Lab 6

ETL, Mongo

DUE: April 2 before class

Part 1
Choose three APIs or data stores/data warehouses/gov't sites with data sets that all provide the same broad genre of data. Create a new collection in your Mongo database (you can name it whatever you'd like).

Download one document from each of the three new APIs/data stores and compare them to each other. Decide what data is important and what data is not important for your application. Devise a schema.

Write a small program in JavaScript that transforms that data so that no matter where it comes from, the result is a single, consistent, format. Load your transformed documents into the new collection.

Repeat this 100 times for each API. That should produce 300 documents that have gone through your ETL pipeline.

Part 2
Extend the React component and Node server/API to be able to GET, POST, PUT, DELETE with this new collection.

Part 3
Add your code from Part 1 into your Node server so that you can make 3 new API endpoints: for each endpoint, running a GET on that endpoint will get data from one of your external sources, perform the whole ETL pipeline right in Node itself, and then put the transformed data into your MongoDB. Extend your component from Part 2 to use these new endpoints.

Part 4
Passwords are important. Here is the gold standard guidelines for digital identity creation: https://pages.nist.gov/800-63-3/sp800-63b.html

Weak Atlas passwords will receive a 50% reduction on your grade.

README.md
As always, please provide a README.md file that has all your citations, a running worklog, where you got stuck, how you got unstuck, etc.

Grading

Part 1:	10 pts.
Part 2:	10 pts.
Part 3:	10 pts.
Creativity:	10 pts.
README.md:	10 pts.
Total	**50 pts.**