Lab Report

Title: **Lab 2 – PART 1**
Notice: Dr. Bryan Runck
Author: Kyle Smith
Date: October 29, 2024

Project Repository: /kylejsmith4/GIS5571/Lab1
Time Spent: 118.5 hours (all of Lab 2)

Abstract
This lab involves several multi-step Extract, Transform, and Load (ETL) processes using Arc GIS Pro and ArcPy in Jupyter Notebooks. First, an ETL pipeline to access LiDAR (.LAS) files from the Minnesota Department of Natural Resources, convert them into Digital Elevation Models (DEM) and Triangulated Irregular Networks (TIN), and output results for review. A side-by-side analysis of 2D & 3D outputs of LiDAR data will help to understand the data and the processes. Finally, a separate ETL pipeline that involved downloading annual 30-Year Normal precipitation data, preparing a spacetime cube in Arc GIS Pro, and creating a simple animation. Part 1 of this Lab explores a variety of spatial data types and processes to help enhance understanding of ETL pipelines.

Problem Statement
This lab focuses on developing spatial data pipelines using fundamental skills such as API queries, creating cubes, and working with TIN and terrain data. The project involves building an Extract, Transfer, and Load (ETL) system for LiDAR data and analyzing multidimensional datasets. The goal is to integrate various geospatial data transformation techniques and leverage ArcGIS Pro's capabilities for 2D and 3D visualization.

Table 1.

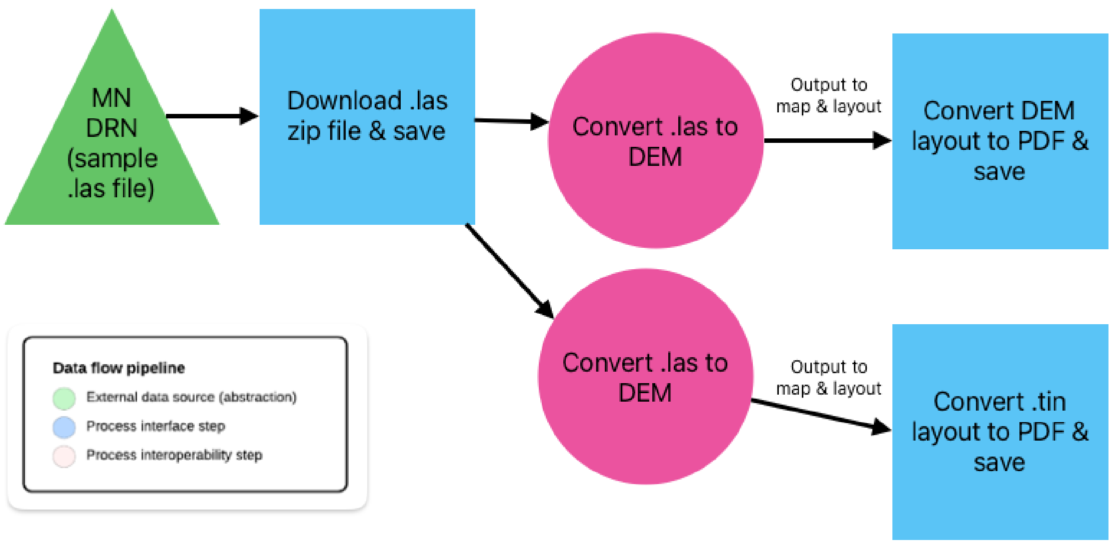
#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Download sample LiDAR (.las) from MN DNR	.las files	LiDAR (.las) points, DEM files (.tif), Triangulated Irregular Networks (.tin) raster	Three Dimensional representation of landforms & elevation	DNR FTP server Specifically sample dataset "4342-12-05.las"	Using ArcPy, do the following: - Download .las file from MN DNR - Convert the .las file in to both a DEM and .tin file - Save and export PDFs of both - Side-by-side analysis of 2D and 3D output of the sample .las file
2	30-Year Normal Precipitation data for the United States	Raw input dataset	Band Interleaved by Line (.bil) rasters, spacetime cube shapefile, animation output	Precipitation and time data	PRISM Climate Group at Oregon State Univ. Specifically, 30-year normal precipitation data.	Using ArcPy, do the following: -Download .bil files from PRISM -Convert data in to spacetime cube -Export an animation

Input Data

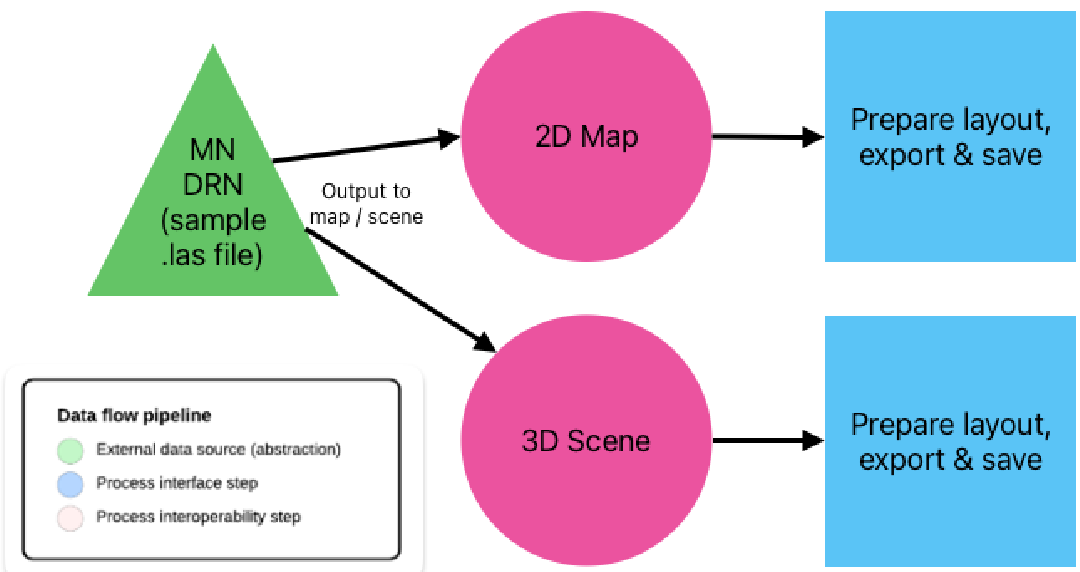
#	Title	Purpose in Analysis	Link to Source
1	LiDAR .LAS Files from MN DNR	Data for elevation modeling, DEM & TIN creation, and multidimensional visualization	DNR FTP server Specifically sample dataset "4342-12-05.las"
2	PRISM 30-Year Normals Precipitation .bil Files	Data for climate analysis (precipitation) and spacetime cube development.	PRISM Climate Group at Oregon State Univ. Specifically, 30-year normal precipitation data.

Methods

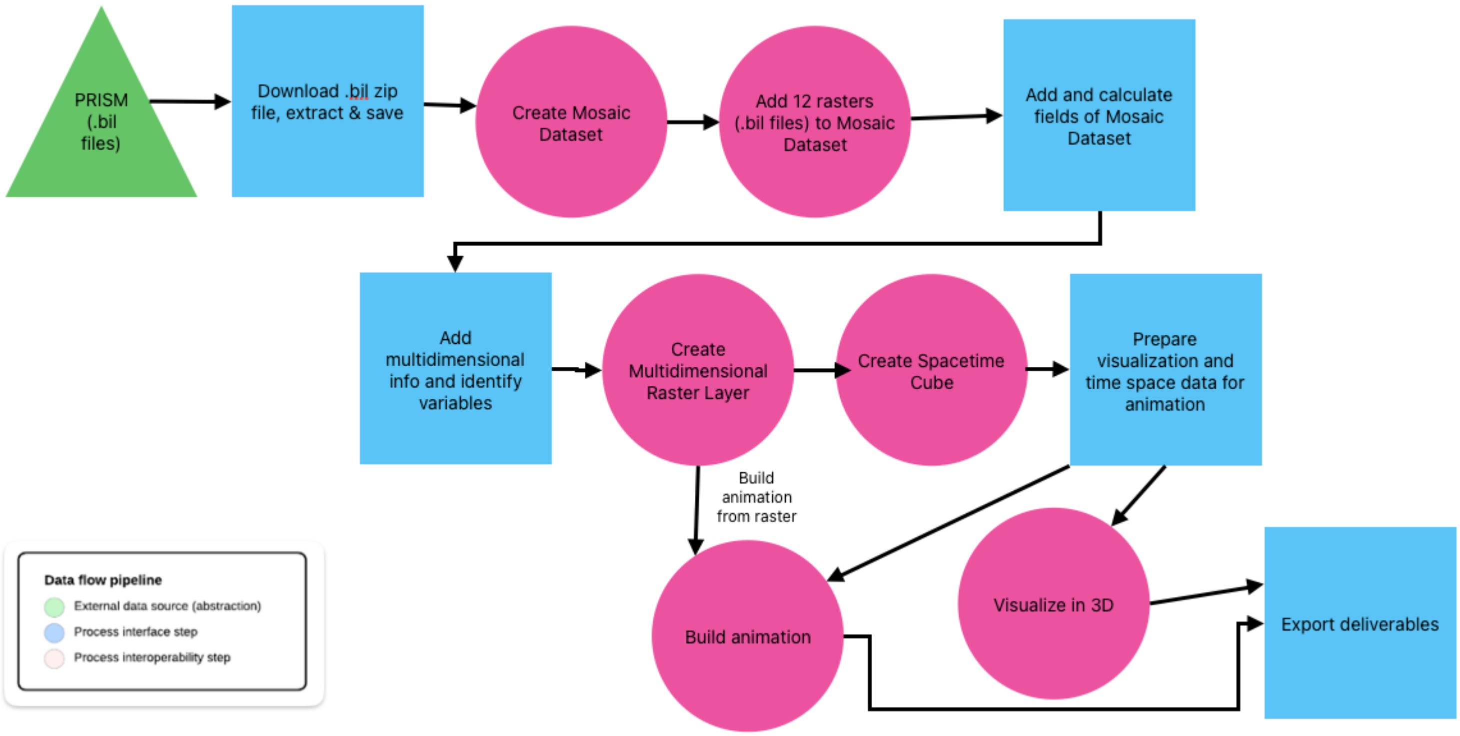
LiDAR ETL



Exploratory data analysis: 2D & 3D visualizations



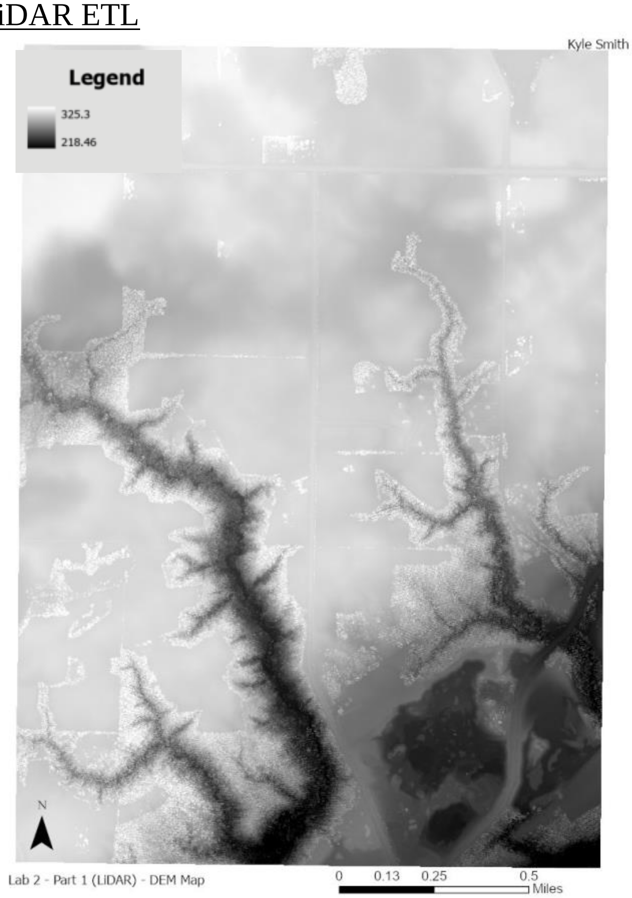
PRISM data: Build ETL and spacetime cube

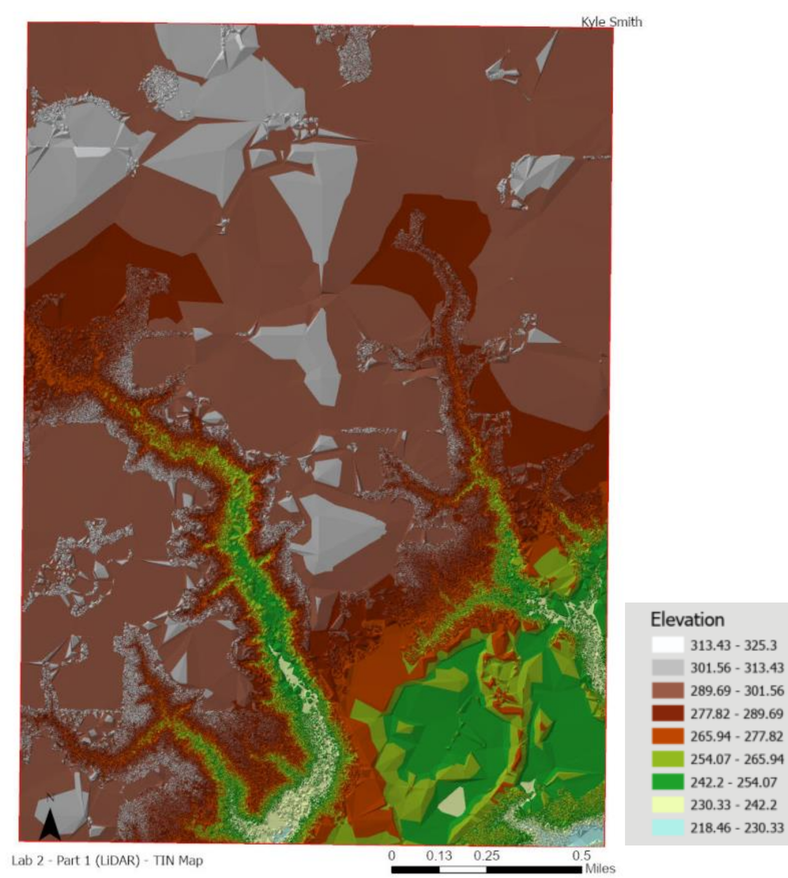


Data flow diagrams from Apple Freeform app – works great!

Results Verification

LiDAR ETL

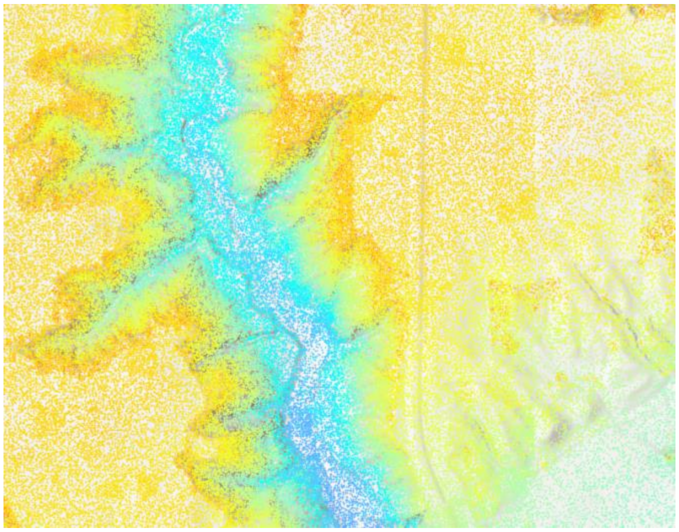




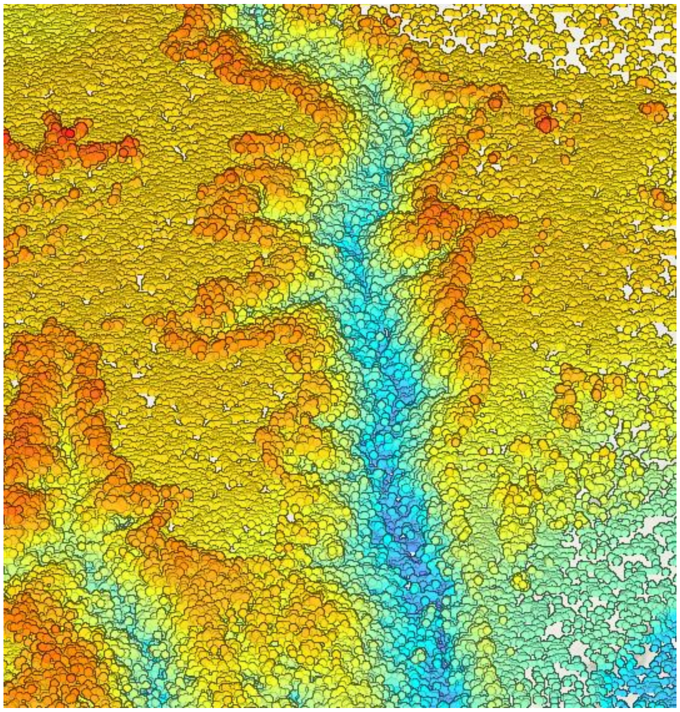
The DEM (Digital Elevation Model) raster map of the selected location represents elevation and landforms in a 2D format. DEMs are much more efficient in analyzing larger areas, however the topography accuracy isn't as great. The .tin map (Triangulated Irregular Network) vector map is a color collection of irregularly points and polygons and can show a more accurate representation of elevation change and landform features. .tin files are often used in precise detailing, such as engineering plans. The accuracy of both models was confirmed by comparing the data with known datasets.

2D / 3D Visualizations

2D Map of .las file

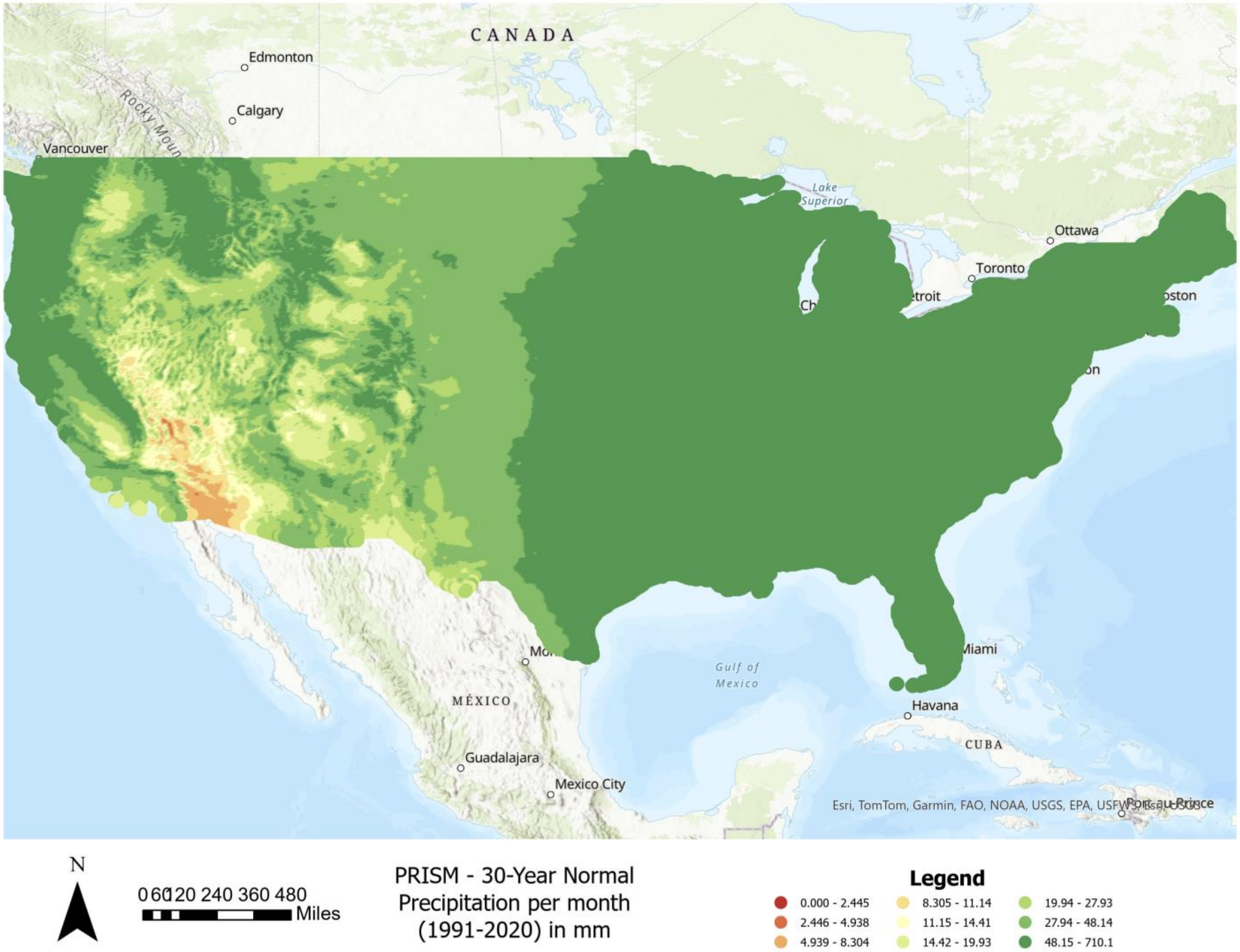


3D Scene of .las file



In reviewing the 2D map and 3D scene of roughly the same location in the same LiDAR file, we can review the different visualizations and methods. The 2D map allows for simple visualization showing elevation through color and density of points. The 3D scene allows for a more comprehensive visualization of the terrain by adding the third dimension. In addition to comparing these two models side-by-side, ArcGIS Pro has geoprocessing tools and other features to provide custom symbology and exaggeration to help better understand complex features in the real world.

PRISM ETL



To analyze the 30-year precipitation normals, we processed the PRISM data and converted it into a Spacetime Cube. This was a complex process which involved transforming .bil files into a single Mosaic dataset and eventually a multidimensional raster. The spacetime cube requires a temporal field, in this case specific years between 1991 and 2020. Further, a variable representing the monthly sum of rainfall at a station in millimeters was prepared and added. After a lengthy period of rendering, the spacetime cube can be viewed in 3D as to demonstrate nationwide shifts in rainfall over the 30-year period. I did find that viewing the timespace cube data in 2D was a somewhat faster way to view the data. As with each of these steps, beast practices in map design were followed, including color ramp, legends, titles, and other labels. To attempt to ensure data is correct, checking datetime format, precipitation units, and consistency in field labeling was checked. My spacetime cube contained over 3,000,000 points – making it challenging in several ways.

One way I addressed this issue of very slow rendering was by adjusting the number of classes and increasing the size of the points/bins. Ultimately, I settled on 9 classes, with a 9-color color ramp, points/ bins at size 8 with no outline drawn around a point. Further, I set data normalization to “Log” and set a Definition Querrey to ensure no negative values resulted. I realize this may not be an ideal representation of this data, and with the benefit of more time I could have played around with this more.

With regard to the animation of the timespace cube, I was not able to successfully complete that task in time for the Lab deadline. I tried!

Discussion and Conclusion

While I did spend a significant amount of time on this project, and it was frustrating due to lengthy rendering and loading times, I did feel satisfied with new familiarity and practice in ARC Pro features. I don't believe my challenges in computing power are a specific problem with my hardware, however I am very much interested in tips and best practices to right-size data and ensure smooth software processing. The ETL pipelines and investigations of various geospatial data in different dimensions in this Lab very much increased my awareness and skills in Arc Pro.

References

Esri. (n.d.). *Explore your raster data with space-time pattern mining*. Retrieved from https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/explore-your-raster-data-with-space-time-pattern-mining/

Esri. (n.d.). *What is ArcPy?*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/arcpy/get-started/what-is-arcpy-.htm

Esri. (n.d.). *Create DEMs and DSMs from LiDAR*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/analysis/3d-analyst/create-dems-and-dsms-from-lidar.htm

Esri. (n.d.). *Use LiDAR in ArcGIS Pro*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/data/las-dataset/use-lidar-in-arcgis-pro.htm

Esri. (n.d.). *BIL format example*. Retrieved from https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/bil-format-example.htm

Esri. (n.d.). *Mosaic*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/mosaic.htm#:~:text=Raster%20Dataset%20toolset,-Copy%20Raster%20Create%20Random%20Raster

Esri. (n.d.). *An overview of multidimensional raster data*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/data/imagery/an-overview-of-multidimensional-raster-data.htm

Esri. (n.d.). *Create Space-Time Cube*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/tool-reference/space-time-pattern-mining/learnmorecreatecube.htm

Esri. (n.d.). *Animate through time*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/mapping/animation/animate-through-time.htm

Minnesota Geospatial Information Office. (n.d.). *LiDAR data resources*. Retrieved from https://resources.gisdata.mn.gov/pub/data/elevation/lidar/

PRISM Climate Group. (n.d.). *PRISM Climate Data*. Retrieved from https://prism.oregonstate.edu/normals/

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	**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	**23**
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	**15**
		100	**85**

Animation of spacetime cube is missing