

GeoAnalytics in ArcGIS Pro & ArcGIS GeoAnalytics Server

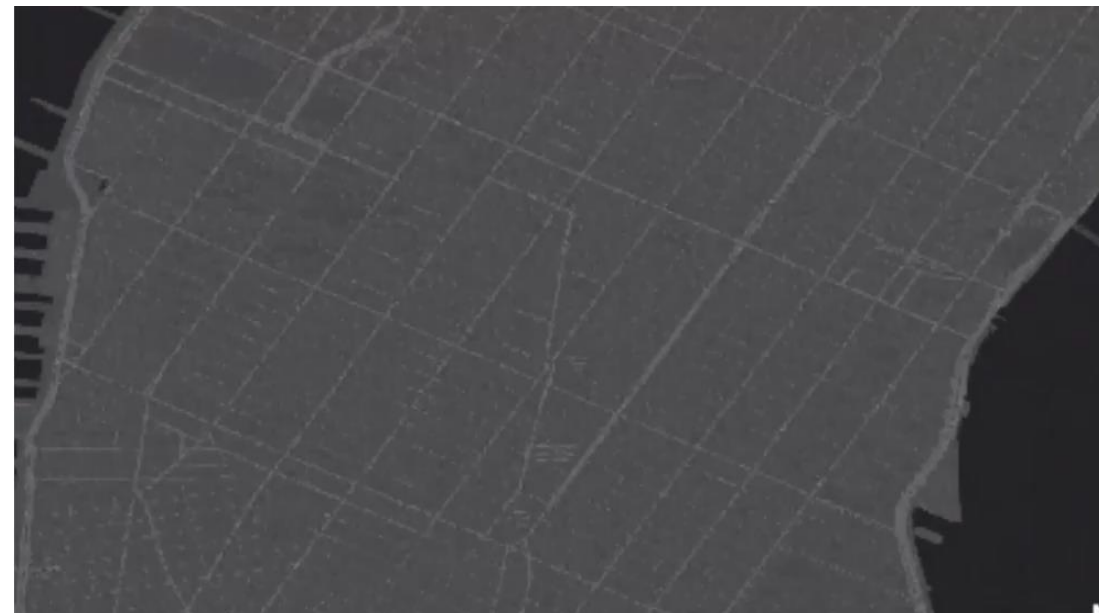
2022年7月
ESRIジャパン株式会社





こんなデータで困ったことはありませんか？

VendorID	lpep_pickup_datetime	lpep_dropoff_datetime	store_and_fwd_flag	RatecodeID	PULocationID	DOLocationID	passenger_count	trip_distance	fare_amount	extra	mta_tax	tip_amount	tolls_amount	ehail_fee	improvement_surcharge	total_amount	payment_type	trip_type
2	1/1/2018 0:18	1/1/2018 0:24	N	1	236	236	5	0.7	6	0.5	0.5	0	0		0.3	7.3	2	1
2	1/1/2018 0:30	1/1/2018 0:46	N	1	43	42	5	3.5	14.5	0.5	0.5	0	0		0.3	15.8	2	1
2	1/1/2018 0:07	1/1/2018 0:19	N	1	74	152	1	2.14	10	0.5	0.5	0	0		0.3	11.3	2	1
2	1/1/2018 0:32	1/1/2018 0:33	N	1	255	255	1	0.03	-9	-0.5	-0.5	0	0		-0.3	-4.3	3	1
2	1/1/2018 0:32	1/1/2018 0:33	N	1	255	255	1	0.03	3	0.5	0.5	0	0		0.3	4.3	2	1
2	1/1/2018 0:38	1/1/2018 1:08	N	1	255	161	1	5.63	21	0.5	0.5	0	0		0.3	23.3	2	1
2	1/1/2018 0:18	1/1/2018 0:28	N	1	189	65	5	1.71	8.5	0.5	0.5	0	0		0.3	9.8	2	1
2	1/1/2018 0:38	1/1/2018 0:55	N	1	189	225	5	3.45	14.5	0.5	0.5	3.16	0		0.3	18.96	1	1
2	1/1/2018 0:05	1/1/2018 0:18	N	1	129	82	1	1.61	10	0.5	0.5	0	0		0.3	11.3	2	1
2	1/1/2018 0:35	1/1/2018 0:42	N	1	226	7	1	1.87	7.5	0.5	0.5	0	0		0.3	8.8	2	1
2	1/1/2018 0:21	1/1/2018 0:39	N	1	145	129	2	4.12	16.5	0.5	0.5	0	0		0.3	17.8	2	1
2	1/1/2018 0:56	1/1/2018 1:04	N	1	7	223	2	1.22	7	0.5	0.5	0	0		0.3	8.3	2	1
2	1/1/2018 0:11	1/1/2018 0:30	N	1	255	189	1	4.67	17	0.5	0.5	0	0		0.3	18.3	2	1
2	1/1/2018 0:57	1/1/2018 1:12	N	1	97	188	1	2.71	11.5	0.5	0.5	3.84	0		0.3	16.64	1	1
2	1/1/2018 0:36	1/1/2018 0:51	N	1	244	75	2	6.01	19	0.5	0.5	4	0		0.3	24.3	1	1
1	1/1/2018 0:07	1/1/2018 0:15	N	1	225	37	1	1.9	8	0.5	0.5	3	0		0.3	12.3	1	1
1	1/1/2018 0:25	1/1/2018 0:42	N	1	36	145	2	4.3	15.5	0.5	0.5	3.35	0		0.3	20.15	1	1
1	1/1/2018 0:42	1/1/2018 1:00	N	1	145	173	1	6.9	22	0.5	0.5	0	0		0.3	23.3	1	1
2	1/1/2018 0:06	1/1/2018 0:08	N	1	49	49	1	0.3	3.5	0.5	0.5	0	0		0.3	4.8	2	1
2	1/1/2018 0:34	1/1/2018 0:52	N	1	40	113	1	4.47	16.5	0.5	0.5	3.56	0		0.3	23.31	1	1
1	1/1/2018 0:25	1/1/2018 0:28	N	1	179	7	1	0.5	4.5	0.5	0.5	0	0		0.3	5.8	1	1
2	1/1/2018 0:36	1/1/2018 0:51	N	1	7	193	1	1.82	9	0.5	0.5	0	0		0.3	10.3	1	1
2	1/1/2018 0:53	1/1/2018 1:26	N	1	97	74	1	11.79	36	0.5	0.5	7.46	0		0.3	46.71	1	1
1	1/1/2018 0:11	1/1/2018 0:22	N	1	255	112	1	1.9	9	0.5	0.5	3.05	0		0.3	13.35	1	1
1	1/1/2018 0:40	1/1/2018 1:01	N	1	255	28	1	10.3	29	0.5	0.5	5	0		0.3	35.3	1	1
2	1/1/2018 0:15	1/1/2018 0:25	N	1	80	80	1	1.66	8.5	0.5	0.5	1.96	0		0.3	11.76	1	1
2	1/1/2018 0:35	1/1/2018 0:48	N	1	255	232	1	2.91	12	0.5	0.5	3.32	0		0.3	16.62	1	1
2	1/1/2018 0:55	1/1/2018 1:28	N	1	256	50	1	6.09	25	0.5	0.5	6.58	0		0.3	32.88	1	1
2	1/1/2018 0:41	1/1/2018 0:56	N	1	179	75	5	5.3	17	0.5	0.5	3.66	0		0.3	21.96	1	1
2	1/1/2018 0:36	1/1/2018 0:44	N	1	41	75	1	1.63	8	0.5	0.5	1.86	0		0.3	11.16	1	1
2	1/1/2018 0:48	1/1/2018 0:51	N	1	75	74	1	0.91	4.5	0.5	0.5	0	0		0.3	5.8	2	1
2	1/1/2018 0:56	1/1/2018 1:00	N	1	74	74	2	0.92	5	0.5	0.5	0	0		0.3	6.3	1	1
2	1/1/2018 0:27	1/1/2018 0:34	N	1	7	223	1	0.98	6	0.5	0.5	0	0		0.3	7.3	2	1
2	1/1/2018 0:41	1/1/2018 0:52	N	1	179	7	1	1.42	9	0.5	0.5	0	0		0.3	10.3	2	1
2	1/1/2018 0:48	1/1/2018 0:53	N	1	254	254	1	0.42	5	0.5	0.5	0	0		0.3	6.3	1	1
2	1/1/2018 0:46	1/1/2018 1:02	N	1	255	36	1	3.15	13	0.5	0.5	2.86	0		0.3	17.16	1	1
2	1/1/2018 1:03	1/1/2018 1:09	N	1	260	226	1	0.85	5.5	0.5	0.5	0	0		0.3	6.8	2	1
2	1/1/2018 0:30	1/1/2018 0:38	N	1	181	190	5	1.28	7.5	0.5	0.5	1.76	0		0.3	10.56	1	1
2	1/1/2018 0:39	1/1/2018 0:48	N	1	190	89	5	1.58	8	0.5	0.5	1.86	0		0.3	11.16	1	1
2	1/1/2018 0:59	1/1/2018 1:25	N	1	188	76	5	4.43	18.5	0.5	0.5	0	0		0.3	19.8	1	1



表形式の大量データ

位置情報を持った大量データ

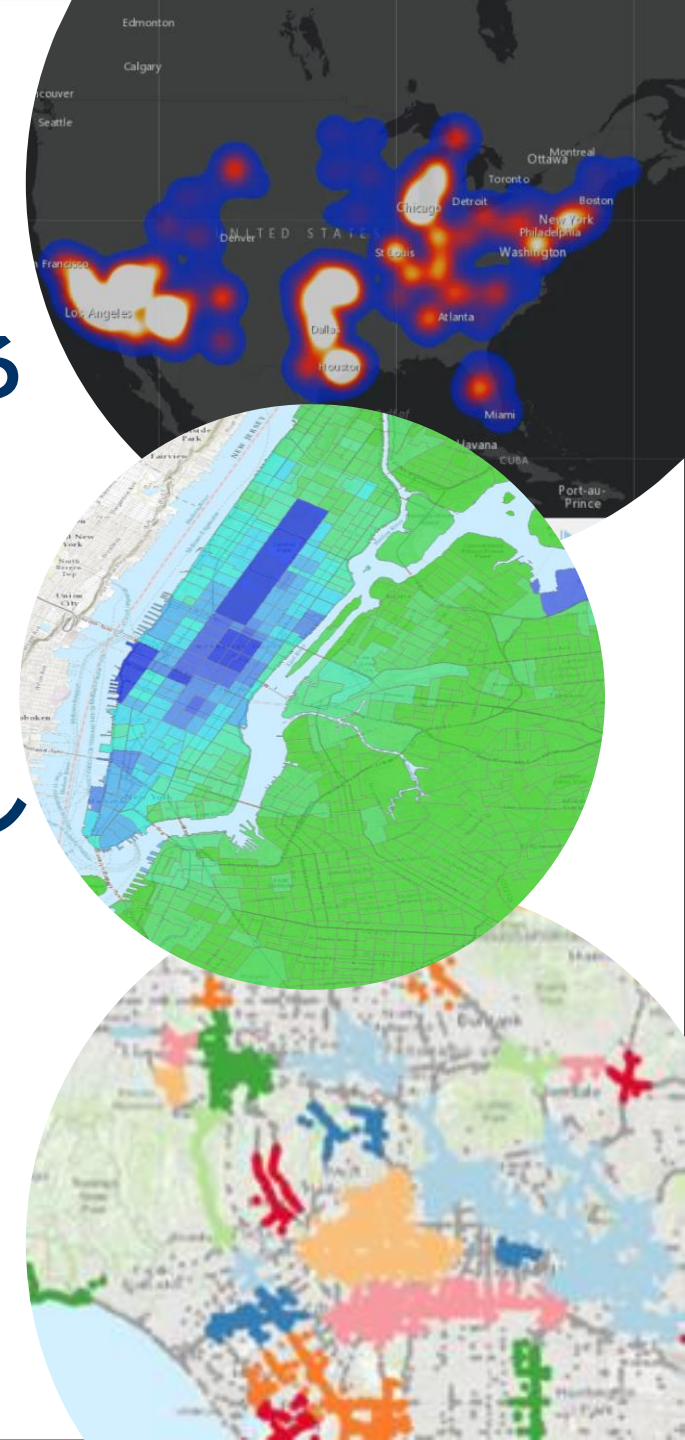
大量データをどのように意味がわかるようにするのか？

GeoAnalytics とは?

大規模なベクターと表形式 データを高速に解析する
並列計算化※されたツールボックス

空間と時間をまたがって大量のデータのパターン、
関係性、異常 やインシデント を同定する分析ツールの
コレクション

※GeoAnalytics Server の場合はコンピュータを並列化

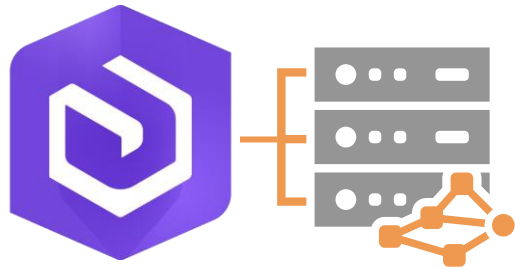




GeoAnalytics を使う理由は?

- 既存のツールやワークフローでは、十分な速度でデータ処理できない
- 他の分析で利用するために、データをより管理しやすいものになりたい
- データにノイズが多く、また重要なものを抽出したい
- GISに読み込むのに苦労しているデータセットがある
- いくつかのツールがとても素晴らしい

GeoAnalytics は以下で利用可能



GeoAnalytics Server

ArcGIS Enterprise による複数のサーバー
コアとマシンでの **分散処理**

ライセンス: Enterprise
+ GeoAnalytics Server License



GeoAnalytics Desktop

ArcGIS Pro によるラップトップもしくはデ
スクトップで複数コアを使う **並列処理**

ライセンス: Advanced License

Desktop と Server の使い分け

- **GeoAnalytics Server** は以下のような場合に利用
 - ビッグ データ解析を組織全体に導入
 - 1台もしくは複数台のサーバーマシンのパワーを活用する
 - 外部のビッグ データ ストレージや既存Web レイヤーに接続する
 - カスタム解析による拡張
- **GeoAnalytics Desktop** は以下のような場合に利用
 - ローカルデータ（ファイル、データベース）を自分のデスクトップマシンで従来より高速に処理する
 - GeoAnalytics Server で利用したいワークフローのプロトタイピング
 - ビッグ データ コネクションへの接続と利用

Desktop と Server の使い分け

	GeoAnalytics Server (Enterprise 10.9.1)	GeoAnalytics Desktop (Pro 2.9)
入力データ	<ul style="list-style-type: none">- ビッグ データ ファイル共有- ホスト フィーチャレイヤー- フィーチャ サービス	<ul style="list-style-type: none">- File + Enterprise ジオデータベース- シェープファイル- ビッグ データ コネクション (BDC)
出力データ	<ul style="list-style-type: none">- ホスト フィーチャレイヤー- ビッグ データ ファイル共有	<ul style="list-style-type: none">- File + Enterprise ジオデータベース- シェープファイル
解析のスケールアウト	<ul style="list-style-type: none">- マシンの台数をコントロール- コアとRAMの比率をコントロール- ビッグ データ ストアでのデータストレージのスケールアウト	<ul style="list-style-type: none">- 1台のマシンのみ- RAMの比率をコントロール (もしくは値)
ツール	<ul style="list-style-type: none">- 30 ツール- Python スクリプトの実行	<ul style="list-style-type: none">- 23 ツール- 7 ツール (BDCの管理)
インタフェース	<ul style="list-style-type: none">- REST + ArcGIS API for Python- Pro と ArcPy (とモデルビルダー)- ポータルの Map Viewer	<ul style="list-style-type: none">- Pro と ArcPy (とモデルビルダー)

[Blog post covering this topic](#)



GeoAnalytics Desktop



解析機能の概要

Desktop で利用可能なツール

パターンの分析

- 密度の計算 | Calculate Density
- ホットスポット分析 | Find Hot Spots
- ポイントクラスターの検索 | Find Point Clusters
- フォレストベースの分類
と回帰分析 | Forest-based Classification and Regression
- 一般化線形回帰分析 | Generalized Linear Regression

近接エリアの分析

- バッファの作成 | Create Buffers
- 近接イベントのトレース | Trace Proximity Events
- 近接性でグループ化 | Group By Proximity

データの管理

- フィールド演算 | Calculate Field
- レイヤーのクリップ | Clip Layer
- 境界のディゾルブ | Dissolve Boundaries
- オーバーレイ | Overlay Layers

GeoAnalytics Desktop ツールボックスの概要

データの集約

- ポイントの集約 | Aggregate Points
- データセットの記述 | Describe Dataset
- フィーチャの結合 | Join Features
- トラッキングの再構築 | Reconstruct Tracks
- 属性の集計 | Summarize Attributes
- 中心と分散の集計 | Summarize Center and Dispersion
- エリア内での集計 | Summarize Within

場所検索

- インシデントの検出 | Detect Incidents
- 滞在場所の検索 | Find Dwell Locations
- 類似フィーチャの検索 | Find Similar Locations

データへの情報付加

- モーションの統計量を計算 | Calculate Motion Statistics

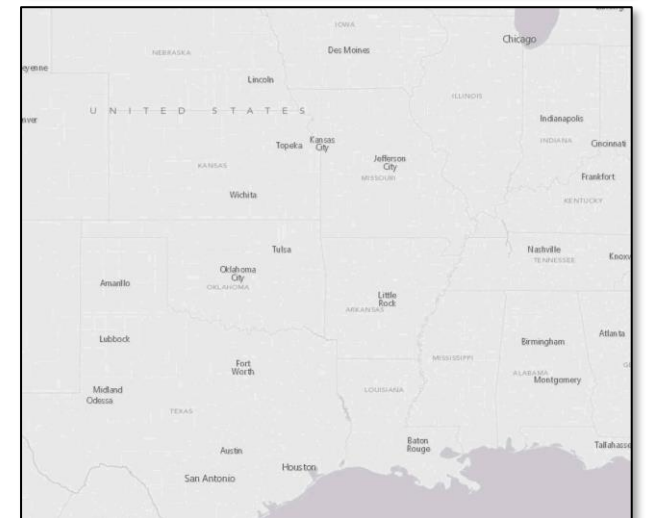
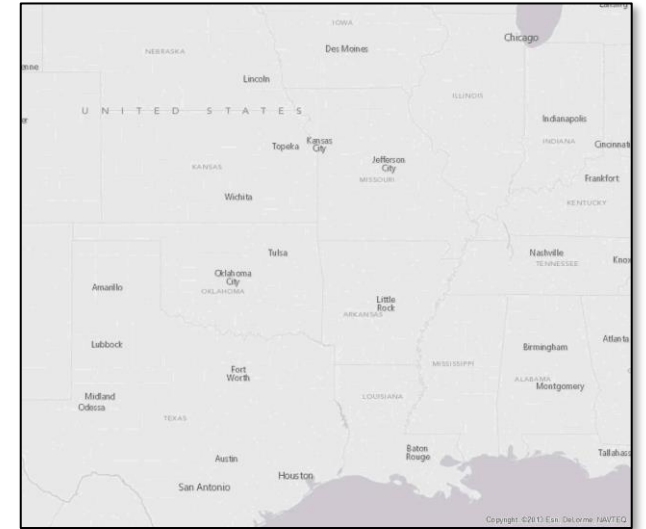
※ Pro 2.7で追加、Pro 2.9 で追加



解析機能の特徴

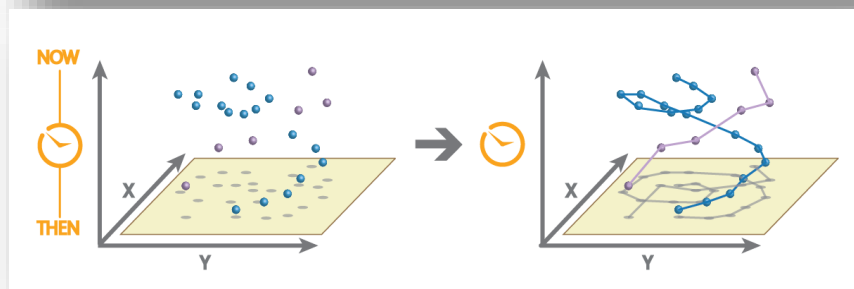
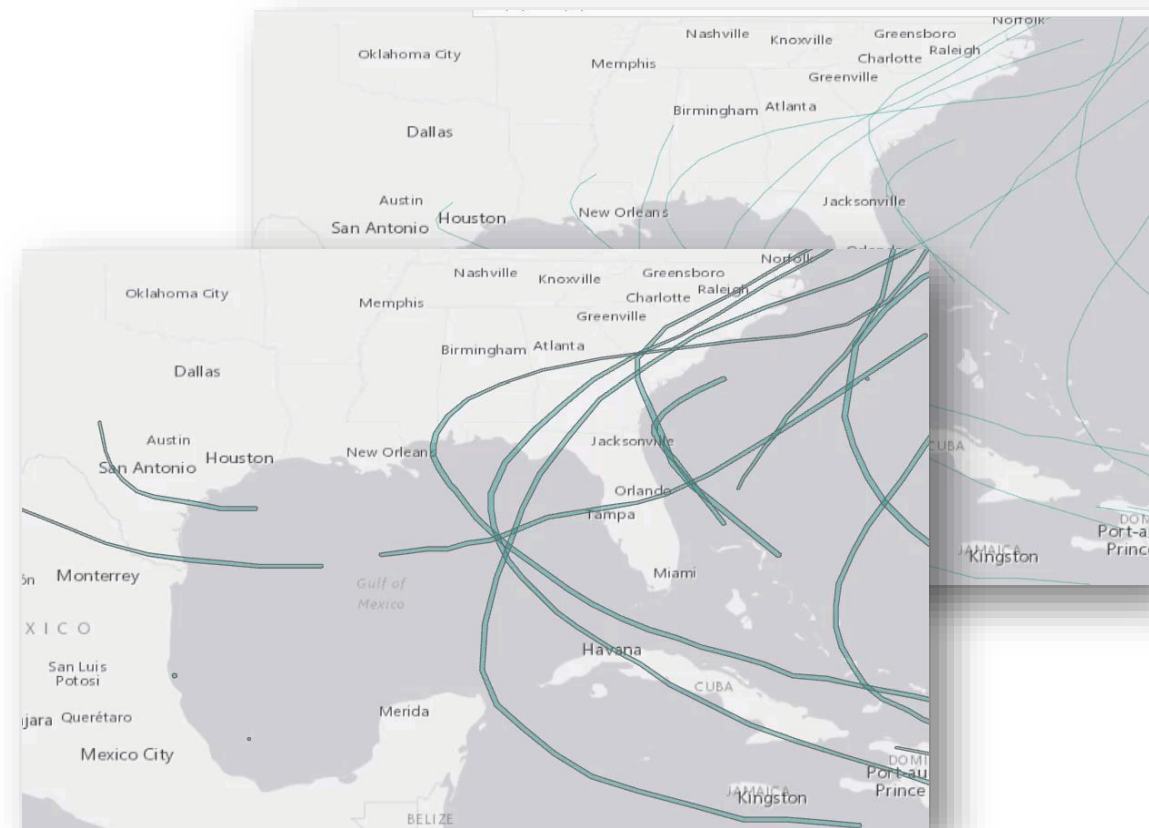
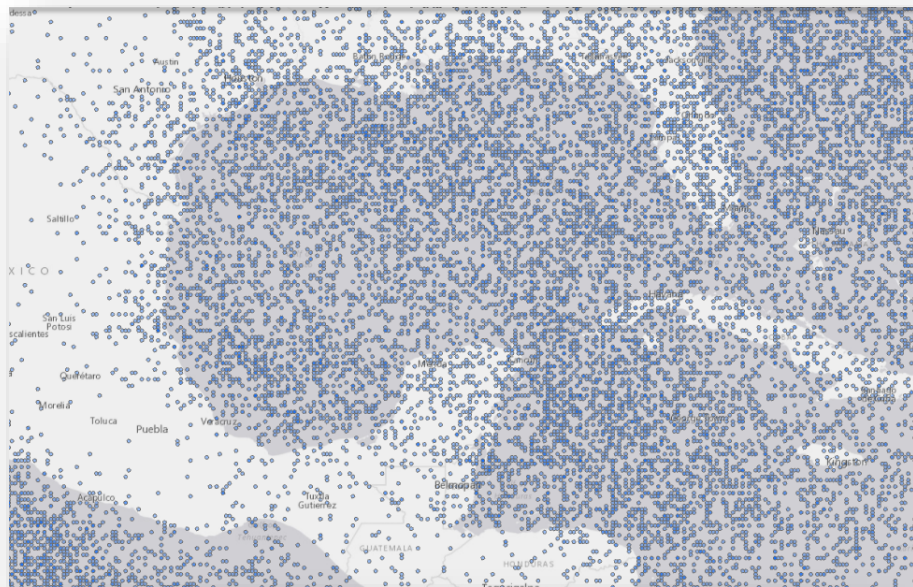
空間 と 時間 の両方でデータを操作

- GeoAnalytics を使って時空間分析を行う
- 時間の入力データを定義
 - インスタント 瞬間的なもの
 - インターバル 時間的に継続するもの
- 時間ステップでデータを分析する
- 結果はPro や Map Viewer を使って
タイム スライダーで可視化



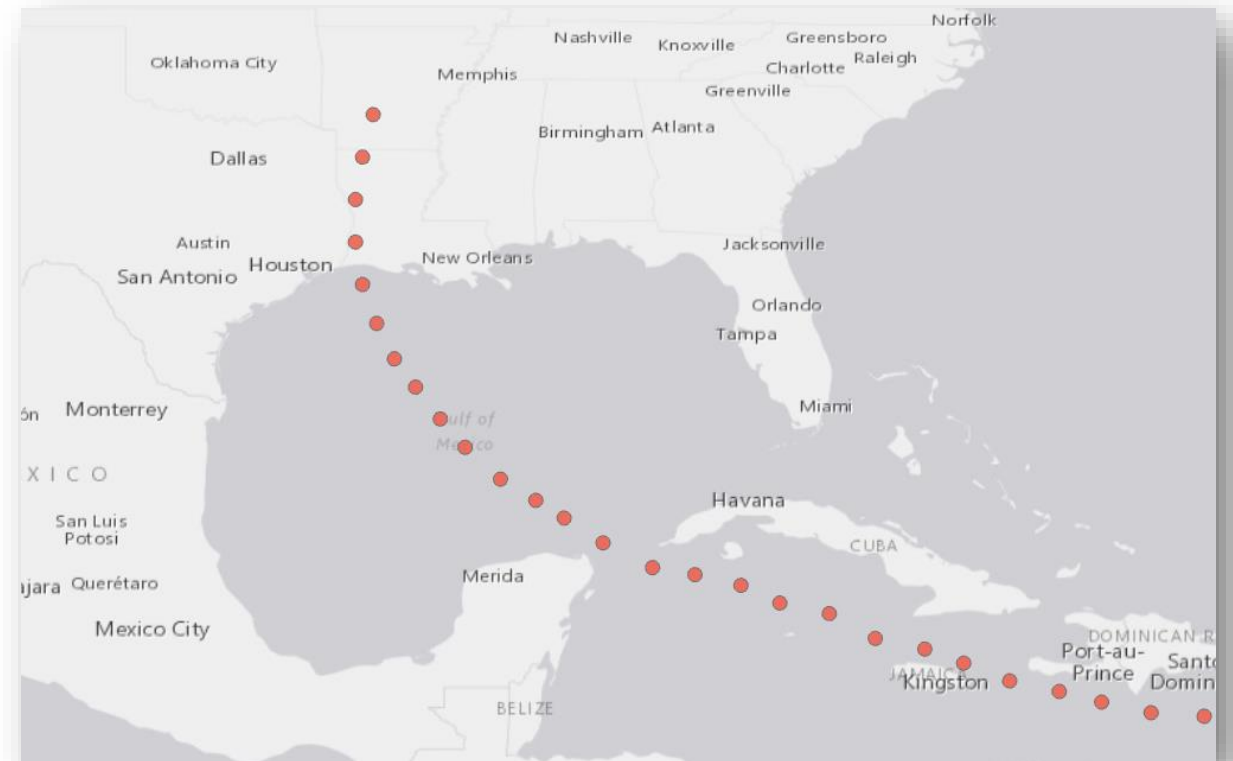
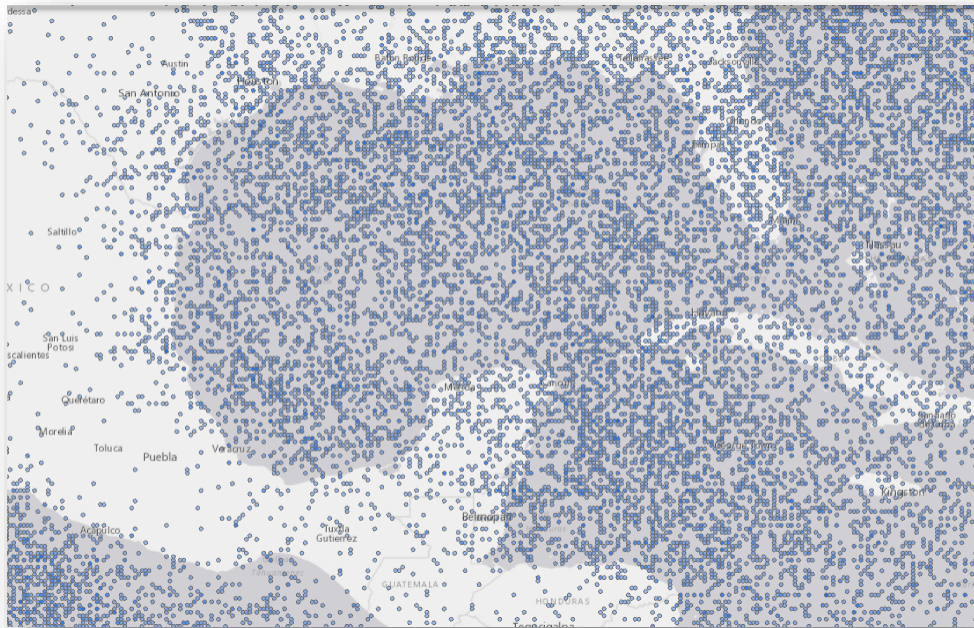
トラッキングの再構築 | Reconstruct Tracks

- 動きがある時系列位置情報をトラック 軌跡 として可視化



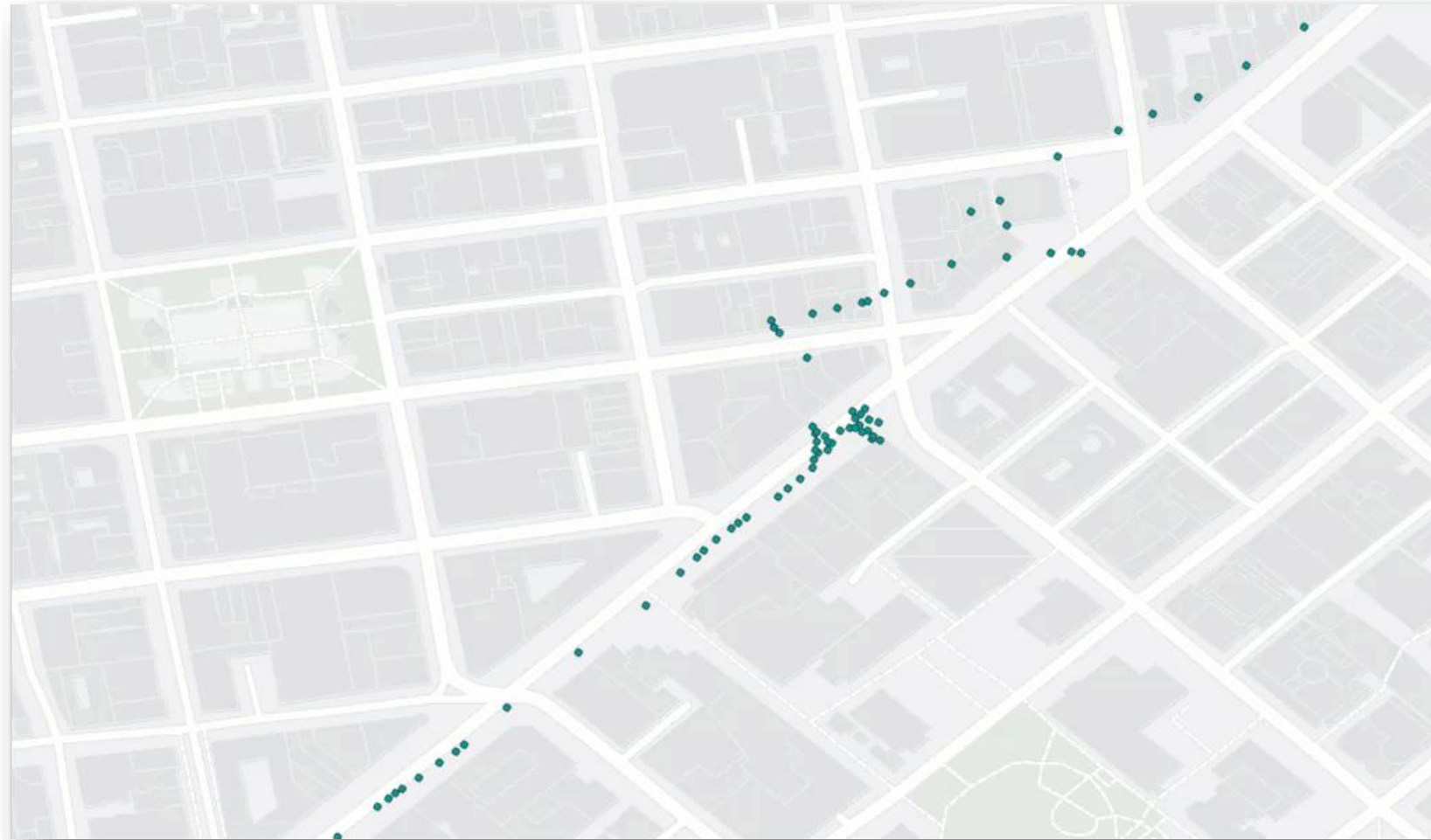
インシデントの検出 | Detect Incidents

- 指定された条件に合致するフィーチャを検出



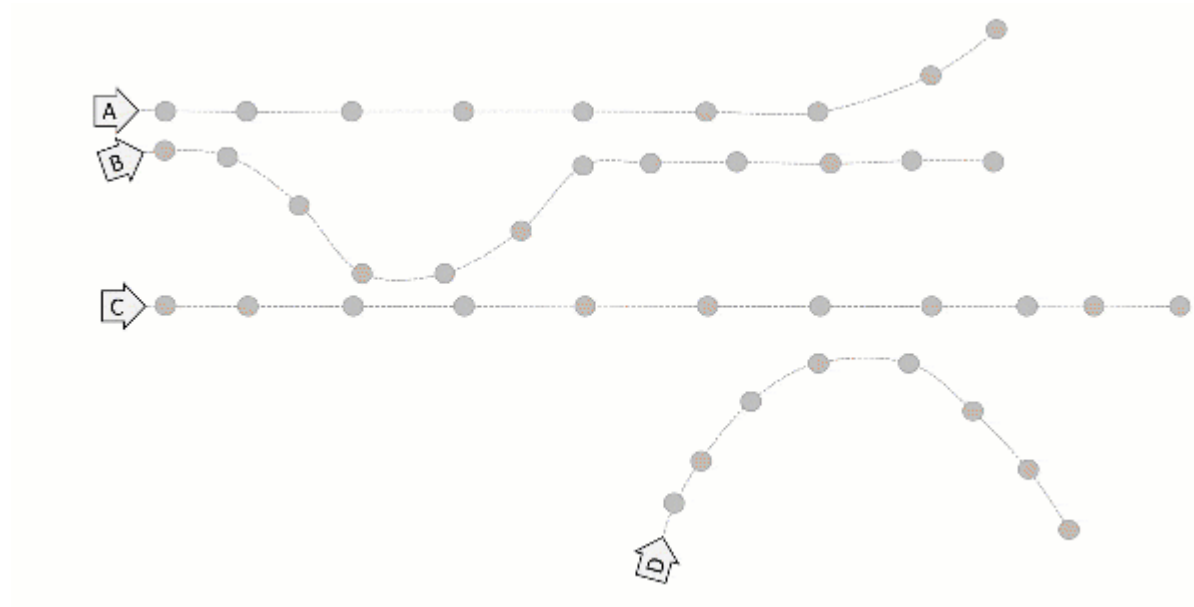
滞在場所の検索 | Find Dwell Locations

- 物体が静止、または静止に近い場所を探す



近接イベントのトレース | Trace Proximity Events

- 空間（場所）と時間で互いに近接するイベントをトレース

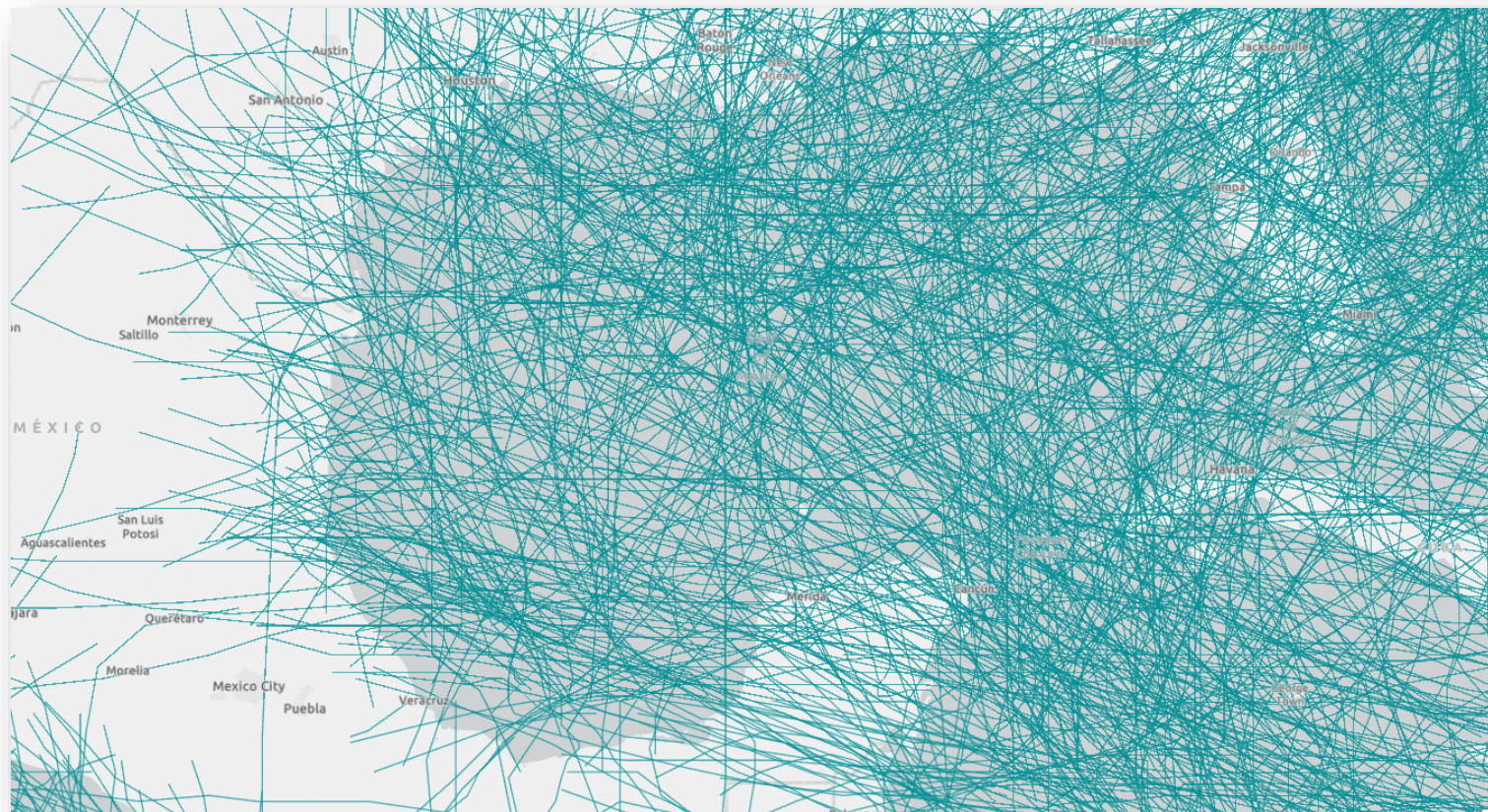


[Use Proximity Tracing to Identify Possible Contact Events](#)

トラックの再構築 + エリア内での集計

Reconstruct Tracks + Summarize Within

- 移動体のトラックを構築し、ビンに集約することによって“トラックヒートマップ”を作成



[Track Heat Mapping with GeoAnalytics Desktop tools](#)

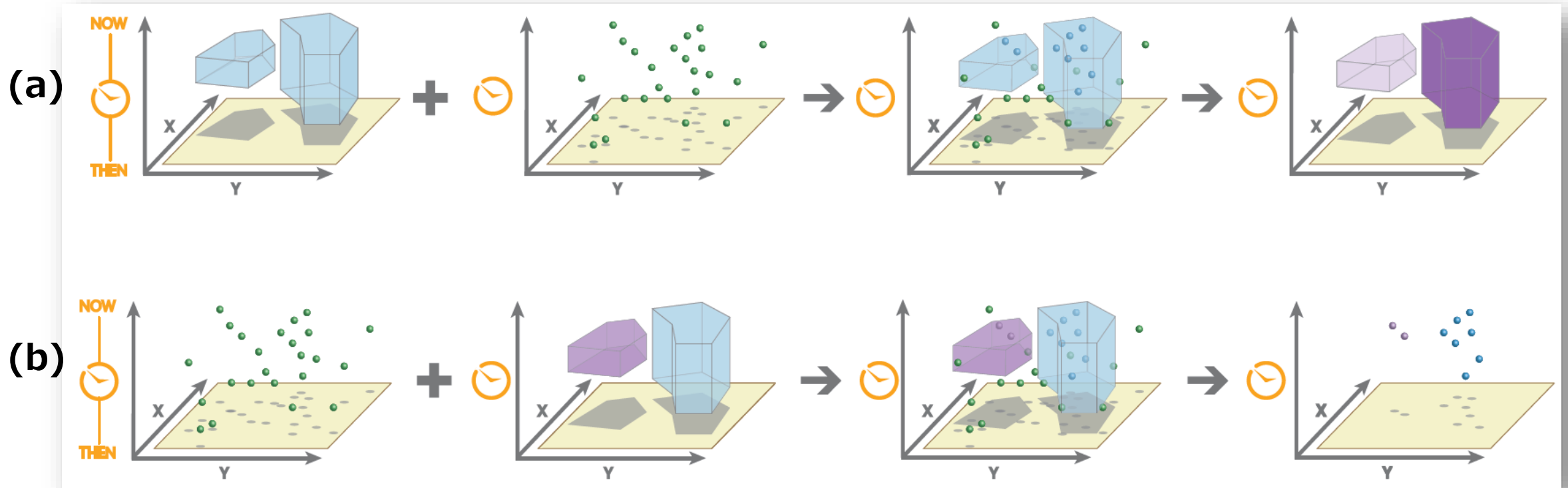


その他の時空間分析ツール

- **時空間結合**
 - **フィーチャの結合 | Join Features**
- **時間ステップ**
 - **ポイントの集約 | Aggregate Points**
 - **密度の計算 | Calculate Density**
 - **ホットスポット分析 | Find Hot Spots**
 - **時空間キューブの作成 | Create Space Time Cube**
- **時空間クラスタリング**
 - **ポイントクラスタの検索 | Find Point Clusters**

フィーチャの結合 | Join Features

時間的な関係性をもとにフィーチャを結合

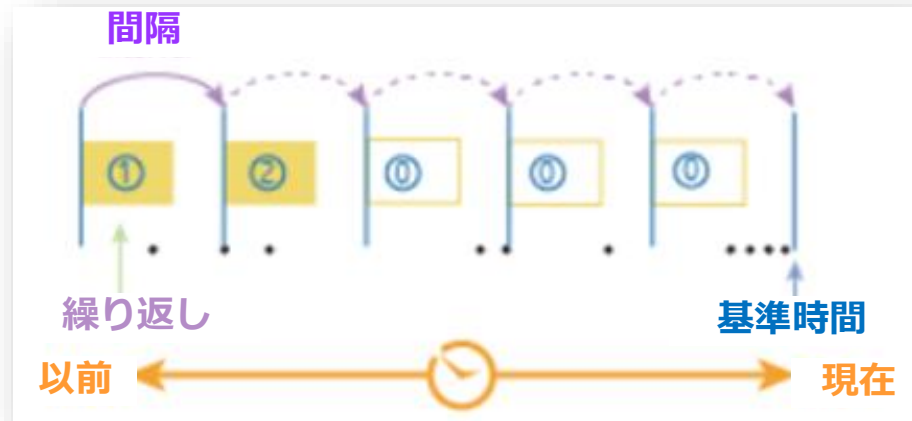


(a) ポリゴンからポイントフィーチャへ、(b) ポイントからポリゴンフィーチャへ の時空間結合

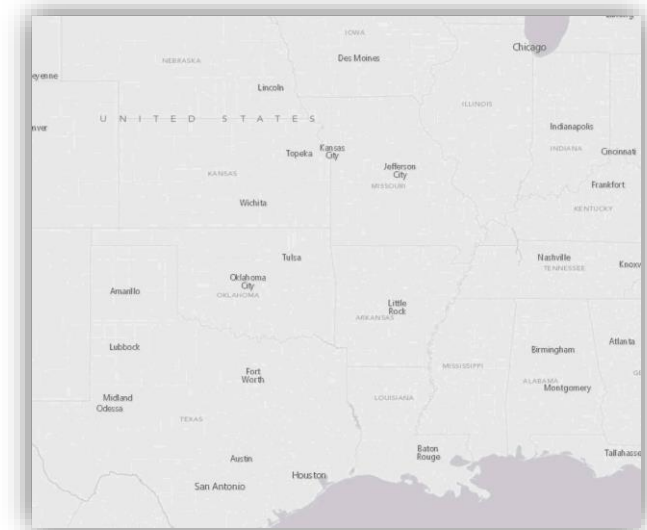
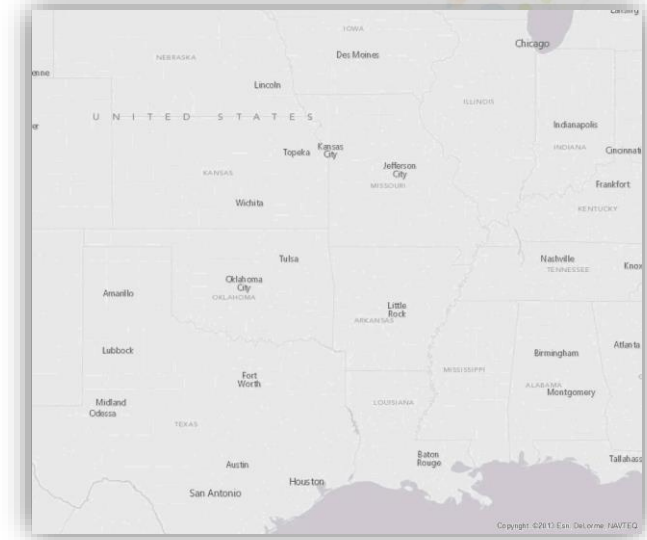
時間ステップ

指定した間隔でデータを分析する

- 時間ステップ間隔 Time step interval
- 時間ステップ繰り返し間隔 Time step repeat interval
- 基準時間 Reference time



時間ステップの仕組み



時空間クラスタリング | Spatiotemporal Clustering

次に基づいてポイントのクラスターを見つける

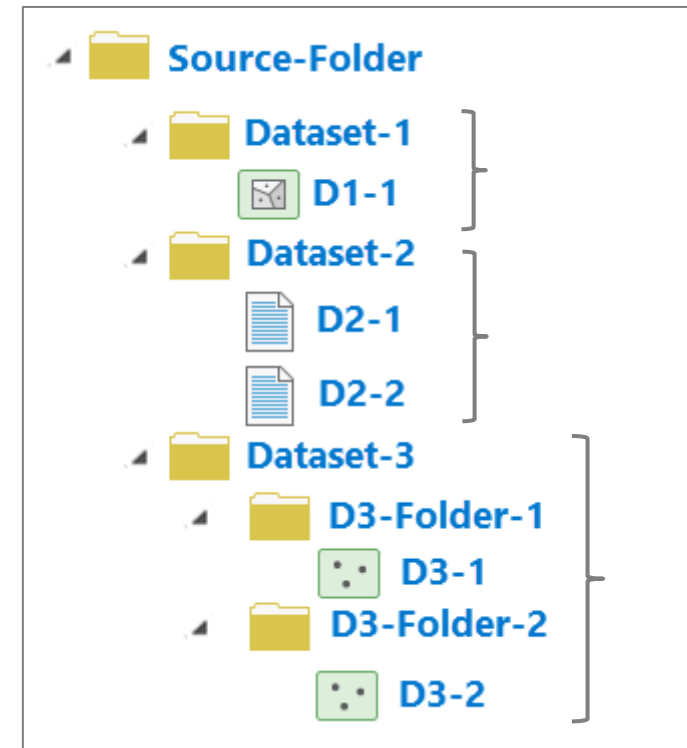
- 各クラスターに含まれる最小限のフィーチャ数
- 距離
- 時間的距離



ビッグデータ コネクション BDC の詳細

ディレクトリ内のデータセット コレクションから直接読み込み

- 区切りファイル (*.csv、*.tsv、*.txt)
- シェープファイル (*.shp)
- パッケージ ファイル (*.gz.parquet)
- ORC ファイル (*.orc.crc)



ビッグデータ コネクション BDC の使い時の例

- 同スキーマの複数のデータセットとファイルを、1つのデータセットとして扱うことができる。
- BDCは分析実行時にデータにアクセスするので、既存データセットへのデータ追加は、再登録や再公開の必要なく対応可能
- どのデータセットが表示されるか、削除、追加、更新をBDCの変更で対応することができる
- BDCは時間とジオメトリの定義が柔軟で、1つのデータセットに複数の時間フォーマットを使用できます

ビッグデータ コネクション の接続例1

経度緯度がある CSVデータ

約136万レコード
ファイルサイズ 約140MB

名前	種類	メタデータ	ジオグラフィ	テーブル
Miami	ビッグデータ ポイントフィーチャ			

◆元データはMarineCadastre.gov のAISデータ :

2009年～のAISデータが月ごとにまとまっている。2015年～は日付ごとのCSVデータとなっている (365日分=365ファイル)。

<https://marinecadastre.gov/ais/>

<ftp://ftp.coast.noaa.gov/pub/MSP/AIS/AIS.SampleData.zip>

BDC を介してCSVは、GeoAnalytics ツールの
インプット時にはポイント フィーチャクラスと
して扱われる

ビッグデータ コネクション の接続例2

経度緯度がない パーケットファイル 複数を同一フォルダーに配置

[合計]
約2千6百万レコード
ファイルサイズ 約318MB

BDC を介してパーケットファイルは、GeoAnalytics ツールのインプット時にはテーブルとして扱われる。テーブルの単位は元のフォルダ名ごと。

◆元データはNYC Yellow Taxi Trips データ (Pickup , Drop-off のTaxi ZoneのIDがあるPARQUETファイル。2009年～。Taxi Zone はシェープファイルで提供されている)

<https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page>

https://s3.amazonaws.com/nyc-tlc/misc/taxi_zones.zip

接続例2 での解析例

フィーチャの結合 | Join Features

The screenshot displays the ArcGIS Pro interface for a project named "GeoAnalytics - NYC_Taxi - ArcGIS Pro". The main map area shows a geographic view of New York City with taxi zones colored in a gradient from light orange to dark brown. The legend on the left side of the map shows the following layers and their FID ranges:

- NYC_Taxi
 - 解析A-OpenData_taxi_pickup_drop
 - 解析B-TLC_taxi_zone
 - nyc_taxi_zones_pickup_cnt (FID: 0 - 51, 52 - 104, 105 - 157, 158 - 209, 210 - 262)
 - nyc_taxi_zones_drop_cnt (FID: 0 - 50, 51 - 101, 102 - 157, 158 - 210, 211 - 262)
 - nyc_taxi_zones (FID: 0 - 50, 51 - 101, 102 - 157, 158 - 210, 211 - 262)
 - Light Gray Canvas Reference
 - Light Gray Canvas Base
 - スタンドアロンテーブル
 - yellow_2014

The interface includes a ribbon with various toolbars such as "プロジェクト", "マップ", "挿入", "解析", "表示", "編集", "画像", "共有", and "時間". The bottom status bar shows a scale of 1:243,060 and coordinates 73.6232175°W 40.5715052°N.

ビッグデータ コネクション の接続例3

経度緯度がある CSVデータ 複数を別々のフォルダーに配置

約1億6千5百万レコード
ファイルサイズ 約25.9GB

名前	種類
yellow_2014_nyc_taxi_data	ビッグデータ ポイントフィーチャ
yellow_2015_nyc_taxi_data	ビッグデータ ポイントフィーチャ

BDC を介してCSVは、GeoAnalytics ツールの
インプット時にはポイント フィーチャクラスと
して扱われる。ポイント フィーチャクラスの単
位は元のフォルダ名ごと。

◆元データはNYC Yellow Taxi Trips データ (Pickup , Drop-off の地点の緯度経度があるCSV。2014年, 2015年1月~6月のみ)

<https://data.cityofnewyork.us/Transportation/2014-Yellow-Taxi-Trip-Data/gn7m-em8n>

<https://data.cityofnewyork.us/Transportation/2015-Yellow-Taxi-Trip-Data/ba8s-jw6u>

接続例3 での解析例

ポイントの集約 | Aggregate Points

The screenshot displays the ArcGIS Pro interface for a GeoAnalytics project named 'NYC_Taxi'. The main map area shows a heatmap of taxi pickup and drop-off locations in Manhattan, New York City, with a concentration of points in the Midtown area. The heatmap is color-coded by count, with a legend on the left side of the map area showing the following ranges:

COUNT
1.000000 - 27.000000
27.000001 - 84.000000
84.000001 - 183.000000
183.000001 - 488.000000
488.000001 - 1316.000000

The interface also shows various toolbars and panels, including the 'Project' pane on the left, the 'Map' pane at the top, and the 'Feature Layer' pane on the right. The 'Feature Layer' pane shows the 'Aggregate Points' tool settings, including a time range from 2014/01/22 to 2014/01/23, a step count of 30, and a step interval of 1 day. The 'Map' pane shows the 'Aggregate Points' tool settings, including a time range from 2013/12/31 to 2015/01/01, a step count of 30, and a step interval of 1 day. The 'Feature Layer' pane shows the 'Aggregate Points' tool settings, including a time range from 2013/12/31 to 2015/01/01, a step count of 30, and a step interval of 1 day.



GeoAnalytics Desktop ツール を実行する際に考慮に入れること

知っておくべきこと

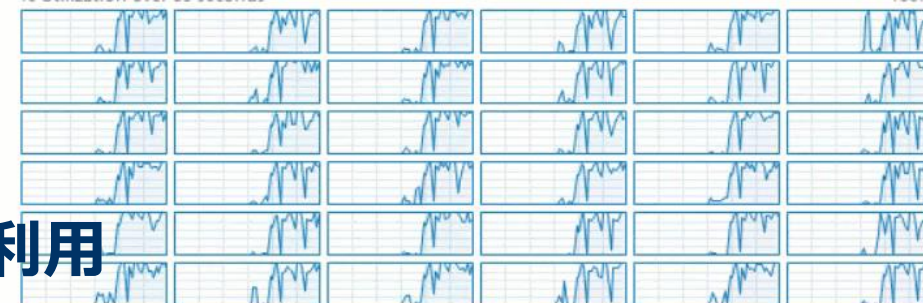
- **GeoAnalytics ツールはSpark を利用**
 - マシンのメモリの95% とすべてのコアを利用
- **並列処理ファクター**
 - コア使用量をコントロール
- **起動時間**
 - GeoAnalytics はSpark を開始するために5-15秒かかる
- **ツールの実装が異なる**
 - Proで同じようなツールを使っても、結果が全く同じ結果になる訳ではない

CPU

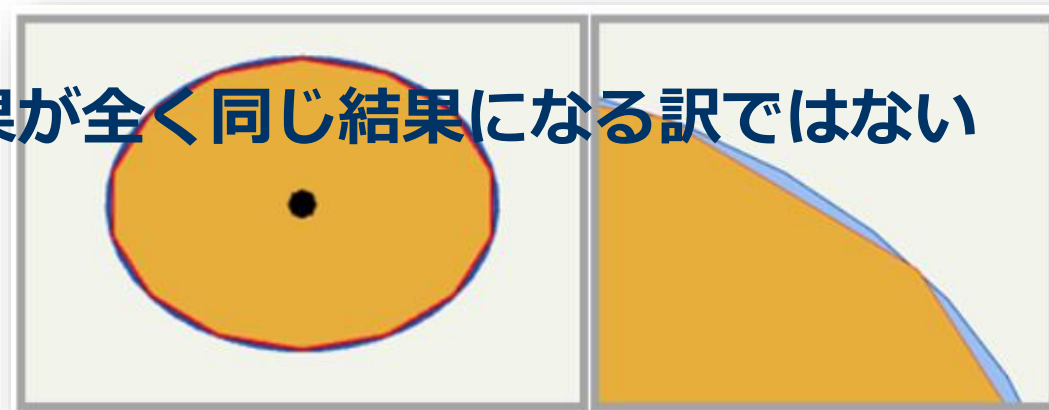
Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz

% Utilization over 60 seconds

100%



Utilization	Speed	Base speed:	2.30 GHz
100%	2.78 GHz	Sockets:	1
		Cores:	18
Processes	Threads	Handles	Logical processors: 36
166	3233	74719	Virtualization: Enabled
Up time		L1 cache:	1.1 MB
6:12:24:45		L2 cache:	4.5 MB
		L3 cache:	45.0 MB



パフォーマンス



ツールでの処理継続時間は複数の要因に左右される

入力データ	データソース	ツール + パラメータ	ハードウェア
フィーチャ数	コロケーション ※	使用している操作	RAM
ジオメトリの種類	ファイルの種類	近接の多さ	ディスク容量と スピード
頂点数		データの配置	コア数

※ コロケーションの例 (マシン上のデータへのBDC, EnterpriseGDB)

RAM についての詳細



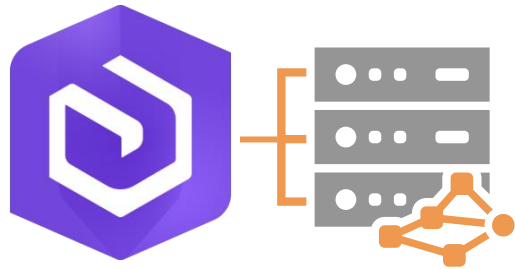
Spark は解析の実行にメモリ (RAM) を使います

- 可能であればメモリを増設 (64 > 32 > 16 GB RAM)
- メモリ上で完結できないジョブは、一時ディレクトリに書き込まれます
 - 一時ドライブを使用
 - 一時ドライブの容量が不足する場合、Windows の設定で変更することが可能です



より大量のデータを処理を実行するには？

より大きなデータの場合は？



GeoAnalytics Server

ArcGIS Enterprise による複数のサーバー
コアとマシンでの **分散処理**

ライセンス: Enterprise
+ GeoAnalytics Server License



GeoAnalytics Desktop

ArcGIS Pro によるラップトップもしくはデ
スクトップで複数コアを使う **並列処理**

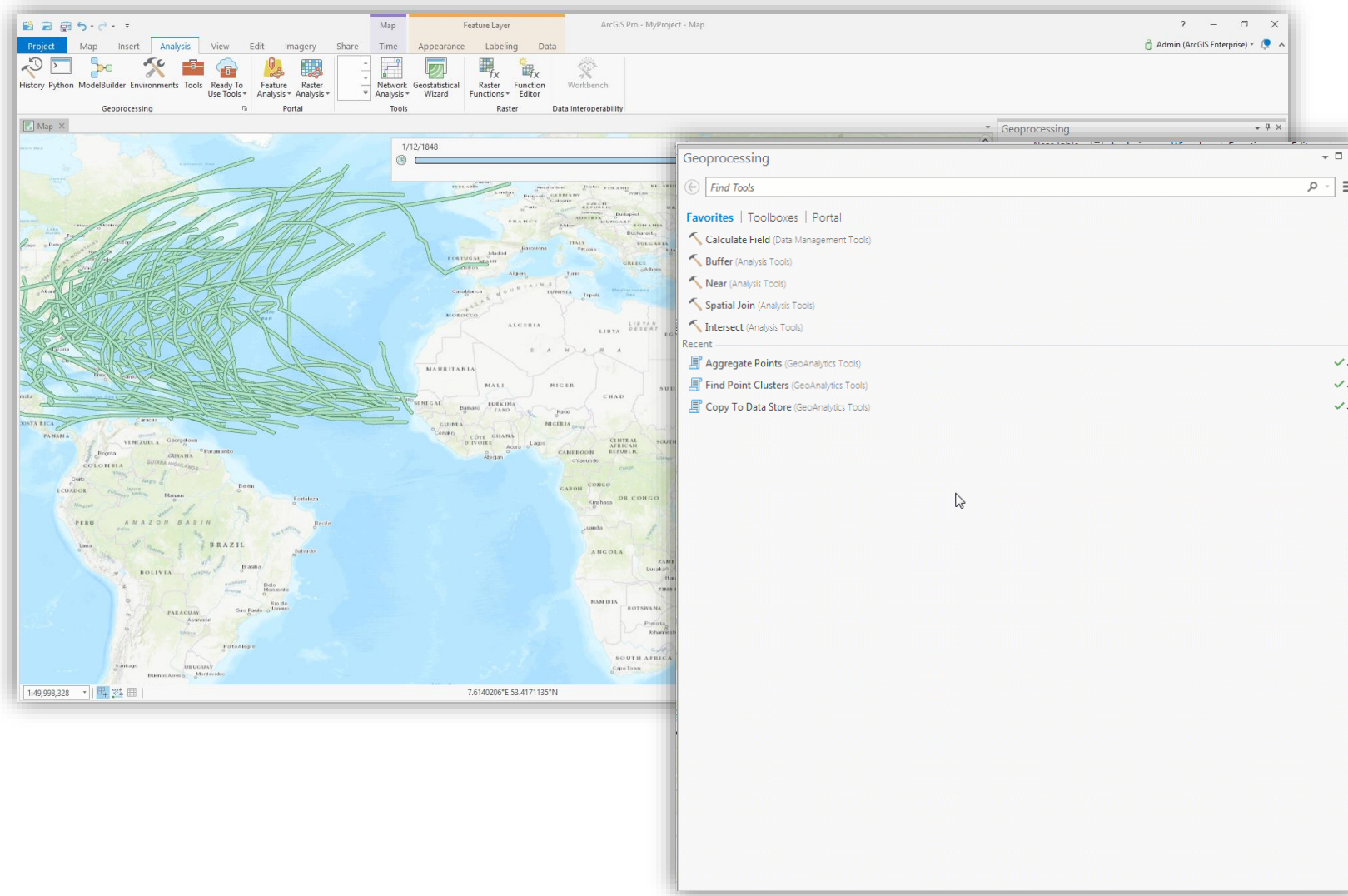
ライセンス: Advanced License



GeoAnalytics Server

GeoAnalytics Server は ArcGIS Pro で利用可能

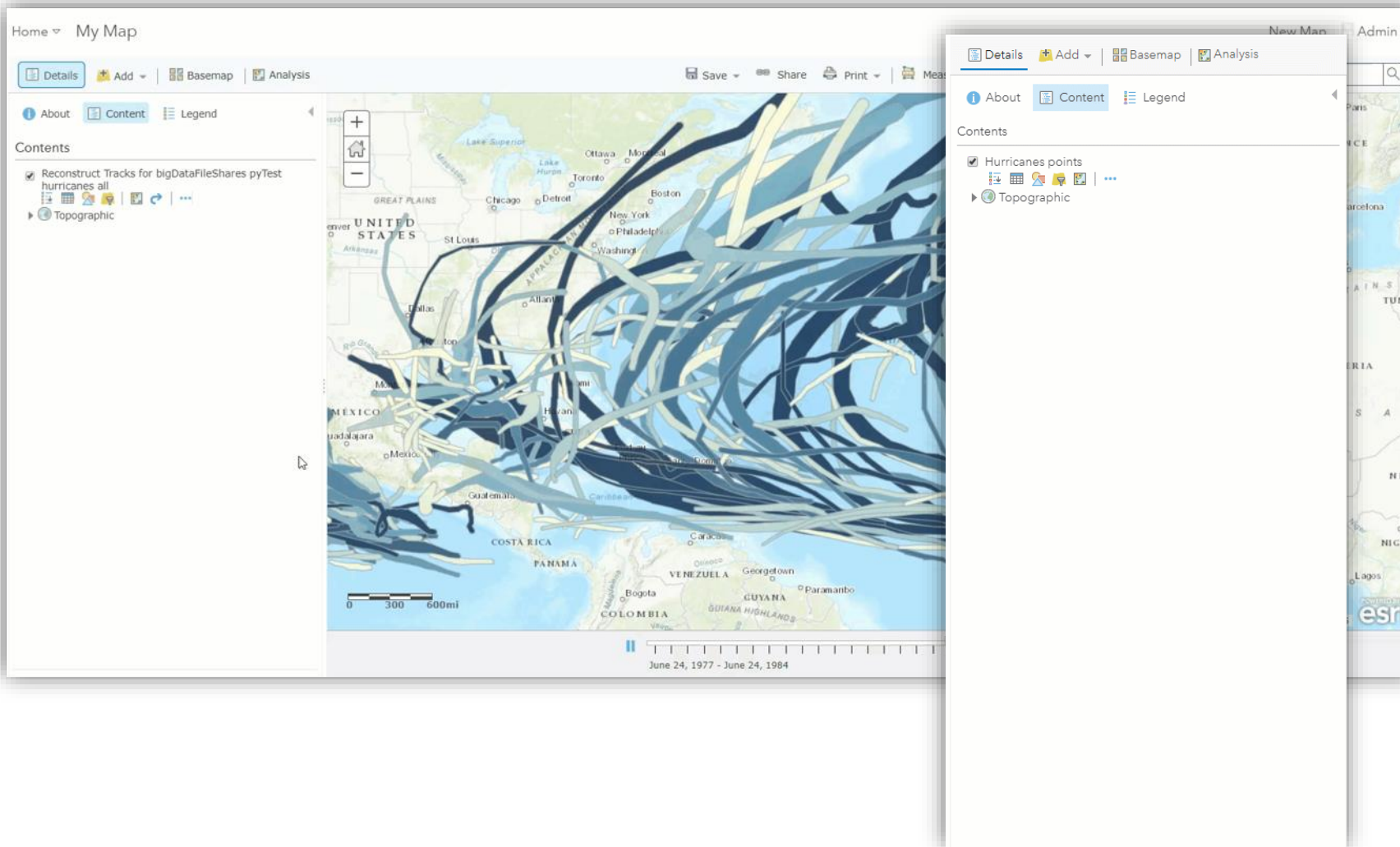
使い慣れたユーザーインターフェイスから実行可能



ArcGIS Pro

GeoAnalytics Server は ArcGIS Enterprise (portal) で利用可能

使い慣れたユーザーインターフェイスから実行可能



ArcGIS Pro

Enterprise portal

GeoAnalytics Server は ArcGIS API for Python で利用可能

開発者向けインターフェイス



jupyter GeoAnalytics_Sample_Notebook (autosaved) Python 3

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

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

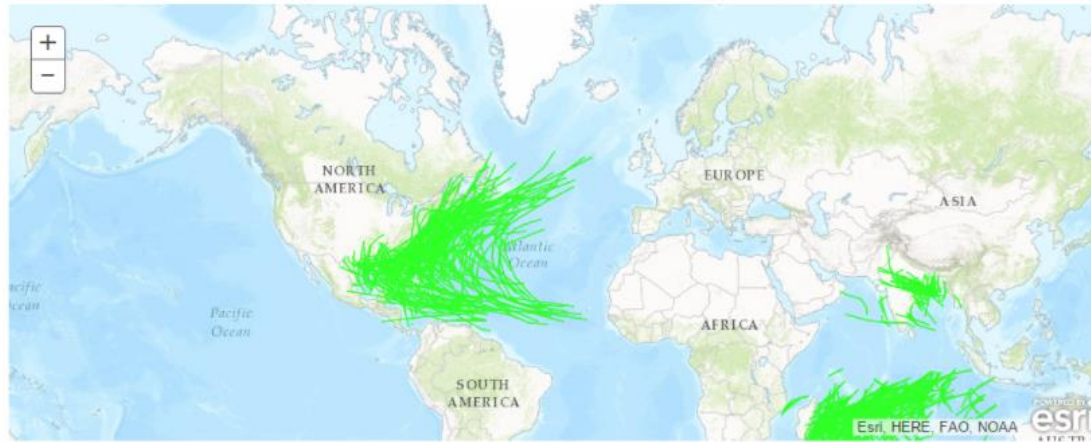
```
In [ ]: # Import the arcpy module
# and the GeoAnalytics module

from arcpy.gis import GeoAnalytics
```

To modify this notebook

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

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



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

ArcGIS Pro

Enterprise portal

ArcGIS API for Python

GeoAnalytics Server は ArcGIS REST API で利用可能

開発者向けインターフェイス



ArcGIS REST Services Directory Logged in user : admin | [Logout](#) |

[Home](#) > [services](#) > [System](#) > [GeoAnalyticsTools \(GPServer\)](#) [Help](#) | [API Reference](#)

[JSON](#) | [SOAP](#)

System/GeoAnalyticsTools (GPServer)

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

Tasks:

- [AggregatePoints](#)
- [DescribeDataset](#)
- [JoinFeatures](#)
- [CreateBuffers](#)
- [CalculateDensity](#)
- [FindPointClusters](#)
- [ReconstructTracks](#)
- [CreateSpaceTimeCube](#)
- [CopyToDataStore](#)
- [SummarizeAttributes](#)
- [SummarizeWithin](#)
- [FindSimilarLocations](#)
- [FindHotSpots](#)
- [CalculateField](#)
- [DetectIncidents](#)
- [GeocodeLocations](#)
- [OverlayLayers](#)
- [BuildMultiVariableGrid](#)
- [AppendData](#)

Execution Type: esriExecutionTypeAsynchronous

Result Map Server Name:

MaximumRecords: 1000

Child Resources: [Info](#)

ArcGIS Pro

Enterprise portal

*ArcGIS API for
Python*

ArcGIS REST API



解析機能

解析機能

パターンの分析

密度の計算 | Calculate Density
時空間キューブの作成 | Create Space Time Cube
ホットスポット分析 | Find Hot Spots
ポイントクラスターの検索 | Find Point Clusters
フォレストベースの分類
と回帰分析 | Forest-based Classification and Regression
一般化線形回帰分析 | Generalized Linear Regression
地理空間加重回帰分析 | Geographically Weighted Regression

近接エリアの分析

バッファの作成 | Create Buffers
近接イベントのトレース | Trace Proximity Events

データの管理

データのアペンド | Append Data
フィールド演算 | Calculate Field
レイヤーのクリップ | Clip Layer
データストアにコピー | Copy to Data Store
境界のディゾルブ | Dissolve Boundaries
レイヤーのマージ | Merge Layers
オーバーレイ | Overlay Layers

データの集約

ポイントの集約 | Aggregate Points
多変数グリッドの構築 | Build Multi-Variable Grid
データセットの記述 | Describe Dataset
フィーチャの結合 | Join Features
トラッキングの再構築 | Reconstruct Tracks
属性の集計 | Summarize Attributes
中心と分散の集計 | Summarize Center and Dispersion
エリア内での集計 | Summarize Within

場所検索

インシデントの検出 | Detect Incidents
滞在場所の検索 | Find Dwell Locations
類似フィーチャの検索 | Find Similar Locations
テーブルから場所をジオコーディング | Geocode Locations

データへの情報付加

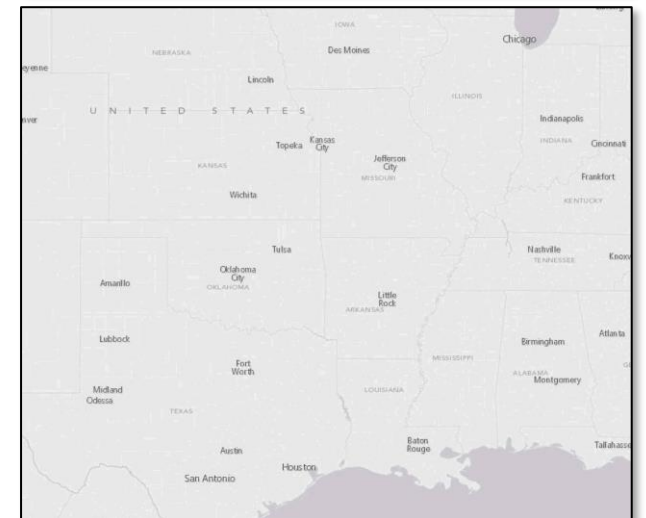
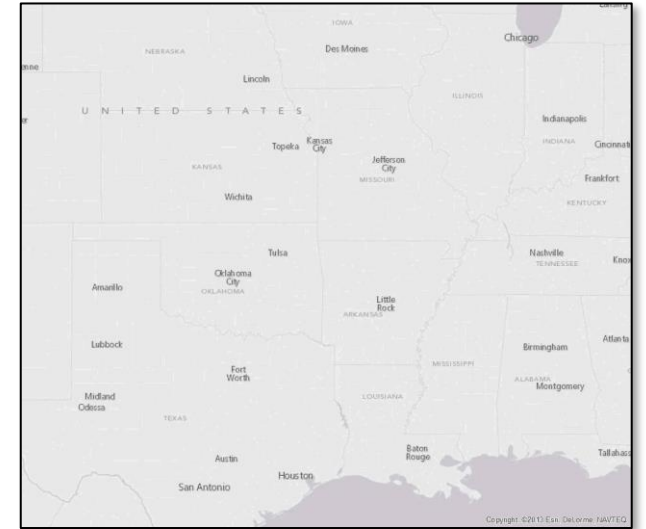
モーションの統計量を計算 | Calculate Motion Statistics
多変数グリッドから情報付加 | Enrich from Multi-Variable Grid

※ 10.9 での新機能

解析機能の特徴

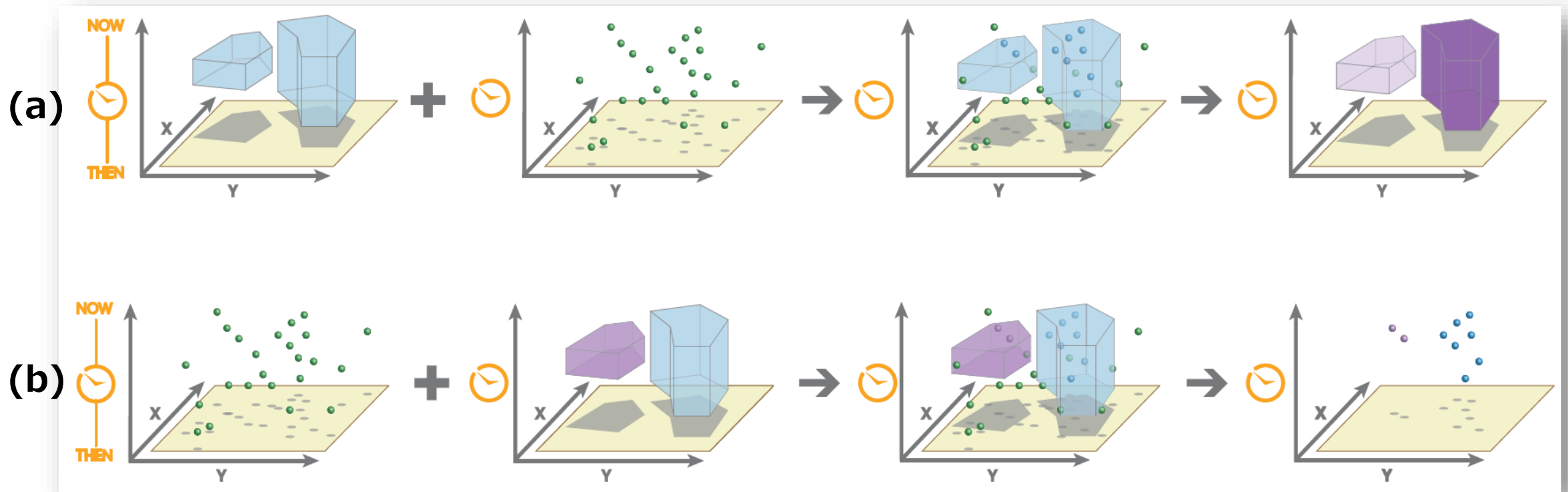
空間 と 時間 の両方でデータを操作

- GeoAnalytics を使って時空間分析を行う
- 時間の入力データを定義
 - インスタント 瞬間的なもの
 - インターバル 時間的に継続するもの
- 時間ステップでデータを分析する
- 結果はPro や Map Viewer を使って
タイム スライダーで可視化



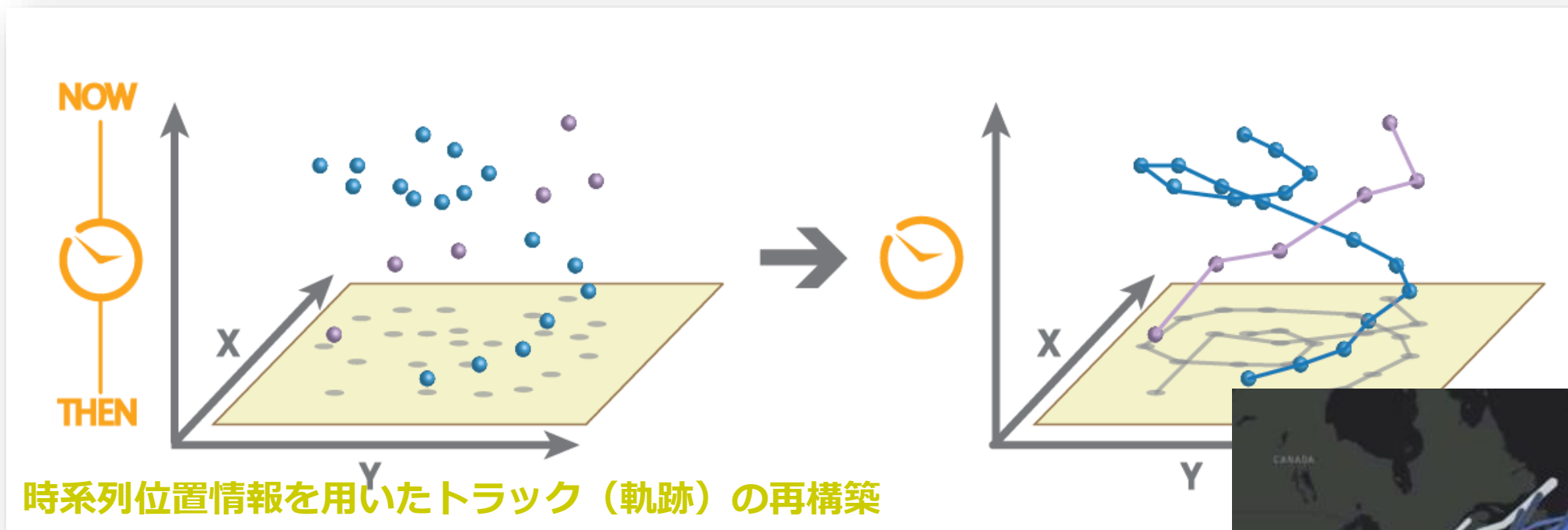
フィーチャの結合 | Join Features

位置, 時間, 属性 でデータを結合

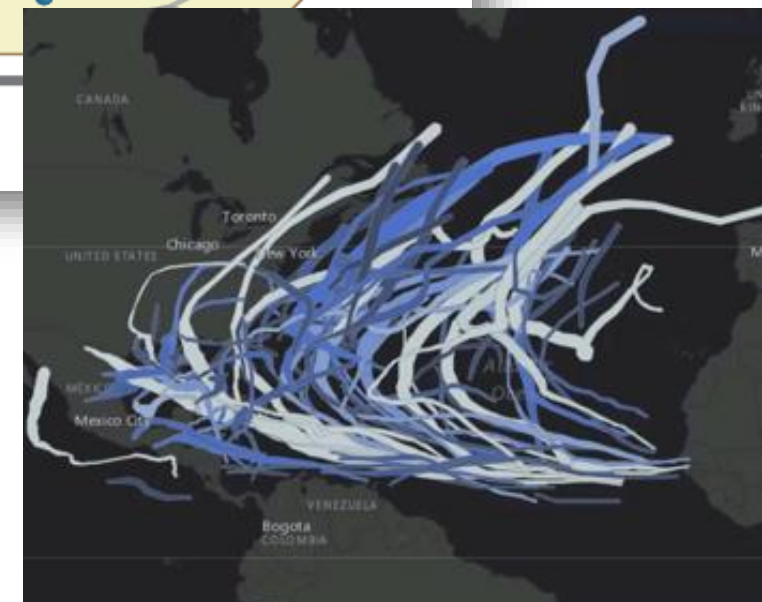


(a) ポリゴンからポイントフィーチャへ、(b) ポイントからポリゴンフィーチャへ の 時空間結合

トラッキングの再構築 | Reconstruct Tracks



- **トラック（軌跡）データの例:**
 - ハリケーン
 - 船舶
 - 航空機

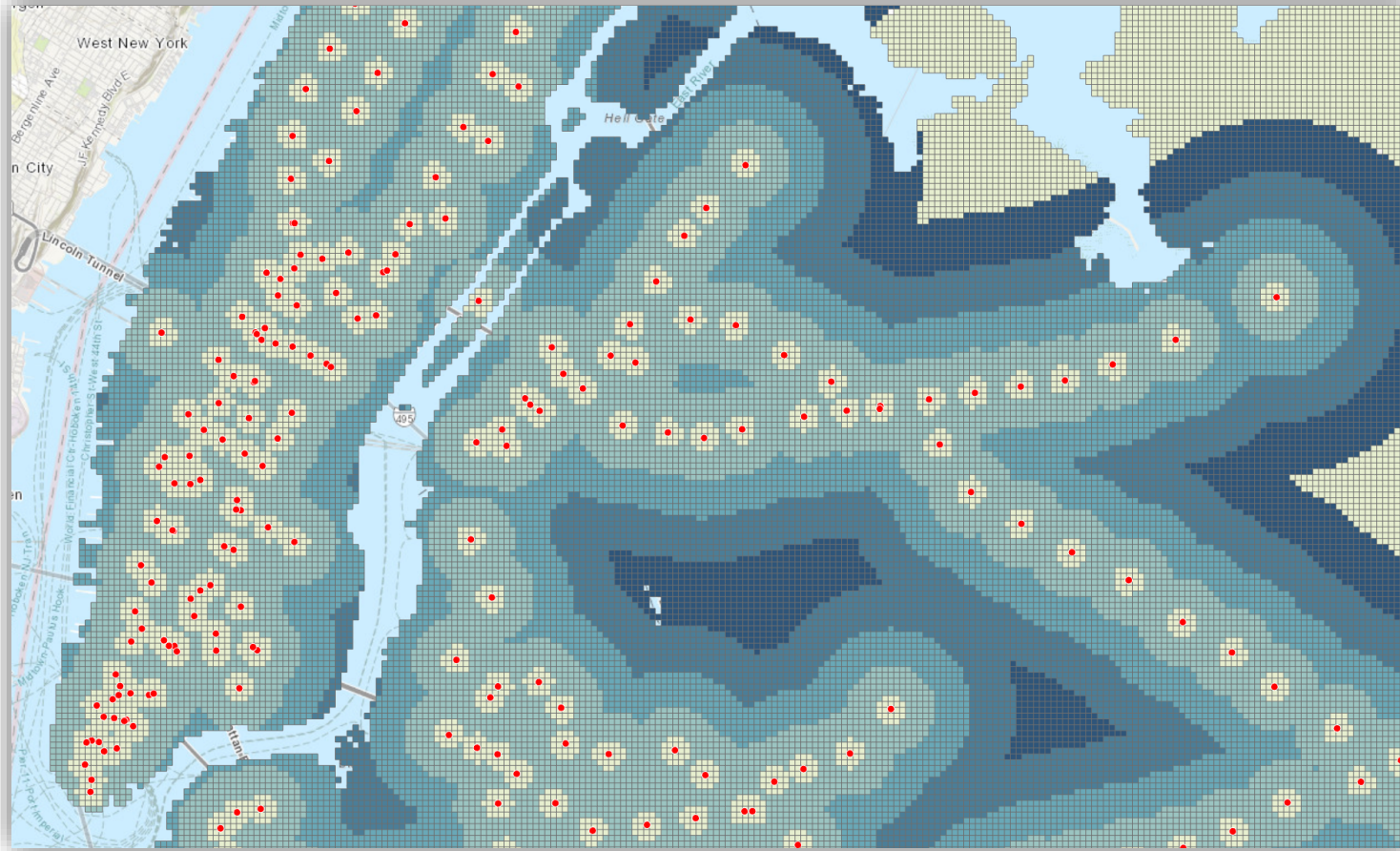


多変数グリッドの構築 | Build Multi-Variable Grid

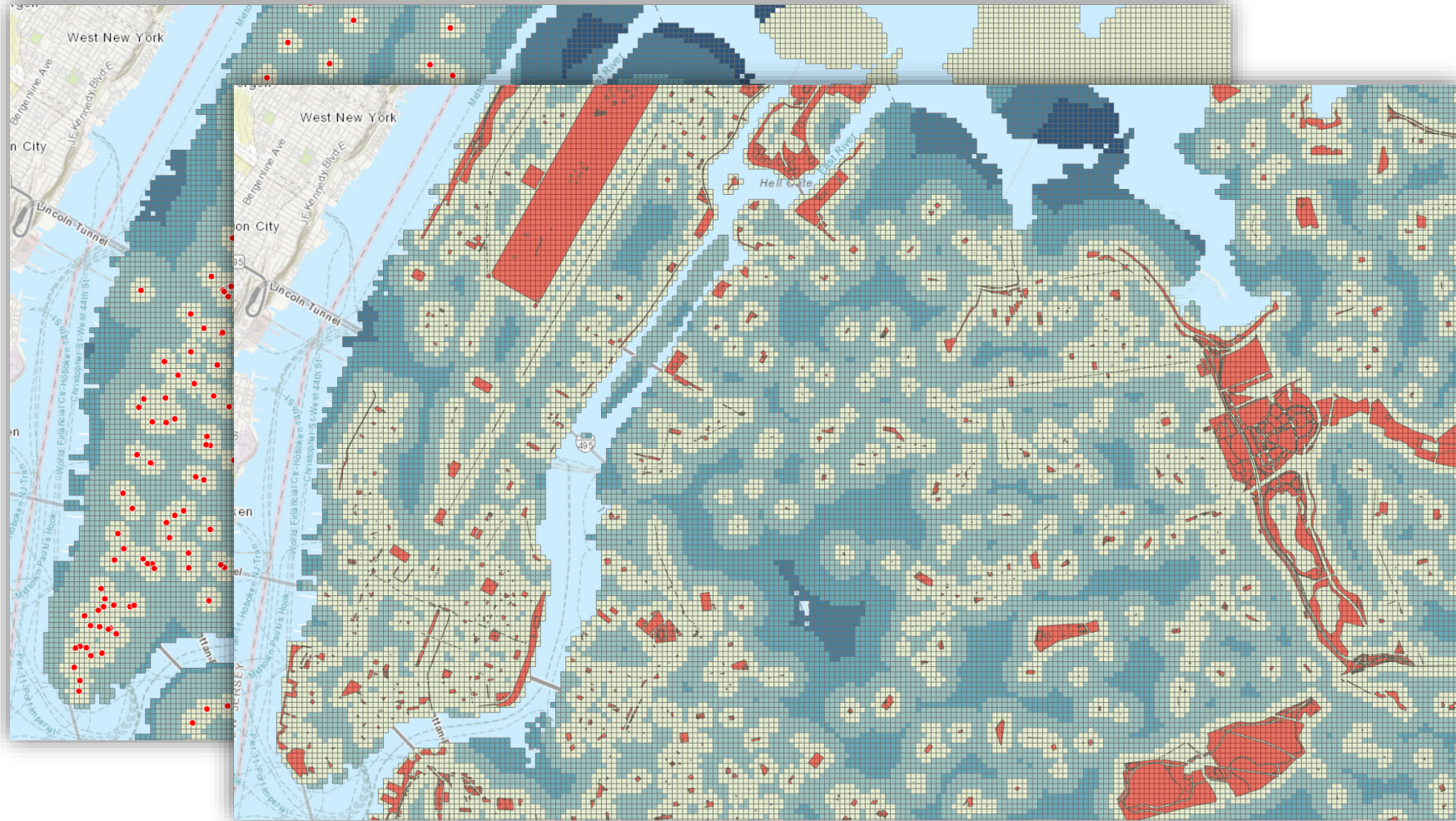


- **次を計算することによって、複数のデータセットを1つに集約するために作られた:**
 - 最寄りまでの距離
 - 最寄りの属性
 - インターセクトのサマリ
 - 一定距離内のサマリ
- …関心のある 1つ、もしくは複数レイヤーに対して
- **このレイヤーは、以下として使用します:**
 - 可視化
 - 統計ツール (GLR + Forest-based Classification and Regression)のインプット
 - レイヤーへの情報付加 (An enrichment layer)

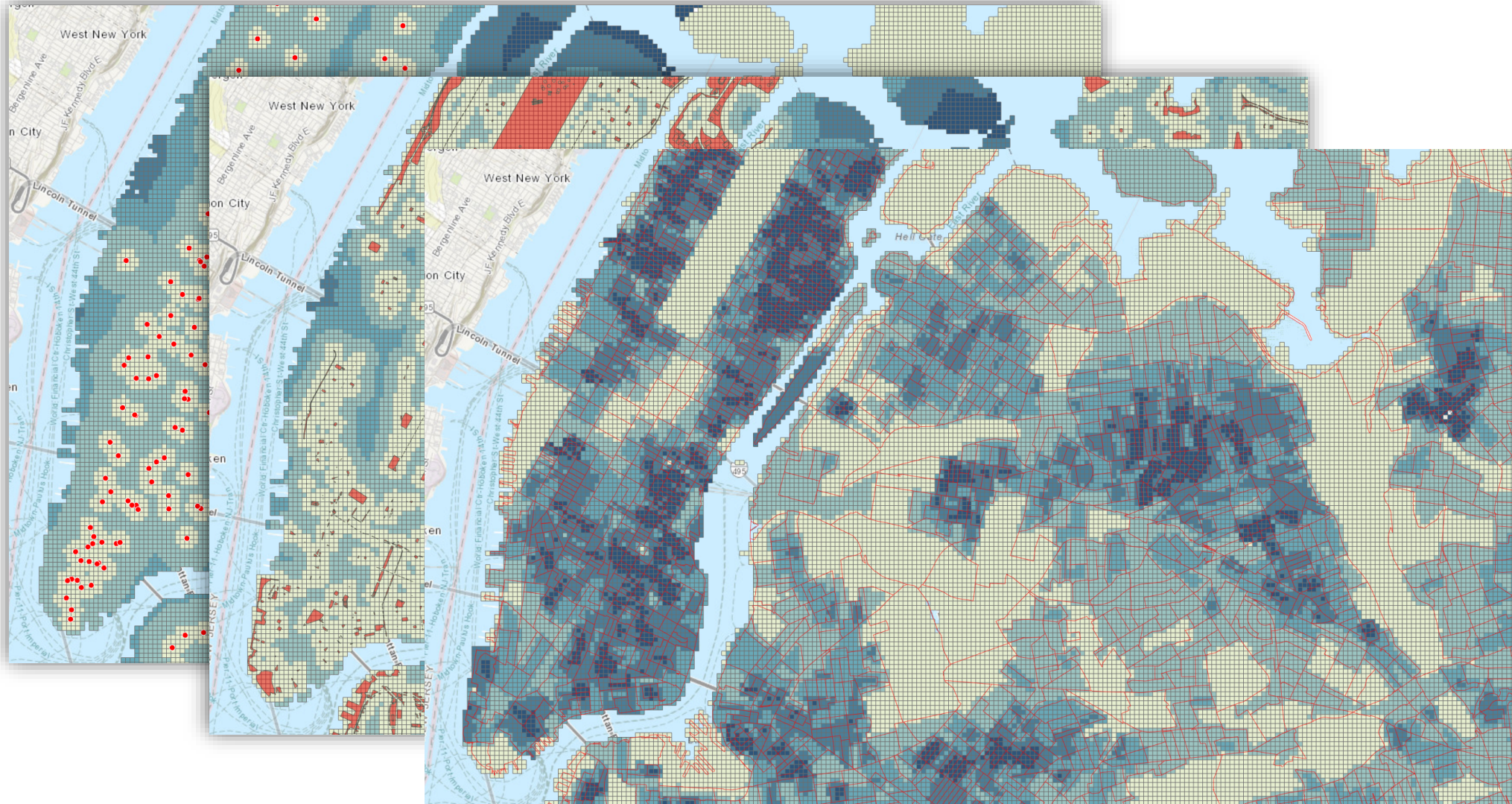
多変数グリッドの構築 – 最寄りの地下鉄駅



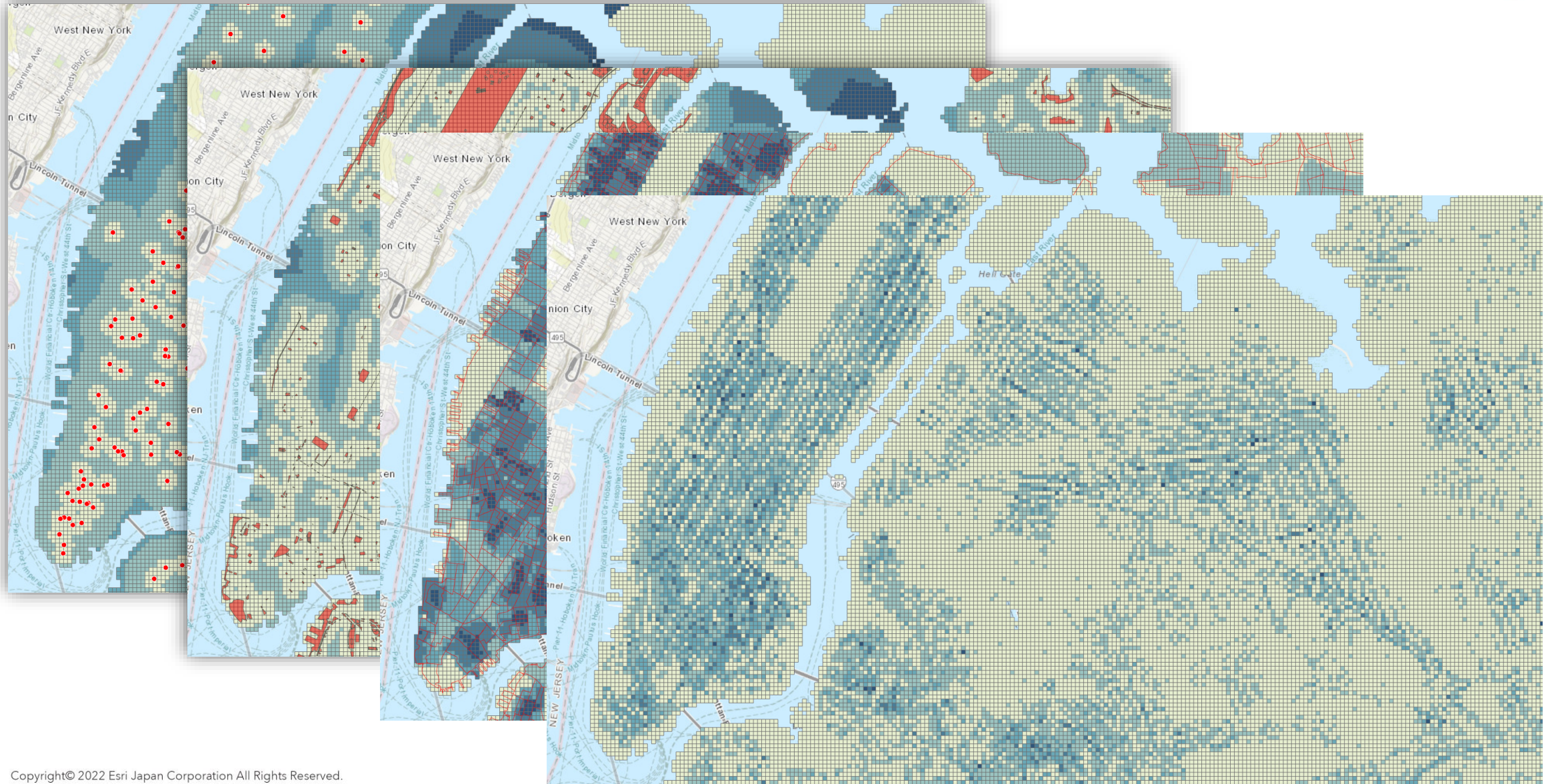
多変数グリッドの構築 – 最寄りの公園



多変数グリッドの構築 – 1平方マイルあたりの人口



多変数グリッドの構築 – 苦情件数



GeoAnalytics Server でPySpark にアクセスし使用する

- **Run Python Script** を使用して分散解析を実行
 - GeoAnalytics Server サイトでカスタム python script を実行する
 - 他の python の機能を利用してサイト全体に解析を分散させることが可能
 - GeoAnalytics ツールと連鎖させるパイプライン処理を作成可能
 - pyspark (ml, sql) と data frames を利用可能





データ統合

GeoAnalytics Server

どのようなデータを分析できますか？ また、どこから？



Enterprise portal で Enterprise 内のデータにアクセスし、共有することが可能



ArcGIS GeoEvent Server で収集したデータをシームレスに分析できる



Hive, HDFS, と **ファイル** のデータを分析できる



クラウドストア で **Azure** や **Amazon** へ接続

Big Data へ簡単にアクセス

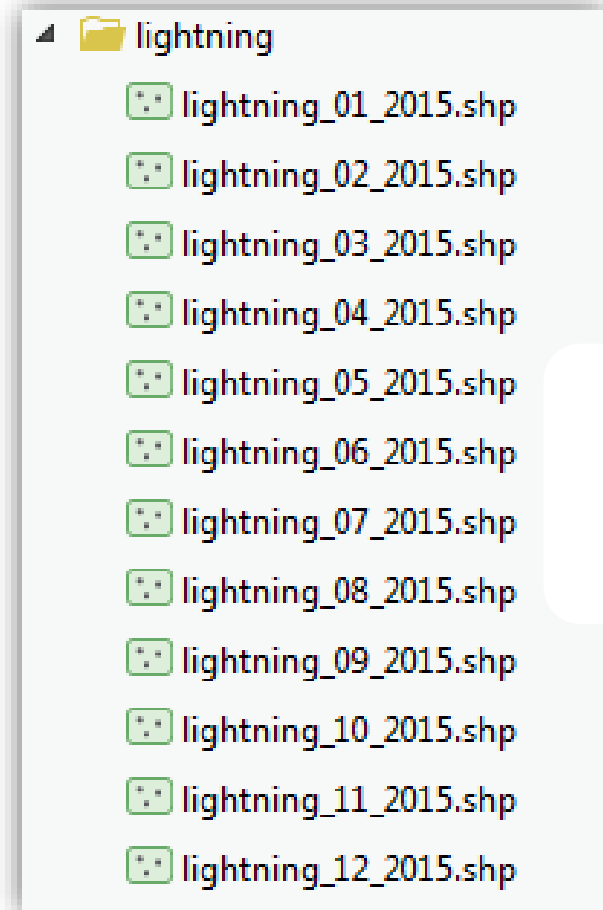
ビッグデータ ファイル共有

次に保存されたファイルから直接読み込む

- Hive (Hadoop 用のデータ ウェアハウス システム)
- HDFS (Hadoop Distributed File System)
- フォルダー 共有
- クラウド ストア

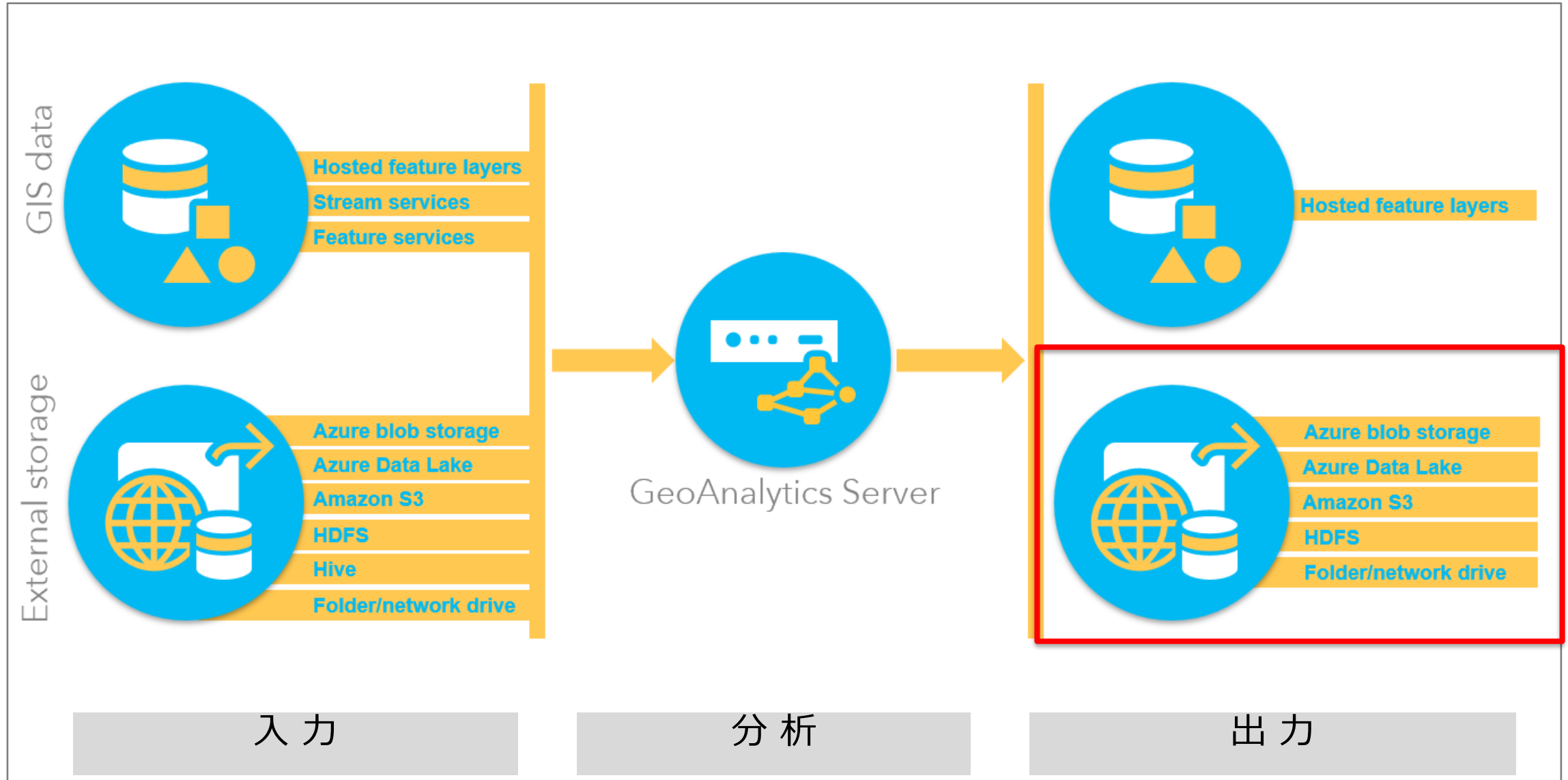
対応するファイル形式

- 区切りファイル
- シェープファイル
- ORC ファイル
- パーケット ファイル



同スキーマの複数ファイルを1つのデータとして扱うことが可能

ビッグデータ ファイル共有への書き込み（10.7 以降）





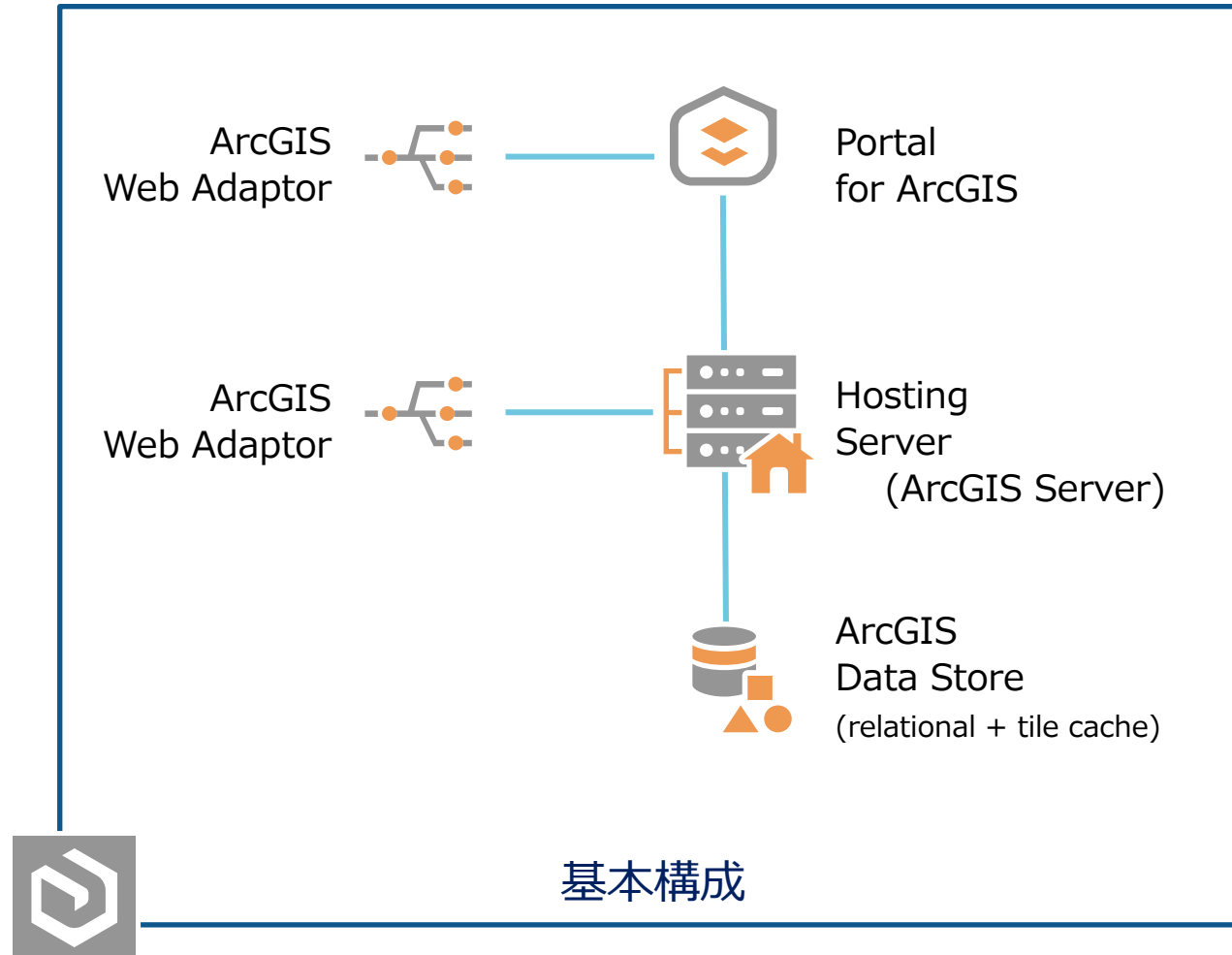
GeoAnalytics Server の配置

ArcGIS GeoAnalytics Server を使用するのにインストールが必要なものは?

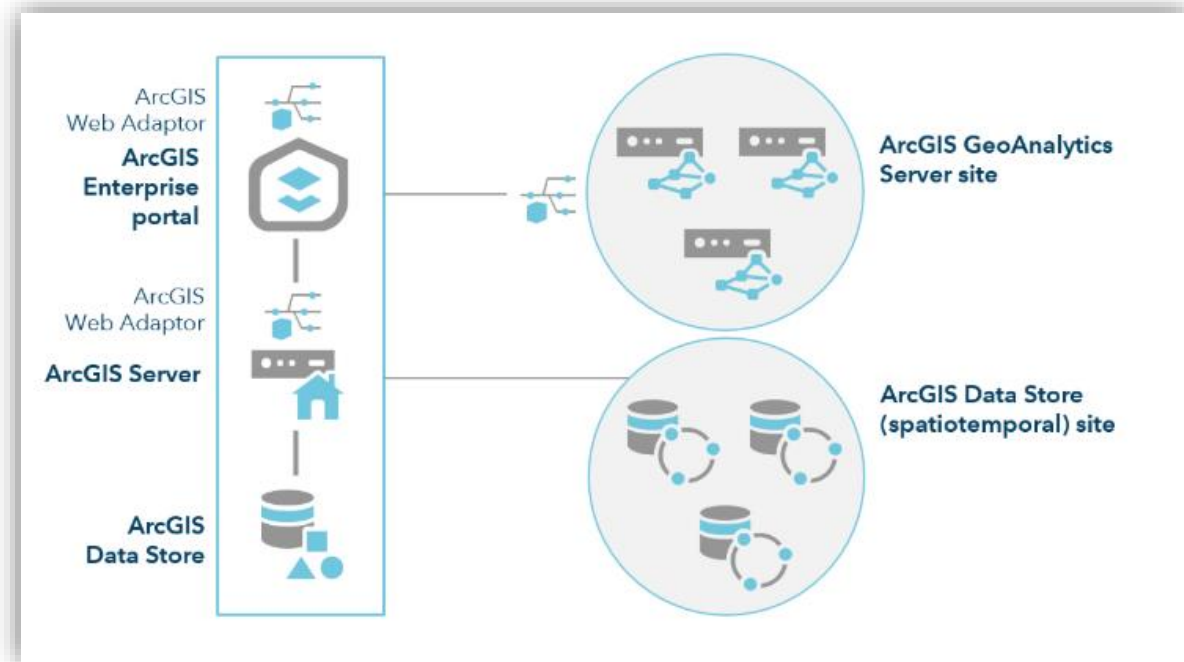
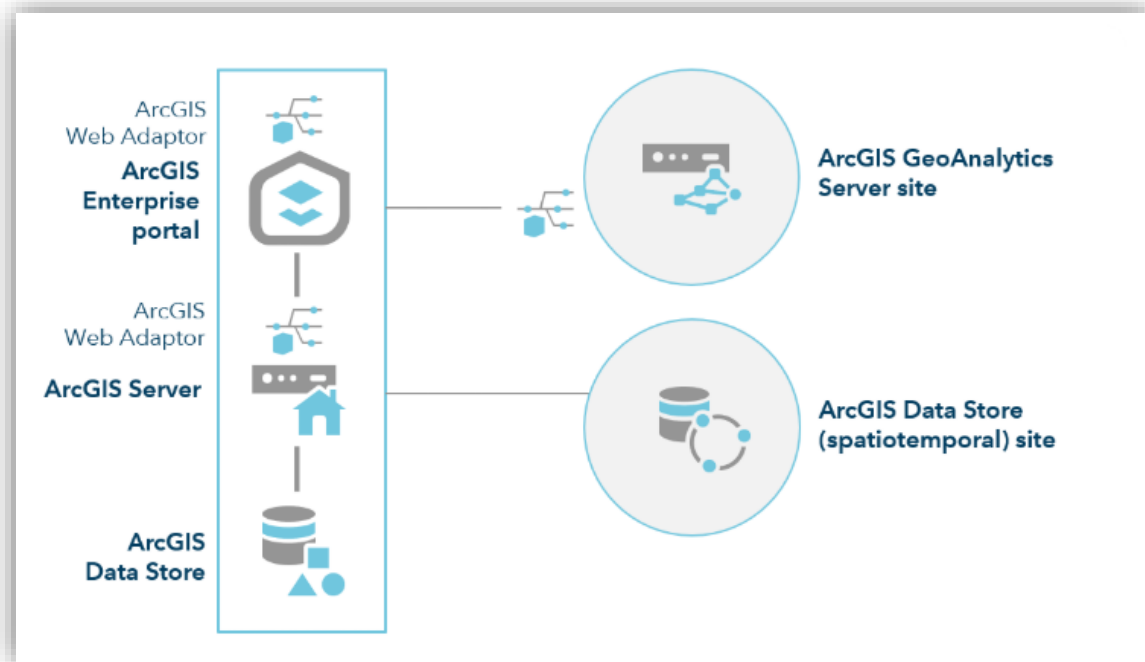


ArcGIS
Enterprise

=



GeoAnalytics Sever サイト | 1台 or 3台 のコンピュータ

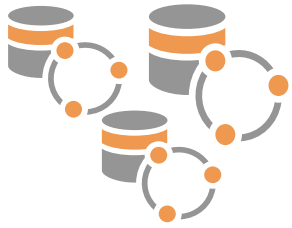


動作環境



- **ArcGIS Enterprise の基本構成**

- 最小 4 コア と 16 GB の RAM



- **ArcGIS ビッグ データ ストア Spatiotemporal Big Data Store**

- それぞれのコンピュータに 最小 16 GB の RAM (**32GB を推奨**)
- 高速なディスク / データ用の十分なディスク容量
- 少なくとも GeoAnalytics Server と同じ台数のコンピュータ



- **ArcGIS GeoAnalytics Server 1台 or 3台のコンピュータ**

- それぞれのコンピュータに 最小 4 コア と 16 GB の RAM
- 計算に十分な一時ディスクの容量

構成



- Windows と Linux の両方に対応
- オンプレミス or クラウド のインフラに配置可能
- ArcGIS Enterprise と GeoAnalytics Server の構成を次で簡略化：
 - ArcGIS Enterprise Builder – 単一コンピュータの基本構成
 - Chef Cookbooks for ArcGIS – 複数コンピュータ と 高可用性(HA)構成
 - ArcGIS Enterprise Cloud Builder – AWS と Azure

まとめ



- **Integrated:** 既存のビッグ データ ストレージ、既存のGIS データ、現在使用しているもの (Desktop or Enterprise)を併用・兼用しながら使える
- **Spatiotemporal:** ツールは空間と時間のデータ分析のために設計された
- **Accelerated:** コンピュータを並列化することによって分析処理の時間を短縮
- **Actionable:** 実用的な洞察とインテリジェンスを生み出すために大量のデータを分析することが可能。大量のデータをより明確に、より意味のある方法で可視化。