



Analysis of Top Reddit Posts

Era Kalaja

The Problem:

Use the PRAW library in Python to pull meaningful data about the most popular Reddit posts in the r/datascience subreddit.

In the bottom left corner, there are decorative geometric shapes: a blue semi-circle and a pink square. A thin white circle is partially visible on the far left edge of the slide.

How I solved it...

Pulling Posts

Step 1

Pulling data from Reddit using PRAW Package



reddit

- Reddit account
- Application in account
 - Client ID and Secret
- PRAW documentation
- Imported PRAW library
 - Connected to creds

Meaningful Data

Step 2

Filtering data to only show relevant posts to our goals

- 100 posts
 - Scores over 50
 - From 2023
 - Post titles > 5 words
 - Top 3 comments



u/gpbuilder • 9 mo. ago

Pretty Accurate Chart to Clear Up Job Title Ambiguities

Career



Date	Post Score	Post Title
2023-04-26 19:26:45	1987	Pretty Accurate Chart to Clear Up Job Title Ambiguities

Exporting



Step 3

Export findings to excel

■ Columns:

B	C	D	E	F	G
Date	Post Score	Post Title	Comment1	Comment2	Comment3



Post Analysis...

Top Posts?

Career based

- Job opportunities
- Industry Questions

	Date	Post Score	Post Title
0	2023-04-26 19:26:45	1987	Pretty Accurate Chart to Clear Up Job Title Ambiguities
1	2023-05-15 21:45:36	1168	I investigated the Underground Economy of Glassdoor Reviews
110	2023-04-12 10:19:05	292	Manager with no data science background wants to bring in people from his team and is asking us to train them end to end to replace existing people

Best Dates/Times?

Months

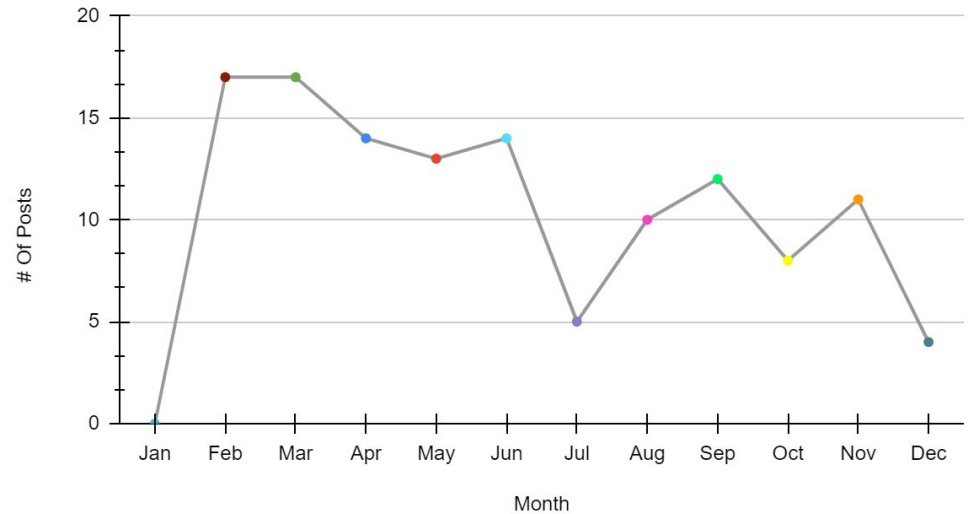
- February
- March

Time

- After 6pm

Why?

Number of Posts Per Month



Top Comments?

Satire

- Insults to Redditor
- Meme Photos

Career Advice

\nChart : Ambiguous\nJob Title Ambiguities :
even more ambiguous

> _Yeah, well, that's just, like, your opinion, man._

Your communications
skills will take you much
farther in your DS career
than your technical skills



Thank You