



Extend Your Big Data Analysis Through Python in GeoAnalytics Tools

Michael Park & Noah Slocum

2021 ESRI
DEVELOPER SUMMIT

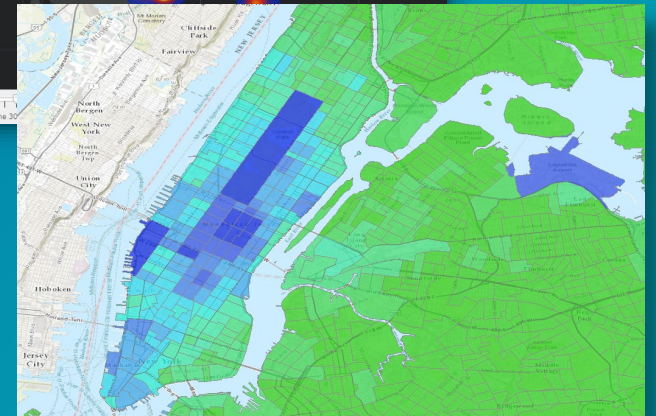
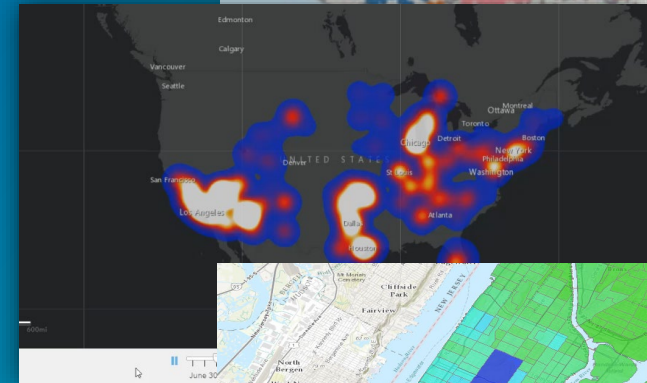
Agenda

- **Overview of GeoAnalytics Server**
- **Intro to Spark**
- **Demos:**
 - **Reading data**
 - **Writing data**
 - **Chaining together GeoAnalytics Tools**
 - **Connect to data outside of ArcGIS**

What is GeoAnalytics Server?

GeoAnalytics Server *distributes computing* to quickly analyze large amounts of *vector and tabular data* across multiple servers

A collection of analysis tools to identify *patterns, relationships, anomalies* and *incidents* in large amounts of data across space and time



Analysis Tools

Summarize Data

- Aggregate Points
- Build Multi-Variable Grid
- Describe Dataset
- Join Features
- Reconstruct Tracks
- Summarize Attributes
- Summarize Within
- Summarize Center and Dispersion

Manage Data

- Append Data
- Calculate Field
- Clip Layer
- Copy to Data Store
- Dissolve Boundaries
- Merge Layers
- Overlay Layers

Analyze Patterns

- Calculate Density
- Create Space Time Cube
- Find Hot Spots
- Find Point Clusters
- Forest-based Classification and Regression
- Generalized Linear Regression
- Geographically Weighted Regression

Data Enrichment

- Enrich from Multi-Variable Grid
- Calculate Motion Statistics

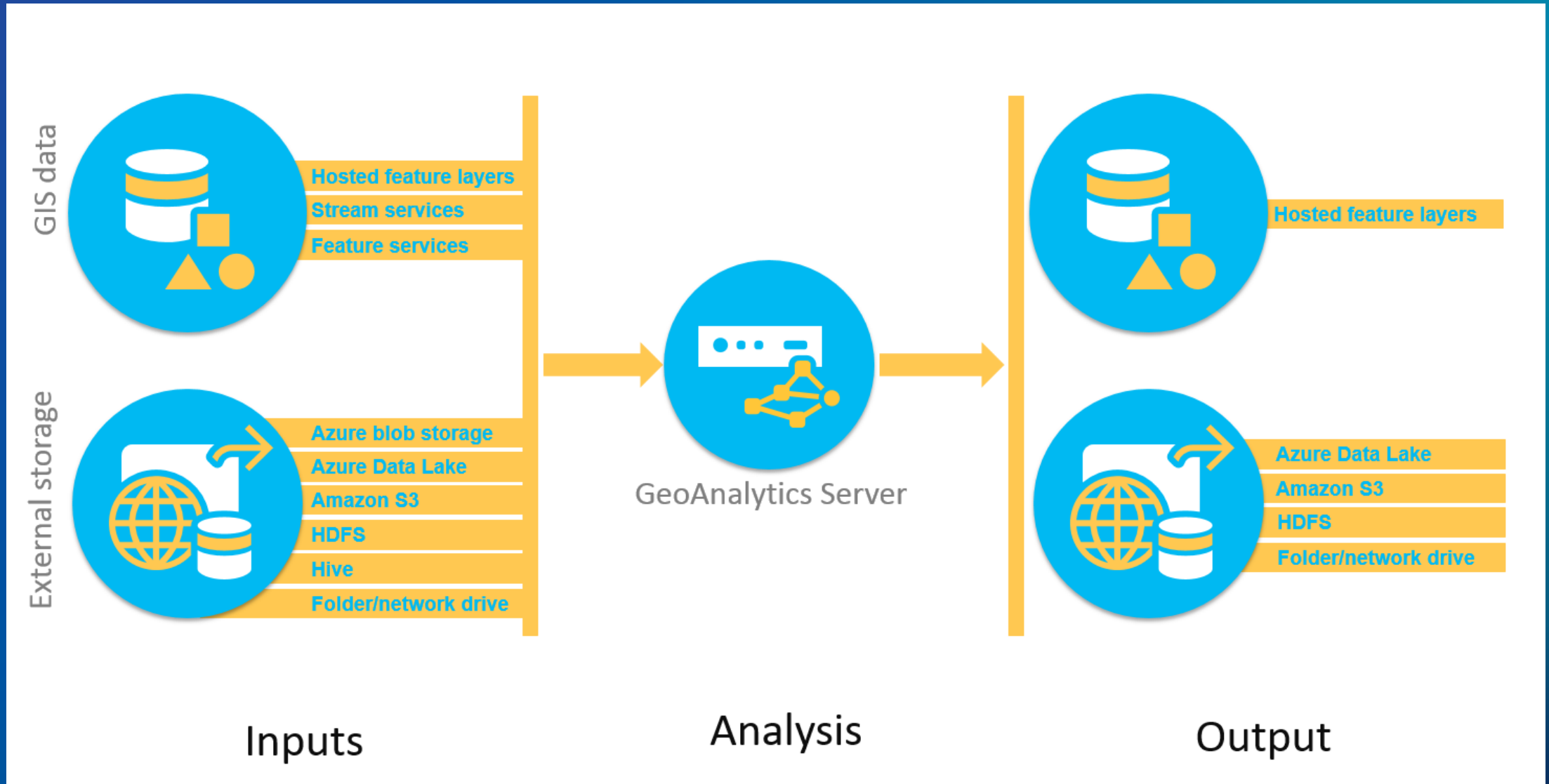
Find Locations

- Detect Incidents
- Find Similar Locations
- Geocode Locations
- Find Dwell Locations

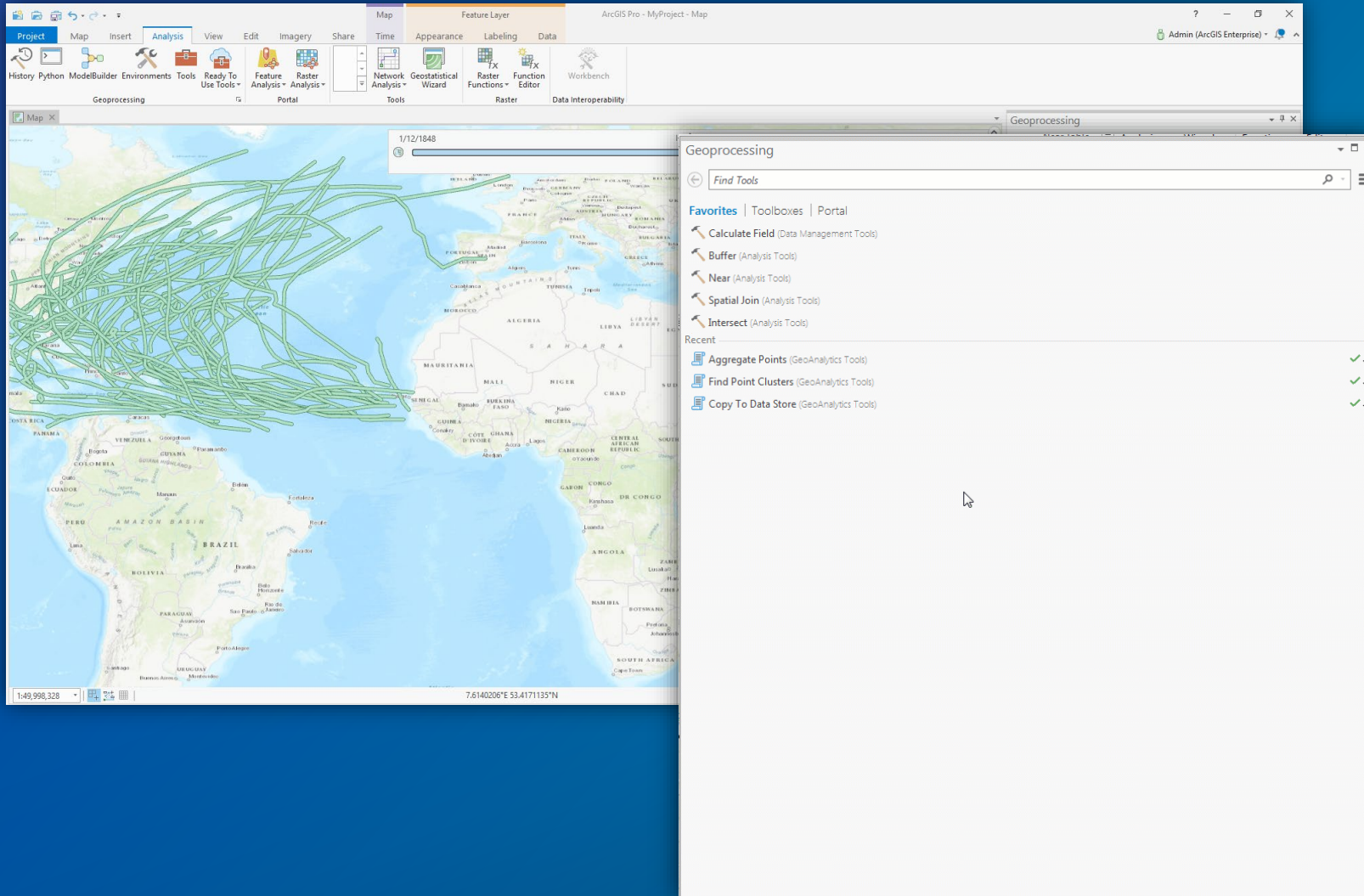
Use Proximity

- Create Buffers
- Trace Proximity Events

What types of data can I analyze? And from where?

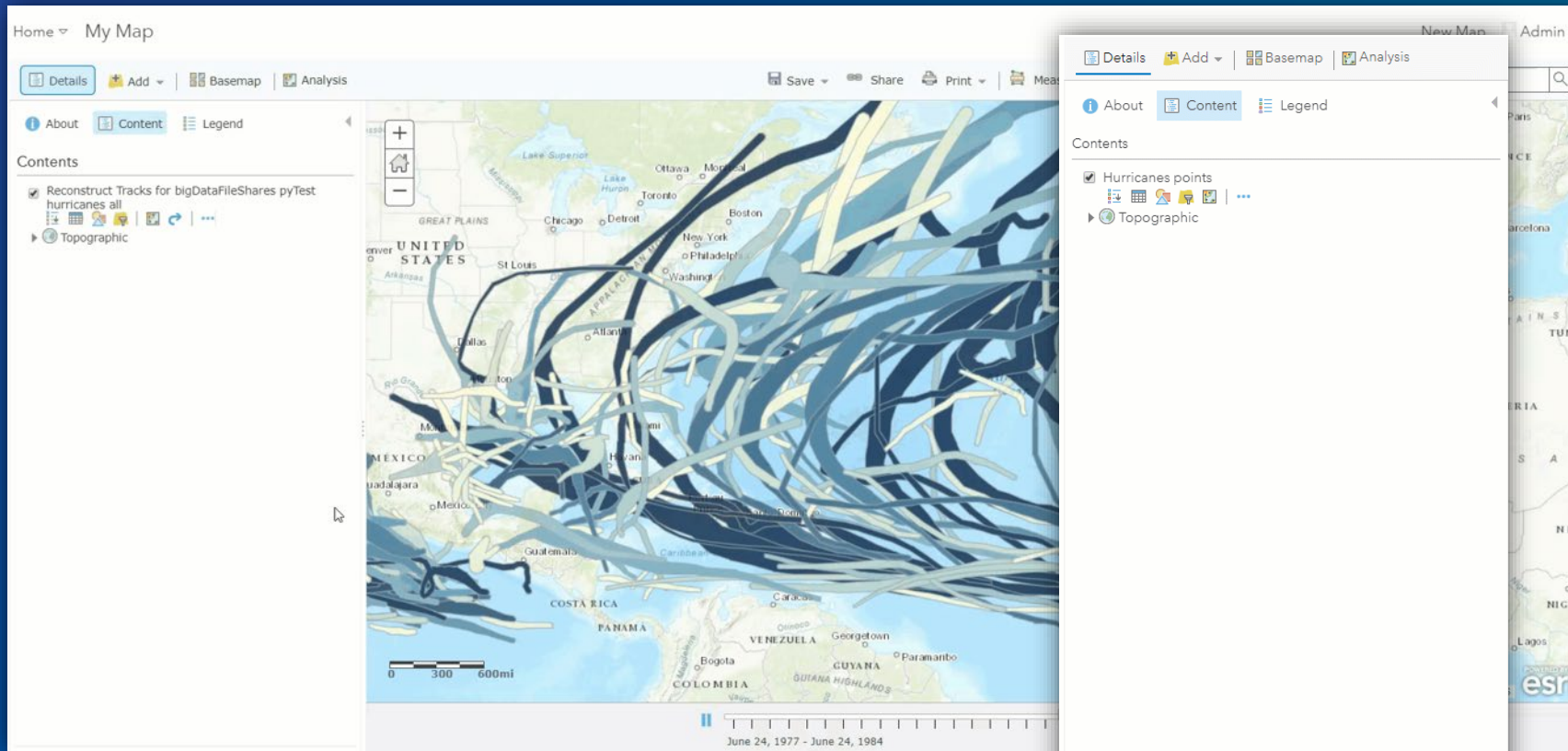


GeoAnalytics Server is ready to use with:



ArcGIS Pro

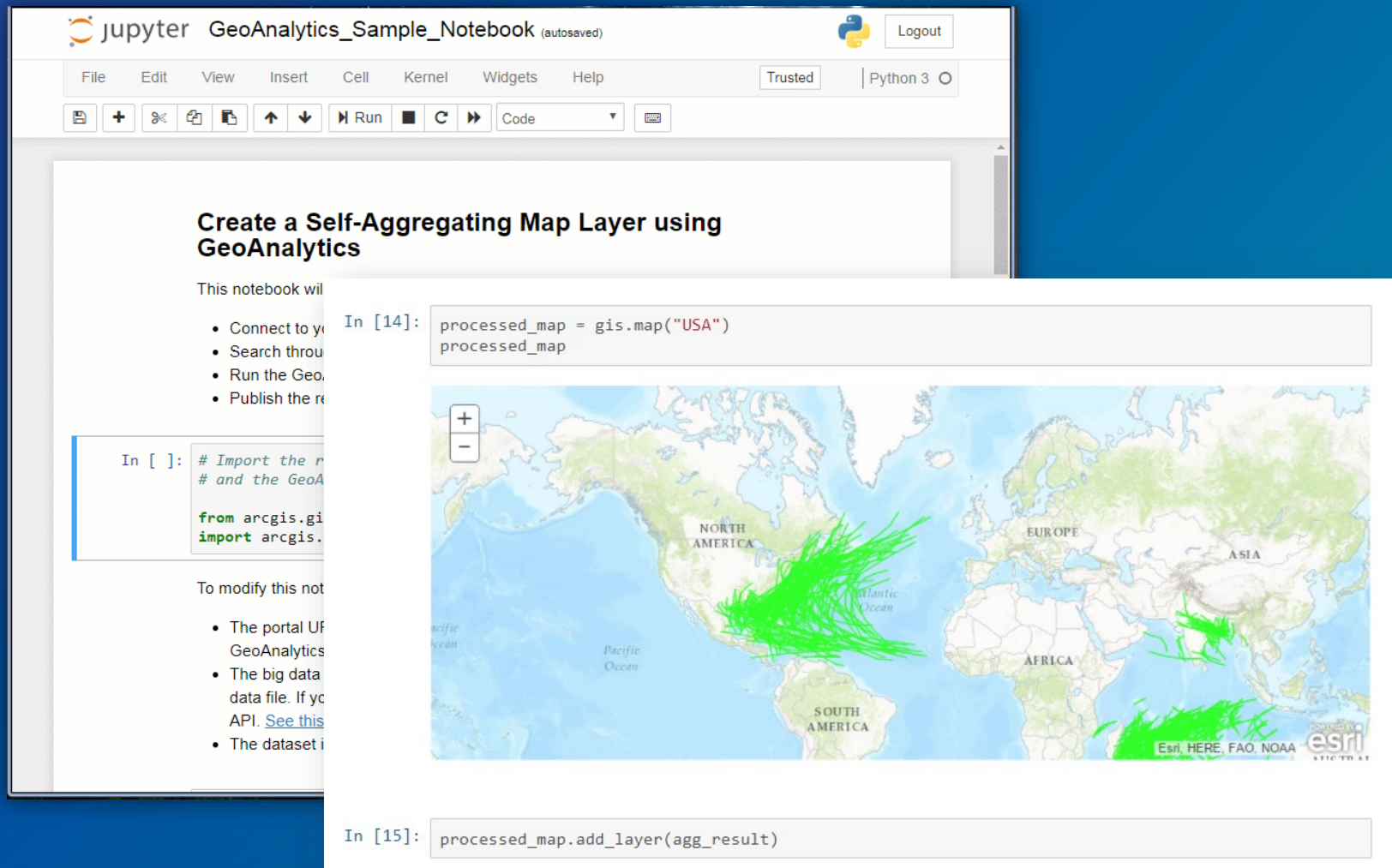
GeoAnalytics Server is ready to use with:



ArcGIS Pro

Enterprise portal

GeoAnalytics Server is ready to use with:



The screenshot shows a Jupyter Notebook interface with the following content:

Create a Self-Aggregating Map Layer using GeoAnalytics

This notebook will

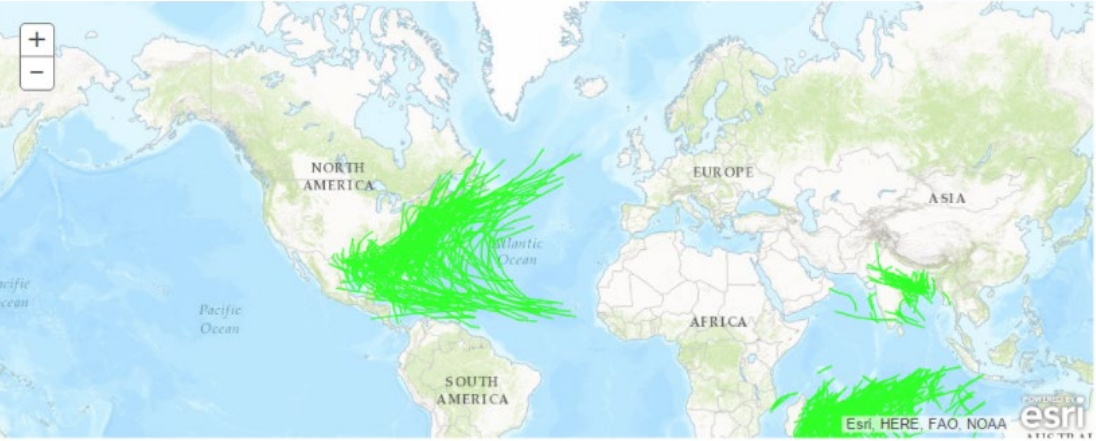
- Connect to your GeoAnalytics Server
- Search through your data
- Run the GeoAnalytics tool
- Publish the results

```
In [ ]: # Import the ArcGIS API for Python  
# and the GeoAnalytics module  
  
from arcgis.gis import GIS  
import arcgis
```

To modify this notebook, you can:

- The portal URL of your GeoAnalytics Server
- The big data file. If you have a local file, you can use the API. [See this](#)
- The dataset ID

```
In [14]: processed_map = gis.map("USA")  
processed_map
```



The map shows a world map with green data points overlaid on the continents of North America, Europe, and Asia. The map includes labels for 'NORTH AMERICA', 'EUROPE', 'ASIA', 'AFRICA', and 'SOUTH AMERICA'. The 'Pacific Ocean' and 'Atlantic Ocean' are also labeled. The map is powered by Esri, HERE, FAO, and NOAA.

```
In [15]: processed_map.add_layer(agg_result)
```

ArcGIS Pro

Enterprise portal

*ArcGIS API for
Python*

Run Python Script (tool)

Batch execution of python code with GeoAnalytics Server



Uses Python 3.6 shipped with ArcGIS Server



Spatial distributed analysis with `geoanalytics`



Non-spatial distributed analysis with `pyspark`



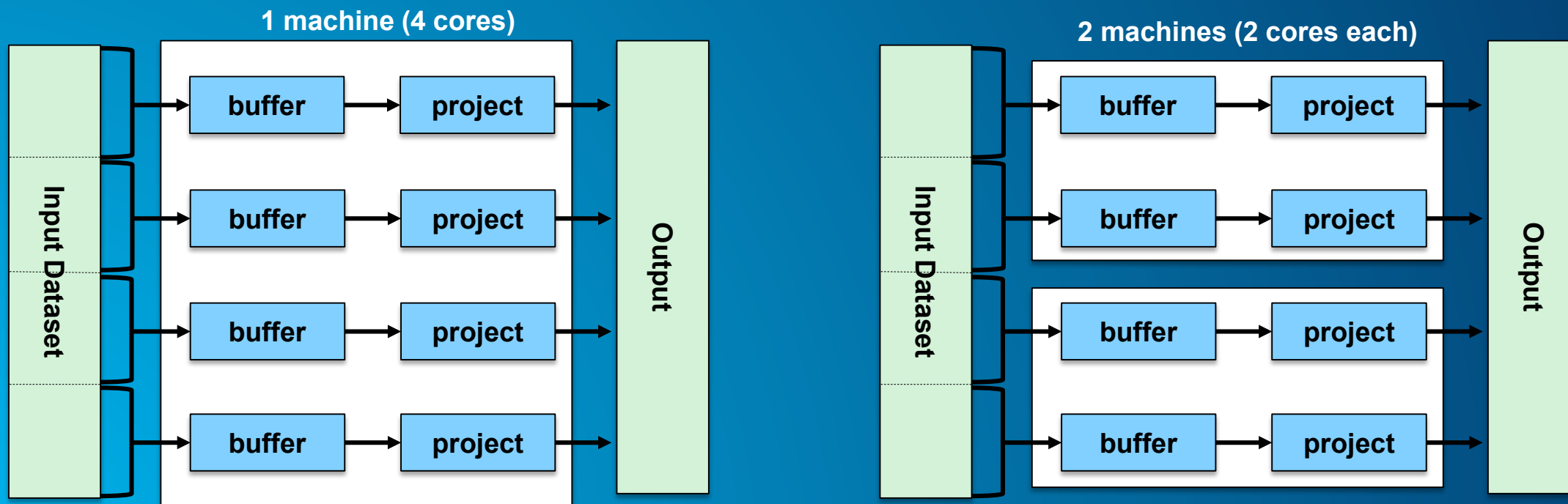
Integration of ArcGIS Enterprise layers and Spark DataFrames



Create custom analysis pipelines

Distributed Processing with Spark

```
df = load("earthquakes")
df_buffered = buffer(df, 10, "miles")
df_projected = project(df_buffered, 3857)
save(df, "output")
```



Links:

- <https://github.com/noahslocum/RunPythonScript-Samples>
- <https://www.esri.com/arcgis-blog/products/geoanalytics-server/analytics/extend-your-big-data-analysis-with-spark>



esri®

THE
SCIENCE
OF
WHERE®