

Group Members: Jacob Bissonette, Austin Franklin, Yuriy Kamenivskyy, Garrett McMerriman, Bill Hoang Pham
Case Study 1

1. What data you collected?
2. Why this topic is interesting or important to you? (Motivations)
3. How did you analyze the data?
4. What did you find in the data?

Problem 1 & 2

For the first two problems we collected 300 tweets containing the word “nutrition” as that was the focus of our inquiry. We chose “Nutrition” as our query keyword because we wanted to find a topic that relates to all of us and would be interesting and useful to explore. We all consider ourselves to be healthy and health-conscious individuals, and also one of our team members happens to work on an MQP which centers around that topic too. Additionally, we hypothesized that this query may provide us with high quality data, and only a small amount of “bad data”. First of all, it is quite niche, and thus we may encounter a lesser number of bots than if we were researching a broader topic. Second of all, we assumed that individuals that are concerned with nutrition and want to post about it on Twitter to popularize their lifestyle, are likely to maintain proper punctuation and grammar in their tweets.

After collecting 300 tweets that contained the keyword ‘nutrition’, the first thing we did was determine the 30 most common words that appeared in these tweets along with our keyword. Figure 1 shows the full list of those thirty keywords.

Word	Count		
nutrition	140	tell	19
i	39	lives	19
food	36	hype	19
when	27	info	19
health	27	oa	19
up	27	have	16
#nutrition	26	amp	16
some	23	how	16
can	23	our	15
heres	21	more	15
sometimes	20	new	15
about	19	not	15
trendyand	19	get	14
trendingits	19	out	14
hard	19	we	14

Figure 1: 30 most frequent words mentioned in tweets with ‘nutrition’

Now, it is important to mention that this result came after several iterations of gradually preventing stopwords from appearing in our final table. We removed words such as “and”, “of”, “on”, “in”, “at”, “then”, etc. Thus, we found that the most frequently occurring and semantically important words were “nutrition”, “I”, “food”, “when”, and “health”. We presume that all of these are related. Nutrition is obviously directly correlated with a person’s health

and wellbeing (“health”), and is a result of proper dietary habits (“food”). The frequent occurrence of the word “I” is first of all, a direct result of the fact that a person’s dietary habits are completely in their control. In a day and age where society is not necessarily the healthiest, a person’s diet is something that is very fixable and comes down to the individual. Secondly, we assume that frequent occurrence of “I” suggests that these tweets may be peoples’ New Year’s resolutions e.g. “I decided to improve my nutrition this year. It’s great for your health.” It is important to mention here that since the keyword ‘nutrition’ itself only came up 140 times in those 300 tweets, this suggests that the other 160 tweets that were collected by our query were probably published by users with “nutrition” in their Twitter name (since if the keyword wasn’t in text, it had to be in the username, to be collected by our query call).

After finishing our inquiry into the most common words we looked for the most famous retweeted tweets. Unsurprisingly, most of the famous retweeted tweets, see Figure 2, all had at least one of the top four most common keywords mentioned in them (not counting the word ‘nutrition’ itself).

Count	Screen Name	Text
31	adityadhimas48	Is nutrition strength. Words could bite. When I look at And imagining comfortable position USE: QQQ : PP https://t.co/JINAwVSajy
9	MotivationQF	One of the top super .. my favorite .. Ginger benefits your health in so many ways #Nutrition #Superfood... https://t.co/dPjwMYXY0b
8	_Georgeobeid	High-Fiber Diet Improve the Response of Melanoma Skin Cancer Patients to Immunotherapy #Cancer #Die601 #Die6037... https://t.co/SupSwbRh37
5	thehealthb0t	According to the Department of Nutrition for Health and Development, iodine deficiency is a public health problem in 54 countries.
5	TeeAirRa_	Food allergies need to be discussed more when we talk about diet and nutrition. Eating food that you're unknowingly... https://t.co/Ky1IG70iiw
3	CNS_SCN	#CNSTHEMATIC22 “Advances in Nutrition and Immune Function” is NOW LIVE! Starting the day off is our conference co-h... https://t.co/mwTEgJbjnQ
2	tariqtramboo	Gout: What to eat ? •Vegetables: All vegetables are fine, including potatoes, peas, mushrooms, eggplants and dark... https://t.co/g0F0SHAKpU
2	healthstepsonly	#smoothie #Food #Diet #dietplan #nutrition #Weight #weightloss #weightlosstips #loseweight... https://t.co/Q529CFzPjX
2	NHLBI_Translate	.@NIHgrants: #NIH awards \$170 million for precision #nutrition study. The Nutrition for #Precision Health study lev... https://t.co/tgZreadQb9
2	MyHealthy_Reach	Show up everyday, put in the work and the result will show. Make this your fitness guide for 2022. Visit for more:.. https://t.co/Za0YoVHYv0

Figure 2: Most Retweeted Tweets

The most retweeted tweets were giving out suggestions for meal plans or educating about the benefits of nutrition for different sorts of health problems. Next, after reviewing the top retweeted tweets, we looked for the top 10 hashtags. We again see the recurring theme of things like “mealprep”, “nutrition”, “food”, “diet”, “health”, and “weightloss”. People want to learn how to eat well, and lose weight with obesity being an issue nowadays. For the user mentions, it was mostly just smaller accounts who were into health and fitness. We found that “nutrition”, “food”, and “diet” were the top 3 hashtag words from our collection of tweets. We found that leoubb, strong_science, and hybrid training were the top 3 user mentions from our collection of tweets. See Figure 3 for hashtag and user mention data.

Word	Count	Word	Count
#nutrition	26	@leoubb	3
#food	10	@strong_science	3
#diet	8	@hybridtraining	3
#dietplan	6	@mondaynightibd	2
#health	5	@dcharabaty	2
#meal	5	@janssenus	2
#mealprep	5	@pfizer	2
#healthy	4	@takedapharma	2
#foodie	3	@fao	2
#weightloss	3	@1bestgrandma	2

Figure 3: Most frequent hashtags and most popular user mentions

Problem 3

Once our nutrition analysis was completed, we then gathered data on a popular Twitter user, being Elon Musk. We decided to choose Elon Musk who is a scientist who is a very active Twitter user and has a very large following. We decided to pair him up with Lex Fridman, a very famous computer scientist, who was recently on a podcast with Elon Musk. We collected usernames and IDs of Twitter accounts that were following Musk's account, were in the account's friends list, and then both following and friends list.

User ID	Username
1485726534740717576	BramVangeneugd3
1485727488437276676	Roman20233765
1485724948912029703	CheilaS48187958
1485725741488689166	Sarfrashawani
1485727191925149702	kazakRealNFT
1485727765626298377	YouTubeCentral9
1485727693442273280	Alonso30327900
1484989604923117576	Its_Only_Chris
1483325028586426370	odetesa97803732
1485726666781564928	o_komnatskyi
1485727660516990981	12Hazper
1485727522901966848	Mush35G
1485727569370664969	SeanCam65693861
1485726470354120704	HariSupercar
1420810161644523520	kamrynlowrey
321641382	TCastillo85
70906623	Taiye12
1485727761968861185	LosPason1
1485727251522019329	Djordje_Bulldog
1485727297860685832	wafleMET

Figure 4: Followers of Elon Musk

User ID	Username
76980293	moxie
958324785229934593	HGJart
1720046887	ilyasut
2997418453	EuropaClipper
14486697	paraga
4914384040	TheBabylonBee
30699048	BillyM2k
177101260	Rainmaker1973
1346194631982247943	inspiration4x
1402156691186241538	CNSAWatcher
3669881952	gigadgets_
1068831	slashdot
1359595748942372867	totalspace360
822825713615175680	redditSpacePorn
1279551125264445448	rookisaacman
16178922	TheMarsSociety
1333467482	CoinDesk
2445883045	dogecoin_devs
2235729541	dogecoin
17060015	cleantechnica

Figure 5: Friends of Elon Musk

When plotting the followers and their IDs of Elon Musk, we noticed that some long-time users of Twitter just started following Elon Musk. We also found that Elon Musk has friends that are also long-time Twitter users. When coding Elon Musk's friends who are also followers, we came up with an empty table. This is due to the fact that we can only gather 5000 users out of his 71.1 million followers and only 107 of them are friends. It would be

nearly impossible to find the followers who are also friends of Elon Musk. Then we decided to find the mutual friends of Elon Musk and another Twitter user named Lex Fridman. We found that Elon Musk and Lex Fridman are both friends with Twitter users, fermatslibrary, and engineers_feed.

Problem 4

For Problem 4 we decided to act as hired data scientists for Netflix and Riot Games. We were hired to determine whether to release the second season of “Arcane” TV-show, and, in case if the series is to be continued, determine any supplemental information that may assist writers or marketing department in ensuring that season two becomes a top-ranked show on Netflix.

In order to determine whether Arcane was a successful show we first collected all the tweets that mentioned the word “Arcane” and had a negative attitude: query = “Arcane :(“. This was a starting point for our research to determine whether, on average, Twitter users liked the series or not. After finding all the “negative attitude” tweets about Arcane in the past week, we searched for 30 most frequent terms that appeared in those tweets. This would help us determine two things: 1) whether people had strong negative emotions towards Arcane and disliked the show (this would determine whether the show should be continued), and 2) what caused most negative emotions about the show (this would help creators understand what to avoid / include in the next season of the show).

Word	Count		
arcane	85	httpstcoxjf0mngial	13
i	66	vi	12
im	48	@beauty_fer2	12
sad	41	hope	9
watching	36	just	8
now	35	dont	8
finished	33	they	8
@thedustyleaves	32	his	8
httpstcorxyrqpqsf	32	3	8
so	22	will	7
#arcane	21	see	7
them	18	was	7
have	15	her	7
me	14	@naeris_arcane	7
#arcanefanart	14	today	6

Figure 6: 30 most frequent words mentioned in tweets with “Arcane”

The first positive was that most “negative attitude” tweets were marked as such because the users were “sad” after watching the series (since “finished” appeared). Thus, we could presume that many users were sad because they finished the series, which was a good signal. However, there remained a possibility that they were sad because of the content they saw which may have been good or bad e.g. they were sad because content didn’t meet their expectation, or they were sad because of events happening in the series, as intended by writers. To understand whether that sadness was “positive” or not we collected the most retweeted and liked tweets and analyzed them.

Count	Screen Name	Text	Count	Screen Name	Text
12	beauty_fer2	them :(#arcane #arcaneafanart https://t.co/XJf0mNgIAL	161	beauty_fer2	them :(#arcane #arcaneafanart https://t.co/XJf0mNgIAL
7	Luc_Arcane	Shattered Light and Shadow. page 194-196 (I made mistake in counting pages again :() #UndertaleAU #UndertaleSLAS...	155	Luc_Arcane	Shattered Light and Shadow. page 194-196 (I made mistake in counting pages again :() #UndertaleAU #UndertaleSLAS...
3	silcothinker	https://t.co/WDhpUBvWAt rllly want to rewatch arcane for them but then i'll have to see the whole of ep 3 and ep 9 again :(93	silcothinker	https://t.co/WDhpUBvWAt rllly want to rewatch arcane for them but then i'll have to see the whole of ep 3 and ep 9 again :(
2	Yumecha36	#silco... https://t.co/8FNdGDrQY5 I miss Silco :((((ALSO, I WAS LOOKING AT HIS ARMS TODAY AND THEY R PERFECT FOR HUGGING??? His hugs must be so goo... https://t.co/tvRMGNTfLs	74	Midnight_Tigerx	@IGotCaughtAgain @papuchochoe LOL fr I can't wait to see his full transformation in arcane and everyone's gonna be like >:((
1	roronoa_mano	episode 3 of arcane is no joke :(56	naeris_arcane	I want to draw Caitvi and i already have like two ideas but finals won't let me :(
1	rinexart	jumping in to a new fandom will never be easy but damn, arcane people are just so so good :(47	naeris_arcane	I'm feeling a little sad today, i wish Vi gave me a hug :(
1	eymlsanjr	dopil set and downoon arcane just got home! 🍷 grabe po yung spark sa poca pero mas grabe yung spark sa wonpil pola f... https://t.co/3UrU9SbdnH	20	piltovertxts	@UnStroy i know :(real sad that it's only a skin line for LoR, they should've added this in league as well but i g... https://t.co/yZzSLmASH1
1	arcane_she	@lamponirossi i don't wanna see hailee doing drugs and smoke cigarettes :(19	rinexart	jumping in to a new fandom will never be easy but damn, arcane people are just so so good :(
1	BroCmon4	I need the 1.5k vbucks for arcane vi :(18	naeris_arcane	I'm so tired :(
1	BroCmon4	@ShiinaBR May ik about the myna pack tho? Sorry guys don't flame me it's my only way to get the arcane vi skin :(14	roronoa_mano	episode 3 of arcane is no joke :(

Figure 7: Most popular tweets with negative attitude about “Arcane”

As it turned out, most users were sad due to some romantic relationship between “them” and they particularly reacted to episodes 3 and 9. This was a good sign, and suggested that the second season should also include a melodramatic romantic storyline to grasp users' attention. Finally, we decided to make a cross section between followers of the League of Legends (LoL) Twitter account and followers of the Netflix Twitter account. These Twitter users would be the prime target market for advertising season 2 of the show. Since, they are interested in both LoL (the game whose content this show was based on) and Netflix (platform which aired the tv-show).

User ID	Username	User ID	Username
1485714614453415936	tftnz_	1485710060894539778	KasieCullen2
1485714246474375169	CondeescuA	1485726139645575174	SexyQuan09
1485695527035486209	Handsomecarrot7	1485697752118005768	mamakendoll
1485693865789116417	__noa31__	1401905264601747459	bremildo1
1485695062281388033	feedthymeme	1485697970926493698	EthanNickyyy
1485692320934764549	Ryan709685621	1485722874992369667	Nebo3Mark
1485725740503113734	Miah80461049	1485722191585062914	genocide458
1485704367235862534	coochslayerr45	1485722151231569923	Armandojo08
1485707895811489798	m0nalise	1485706379763798016	sldvplandr
1485704412320444420	rsyao1	1485717529700048901	9rnfla
1485692638162567170	myaz_Phillly	1485694785298079749	Jojo82576530
1001437240147431425	440Berry	1485721904220712965	daku_rairakku
1485719394886033420	JayzinYt	1485713055908610054	lilian_tunnell
		1485701774946320392	_rafabuassali
		1220668400319352833	KVsyhmyHEyMYDf7
		1485697012448182273	VanessaEmery11

Figure 8: Twitter users that follow both LoL and Netflix Twitter accounts