Lab Report

Title: GIS 5572 Lab1
Notice: Dr. Bryan Runck
Author: Haihan Wang
Date: Feb 11, 2021

Project Repository: https://github.com/HaihanW/GIS5572.git
Google Drive Link: None

Abstract

The Goal of this lab is to practice decomposing interfaces for spatial web API's into informal conceptual models. At same time, comparing and contracting different web API's using informal conceptual models and custom-built extract, transform, and load (ETL) routines. Meanwhile, building an ETL pipeline with Open Source Tools in Esri's Online and ArcPro Jupyter Notebook. The data are getting from the conceptual models for the Minnesota Geospatial Commons API, Google Places API, and NDAWN API.

Problem Statement

Practicing decomposing interfaces for spatial web API's into informal conceptual models. Comparing and contracting different web API's using informal conceptual models and custom-built extract, transform, and load (ETL) routines. Building an ETL pipeline with Open Source Tools in Esri's Online and ArcPro Jupyter Notebook.

Table 1. Data table

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Load data.	Data from online to notebook.	Data of Minnesota		Minnesota Geospatial Commons	
2	Load data.	Data from online to notebook.	Data of places		Google Places	
3	Load data.	Data from online to notebook.	Data of North Dakota weather		NDAWN	

Input Data

The conceptual models for the Minnesota Geospatial Commons API, Google Places API, and NDAWN API.

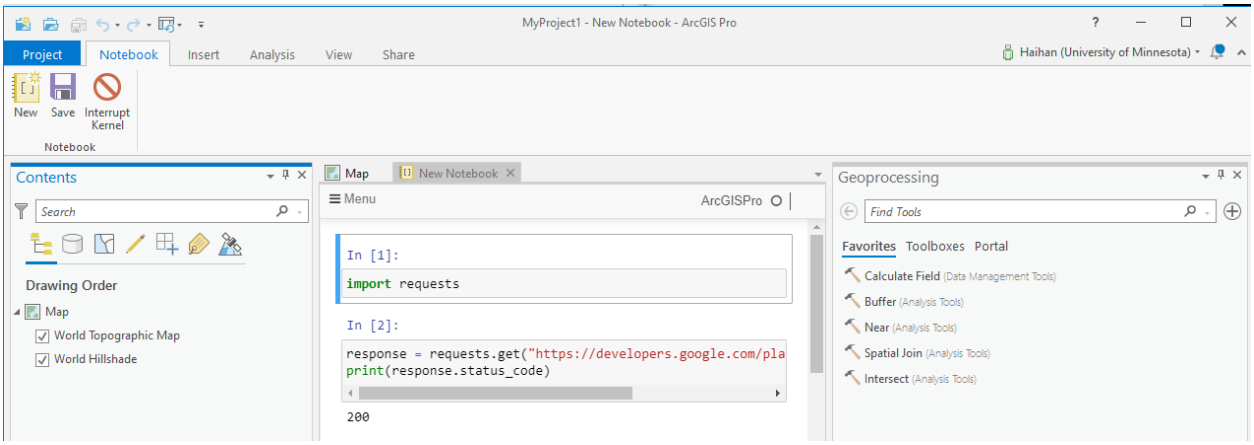
Table 2. Data table

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Geospatial Commons API	Programmatically get data from the API.	Minnesota Geospatial Commons

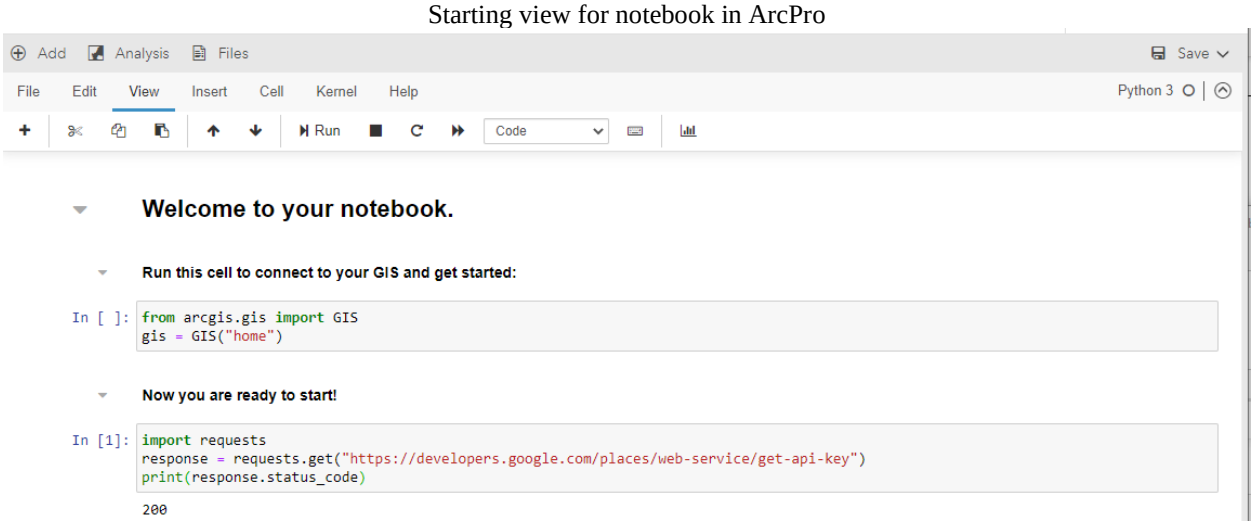
2	Google Places API	Programmatically get data from the API.	Google Places
3	NDAWN API	Programmatically get data from the API.	NDAWN

Methods

The major process is setting notebooks in both online and ArcPro to write code to process the data into the notebook with different APIs.



Starting view for notebook in ArcPro



Starting view for notebook in online

With the code, we could input any online data into the notebook in both environment with APIs.

Results

The results shows that the code is working with the different APIs from the different data website.

```
response = requests.get("https://developers.google.com/places/web-service/get-
print(response.status_code)
```

```
200
```

```
nd = requests.get("https://ndawn.ndsu.nodak.edu/")
print(nd.status_code)
```

```
200
```

```
mgc = requests.get("https://resources.gisdata.mn.gov/pub/gdrs/data/pub/us_mn_s
print(mgc.status_code)
```

```
200
```

```
gpcurl = "https://maps.googleapis.com/maps/api/"
mgurl = "https://gisdata.mn.gov/"
ndurl = "https://ndawn.ndsu.nodak.edu/get-table.html?"
```

```
mgcapi = "pub/gdrs/data/pub/us_mn_state_dot/bdry_jurisdiction/metadata/metadat
mgcdata = requests.get(mgcurl + mgcapi)
print(mgcdata.status_code)
```

```
200
```

```
ndapi = "station=11&variable=mdapet&year=2021&ttype=monthly&quick_pick=&begin_
nddata = requests.get(ndurl + ndapi)
print(nddata.status_code)
```

```
200
```

Results Verification

To verify the results, we can see the print status_code function showing the number 200 which prove the data is available to work and use. The input data can be print out to see if it matches the data shows on the websites.

Discussion and Conclusion

There is no major difference between get the data in ArcPro or online. The different types of APIs have different way to present the data and they could all be processed with similar code with small changes to the parts that need different APIs.

References

Waibel, X. (2020, March 06). Integrate Jupyter notebook into data pipelines. Retrieved February 11, 2021, from https://towardsdatascience.com/integrate-jupyter-into-your-data-pipeline-9a02fab3cee5

Self-score

Fill out this rubric for yourself and include it in your lab report. The same rubric will be used to generate a grade in proportion to the points assigned in the syllabus to the assignment.

Category	Description	Points Possible	Score
Structural Elements	All elements of a lab report are included **(2 points each)**: Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score	28	**25**
Clarity of Content	Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**.	24	**21**
Reproducibility	Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified.	28	**25**
Verification	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**.	20	**20**
		100	**91**