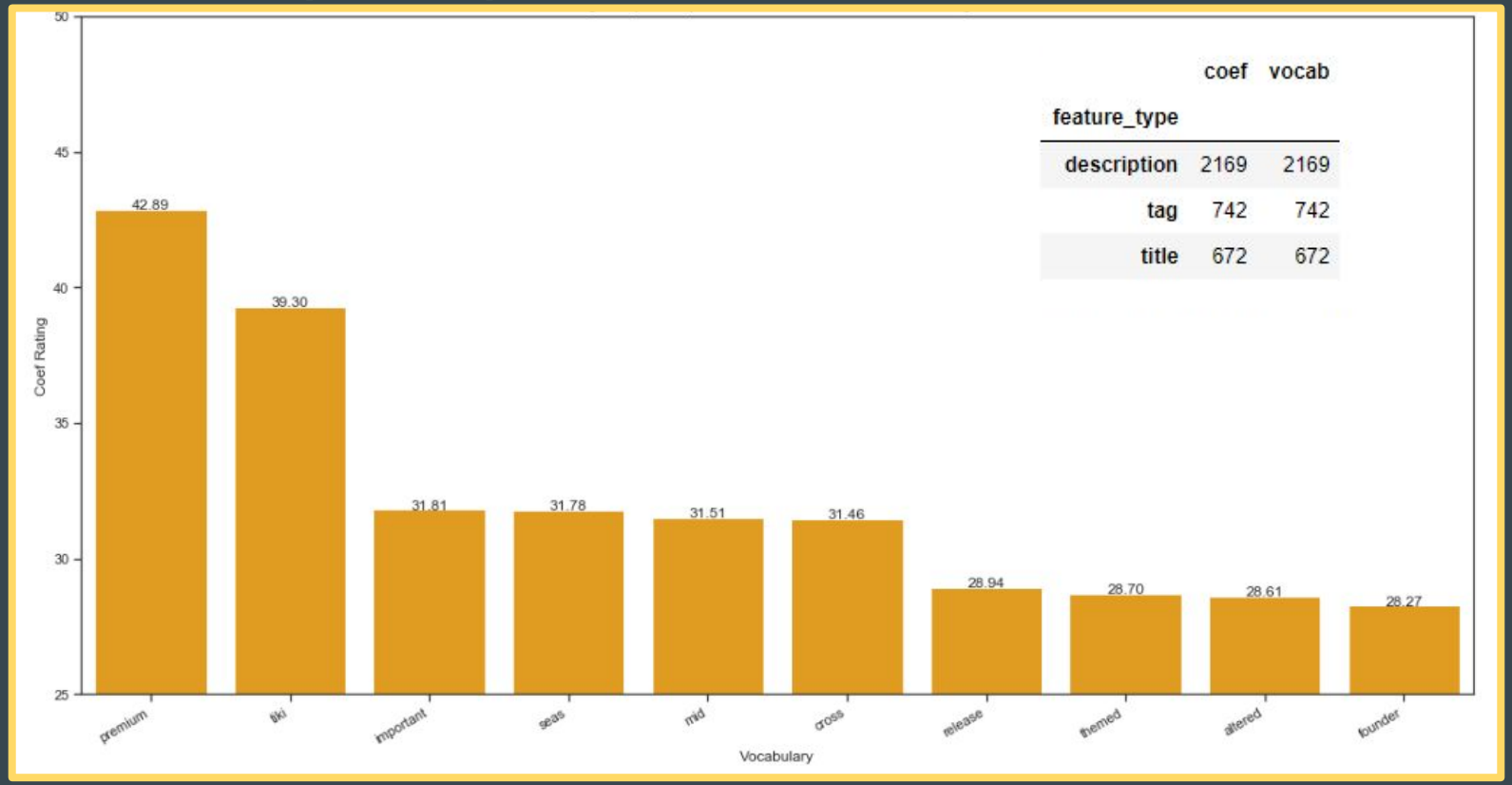
Bottom 10 vocabulary:

Descriptions: 'return', 'buy', 'item', 'cost', 'hang', 'pictures', 'mail', 'shirt', 'shipping', 'design

Messages: None (Makes sense because unlikely to fav and review a disliked item)

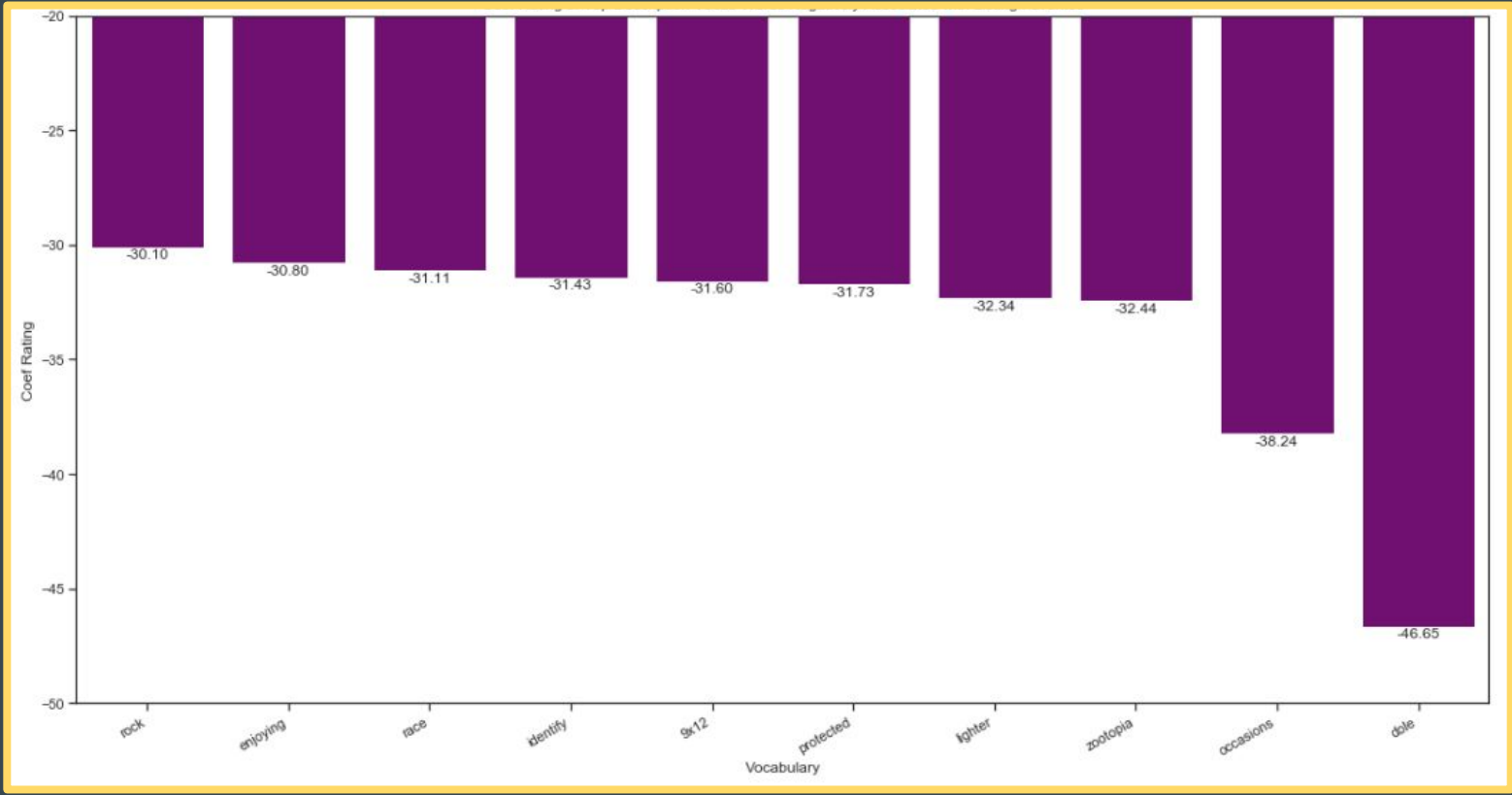
Tags: 'stitch', 'new', 'fantasy', 'vacation', 'wars', 'star', 'girls', 'gold', 'vintage', 'book'

Coef Rating of Top Description Words Associated with Listing Favorites



feature_type	coef	vocab
description	2169	2169
tag	742	742
title	672	672

Coef Rating of Top Description Words Negatively Associated with Listing Favorites



Outcomes - Listing Level

Top 10 vocabulary:

Descriptions: 'premium', 'tiki', 'important', 'seas', 'mid', 'cross', 'release', 'themed', 'altered',
'founder'

Titles: 'stripes', 'favors', '1982', 'lapel', 'ring', 'dole', 'whip', '5x5', 'autograph', 'tramp'

Tags: 'make', 'theme', 'solo', 'lions', 'treat', 'miss', 'unisex', 'company', 'california',
'illustration'

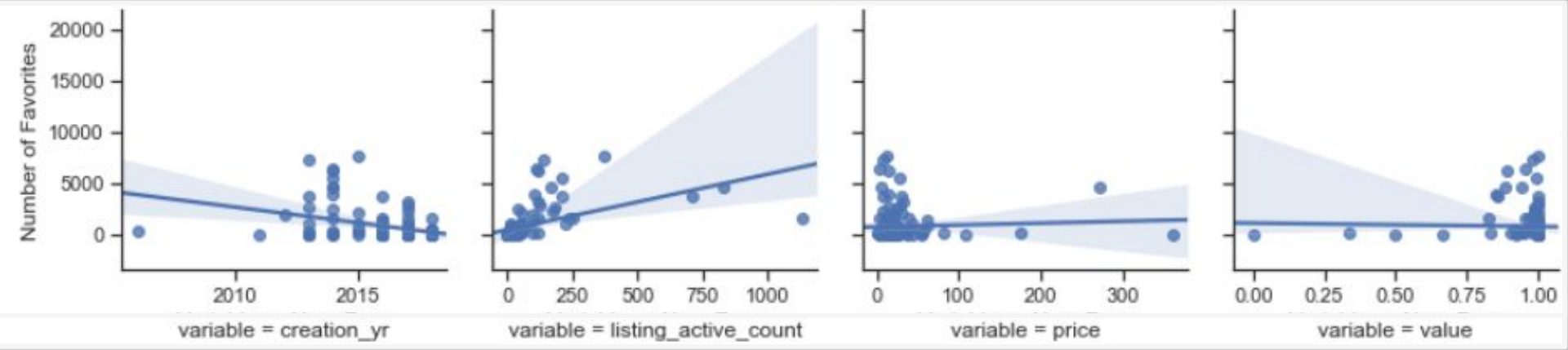
Bottom 10 vocabulary:

Descriptions: 'rock', 'enjoying', 'race', 'identify', '9x12', 'protected', 'lighter', 'zootopia',
'occasions', 'dole'

Titles: 'years', 'painting', 'crown', 'boho', 'youth', 'add', 'spider', 'retrocot', 'card', 'instant'

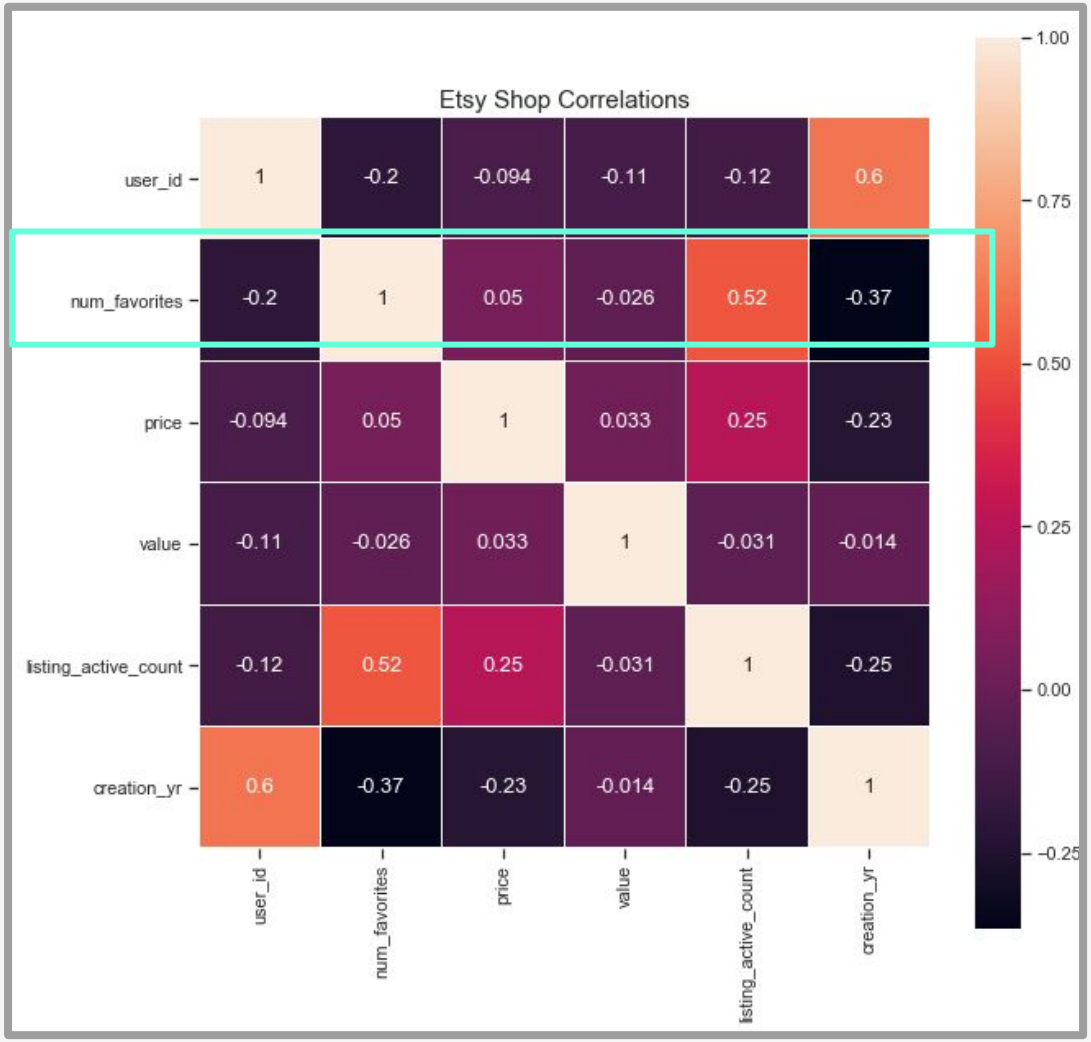
Tags: '1982', 'mickeys', 'lapel', 'paper', 'bob', 'han', 'lightyear', 'ironman', 'terror', 'boy'

Numerical Variable vs. Number of Favorites

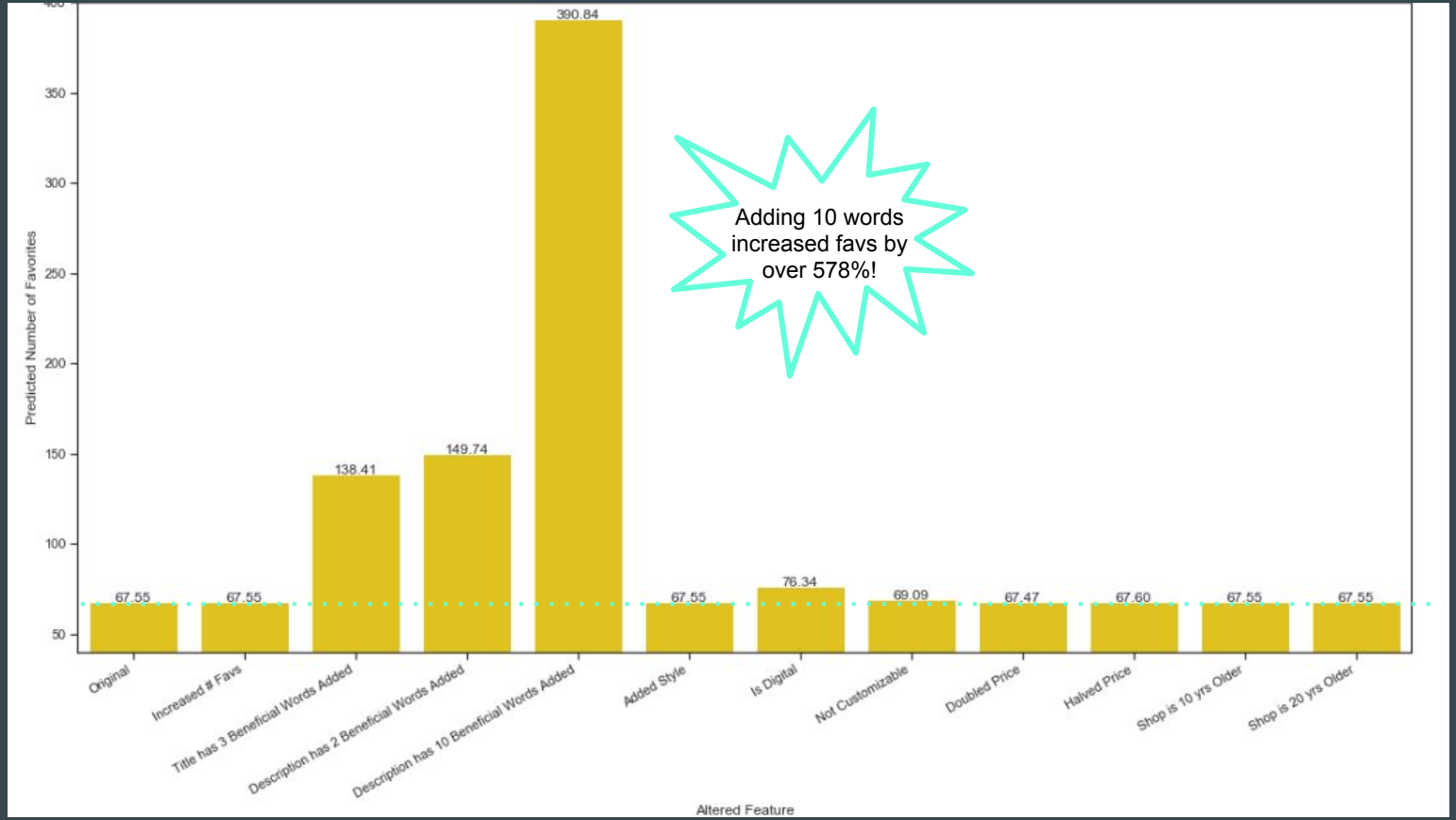


It seems:
- As time ↑, num favs ↓
- As num listings ↑, num favs ↑
- As price ↑, num favs ↓ (slightly)
- As rating value ↑, num favs ↑ (slightly)

Strongest correlation=0.52, between the number of favorites and listing active count.



Predicted # Favorites when Altered Shop Features



Proposed Deliverables - Shop Level

Qualities	Fast shipping, Christmas themed, magical, Disney family shirts, pins, accents
Keywords to use	Magic, ornament, button, band, vinyl, Epcot
Important features	Description words, # of Active listings
Keywords to avoid	Shirt, mail, book, vintage, gold, design

Proposed Deliverables - Listing Level

Qualities	Themed, unisex, autographed, striped, rings
Keywords to use	Premium, tiki, important, seas, cross, release, themed, altered, founder
Important features	Descriptions, Title, Digital copy
Keywords to avoid	rock, enjoying, race, identify, 9x12, protected

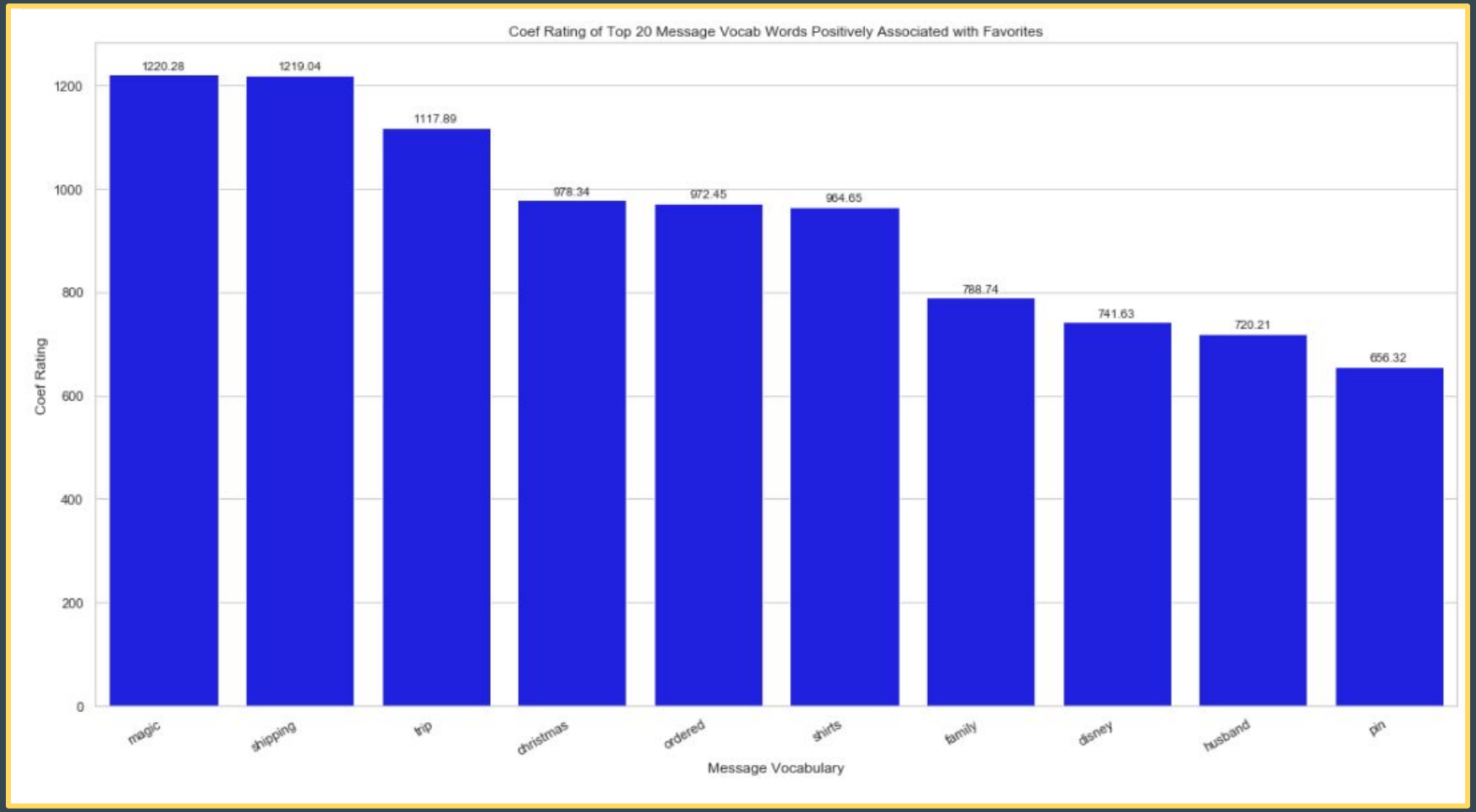
Future Directions

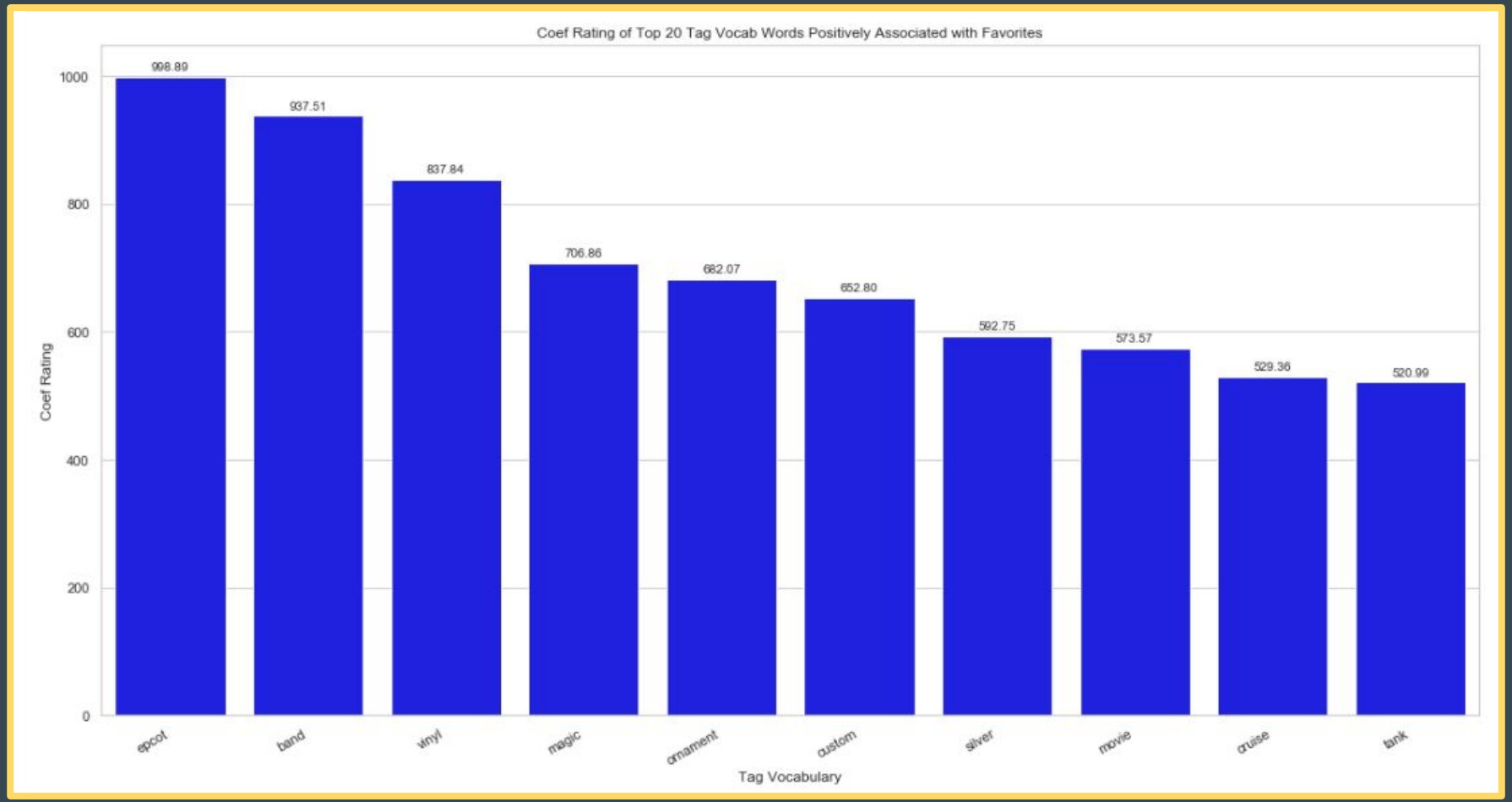
- Learn more about web scraping and do sentiment analysis and mapping for listing level of reviews
- Increase sample size
- Go through every listing in my mom's shop to see how she can change her descriptions/titles/tags to hopefully increase the number of favorites
- Try the application with non-Disney themed shops

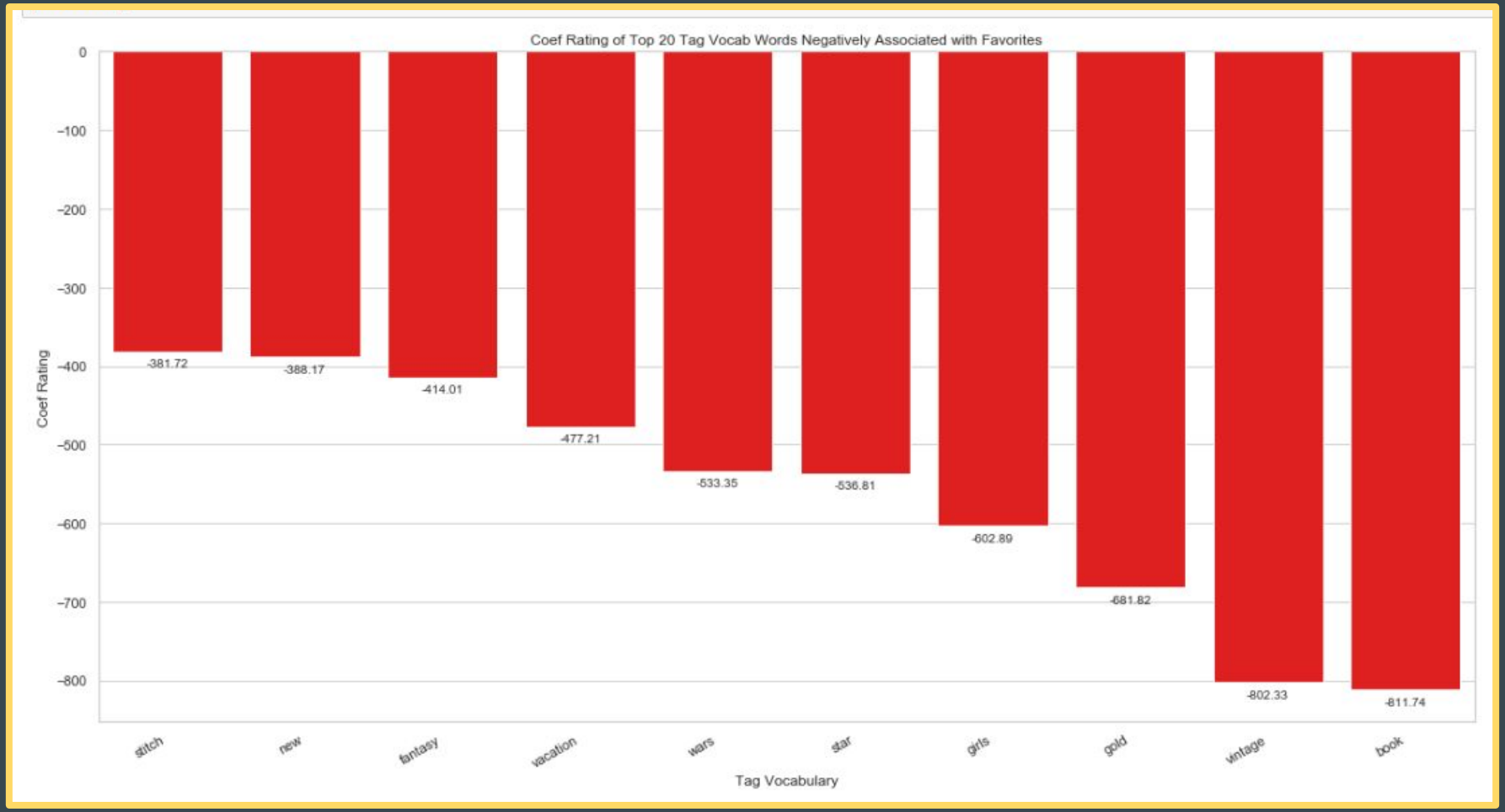
Acknowledgements

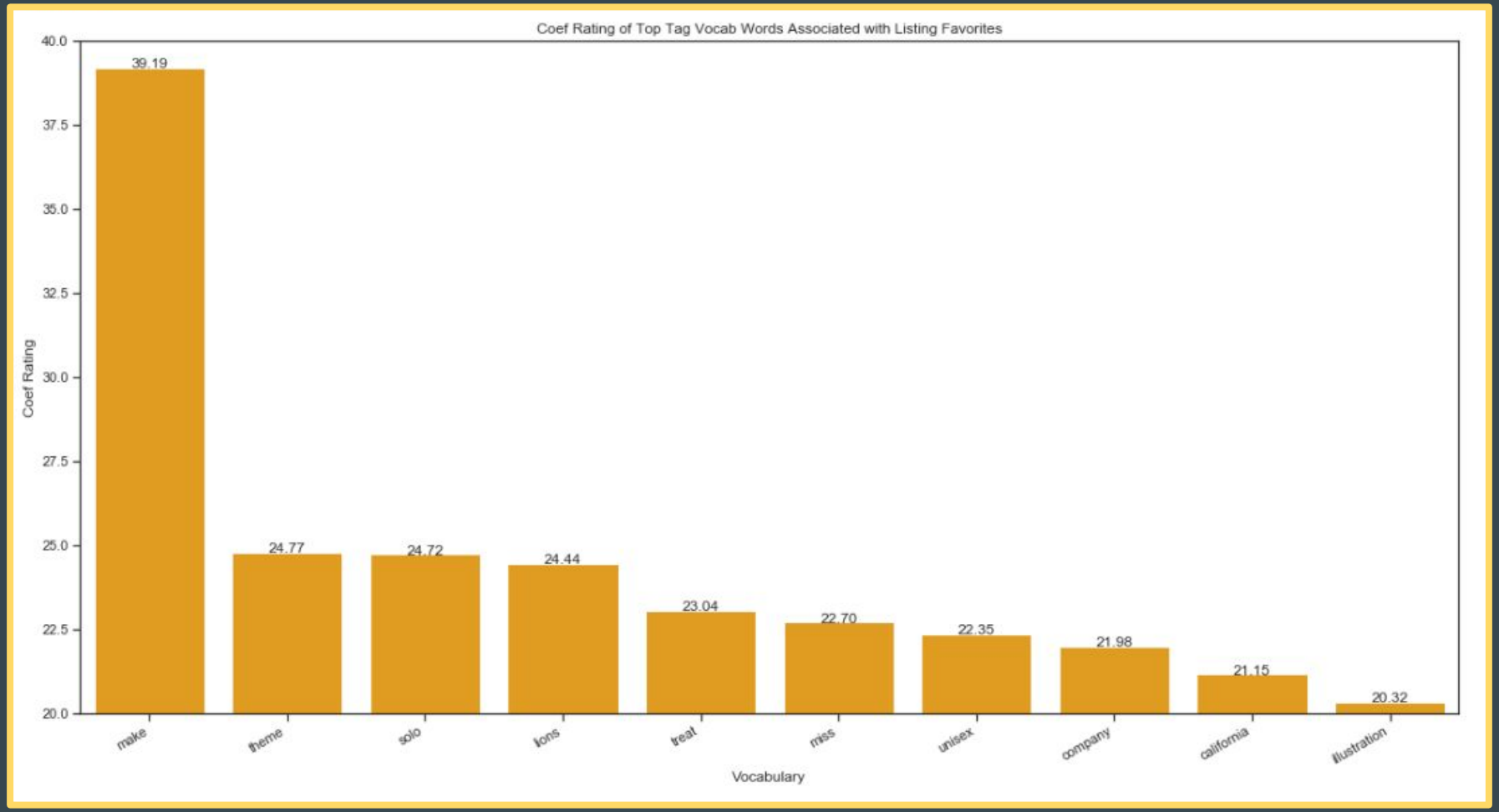
Thinkful mentor: Drew Waters

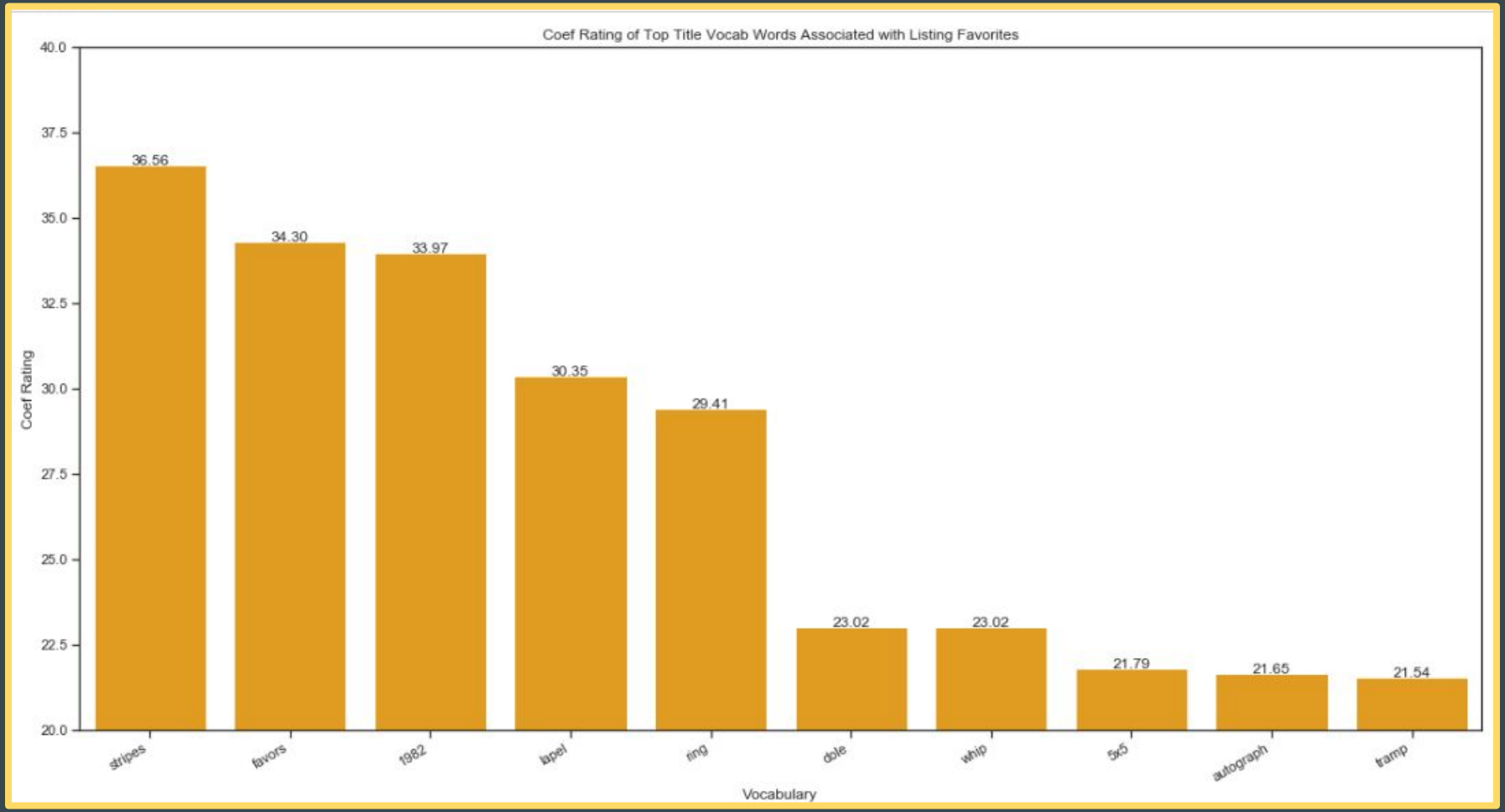
Data and Images Source: Etsy.com

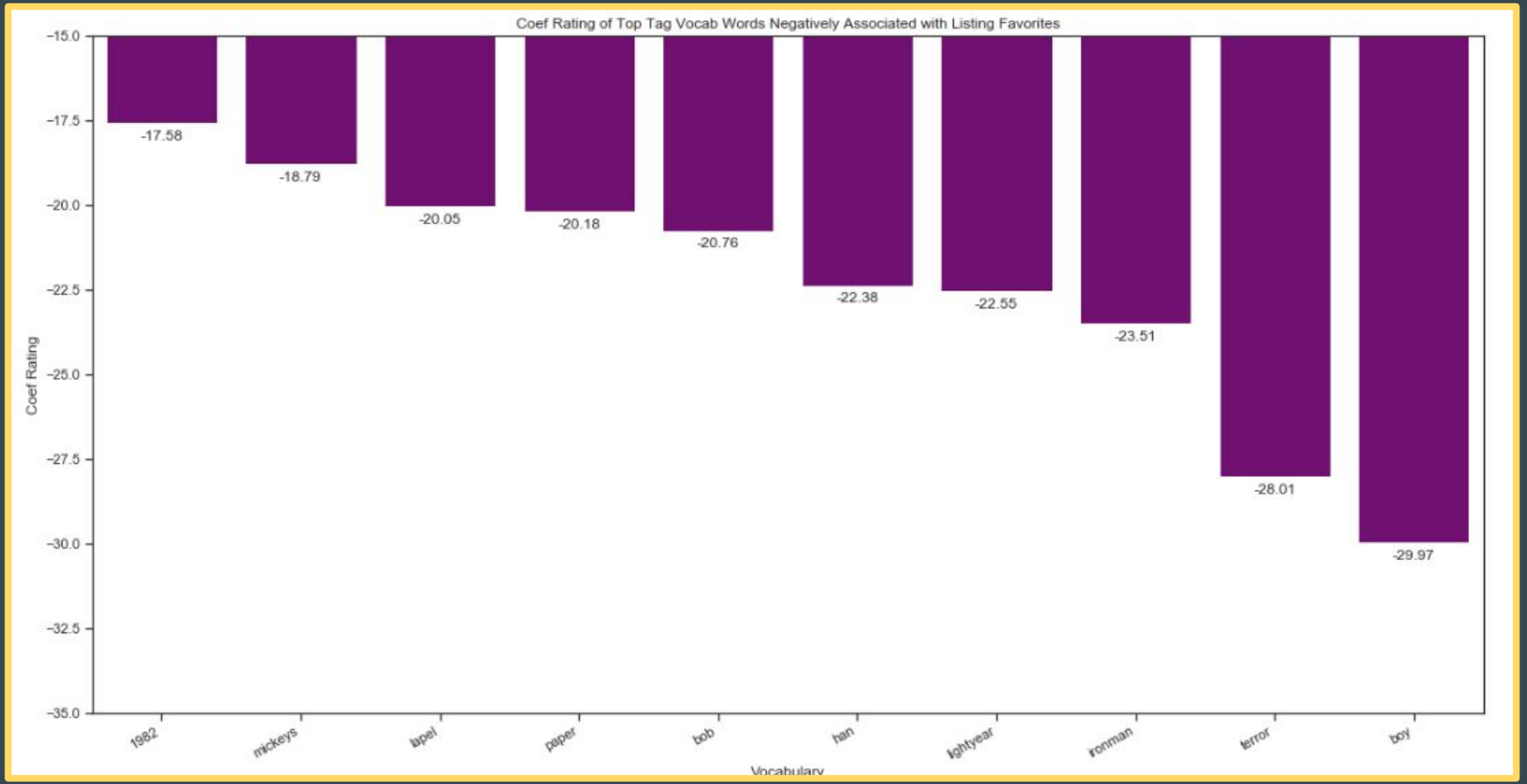

Coef Rating of Top 20 Message Vocab Words Positively Associated with Favorites
Coef Rating
0
200
400
600
800
1000
1200
1220.28
1219.04
1117.89
978.34
972.45
964.65
788.74
741.63
720.21
656.32
magic
shipping
trip
christmas
ordered
shirts
family
disney
husband
pin
Message Vocabulary

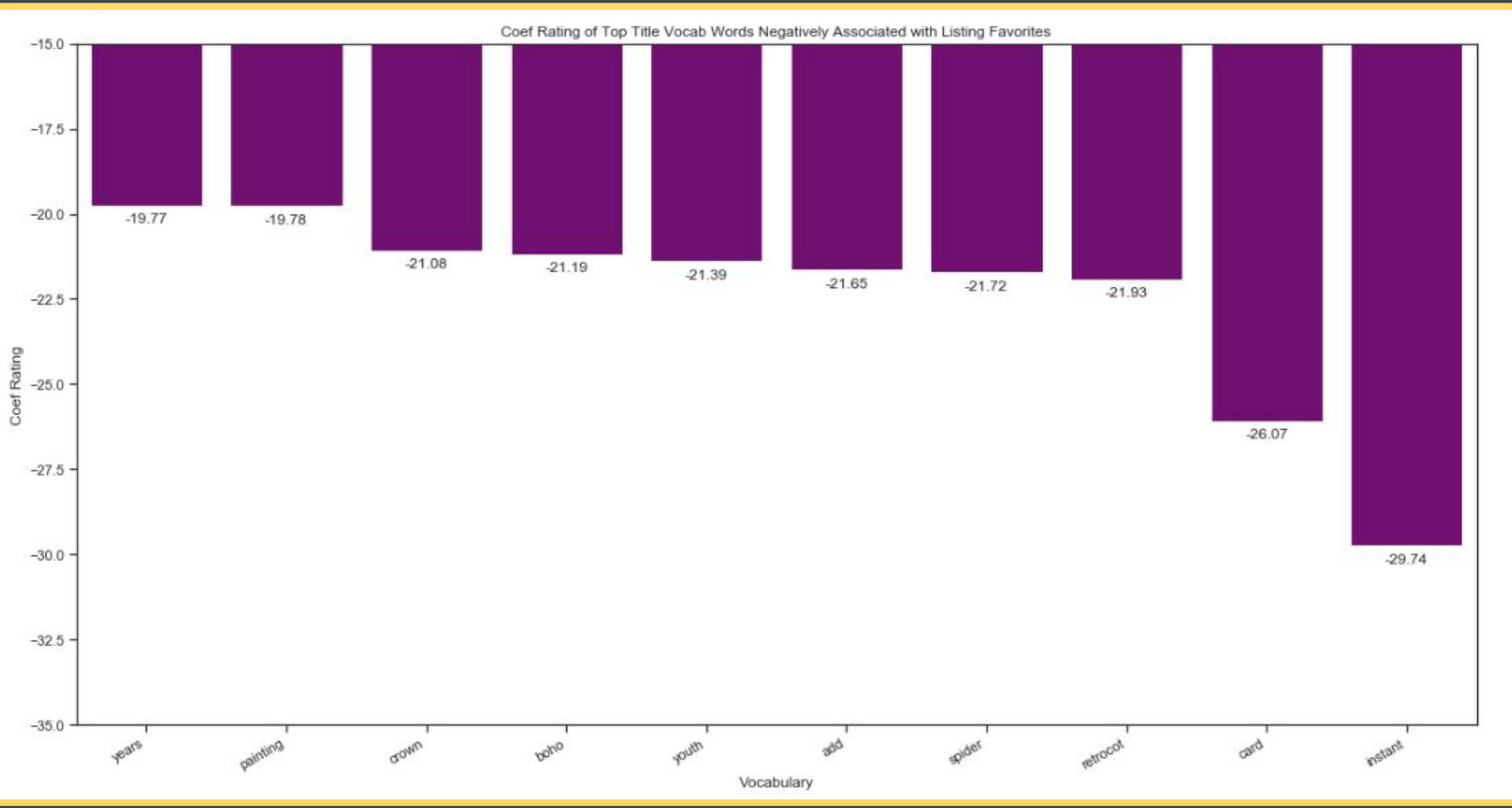

Coef Rating of Top 20 Tag Vocab Words Positively Associated with Favorites
Coef Rating
1000
800
600
400
200
0
998.89
937.51
837.84
706.86
682.07
652.80
592.75
573.57
529.36
520.99
epcot
band
vinyl
magic
ornament
custom
silver
movie
cruise
tank
Tag Vocabulary


Coef Rating of Top 20 Tag Vocab Words Negatively Associated with Favorites
Coef Rating
0
−100
−200
−300
−400
−500
−600
−700
−800
-381.72
-388.17
-414.01
-477.21
-533.35
-536.81
-602.89
-681.82
-802.33
-811.74
stitch
new
fantasy
vacation
wars
star
girls
gold
vintage
book
Tag Vocabulary


Coef Rating of Top Tag Vocab Words Associated with Listing Favorites
Coef Rating
40.0
37.5
35.0
32.5
30.0
27.5
25.0
22.5
20.0
39.19
24.77
24.72
24.44
23.04
22.70
22.35
21.98
21.15
20.32
mike
theme
solo
lions
treat
miss
unisex
company
california
illustration
Vocabulary


Coef Rating of Top Title Vocab Words Associated with Listing Favorites
Coef Rating
40.0
37.5
35.0
32.5
30.0
27.5
25.0
22.5
20.0
36.56
34.30
33.97
30.35
29.41
23.02
23.02
21.79
21.65
21.54
stripes
favors
1962
lapel
ring
dole
whip
5x5
autograph
tramp
Vocabulary


Coef Rating of Top Tag Vocab Words Negatively Associated with Listing Favorites
Coef Rating
−15.0
−17.5
−20.0
−22.5
−25.0
−27.5
−30.0
−32.5
−35.0
-17.58
-18.79
-20.05
-20.18
-20.76
-22.38
-22.55
-23.51
-28.01
-29.97
1982
mickeys
label
paper
bob
han
lightyear
ironman
terror
boy
Vocabulary


Coef Rating of Top Title Vocab Words Negatively Associated with Listing Favorites
Coef Rating
−15.0
−17.5
−20.0
−22.5
−25.0
−27.5
−30.0
−32.5
−35.0
-19.77
-19.78
-21.08
-21.19
-21.39
-21.65
-21.72
-21.93
-26.07
-29.74
years
painting
crown
boho
youth
add
spider
retrocot
card
instant
Vocabulary