

Exercise 01

Mongo data

I'd selected "*Helpful Life Tips from Reddit Dataset (13K Tips)*", a CC0 collection of scraped Reddit *r/lifeProTips* and *r/YouShouldKnow* text posts, as it tickled my fancy for something both relatively lighthearted and with a lot of accompanying metadata.

LPT Post Titles and Karma Scores:

This query extracts the titles of all posts from the "LifeProTips" subreddit, sorts them by their karma score in ascending order, and returns only the score and post title. It uses the find method with a sort and project in the cursor to achieve this.

Multiple Posters with 1000+ Karma:

This query uses the aggregation framework to find users who have posted more than once with a karma score of 1000 or more. It groups posts by the author, filters for authors who are not null and have posted more than once, sorts them by post count, and projects only the author. The result shows users with multiple high-karma posts, with an example of the top author being "[deleted]" and "u/zazzlekdazzle" with 25 posts.

Search for Posts Containing Specific Text:

This query creates a text index on the "postBody" field and searches for posts containing the text "thanks for the gold." It projects only the "postBody" field and returns the results. The number of matching results is logged. An additional query searching for the exact phrase ""thanks for the gold"" is mentioned, which returns a different count of results.

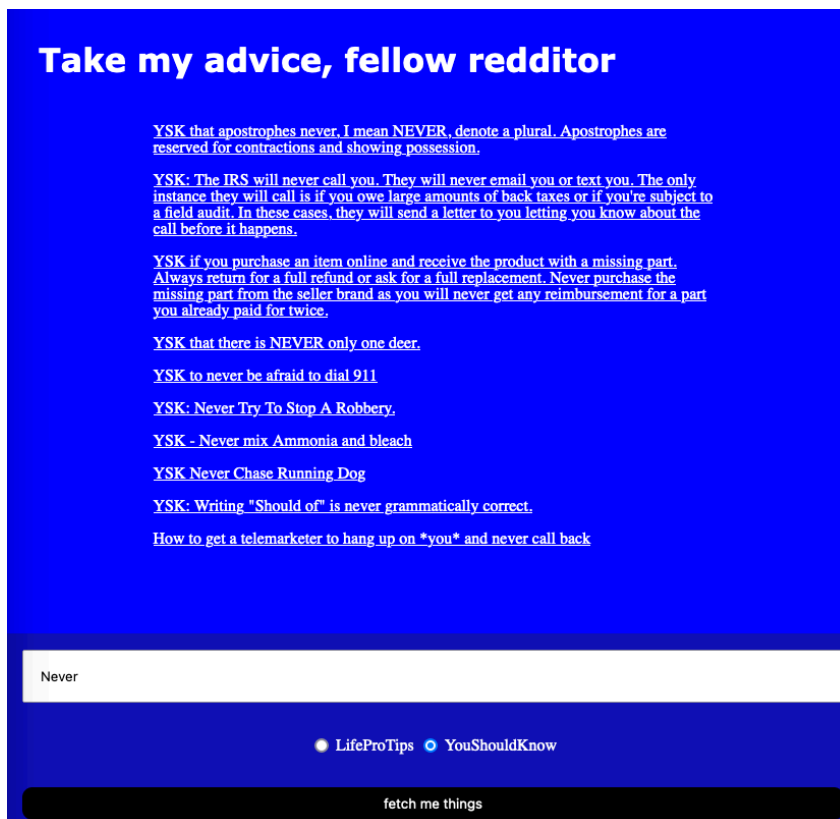
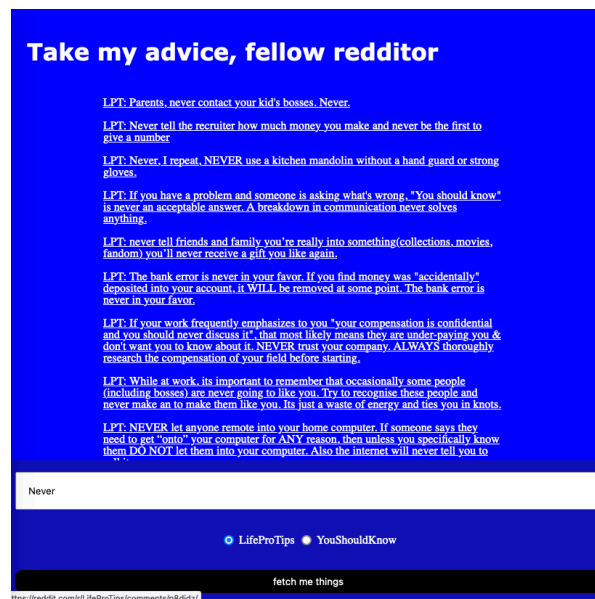
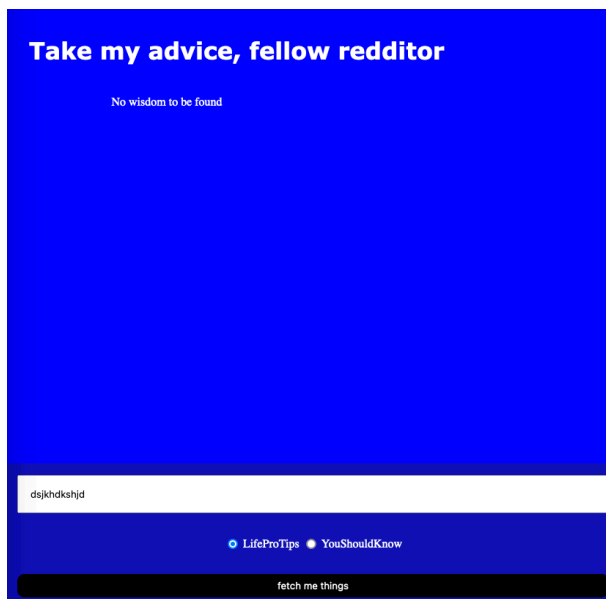
Finding Short LPTs (Less than 40 Characters):

This query finds LPTs (LifeProTips) with short titles (less than 40 characters). It uses \$expr to compare the length of the "postTitle" with 40 characters and checks if "postTitle" exists. The results include LPTs with short titles.

Creating a Composite Index of Popularity:

This query creates a composite index that combines the sum of the "numComments" and "score" fields for each post, creating a "totalAmount" field. It then sorts the results by the "totalAmount" field in ascending order and returns the post titles and their total popularity.

I wanted to do 3d animation, but the design and setup took too long, instead I build an oracle, given that all the data is essentially small edicts to follow.



<https://github.com/stockhuman/c451>