Lab Report

Title: Lab 2 Part 1
Notice: Dr. Bryan Runck
Author: Carl Sones
Date: 11/2/22

Project Repository: *https://github.com/casones/GIS5571.git*
Time Spent: 4.5 hours

Abstract
The goal of this lab was to get more experience with building ETL pipelines for GIS data analysis. This lab focused on using an ETL pipeline to get LiDAR and PRISM data from online directories and store them locally in order to perform analysis. One goal of the analysis was to convert LiDAR data from MN DNR and convert it to raster data in the form of a DEM and a TIN file. The other goal was to use PRISM data to create a multidimensional raster layer that would then be used to generate a Space Time Cube.

Problem Statement
There are two problems to be dealt with in this lab. Both involve retrieving data from an API and storing it locally with an ETL pipeline. One of the problems requires using LiDAR data and converting it from LAS files to a DEM and a TIN. The other problem requires converting PRISM data to a multidimensional raster layer, then to a Space Time Cube showing average precipitation in the United States (Table 1).

Table 1. Part 1 requirements for analysis

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	LiDAR Data	LAS Files	Point Cloud		MN DNR	Extract to local folder and convert to raster data
2	PRISM Data	Raster dataset containing precipitation data for the contiguous United States	Raster	Precipitation	PRISM	Convert BIL data to Multidimensional Raster, then Space Time Cube

Input Data

The data needed for this lab was LiDAR data from the MN DNR to be used for creating a DEM and TIN, as well as PRISM data containing precipitation information from the Oregon State directory to be used for creating a multidimensional raster dataset, then a Space Time Cube.

Table 2. Part 1 input data

#	Title	Purpose in Analysis	Link to Source
1	LiDAR Data	Raw LiDAR data as an input for analysis	[MN DNR]
2	PRISM Data	Data containing precipitation information in the U.S.	[PRISM]

Methods

For the LiDAR data analysis, I started with creating an ETL pipeline in ArcGIS Pro Jupyter Notebooks for retrieving the raw LAS files from the MN DNR, then extracting the zipped files into a local folder. With this data now stored locally, I then began the analysis in my script by using the LAS file as an input for both the LAS Dataset to TIN and LAS Dataset to Raster tools in arcpy. With the output feature layers, I then attempted to export them as PDF's using arcpy.mp.ArcGISProject and lyt.exportToPDF.

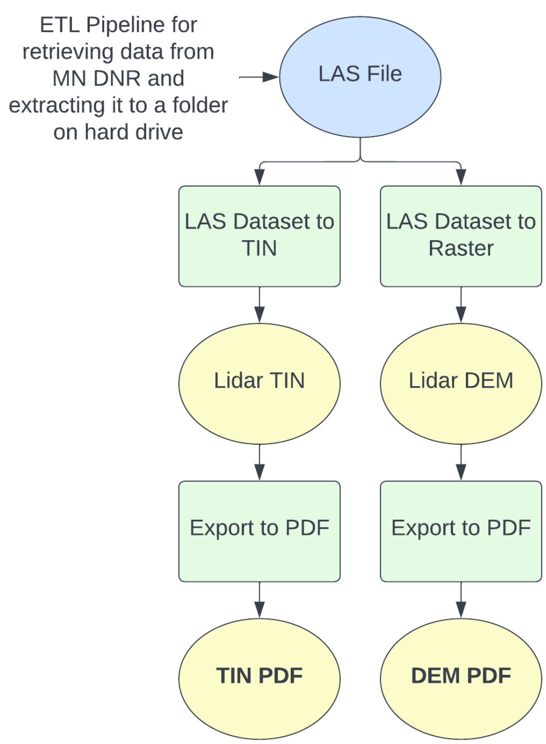

Figure 1: Data flow diagram for LiDAR analysis

For the PRISM data analysis I began by using an ETL pipeline in Jupyter Notebooks to retrieve the data from the Oregon State directory online and extract it to a local folder created just for the BIL files. I then created a mosaic dataset in the dataset using the Create Mosaic Dataset tool in arcpy and used the Add Rasters to Mosaic Dataset tool in order to add the PRISM data to the mosaic dataset I had created. With the PRISM data now in the mosaic dataset, I then used the Calculate Field tool to add precipitation as a variable in the attributes for the footprint. I also used Add Field to create a field for the timestamp to provide temporal information to be used in the process of generating a Space Time Cube. I then used the Build Multidimensional Info tool on the PRISM dataset with updated fields, and used the Make Multidimensional Raster Layer tool on the output layer to create a multidimensional PRISM layer. This layer was then the input for the Create Space Time Cube from Multidimensional Raster Layer tool. I then created a 3D PRISM Space Time Cube from the output layer using the Visualize 3D Space Time Cube tool.

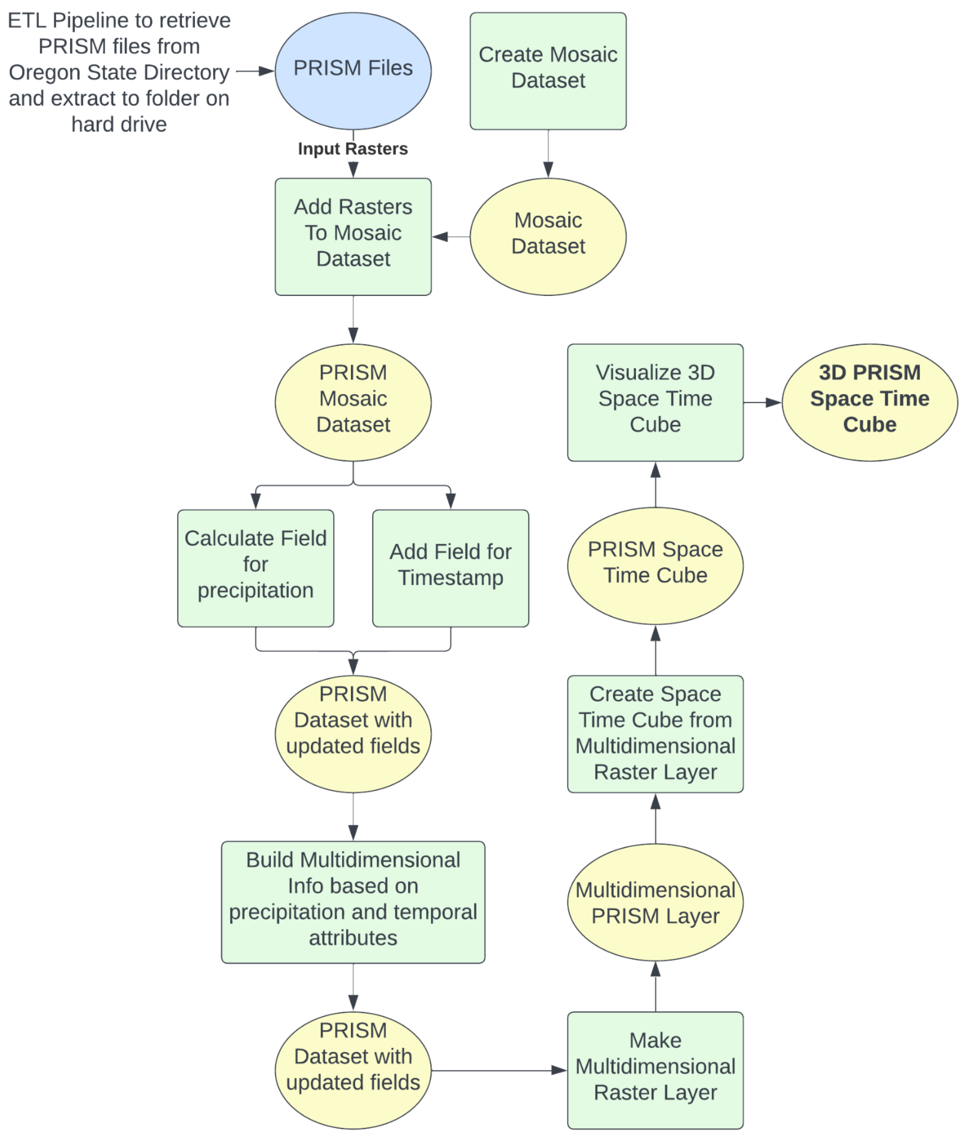

Figure 2: Data flow diagram for PRISM data analysis

Results
The results for the LiDAR data analysis were two raster layers, a DEM and TIN, shown in Figures 3 and 4.

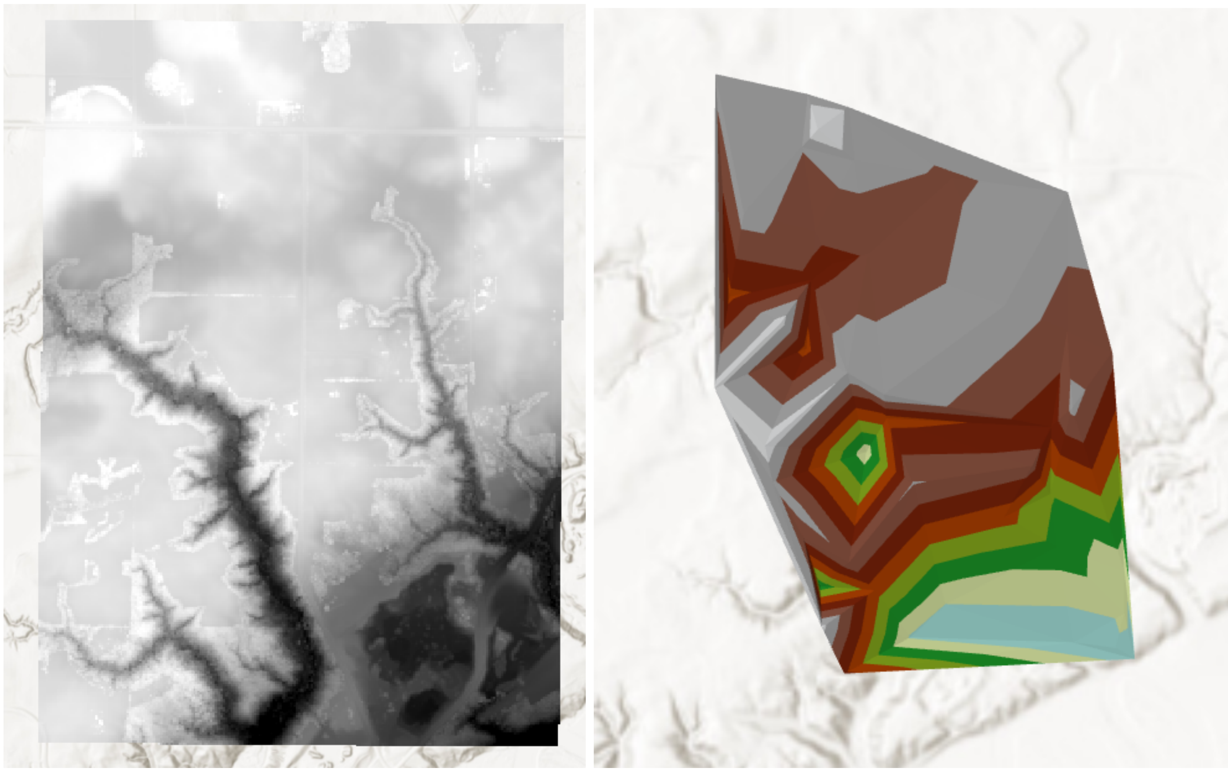

Figure 3 & 4: Results of LiDAR data analysis

The final 3D Space Time Cube resulting from the PRISM data analysis can be seen in Figure 5.

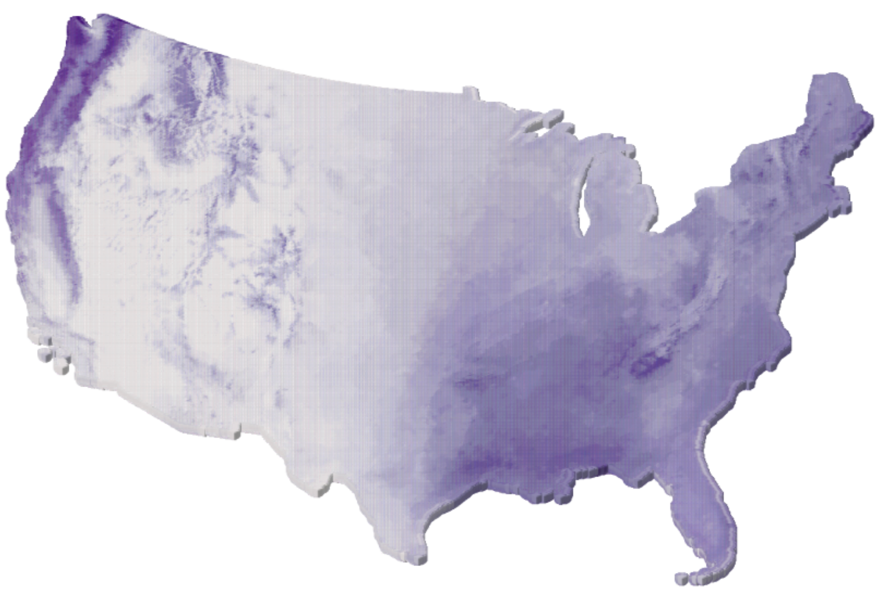

Figure 5: Results of PRISM data analysis

Results Verification

To verify my results I manually inspected the feature layers I had created in a map in the ArcGIS Pro project to confirm that my analysis was successful. I also compared my results with other classmates to make sure that I was getting similar results to others.

Discussion and Conclusion

Besides getting more experience building ETL pipelines to retrieve data from APIs, I also learned more about how LiDAR data can be used to create different types of raster data and got an idea for how it can be utilized for a variety of applications. I also learned what Space Time Cubes are and how simple it is to create them.

References

N/A

Self-score

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	**27**
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	**22**
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	**18**
		100	**95**