



# بررسی شبکه یک گروه تاکام

## علم شبکه - پاییز ۱۴۰۲

علیرضا حبیب زاده، یاسمن کاتبی، علی فیاض

# جمع‌آوری داده‌با استفاده از API تلگرام ساخت خزنده (crawler) تلگرام

- گرفتن کد API کلاینت شخصی
- لاین به اکانت تلگرام یکی از اعضای گروه
- ذخیره‌ی لیست اعضای گروه به صورت فایل CSV
- ذخیره‌ی پیام‌های گروه به صورت فایل CSV
- ذخیره‌ی عکس‌های پروفایل افراد در دیتابیس
- ادغام دو گروه قدیمی و جدید (؟؟؟)

```
telegram_crawler.py ×

3   from telethon.errors import SessionPasswordNeededError
4   from telethon.sync import TelegramClient
5   from telethon.tl.functions.messages import GetDialogsRequest, GetHistoryRequest,
6   from telethon.tl.types import InputPeerEmpty, InputPeerChat
7   from getpass import getpass
8   import csv

9
10  from credentials import api_id, api_hash, phone_number
11
12  client = TelegramClient(phone_number, api_id, api_hash)

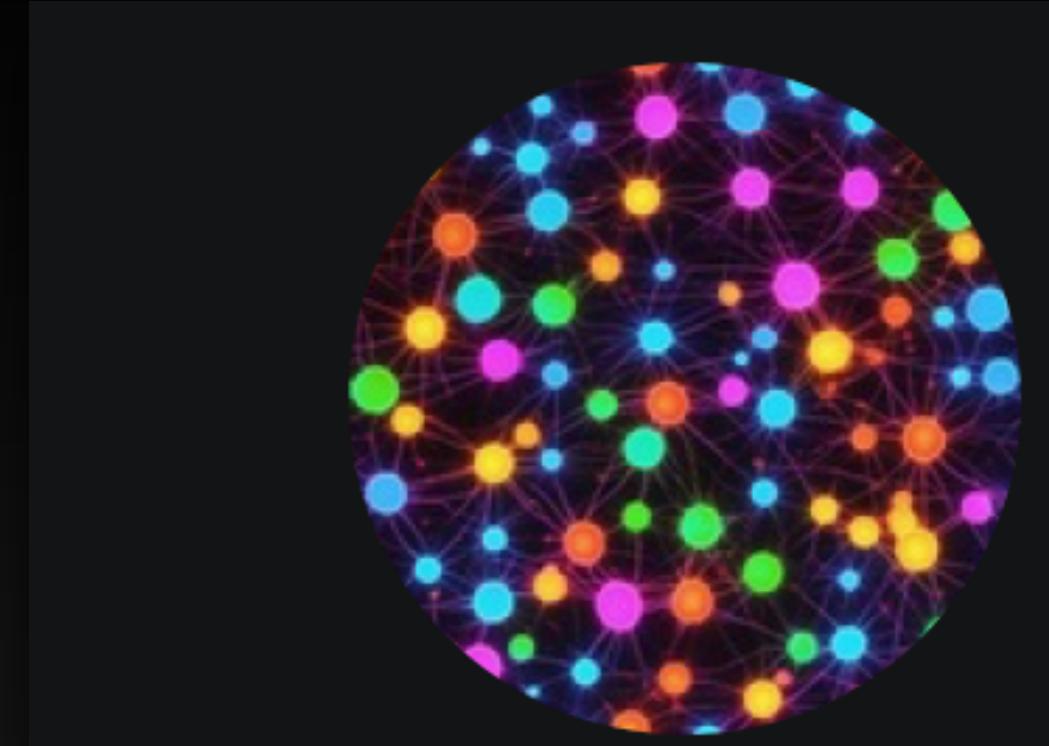
13  client.connect()
14  if not client.is_user_authorized():
15      client.send_code_request(phone_number)
16      try:
17          client.sign_in(phone_number, input('Enter the code: '))
18      except SessionPasswordNeededError:
19          client.sign_in(password=getpass('Enter 2-step verification password: '))

20
21
22  1 usage new *
23  def list_users_in_group(group_id=None, update_photos=False):
24      if group_id:
25          target_group = client.get_entity(group_id)
26      else:
27          chats = []
28          last_date = None
29          chunk_size = 200
30          groups = []

31
32          result = client(GetDialogsRequest(
33              offset_date=last_date,
34              offset_id=0,
```

# گروه تلگرام درس

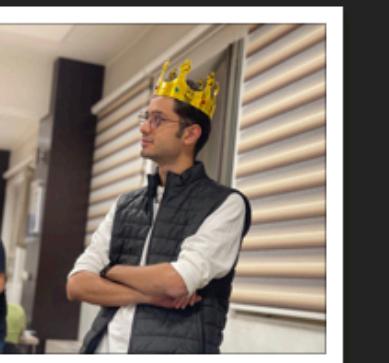
## (نتیجه‌ی خزندۀ برای عکس اعضای گروه)



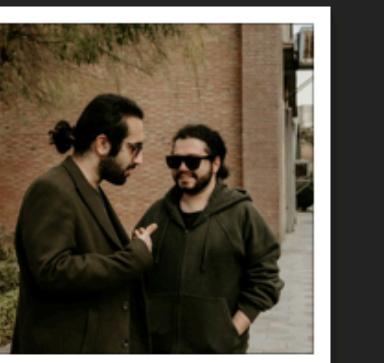
درس علم شبکه  
62 members

< > profile\_pics

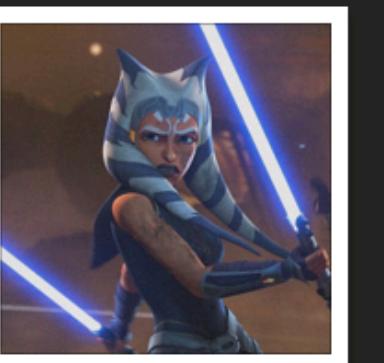
grid list thumbnail details search filter refresh



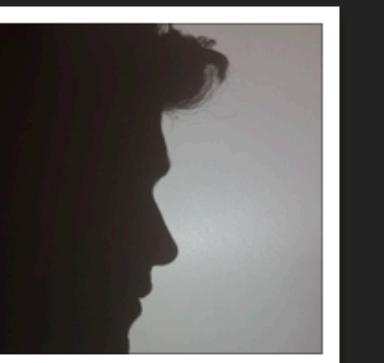
51379205.jpg



72054777.jpg



74259726.jpg



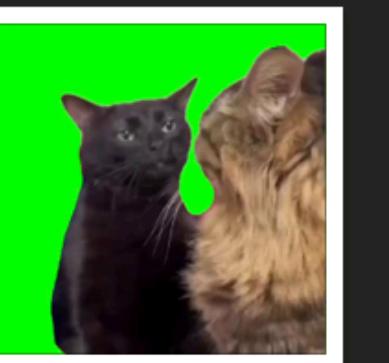
90300310.jpg



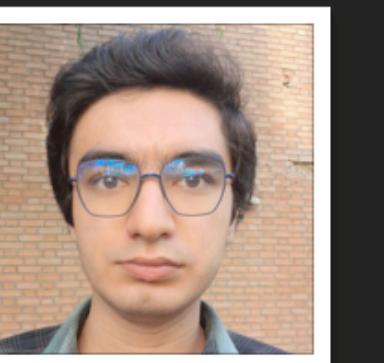
90438363.jpg



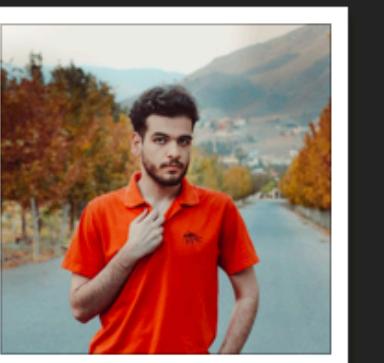
94302343.jpg



103283273.jpg



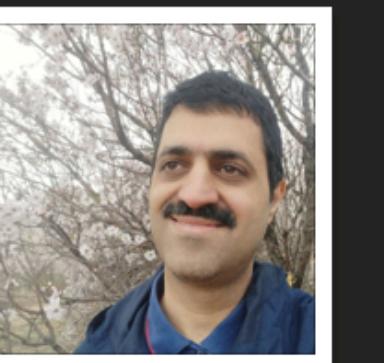
110020679.jpg



110075745.jpg



113130045.jpg



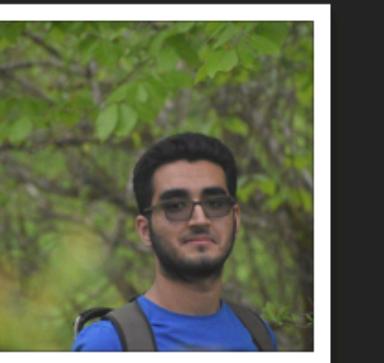
113596538.jpg



114019526.jpg



151119336.jpg



169078873.jpg



182804079.jpg



225569842.jpg



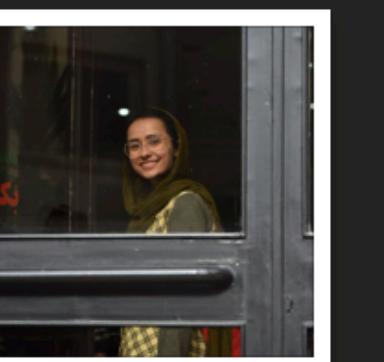
250701125.jpg



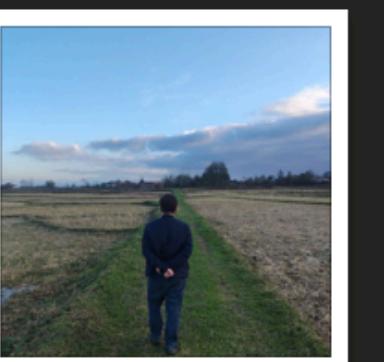
254957240.jpg



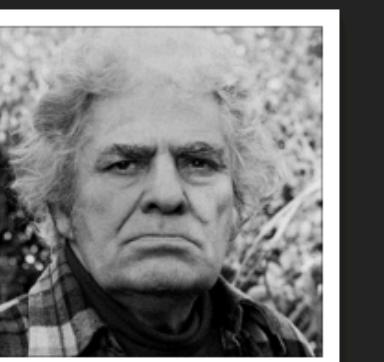
334531588.jpg



345082731.jpg



375997083.jpg



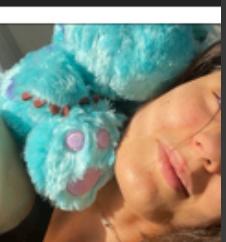
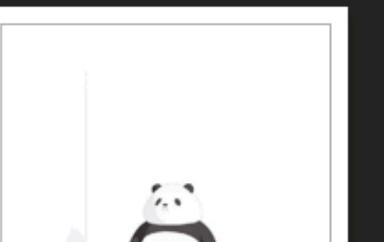
389058712.jpg



417731735.jpg



502771921.jpg



# لیست اعضاي گروه

فیلد ها

- user\_id
- username
- access\_hash
- first\_name
- last\_name
- group\_id
- group

members.csv

Open with Numbers

user_id	username	access_hash	first_name	last_name	group_id	group
113596538	Samanimi	694293864374282792	Saman	Moghimi	1952093821	درس علم شبکه
676320722	YShahmari	6876743246295805006	Yekub		1952093821	درس علم شبکه
254957240	patak_19	709468690113790865	Parsa	Zarezadeh	1952093821	درس علم شبکه
502771921	Armhmdzd	-1882420116839448254	Amirreza	Mahmoudzadeh	1952093821	درس علم شبکه
375997083	Moein_Derakhshan	3306814744382472893	Moein		1952093821	درس علم شبکه
103283273	sepehr_s_y	-1494875107415876362	SEPEHR.S.Y		1952093821	درس علم شبکه
151346069		9173263705290824918	مژگان		1952093821	درس علم شبکه
116908618	slhshamloo	8077894234555670809	Saleh	Shamloo	1952093821	درس علم شبکه
417731735	CetOlivier	7156631571320723439	Sara		1952093821	درس علم شبکه
686550448	alnesari	155581392785975692	Ali	Ans	1952093821	درس علم شبکه
175840468	marziekamkar	-5613171634187709112	Marzie	Kamkar	1952093821	درس علم شبکه
51379205	habibz	3136676058508021306	Alireza	Habibzadeh	1952093821	درس علم شبکه
110075745	mahdi_ghoraishy	224289987459224339	mahdi	Ghoraishy	1952093821	درس علم شبکه
869332700	misslibrainpractice	-2997495962747930194	Fatemeh		1952093821	درس علم شبکه
94302343	Sina_Moammar	-2995410570962395065	Sina	Moammar	1952093821	درس علم شبکه
95281876	M_Shaker_A	-1180662390739901298	Morteza	Shaker Arani	1952093821	درس علم شبکه
247070571	AstonishingAstane	-1835730195362837091	Alireza	Astane	1952093821	درس علم شبکه
1383984423	yacmnkatebi	8047108114705716681	Yaci	Katebi	1952093821	درس علم شبکه
90300310	shojaie	5946570647299950738	MA	Shojaie	1952093821	درس علم شبکه
669168008	EZaapp	8619995061780956951	Zapiv		1952093821	درس علم شبکه
862205162		5060801375685692694	Saman		1952093821	درس علم شبکه
1329060682	Haniehtm	3586662505231065662	Hanie		1952093821	درس علم شبکه
93626156	siiiaaa33	6367184400380700033	Saii		1952093821	درس علم شبکه
586966688	k_b_s_b	-877804500627885371	N		1952093821	درس علم شبکه
99490341	Narges8k	-1847283371849998184	Narges		1952093821	درس علم شبکه
1489741954	fateme_ramezanade81	-6957288686418614026	Fateme	Rmz	1952093821	درس علم شبکه
272594823	Morphyoos	7661645363238503873	Yoosef		1952093821	درس علم شبکه
250701125	AF6ECHO	-435451186291175638	Ali	Fayyaz	1952093821	درس علم شبکه
345082731	yasihakim	-6721437585941062197	Yasamin	Hakiminejad	1952093821	درس علم شبکه
327881189	Parsa_Ghafaripour	347346184661019559	Parsa	Ghafaripour	1952093821	درس علم شبکه
94918358	Mahdi_N_13	-1733650185903589522	Mahdi	Naghiloo	1952093821	درس علم شبکه
182804079	kamyaritto	-1663540206869695727	Kamyar		1952093821	درس علم شبکه

# پیامهای گروه

## فیلد ها

- message\_id
- from\_user\_id
- reply\_to
- pinned
- message
- reactions

message_id	from_user_id	reply_to	pinned	message
471	113596538		False	
470	113596538		False	لطفاً زودتر برام بفرستید.
469	113596538		False	یه جا گذاشتام که ارایه هاتون رو بفرستید CW تو
468	113596538		False	من امروز ف ۳ (کلاس خودمون) رو برای صبح و بعد از ظهر رزرو می کنم.
467	113596538		False	هر ارایه دوازده دقیقه فرصت داره. صبح ساعت نه شروع می کنیم
466	113596538		False	ارایه های فردا با این زمان بندی انجام می شه.
464	676320722	462	False	ولی مرسى
463	676320722	461	False	من گفته باشم من چیزی ننوشتم 😂👉
462	113596538		False	برنامه رو می گذاریم دوشنبه نه تا یازده و یک تا دو و نیم
461	113596538		False	تو خصوصی چیز هایی برام نوشته شده، اگه کس دیگه ای هم چیزی داره بهم بگه
460	676320722	458	False	عیت دانشجو های شدیداً بی کیفیت خیلی بالاست. از اقلیت های جنسیتی و نژادی و هر کی عقش پاره سنگ برداشته ولی با باش پولدار بوده پیدا می کنید
459	676320722		False	البته سر تصمیم گیری معمولاً دانشجو خیلی نظرش تو نتیجه نهایی موثره اگه خودت بگی فلان موقع بعد ددلاين رو رد بکنی آره از چشم استاد می فتی
458	676320722	454	False	بعد کرونا انعطاف پذیری به شدت رشد داشته تو سیستم آکادمیک غرب. کلا فشار و تا حدودی کیفیت و بازدهی کار ها هم تو دانشگاه ها کاهش داشته
457	254957240	454	False	پس همون برنامه قبلی نهایی است درسته؟ صرفاً پرسیدم که اماده بشیم برآش
456	502771921		False	ت. صرفاً شاید جالب باشه که یه علوم انسانی خوندهای می گفت مشکل ما اینه که دموکراسی مون از بوروکراسی پیشی گرفته! حالا بحث مفصل بود)
455	502771921	454	False	انگار دموکراسی، در قالب تفاهم رو ما بیشتر داریم
454	113596538		False	بین داره. می دونم که شرایط اینجا هم سختی های خودش رو داره و من می فهمم تا حدی شرایط رو. ولی به نظرم رسید که خوبه اینها رو شنیده باشید
453	113596538		False	فوقش می تونی بگی زودتر اعلام کردی. تازه اگر زودتر و دیرتر بودن اهمیتی داشته باشه
452	676320722	451	False	من زودتر خواستم 😊
451	113596538	449	False	همه همین رو می خوان!
450	676320722	419	False	.
449	676320722	426	False	دکتر مارو بذارید دوشنبه شنبه من نیستم
448	375997083	447	False	چه بهتر 😂
447	113596538		False	مخالف هم باشه جرات نمی کنه بگه 😊😂

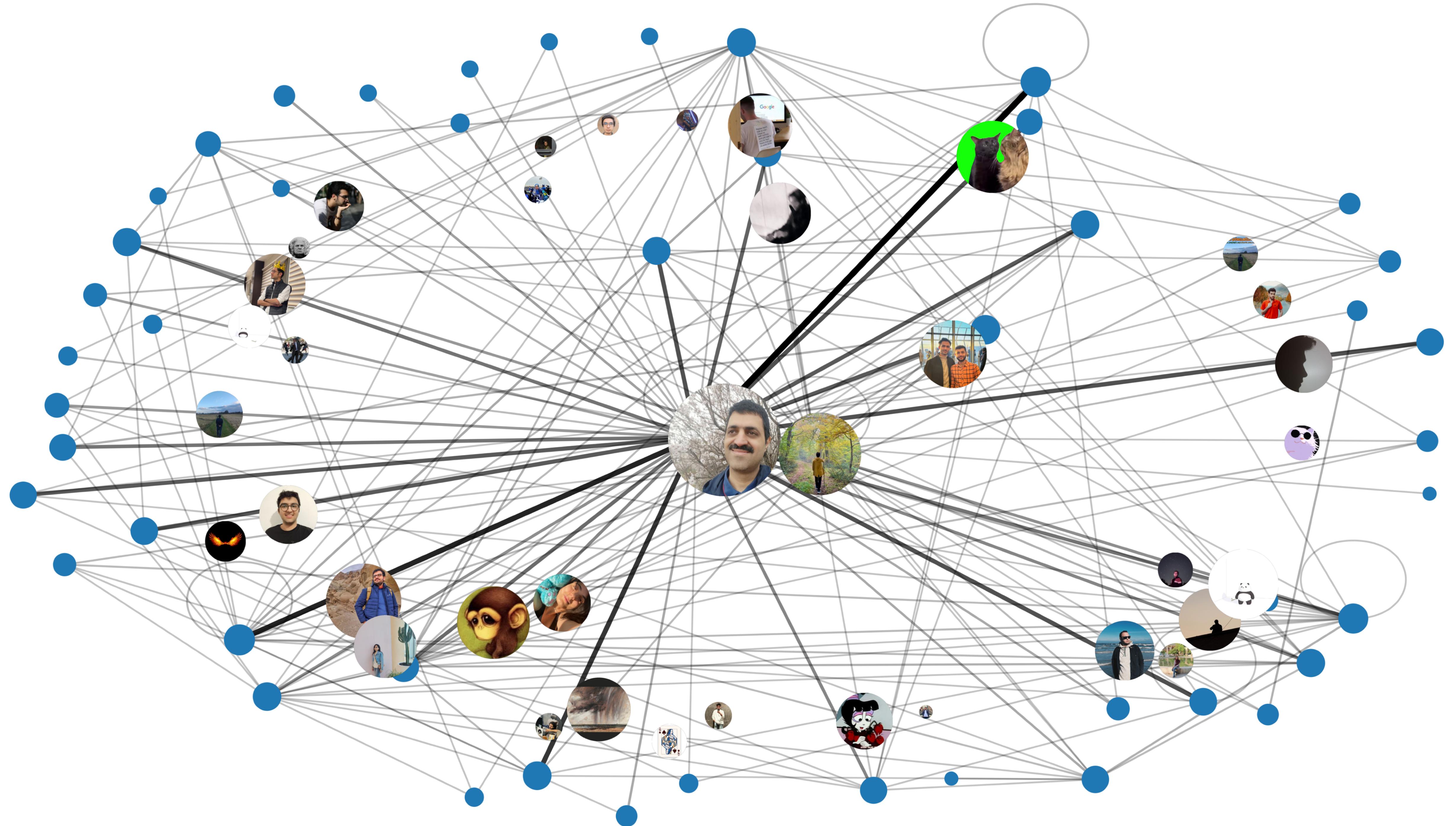
# تشکیل شبکه گروه

ارتباط بین افراد چگونه است؟

- هر پیام می‌تواند ریپلی به پیام دیگری باشد (یال از فرستنده پیام به فرستنده پیام اول)
- هر ریکشن یک یال از فرستنده ریکشن به فرستنده پیام (سخت‌ترین قسم پیاده‌سازی در API !)
- پیام‌های پین شده یال از فرستنده به همه‌ی اعضا است. (نتیجه خوب نشد، فقط استار می‌تواند پین کند و گراف شلوغ می‌شود)

```
telegram_crawler.py group_network2.py x

1 usage
2
3 class User:
4     def __init__(self, user_id: int, username, access_hash, first_name, last_name,
5                  self.user_id = user_id
6                  self.username = username
7                  self.access_hash = access_hash,
8                  self.first_name = first_name
9                  self.last_name = last_name
10                 self.group_id = group_id
11                 self.group = group
12                 self.profile_pic = None # Set if a profile picture exists
13
14
15
16
17
18
19
20
21
22 1 usage
23 class Message:
24     def __init__(self, message_id: int, from_user_id: int, reply_to: int, pinned:
25                  self.message_id = message_id
26                  self.from_user_id = from_user_id
27                  self.reply_to = reply_to
28                  self.pinned = pinned
29                  self.message = message
30                  self.reactions = [int(x) for x in str(reactions).split(',') if pd.notnull(
31
32 # Load data from CSV files
33 1 usage
34 def load_data(group_id):
35     members_path = f"database/group-{group_id}/members.csv"
36     messages_path = f"database/group-{group_id}/messages.csv"
37
38     members_df = pd.read_csv(members_path)
39     messages_df = pd.read_csv(messages_path)
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
59 fit_exponential()
```



# برخی نکات

## در کشیدن گراف

- افراد می‌توانند به پیام خود ریپلی بزنند.
- به جای اضافه کردن چند یال به یال‌ها وزن داده شده و هم پرنگ و کمرنگ شدند و هم ضخامت‌شان عوض شده.
- نتیجه برای گراف جهت‌دار خیلی خوب نبود! وزن دو جهت را در هم ادغام کردیم.
- طبعاً دکتر مقیمی شاهراس (یا یکی از شاهراس‌های؟) گراف است. (+درخواست نمره)



# بررسی توزیع راس ها

آیا ساختار توانی و بی مقياس داریم؟

پاسخ:  
بله!

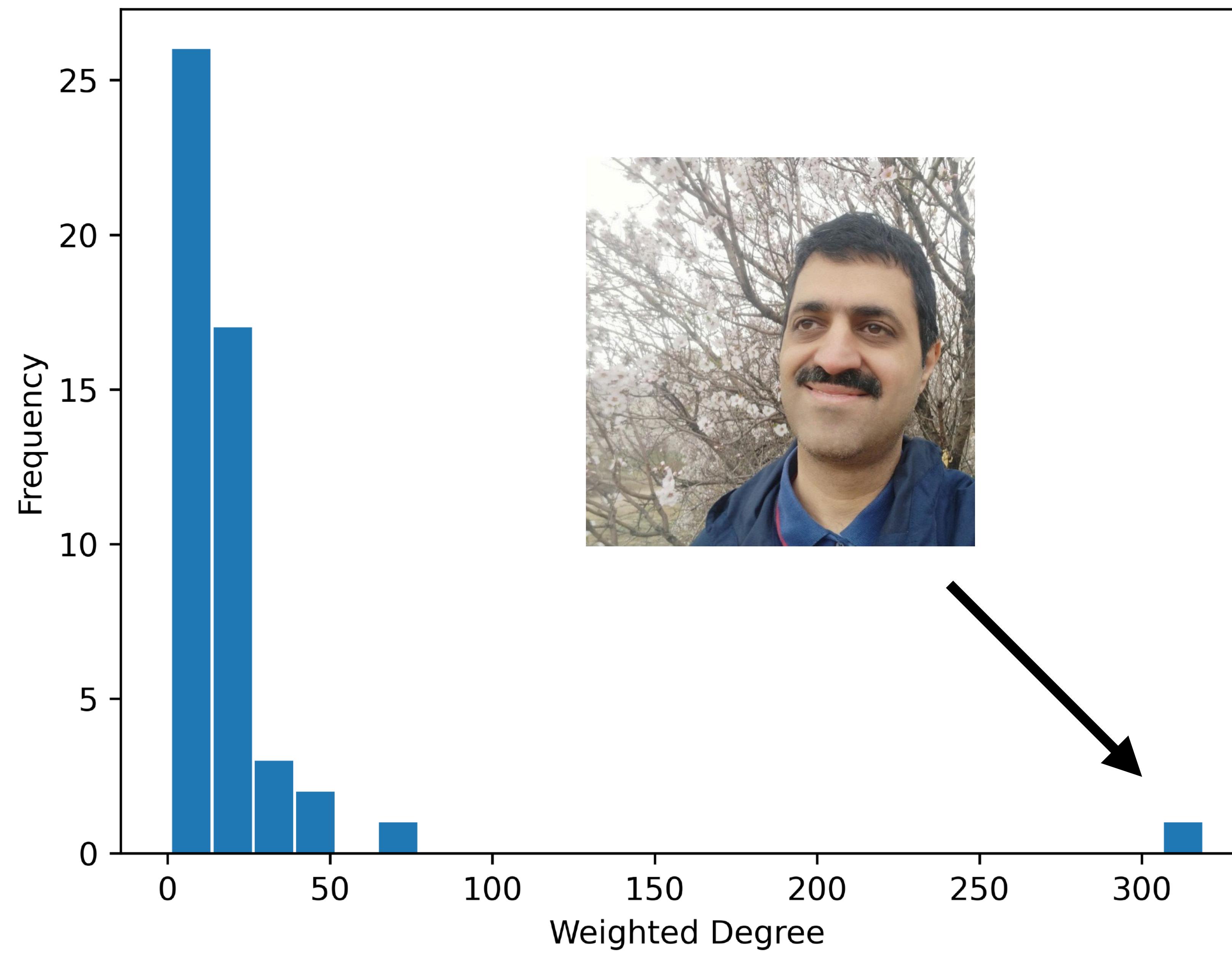
```
telegram_crawler.py group_network2.py x

143
144
145
146 def plot_hist(G):
147     G.remove_nodes_from(list(nx.isolates(G)))
148     degrees = G.degree()
149     degrees = [degree for _, degree in degrees]
150     plt.hist(degrees, bins=len(degrees) // 2, rwidth=0.93)
151     plt.title('Degree Distribution (isolated nodes removed)')
152     plt.xlabel('Weighted Degree')
153     plt.ylabel('Frequency')
154     plt.xscale('log')
155     plt.yscale('log')
156     plt.savefig(*args: './hist.png', dpi=600)
157     plt.show()

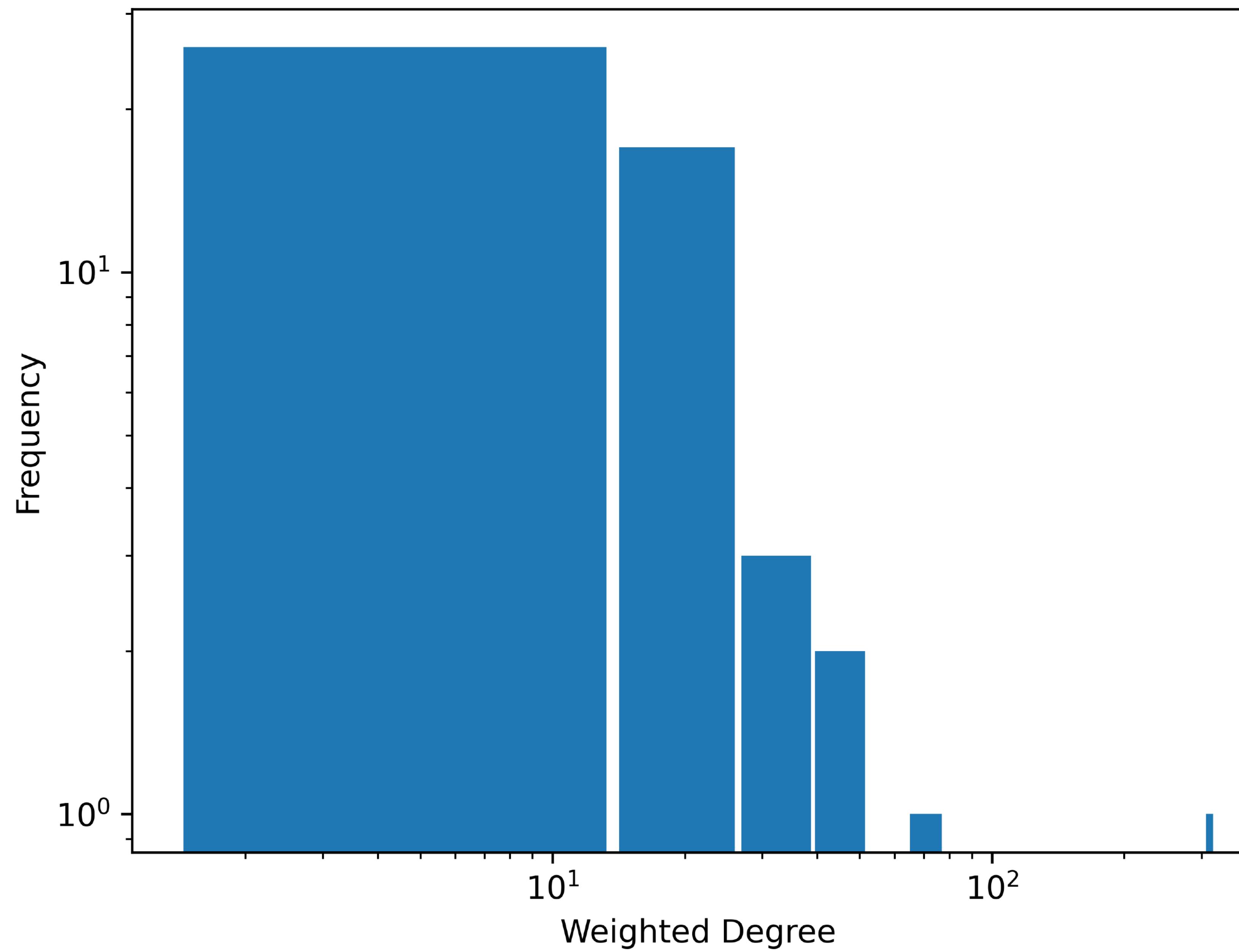
158
159
160 import powerlaw
161
162 1 usage
163 def fit_exponential(G):
164     avalanches = [G.degree(node) for node in G.nodes()]
165     results = powerlaw.Fit(avalanches, discrete=True, xmin=5)

166     plt.figure()
167     results.plot_pdf(marker='o', linestyle='None')
168     results.power_law.plot_pdf(linestyle='--', ax=plt.gca(), label=f'p ∝ s^{-b}, b = {results
169
170     plt.title('Degree Fit')
171     plt.xlabel(f'Degree, xmin={results.power_law.xmin:.0f}')
172     plt.ylabel('Frequency')
173     plt.xscale('linear')
174     plt.yscale('linear')
175     plt.legend()
176     plt.savefig(*args: f'./fit.png', dpi=600)
177
178 fit_exponential()
```

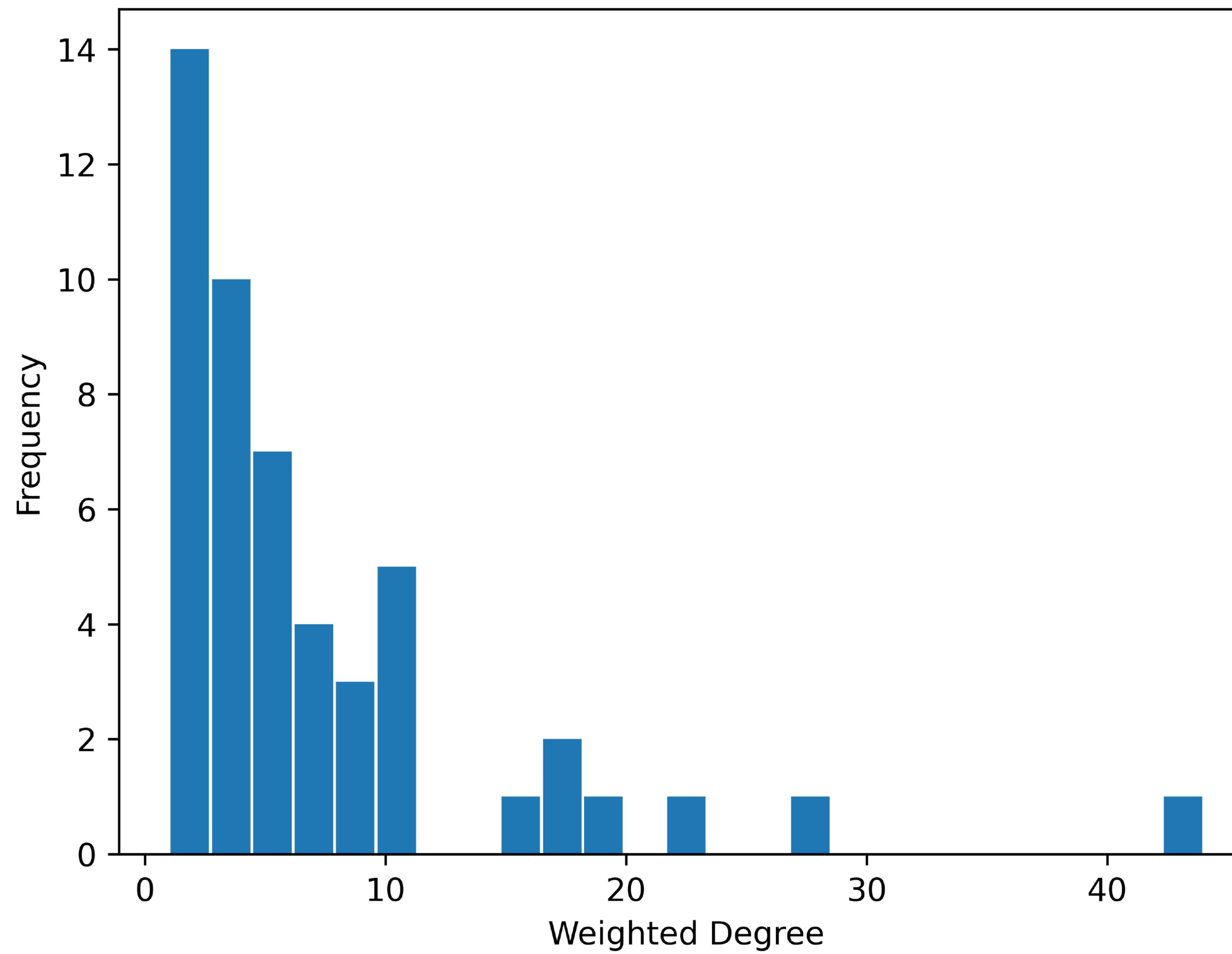
## Weighted Degree Distribution (isolated nodes removed)



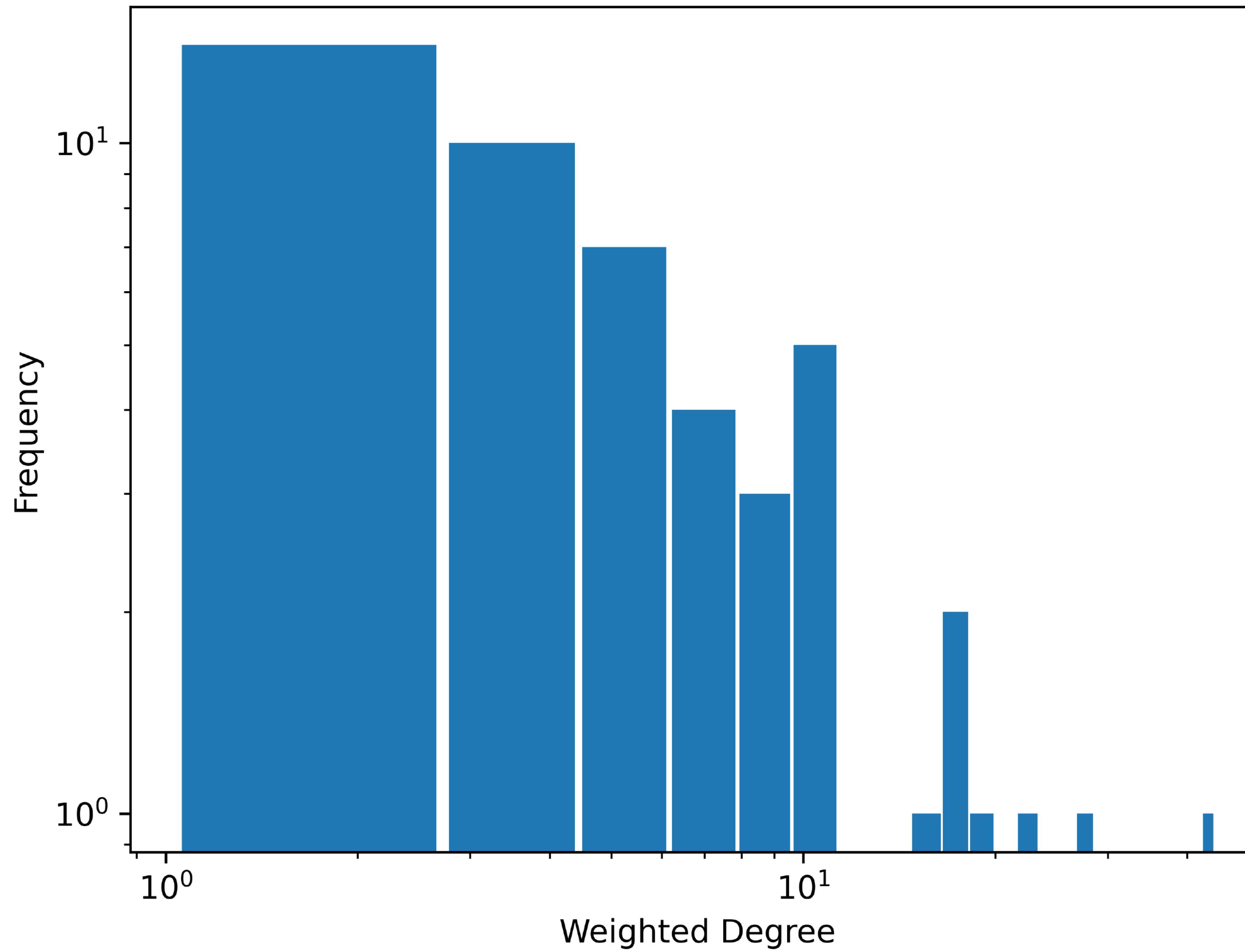
Weighted Degree Distribution (isolated nodes removed)



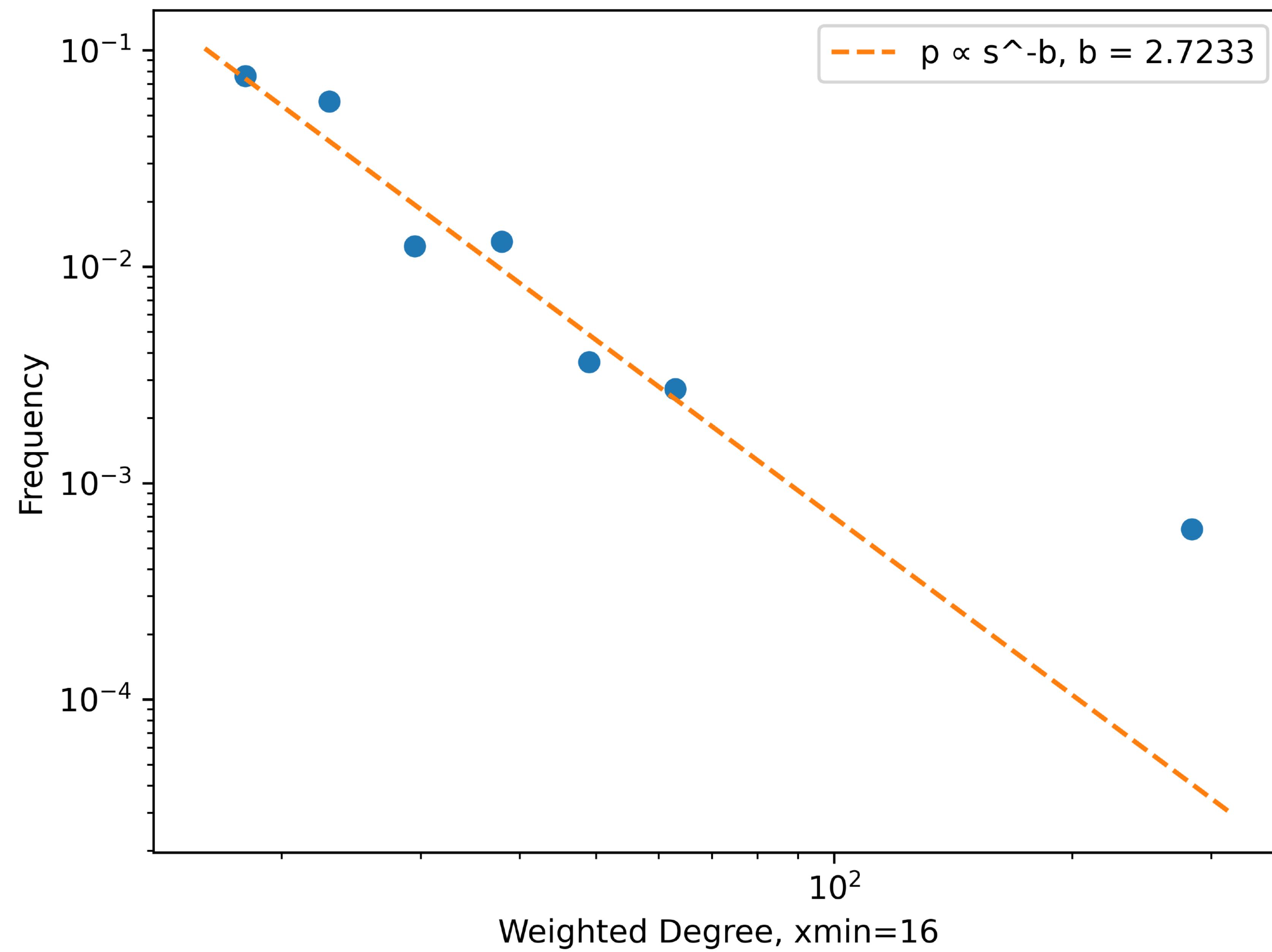
Degree Distribution (isolated nodes removed)



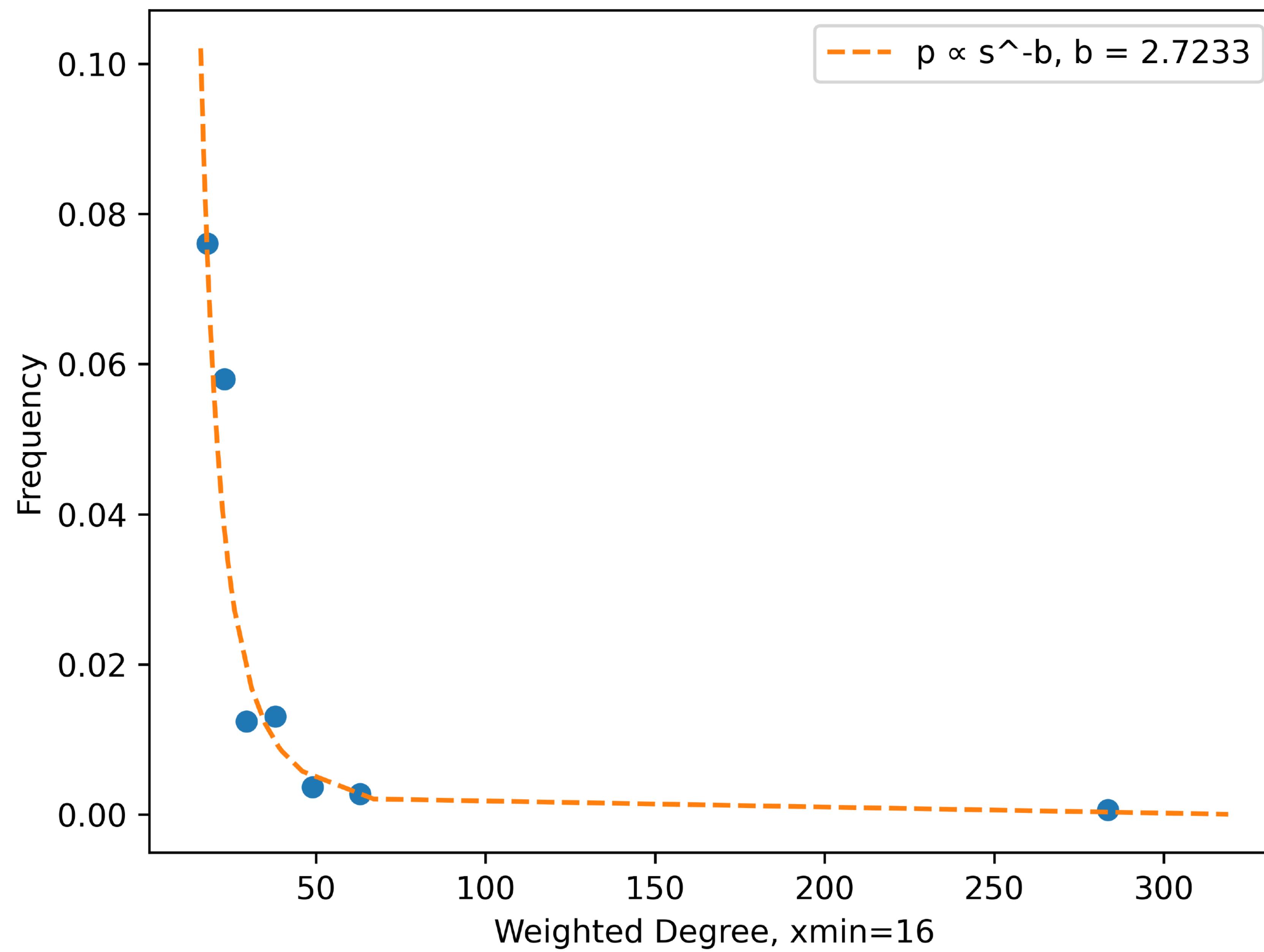
Degree Distribution (isolated nodes removed)



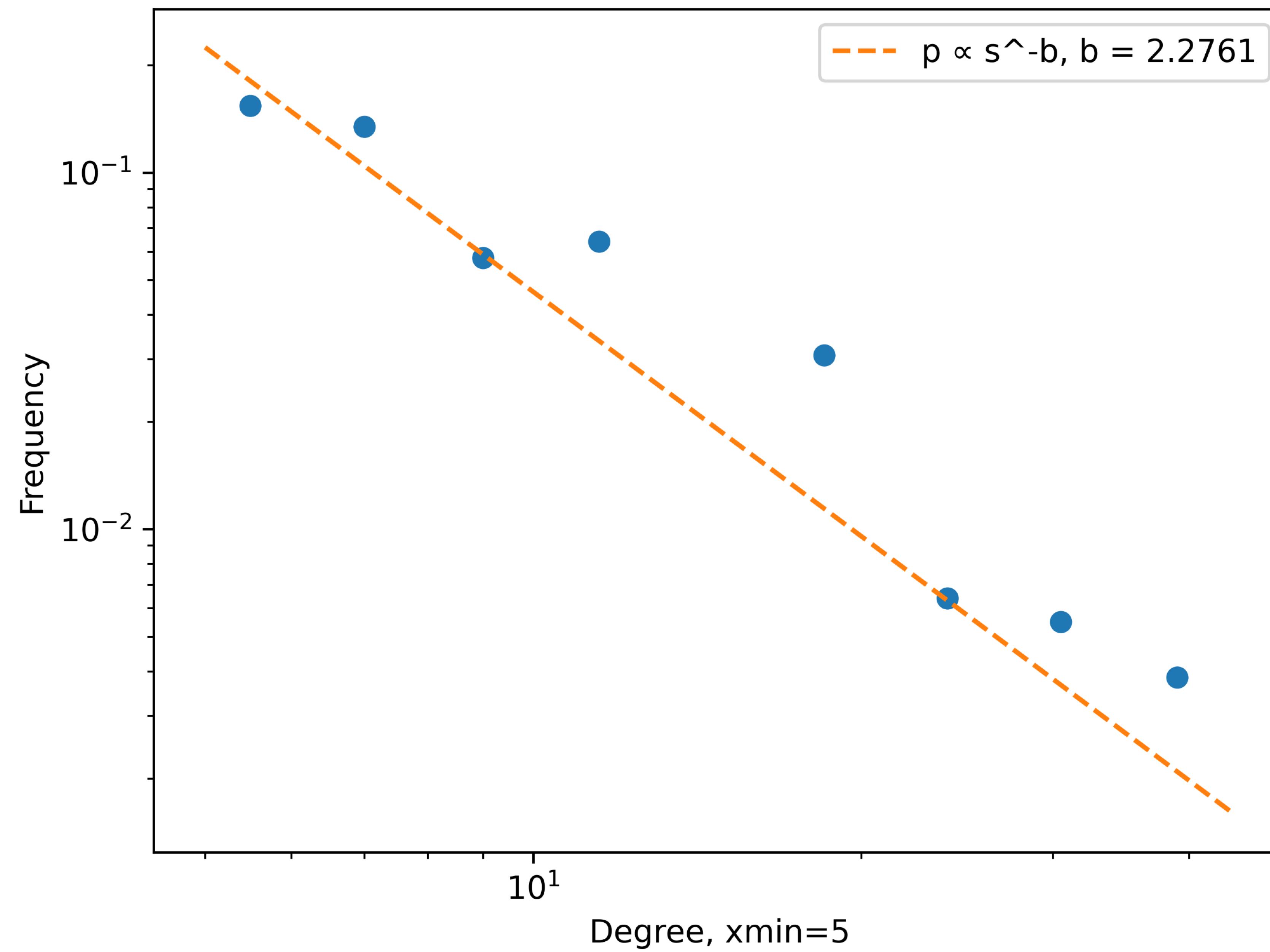
## Weighted Degree Fit



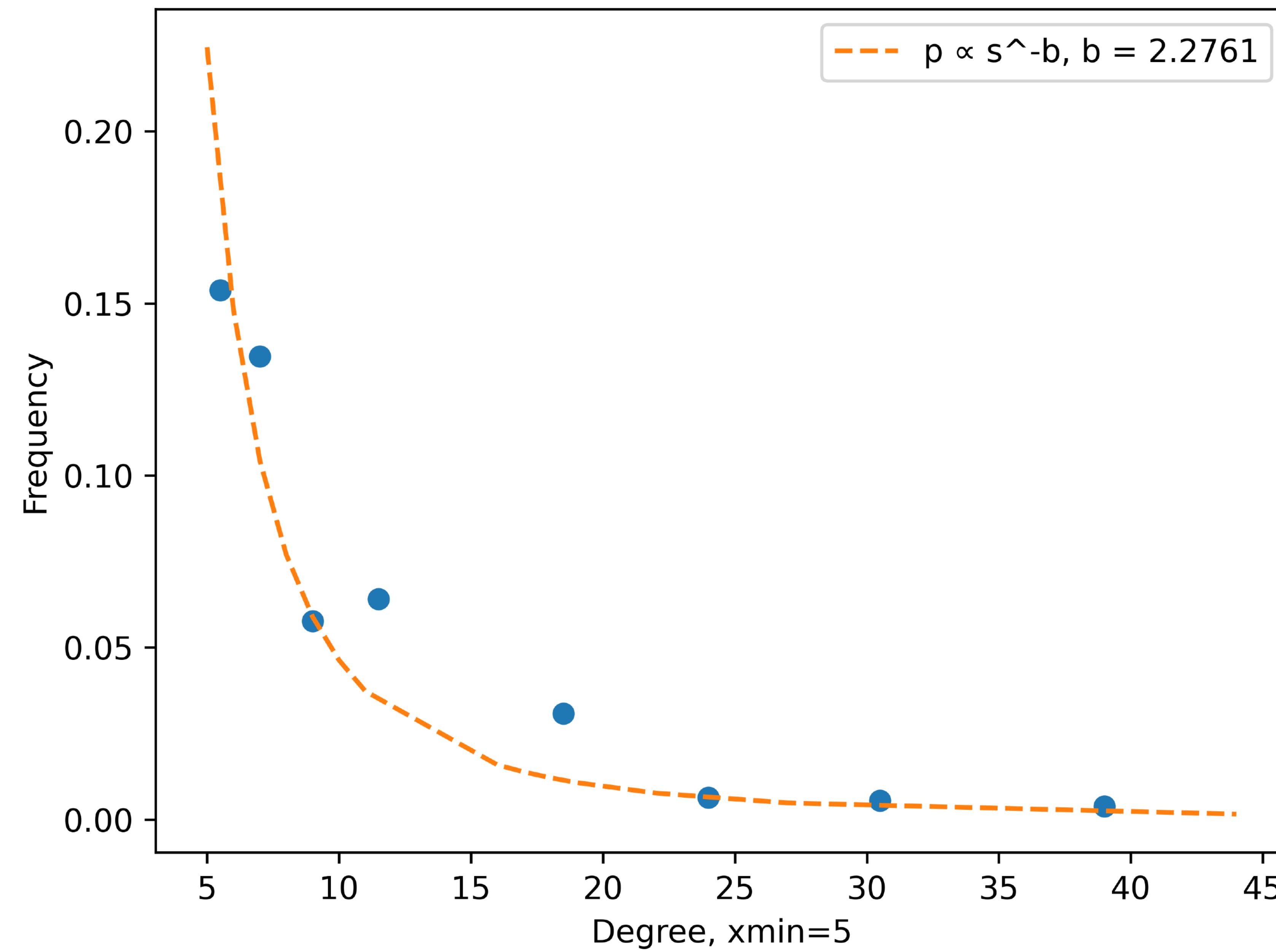
## Weighted Degree Fit



## Degree Fit



## Degree Fit



# سپاس از شما

برخی کارهای ممکن برای آینده:

- بررسی گروههای بزرگتر
- بررسی منشن‌ها (@)
- بهتر کشیدن شبکه
- بررسی و ادغام چندین گروه
- ...