SGO1910 – Introduction to Geographical Information Systems (GIS)

Written school exam

November 20 at 9:00 AM (3 hours).

The exam consists of three parts:

- Part 1 involves five questions (30 % of grade)
- Part 2 involves two questions (35 % of grade)
- Part 3 involves answering one questions (35 % of grade)

You may submit your responses in Norwegian, Swedish, Danish or English.
It is allowed to use English terms if you use one of the Scandinavian languages.

Examination support material

Dictionaries handed in before the examination.

Questions during the exam

If you have any questions during the exam, please contact the senior supervisors in the room.

After the exam

After the examination you will see your submission under **Archive**.

Good luck!

Part 1: Short Answer Questions (30% of grade)

Give a short answer to *all* five questions. You are welcome to use examples.

Q1: Describe what makes spatial data unique compared to non-spatial data. Provide one example to illustrate your point (max 150 words).

- Student clearly differentiates spatial data from non-spatial data (i.e. tables), highlighting key characteristics like:
  - location referencing (data has coordinates),
  - spatial relationships (absolute, but also relative position to other objects)
  - dimensionality (points 1d, lines 2d, polygons 3d).
- Conciseness and Relevance: The answer is focused, within the word limit, and directly addresses the question.
- A brief example is positive: Shows a deeper understanding or real-world application of spatial data's uniqueness. For instance, an object or phenomenon can be represented in a table, but we can visualise it on a map once we add coordinates. This allows new types of understanding and analysis.

Q2: Explain what vector and raster data are with an example for each, and note one key difference between them (max 150 words).

- Definition: Accurate and precise definitions of vector and raster data.

- Key Difference: Clearly explains a fundamental difference, such as data representation (discrete for vector, continuous for raster) or scalability.
- Examples: Relevant examples provided for each (e.g., vector data like roads, raster data like satellite imagery).
- Conciseness and Clarity: The answer is concise, clear, and within the word limit.

Q3: Summarize the concept of the Modifiable Areal Unit Problem (MAUP) in GIS (max 150 words).

- Accurately and concisely summarizes the MAUP:
  - i.e. “statistical results can significantly vary based on the scale and boundaries of the spatial units used in analysis. This issue arises because the way we divide areas into units (like districts or zones) can affect the outcomes of spatial analysis.”
  - Scaling and zoning
  - Impact on spatial analysis (changing scale or zone changes the results)
- Examples: Illustrates with an example or explains the significance in data analysis.

Q4: Outline Tobler’s first law of geography and its significance in spatial analysis (max 150 words).

- Definition: i.e. Tobler’s first law of geography states that 'everything is related to everything else, but near things are more related than distant things.' This law is significant in spatial analysis as it underpins many spatial phenomena and GIS techniques.
- Examples are virtually everywhere in social sciences. Students should provide an example. Talked about many things in lectures: income, language, voting behavior.
- Can also mention relevance to spatial autocorrelation, often equated with positive spatial autocorrelation.

Q5: Briefly explain the necessity of coordinate systems and map projections in GIS (max 150 words).

- Coordinate systems provide a framework for locating features on the Earth's surface.
- Map projections, on the other hand, are methods for transferring the Earth's spherical surface onto a flat map.
- Importance of having data in the same coordinate system.
- Appropriate map projection is necessary for correct representation and accuracy.
- Example of a map projection.

Part 2: Applied Questions (35% of grade)

Answer *both* questions.

Q1: Identify and describe common sources of spatial data in raster and vector formats, with one example for each. Discuss one advantage and one limitation of each format in GIS (max 300 words).

- Student should accurately identify common sources of spatial data for both raster and vector formats.
- I.e.: Raster Data Sources: Satellite imagery (e.g., Landsat, Sentinel), aerial photography, digital elevation models (DEM). Example: Satellite imagery used for land cover classification.

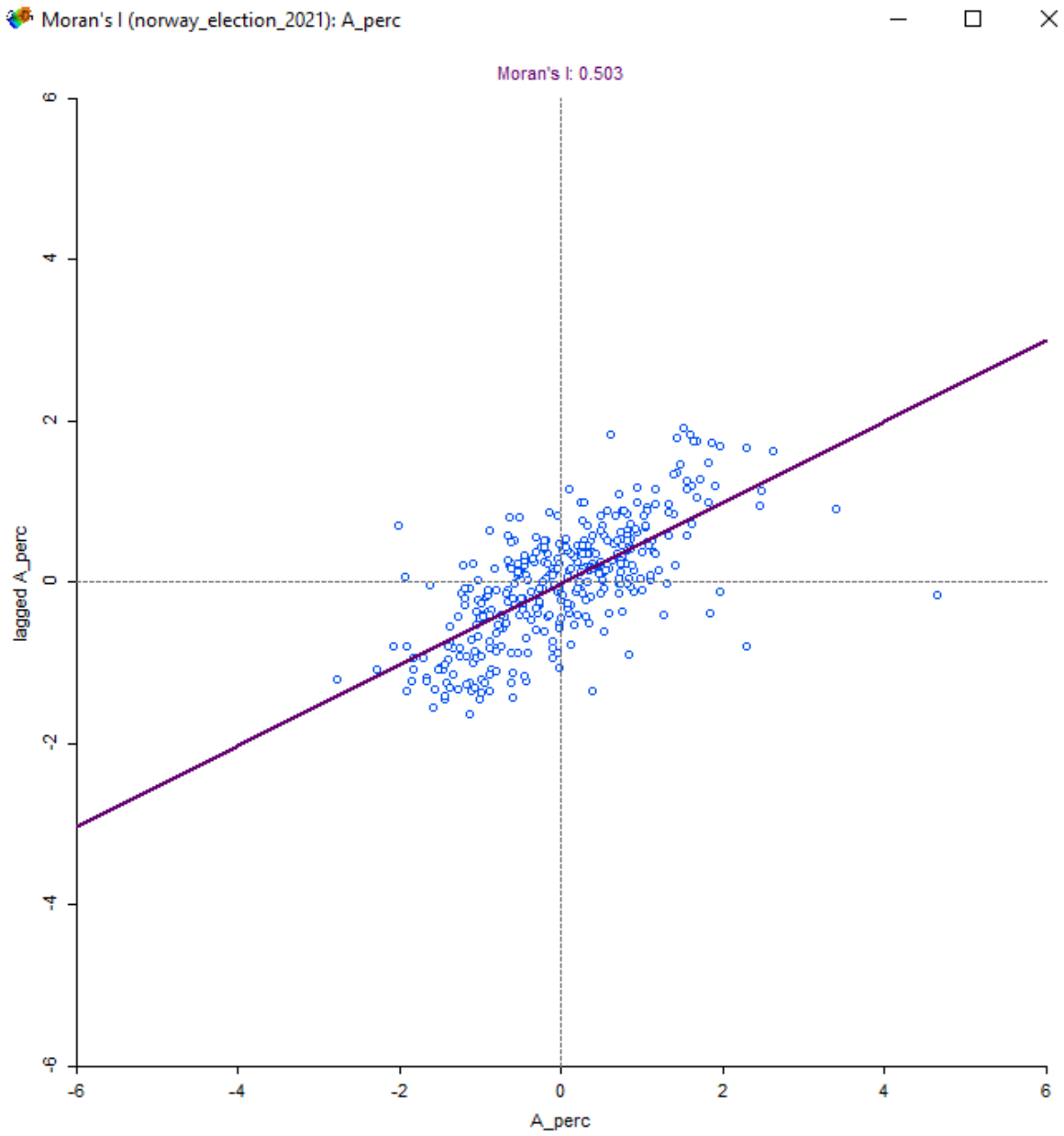
- Vector Data Sources: Cadastral maps, road networks from government databases, OpenStreetMap data. GPS. Example: Road network data used for urban planning or traffic network analysis.
- Advantages and Limitations:
  - Raster Data:
  - Advantage: Suitable for representing continuous data, such as elevation or temperature. Representations well suited for satellite imagery analysis. Ideal for spatial analysis of continuous phenomena. Resolutions and pixel size can be adjusted.
  - Limitation: File size with large coverage and high resolution. Not precise representations of discrete objects with well defined boundaries, and resolution fixed. Generalization of objects or phenomena. For instance what is in between land type areas, transitions is often lost (i.e. between agricultural and forest areas).
  - Vector Data:
  - Advantage: Precise in representing discrete features like boundaries and networks. Efficient storage. Can store attributes.
  - Limitation: Can be complex to manage with many attributes and relationships, and less effective for representing continuous phenomena.

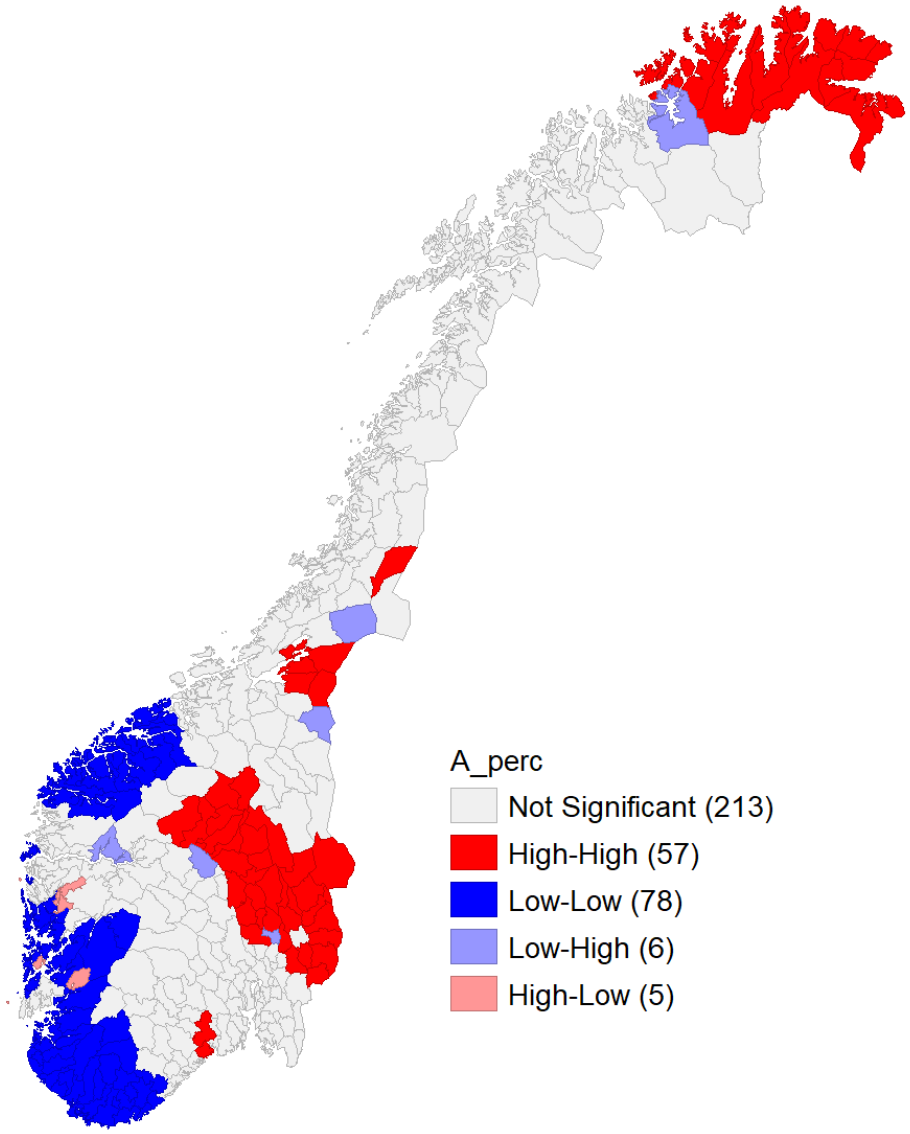
Q2: Discuss the potential sources of error in spatial data and explain how uncertainty can affect different stages of GIS analysis (max 200 words).

- Error in spatial data is inherent, and is introduced at all stages in a GIS process.
- Potential Sources of Error:
  - Measurement errors (e.g., GPS inaccuracies, sensor errors).
  - Data processing errors (e.g., incorrect data input, software bugs).
  - Projection and scale issues (e.g., distortions in map projections, scale-dependency of spatial analysis).
  - Human factors (e.g., subjective interpretation in manual data classification).
- Impact on GIS Analysis:
  - Data Collection Stage: Inaccurate GPS measurements can lead to errors in the location of features.
  - Data Processing Stage: Software bugs or incorrect data entry can result in incorrect calculations or misinterpretation.
  - Analysis Stage: Projection distortions can affect spatial analysis, like distance calculations or area estimations.
  - Interpretation Stage: Errors from earlier stages can compound, leading to misleading conclusions, such as overestimating the extent of an environmental hazard.

Part 3: Essay Question (35% of grade)

Q1: The figures provided display results from two types of spatial autocorrelation analyses of the 2021 election results for the Labour Party in Norway. Describe what each figure represents, and interpret the outcomes shown. Discuss the implications of these results for understanding spatial patterns in the data (max 500 words).


Moran's I (norway_election_2021): A_perc
Moran's I: 0.503
lagged A_perc
A_perc
6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6



- Both figures are outputs from GeoDa and two types of spatial autocorrelation analysis of election results for Arbeiderpartiet in 2021: global morans I and local morans I.
- They can use this opportunity to show their understanding of spatial autocorrelation, which has been covered extensively in the course, using GeoDa in seminars.
- The first figure is a scatterplot commonly used in Moran's I analysis, which measures spatial autocorrelation and indicates whether the pattern expressed is clustered, dispersed, or random. The student should interpret the scatterplot's Moran's I value and the slope of the fitted line to assess the level and direction of spatial autocorrelation.
    - The X axis shows the percentage votes for the Labour party in each unit, and the Y axis shows the average percentage votes for Labour Party in neighbouring units.
    - The points represent the distinct observations.
    - Line of best fit illustrates the relationship between the value of an observation and the average value of its neighbours (spatial lag).

- The second figure is a Local Indicator of Spatial Association (LISA) cluster map, which identifies clusters of similar values (high values near high, low values near low) and outliers (high values near low and vice versa).
- The student should interpret the LISA map clusters (high-high, low-low) and outliers, explaining what the colours represent in terms of the spatial distribution of the data.
- Discusses how the global Moran's I result informs us about the overall spatial distribution of the data (e.g., clustered, dispersed, or random).
- Elaborates on what the LISA results tell us about local patterns, such as areas of significant high-value or low-value clusters.
- Discuss and show understanding of spatial patterns and their possible causes or consequences in relation to the variable: share of labour party votes. What does it tell us about the distribution of votes for the labour party?