PAST PAPER FOR GIS

according to geographical information system explain the following terms and phrases:

i. **Spatial data:** Spatial data are any data that have a location, such as coordinates, addresses, or boundaries. Spatial data can be represented in different formats, such as points, lines, polygons, or raster12.

ii. **Ground truthing:** Ground truthing is the process of verifying the accuracy and quality of spatial data by comparing them with the actual conditions on the ground34. Ground truthing can involve field surveys, observations, or measurements.

iii. **Spatial referencing/georeferencing:** Spatial referencing or georeferencing is the process of assigning a coordinate system and a projection to spatial data, so that they can be located and aligned with other spatial data on a map12. Spatial referencing or georeferencing can also involve transforming the coordinates of spatial data from one system to another.

iv. **Attribute data:** Attribute data are any data that describe the characteristics or properties of spatial data, such as names, types, or values12. Attribute data can be stored in tables or databases, and linked to spatial data by a common identifier or key.

v. **Spatial entities:** Spatial entities are any objects or phenomena that have a spatial dimension, such as buildings, rivers, or climate. Spatial entities can be classified into discrete or continuous, depending on whether they have well-defined boundaries or not.

vi. **GIS**: GIS stands for geographic information system, which is a system that creates, manages, analyzes, and maps all types of spatial data12. GIS can integrate data from various sources, such as maps, GPS, and imagery, and apply various methods and tools, such as classification, modeling, and visualization, to extract useful information and generate outputs, such as maps, reports, and dashboards.

Some of the main functions of GIS are[23]:

a. **Data input**: This is the process of encoding data into a digital format and writing it to the GIS database. Data input requires both spatial data (the location of features on a map) and non-spatial data (descriptive or numeric information about features) to be encoded into a GIS. Data input can be done using primary methods (direct measurement or capture of data) or secondary methods (conversion or derivation of data from other sources).
b. **Data storage**: This is the process of organizing and managing data in the GIS database. Data storage involves choosing appropriate data models (such as raster or vector) and data structures (such as tables or arrays) to represent the data in the GIS. Data storage also involves creating indexes, metadata, and schemas to facilitate data retrieval and manipulation.
c. **Data manipulation**: This is the process of modifying, transforming, or integrating data in the GIS database. Data manipulation involves performing operations such as editing, updating, merging, clipping, buffering, projecting, or generalizing data to suit the needs of the GIS project.
d. **Data analysis**: This is the process of applying various methods and techniques to extract information from data in the GIS database. Data analysis involves performing tasks such as querying, measuring, overlaying, interpolating, classifying, modeling, or simulating data to answer spatial questions or solve spatial problems.
e. **Data output**: This is the process of presenting or communicating the results of data analysis in the GIS database. Data output involves creating products such as maps, charts, tables, reports, or animations to display or share the information derived from data in the GIS.

vii. **Transformation:** Transformation is the process of changing the shape, size, orientation, or position of spatial data, to make them compatible or consistent with other spatial data or a reference system. Transformation can involve operations such as translation, rotation, scaling, or warping.

Describe the following dimensions of primary and secondary data

i. **Temporal dimension:** This refers to the time period when the data was collected or updated. Primary data is usually more recent and up-to-date than secondary data, which may be outdated or obsolete. The temporal dimension affects the relevance and accuracy of the data for the research purpose1.
ii. **Thematic dimension:** This refers to the topic or subject matter of the data. Primary data is usually more specific and focused on the research question, while secondary data may be more general or broad. The thematic dimension affects the suitability and applicability of the data for the research purpose2.
iii. **Spatial dimension:** This refers to the geographic area or location of the data. Primary data is usually more localized and representative of the target population, while secondary data may be more aggregated or dispersed. The spatial dimension affects the validity and reliability of the data for the research purpose3.

Describe by giving examples transformation in GIS

Transformation in GIS is the process of changing the geometry of a raster dataset from one coordinate system to another. This is necessary to ensure that data from different sources can be aligned and compared on a common map. There are two main types of transformation in GIS: map-to-map and image-to-map1.

i. **Map-to-map transformation:** This is the conversion of coordinates from one map projection to another. For example, if you have a map of Kenya in the Universal Transverse Mercator (UTM) projection, and you want to overlay it with a map of Africa in the Robinson projection, you need to apply a map-to-map transformation to convert the UTM coordinates to Robinson coordinates. This can be done using mathematical formulas or tables of parameters that define the relationship between the two projections1.

ii. **Image-to-map transformation:** This is the registration of an image, such as a satellite image or an aerial photograph, to a map projection. For example, if you have an image of Taita–Taveta county, and you want to georeference it to a map of Kenya in the UTM projection, you need to apply an image-to-map transformation to assign UTM coordinates to the pixels of the image. This can be done using control points, which are pairs of known coordinates on the image and the map, and transformation methods, such as affine, polynomial, or spline, that calculate the best fit between the control points1.

Digitizing is the process of converting information into a digital format that can be read and processed by computers.

Data input in GIS is the process of encoding data into a digital format and writing it to the GIS database.

There are two main types of data input in GIS: primary and secondary2.

i. **Primary data input:** This is the direct measurement or capture of data in digital format for use in a GIS project. Examples of primary data input are remote sensing, surveying, GPS, and field data collection2.

ii. **Secondary data input:** This is the conversion or derivation of data from other sources that were originally captured for another purpose. Examples of secondary data input are scanning, manual digitizing, vectorization, photogrammetry, and COGO.

Examples of GIS software:

i. **ArcGIS Pro:** Powerful and comprehensive professional GIS software from Esri, offering advanced analytical tools, 3D capabilities, and strong data management features.

ii. **QGIS:** Free and open-source GIS software with a vibrant community, offering a wide range of tools for analysis, mapping, and geoprocessing.

iii. **MapInfo Pro:** Industry-standard GIS software for asset management, infrastructure planning, and public safety applications.

iv. **Esri ArcGIS Online:** Cloud-based GIS platform providing map creation, spatial analysis, and data sharing capabilities for organizations of all sizes.
v. **Mapbox:** Platform for building custom web maps and applications with powerful interactive features and APIs.
vi. **Leaflet**: Open-source JavaScript library for creating lightweight and interactive maps on the web.

Discuss GIS as a hardware, software, liveware and data

i. **Hardware:** This refers to the physical devices and equipment that are used to run GIS applications. Hardware includes computers, servers, scanners, printers, plotters, GPS receivers, and other devices that can input, output, process, and store geographic data3.
ii. **Software:** This refers to the programs and applications that are used to create, manage, analyze, and visualize geographic data. Software includes operating systems, database management systems, GIS software packages, web browsers, and other tools that can handle geographic data4.
iii. **Liveware:** This refers to the human users and support staff that are involved in GIS projects. Liveware includes data collectors, data analysts, data managers, data users, programmers, technicians, and other professionals that can operate, maintain, and improve GIS systems5.
iv. **Data:** This refers to the information that is stored and manipulated in GIS systems. Data includes both spatial data (the location and shape of features on a map) and non-spatial data (the attributes and characteristics of features) that can be represented in different formats, such as raster, vector, image, or text.

Reason of GIS as a Hardware

Geographic Information System, is a hardware because it requires physical devices and equipment to run GIS applications. Hardware includes computers, servers, scanners, printers, plotters, GPS receivers, and other devices that can input, output, process, and store geographic data.

How to GIS acquire skills as liveware

i. **Learn the fundamentals of GIS:** You should have a solid understanding of the basic concepts and principles of GIS, such as spatial data types, coordinate systems, map projections, spatial analysis, and cartography. You can learn these fundamentals through formal education, such as a degree or certificate program in GIS, or through online courses and tutorials, such as those offered by Esri2 or USC3.

ii. **Practice using GIS software:** You should be proficient in using various GIS software applications, such as ArcGIS, QGIS, GRASS GIS, and Maptitude, to perform common GIS tasks, such as data input, data storage, data manipulation, data analysis, and data output. You can practice using GIS software by following exercises and projects from books, websites, or online courses, or by creating your own GIS projects based on your interests or needs.

iii. **Expand your knowledge of data formats and sources:** You should be familiar with different types of spatial and non-spatial data formats, such as CAD, databases, remote sensing, satellite imagery, LiDAR, IoT and sensors, building information modeling (BIM), and cloud services, and how to work with them in GIS. You should also know where to find and access reliable and relevant data sources for your GIS projects, such as government agencies, open data portals, or web services.

iv. **Learn database management and SQL:** You should be able to organize and manage your GIS data in a relational database management system (RDBMS), such as Oracle, MySQL, or PostgreSQL, and use SQL (Structured Query Language) to query and manipulate your data. You can learn database management and SQL through online courses and tutorials, such as those offered by Oracle4 or W3Schools.

v. **Learn programming and scripting languages:** You should be able to automate, customize, and extend your GIS workflows and applications by using programming and scripting languages, such as Python, R, JavaScript, or C#. You can learn programming and scripting languages through online courses and tutorials, such as those offered by Codecademy or Coursera.

vi. **Keep up with the latest trends and developments in GIS:** You should be aware of the current and emerging issues, challenges, and opportunities in the GIS field, such as big data, artificial intelligence, machine learning, web mapping, and geo-visualization. You can keep up with the latest trends and developments in GIS by reading blogs, newsletters, journals, and magazines, such as GIS Resources, GIS Lounge, or Directions Magazine, or by attending webinars, workshops, or conferences, such as Esri User Conference or FOSS4G.

Types of GIS data and how to create them

There are two main types of data in GIS: vector and raster. Vector data consists of points, lines, and polygons that represent the location and shape of features on a map. Raster data consists of pixels that represent the value or color of an area on a map.

Some of the common methods and sources for creating vector data are:

i. **Digitizing:** This is the process of converting analog data, such as paper maps or aerial photos, into digital vector data by tracing the features on a screen or a digitizing tablet. Digitizing can be done manually or automatically using software tools3.

ii. **Surveying:** This is the process of measuring the position and elevation of features on the ground using instruments, such as GPS, total station, or theodolite. Surveying can produce accurate and precise vector data for specific locations4.

iii. **Geocoding:** This is the process of assigning geographic coordinates to features based on their address or name. Geocoding can use existing databases, such as postal codes or street names, or online services, such as Google Maps or Bing Maps, to generate point vector data.

iv. **COGO:** This is the process of creating vector data based on coordinate geometry, such as bearings, distances, and angles. COGO can be used to create accurate and precise vector data for features that have regular shapes, such as parcels or boundaries.

Some of the common methods and sources for creating raster data are:

i. **Scanning**: This is the process of converting analog data, such as paper maps or photographs, into digital raster data by capturing the image using a scanner or a camera. Scanning can produce high-resolution raster data, but may require georeferencing to align the image with a map projection.
ii. **Remote sensing:** This is the process of acquiring raster data from a distance using sensors, such as satellites, airplanes, drones, or radar. Remote sensing can produce raster data for large areas and different wavelengths, such as visible, infrared, or microwave.
iii. **Interpolation:** This is the process of creating raster data from scattered points by estimating the value of the pixels between the points using mathematical methods, such as inverse distance weighting, kriging, or splines. Interpolation can produce continuous raster data, such as elevation or temperature.
iv. **Classification:** This is the process of creating raster data from existing raster data by grouping the pixels into categories based on their value or color. Classification can produce discrete raster data, such as land use or vegetation.

Why data collection is expensive:

i. **Data quality:** The quality of the data affects the accuracy and reliability of the GIS analysis and output. To ensure high-quality data, GIS professionals need to use appropriate techniques, instruments, and standards to capture, store, manipulate, and validate data. This may require investing in expensive hardware, software, and training.
ii. **Data quantity:** The quantity of the data affects the coverage and detail of the GIS project. To ensure sufficient data, GIS professionals need to collect data from various sources, such as remote sensing, surveying, geocoding, or scanning, and integrate them into a common

format and coordinate system. This may require spending a lot of time, bandwidth, and storage space.

iii. **Data availability:** The availability of the data affects the accessibility and suitability of the GIS project. To ensure relevant data, GIS professionals need to find and obtain data from reliable and updated sources, such as government agencies, open data portals, or web services. This may require paying for licenses, fees, or subscriptions.

Types of analysis procedures

i. **Queries and reasoning:** These are procedures that allow users to select, filter, or classify data based on their attributes or location. Queries and reasoning can be used to explore data, identify patterns, or test hypotheses. Examples of queries and reasoning are attribute queries, spatial queries, spatial joins, and spatial reasoning.

ii. **Measurements:** These are procedures that allow users to calculate geometric or topological properties of data, such as distance, area, perimeter, direction, angle, or connectivity. Measurements can be used to describe data, compare data, or perform analysis. Examples of measurements are buffer, point-in-polygon, polygon overlay, and network analysis.

iii. **Transformations:** These are procedures that allow users to modify, convert, or project data from one format, structure, or coordinate system to another. Transformations can be used to prepare data, integrate data, or change data. Examples of transformations are geocoding, digitizing, vectorization, and map projection.

iv. **Descriptive summaries:** These are procedures that allow users to summarize, aggregate, or generalize data using statistical or graphical methods. Descriptive summaries can be used to characterize data, simplify data, or display data. Examples of descriptive summaries are zonal statistics, spatial autocorrelation, spatial interpolation, and cartography.

v. **Optimization:** These are procedures that allow users to find the best or most efficient solution to a spatial problem or objective, subject to some constraints or criteria. Optimization can be used to plan, allocate, or evaluate data. Examples of optimization are site selection, location-allocation, routing, and spatial modeling.

vi. **Hypothesis testing:** These are procedures that allow users to test the validity or significance of a spatial hypothesis or relationship, using statistical or inferential methods. Hypothesis testing can be used to confirm, reject, or refine data. Examples of hypothesis testing are spatial regression, spatial clustering, spatial sampling, and spatial simulation.