**Lab Report**

Title: Comprehensive flood assessment of the Red River

Notice: Dr. Bryan Runck

Author: Maochuan Wang

Date: Sep 27, 2022

**Project Repository:** *https://github.com/MaochuanW/GIS5571/tree/main/Final%20Project*

**Google Drive Link:** N/A

**Time Spent:** 5 hrs

**Abstract**

This final project report focuses on the flood assessment of Red River. The data used for this lab includes Digital elevation models, WorldView 2-meter high-resolution imagery, precipitation data from NOAA, and housing value API from the housing website. First, an ETL was developed to retrieve precipitation data from NOAA. Then, for the dem analysis, a jupyter notebook will be developed by using various python packages. For land cover changes, land classification tools from arc pro will be used. In the end, with all the data being used, I will be able to create a vulnerability map for the Red River.

**Problem Statement**

Flooding is one of the most detrimental natural disasters in this world. Using GIS technology to conduct flood assessments is crucial to prevent future flooding and mitigate risks. With the combination of using ETL, python, jupyter notebook, and data pipeline, I am going to create a project that takes in data and conducts geospatial analysis to visualize the impact of flooding.

Table 1. data requirement

| **#** | **Requirement** | **Defined As** | **(Spatial) Data** | **Attribute Data** | **Dataset** | **Preparation** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | DEM | Digital Elevation Model | Raster data |  | [EarthDEM](https://www.pgc.umn.edu/data/earthdem/) |  |
| 2 | API | Application programming interface | code |  | [Realty Mole API](https://www.realtymole.com/) |  |
| 3 | Precipitation | Precipitation data | shapefile |  | [NOAA](https://www.cpc.ncep.noaa.gov/products/GIS/GIS_DATA/) |  |
| 4 | Imagery | Satellite Imagery | raster data |  | [PGC](https://www.pgc.umn.edu/data/commercial-imagery/) |  |

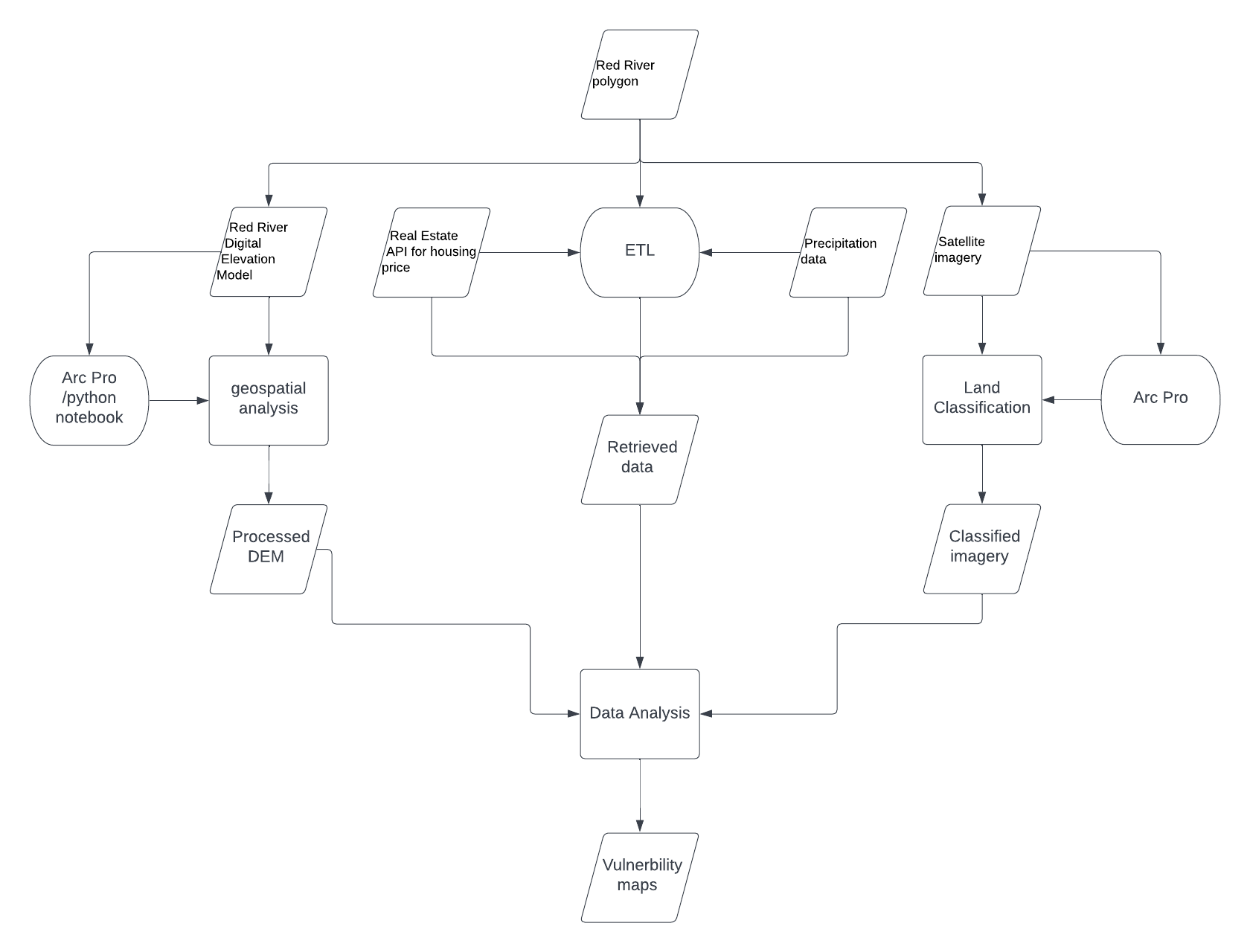
**Input Data**

There are 5 main data I will be using for this project. The first one is the Red River DEM to identify flood areas and find routes for draining water. The second data is the real estate API to calculate the housing price. The third data is the precipitation data to assess the impact on how rainfall and other precipitation can affect flood intensity. The 4th data is a polygon that provides the location. The 5th data is the satellite imagery that can use for landcover analysis for flooding.

Table 2. Data will be used for this project

| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| --- | --- | --- | --- |
| 1 | Red River Digital Elevation Model | Identify flood-affected locations, suitable places for shelter, and routes for draining water…. | [EarthDEM](https://www.pgc.umn.edu/data/earthdem/) |
| 2 | Real Estate API for housing price | Calculates how flood can affect property value | [Realty Mole API](https://www.realtymole.com/) |
| 3 | Precipitation data | Assess the impact on how precipitation can affect flood season and flood intensity | [NOAA](https://www.cpc.ncep.noaa.gov/products/GIS/GIS_DATA/) |
| 4 | Red River polygon | Provide the location of the river in relation to other datasets | Arc pro |
| 5 | Satellite imagery | Landcover analysis pre and after flood change | [PGC](https://www.pgc.umn.edu/data/commercial-imagery/) |

**Methods**



As shown in the flow chart above, this project involves with multiple processes that intertwined with each other. Starting with the Red River polygon, this vector file will serve as the location shapefile that tells me the location of the data that I should look for. The next step is to create an ETL that takes in data from real estate and precipitation API, this step will be conducted by using python jupyter notebook. For the Red River DEM, I will be using some geospatial analysis tools in jupyter notebook or arc pro to get the processed DEM that shows the desired results. The satellite imagery will be used to assess land cover change, and I will be using Arc Pro to do the land cover classification. In the end, I will combine all the processed data and display them on the vulnerability maps I will create in order to show the final result.

**Results**

The result of my final project will be demonstrated in vulnerability maps. I’m hoping it will be multiple maps with different categories, such as flood prediction map, areas prone to flood map…etc.

**Results Verification**

I can compare my results to other flood assessment results on the internet, or in other journal articles. For example, I could compare my precipitation data to some weather websites, and compare the flood area with historical satellite imagery. I could also verify my result by checking thoroughly with my code in jupyter notebook.

**Discussion and Conclusion**

I believe that this final project idea is a great way to utilize what I’m learning in this class. It uses a lot of data and a lot of analysis both from python jupyter notebook and arc pro. There is also a way to build an ETL to retrieve data from some website APIs, which is a crucial part of the final project. Overall, I think it will be an interesting project to work on, and the results will be significant and meaningful to the communities that suffer from flooding every year near the Red River.

**References**

Earthdem. Polar Geospatial Center. (n.d.). Retrieved September 28, 2022, from <https://www.pgc.umn.edu/data/earthdem/>

Realty mole. Realty Mole. (2020, July 22). Retrieved September 28, 2022, from <https://www.realtymole.com/>

Climate prediction center - GIS data (Shapefile and raster). (n.d.). Retrieved September 28, 2022, from https://www.cpc.ncep.noaa.gov/products/GIS/GIS\_DATA/

**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 | **28** |
| **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 | **24** |
| **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 | **28** |
| **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 | **100** |