



UChicago

Social Media Trend Analysis

Big Data Final Project
Mohd Kamaludin - 6/9/2018

Executive Summary

1. **Tweets** from University of Chicago, as well as other 3 other major universities in Chicago, namely the Northwestern University, Loyola Chicago, and University of Illinois-Chicago, are **carefully filtered from the collection to separate noise** from useful tweets for analysis
2. Size of original tweets and retweets activities are **compared to determine effects of time** and day with user publishing patterns. In addition, tweets are classified by **mobile devices** such as ios and android to determine influences of platform with **tweeting behavior**
3. Word cloud is used to **visualized most used words** and predict possible tweet narrative
4. UChicago official accounts (i.e. @UChicago, @ViewChicago, @UChicagoCollege, @HarrisPolicy, @UChicagoMed @BoothFullTime) have strong followers as well as retweeted activities. However, further improvement can be made to **improve audience engagement**

Tweets 17K Following 1,465 Followers 46.2K Likes 2,911 Lists 9 Moments 6

Tweets Tweets & replies Media

Pinned Tweet
UChicago @UChicago · 6h #UChicago2018 Convocation twitter.com/i/moments/1005...

#UChicago

#UChicago2018 Convocation
UChicago @UChicago
Graduation weekend is finally here, follow along and contribute your experience with #UChicago2018

Moments

UChicago Retweeted
The Nobel Prize @NobelPrize · 9h "Newton had his apple - and I had my cashews!"

Economics Laureate @R_Thaler discusses how a bowl of cashews inspired his early research - and much more - in @UChicago's Big Brains podcast. Listen here:

Big Brains podcast welcomes Nobel-winning economist Richard Thaler
news.uchicago.edu

UChicago @UChicago · 3h Gathering on the Main Quadrangle to celebrate, #UChicago2018 graduates receive advice on future leadership. Convocation Day Recap: ms.spr.ly/6014rNnxg #UChicago

Methodology

1. 70,439,383 tweets is extracted for the period between 6/22/2017 - 6/3/2018. This collection is then narrowed down using careful combination of keywords of unique university names and Chicago to create a subset of data for each of the selected Universities, and avoiding:
 - a. UofCincy
 - b. Northwestern State University
 - c. Loyola Maryland
2. Feature Engineering
 - a. Mobile platform users can be identified with iOS and Android keywords from ‘source’
 - b. **Retweet Indicator** is created by identifying tweets starting with “RT @%”
 - c. **Original authors** of the tweets can be identified between ‘RT @’ and ‘.’ characters
 - d. Combined tweets and **expanded tweets** using ‘Truncated’ = True flag
3. Word frequencies are generated using Natural Language Toolkit package
 - a. **Stop words are removed**
 - b. Words less than 3 characters long are excluded

University	Variation of Keywords	# of Tweets
UChicago	uchicago, uofc_, uofchicago, chicagobooth	188,178
Northwestern	northwesternu_, northwestern university, kellogg	97,541
Loyola Chicago	loyola chicago, quinlan business	370,329
UIC	uic, gies business, uofi chicago	91,701

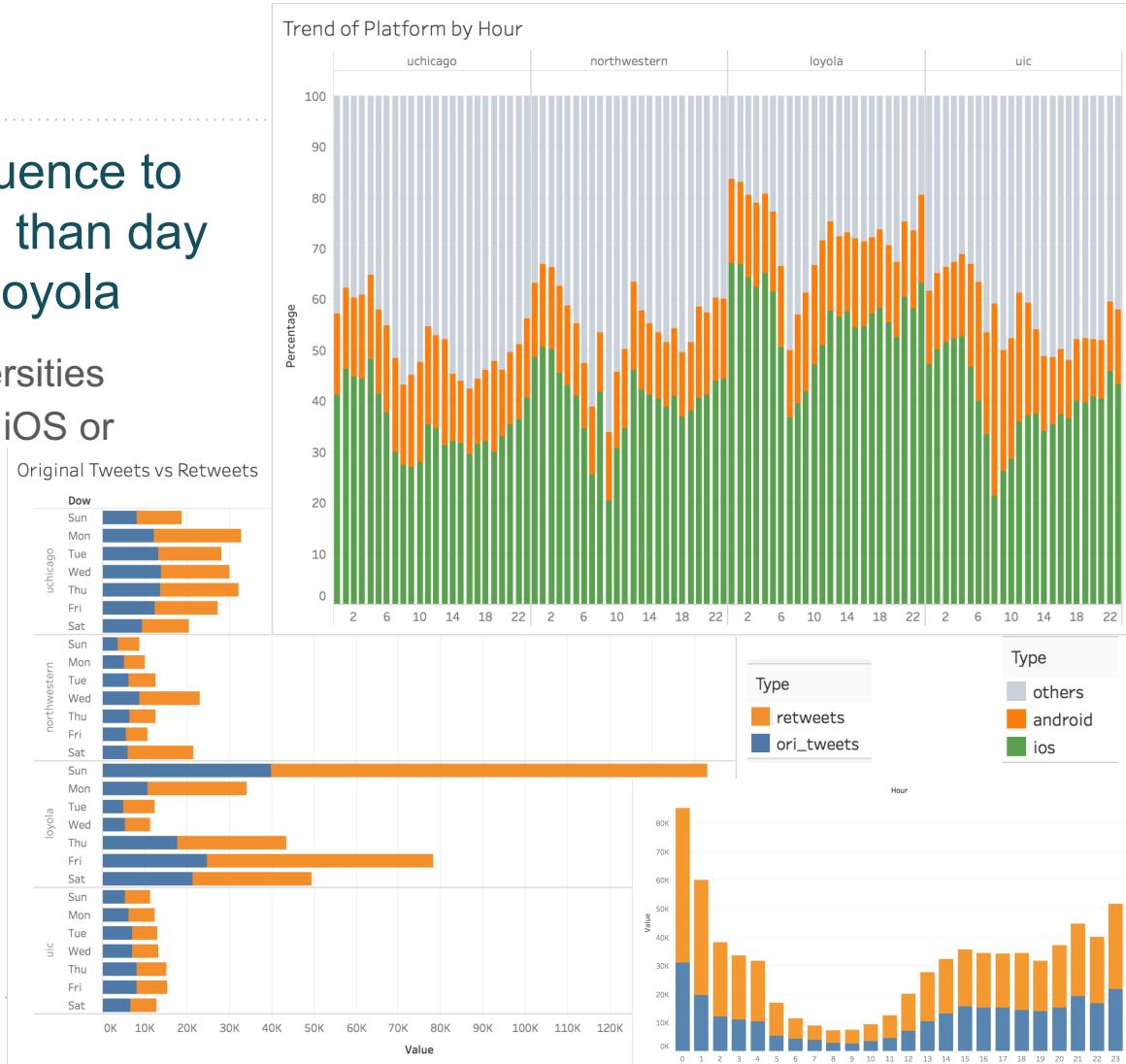
User Behavior Trend

Time of the day has a stronger influence to the number of tweets and retweets than day of the week with the exception of Loyola

1. More than half of the users in all Universities data sets are mobile users either using iOS or Android devices

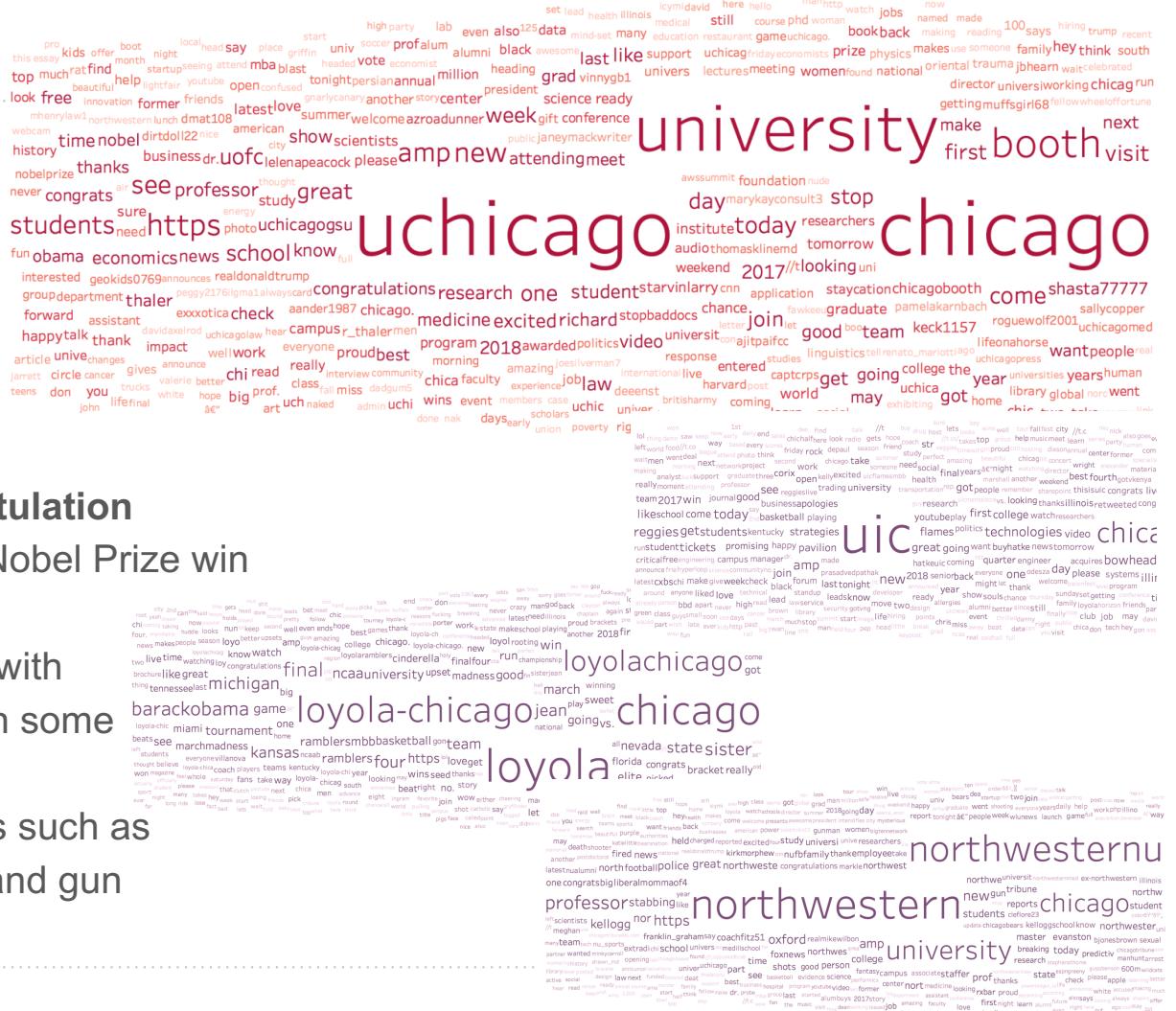
- a. More **retweets** are occurring in **late evenings and early morning**
- b. More retweeters are on mobile

2. UChicago and Loyola seem to have **more active following** as **retweets activities** are more than half of overall tweet footprints



Word Frequencies

Significant events -- such as major sports development, breaking news, International Awards -- influence the choice of words in the cloud



Influencer

Accounts with high followers
are likelier to be retweeted

1. UChicago has a lot of tweeter presence when compared to

Other schools:

- a. @UChicago, the official UChicago twitter account, **have the most total tweets and followers**

- b. @UChicagoGSU, the UChicago Graduate Student United student body also has a huge visibility with a **high number of tweets and also retweeted tweets**

- c. @ChicagoMaroons, the official UChicago Athletics news, also has a sizeable following

2. For other schools:

- a. @ResearchNU, Northwestern University Research voice leads for Northwestern followers
- b. @UICnews leads among University of Illinois Chicago published tweets with a big followers

university	rank	screen_nm	ori_tweet	retweets	total	likes_cnt	followers	following	loc
uchicago	0	uchicagogsu	447	574	1021	2703	2573	1198	Chicago, IL 60637
uchicago	1	UChicago	696	192	888	2760	46085	1463	Chicago, Illinois
uchicago	2	ChicagoMaroons	550	103	653	722	3128	193	Chicago
uchicago	3	LyoLouisJacques	35	469	504	1124	1475	2753	Chicago, Illinois
uchicago	4	haileysecretary	350	17	367	NaN	NaN	NaN	Ft. Lauderdale Fl
uchicago	5	DanW_BCI	344	0	344	NaN	NaN	NaN	Nevada City, CA
northwestern	0	ISENatNU	277	27	304	4305	1582	2255	Evanston, IL
northwestern	1	GamerGeekNews	266	1	267	3670	5095	5556	None
northwestern	2	ResearchNU	86	120	206	3829	2803	494	Evanston, IL
northwestern	3	NotreDameSpyder	170	0	170	0	427	0	None
northwestern	4	LIVE_COVERAGE	1	161	162	405528	8717	9580	NYC↔#LongIsland-Atlantic~~
northwestern	5	NorthwesternSS	139	0	139	0	4	0	None
loyola	0	tmj_il_vets	966	0	966	0	203	119	Illinois Non-Metro
loyola	1	azim382838	453	0	453	NaN	NaN	NaN	Bangladesh
loyola	2	KuntrinaPierson	349	0	349	NaN	NaN	NaN	None
loyola	3	loyolaforus	250	64	314	7127	680	1354	Chicago, IL
loyola	4	RUWTncaab	267	0	267	NaN	NaN	NaN	None
loyola	5	BillDaley	250	0	250	3111	26344	23240	Chicago
uic	0	illinivote2016	1	468	469	20133	1170	5003	Illinois, USA
uic	1	UICnews	217	163	380	1522	8976	247	Chicago, IL
uic	2	bfloodUIC	286	39	325	832	777	1015	Chicago
uic	3	UICFlamesMBB	248	2	250	560	2445	444	Chicago
uic	4	VoteForDoctorY	14	235	249	2889	239	1872	Illinois, USA
uic	5	dabyronicred	210	38	248	1529	508	5002	Chicago

*Table: Number of accounts with the highest number of tweets, and their followers count

Conclusion

Recommendation

1. UChicago official accounts have strong followers, and paying **closer attention to current events** can help shapes the social media narrative for UChicago. For example, **any positive events can be boosted** even more with the help of other influencers such as Barack Obama for Loyola. On the other hand, negative events such as active shooting in Northwestern can also be quickly spread
2. Careful **timing of tweet** as publishing late in the evening can encourage more retweets and improve reach
3. **Specialized topic** such as Sports and Research should have their own twitter handle as more users will be interested to subscribe
4. Audience are primarily on mobile devices, **shorter tweets** may be better

university	rank	ori_author	retweets	ios	android	others
uchicago	1	RT @DJBooth	5971	71.998	20.700	7.302
uchicago	2	RT @UChicago	5375	45.637	16.744	37.619
uchicago	3	RT @NobelPrize	4040	37.005	30.223	32.772
uchicago	4	RT @uchicagogsu	3511	47.536	20.649	31.814
uchicago	5	RT @sofiorden	2961	0.000	0.034	99.966
uchicago	6	RT @davidaxelrod	2245	52.695	25.078	22.227
uchicago	7	RT @haileyssecretary	1840	0.054	0.217	99.728
northwestern	1	RT @chicagotribune	13523	63.403	19.692	16.905
northwestern	2	RT @ScholarshipsGL	908	1.542	4.185	94.273
northwestern	3	RT @CBSNews	847	49.941	23.613	26.446
northwestern	4	RT @EmmWald	769	55.267	25.098	19.636
northwestern	5	RT @eb4prez	758	76.121	15.567	8.311
northwestern	6	RT @ChicagoSports	747	53.949	11.914	34.137
northwestern	7	RT @AAIHS	657	35.160	23.744	41.096
loyola	1	RT @marchmadness	19458	73.995	17.139	8.865
loyola	2	RT @BleacherReport	16860	82.325	11.833	5.842
loyola	3	RT @SportsCenter	15728	81.409	13.149	5.443
loyola	4	RT @BarackObama	11019	64.516	24.676	10.809
loyola	5	RT @JeffBarker_	7556	83.549	12.401	4.050
loyola	6	RT @darrenrovell	7425	70.465	17.630	11.906
loyola	7	RT @ESPNStatsInfo	5776	69.235	19.823	10.942
uic	1	RT @illinivote2016	1122	45.455	2.050	52.496
uic	2	RT @ACLUofIL	992	56.351	23.085	20.565
uic	3	RT @jeremymcinnon	742	63.342	31.941	4.717
uic	4	RT @chicagotribune	584	53.082	26.370	20.548
uic	5	RT @anshumanjohriAJ	533	3.940	85.553	10.507
uic	6	RT @uic_ph	497	44.869	45.674	9.457
uic	7	RT @bfloodUIC	445	29.888	6.067	64.045

*Table: Accounts with the most retweeted tweets and their retweeters' platform of choices

Future Work

1. Further analysis can be performed to determine the similarities of tweets using **SinHash** or Jaccard, as well as **identify better topic tweets**
2. Tweets metric such as **likes, replies, and retweets** can be incorporated for deeper analysis of **influencers' engagement with their followers**