Open Street Map Data Wrangling Process

Analysis Process:

Questions:

Given an assumed valid XML file:

What elements exist within that file?

What attributes do those elements have?

What data types are found for those attributes?

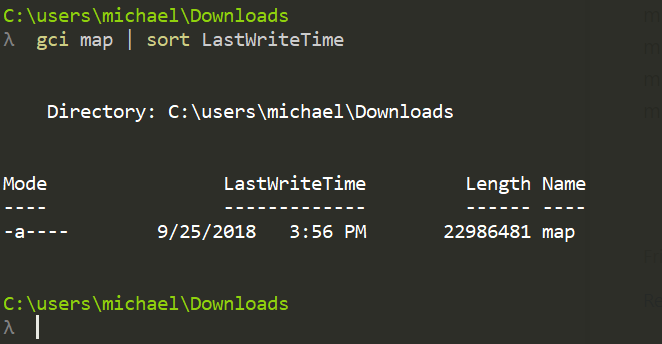
Which elements are nested?

Wrangle

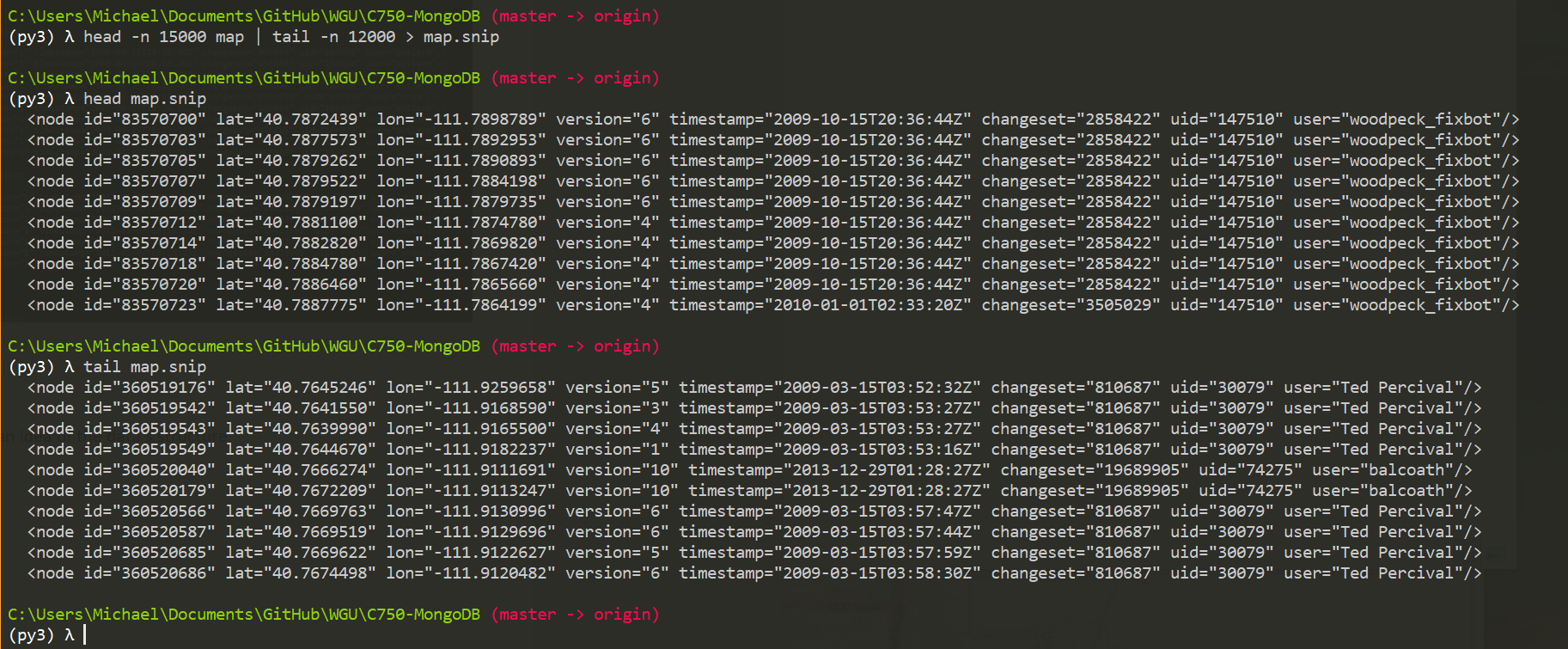
Gather

<https://www.openstreetmap.org/?edit_help=1#map=15/40.7770/-111.8740>

Data source: <https://www.openstreetmap.org/#map=13/40.7478/-111.8452>



How I made the 1.2mb Snip.



Extract

Clean

Store

Explore

Started with a head and tail to get an idea of the data’s structure



This leads me to question what tag structures exist within this data set.

> db.map.find({}, {"\_id":0, "name":1, "nested\_elements":1}).pretty()

{ "name" : "note", "nested\_elements" : null }

{ "name" : "meta", "nested\_elements" : null }

{ "name" : "bounds", "nested\_elements" : null }

{ "name" : "node", "nested\_elements" : [ "tag" ] }

{ "name" : "tag", "nested\_elements" : null }

{ "name" : "nd", "nested\_elements" : null }

{ "name" : "way", "nested\_elements" : [ "nd", "tag" ] }

{ "name" : "member", "nested\_elements" : null }

{ "name" : "relation", "nested\_elements" : [ "tag", "member" ] }

{ "name" : "osm", "nested\_elements" : null }

>

Document Descriptors

Map Elements

Element Descriptors

Conclude/Predict

Communicate