

# Part II: Spatial Data Analysis

Application of ArcGIS in Public Health

---

Tesfahun T. GEREMEW , Epidemiologist and Data Analyst

Amhara Public Health Institute (APHI)

 May 20, 2024

# Contents

---

## 1 Objectives

## 2 Introduction

- Spatial Data Analysis
- Important Terminologies

## 3 Spatial Hypothesis Testing

- Spatial Statistics
- Spatial Interpolation
- Proximity Analysis

# Objectives

➔ At the end of this session, you may able to:

- Describe spatial data analysis and statistical tools
- Identify spatial statistics.
- Evaluate the spatial patterns of attributes (**spatial autocorrelation**)
- Identify local clusters of areas with **high rate** (hotspots) or **low rate** (cold spots)
- Do spatial interpolation and buffer analysis

## Spatial Data analysis:

- The process of examining the locations, attributes, and relationships of features in spatial data through overlay and other analytic techniques to address a question or gain useful knowledge.
- A set of techniques for analyzing spatial data
- Extracts or creates new information from spatial data.

# Introduction...

## Spatial analysis helps in:

- Understanding the association between prevalence of certain events/conditions and specific geographic features.
  - **Spatial Autocorrelation.**
- Detecting local clusters
  - **Hot spot analysis.**
- Mapping the **physical accessibility** or the impact of zones of locations or sites, where the controlled activity needs to be strengthened
  - **Buffer or proximity analysis.**
- Predicting the unknown values of geographic areas/locations from limited sampled point data.
  - **Spatial Interpolation**

# Important Terminologies

## Spatial clustering

- An excess of events/conditions at specific locations (e.g., local or focused clustering) or by an unusual pattern over the entire study area (e.g., global clustering).

## Hot spots:

- Locations with *many incidents*
- Statistically significant spatial clusters of **high values**.

## Cold spots:

- Locations with *very few incidents*
- Statistically significant spatial clusters of **low values**.

# Spatial Hypothesis Testing

- The goal of spatial statistics is to assess for significant spatial patterns or trends.
- There are two broad categories of tests:
  - **Tests of GLOBAL trends:** looks for evidence of spatial heterogeneity across the entire study area
  - **Tests of LOCAL trends:** to identify specific/local clusters

# Spatial Statistics

## Global Tests » Global Spatial Autocorrelation

- Measures spatial autocorrelation based on feature locations and attribute values using the Global Moran's I statistic.
  - **Spatial Patterns:**
    - NON-RANDOM (Clustered or dispersed)
    - RANDOM (No spatial autocorrelation)
  - Detect the presence or absence of clustering over the whole study region without specifying the spatial location.

## Local Tests » Cluster analysis

- A focused test which is used to detect raised incidence of events/conditions in specific areas.
- Additionally, specify the location(s) or local clusters.

## ❖ Global spatial auto-correlation

- Does NOT suggest the location of the clusters.
- Only suggest the null hypothesis of spatial randomness can be rejected.

❓ To know where these clusters are then, we need to employ another method called **Cluster detection**

## Global Measures

- Global Moran's I statistic

## Local Measures

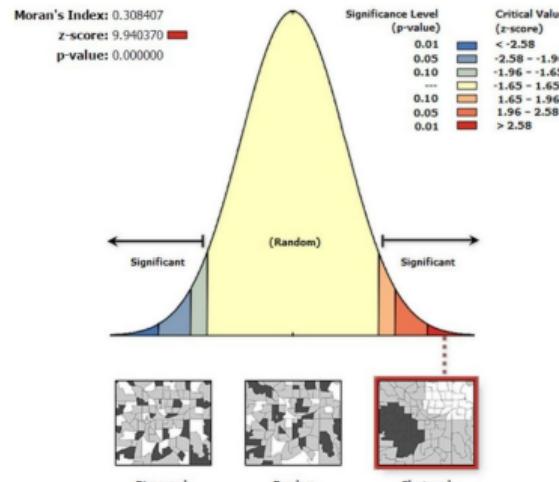
- Getis-Ord Gi\* statistic (pronounced G-i-star)
- LISA or Local Moran's I
- Kulldorff's Spatial Scan Statistic

# Common softwares for spatial data handling

- ArcGIS
- QGIS
- 
- GeoDa
- SaTScan
- Advantage of performing geospatial analysis in :
  - Statistical computing capabilities and impressive graphics engine makes it adept for spatial analysis and mapping.
  - Command-line interface (CLI) enables analysis and visualization to be customizable, transparent and **reproducible**

# Spatial statistic ...

## Global Moran's I Statistic



Given the z-score of 9.94036978752, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

## Getis-Ord Gi\* Statistic

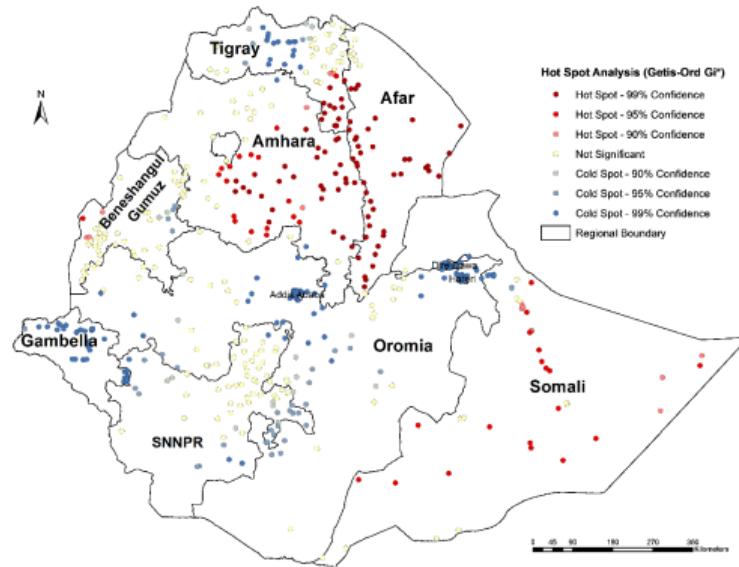


Figure: Global spatial autocorrelation in childhood MCV1 in Ethiopia, EDHS 2016.

Figure: Hot spot and cold spot areas of FGM/C among girls in Ethiopia, EDHS 2016.

# Spatial statistic ...

## Kulldorff's Spatial Scan Statistic

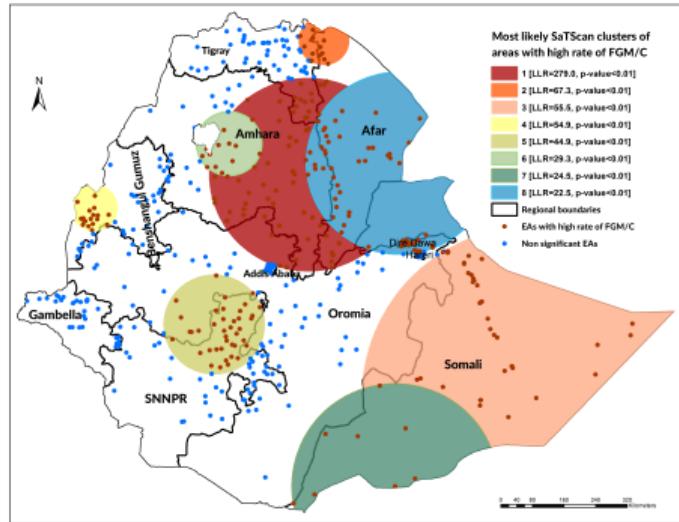


Figure: Clustering of areas with high FGM/C among girls in Ethiopia, 2016.

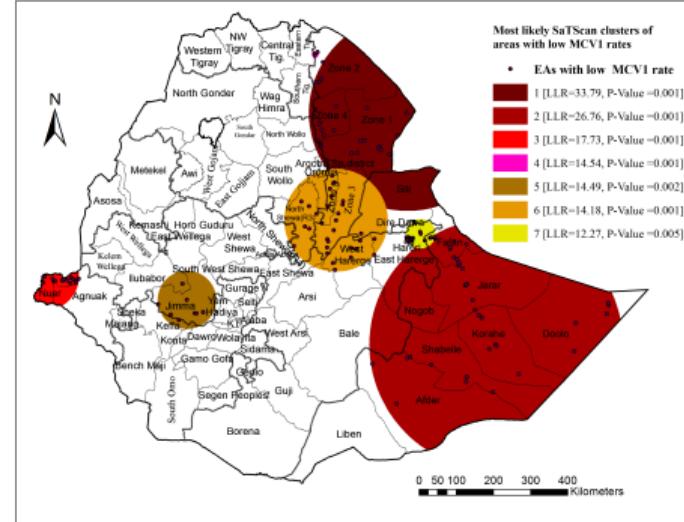


Figure: The spatial clustering of areas with low childhood MCV1 in Ethiopia, 2016.

# Spatial Interpolation

- Spatial interpolation is the procedure of estimating the value of properties at unsampled sites within the area covered by existing observations.
- Interpolation predicts the values for cells in a raster from a limited number of sample data points.
- It can be used to predict unknown values for any geographic point data, such as elevation, rainfall, risk levels, and so on.
- Almost all cases of the attributes must be **interval** or **ratio** scaled.

# Spatial statistic ...

## Interpolation using an inverse distance weighted (IDW) technique

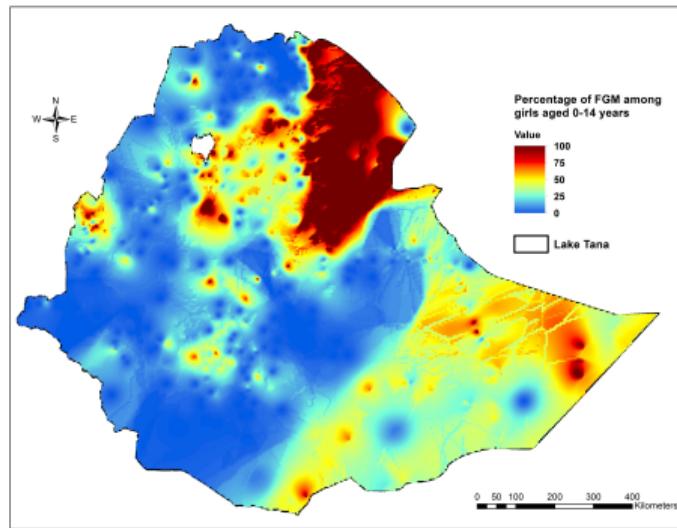


Figure: Spatial distribution of FGM/C among girls in Ethiopia, 2016.

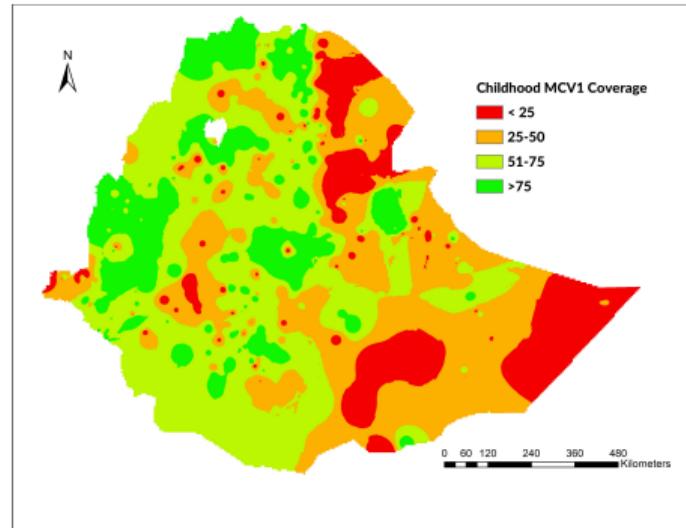


Figure: Spatial distribution of childhood MCV1 coverage in Ethiopia, 2016.

## ① What is buffer?

- ▶ A buffer is a zone that is drawn around any point, line, or polygon
  - Buffering    
  - Multiple ring buffer    
  - Create Thiessen polygon
- ▶ **Buffer analysis** is used for identifying areas surrounding geographic features.



# Practice Session



# Thank You !!!

- For More Information, Please Contact:

 Tesfahun Taddege GEREMEW<sup>id</sup>

 Amhara Public Health Institute (APHI)

 ttaddege@gmail.com

 +251-928-51-6278

 Bahir Dar, Ethiopia