Venmo money
Venmo problems

an investigation by Allen Chen

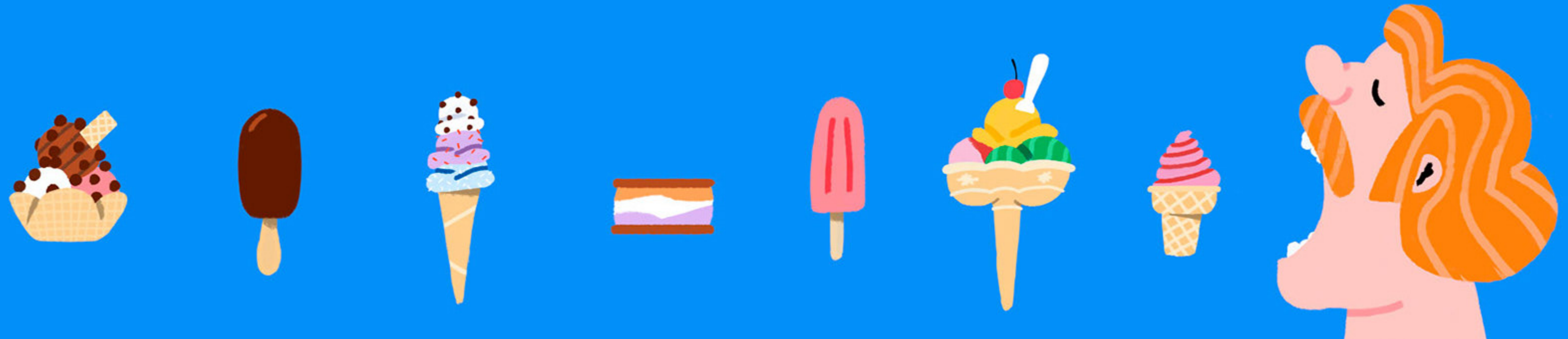
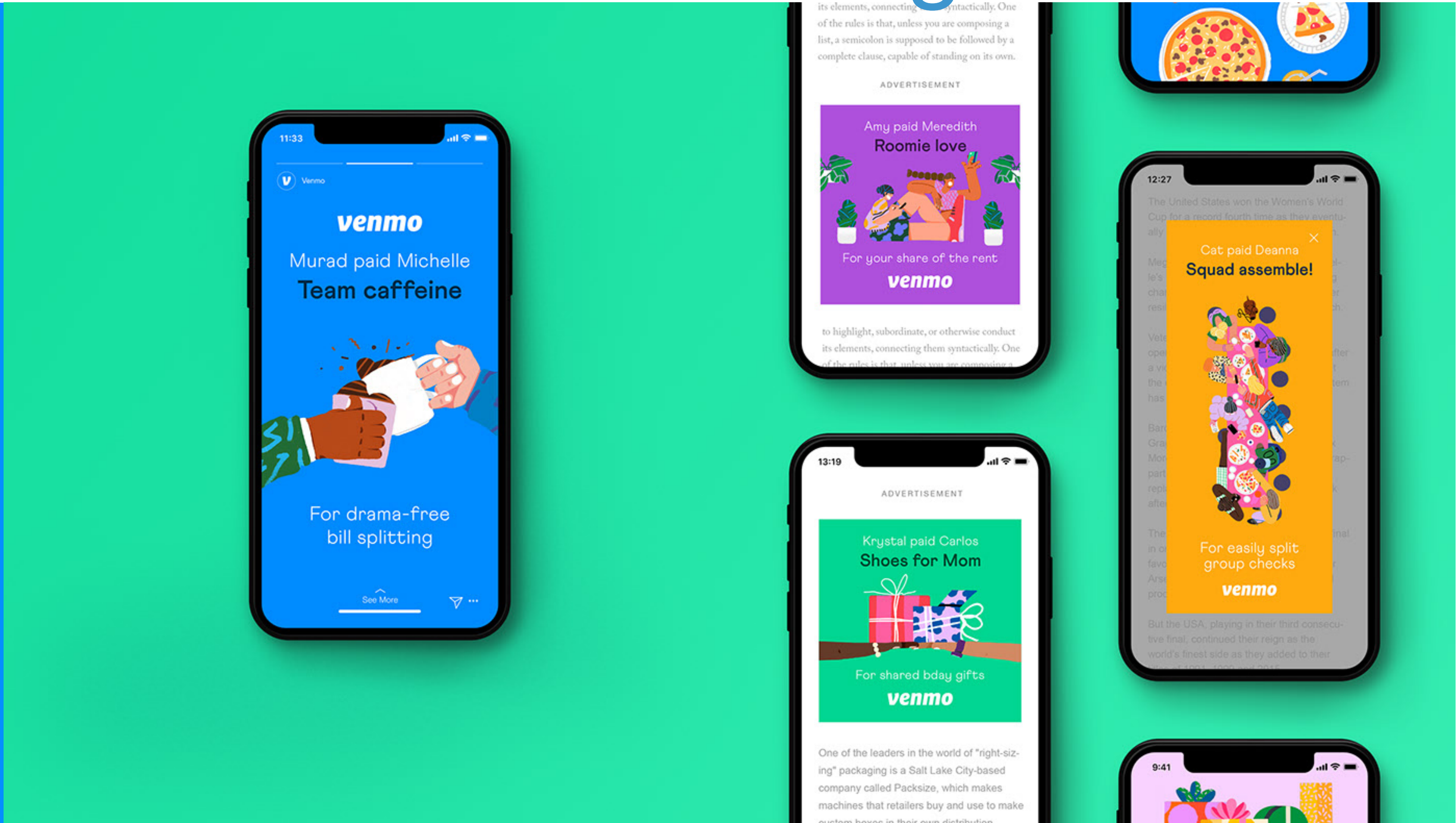


Table of Contents

Venmo: What is it good for?



Data Collection

1. Get Common Names
2. Search for Usernames
3. Pull public transactions

286k transactions | ~30k users

Data Processing

1. Convert transactions into DataFrame

2. Clean text

3. “Demojize” Emojis (e.g. pizza)

4. Manual adjustments for Google’s pre-trained word embeddings (e.g. french_fries)

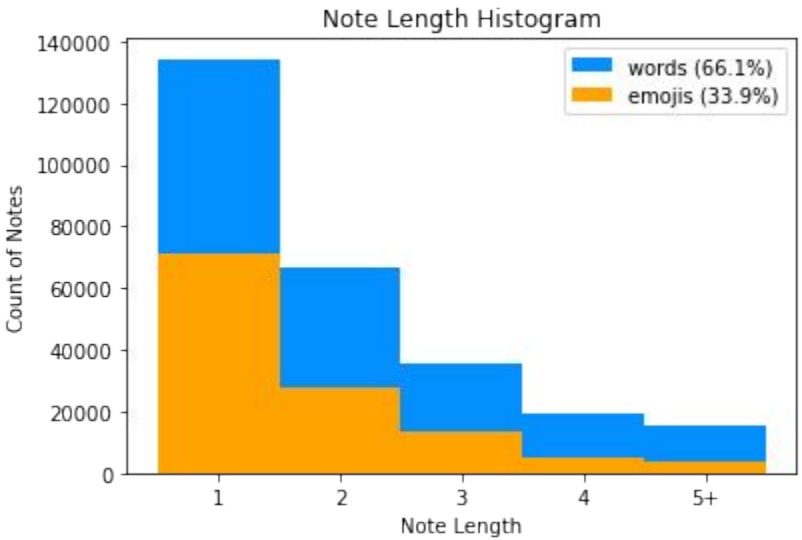
Motivation

Data

Modeling

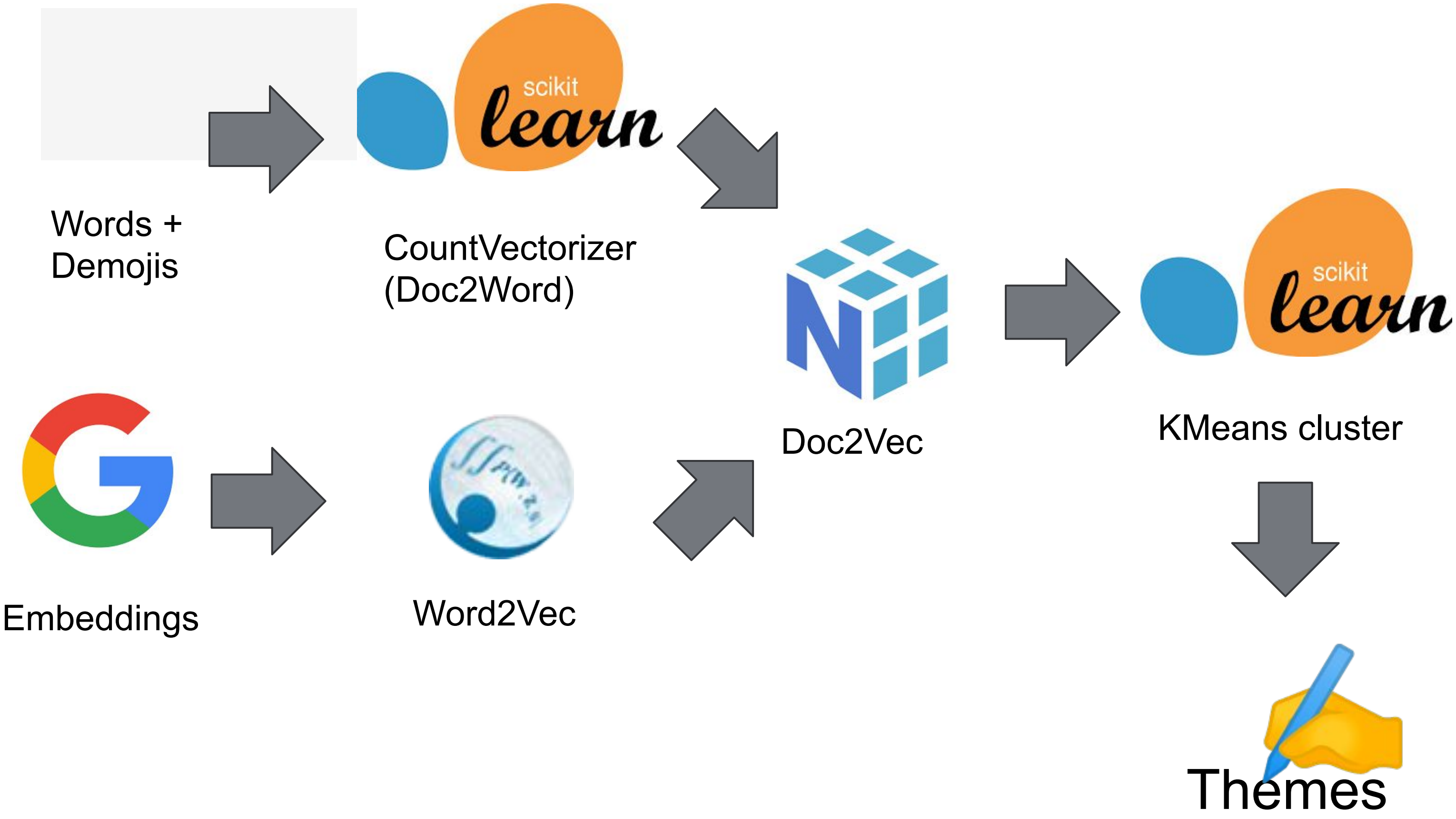
Conclusion

Exploratory Data Analysis



- Notes are short!
- Emoji use is very prevalent!

Modeling Workflow



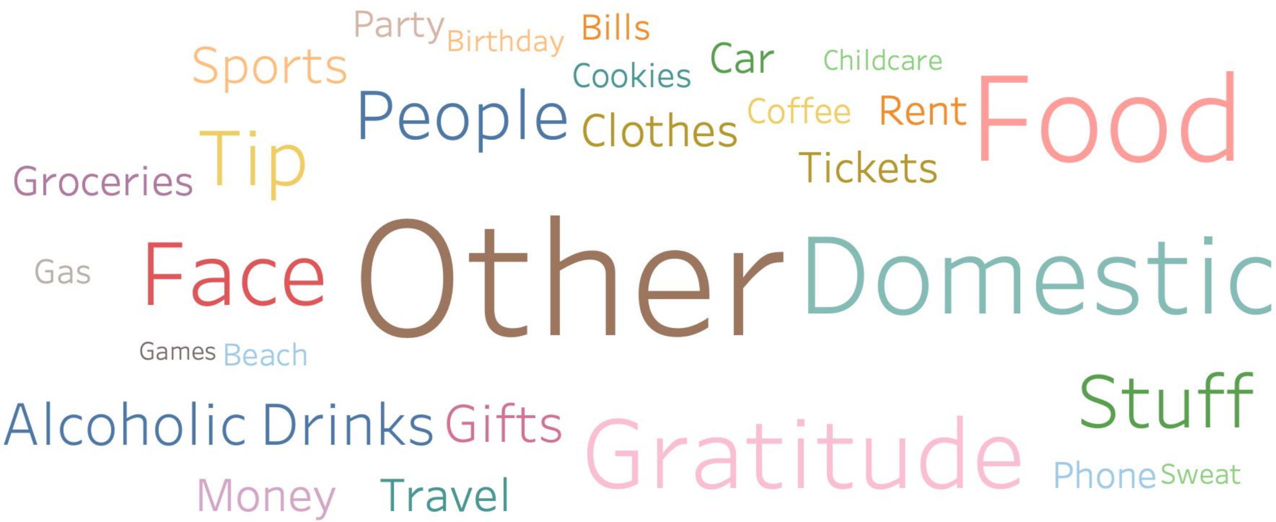
Motivation

Data

Modeling

Conclusion

Venmo Themes



Future Work

- Time Analysis
    - Impact of COVID?
    - Seasonality of sports vs rent/utilities
    - Taco Tuesday?
- Cohort Comparison
    - iPhone vs Android
    - Skin Tones (though be careful of conclusions)
- Suggest Emojis
    - Identify common words and themes that don’t have emojis

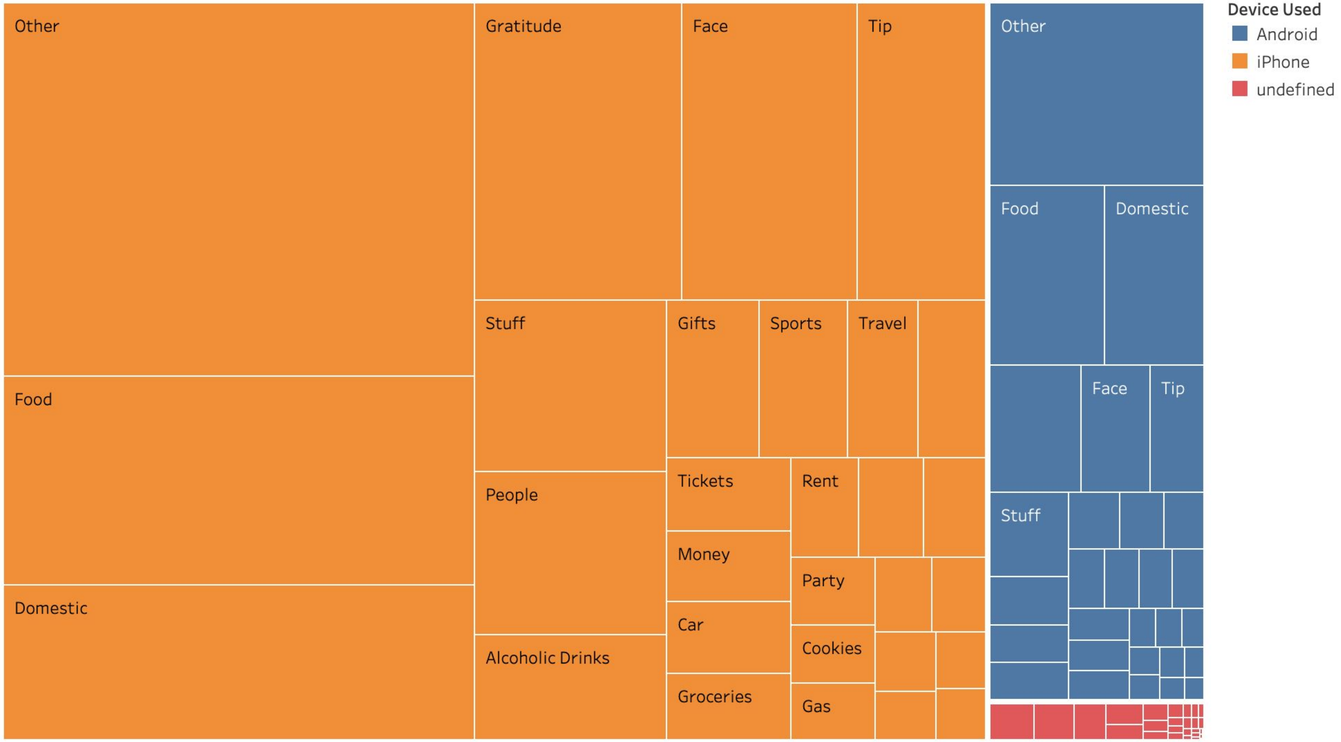
Motivation

Data

Modeling

Conclusion

More iPhone than Android (but the themes are the same)



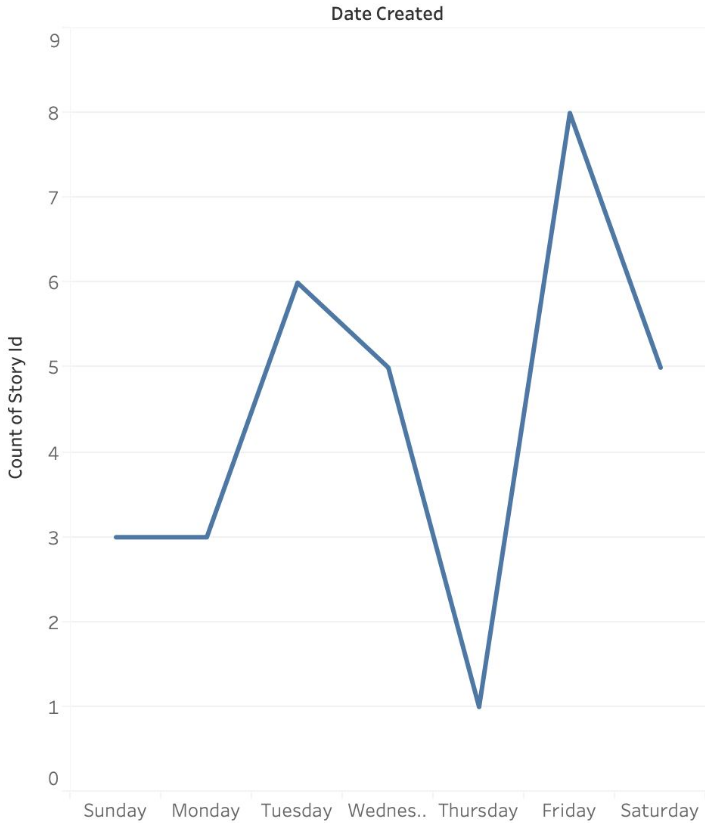
Motivation

Data

Modeling

Conclusion

Taco Tuesdays and Fridays



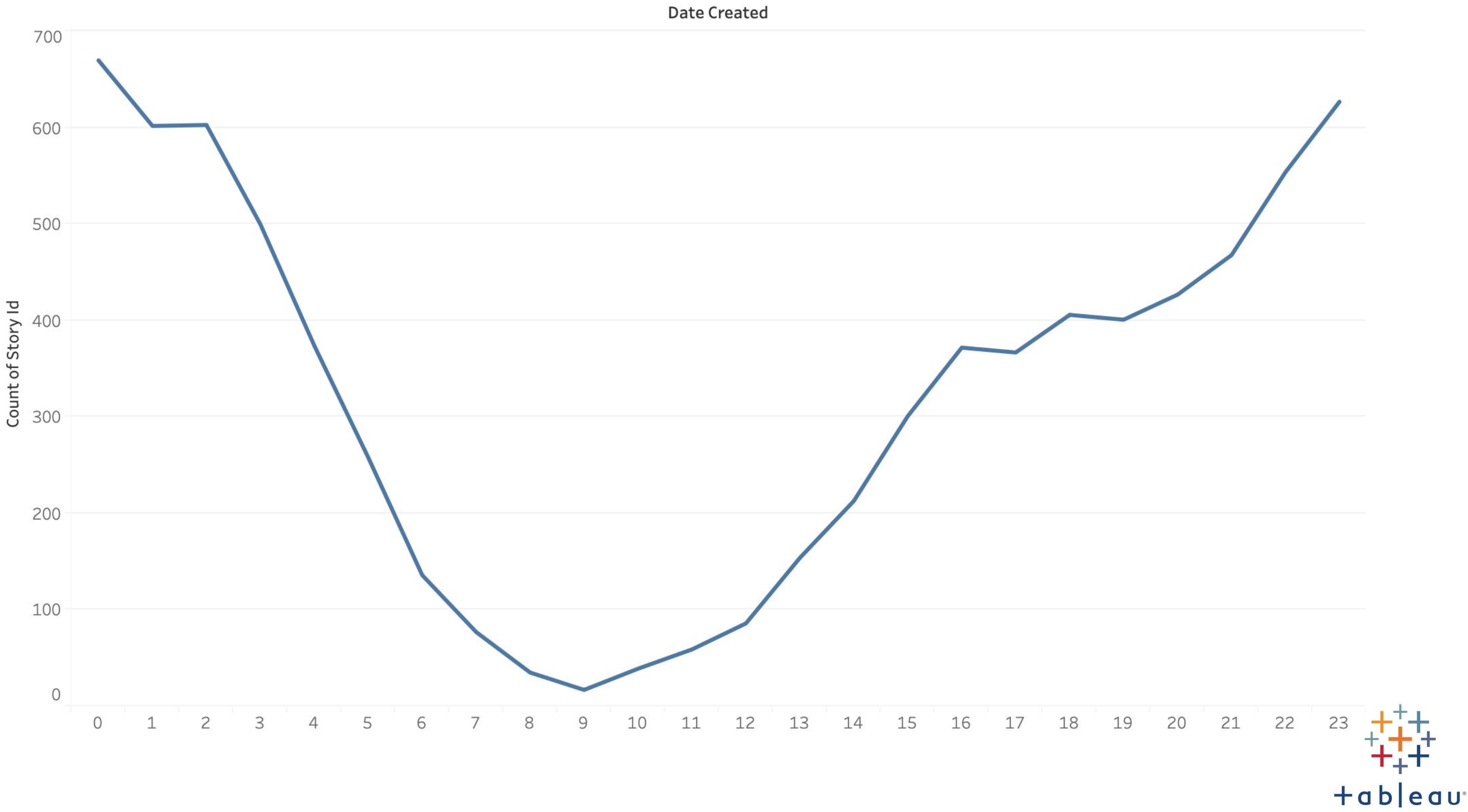
Motivation

Data

Modeling

Conclusion

Drink O Clock is Late at Night



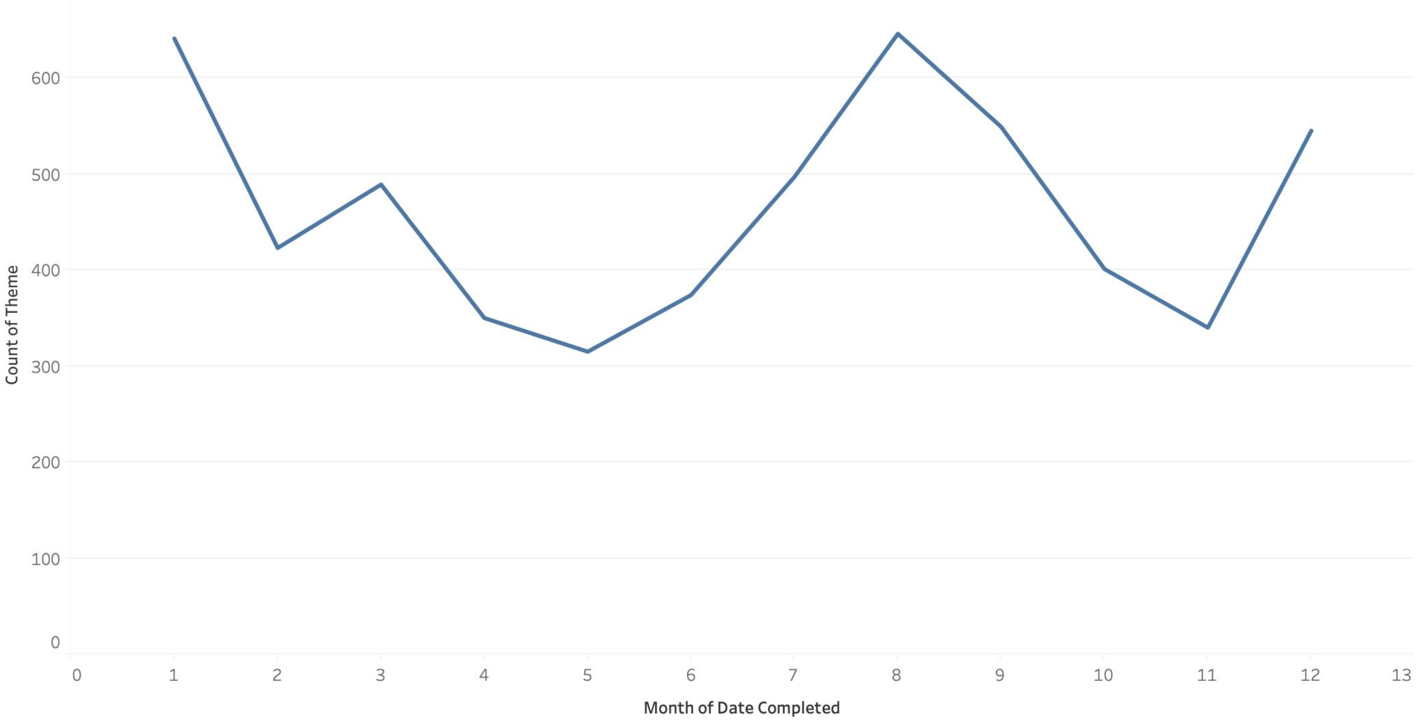
Motivation

Data

Modeling

Conclusion

Sports: Fantasy Football and March Madness



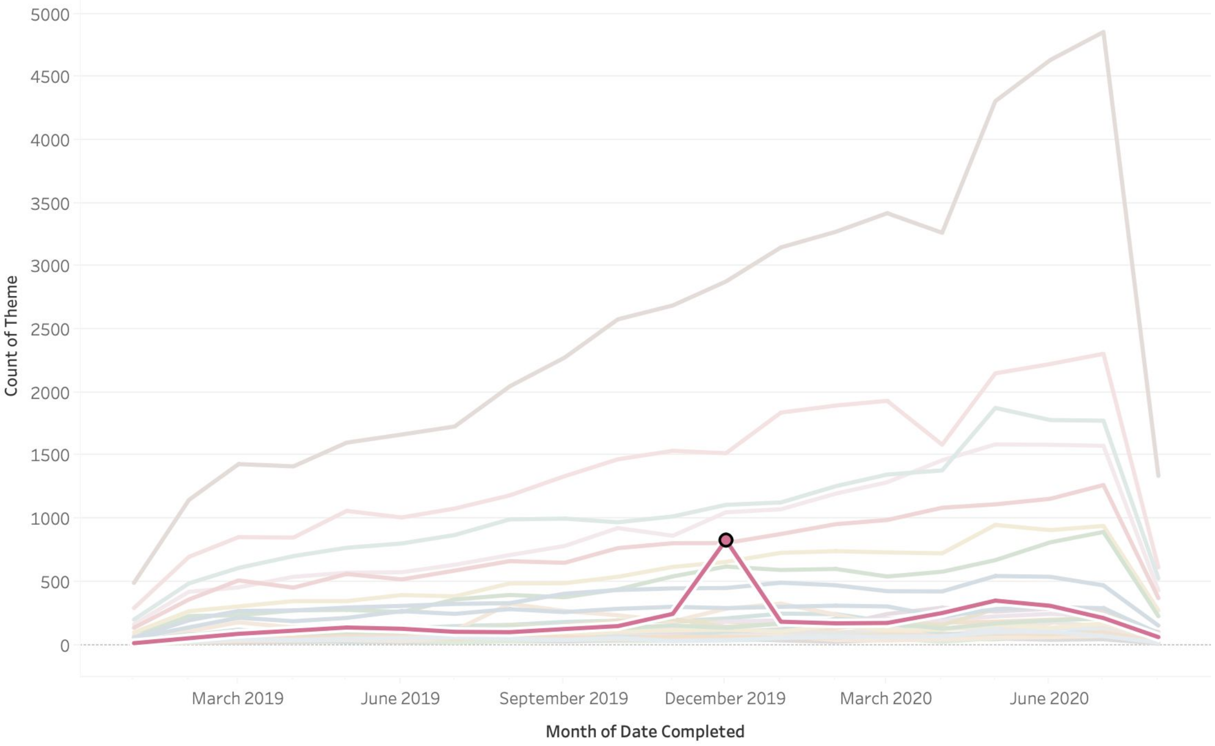
Motivation

Data

Modeling

Conclusion

Gifts: Christmas is the time for gifts



Thank you!

venmo

@Allen-Chen-22

allenchen@gmail.com
https://github.com/mrallenchen

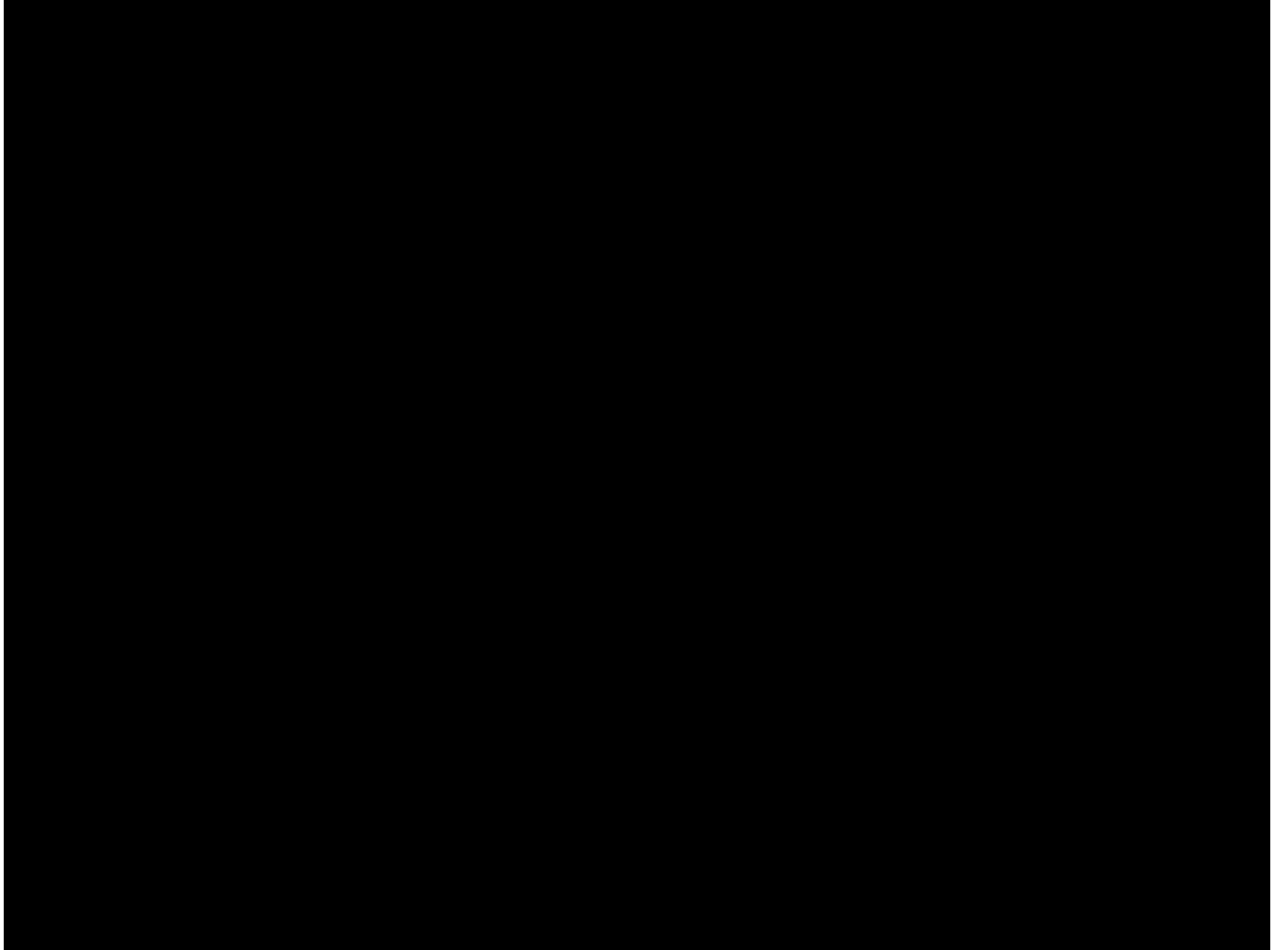
Credits

Art and graphics:
Sebastian Curi and Koto Studio

Appendix

(undetected) Latent Theme

"When my family asks how I'm affording life when I have no job"



Themes

			most common word/emoji			
rank	**theme**	**count of notes**	**1**	**2**	**3**	**4**
1	Other	70573	(services, 108)	(head, 96)	(din, 84)	(sale, 71)
2	Domestic	21466	(house, 1208)	(car, 791)	(love, 770)	(baby, 764)
3	Gratitude	14830	(thanks, 1533)	(good, 583)	(know, 420)	(day, 385)
4	Tip	14599	(tip, 518)	(nail, 503)	(face, 502)	(party, 472)
5	Stuff	12871	(stuff, 1676)	(wifi, 458)	(lol, 437)	(ass, 366)
6	Food	12744	(lunch, 1111)	(dinner, 992)	(chicken, 720)	(steaming_bowl, 469)
7	Face	11362	(face, 6417)	(person, 1030)	(dog, 960)	(getting, 888)
8	Domestic	8597	(money, 3478)	(house, 3326)	(deposit, 638)	(garden, 631)
9	Face	7694	(face, 10153)	(smiling, 4467)	(eyes, 2166)	(kiss, 1671)
10	Food	6611	(food, 6414)	(foods, 246)	(chinese, 96)	(thai, 72)
11	People	6339	(skin, 6299)	(tone, 6298)	(light, 4718)	(medium, 3583)
12	Gifts	5752	(gift, 3455)	(birthday, 1121)	(christmas, 823)	(tree, 561)
13	Gratitude	5107	(heart, 4526)	(red, 2294)	(hearts, 856)	(suit, 305)
14	Alcoholic Drinks	5047	(glass, 1849)	(wine, 1201)	(beer_mugs, 1158)	(clinking, 1158)
15	Travel	4274	(uber, 1687)	(hotel, 777)	(taxi, 736)	(lyft, 321)
16	Food	4082	(taco, 2024)	(sushi, 1096)	(hamburger, 890)	(tacos, 197)
17	Clothes	4052	(shirt, 1425)	(clothes, 534)	(shoe, 457)	(woman’s, 373)

Themes (page 2)

			most common word/emoji			
rank	theme	count of notes	1	2	3	4
18	Car	3673	(automobile, 1705)	(gas, 1197)	(electric, 543)	(car, 172)
19	Sports	3404	(baseball, 550)	(basketball, 521)	(soccer, 442)	(ball, 402)
20	Groceries	3380	(groceries, 1304)	(drinks, 446)	(shopping, 316)	(supplies, 263)
21	Food	3338	(food, 820)	(bread, 557)	(face, 535)	(meat, 438)
22	Food	3275	(pizza, 3274)	(thanks, 7)	(team, 2)	(points, 1)
23	Gratitude	3248	(thank, 3207)	(heart, 172)	(face, 118)	(red, 82)
24	Rent	3160	(rent, 3033)	(house, 135)	(rental, 130)	(july, 129)
25	Food	2861	(green, 497)	(cheese, 480)	(wedge, 456)	(bread, 410)
26	Phone	2618	(phone, 2350)	(mobile, 1531)	(money, 917)	(telephone, 353)
27	Tickets	2572	(tickets, 1438)	(ticket, 1151)	(admission, 736)	(tix, 204)
28	Money	2500	(dollar, 2827)	(banknote, 1693)	(venmo, 1142)	(arrow, 353)
29	Party	2429	(festival, 590)	(trip, 491)	(floaty, 303)	(camping, 286)
30	Coffee	2351	(coffee, 2486)	(cooking, 177)	(tea, 165)	(shake, 28)
31	People	2309	(baby, 762)	(woman, 633)	(face, 375)	(man, 355)
32	Food	2259	(knife, 2333)	(fork, 2212)	(plate, 540)	(kitchen, 122)
33	Sports	2176	(football, 2347)	(american, 2020)	(fantasy, 482)	(basketball, 377)
34	People	2136	(skin, 4718)	(tone, 4718)	(light, 3588)	(medium, 2818)
35	Cookies	1946	(cookie, 1267)	(cookies, 456)	(rocket, 265)	(popsicle, 234)

Themes (page 3)

			most common word/emoji			
rank	**theme**	**count of notes**	**1**	**2**	**3**	**4**
36	Gas	1849	(fuel, 1931)	(pump, 1928)	(gas, 81)	(automobile, 59)
37	Alcoholic Drinks	1676	(beer, 1836)	(mug, 1808)	(pizza, 75)	(hamburger, 54)
38	Money	1623	(money, 1853)	(bag, 831)	(computer, 486)	(laptop, 479)
39	Beach	1515	(fish, 357)	(tropical, 239)	(tree, 189)	(sun, 171)
40	Birthday	1383	(party, 1496)	(popper, 1430)	(birthday, 1093)	(cake, 824)
41	Bills	1324	(bills, 1281)	(splitwisecom, 507)	(states, 73)	(united, 73)
42	Sweat	1190	(sweat, 1390)	(droplets, 1385)	(electric, 482)	(plug, 481)
43	Tickets	1099	(camera, 1027)	(movie, 742)	(flash, 130)	(film, 108)
44	Childcare	1093	(babysitting, 610)	(tutoring, 291)	(daycare, 145)	(school, 26)
45	Gratitude	1040	(heart, 2925)	(red, 1576)	(purple, 222)	(blue, 199)
46	Alcoholic Drinks	1023	(glass, 1303)	(wine, 742)	(bottle, 702)	(cork, 681)
47	Games	1001	(game, 689)	(video, 236)	(musical, 155)	(die, 107)
48	Food	924	(pizza, 917)	(french_fries, 490)	(hamburger, 432)	(doughnut, 64)
49	Food	899	(pizza, 1234)	(face, 38)	(party, 35)	(chicken_drumstick, 26)
50	Food	885	(burrito, 999)	(taco, 74)	(face, 14)	(drink, 11)
51	Face	807	(face, 2890)	(smiling, 1281)	(eyes, 725)	(middle, 501)
52	Bills	784	(voltage, 836)	(high, 836)	(electric, 724)	(light_bulb, 690)
53	Food	714	(taco, 1012)	(pizza, 937)	(hamburger, 335)	(sushi, 265)

Popular Emojis

rank	emoji	demoji	count	rank	emoji	demoji	count	rank	emoji	demoji	count
1	🍕	:pizza:	5813	16	😘	:face_blowing_a_kiss:	2076	31	🍣	:sushi:	1377
2	💸	:money_with_wings:	5681	17	🍔	:hamburger:	2034	32	🔌	:electric_plug:	1327
3	❤️	:red_heart:	4640	18	⛽	:fuel_pump:	2023	33	🍟	:french_fries:	1326
4	🏻	:light_skin_tone:	4367	19	🍻	:clinking_beer_mugs:	1924	34	💕	:two_hearts:	1287
5	🏼	:medium_light_ skin_tone:	4178	20	🏽	:medium_skin_tone:	1800	35	💇	:person_getting_ haircut:	1275
6	🏠	:house:	3995	21	🍴	:fork_and_knife:	1760	36	🌯	:burrito:	1256
7	🌮	:taco:	3053	22	💵	:dollar_banknote:	1667	37	🍾	:bottle_with_ popping_cork:	1224
8	♀	:female_sign:	2901	23	🍹	:tropical_drink:	1662	38	🐶	:dog_face:	1183
9	🎁	:wrapped_gift:	2669	24	🚗	:automobile:	1652	39	🍳	:cooking:	1156
10	🎉	:party_popper:	2599	25	💦	:sweat_droplets:	1497	40	🎂	:birthday_cake:	1155
11	☕	:hot_beverage:	2456	26	📱	:mobile_phone:	1478	41	💃	:woman_dancing:	1074
12	🍺	:beer_mug:	2311	27	🍗	:poultry_leg:	1465	42	⚡	:high_voltage:	1071
13	🏈	:american_football:	2163	28	👶	:baby:	1458	43	👕	:t-shirt:	1068
14	🔥	:fire:	2123	29	♂	:male_sign:	1436	44	🍆	:eggplant:	1029
15	🍷	:wine_glass:	2091	30	🍪	:cookie:	1391	45	🏀	:basketball:	1007

*Note that counts will include multiple instances that occur within a single note
** Skin tones shown are used to modify the color of certain emojis

Word Embedding Adjustments

```
    #custom definition for interprebility against Google training set
    text = text.replace('hot_dog', 'hotdog')
    text = text.replace('hot_beverage', 'coffee') #this
    text = text.replace('sport_utility_vehicle', 'SUV')
    text = text.replace('shallow_pan_of_food', 'paella')
https://emojipedia.org/shallow-pan-of-food/
    text = text.replace('cup_of_water', 'soda')
    text = text.replace('money_with_wings', 'money')
    text = text.replace('wrapped_gift', 'gift')

    #remove punctuations, especially key to get rid of semicolons and
underscores
    text = re.sub('[\.\,\!\?\:\;\-\=\_\'\%\(\)\¯\\\¯]', ' ', text)

    #add back underscore for connected words/emojis that Google data can
understand
    text = text.replace('french fries', 'french_fries')
    text = text.replace('bento box', 'bento_box')
    text = text.replace('light bulb', 'light_bulb')
    text = text.replace('beer mugs', 'beer_mugs')
    text = text.replace('clinking glasses', 'clinking_glasses')
    text = text.replace('steaming bowl', 'steaming_bowl')
    text = text.replace('fried shrimp', 'fried_shrimp')
    text = text.replace('santa claus', 'santa_claus')
    text = text.replace('palm tree', 'palm_tree')
    text = text.replace('cherry blossom', 'cherry_blossom')
```

```
#custom definition for interprebility (done after the fact because of
the underscore)
text = text.replace('poultry leg', 'chicken_drumstick')

#custom stop words
text = text.replace('oncoming', '') #oncoming_automobile
text = text.replace('selector', '') #airplane_selector

#other notes:
# dog_face and pig_face will generate the terms: dog, pig, face.
leaving this as is to get the face
# ok_hand will become ok, hand. Then ok will be removed in later
vectorizer option for only 3 letter words.
# other strange emoji definitions: telephone_receiver,
zany_hearts, rocket_popsicle. last of which is a custom venmo
made for festivals
```

Word Embedding Replacements

'airbnb':'guesthouse',
'bnb':'guesthouse',
'doughnut':'donut',
'dranks':'drinks',
'facepalming':'facepalm',
'ibotta': "rebate", #ibotta is a company that gives cashback rewards through venmo.
'uber':'cab', #google_vec_file won't treat uber correctly
'lyft':'taxi_cab', #google_vec_file won't treat lyft correctly
'pge':'electric_utilities',
'splitwise':'split_evenly',
'splitwisecom':'split_evenly',
'spotify':'iTunes', #iTunes is the old version of spotify
'venmo':'PayPal', #paypal was venmo before venmo was venmo (and now currently the parent company)

Model Exploration

Explored

Demojize:

- Words + Emoji
- Words + Emoji + Demoji
- Words + Demoji

Vectorizer:

- CountVectorizer
- TF-IDF Model

Dimensionality Reduction

- LSA
- NMF
- LDA

Champion

Demojize:

- Words + Demoji

Initial Document Vectorizer:

- CountVectorizer

Transformer:

Word (+ Demoji) Embeddings:

- Google-news-300 (2013)

Gensim:

- Word2Vec

Clustering:

- KMeans

Venmo Data Fields

Transaction
story_id
payment_id
date_completed
date_created
date_updated
payment_type
amount
audience
status
note
device_used
actor_user
target_user

User
id:
username
first_name
last_name
display_name
phone
profile_picture_url
about
joined
is_group
is_active