**Lab Report**

Title: Exploratory Data Analysis with .lis and .bil Filetypes

Notice: Dr. Bryan Runck

Author: Jake Ford

Date: 10-12-2022

**Project Repository:** [**https://github.com/ThisFord/GIS5571-arc1.git**](https://github.com/ThisFord/GIS5571-arc1.git)

**Google Drive Link:**

**Time Spent:** *4+*

**Abstract**

*<Delete this text in light grey throughout>*

*250 words max. Clearly summarize the following major sections. Each gets one or two sentences.*

**Problem Statement**

Working with data in ETL pipelines is essential for proficiency in GIS. This project builds on previous working knowledge of ETL pipelines by incorporating new datasets, implementing exploratory data analysis, and producing a simple geovisualization in the form of a PDF export through ArcPro based Jupyter Notebooks.

*Table 1. <insert caption>*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **(Spatial) Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | Road network | Raw input dataset from MNDOT | Road geometry |  | [Mn GeoSpatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) |  |
| 2 | High volume traffic | > 100 cars per hour |  | Volume | AADT Data |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |

**Input Data**

*Describe the data in two paragraphs max. Fill out the table.*

*Table 2. <insert caption>*

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | Minnesota Roads | Raw input dataset for routing analysis from MNDOT | [Mn GeoSpatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) |
| 2 |  |  |  |
| 3 |  |  |  |
|  |  |  |  |

**Methods**

*Include a data flow diagram or screenshot from model builder. Do references in line (Rammankutty, 2033). Document any and all steps that you did to the input data in the data flow diagram. Provide natural language description of the most important steps, giving a narrative arc and provide well formatting screenshots with a boarder and centered throughout.*

*Resources on Data Flow Diagrams:*

* [*https://www.visual-paradigm.com/tutorials/data-flow-diagram-dfd.jsp*](https://www.visual-paradigm.com/tutorials/data-flow-diagram-dfd.jsp)
* [*https://www.lucidchart.com/pages/data-flow-diagram/how-to-make-a-dfd*](https://www.lucidchart.com/pages/data-flow-diagram/how-to-make-a-dfd)

*Figure 1. Data flow diagram.*

*If appropriate, add in pseudo-code describing model algorithms and/or objects. If using mathematical equations, create a clear mapping between the reference equation, pseudo-code, and actual implementation in a programming language.*

**Graphical user interface, application

Description automatically generatedDiagram, schematic

Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generated**

**Results**

*Show the results in figures and maps. Describe how they address the problem statement.*

*Follow best practice for map design, coloring, etc.*

**Results Verification**

*How do you know your results are correct? This can be a qualitative or quantitative verification.*

**Discussion and Conclusion**

*What did you learn? How does it relate to the main problem?*

**References**

*Use a common format*

**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 |  |
| **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 |  |
| **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 |  |
| **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 |  |
|  |  | 100 |  |