scrape data / api to get data

clean data

insert into database

query from database & conduct analysis

display results on website

highlight corona recoveries

- daily download of csv from internet

- upload new data into database

- pandas data manipulation to create visualization + summary stats

serve cute uplifting content

- use Instagram API to gather urls of favorite posts (wholesome comics, shibes, ducks)

- store into a separate database table

- randomly serve to user (embedded in webpage)

As a baker who prefers following YouTube videos to written recipes, I want to create an index of YouTube videos from my favorite baking channels and build some helper tools on top of that. My goal is to provide a useful tool for bakers like myself to discover new recipes to follow and cut down on wasted ingredients.

My vision is a web app built using Flask and MongoDB where users can choose to start from…

* the ingredients: the user enters what ingredients they have and then gets a list of baking videos that they can follow using those ingredients
* the video: the user selects the video that they wish to follow and then gets a list of baking videos utilizing similar ingredients (allowing for more efficient purchase & allocation of ingredients)

After the user selects the final video(s) they wish to follow, the app would then generate a shopping list compiling all the ingredients needed in their aggregate amounts.

To curate my database, I plan to…

1. use the YouTube API to get basic video data (id, title, description) from my favorite channels (e.g. Cooking Tree, Hanse, AriKitchen)
2. use a combination of conditional statements and regex to pull out a list of likely ingredient strings from the video description
3. run each potential ingredient through the Spoonacular API to parse out the ingredient name, amount, and unit
4. insert each video's data into a MongoDB database as a video object containing an id, title, channel source, and ingredient list
5. update the database on a weekly basis using a crontab or something similar

Additional functionalities if time permits include:

* calculating the most used ingredients across all videos
* determining the most popular dessert categories according to view count

additional aggregate stats – most used ingredients ? most popular dessert categories based on view count (use youtube videos api -- <https://developers.google.com/youtube/v3/docs/videos/list?apix_params=%7B%22part%22%3A%22statistics%22%2C%22id%22%3A%22m1QN6GnWXWI%22%7D>)

create a cron application in order to update database with new videos