

Lydia Graveline  
CART451  
Exercise 1  
October 2023

With an interest in crowdsourcing, I decided to work with a dataset sourced from Reddit: [Helpful Life Tips from Reddit Dataset \(13K Tips\)](#). This dataset comprises over 13,000 tips extracted from posts with more than 1,000 upvotes on the subreddits r/lifeProTips and r/YouShouldKnow, spanning the years 2005 to 2022.

The dataset offered a straightforward set of attributes, including 'username,' 'number of comments,' and 'post score.' The 'helpful tips' could be found in the 'postTitle' and 'postBody', I figured that these would be the main field I would be working with. The dataset was obtained in CSV format, and I used MongoDB Compass to import the data into Visual Studio Code. This process was pretty easy once I understood how it worked. The dataset featured simple, non-nested data. The only tweak was that I had to convert the data-type to a string for the 'id' and 'author' fields to ensure integration.

### Querying & Sorting:

Drawing inspiration from class examples, I employed regular expressions to query the database. Specifically, I began by searching for posts with the term 'sleep' in the 'postTitle.' I then extended my search to the 'postBody' and, subsequently, searched for posts that contained 'sleep' in both 'postTitle' and 'postBody.' Comparing the different results of each search informed the rest of my process.

'sleep' in postTitle:



A screenshot of the MongoDB Compass interface. At the top, there is a search bar containing the text "sleep" and a "Get Data!" button. Below the search bar, the heading "the response" is displayed. The main area shows a list of search results, each consisting of a title, body, and comments count. The titles are truncated versions of the original posts, such as "LPT If you feel tired and want to sleep with kids in the house. Tell them to wake you up in about 30 minutes so we can start cleaning the house and they will do literally anything to avoid waking you up.", "LPT: If your pet uncharacteristically starts having random "accidents," do not start scolding as it could be a sign of a serious issue. Mine starting having accidents last week. Today he was put to sleep and all I can think about was how tough I was on him because of things he had no control over.", and "LPT: if you're unable to fall asleep at night instead of closing your eyes do the opposite. Keep your eyes wide open. You'll feel drowsy and will automatically close your eyes. If your mind starts racing again open your eyes again. Keep repeating this process and you'll fall asleep quickly.". The bodies and comments counts are also truncated versions of the original data.

title	body	comments
LPT If you feel tired and want to sleep with kids in the house. Tell them to wake you up in about 30 minutes so we can start cleaning the house and they will do literally anything to avoid waking you up.	undefined	1129
LPT: If your pet uncharacteristically starts having random "accidents," do not start scolding as it could be a sign of a serious issue. Mine starting having accidents last week. Today he was put to sleep and all I can think about was how tough I was on him because of things he had no control over.	undefined	1937
LPT: if you're unable to fall asleep at night instead of closing your eyes do the opposite. Keep your eyes wide open. You'll feel drowsy and will automatically close your eyes. If your mind starts racing again open your eyes again. Keep repeating this process and you'll fall asleep quickly.	I sometimes have trouble falling asleep and this works like a charm everytime.	1502
LPT - the best way to fix a bad sleeping pattern isn't by going to bed early, it's by getting up early	I'm one of those people who constantly find themselves up at 5am watching some men build a swimming pool out of dirt. I used to try and fix my sleeping schedule by going to bed early the next day but obviously I couldn't sleep because I had slept late previous morning. It then took me several days and potentially all-nighter to fix my sleeping pattern. Recently, I've discovered that the best way to recover from this is by setting an alarm for the morning. Yes you may be a bit tired during the day, but it means you will fall asleep earlier that night and hey presto you are back to being awake in the day time. Edit - thank you for the awards! I didn't expect my first big post to be about sleep deprivation but I guess it does make sense.	2378
LPT: If you have a kid who wakes up everyone way too early on Christmas, wrap a small lego set for them in their room. They can wake up and start their Christmas early with building a toy while everyone else can get some extra time to sleep	Edit: thanks for the awards kind strangers! I was absolutely a day/week late with this but while I don't have kids I wanted to share something my dad did for me growing up that made the day really special- the Lego set was something I always looked forward to and thought others might like to try as well- happy holidays!	1755
LPT - When a baby is unhappy we understand they're probably missing something basic: food, water, sleep, warmth, etc. This is true of you, too. If you find yourself unexpectedly in a bad mood, chances are it's a simple fix.	undefined	1370
LPT: Each person's brain has a set number of hours of sleep that is required for proper functioning. Don't listen to your parents, co workers or boss telling you that a human only needs 4-6 hours of sleep. Less sleep over long period can lead to poor memory, mental health issues and even Alzheimer's	For example, I require 7 hours of sleep. On days where I sleep less, I'm annoyed, my memory and concentration ability is affected. I feel mentally sick through the day. Once I went a few days like this and then one day I had a good sleep. I realised how important sleep was. Your brain functions so much better. Everything is more clear. Just pay attention to how you perform on less sleep to understand this. There are many studies showing association of poor sleep with dementia and Alzheimer's. There are studies that show association of poor sleep with high blood pressure and cardiovascular diseases. Edit 1: Many had asked about source for my claims <a href="https://www.nih.gov/news-events/nih-research-matters/sleep-deprivation-increases-alzheimers-protein">https://www.nih.gov/news-events/nih-research-matters/sleep-deprivation-increases-alzheimers-protein</a>	1370

'sleep' in postBody:

sleep  
Get Data!

the response

**title:** LPT If you're planning on visiting San Francisco please for the love of God do not leave ANYTHING of even a vague resemblance of value in your car, or your windows will get smashed and you'll lose it.  
**body:** I'm not talking about a laptop or a purse. I'm talking about a hoodie, a blanket, a travel mug, a USB cable, or heaven forbid a few coins in plain sight. Hell, even kids toys aren't safe. Tinted windows are practically a guarantee your windows will get smashed. The biggest pain in the ass is getting the windows replaced, not necessarily whatever gets stolen. Buddy of mine who used to live in lower Haight got his car windows smashed so often he decided to just leave them down one night. He woke up to find THREE homeless people sleeping in his car.  
**comments:** 12104

**title:** LPT: Pay Attention to the smell of your home when you come back from a trip - that's what it smells like to guests all the time, you just get used to it.  
**body:** Whoa! Front page! And all because I stumbled back in my house, half-asleep and jet lagged from Vegas and noticed my house really smells like my dog. I have so many people to thank! And so many dogs to thank! Like mine, for example. Edit: Thanks for the gold! And to the rest of you, uh, go Bears and Trubisky is the future.  
**comments:** 3301

**title:** LPT: Birds can give you a lot of useful information just by listening to them.  
**body:** As I'm listening to the birds, I realize not a lot of people know this. You can tell a lot about weather and other animals just by listening to birds. I'm not sure how city birds such as pigeons work with this tip, so let me know! A few things you can tell are: 1. Birds go CRAZY when there is a snake nearby. I don't mean like mating season crazy, I mean constantly screaming and sometimes dive-bombing the actual snake. Use this to locate the snake and keep yourself or your family safe. 2. Birds 'go to sleep' when rain is coming and stay quiet in their nests. 3. If it is raining and the birds are still loud time, the rain is going to continue for a long time (8-10+ hours). They're basically panicking and trying to solve the problem of running out of food soon.  
**comments:** 2999

**title:** YSK: the easiest way to make sure your kid fights you on every chore is to avoid saying "thank you."  
**body:** I'm a teacher. Every year, I get at least one parent who tells me that they just expect their kids to do their chores, and they won't ever say "thank you" for doing them. Then these parents wonder why their kids fight them on chores. These same parents often tell me they don't understand why their kids put so much effort into my class. It's really very easy. I thank them for their answers in class. I thank them for helping their classmates. I thank them for picking up after a lab. I thank them when they give their presentations. Every opportunity I have to do it, I say "thank you." My mom always gave me an enthusiastic "thank you!" when my sister and I did chores. As a result, we always did them, often without being asked. We said "thank you" when she made dinner. A home culture of saying "thank you" is a very easy and effective way to improve adult/teen relationships, and you'll have happier teens too. Edit: it's after 10pm here, and I'm off to sleep. Thank you for all your comments and the award. You guys are awesome. Edit: Well, RIP my inbox! I can't reply to all of you, but thank you so much for commenting and thank you to everyone who gave me an award! This is my most popular post ever, and I've loved seeing some of the other perspectives. You're all amazing, and thank you so much!  
**comments:** 1721

**title:** LPT: if you're unable to fall asleep at night instead of closing your eyes do the opposite. Keep your eyes wide open. You'll feel drowsy and will automatically close your eyes. If your mind starts racing again open your eyes again. Keep repeating this process and you'll fall asleep quickly.  
**body:** I sometimes have trouble falling asleep and this works like a charm everytime.  
**comments:** 1502

**title:** YSK: that if something's worth doing, it's worth doing poorly.  
**body:** A wise man once said that, I can't remember where I heard it, but it boils down to this: Why YSK: If something is important, it's usually better to half-ass it than it is to not do it at all. Some examples: Got a big test tomorrow and not up to studying? Study for 15 minutes. A 40 is much better than a 0. Don't feel like working out today? Go for a walk. Too busy for breakfast? Drink a cup of milk. Stayed up too late? Go to bed right now. 3 hours of sleep is superior to none. The pile of dirty dishes too daunting? Wash 5. Can't do the whole yard? Mow the front. The principle can be applied to

Both postTitle \$and postBody:

sleep  
Get Data!

the response

**title:** LPT: If you're unable to fall asleep at night instead of closing your eyes do the opposite. Keep your eyes wide open. You'll feel drowsy and will automatically close your eyes. If your mind starts racing again open your eyes again. Keep repeating this process and you'll fall asleep quickly.  
**body:** I sometimes have trouble falling asleep and this works like a charm everytime.  
**comments:** 1502

**title:** LPT - the best way to fix a bad sleeping pattern isn't by going to bed early, it's by getting up early  
**body:** I'm one of those people who constantly find themselves up at 5am watching some men build a swimming pool out of dirt. I used to try and fix my sleeping schedule by going to bed early the next day but obviously I couldn't sleep because I had slept late previous morning. It then took me several days and potentially and all-nighter to fix my sleeping pattern. Recently, I've discovered that the best way to recover from this is by setting an alarm for the morning. Yes you may be a bit tired during the day, but it means you will fall asleep earlier that night and hey presto you are back to being awake in the day time. Edit - thank you for the awards! I didn't expect my first big post to be about sleep deprivation but I guess it does make sense.  
**comments:** 2378

**title:** LPT: Each person's brain has a set number of hours of sleep that is required for proper functioning. Don't listen to your parents, co workers or boss telling you that a human only needs 4-6 hours of sleep. Less sleep over long period can lead to poor memory, mental health issues and even Alzheimer's  
**body:** For example, I require 7 hours of sleep. On days where I sleep less, I'm annoyed, my memory and concentration ability is affected. I feel mentally sick through the day. Once I went a few days like this and then one day I had a good sleep, I realised how important sleep was. Your brain functions so much better. Everything is more clear. Just pay attention to how you perform on less sleep to understand this. There are many studies showing association of poor sleep with dementia and Alzheimer's. There are studies that show association of poor sleep with high blood pressure and cardiovascular diseases. Edit 1: Many had asked about source for my claims <https://www.nih.gov/news-events/nih-research-matters/lack-sleep-middle-age-may-increase-dementia-risk> <https://www.npr.org/2020/11/16/935475284/scientists-discover-a-link-between-lack-of-sleep-and-alzheimers-disease> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4651462/#:~:text=More%20specifically%20%20when%20one%20sleeps,help%20maintain%20its%20normal%20functioning.> "Until recently, the latest research developments have concluded that sleeping has much more impact in the brain than previously thought. More specifically, when one sleeps, the brain resets itself, removes toxic waste byproducts which may have accumulated throughout the day [2]. This new scientific evidence is important because it demonstrates that sleeping can clear "catabolic" in the brain and help maintain its normal functioning. More importantly speaking, this paper illustrates the different principles of sleep, starting from the non-rapid eye movement (NREM) to the behavioral as well as mental patterns with chronic sleep loss as well as the importance of sleeping acting as a garbage disposal in the body." Edit 2: Yes I agree. Not just Quantity of sleep but Quality of sleep matters as well  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5449130/> Edit 3: Amount of sleep required varies from individual to individual  
<http://healthysleep.med.harvard.edu/healthy/science/variations/individual-variation-genetics> Edit 4: For people saying nobody says that. My mom did. I followed the 6 hour thing for very long till I realised, that wasn't true and I needed 7 hours. I used to wake up at 4.30 AM to push more hours of studies (after 6 hours sleep) man let me tell you. I was extremely sleepy and tired in class. I stopped doing that later. Couldn't keep doing that. When I was a teenager, they never let me sleep over 8AM, even during summer holidays. About Boss and Coworkers....In 5 months I'll become a doctor. Healthcare, depending on your specialty and job is one sector where sleep and mental health is actually ignored. I see my interns/ house surgeons staying awake 36 hours. Sometimes the job requires it. Night duties are a part of the job. Even during our undergraduate it's considered very normal to lose sleep over studying for tests and exams. Most of them sleep hardly 3 - 5 hours before University exams. It has kinda become the norm. And yes I've heard my own friends bragging about how less they slept the previous day. It's pathetic. In our student life these kinda extreme situations happen just before the exams. During internship, half the interns I see are sleep deprived and stressed. Brings me to another point. It's not possible to have a good sound sleep all the time, but we can have good sleep atleast most of the time.  
**comments:** 2254

Recognizing that each post had a 'score,' which according to google is the result of upvotes and downvotes for that post, I aimed to prioritize the most relevant and helpful posts. Initially, I sorted results based on their 'score,' making the results collapsible for easier navigation.

sorted by score:

LPT If you're planning on visiting San Francisco please for the love of God do not leave ANYTHING of even a vague resemblance of value in your car, or your windows will get smashed and you'll lose it.

author: [deleted]  
I'm not talking about a laptop or a purse. I'm talking about a hoodie, a blanket, a travel mug, a USB cable, or heaven forbid a few coins in plain sight. Hell, even kids toys aren't safe. Tinted windows are practically a guarantee your windows will get smashed. The biggest pain in the ass is getting the windows replaced, not necessarily whatever gets stolen. Buddy of mine who used to live in lower Haight got his car windows smashed so often he decided to just leave them down one night. He woke up to find THREE homeless people sleeping in his car.  
score: 95703  
comments: 12104

LPT: Pay Attention to the smell of your home when you come back from a trip - that's what it smells like to guests all the time, you just get used to it.

LPT If you feel tired and want to sleep with kids in the house. Tell them to wake you up in about 30 minutes so we can start cleaning the house and they will do literally anything to avoid waking you up.

LPT: Birds can give you a lot of useful information just by listening to them.

LPT: If your pet uncharacteristically starts having random "accidents," do not start scolding as it could be a sign of a serious issue. Mine started having accidents last week. Today he was put to sleep and all I can think about was how tough I was on him because of things he had no control over.

YSK the easiest way to make sure your kid fights you on every chore is to avoid saying "thank you."

LPT: if you're unable to fall asleep at night instead of closing your eyes do the opposite. Keep your eyes wide open. You'll feel drowsy and will automatically close your eyes. If your mind starts racing again open your eyes again. Keep repeating this process and you'll fall asleep quickly.

YSK: that if something's worth doing, it's worth doing poorly.

LPT - the best way to fix a bad sleeping pattern isn't by going to bed early, it's by getting up early

LPT: If you have a kid who wakes up everyone way too early on Christmas, wrap a small lego set for them in their room. They can wake up and start their Christmas early with building a toy while everyone else can get some extra time to sleep

LPT - When a baby is unhappy we understand they're probably missing something basic: food, water, sleep, warmth, etc. This is true of you, too. If you find yourself unexpectedly in a bad mood, chances are it's

```
{ $match: { $and: [{ postTitle: regexM }, { postBody: regexM }] } },
{ $project : { _id : 0, postTitle : 1, postBody : 1, author : 1, score : 1, numComments : 1 } },
{ $sort : { score : -1 } },
{ $limit:20 }
```

While it was interesting to see top posts, these results felt less relevant to the search. Additionally, trying other search terms (for example 'travel') would often yield the same top result because of its high rating. I noticed that not all posts had content in the 'postBody,' and using '\$and' in my queries might be excluding potentially valuable posts. To address this, I experimented with '\$or' to broaden the search. However, I found that this approach resulted in even less relevant results. So now, my goal was to find a way to prioritize posts that contained the search criteria more frequently and reduce the effect of high ratings on search outcomes.

To enhance the search accuracy, I wanted to find a way to count the occurrences of the match in each post. I spent a long time trying to do this, but ultimately couldn't figure it out. At this point I desperately turned to chatGPT for answers, which only provided me with broken code and several deadends. Ultimately, I created a text index that covered both 'postBody' and 'postTitle.' Instead of relying on regular expressions, I began utilizing MongoDB's built-in text search. The results were sorted by 'textScore,' which uses Mongo's word frequency and term proximity detection to determine relevance.

Sorted by searchScore:

The screenshot shows a search interface with a purple header bar. In the top left corner of the header, there is a search bar containing the text "how to sleep better". Below the search bar is a button labeled "Find Advice". The main content area is a dark grey box containing several text snippets from various sources. The snippets are as follows:

- LPT: Each person's brain has a set number of hours of sleep that is required for proper functioning. Don't listen to your parents, co workers or boss telling you that a human only needs 4-6 hours of sleep. Less sleep over long period can lead to poor memory, mental health issues and even Alzheimer's
- LPT: Sleep better by either avoiding technology or changing it slightly for an hour before you go to bed. Some info inside, please comment for other suggestions.
- LPT : The ultimate sleep guide!
- LPT: To Get Better Sleep, Create a Restful Environment and Limit What you Do in Bed to Just Sleeping!
- YSK: if you have trouble getting out of bed in the morning, try sleeping with your curtains open.
- YSK that sleeping more isn't always better (and I'm not talking about oversleeping)
- LPT: Waking up early isn't necessary for career success, and can be counterproductive for some. What's better is to find a schedule that fits your own chronotype, to maximize health and productivity.
- YSK about sleep apnea. A common medical issue, sleep apnea causes you to stop breathing while you sleep. The disorder affects millions and often goes undiagnosed. Serious sleep apnea can lead to an elevated risk for high blood pressure, diabetes, and more.
- LPT: If you wake up naturally 2 hours or less before your alarm goes off, get up. You are going to feel way more rested and energized than if you slept another hour and let the alarm break your sleep cycle.
- YSK how crucial SLEEP, DIET and EXERCISE are for your quality of life
- LPT: Two secrets to shutting your brain off and going to sleep, from a chronic insomniac.

I discovered that sorting by 'searchScore' based on word frequency and term proximity (`{$meta: "textScore"}`) consistently delivered more accurate results than sorting by postScore. I liked these results, but I felt they could be better. Trying out more complicated searches, multi-word phrases or questions, worked pretty well when sorted by searchScore, but demonstrated the limitations of Mongo's index search system. I still wanted to incorporate the postScore as a factor in the sorting process. Currently, the closest match to the search query may not necessarily provide the best advice or tip. This was becoming difficult, so I used a different method for incorporating postScore: a dropdown menu that enables users to select their preferred sorting method, be it by relevance or post score. This enhanced the user-experience, and I found it fun and interesting to easily compare the different results.

The screenshot shows a search interface with a purple header bar. In the top left corner of the header, there is a search bar containing the text "quit smoking". Below the search bar is a dropdown menu with two options: "relevance" (which is checked) and "post score". The main content area is a dark grey box containing several text snippets from various sources. The snippets are as follows:

- LPT: The one thing that helped me finally quit smoking for good
- LPT: Quitting a habit
- LPT: This is a bit late, but for your new years resolution, dont just set goals that are abstract and vague, set something quantifiable! (examples in description)
- LPT If your boss fires you, do not fire back with, "I quit!"
- YSK: Why do all airplanes still have ashtrays in the lavatories, even though smoking is not allowed.
- LPT: Clear a room full of smoke in just a few minutes.
- YSK: Smoking (and vaping) significantly increases mens' risk of erectile dysfunction
- LPT: when you say or think "X, but Y," change your perspective by thinking about "Y, but X"
- LPT: If you are planning a New Year's resolution, make a calendar of daily goals
- LPT: If it's raining ash, keep track of the size of ash flakes. If they're small, it's more likely to be a nuisance than dangerous. If the flakes grow in size, EVACUATE IMMEDIATELY from wherever you currently are. It means the fire is getting closer and it might be doing so quickly.

```
{
  $match: {
    $text: { $search: searchText }
  },
  {
    $addFields: {
      searchScore: { $meta: "textScore" }
    }
  },
  { $project : { _id : 0, postTitle : 1, postBody : 1, author : 1, score: 1, numComments : 1, searchScore: 1 } },
  {
    $sort: {
      [sortField]: -1
    }
  },
  {$limit:10}
}
```

## Final Thoughts & Future Enhancements:

I would like to reintroduce regular expressions for better accuracy in my searches, which would also entail developing a unique method for determining relevance based on word and phrase frequency. Additionally, I could implement user input options for sorting results in ascending or descending order. While it was sometimes hard to find advice directly relevant to the search, the inaccuracy was actually very insightful. Attempting to find life advice for someone like me, a ‘girl’, mostly just highlighted the disproportionate amount of male Reddit users, and some harsh realities of girlhood.

Search for ‘girl’ sorted by postScore (left) and textScore (right):

The image shows two side-by-side screenshots of a Reddit search results page. Both screenshots have a pink header bar with a search bar containing 'girl'. Below the header, there are two sorting dropdown menus: 'post score' on the left and 'relevance' on the right.

**Left Screenshot (post score sorted):**

- LPT: I'm 43. By your late 20's/early 30's, make sure physical fitness becomes an absolute top priority.
- LPT: when your child tells you anything about a friend of opposite sex don't make jokes about them bleeping together
- LPT: Scratch your girl's back underneath her bra band
- YSK: The police can AND WILL lie to you
- LPT: After a miscarriage, do a few online searches for "where to buy cute dog clothes" so instead of being targeted with ads that remind you of your loss, you will be comforted by cats in tuxedos and puppies in panda costumes
- LPT : stay away from \*that\* older guy
- author: viberson  
mainly directed to teen girls/ young adults (15-20-ish): that older guy who thinks you're 'mature' is lying to you. he knows you're young and finds you vulnerable. stay away, there's a reason they don't go for girls their age. disclaimer: once you're above the age of consent, it's fine to sleep with those also above age. I'm on about very specific people who prey on teens or new adults who can't spot the red flags.  
score: 47263  
score: 0.7875000000000001  
comments: 3508
- LPT: Guys if you're in the drug store buy a 10 pack (\$2) of those black hair rubber bands and keep them in your glove box. I've been the hero to countless women this way.
- LPT: Give your man some physical love and attention

**Right Screenshot (relevance sorted):**

- LPT: How to compliment a girl without coming off as creepy
- LPT: Scratch your girl's back underneath her bra band
- LPT: Guys, if you want to kiss a girl, just ask her first.
- YSK getting a girl really drunk so she'll have sex with you IS RAPE
- LPT: Period guide for dudes
- YSK that the hymen is not an accurate indicator of virginity.
- [LPT] Don't have friends in a quiet apartment complex? Plan a fun event and go door to door knocking inviting people over. It turns out a lot of people are super lonely and it's fairly easy to get a group of random people to do stuff.
- YSK that digging large holes on the beach or elsewhere can be dangerous and you must cover or fill them back in before you leave.
- YSK common misconceptions about sexual consent
- YSK That the HPV vaccine is just as useful for men as it is for women.

## Random Exploration:

Lastly, just for fun, I included a way to retrieve a completely random post from the database. This feature has quickly become my favorite way to explore the collection.

relevance ▼

search Random

YSK "Flushable" wet wipes are not flushable.

author: JaySayMayday  
It's a misrepresentation that continues to be allowed in many nations worldwide. They can and will clog sewage systems. Here's a link with more info for those interested.  
<https://www.chemistryworld.com/news/explainer-are-flushable-wet-wipes-really-flushable/1017594.article>

score: 11288  
score: undefined  
comments: 991