

GRADING GUIDELINE!

SGO1910 – Introduction to Geographical Information Systems (GIS)

Written school exam 2024

Exam date: November 18 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)

Provide concise answers to all five questions. Use examples where relevant.

1. Describe the Modifiable Areal Unit Problem (MAUP) in GIS and explain why it can affect the results of spatial analysis. (max 150 words)

Guideline: Student provides a clear and accurate description of MAUP, recognizing it as the issue where "statistical results can vary significantly depending on the scale (size of spatial units) and zoning (arrangement of boundaries) in spatial data."

Response should be concise (max 150 words) but cover the essential impact MAUP has on spatial analysis. Student effectively (taking into account that this is BA level) explains how both scaling and zoning influence spatial analysis, showing an understanding that different scales (e.g., neighborhood vs. regional) or zone arrangements (e.g., grid-based vs. administrative boundaries) can lead to different analytical results.

Student discusses the significance of MAUP in data analysis, noting how it can lead to varying conclusions depending on the chosen scale and zoning. An example or brief

application to a real-world context (e.g., how MAUP might influence crime rate analysis by neighborhood or election outcomes by voting district) strengthens their response.

2. Explain the purpose of coordinate systems and map projections in GIS. Describe one example of a map projection and how it impacts spatial accuracy. (max 150 words)

Student provides a clear, concise explanation of the purpose of coordinate systems and map projections in GIS, noting that they allow us to represent the Earth's curved surface on flat maps and locate features accurately.

Explanation should highlight the need for coordinate systems to standardize locations and for map projections to convert spherical data to a two-dimensional format, acknowledging that projections manage distortions but introduce trade-offs in accuracy.

Student describes a specific map projection (e.g., Mercator, Albers, or Robinson) and discusses how it affects spatial accuracy in one or more aspects (area, shape, distance, or direction). For example, they may explain that the Mercator projection preserves direction, making it useful for navigation, but distorts area, particularly near the poles, making regions like Greenland appear much larger than they are.

3. Differentiate between raster and vector data in GIS. Provide one example of each and discuss one key advantage of each format. (max 150 words)

Student provides a clear, concise differentiation between raster and vector data, noting that raster data represents spatial information in a grid of pixels (ideal for continuous data like elevation or temperature), while vector data represents features with discrete points, lines, and polygons (suitable for precise boundaries like roads or land parcels).

Student provides one relevant example of each data format, such as satellite imagery or digital elevation models for raster data, and city boundaries or road networks for vector data.

For raster data, an example advantage might be its suitability for representing continuous surfaces and performing spatial analyses like overlaying, calculating gradients or representing variation within administrative areas, where vector data represents homogeneous values (i.e. one value per polygon). For vector data, an advantage might be its precision in representing discrete features and efficiency in storing data with clear boundaries.

4. Summarize Tobler's first law of geography and explain its relevance in spatial analysis. (max 150 words)

Student provides an accurate summary of Tobler's First Law, noting that it states: "Everything is related to everything else, but near things are more related than distant things." Explanation should capture the essence of spatial dependency, where proximity often implies stronger similarity or correlation between spatial entities.

Student discusses the importance of Tobler's First Law in spatial analysis, explaining how it underpins many GIS techniques, such as spatial interpolation, clustering, and predictive modeling. The answer could also highlight that understanding spatial relationships (especially proximity) helps analysts make more accurate predictions and better understand patterns in geographic data.

5. Describe the types of distortions that can occur with different map projections. (max 150 words)

Student accurately identifies the four main types of distortions that can occur with map projections: area, shape, distance, and direction.

Explanation should briefly define each type of distortion, such as area distortion (where some regions appear larger or smaller than in reality) and shape distortion (where the actual shape of features may be altered).

Bonus if the student provides a concise explanation of why these distortions occur, noting that projecting a spherical surface onto a flat map requires trade-offs, leading to unavoidable distortions in some properties. Student briefly mentions the importance of selecting the right projection based on the purpose of the map, as some projections are better suited for certain tasks (e.g., preserving area in thematic maps or preserving shape in navigation maps).

Part 2: Applied Questions (35% of grade)

Answer both questions.

1. **When working with continuous data in GIS we usually apply classification methods for visualisation in maps (max 400 words)**

Here are some of the main classification methods. Please explain what each of them do, and how they differ from each other:

1. Natural breaks (Jenks).
2. Equal interval
3. Quantiles

Students can get 10 points on this assignment. Describing each classification methods gives 2 points each, and 4 points if they can discuss the comparison. Description and discussion at the basic level is sufficient given that this is BA course.

Explanation of Natural Breaks (Jenks) (2 points):

- Student provides a clear description of the **Natural Breaks (Jenks)** method, explaining that it divides data based on natural groupings inherent in the data distribution.
- Response should mention that the Jenks algorithm minimizes variance within classes and maximizes variance between classes, making it ideal for datasets with distinct clusters.
- Explanation should include that this method can provide more visually meaningful categories, especially for data with unevenly distributed values.

Explanation of Equal Interval (2 points):

- Student clearly explains the **Equal Interval** method, noting that it divides the range of data values into equally sized intervals, regardless of the data distribution.

- They should mention that each class represents an equal range of values, making it suitable for comparing data against uniform thresholds.
- Explanation should also include that this method may not highlight variability well if data is highly skewed or clustered.

Explanation of Quantiles (2 points):

- Student provides a concise explanation of the **Quantiles** method, stating that it divides data into equal-sized groups based on the number of data points in each class.
- Explanation should include that each class contains approximately the same number of observations, making it useful for ensuring balanced class sizes in visual representations.
- They should note that quantiles may place very different values in the same class if data is skewed, which can obscure differences in data distribution.

Comparison of Methods and When to Use Each (4 points):

- Student effectively compares the methods, highlighting how **Natural Breaks** best represents data clusters, **Equal Interval** is ideal when comparing data to uniform standards, and **Quantiles** is useful for balancing class sizes.
- Discussion should emphasize that the choice of method depends on the data distribution and the purpose of the map: Natural Breaks for highlighting patterns, Equal Interval for showing equal ranges, and Quantiles for evenly distributed observations.

2. Global and Local (LISA) Spatial Autocorrelation in GIS (max 400 words):

Define spatial autocorrelation (global and local) and how we can use it to examine spatial features or phenomena. Use real-world examples where appropriate to illustrate your points.

Definition of Spatial Autocorrelation:

- Student provides a clear, concise definition of spatial autocorrelation, noting that it measures the degree to which a spatial feature and its surrounding areas are similar or dissimilar.
- The response should explain that spatial autocorrelation helps assess the spatial patterns within data, determining whether patterns are clustered, random, or dispersed.

Explanation of Global Spatial Autocorrelation:

- Student describes Global Spatial Autocorrelation as an overall measure of spatial dependence across the entire study area, often using statistics like Moran's I.
- They should mention that Moran's I provides a single value indicating whether a dataset has clustering (positive autocorrelation), dispersion (negative autocorrelation), or randomness.

- Explanation should include that global autocorrelation is useful for understanding broad spatial patterns but may overlook localized variations.

Explanation of Local Spatial Autocorrelation (LISA):

- Student defines Local Spatial Autocorrelation (LISA) as a measure that identifies clustering or dispersion at a localized level, detecting "hot spots" (high-value clusters) and "cold spots" (low-value clusters).
- They should explain that LISA values allow for examination of spatial autocorrelation at individual locations, which is useful for identifying localized patterns that may not be evident in global measures.
- Bonus if the student include an example of a LISA statistic, such as the Getis-Ord G_i^* statistic, which helps identify statistically significant clusters.

Bonus if the student provides a Real-World Examples for Illustration:

- Student provides relevant real-world examples to illustrate both global and local spatial autocorrelation.
- For global autocorrelation, they might mention an analysis of global temperature patterns, where a high Moran's I value would suggest similar temperatures cluster together.
- For local autocorrelation, they could refer to urban crime analysis, where LISA might reveal high-crime clusters (hot spots) and low-crime areas (cold spots) within a city, providing actionable insights for resource allocation in policing.

Part 3: Essay Question (35% of grade)

Interpret the maps in Figure 1 and Figure 2 on the next page. They show the popular vote in the 2020 election in the USA, in which Joe Biden and the Democratic Party won over the Republican Party and Donald Trump.

1. Use your knowledge of GIS, spatial analysis, and cartography. Explain how the maps follow guidelines of good mapping and what can be improved, or are anything missing that should have been included? (max 400 words).

Student discusses key cartographic principles applied in the maps, such as color use (red for Trump, blue for Biden), lack of legend, and title, in map 1. Maps are also lacking scale bar and north arrow. Map 2 is missing Alaska.

Student addresses strengths of the county-level map, such as providing a detailed spatial resolution of voting patterns and showing local voting outcomes.

They should also discuss limitations, such as misleading perceptions due to large areas with low populations (e.g., rural areas), which might exaggerate support for Trump in terms of land area rather than population.

Student explains how the cartogram addresses the population disparity by resizing counties based on population, giving a more accurate picture of voter distribution.

They should discuss how this map prioritizes population over geographic area, thus conveying a more realistic sense of each candidate's support based on population density.

Student suggests improvements, such as adding population data, urban/rural markers, or more detailed legends.

2. Why do the two maps look different? (max 400 words).

Figure 1 is a traditional choropleth map that displays counties won by each candidate in red or blue, with the color intensity indicating the degree of support. This map emphasizes the geographic area of each county, showing the spatial distribution of support across the United States. As a result, large, sparsely populated rural counties, which tend to lean Republican, occupy significant portions of the map, visually amplifying Trump's support over large regions. This leads to the perception of extensive Republican support due to the dominance of red-colored areas, even though these areas may have fewer people.

In contrast, Figure 2 is a cartogram, where counties are resized according to their population rather than their geographic size. This type of map prioritizes population density, making highly populated urban areas appear larger. Consequently, areas with large populations, where Biden generally performed well, become more prominent on the map. By focusing on population rather than land area, the cartogram more accurately reflects where votes were concentrated and provides a better sense of the overall voter distribution for each candidate.

These differences illustrate the challenge in mapping election data: geographic-based maps can skew perceptions toward land area, while population-based maps offer a more realistic view of voter distribution. Together, these maps provide a fuller understanding of the spatial dynamics in U.S. elections, with one map showing the physical spread of support and the other representing the actual distribution of voters across the country.

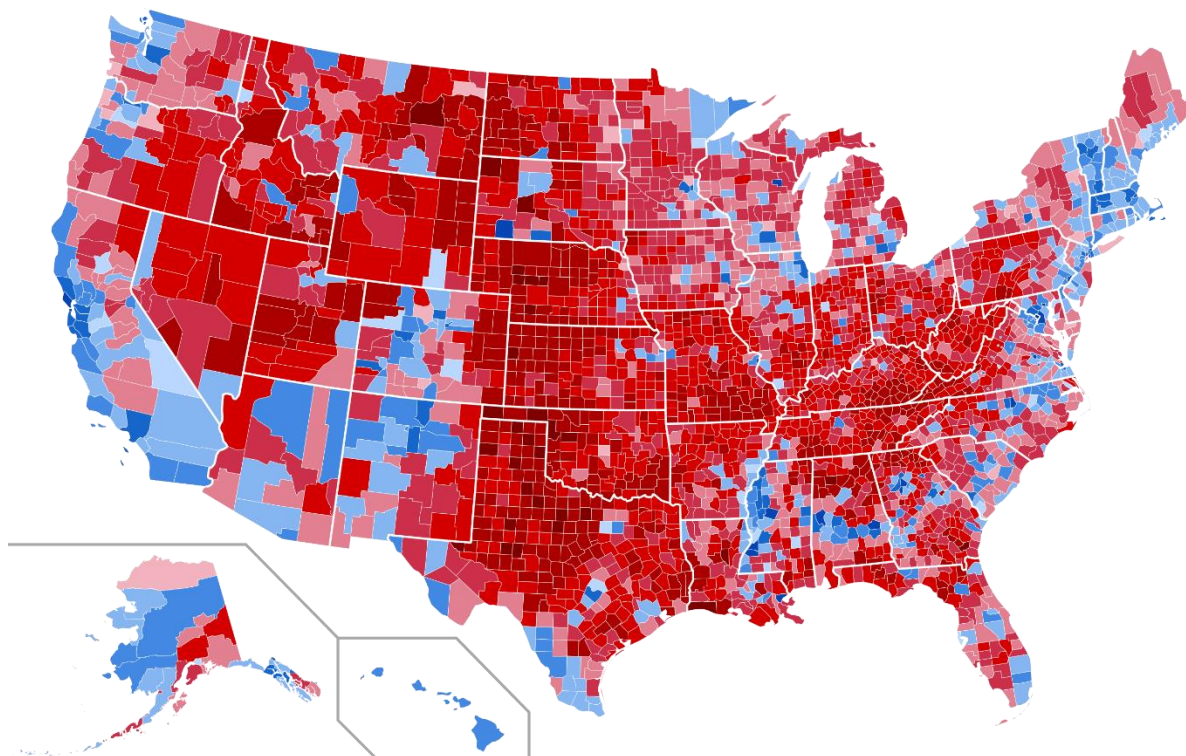


Figure 1. Election results at the county level. Red counties won by Trump and blue counties won by Biden.

US Presidential Election 2020

Results mapped at county level showing the candidate with the largest vote share in each area

Preliminary results*

Biden
78,780,121 votes (50.9%)
306 electoral votes

Trump
73,163,140 votes (47.3%)
232 electoral votes

* not confirmed final result, last updated 16 Nov 2020
Source: <https://github.com/levelatus/Election2020-NYT-Results/>

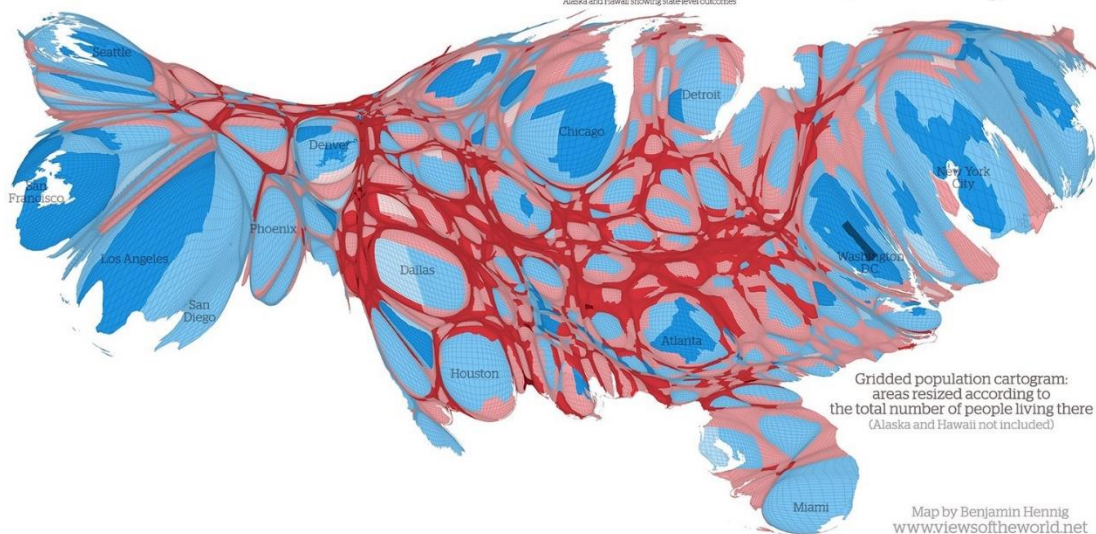
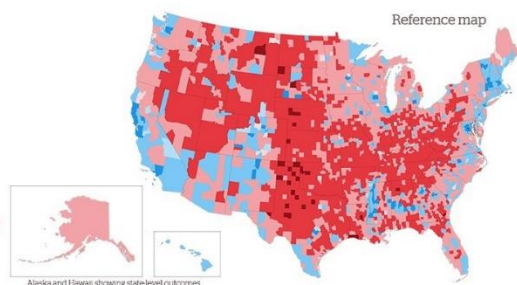
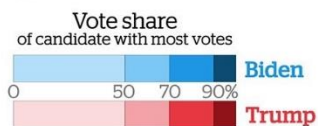


Figure 2. Cartogram showing the election results weighted by population. Red counties won by Trump and blue counties won by Biden.