Overview

- Introduction
- Demos
- Analysis Concepts using GeoAnalytics Server
- GeoAnalytics Data Sources
- GeoAnalytics Server Administration
What is GeoAnalytics Server?

ArcGIS has a new way of processing **large vector and tabular data** with both spatial (location) and temporal (time) components that is designed for **fast distributed analytics** and storage.
Answer questions like…

- Using millions of emergency calls accumulated over the last decade, which areas had the highest rates of emergency calls?

- What are the most popular locations for taxi pickups in New York City, and how is this trend changing weekly?

- What is the flight path of recorded GPS tracks, and how many of those paths occurred within 100 km of a no fly zone in 2015?
GeoAnalytics Server | Extends ArcGIS Analysis Capabilities

- Geoprocessing
- Powerful analytics

- GeoAnalytics
- Distributed analytics and distributed storage

- Web GIS Layers
- Rich geoinformation model
GeoAnalytics is out of the box and ready to use within ArcGIS Enterprise
GeoAnalytics Server | Why?

Run analysis against data that is too big for a single desktop
GeoAnalytics Server | Familiar User Interfaces

ArcGIS Pro
Portal Map Viewer
REST API
ArcGIS API for Python

System/GeoAnalyticsTools (GPServer)

Service Description: The GeoAnalyticsTools service is provided for distributed analysis of large datasets.

Tasks:
- AggregatePoints
- DescribeDataset
- JoinFeatures
- CreateBuffers
- CalculateDensity
- ReconstructTracks
- CreateSpaceTimeCube
- CopyToDataStore
- SummarizeAttributes
- SummarizeWithin
- FindSimilarLocations
- FindHotSpots

Execution Type: esriExecutionTypeAsynchronous
REST API
ArcGIS API for Python

Analyzing New York City taxi data using big data tools

At 10.5, ArcGIS Enterprise introduces ArcGIS GeoAnalytics Server which provides you the ability to perform big data analysis on your infrastructure. This sample demonstrates the steps involved in performing an aggregation analysis on New York city taxi point data using ArcGIS API for Python.

The data used in this sample can be downloaded from the NYC Taxi & Limousine Commission website. For this sample, data for the months January & February of 2015 were used, each averaging 12 million records.

Note: The ability to perform big data analysis is only available on ArcGIS Enterprise 10.5 licensed with a GeoAnalytics server and not yet available on ArcGIS Online.

The NYC taxi data
To give you an overview, let us take a look at a subset with 2000 points published as a feature service.

In [1]:
import aragis
from aragis.gis import GIS

ago_gis = GIS() # Connect to ArcGIS Online as an anonymous user
search_subset = ago_gis.content.search("NYC_Taxi_subset", item_type = "Feature Layer")
subset_item = search_subset[0]
subset_item

Out[1]:
{
    "id": "NYC_taxi_subset",
    "title": "A subset of NYC taxi data"
}
Gain insight into both spatial and temporal patterns
GeoAnalytics Server

Run analysis against massive collections in a scalable manner

Data: Flight Aware
Aggregate Points

Pro UI
Create Space
Time Cube
Pro UI
Reconstruct Tracks
Portal UI
GeoAnalytics Architecture | Inputs and Results

Use your layers through Pro, Portal, ArcGIS API for Python, or the REST API

Feature Services

Big Data File Share

ArcGIS Relational Data Store

ArcGIS Spatiotemporal Big Data Store

GeoAnalytics Server

Portal for ArcGIS

New Web GIS Layers

Web GIS Layers

ArcGIS Data Store
Analysis Overview

- Analysis capabilities patterned from ArcGIS Online Spatial Analysis
- Tools useful for big data with an emphasis on:
  - Spatiotemporal analysis
  - Summarization and aggregation
GeoAnalytics Server | Rich Collection of Analysis Tools

Summarize Data
Aggregate Points
Join Features
Reconstruct Tracks
Summarize Attributes
Summarize Within

Find Locations
Find Similar Locations

Analyze Patterns
Calculate Density
Create Space Time Cube
Find Hot Spots

Use Proximity
Create Buffers

Manage Data
Copy to Data Store
Aggregation

- Aggregate into polygons

Target Features + Join Features → Intermediate Result → Final Result

- Aggregate into time steps
Points into bins (square or hexagon)
Aggregation of Lightning Strikes
Aggregation of Lightning Strikes | 6 Hour Intervals
Time Stepping

• Three parameters to define a time step:
  - **Interval** (duration of time in a step)
  - **Repeat** (frequency of a step)
  - **Reference time** (alignment)

• Examples:
  - Hourly steps
    - **Interval**: 1 hour
  - Every 12th hour
    - **Interval**: 1 hour
    - **Repeat**: 12 hours
  - Every Monday
    - **Interval**: 1 day
    - **Repeat**: 1 week
    - **Reference**: Some Monday
Joining Features in Space and Time

- **Spatiotemporal joins**

<table>
<thead>
<tr>
<th>Target Features</th>
<th>Join Features</th>
<th>Intermediate Result</th>
<th>Final Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Target Features" /></td>
<td><img src="image2" alt="Join Features" /></td>
<td><img src="image3" alt="Intermediate Result" /></td>
<td><img src="image4" alt="Final Result" /></td>
</tr>
<tr>
<td><img src="image5" alt="Next Frame" /></td>
<td><img src="image6" alt="Next Frame" /></td>
<td><img src="image7" alt="Next Frame" /></td>
<td><img src="image8" alt="Next Frame" /></td>
</tr>
</tbody>
</table>
Aggregate Points into Tracks

- Reconstruct Tracks
  - Summarize time-enabled points into tracks
GeoAnalytics Server and your Data

- Use your GIS data
  - Works with layers already in your Web GIS

- Use your own big data sources: big data file shares
  - Easily attach folders located locally on a disk or file share
  - Data stored in HDFS or Hive

- Output:
  - Hosted feature service
  - By default, results are stored in the spatiotemporal data store
  - Local NetCDF file (Create Space Time Cube)
Registering a Big Data File Share

In ArcGIS Server Manager, register a big data file share on your ArcGIS Server.

**Name:**

**Type:**
- File Share
- HDF5
- Hive
- Cloud Store

Click **Create** to register the file share.
GeoAnalytics Server Setup and Administration

- GeoAnalytics Server Setup
- Deployment Patterns
Components Required for GeoAnalytics Server

- Portal for ArcGIS
- ArcGIS for Server
- ArcGIS Web Adaptor
- ArcGIS Data Store (relational and spatiotemporal)
Step 1: Set up a Base Enterprise Deployment

1. Install Portal for ArcGIS
2. Install ArcGIS Server (license as standard or advanced GIS server)
3. Install ArcGIS Data Store and register relational data store with Server
4. Install and configure Web Adaptors (WA) for Server and Portal
5. Federate Server with Portal and configure as the Hosting Server
Step 2: Install and Configure Spatiotemporal Data Store

1. Install ArcGIS Data Store
2. Configure a spatiotemporal data store with the hosting server
Step 3: Set up a GeoAnalytics Server

1. Install ArcGIS Server (licensed as a GeoAnalytics Server)
2. Install and configure a Web Adaptor for the Server
3. Federate the server with the Portal and configure as the GeoAnalytics server
Deployment patterns | Simple deployment

Base Deployment (Machine 1)
- Portal for ArcGIS
- Hosting Server
- ArcGIS Relational Data Store

ArcGIS Spatiotemporal Data Store (Machine 2)

GeoAnalytics Server (Machine 3)
- Big Data File Share
Deployment patterns | Multi-machine deployment

- **Base Deployment (Machine 1)**
  - WA (Server)
  - WA (Portal)
  - Portal for ArcGIS
  - Hosting Server
  - ArcGIS Relational Data Store

- **ArcGIS Spatiotemporal Data Store (multi-machine)**

- **ArcGIS GeoAnalytics Server (multi-machine)**
  - WA (GeoAnalytics)
  - GeoAnalytics Server
  - Big Data File Share
What can I run GeoAnalytics on?

- Supported on Windows and Linux platforms
  - On-premises
  - Amazon (easy to configure with Amazon Cloud formation templates)
  - Azure (easy to configure with Azure Cloud Builder)
Minimum System Requirements

• **Base ArcGIS Enterprise deployment**
  - Minimum 4 cores and 16 GB RAM

• **ArcGIS Spatiotemporal Data Store**
  - Minimum 16 GB RAM for each machine
  - Fast disk/sufficient disk space for data

• **ArcGIS GeoAnalytics Server**
  - Minimum 4 cores and 16 GB RAM for each machine
  - Sufficient temp space for compute
Recommended System Requirements

- **Base ArcGIS Enterprise deployment**
  - Minimum 4 cores and 32 GB RAM
  - Or install the base deployment across multiple machines

- **ArcGIS Spatiotemporal Data Store**
  - 32 GB RAM for each machine

- **ArcGIS GeoAnalytics Server**
  - 32 GB RAM for each machine

- Recommend an equal number of GeoAnalytics to spatiotemporal data store machines
Summary

• ArcGIS GeoAnalytics “out of the box” and “ready to use” within ArcGIS Enterprise

• Distributed Analytics - helps you get “big jobs” done faster

• Makes big data easier to use and easier to share

• Focuses on space and time analysis for massive dataset processing

• Lets you see the big data picture in a clearer and more meaningful way