Bringing GIS Analysis to Life Using Python Notebooks

Workshop Overview

This workshop will introduce you to using Python in Jupyter Notebook, an open source web application that enables you to create and share documents that contain rich text, equations, multimedia, alongside executable code and visualization of analysis outputs. The workshop will step you through the basics of setting up and being productive with Python notebooks. You will be introduced to ArcGIS Notebooks, which are Python Notebooks that are well-integrated within the ArcGIS platform. Finally, you will be guided through a series of ArcGIS Notebooks that illustrate how to create compelling data notebooks for data science that integrate your own Python scripts using the ArcGIS API for Python and ArcPy in combination with thousands of open source Python libraries to enhance your analysis and visualization.

Learning Outcomes

By completing this workshop, you will become comfortable with the following tasks:

- Setting up Jupyter Notebook server and different variations of ArcGIS Notebooks
- Basic interaction with a notebook, and techniques for using them effectively
- Composing effective notebooks for data science

Time Required

The following classroom time is required to complete this tutorial:

- 90 minutes

Materials Required

Technology:

- ArcGIS Pro 2.5+
- Notebook Server for ArcGIS Enterprise 10.7.1+ (optional)

Data:

- Sample data for this tutorial are included as part of the download.

Data Sources

- Toronto Police Service: Major Crime Indicators: http://data.torontopolice.on.ca/datasets/mci-2014-to-2019
- NOAA National Centers for Environmental Information: NCDC International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 3: https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.ncdc:C00834

Production Date

The Education and Research Group at Esri Canada makes every effort to present accurate and reliable information. The Web sites and URLs used in this tutorial are from sources that were current at the time of production, but are subject to change without notice to Esri Canada.

- Production Date: March 2020

Background Information

Python was first introduced to the ArcGIS community in 2004 with the release of version 9.0, as one of the scripting languages that provided access to the ArcGIS geoprocessing framework. Since then, Python has become the scripting language of choice for ArcGIS users. At the same time, the ecosystem of Python libraries designed for specific purposes has continued to grow (e.g., NumPy or SciPy for advanced mathematical and scientific processing, or more recently modules for integrating deep learning and machine learning tools such as Tensorflow). With the introduction of the notebook interface, initially with IPython, and more recently with the Jupyter Notebook server, Python and specifically Python Notebooks have become a favoured environment for enabling data science.

While using Python notebooks with ArcGIS has been possible for many years, it has become a much more integrated experience with the release of the ArcGIS API for Python in 2016, and more recently Notebook Server for ArcGIS Enterprise (2019), ArcGIS Notebooks in ArcGIS Pro (2020), and ArcGIS Notebooks in ArcGIS Online (coming soon).

By completing this workshop, you will have learned how to setup and interact with ArcGIS Notebooks, how to load, analyze, and visualize GIS datasets in a Python Notebook into a compelling document that communicates your GIS data processes or data science workflows.

References and Reading

- Introducing ArcGIS Notebooks
  https://www.esri.com/arcgis-blog/products/arcgis-enterprise/analytics/introducing-arcgis-notebooks/

Workshop Folder Contents

The following is a summary of the folders and files provided to you for this workshop – copy these files to your hard drive. With these files copied to your hard drive, and ArcGIS Pro licensed for offline use, you will be able to complete the entire tutorial without a connection to the Internet. All references to files later in this document will be relative to the top `arcgis_notebooks` workshop folder:

Folder Name / File Name	Description
arcgis_notebooks/	
documents/	A variety of reference documents for the workshop, (cheatsheets, presentation slides, this document)
hurricane_analysis/	Sample Python notebooks - hurricane track analysis
data/	Pre-packaged sample data and required files
part1_prepare_hurricane_data.ipynb	Part 1: Data wrangling
part2_explore_hurricane_data.ipynb	Part 2: Data exploration and analysis
part3_analyze_hurricane_data.ipynb	Part 3: Statistical correlation analysis
notebook_basics/	Sample Python notebooks
data/	Pre-packaged sample data and required files
starter_notebook.ipynb	Basic overview of working with Jupyter Notebook
notebook_basics.aprx	ArcGIS Pro project configured for working with the arcgispro_notebook.ipynb notebook
arcgispro_notebook.ipynb	Overview of notebooks integrated into ArcGIS Pro
hosted_notebook.ipynb	Overview of hosted notebooks in ArcGIS Online or ArcGIS Enterprise

Hands-on

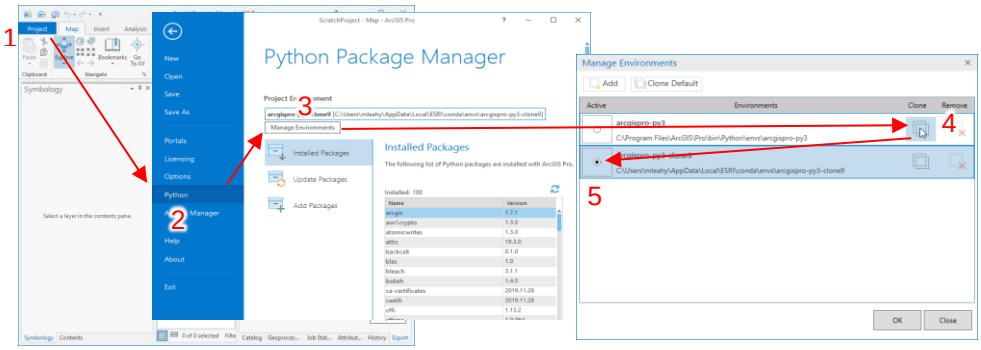
Part 1: Getting Started

Ensure that you have the following software installed:

- **ArcGIS Pro** 2.5+
  - *If necessary*: obtain a trial copy of ArcGIS Pro (https://www.esri.com/en-us/arcgis/trial)
  - *Optional*: authorize ArcGIS Pro for offline use (https://pro.arcgis.com/en/pro-app/get-started/start-arcgis-pro-with-a-named-user-license.htm#ESRI_SECTION1_15AD453E27C446CE9B51D45C021E8067)

With ArcGIS Pro 2.5 installed and licensed, prepare a clone of the default Python environment installed with ArcGIS Pro:

1. Open the ArcGIS Pro settings (or when working on a project in Pro, click the **Project** tab)
2. Select **Python** in the left-hand menu to access the Python Package Manger interface
3. Click **Manage Environments**
4. Click the clone button ( ) next to the default environment named **arcgispro-py3**
5. Wait for the clone operation to complete, ensure that your new cloned environment is set as active, and close ArcGIS Pro



6. Now that you have your ArcGIS Pro python environment cloned, you are free to modify it as you see fit. This workshop will make use of several packages that are not installed by default. To install these, open the **Python Command Prompt** installed with ArcGIS Pro (Windows Start Menu > ArcGIS > Python Command Prompt). When the command prompt opens, you should see the name of the Python environment you just cloned in ArcGIS Pro to the left of the cursor and current path.

   Ensure your PC is connected to the Internet, then execute the following two commands:

```
conda install dask graphviz python-graphviz seaborn
```

```
conda install -c conda-forge jupyter_contrib_nbextensions
jupyter_nbextensions_configurator
```

   The first of the two commands above will install a set of packages that are required for some of the sample notebooks included in this workshop. The second command will install some additional packages that enable you to configure and use extensions that enhance the Jupyter Notebook system.

This installation process will take some time to complete, depending on your Internet connection and PC's hardware capacity. If you see any errors/warnings, you can safely ignore them. Do not close the command prompt until the process is completed, and the cursor has returned. When the process is completed, you can open the Jupyter Notebook server.

7. Once you have completed the installation of the above packages, you will want to change the current drive and directory to the folder on your hard drive that contains the files for this workshop (please place this in a location with a simple path, that does not include spaces or special characters), and then launch the Jupyter Notebook server. You can do this with the following commands (adjust the drive letter and path in the first two commands as appropriate):

```
D:\
cd D:\arcgis_notebooks
```

```
jupyter notebook
```

The last command above will launch the Jupyter web interface, and you will see a view of the contents of the `arcgis_notebooks` directory.

8. Consider enabling the table of contents extension in Jupyter by navigating to the 'Nbextensions' tab in the Jupyter web interface, and clicking the checkbox next to the 'Table of Contents (2)' item (if this checkbox is greyed-out, you may need to uncheck the option near the top labelled 'disable configuration for nbextensions without explicit compatibility').

   With this extension enabled, you can easily navigate within your notebooks with a table of contents that displays and synchronizes with headers in Markdown cells of your notebook.

Part 2: Python Notebook Basics

For the first portion of the workshop, we will start by reviewing basic aspects of the Python Notebook environment, how to compose rich text and multimedia, and how to interact with code to execute Python commands and visualize outputs.

1. **Using the Jupyter Notebook web interface:** within the web interface launched at the end of the steps in Part 1 (above), navigate into the `notebook_basics` folder, and open the `starter_notebook.ipynb` file.
   - At this point, you may choose to open the `documents/Jupyter_Cheatsheet.pdf` document, and keep it open for reference as you work with your notebooks in the Jupyter environment.
   - For a tour of the notebook interface, you can select 'User Interface Tour' in the Help menu.
2. To start the notebook from scratch, you can choose "Restart & Clear Output" from the Kernel menu.
3. You can follow the notebook by running code inside each code block in one of the following ways:
   - Select a cell with python code in it and do one of the following:
     i. Click the 'Run' button in the main toolbar.
     ii. Press CTRL+Enter
     iii. Press SHIFT+Enter (this automatically selects the next cell; repeat the SHIFT+Enter keystroke to execute multiple cells in sequential order)
     iv. In the main Cell menu at the top, select one of the 'Run …' options
     v. Some of the cells in the notebook deliberately generate errors, so this notebook will not successfully run if you choose the 'Restart & Run All' option from the 'Kernel'

menu.

4. **Using ArcGIS Pro:** open the `notebook_basics.aprx` project file located in the `notebook_basics` folder in ArcGIS Pro. In the Catalog pane, expand the Folders item, then the `notebook_basics`, then right-click on the `arcgispro_notebook.ipynb` file and choose 'Add To Project'. Expand the 'Notebooks' item in the Catalog pane, then double-click on the notebook file to open it.
5. Once the notebook is opened in ArcGIS Pro, you can follow its workflow by executing the code cells as described in step 3 above, or optionally choose 'Restart & Run All' option from the 'Kernel' menu.

6. **Using ArcGIS Enterprise or ArcGIS Online:** open a web browser and login to your ArcGIS Enterprise that has a Notebook Server enabled with it, or login to your ArcGIS Online organization that has hosted notebooks enabled (currently in private beta).

   Note: if you do not have access to an ArcGIS Enterprise or ArcGIS Online portal with Notebooks enabled, you may open the `hosted_notebook.ipynb` file below using the standard Jupyter Notebook server, or using ArcGIS Pro, and follow the notebook as described in Step 3 above. In either of these environments, you will need to modify the python code that connects to an ArcGIS Enterprise or ArcGIS Online, and you must use an account with a role that has sufficient permissions and access to service credits in order to perform certain processes.

7. Navigate to the Content section, click 'Add Item' -> 'From my computer'. In the dialog that appears, choose the `hosted_notebook.ipynb` file from the `notebook_basics` folder, enter one or more tags, and click 'Add Item'.
8. When your notebook has been uploaded as a new item in your content, you will be taken to its details page. Click the 'Open Notebook' button near the top of the page to launch the hosted notebook in your web browser.
9. Follow this notebook by executing code cells as described in step 3 above, or optionally choose the 'Restart & Run All' option from the 'Kernel' menu.

Part 3: Data Science with ArcGIS and Python Notebooks

1. In the Jupyter Notebook web interface, navigate into the `hurricane_analysis` folder, and open the `part1_prepare_hurricane_data.ipynb` file.
2. Clear the notebook and follow the notebook as you did with the notebooks in the previous section.
3. Repeat steps 1 & 2 for `part2_explore_hurricane_tracks.ipynb` and `part3_analyze_hurricane_tracks.ipynb` notebooks.

Resources:

- ArcGIS API for Python (API reference, guides, sample notebooks)
  https://developers.arcgis.com/python/

- Hosted ArcGIS Notebooks Samples by Esri:
  https://www.arcgis.com/home/search.html?q=owner%3A%22ArcGISPyAPIBot%22

- ArcGIS Pro Python reference (ArcPy)
  https://pro.arcgis.com/en/pro-app/arcpy/main/arcgis-pro-arcpy-reference.htm

- Jupyter Notebook project (community, documentation, resources): https://jupyter.org/

- Coding Standards for Jupyter Notebook https://www.esri.com/about/newsroom/arcuser/coding-standards-for-jupyter-notebook/