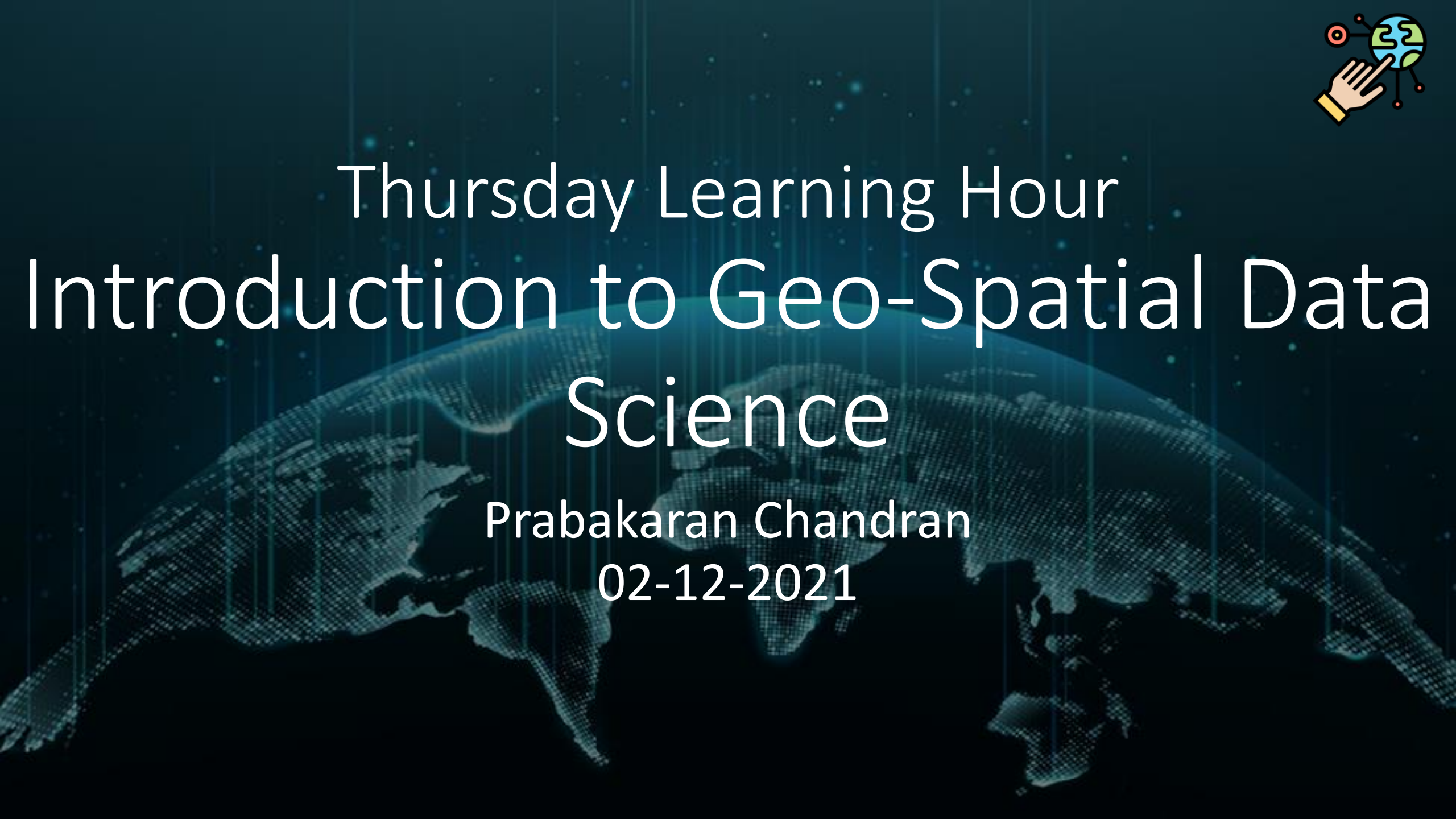


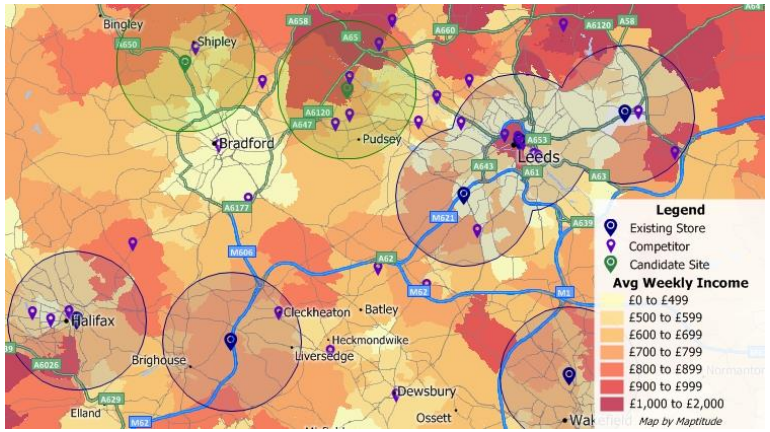
Thursday Learning Hour

Introduction to Geo-Spatial Data Science

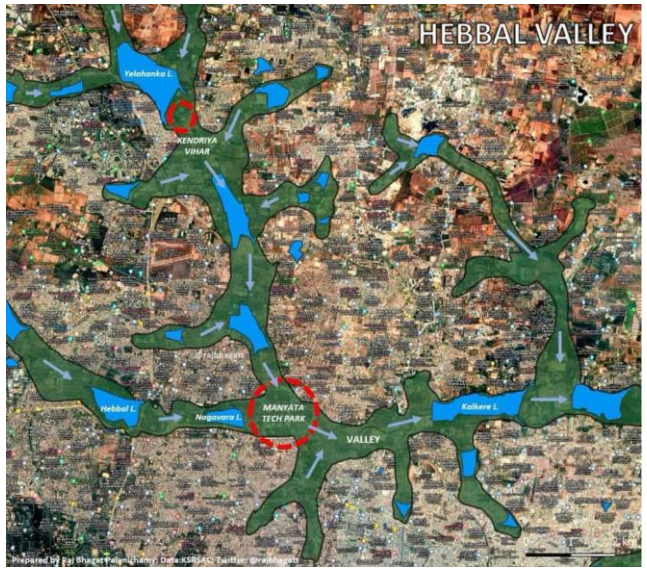
Prabakaran Chandran

02-12-2021

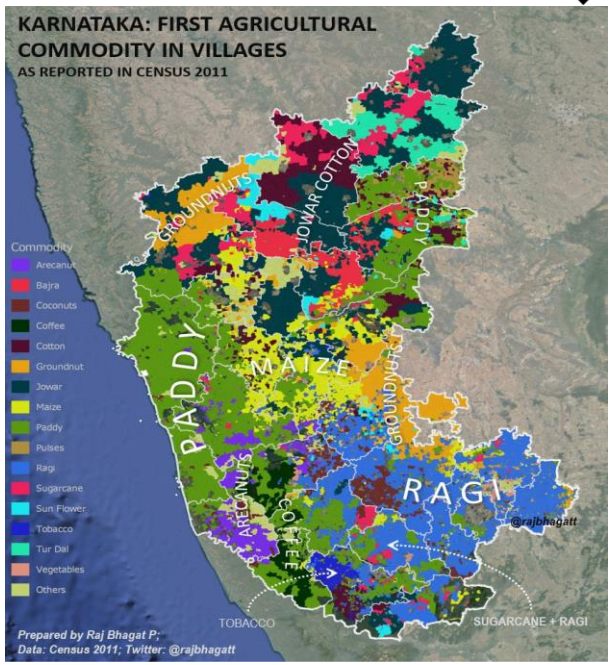
Let’s get started with Few real word Problems



1. KFC wants to setup a new store in an UK city.
2. How do they select the desired location subject to the revenue criteria?



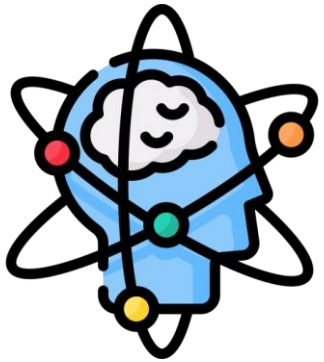
1. BBMP wants to regulate the construction activities in order to avoid manmade flood scenario in Bengaluru corporation
2. How can they Identify the Flood Prone area and reason for that?



1. Government of Karnataka wants to setup Research centers across the state to conduct researches on specific Agricultural commodities
2. How can they decide where to setup and What to Focus on?

How to solve these problems?

With the help of Science?



Concentrates more on Causal relationships to explain events or phenomena of the universe

With the help of Data Science?

Concentrates more on Correlation / pattern mining

With the help of Spatial Science?

Concentrates more on Geographical phenomena and their causal relationships

Combination of above disciplines ?

Spatial data science is similar to data science in definition. But, there is a major difference that spatial data science equally focuses on both causality and correlation.

So , What exactly is Spatial Data Science?

Data Mining and Pattern Matching Approach

Data Science

Spatial Data Science

Spatial Science

Casual Approach

There are four main disciplines of spatial data science,

- GIS,
- spatial DBMS,
- spatial big data systems
- spatial data analytics.

components of SDS

- GIS
- Spatial DBMS and Big data
- Spatial Analytics
- Topology
- Uncertainty
- Spatial Relationship
- Coordinate system

Why is spatial special?

Spatial Information

Business Perspective

Map or spatial is a platform and interface on which human,
space, advertisement, hardware, software, contents, service, and future robot (car and drones), and IOT can be integrated all together

Data Perspective

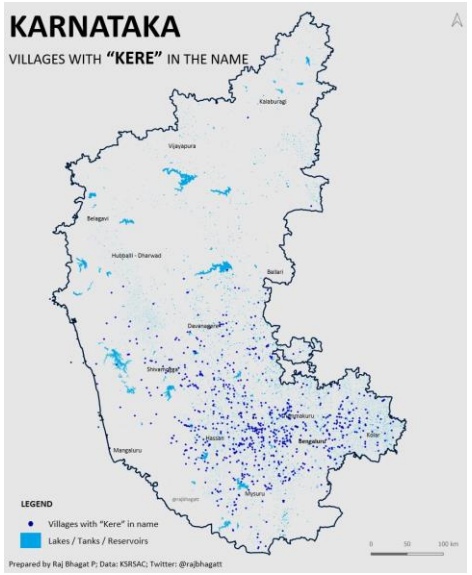
1. Uncertainty and Specific topology in Spatial Data
2. Change in perspective during different level of aggregation
3. Topology of the Data

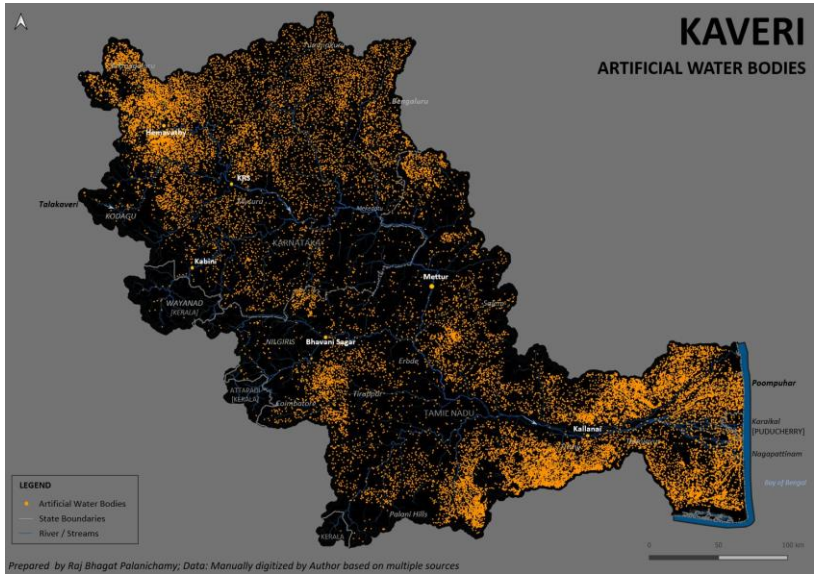
Technology Perspective

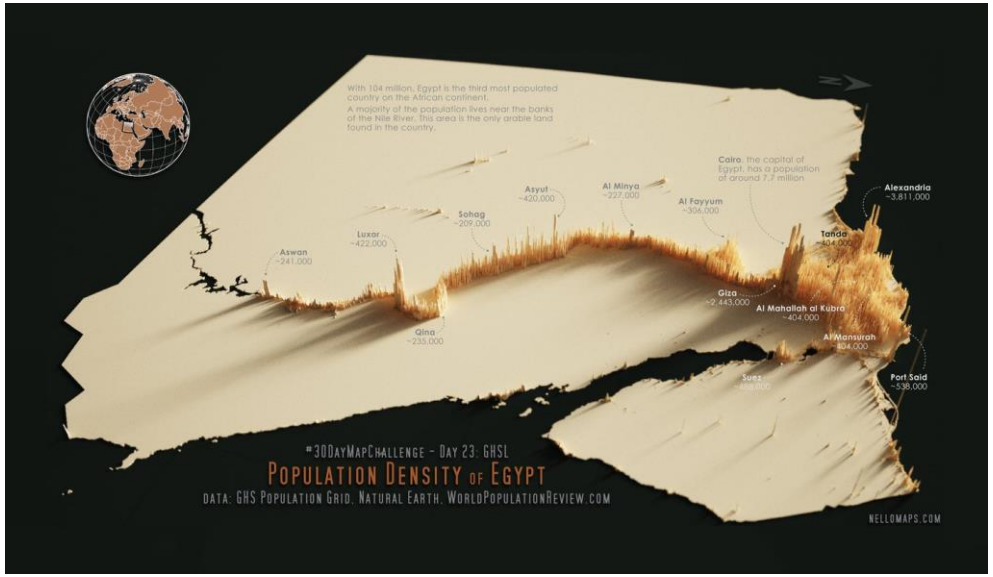
1. Spatial information cannot be directly stored as DBMS easily
2. Requires Special Data based management systems and topology
3. PostGis , GeoHadoop

Let’s Get started with GIS!

What is GIS (Geographical Information System) ?



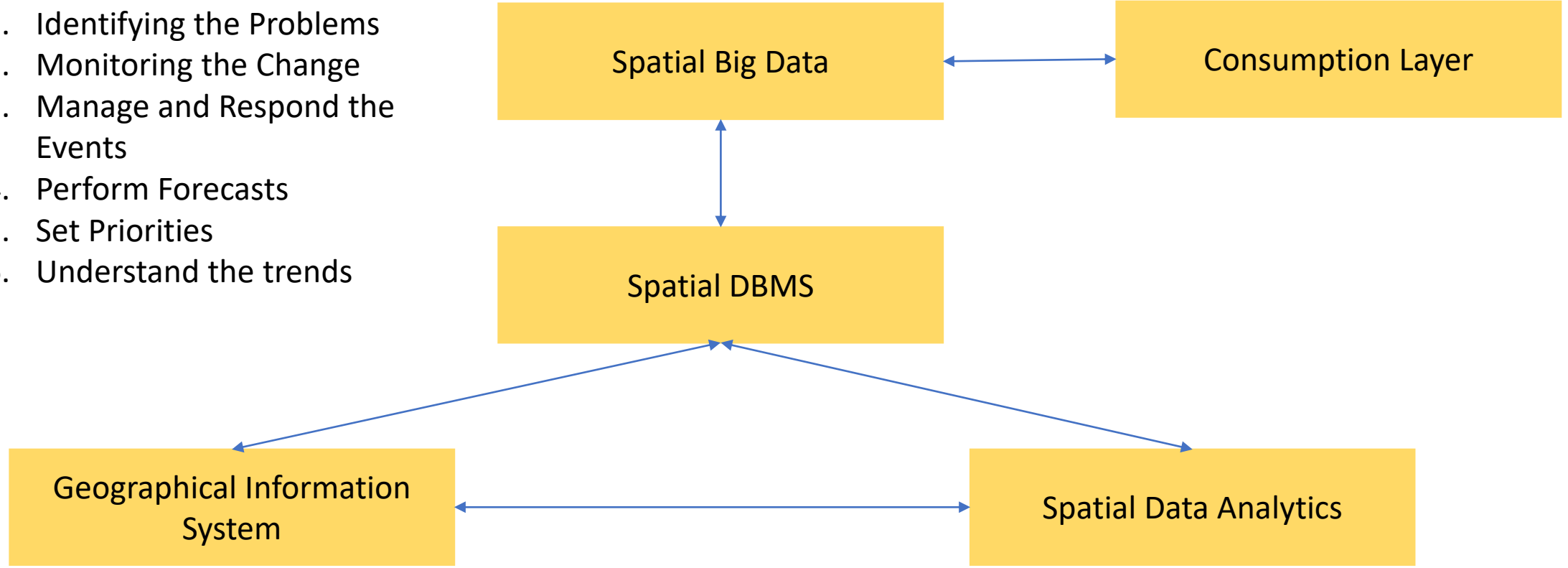




- An Information system that creates, manages, analyzes, and maps all types of data.
- GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there).
- This provides a foundation for mapping and analysis that is used in science and almost every industry.
- GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication and efficiency as well as better management and decision making.

How to solve with the help of Spatial Data Science

1. Identifying the Problems
2. Monitoring the Change
3. Manage and Respond the Events
4. Perform Forecasts
5. Set Priorities
6. Understand the trends



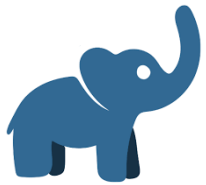
4 Disciplines of Spatial Data Science constitutes the Spatial Data Management and Analytics System which can solve the Spatial Problems.

Opensource technologies in Spatial Data Science

Cloud / On premise

Spatial Hadoop/Hive

Java / Python / Javascript



PostGIS

QGIS

R / Python

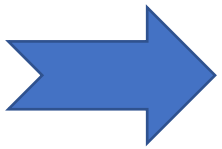
Other than these opensource technologies , there are many SAAS , PAAS Players in this field makes life easier

What is Spatial Data ? How does it become Spatial Big data

Spatial data comprise the relative geographic information about the earth and its features. A pair of latitude and longitude coordinates defines a specific location on earth. Spatial data are of two types according to the storing technique, namely, raster data and vector data.

If I want to store the Mobility information of myself , How does the data look like?

Time stamp	Location
1.30 pm	Park Square mall
2.30 pm	Aviator
3.30 pm	Spice Garden
4.30 pm	MTR

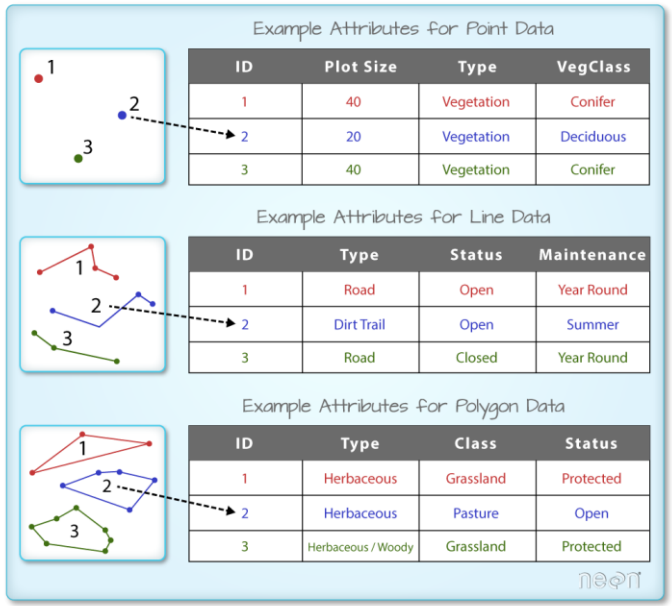


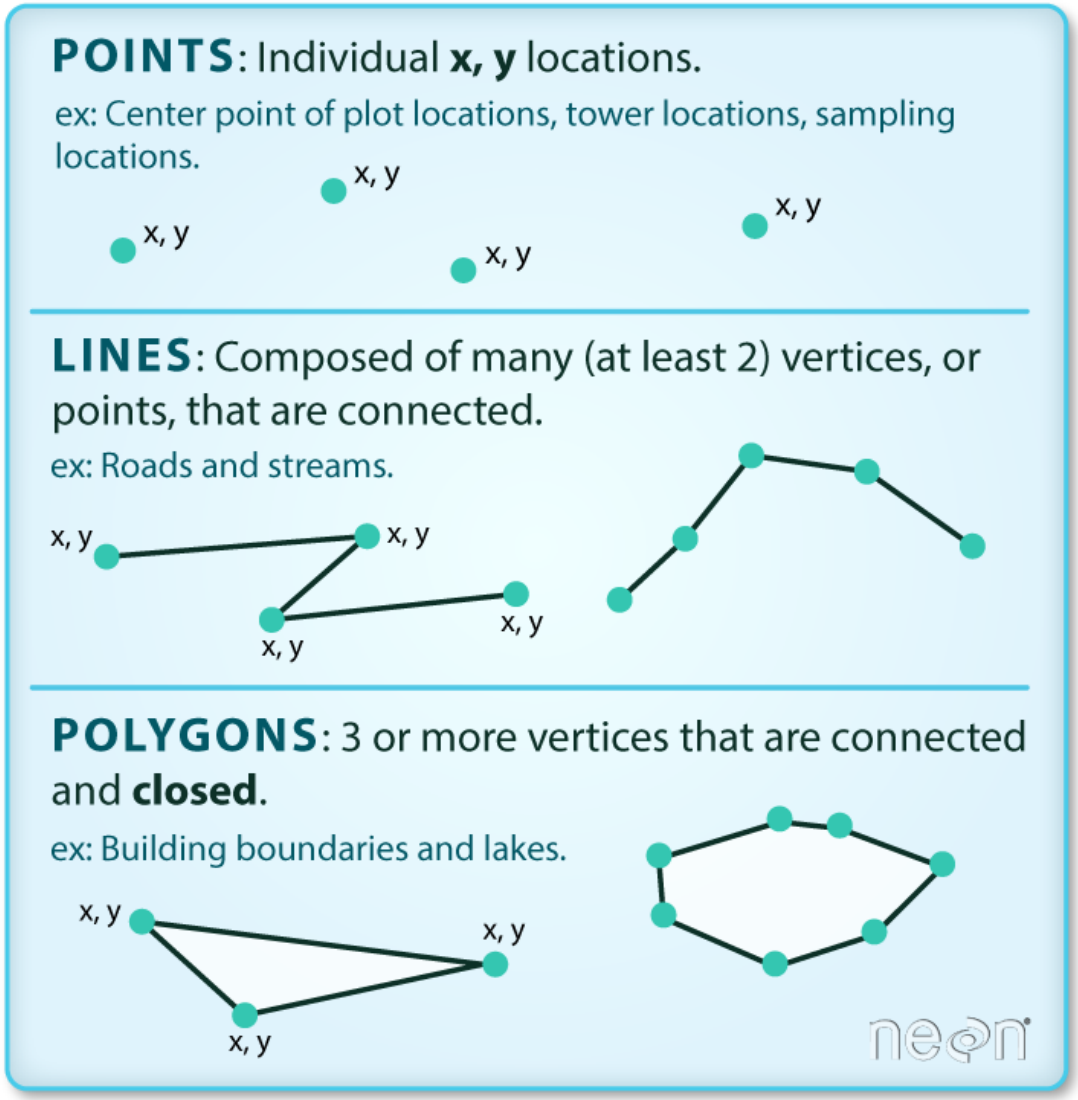
Time stamp	Location
1.30 pm	12.987332141928313, 77.7360416126701)
2.30 pm	12.957332141928313, 77.7360416126701
3.30 pm	12.887332141928313, 77.6360416126701
4.30 pm	12.787332141928313, 77.2360416126701

What is Spatial Data ? How does it become Spatial Big data

Vector Data :

Vector data structures represent specific features on the Earth's surface, and assign attributes to those features. Vectors are composed of discrete geometric locations (x, y values) known as vertices that define the shape of the spatial object. The organization of the vertices determines the type of vector that we are working with: point, line or polygon.





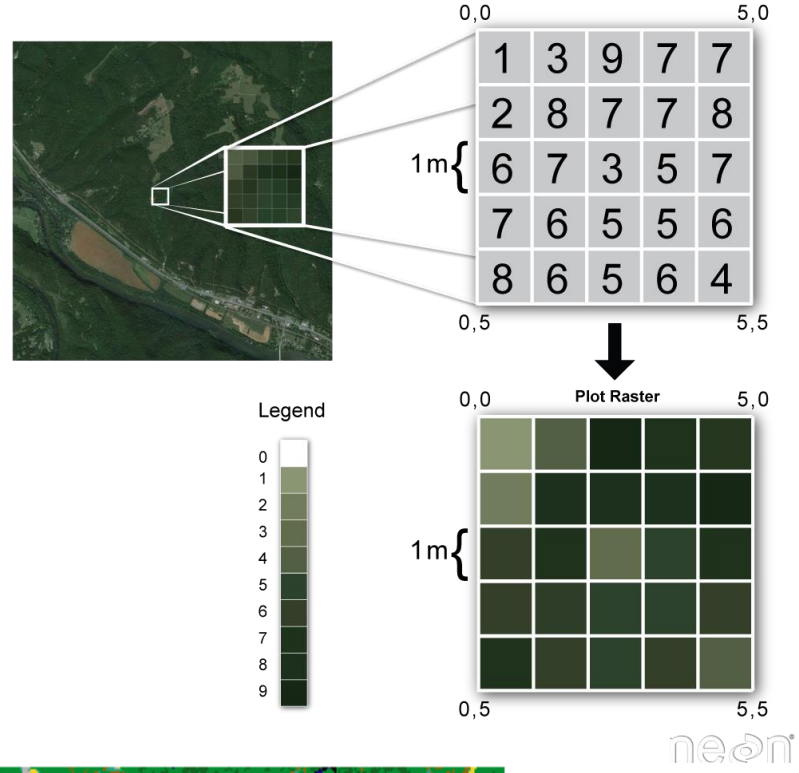
What is Spatial Data ? How does it become Spatial Big data

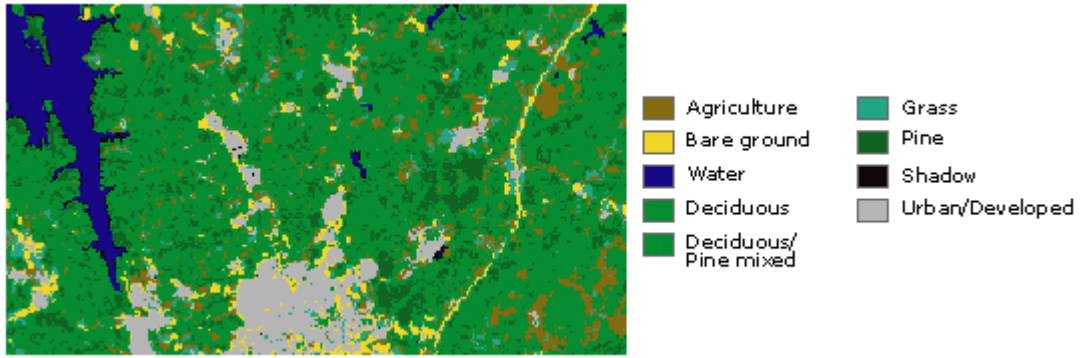
Raster Data :

Raster data is any pixelated (or gridded) data where each pixel is associated with a specific geographical location. The value of a pixel can be continuous (e.g. elevation) or categorical (e.g. land use).

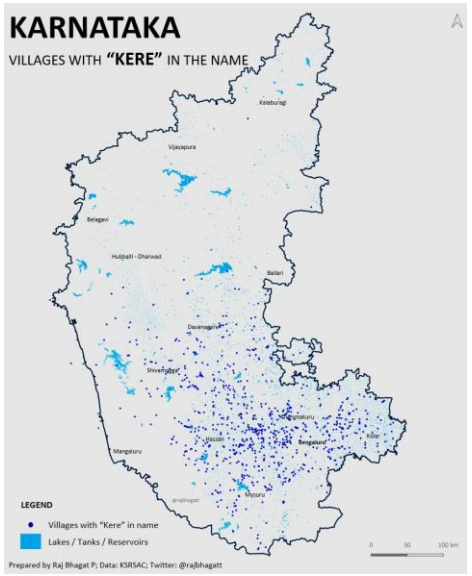
If this sounds familiar, it is because this data structure is very common: it's how we represent any digital image.

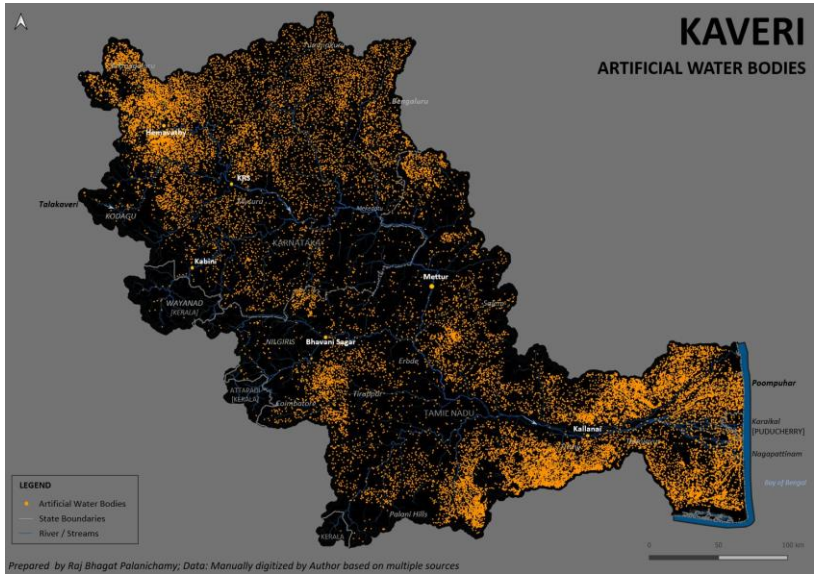
A geospatial raster is only different from a digital photo in that it is accompanied by spatial information that connects the data to a particular location. This includes the raster's extent and cell size, the number of rows and columns, and its coordinate reference system (or CRS).

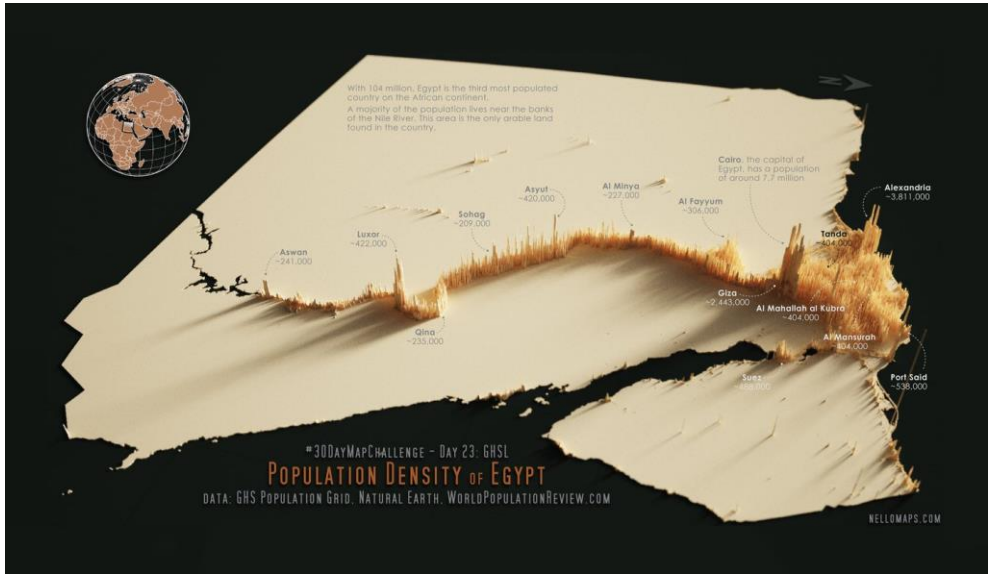




How is it collected and Stored ?



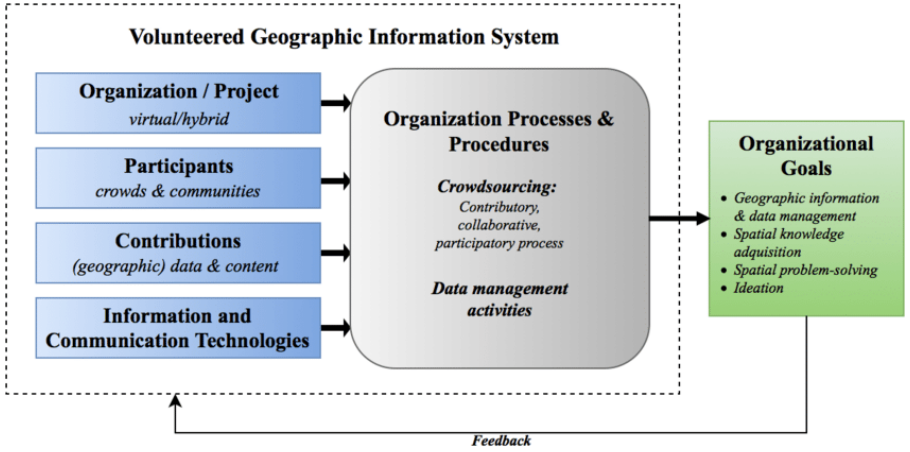




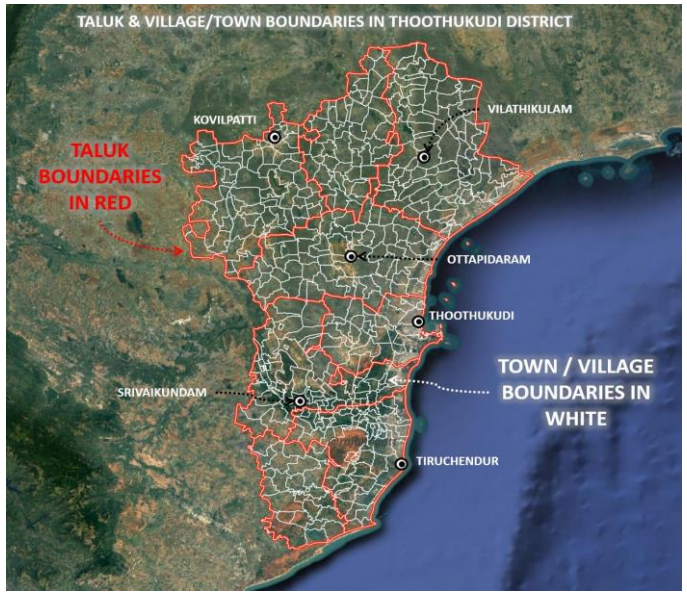
- An Information system that creates, manages, analyzes, and maps all types of data.
- GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there).
- This provides a foundation for mapping and analysis that is used in science and almost every industry.
- GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication and efficiency as well as better management and decision making.

How is it collected and Stored ?

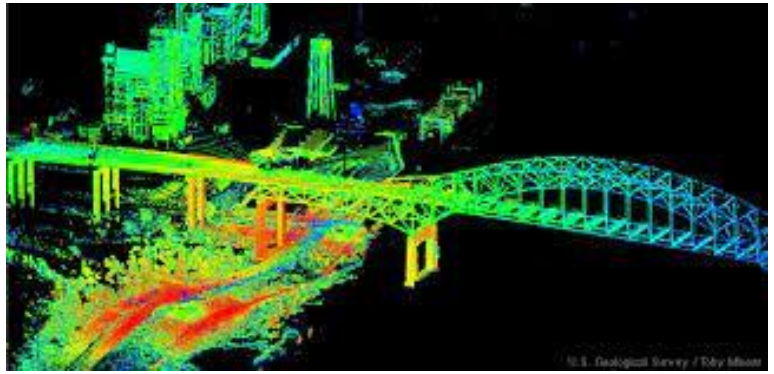
1. Surveying
2. Remote Sensing
3. Maps
4. Cartography
5. Global Positioning System (GPS) and Local Tracking systems like Beacons , Wifi
6. Volunteered Geographic Information



Map Digitization



Lidar mapping



How to get the insights from the Spatial Data?

Spatial Data → Geo Processing → Analytics and Modeling

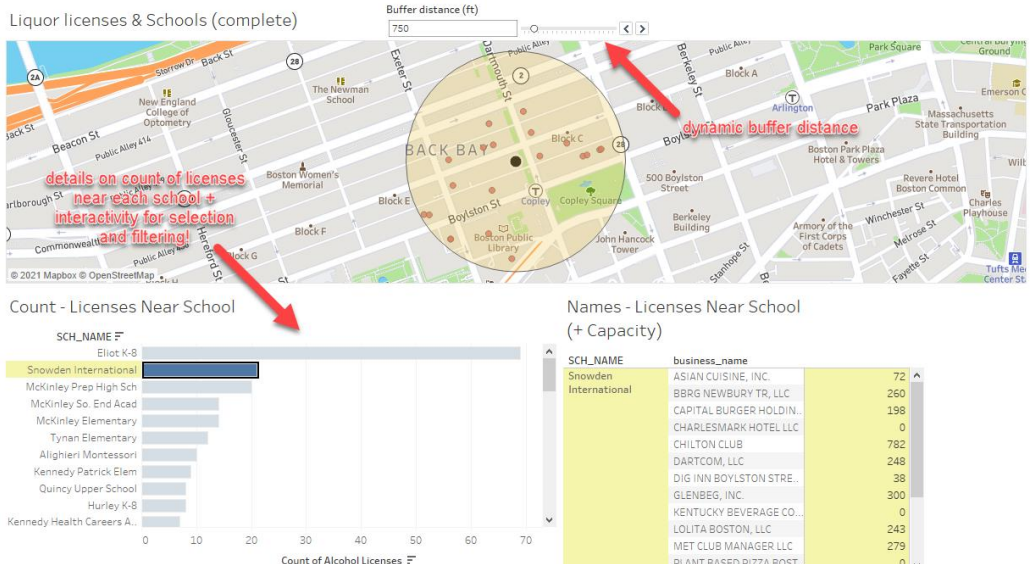
1. Spatial Joins
2. Overlaying Different Layers
3. Measurement of KPIs
4. Integrating with other attributes
5. Reference system changes
6. Extraction of Necessary Information

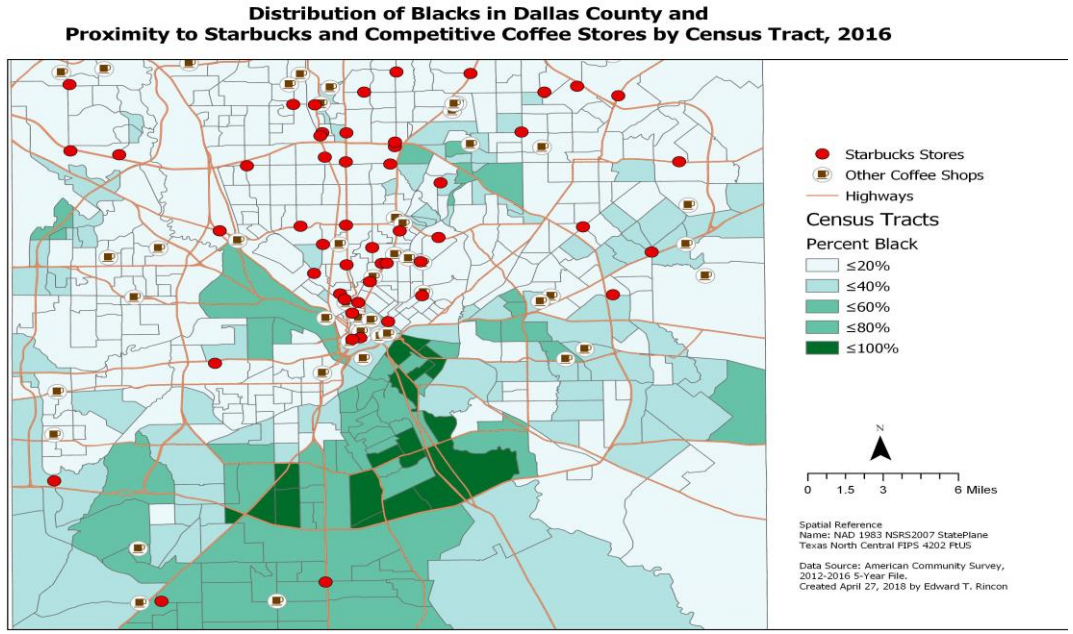
Analytics and Modeling →
- Proximity and Accessibility analysis
- Spatial Auto Correlation and Interpolation
- Spatial Categorization
- Spatial Factor Analysis
- Network Analysis

How to get the insights from the Spatial Data?

Broad Street was notorious as the centre of an 1854 outbreak of cholera. Dr John Snow traced the outbreak to a public water pump on the street, and disabled the pump. Before this time, the disease was widely thought to be caused by air-borne 'miasma'; Snow's findings showed it to be water-borne.

Proximity Analysis

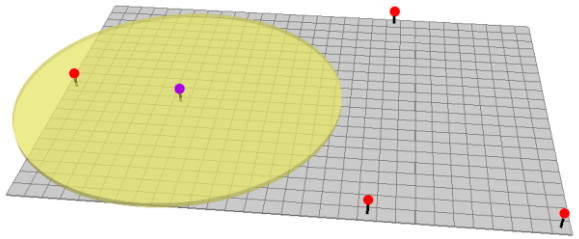




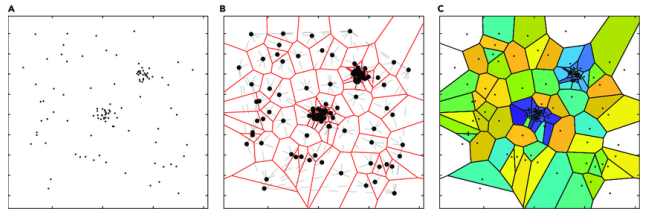
Proximity analysis [spatial analysis] — A type of analysis in which geographic features (points, lines, polygons, or raster cells) are selected based on their distance from other features or cells.

1. *Regulating Liquor Licenses around the educational Institutions*
2. *Selecting the new site for the Starbucks store*

Proximity is more than the simple linear distance.

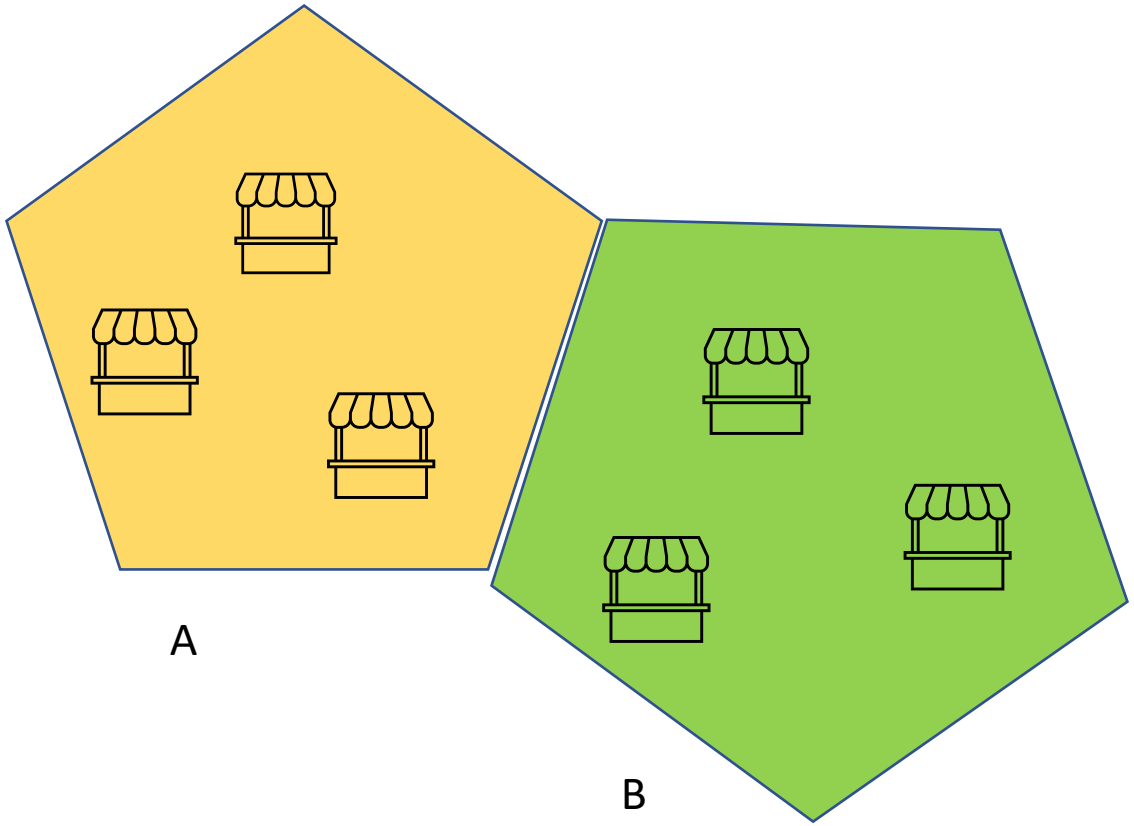


Buffers based Proximity



Voronoi based Proximity

Accessibility Analysis



We need to open a new store ,

In the both Areas , we have 3 competitors' shops , Which one to select

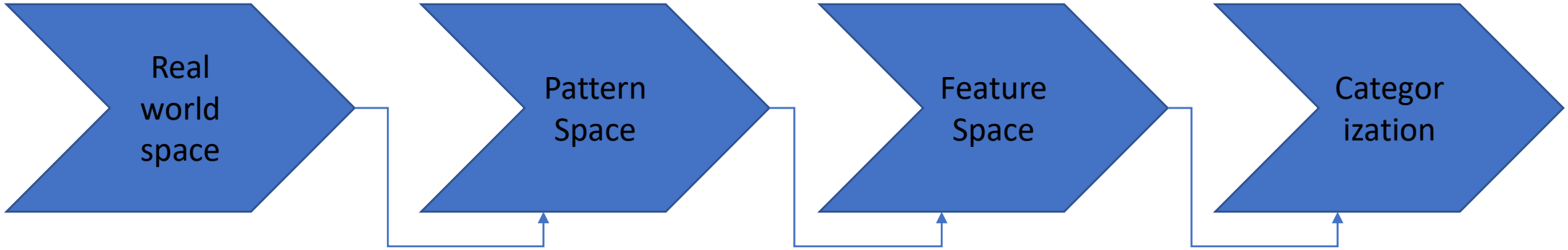
Location A : 3 stores , 30,000 Population

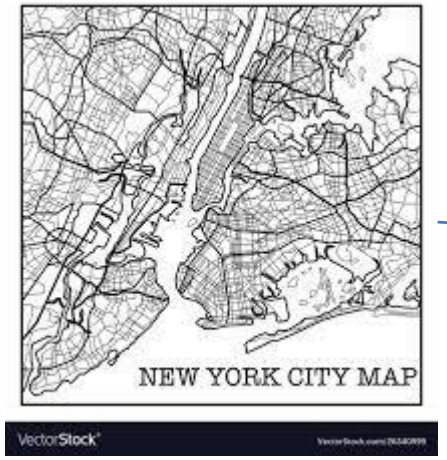
Supply / Demand = 3/30000

Location B : 3 Stores , 3000 Population

Supply / Demand = 3/3000

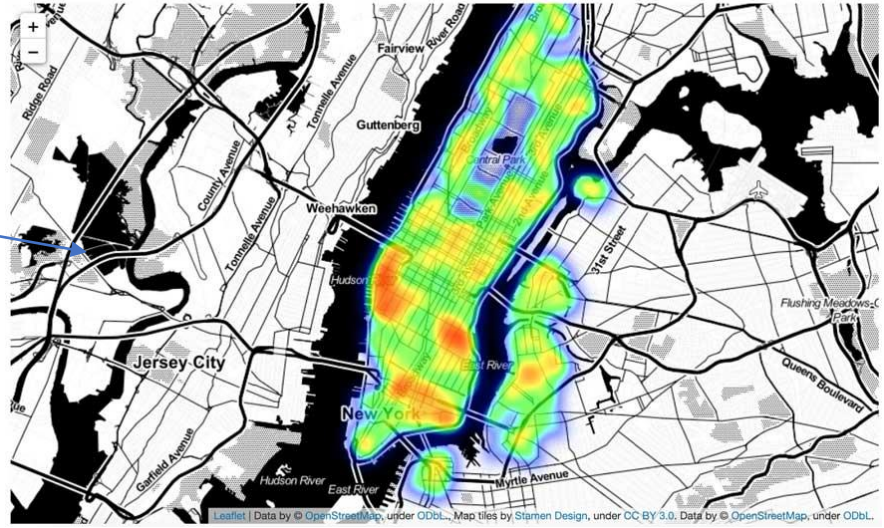
Spatial Categorization



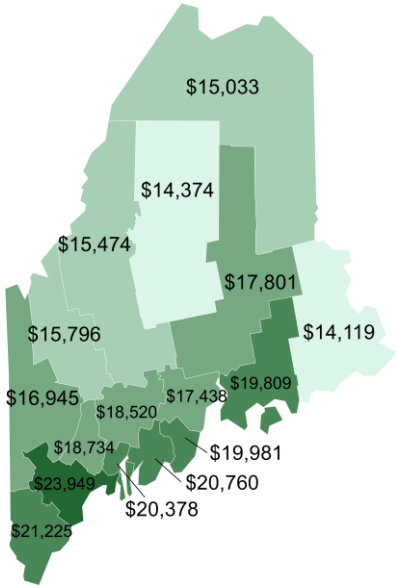


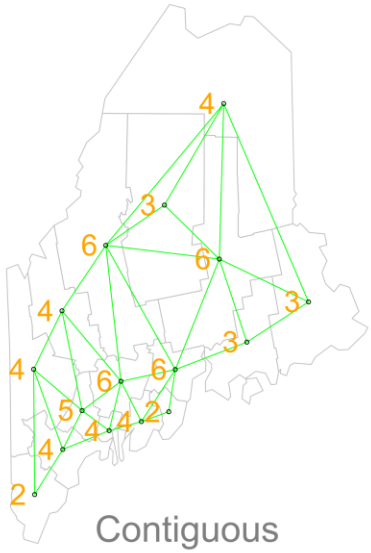
No of Pickups and Drops
Total Population
Average income
No of stores
Age wise population
Point of Interests

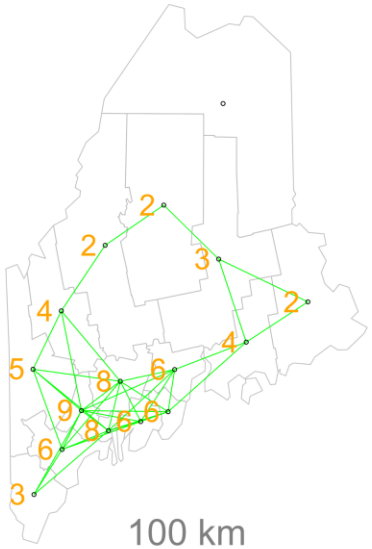
No of Pickups and Drops
No of stores
Point of Interests

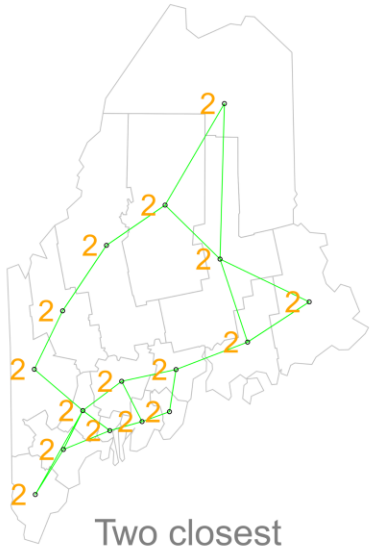


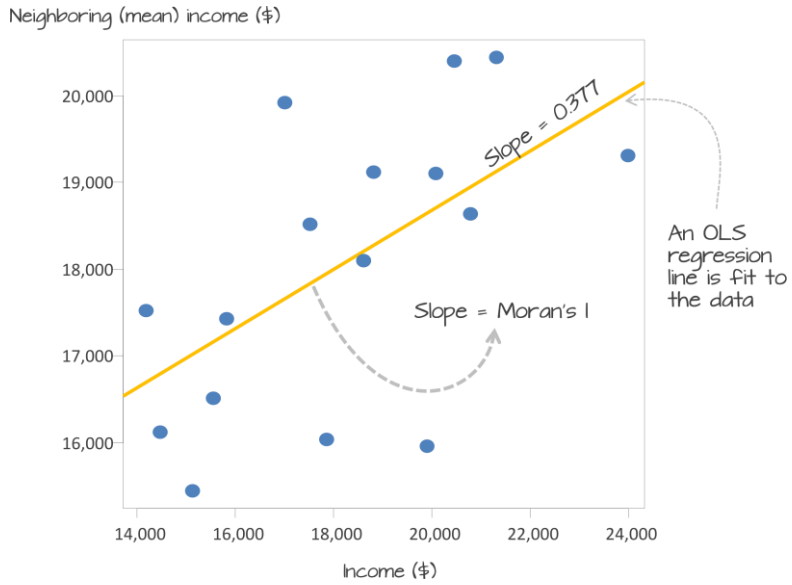
Spatial Autocorrelation :











Let's 2010 per capita income for the state of Maine. It may look like North south trend in Per capita income

We need a Statistical Evidence to measure the relationship between space and income.

Autocorrelation (whether spatial or not) is a measure of similarity (correlation) between nearby observations

Questions

Thanks for Joining

5 Layer of GIS :

Geo Viz and Information Delivery

Spatial Data Analytics

Spatial Data Acquisition

Spatial Data Model

Spatial Reference System

Offers Coordinate Reference system