

WOMEN IN DATA SCIENCE
CALGARY @SCHULICH
SCHOOL OF ENGINEERING

AUTOMATING GEOPROCESSING TASKS TO CREATE CITY WIDE DATA LAYERS

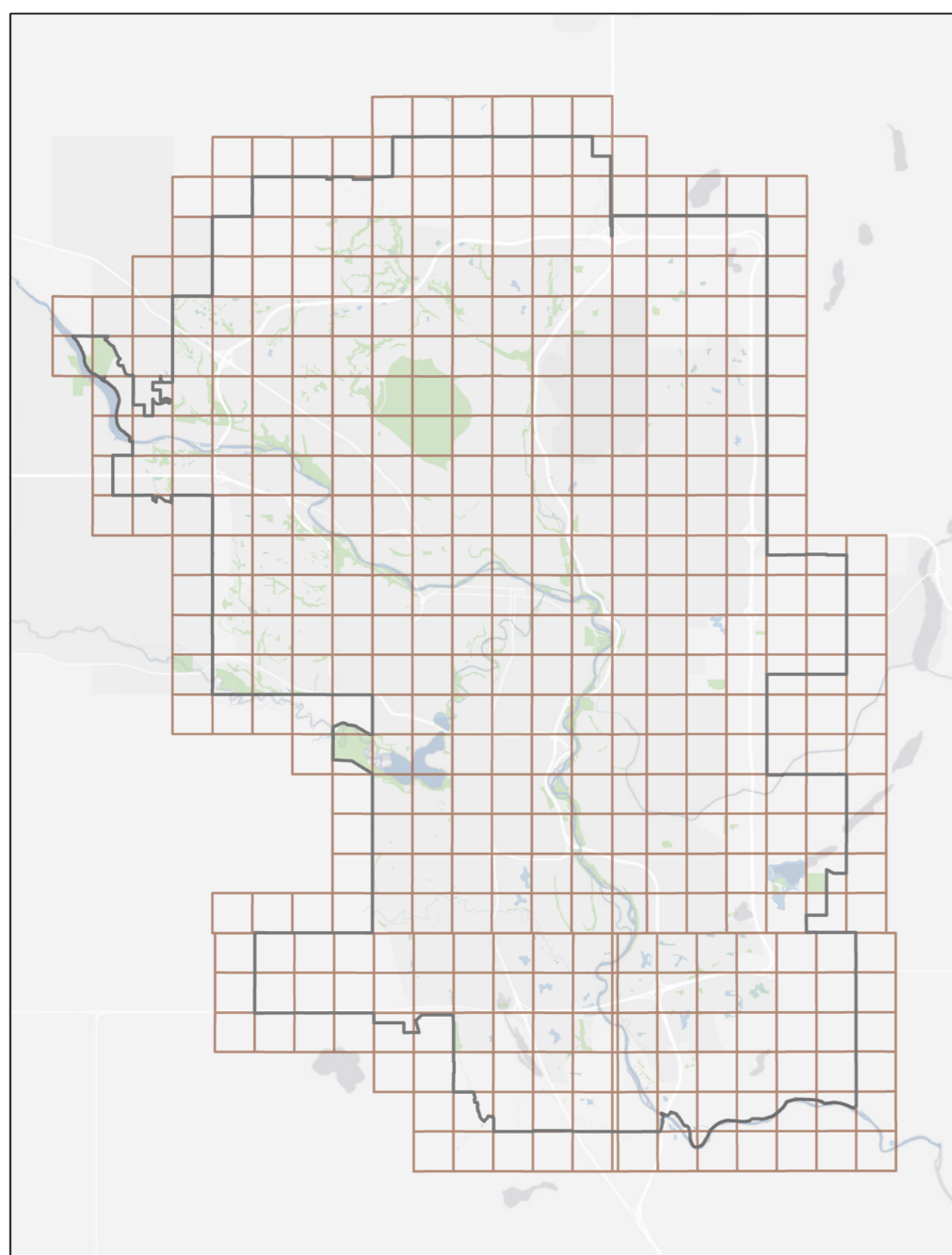
Using Python and ArcPy for ArcGIS

Renna Truong

Introduction

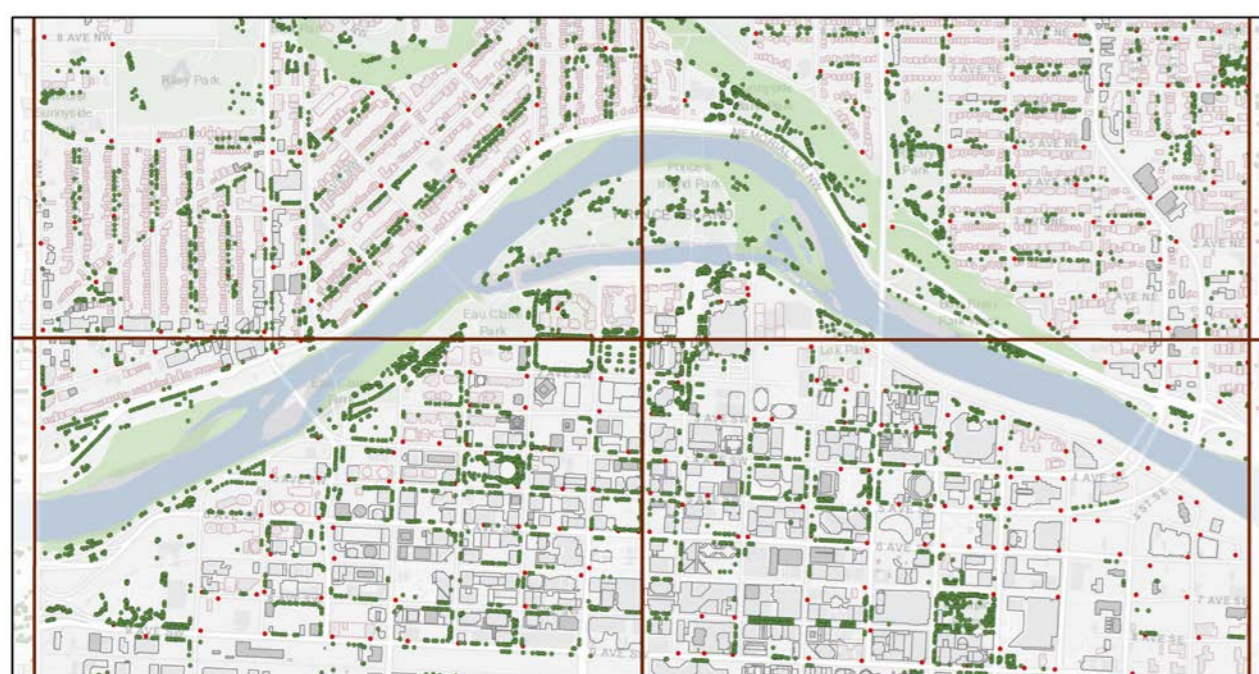
Annually, Spatial and Numeric Data Services (SANDS) receives 434 Digital Aerial Survey (DAS) datasets, in AutoCAD format, from the City of Calgary. Each dataset covers an Alberta Township System (ATS) section (Figure 1) and is made up of five layers containing “surface features and topography” information “derived from 1:5000 aerial photos”¹.

University of Calgary students regularly utilize these files in their research, however, in some cases they are looking for specific features covering an area much larger than an ATS section and in a more geographic information system (GIS) friendly format. To derive city wide products, from the DAS files, required reiterating through numerous geoprocessing operations. These repetitive and time consuming tasks were automated and accomplished within a day, using Python and ArcPy, instead of weeks if executed manually. The next section details the steps that were taken to create the outputs.



▲ Figure 1. Coverage of the 434 datasets by ATS sections³.

Figure 2. Flow chart illustrating the geoprocessing workflow of the developed Python script.

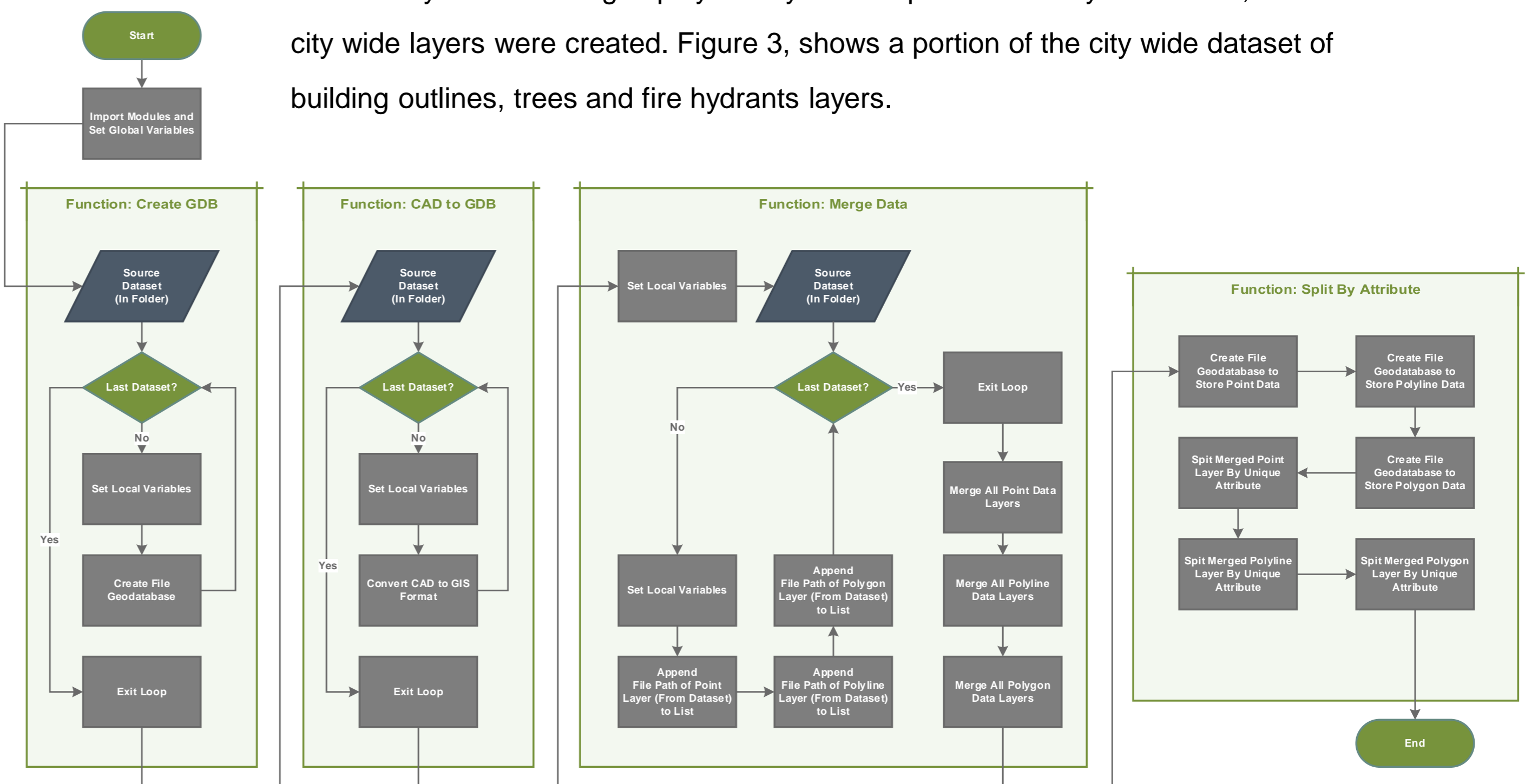


▲ Figure 3. A portion of the city wide dataset of commercial (grey) and residential (pink) building outlines, trees (dark green) and fire hydrants (red)⁴.

Methods

The data of interest resided in three out of the five layers per DAS dataset (Point, Polygon, and Polyline). As a result, the total number of layers to process came to 1,302 (434 datasets X 3 layers). The script developed was written in Python and utilized ArcPy, a Python site package that provides tools to work with spatial data². The geoprocessing tasks that were carried out are summarized below and Figure 2 illustrates the workflow of script, excluding information on parameters, inputs, and outputs.

1. Created 434 geodatabases (GDB) based on each DAS dataset name.
2. Converted 1,302 CAD layers into GIS format and set a spatial reference for each dataset. The converted layers are stored in the corresponding GDB.
3. Merged the converted layers, based on the geometry type (point, polygon, polyline), into one dataset. The three merged layers were then stored in a new GDB.
4. Split the merged datasets based on unique attributes from the Layer field.
 - The merged point layer was split into 72 layers. The merged polygon layer was split into 130 layers. The merged polyline layer was split into 162 layers. In total, 364 GIS city wide layers were created. Figure 3, shows a portion of the city wide dataset of building outlines, trees and fire hydrants layers.



Conclusion

Automating the multiple geoprocessing tasks, to create city wide products, increased efficiencies by executing tasks quicker and more accurately. This frees up time for the operator to focus on other responsibilities.

Since SANDS receives updated datasets from the City of Calgary annually, the script will be invoked time and time again.

References

1. City of Calgary. (2018). Digital Aerial Survey. Retrieved February 20, 2019, from <http://www.calgary.ca/CS/IIS/Pages/Mappingproducts/Digital-Aerial-Survey.aspx>
2. ESRI. (2018). What is ArcPY?. Retrieved March 2, 2019, from <http://desktop.arcgis.com/en/arcmap/latest/analyze/arcpy/what-is-arcpy-.htm>
3. Open Calgary (2018), ESRI, HERE, Garmin, OpenStreetMap contributors, and the GIS user community.
4. City of Calgary (2017), Open Calgary (2018), ESRI, HERE, Garmin, OpenStreetMap contributors, and the GIS user community.