**1.1 Understanding APIs**

***Note, Part 1 questions can be answered by carefully reading the***[***documentation of the PRAW library carefully***](https://praw.readthedocs.io/en/v3.6.2/pages/getting_started.html)***and/or the***[***reddit API documentation***](https://github.com/reddit-archive/reddit/wiki/API#rules)***.***

* **1.1.1** How many API calls were required to collect the submissions?
* **Ans: 30. Since there is a limit of 100 for a single API call, the 1000 limit results in 10 calls. Since we made 3 praw calls for the 3 subreddits, the total calls made is 30.**
* **1.1.2** Why did we set the submission limit at 1000?
* **Ans: The calls will take longer as the limit is increased. Between calls, the praw wrapper puts in a 2-second delay to confirm with API rules (Ref:** [**https://praw.readthedocs.io/en/v3.6.2/pages/getting\_started.html**](https://praw.readthedocs.io/en/v3.6.2/pages/getting_started.html) **). With the limit at 1000, there are 10 API calls with a 2-second delay which would take at-least 19 seconds.**
* **1.1.3** How long, in minutes, would it take you to collect 1000 posts from 25 different subreddits? What about from 500 different subreddits? *Hint: You'll have to consider how many API requests you are allowed to make in a given time period.*
* **Ans: With 3 subreddits, it takes at-least 57 seconds. So, for 25 subreddits, it would take at-least ~8 minutes (57 \* 25 / (3 \* 60)). For ~500 subreddits, it would take at-least ~160 minutes (57 \* 500 / (3 \* 60)).**

**1.2 Thinking about your sample**

You collected (approximately) the top 1000 submissions from 3 different subreddits.

* **1.2.1** Do you think these posts are representative of **all** the posts on that subreddit? (Yes or no, only)
* **Ans: No.**
* **1.2.2** Why or why not? That is, if you think so, why do you think there's not much sampling bias here? If not, what do you think might be different about these top posts than other posts?
* **Ans: A representative sample of all the posts has to adequately replicate the subreddit’s posters and selecting 1000 top posts from a subreddit does not do this because of the following factors: a) The ‘non-top’ posters are not represented in the sample, and b) The sample size would be too small compared to the total posts in most subreddits.**