Tara Paranjpe
CSE 472 Project 1
Platform: Reddit
Project Report

An Analysis of Top Authors of the ASU Subreddit and Their Subreddit Interests

---

Things to Note for Script Execution:
- **Please provide username, password, clientID, and client secret in the script (for auth request and reddit instance).**
- **Please note I have given credit to a github code example for retrieving the Reddit Auth Token.**
- **When executing script, if execution pauses and no other output shows on the terminal after 3-5 min and after authorization token information is output, please terminate and restart the script. Script execution gets hung randomly.**

  Thank you for your patience.

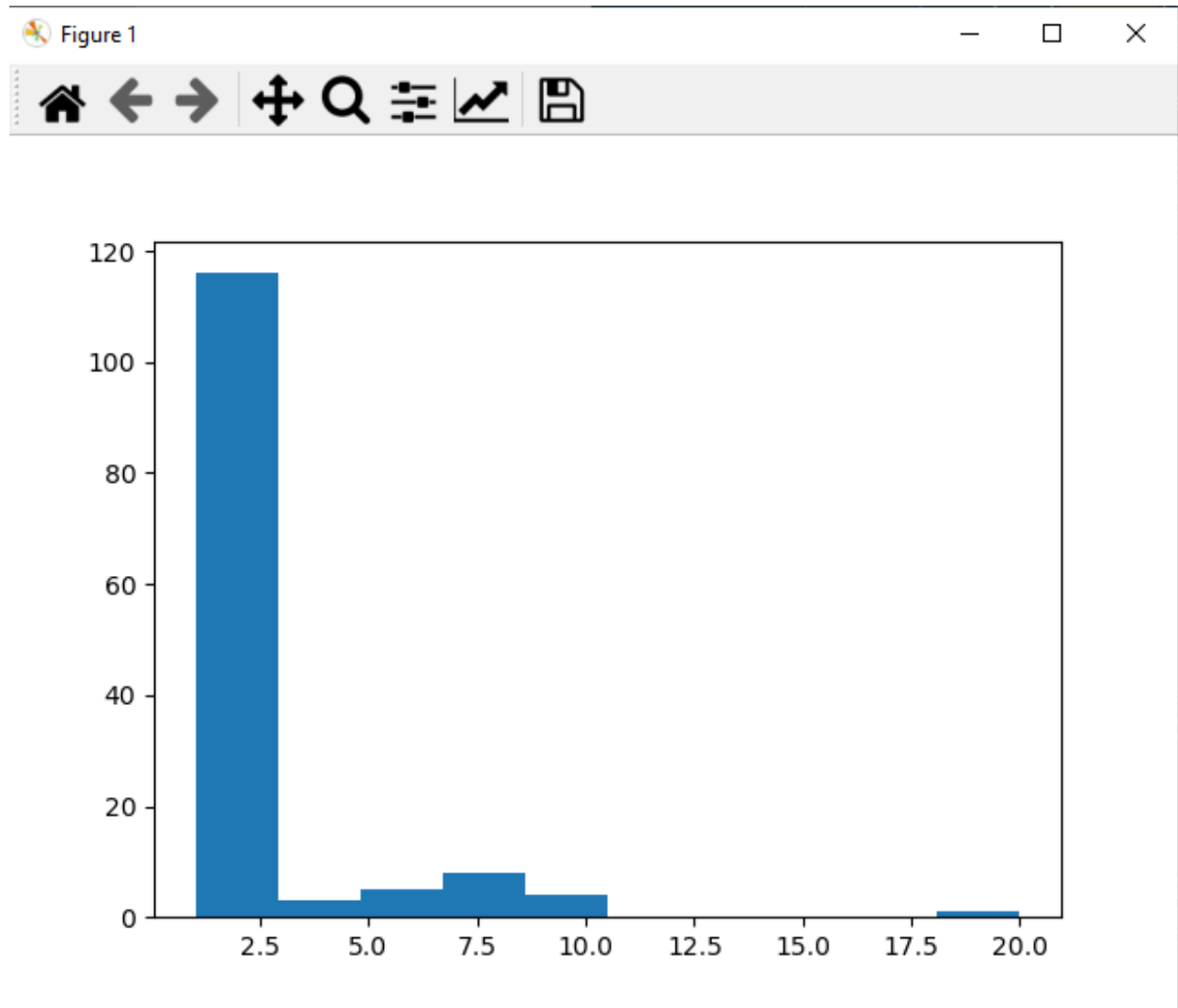---

Step 1: Crawling for JSON Data

Please see crawled data in the file called "dataCrawl.json".

The data crawled contains JSON data related to post data for specific subreddits. I began by making a get request for the most recent posts in the ASU subreddit. I analyzed the data that came from this get request, and found that we can see all sorts of information about the post, upvotes, and author. I then leveraged PRAW (Python Reddit API Wrapper) to get more information about each author and their related comments. For the most recent comments made for each author in the JSON body, I gathered the associated subreddits. Once I laid the groundwork, I crawled for more data about each of the relevant subreddits, and appended them to dataCrawl.json. This way, I have the ASU subreddit posts, as well as common ASU subreddit authors's subreddits of interest.

Step 2: Creating the Network

      The network I created shows the relationships between subreddits. ASU subreddit is the starting node, as we are using this as an analysis of students' subreddit interests. The graph only expands 2 levels, meaning the relevant subreddits to the relevant subreddits of the ASU subreddit. The number of out degrees for each relevant subreddit is randomized, some have as little as 3 nodes and as many as 8 nodes.

      I would classify this network graph as a Subreddit Relationship and Diffusion network, since each topic is connected to other related topics. I use "related" loosely, as it's more about the interests of the ASU subreddit authors, rather than topics related to the topic of "ASU" itself. This network graph also shows the subreddits where ASU subreddits authors are most active. This means rather than ingesting information across these subreddits, these authors contribute to the spread of information and add to information propagation.

Step 3: Network Measures and Analysis

Histogram Analysis of Degree Distribution

This histogram shows the number of connections of each node in the graph. The degree distributions vary based on the out degree. Most of the leaf nodes have a degree distribution of 1, while just the 'ASU' node has a degree distribution of 20, since it has 20 associated nodes.

I also measured Degree Centrality. This type of analysis gives us the importance score related to the number of links that hold each node together. Here are the results for the graph shown above.

Degree Centrality Analysis
* Please be aware some terms are not professionally appropriate.

ASU 0.14705882352941177
nflmemes 0.0661764705882353
TheThiccness 0.0661764705882353
AnalGW 0.0661764705882353
Veterans 0.0661764705882353
StudentLoans 0.058823529411764705
shitposting 0.058823529411764705
BreedingMaterial 0.058823529411764705
BokuNoHeroAcademia 0.058823529411764705
eBaySellerAdvice 0.058823529411764705
SluttyOnlyfans 0.051470588235294115
amex 0.051470588235294115
Cringetopia 0.051470588235294115
AnketTR 0.044117647058823525
phoenix 0.044117647058823525
Frat 0.03676470588235294
GunAccessoriesForSale 0.03676470588235294
dividends 0.03676470588235294
booty_queens 0.029411764705882353
battlestations 0.029411764705882353
AskMenOver30 0.029411764705882353
nfl 0.007352941176470588
formuladank 0.007352941176470588
Patriots 0.007352941176470588
fut 0.007352941176470588
memes 0.007352941176470588
AZCardinals 0.007352941176470588
Libertarian 0.007352941176470588
marvelstudios 0.007352941176470588
funny 0.007352941176470588
CivVI 0.007352941176470588
UWMadison 0.007352941176470588
Drugtests 0.007352941176470588
Debate 0.007352941176470588
Divorce 0.007352941176470588
Eyebleach 0.007352941176470588
AskReddit 0.007352941176470588
prephysicianassistant 0.007352941176470588

curvy 0.007352941176470588
pussy 0.007352941176470588
fuckdoll 0.007352941176470588
tiktokporn 0.007352941176470588
AdultNeeds 0.007352941176470588
pornID 0.007352941176470588
u_renawhite1997 0.007352941176470588
sensualcurves 0.007352941176470588
BestBooties 0.007352941176470588
Milfie 0.007352941176470588
RealGirls 0.007352941176470588
u_LatifahLove__ 0.007352941176470588
ModSupport 0.007352941176470588
Slutsofonlyfans 0.007352941176470588
LabiaGW 0.007352941176470588
NSFW_ANALQUEENS 0.007352941176470588
FaceFuck 0.007352941176470588
anal 0.007352941176470588
u_fluffy_faerie 0.007352941176470588
AskRedditAfterDark 0.007352941176470588
CryptoMoonShots 0.007352941176470588
JuicyOnlyfans 0.007352941176470588
OnlyFansSpecials 0.007352941176470588
AdorableOnlyfans 0.007352941176470588
onlyfansgirls101 0.007352941176470588
ThickThighs 0.007352941176470588
manga 0.007352941176470588
OnePunchMan 0.007352941176470588
movies 0.007352941176470588
translator 0.007352941176470588
poshmark 0.007352941176470588
legostarwars 0.007352941176470588
u_peachnecctar 0.007352941176470588
Onlyfanspromoss 0.007352941176470588
SheFucksHim 0.007352941176470588
macsetups 0.007352941176470588
nursing 0.007352941176470588
Romania 0.007352941176470588

Costco 0.007352941176470588
churning 0.007352941176470588
kurzgesagt 0.007352941176470588
SiyasiPusulaMemeleri 0.007352941176470588
WeebTurks 0.007352941176470588
KGBTR 0.007352941176470588
TurkeyJerky 0.007352941176470588
reloading 0.007352941176470588
Reno 0.007352941176470588
gundeals 0.007352941176470588
Unexpected 0.007352941176470588
Amnesia 0.007352941176470588
dankvideos 0.007352941176470588
PoliticalCompassMemes 0.007352941176470588
NonPoliticalTwitter 0.007352941176470588
ZargoryanGalaksisi 0.007352941176470588
tf2 0.007352941176470588
LeagueOfMemes 0.007352941176470588
redditdev 0.007352941176470588
BustyPetite 0.007352941176470588
milf 0.007352941176470588
FreeCompliments 0.007352941176470588
thepromisedneverland 0.007352941176470588
Komi_san 0.007352941176470588
povertyfinance 0.007352941176470588
RandomActsofCards 0.007352941176470588
NYKnicks 0.007352941176470588
Sonsofanarchy 0.007352941176470588
Ebay 0.007352941176470588

aznidentity 0.007352941176470588
Fibromyalgia 0.007352941176470588
RangersApprentice 0.007352941176470588
PS5 0.007352941176470588
flightsim 0.007352941176470588
VinylWrap 0.007352941176470588
PoliticalDiscussion 0.007352941176470588
stocks 0.007352941176470588
tes3mods 0.007352941176470588
Detailing 0.007352941176470588
berserklejerk 0.007352941176470588
OnePiece 0.007352941176470588
TikTokCringe 0.007352941176470588
teenagers 0.007352941176470588
Tinder 0.007352941176470588
insaneparents 0.007352941176470588
Boxing 0.007352941176470588
AskWomenOver30 0.007352941176470588
Jewish 0.007352941176470588
Volvo 0.007352941176470588
VeteransBenefits 0.007352941176470588
PokemonTCG 0.007352941176470588
AskTrumpSupporters 0.007352941176470588
shorthairedhotties 0.007352941176470588
numerology 0.007352941176470588
loperamide 0.007352941176470588
uscg 0.007352941176470588
ibs 0.007352941176470588
CFB 0.007352941176470588
WTF 0.007352941176470588
news 0.007352941176470588
personalfinance 0.007352941176470588

Finally, I measured betweenness centrality. Betweenness centrality measures if a node lies on a path between other nodes. It indicates its importance.

Betweenness Centrality Analysis
* Please be aware some terms are not professionally appropriate.

ASU 0.9537037037037037	BestBooties 0.0
nflmemes 0.11459694989106753	Milfie 0.0
TheThiccness 0.11459694989106753	RealGirls 0.0
AnalGW 0.11459694989106753	u_LatifahLove__ 0.0
Veterans 0.11459694989106753	ModSupport 0.0
StudentLoans 0.10065359477124183	Slutsofonlyfans 0.0
shitposting 0.10065359477124183	LabiaGW 0.0
BreedingMaterial 0.10065359477124183	NSFW_ANALQUEENS 0.0
BokuNoHeroAcademia 0.10065359477124183	FaceFuck 0.0
eBaySellerAdvice 0.10065359477124183	anal 0.0
SluttyOnlyfans 0.08660130718954248	u_fluffy_faerie 0.0
amex 0.08660130718954248	AskRedditAfterDark 0.0
Cringetopia 0.08660130718954248	CryptoMoonShots 0.0
AnketTR 0.0724400871459695	JuicyOnlyfans 0.0
phoenix 0.0724400871459695	OnlyFansSpecials 0.0
Frat 0.05816993464052288	AdorableOnlyfans 0.0
GunAccessoriesForSale 0.05816993464052288	onlyfansgirls101 0.0
dividends 0.05816993464052288	ThickThighs 0.0
booty_queens 0.04379084967320261	manga 0.0
battlestations 0.04379084967320261	OnePunchMan 0.0
AskMenOver30 0.04379084967320261	movies 0.0
nfl 0.0	translator 0.0
formuladank 0.0	FreeCompliments 0.0
Patriots 0.0	thepromisedneverland 0.0
fut 0.0	Komi_san 0.0
memes 0.0	povertyfinance 0.0
AZCardinals 0.0	RandomActsofCards 0.0
Libertarian 0.0	NYKnicks 0.0
marvelstudios 0.0	Sonsofanarchy 0.0
funny 0.0	Ebay 0.0
CivVI 0.0	poshmark 0.0
UWMadison 0.0	legostarwars 0.0
Drugtests 0.0	u_peachnecctar 0.0
Debate 0.0	Onlyfanspromoss 0.0
Divorce 0.0	SheFucksHim 0.0
Eyebleach 0.0	macsetups 0.0
AskReddit 0.0	nursing 0.0
prephysicianassistant 0.0	Romania 0.0
Costco 0.0	aznidentity 0.0
churning 0.0	Fibromyalgia 0.0
kurzgesagt 0.0	RangersApprentice 0.0
SiyasiPusulaMemeleri 0.0	PS5 0.0
WeebTurks 0.0	flightsim 0.0
KGBTR 0.0	VinylWrap 0.0

TurkeyJerky 0.0
reloading 0.0
Reno 0.0
gundeals 0.0
Unexpected 0.0
Amnesia 0.0
dankvideos 0.0
PoliticalCompassMemes 0.0
NonPoliticalTwitter 0.0
ZargoryanGalaksisi 0.0
tf2 0.0
LeagueOfMemes 0.0
redditdev 0.0
BustyPetite 0.0
milf 0.0
curvy 0.0
pussy 0.0
fuckdoll 0.0
tiktokporn 0.0
AdultNeeds 0.0
pornID 0.0
u_renawhite1997 0.0
sensualcurves 0.0

PoliticalDiscussion 0.0
stocks 0.0
tes3mods 0.0
Detailing 0.0
berserklejerk 0.0
OnePiece 0.0
TikTokCringe 0.0
teenagers 0.0
Tinder 0.0
insaneparents 0.0
Boxing 0.0
AskWomenOver30 0.0
Jewish 0.0
Volvo 0.0
VeteransBenefits 0.0
PokemonTCG 0.0
AskTrumpSupporters 0.0
shorthairedhotties 0.0
numerology 0.0
loperamide 0.0
uscg 0.0
ibs 0.0
CFB 0.0
WTF 0.0
news 0.0
personalfinance 0.0