0

Last Updated 19 September 2019
Version 2.0

GISCI Geospatial Core Technical Knowledge Exam® Unofficial Study Guide

Primary Editor: Miklos Nadas, GISP

Disclaimer: The GIS Certification Institute (GISCI) and Miklos Nadas are not affiliated. GISCI requested Miklos Nadas to compile his publicly shared GISCI Geospatial Core Technical Knowledge Exam® notes into a comprehensive study guide. The information in this document is supposed to supplement an individual in preparation for the GISCI Geospatial Core Technical Knowledge Exam®. The information was not reviewed by GISCI for its validity or accuracy. Instead, others in the GIS community have assisted with contributing to this study guide and reviewing its content. It is recommended to utilize all the Study Resources documented in the GISCI Geospatial Core Technical Knowledge Exam® Candidate Material.

CONTENTS

MESSAGE FROM THE EDITOR

In 2015, I took the brand new GISCI Geospatial Core Technical Knowledge Exam®. I did not know what to expect on the exam, so I expected everything. I took a few months before the exam to go through all my notes and text books supplemented with online resources to create myself a study guide. After creating over 70 pages of material to review, I whittled it down to about 20 pages in a condensed version. While working on the study guide, I shared it with the GIS community and posted it to social media sites. Only I contributed to the study guide in that first year, but I decided to keep the study guides available to the public.

In the past few years, I get the occasional email requesting editing rights to the document, which I gladly approve. I am happy to hear that these documents have been shared, updated, and referenced by multiple individuals. To my surprise, Bill Hodge from GISCI reached out to me to discuss these study guides. GISCI requested that I formalize these study guides into a comprehensive format for others to use.

I am dedicated to keeping this study guide available for free to the public and updated. I will also keep the documents on Google Drive available as another source for studying. For all those who want to help keep this unofficial study guide up to date, accurate, and one of the best resources for the studying for the GISP Exam, please reach out to me at miklos.nadas@gmail.com. I am looking for content enhancement, non-copyrighted images (images you create) to supplement the information, grammar proofreading, validity of the information, and notification on broken URL links.

Also, thank you to everyone who contributed to the original notes on Google Drive and this study guide.

Good luck to everyone taking the GISP exam!

STUDY GUIDE VERSIONS

V1 Notes: Version 1 of the study guide was published in October of 2017 following the set of KSAs created in 2015

V2 Notes: In July 2019, a new format for KSAs were devised by GISCI. These new KSAs were matched with the older KSAs in the following link. The version 1 study guide was updated to follow the new format KSA. Some slight updates were also made to the study guide.

https://www.gisci.org/Portals/0/PDF's/BP%20crosswalk%20for%20dist%20w%202015%20Blueprint%207-19.pdf

EDITORS

PRIMARY EDITOR - MIKLOS NADAS, GISP

Miklos Nadas graduated with a master's degree in Geography and GIS from the University of Akron in 2009. He has over eleven years of work experience using multiple GIS platforms for spatial analysis and data management. Currently, he works for Brown and Caldwell, a consulting firm which provides water, storm water, wastewater, and environmental services for clients. He supports clients with GIS analysis, data management, data collection, software development, and script automation.

CONTRIBUTING EDITORS

Charles Altendorf

Becca Decker

Shital Upanmanyu Dhakal, GISP

Mir Hashmi

CONCEPTUAL FOUNDATIONS

101. UNDERSTANDING OF DATUMS, COORDINATE SYSTEMS, AND PROJECTIONS

A. Georeferencing – associating a map (such as a pdf without spatial information) or image (such as an aerial image without spatial information) with spatial locations
B. Control points – consisting of multiple points, points come in pairs that match the spatial location with a point on an unreferenced image or map
C. Spatial reference system (SRS) or coordinate reference system (CRS) is a coordinate-based local, regional or global system used to locate geographical entities
D. International Terrestrial Reference System (ITRS). It is a three-dimensional coordinate system with a well-defined origin (the center of mass of the Earth) and three orthogonal coordinate axes (X,Y,Z)
E. Map projection - transforming coordinates from a curved earth to a flat map
    a. UTM - Universal Transverse Mercator - a global coordinate system - UTM zones are 6 degrees
F. Horizontal datum - model of the earth as a spheroid (2 components, reference ellipsoid and a set of survey points both the shape of the spheroid and its position relative to the earth)
G. Vertical datum - reference point for elevations of surfaces and features on the Earth - could be based on tidal, seas levels, gravimetric, based on a geoid
H. NAVD88 - gravity based geodetic datum in North America
I. Geodetic datum - set of control points whose geometric relationships are known, either through measurement or calculation
J. WGS 84 - World Geodetic System - reference coordinate system used by the Global Positioning System (GPS)
K. SRID integer - spatial reference system id numbers, including EPSG codes defined by the International Association of Oil and Gas Producers
L. 4 distortions - Distance - Direction - Shape – Area
M. Mercator Projection - Preserves shape and direction, area gets distorted - projecting earth onto a cylinder tangent to a meridian
N. Azimuthal Equidistant - planar (tangent) - used for air route distances - distances measured from the center are true - distortion of other properties increases away from the center point
O. Cylindrical equal-area projections - preserves area, shape and distance gets distorted near the upper and lower regions of the map - straight meridians and parallels - meridians are equally spaced and the parallels are unequally spaced
P. Conic projections - preserves directions and areas in limited areas - distorts distances and scale except along standard parallels - generated by projecting a spherical surface onto a cone
Q. Choosing a projection:
    a. Latitude: Low-latitude areas (near equator) use a conical projection; Polar regions use a azimuthal planar projection

b. Extent: Broad in East-West (e.g., the US) use a conical projection; Broad in North-South (e.g., Africa) use a transverse-case cylindrical projection
c. Thematic: If you are doing an analysis that compares different values in different locations, typically an equal-area projection will be used

Additional Information:

http://www.geog.ucsb.edu/~kclarke/G176B/Lecture04.ppt

https://en.wikipedia.org/wiki/Spatial_reference_system

https://en.wikipedia.org/wiki/Geodetic_datum

102. UNDERSTANDING OF REPRESENTATION OF DISCRETE FEATURES AND CONTINUOUS PHENOMENA IN GIS

A. Discrete features – a feature that has a definable boundary, begins and ends, for example a highway or lake
B. Continuous phenomena – each location is a measure of something, for example elevation
   a. Measure of concentration level
   b. Measure of a value in terms of a fixed point (like elevation in terms of sea level)
C. Be able to indicate if a geographic feature is either discrete or continuous

Additional Information:

http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Discrete%20and%20continuous%20data%20in%203D%20Analyst

103. KNOWLEDGE OF EARTH GEOMETRY AND ITS APPROXIMATIONS

A. Geoid is the shape that the surface of the oceans would take under the influence of Earth's gravitation and rotation alone, in the absence of other influences such as winds and tides - used to reference heights, by registering ocean's water level at coastal places using tide gauges - this is how the mean sea level is determined
B. Reference ellipsoid is a mathematically defined surface that approximates the geoid, the truer figure of the Earth, or other planetary body
C. Oblate ellipsoid - fits the geoid to a first order approximation - formed when an ellipse is rotated about its minor axis
D. Sphere - As can be seen from the dimensions of the Earth ellipsoid, the semi-major axis a and the semi-minor axis b differ only by a bit more than 21 kilometers
E. First (direct) geodetic problem - Given a point (in terms of its coordinates) and the direction (azimuth) and distance from that point to a second point, determine (the coordinates of) that second point.
F. Second (inverse) geodetic problem - Given two points, determine the azimuth and length of the line (straight line, arc or geodesic) that connects them
G. For more information on datums, see Section 101

104. KNOWLEDGE OF BASIC GEOMATICS AND RELATIONSHIPS TO GIS

A. Geomatics - science and technology of gathering, analyzing, interpreting, distributing, and using geographic information (includes surveying, mapping, remote sensing, GIS, GPS)
B. For more information on GPS, see the Section on GPS

Additional Information:

http://wiki.gis.com/wiki/index.php/Geomatics

http://wiki.gis.com/wiki/index.php/Global_Positioning_System

GEOSPATIAL DATA FUNDAMENTALS

201. UNDERSTANDING OF SPATIAL DATA MODELS AND THEIR ASSOCIATED PLANAR GEOMETRIES

A. Spatial model - Basic properties and process for a set of spatial features
   a. According to Bolstad
      i. Cartographic Model – temporally static, combined spatial datasets, operations and functions for problem-solving
      ii. Spatio-temporal models – dynamics in space and time, time-driven processes
      iii. Network models - modeling of resources (flow, accumulation) as limited to networks
   b. According to Goodchild
      i. Data models - entities and fields as conceptual models
      ii. Static modeling - taking inputs to transform them into outputs using sets of tools and functions
      iii. Dynamic modeling - iterative, sets of initial conditions, apply transformations to obtain a series of predictions at time intervals
   c. According to DeMers
      i. Based on purpose descriptive - passive, description of the study area prescriptive - active, imposing best solution
      ii. Based on methodology stochastic - based on statistical probabilities deterministic - based on known functional linkages and interactions
      iii. Based on logic inductive - general models based on individual data deductive - from general to specific using known factors and relationships
B. Vector - coordinate based data model that represents points, lines, polygons
   a. Points – discrete locations on the ground
      i. Represented by a coordinate pair
   b. Lines – linear features, such as rivers, roads, and transmission cables
      i. Composed of vertices
      ii. Begin and end at vertices
      iii. Represented by an ordered list of vertices
   c. Polygons – form bounded areas, such as islands, land masses, and water features
      i. Composed of nodes and vertices
      ii. The start node is the same as the end node
   d. Attributes associated with each feature
C. Raster - composed of rectangular arrays of regularly spaced square grid cells and each cell has a value (attribute)
   a. Examples include soil pH, elevation, salinity of a water body
   b. Single or multiple bands
   c. Each cell has 1 attribute value
   d. Raster coordinates are stored by ordering the matrix

D. Pixel - smallest resolvable piece of scanned image - pixel is always a cell but a cell is not always a pixel
E. Geodatabase - object oriented spatial model (feature classes, feature datasets, non-spatial tables, topology, relationship classes, geometric networks)
    a. Basic components include feature classes, feature datasets, non-spatial tables
    b. Complex components include topology, relationship classes, geometric networks
    c. Relationship classes – model real-world relationships that exist between objects such as parcels and buildings
        i. More information on relationships in Section 537
F. Grid - parallel and perpendicular lines for reference as a map projection or coordinate system
G. TIN - Triangulated Irregular Network - portions vector data into contiguous, nonoverlapping triangles
    a. Create Delaunay triangles
    a. Advantages of TIN - small areas with high precision elevation data
        i. More efficient storage than DEM or contour lines
    b. Disadvantage of TIN - requires very accurate data source and costs are expensive, TIN production and use are very computer intensive)
H. Topological - features need to be connected using specific rules
I. Hierarchical - database that stores related information in a tree-like structure
    b. Records can be traced to parent records to a root record
J. Network - collection of topologically connected network elements (edges, junctions, turns)
    c. Each element is associated with a collection of network attributes
K. Object Oriented - data management structure stores data as objects (classes) instead of rows and tables as a relational database
    d. Examples include SQL Server, Oracle, PostgreSQL

Additional Information:

https://books.google.com/books?id=dk9Ebkf5x7UC&lpg=PA74&ots=F5Ruo0IoWi&dq=gis%20planar%20geometry&pg=PA75#v=onepage&q=gis%20planar%20geometry&f=false

https://en.wikipedia.org/wiki/Plane_(geometry)

http://www.directionsmag.com/entry/geography-and-everyday-life/129601

http://www.directionsmag.com/entry/thinking-spatially/123985

http://www.dbdebunk.com/2017/03/what-is-true-relational-system-and-what.html

https://courses.washington.edu/gis250/lessons/introduction_gis/spatial_data_model.html

http://www.esri.com/news/arcnews/summer99articles/ai8special/ai8_object.html

202. UNDERSTANDING OF SPATIAL DATA RELATIONSHIPS

A. Types of relationships

a. 1-1 - each object of the origin table can be related to 0 or 1 object of the destination table
b. 1-Many - each object in the origin table can be related to the multiple objects in the destination table
c. Many-Many - multiple objects of the origin table can be related to multiple objects of the destination table

B. Equals - a = b - topologically equal
C. Disjoint - $a \cap b = \emptyset$ - no point in common
D. Intersects - $a \cap b \neq \emptyset$ - some common interior points
E. Touches - $(a \cap b \neq \emptyset) \wedge (a^o \cap b^o = \emptyset)$ - a touches b, at least one boundary point in common but no interior points
F. Contains - $a \cap b = b$ - feature b is within a
G. Covers - $a^o \cap b = b$ - every point of b is a point of a
H. Covered By - Covers(b,a) - every point of a is a point of b
I. Within - $a \cap b = a$ - a is within b
J. Crosses - a crosses b at some point
K. Overlaps - a and b have common interior points
L. Basic topology rules
    a. Polygon rules
        i. Must be larger than cluster tolerance
        ii. Must not overlap
        iii. Must not have gaps
        iv. Must not overlap with
        v. Must be covered by feature class of
        vi. Must cover each other
        vii. Must be covered by
        viii. Boundary must be covered by
        ix. Area boundary must be covered by boundary of
        x. Contains point
        xi. Contains one point
    b. Line rules
        i. Must be larger than cluster tolerance
        ii. Must not overlap
        iii. Must not intersect
        iv. Must not intersect with
        v. Must not have dangles
        vi. Must not have pseudo nodes
        vii. Must not intersect or touch interior
        viii. Must not intersect or touch interior with
        ix. Must not overlap with
        x. Must be covered by feature class of
        xi. Must be covered by boundary of
        xii. Must be inside

xiii. Endpoint must be covered by
xiv. Must not self-overlap
xv. Must not self-intersect
xvi. Must be single part

c. Point rules
    i. Must coincide with
    ii. Must be disjoint
    iii. Must be covered by boundary of
    iv. Must be properly inside
    v. Must be covered by endpoint of
    vi. Point must be covered by line

Additional Information:

https://en.wikipedia.org/wiki/Spatial_relation

https://apollomapping.com/blog/g-faq-spatial-topology-gis-part/topology_rules_poster-2

203. UNDERSTANDING OF DATA QUALITY

A. Geometric accuracy – The closeness of a measurement to its true value
B. Root Mean Squared Error (RMS) – a calculation to describe the difference between the measurement and the true value
    a. This can apply to georectification
    b. RMS = the square root of the average of squared errors
C. Thematic Accuracy - accuracy of the non-spatial data
    a. Such as the is the street name accurate on a street feature class
D. Resolution – smallest separation between two coordinate values
    a. For rasters this refers to the cell size
E. Precision – level of measurement and exactness of attribute data
F. Fitness for use – Does the data fulfill the needs of the project
G. Confusion matrix – assesses accuracy of image classification based on additional ground truths
H. Quality Assurance - process oriented and focuses on defect prevention
    a. Establishment of good quality management system and assessment of its adequacy - periodic audits - managerial tool
I. Quality Control - product oriented and focuses on defect identification
    a. Finding and eliminating sources of quality problems through tools and equipment - corrective tool
J. Imprecision - all data is taken from a 3D globe and transferred to a 2D surface through spatial transformations (projections and datums) which causes distortions with the data
K. Uncertainty - The GIS data was created/collected at a certain point of time, may already be out of date

Additional Information:

https://en.wikipedia.org/wiki/Root-mean-square_deviation http://www.ce.utexas.edu/prof/maidment/giswr2005/visual/Scale.ppt

http://spatial-analyst.net/ILWIS/htm/ilwismen/confusion_matrix.htm

http://wiki.gis.com/wiki/index.php/Quality_control

http://www.diffen.com/difference/Quality_Assurance_vs_Quality_Control

204. UNDERSTANDING OF DATA RESOLUTION

A. Data resolution – the cell size of a raster – the area covered by the ground represented by just one cell

Additional Information:

http://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/cell-size-of-raster-data.htm

https://ibis.geog.ubc.ca/~brian/Course.Notes/gisscale.html

205. UNDERSTANDING OF DATA VALIDATION AND UNCERTAINTY

A. Uncertainty – differences between real world and GIS
    a. May be visible from the original data or measuring that data
    b. Assumptions made when creating the data
    c. Model structure
    d. Retrieval errors, sampling errors, and inadequate ground observations
B. Validation – to ensure the accuracy of the data is preserved
    a. Ground observations to ensure data accuracy
    b. Can also be compared to model generated data, but less accurate

Additional Information:

https://dusk.geo.orst.edu/gis/Chapter6_notes.pdf

https://www.gislounge.com/determining-uncertainty-in-gis-data/

https://cpb-us-e1.wpmucdn.com/blogs.uoregon.edu/dist/f/11715/files/2015/08/Lecture-17-19aar17.pdf

https://solargis.com/docs/accuracy-and-comparisons/overview

https://elibrary.worldbank.org/doi/abs/10.1596/978-1-4648-0475-5_ch8

206. UNDERSTANDING OF METADATA

A. Metadata - information that describes the content, quality, condition, origin and other characteristics of data or other pieces of information

Additional Information:

https://en.wikipedia.org/wiki/Geospatial_metadata

207. KNOWLEDGE OF TEMPORAL DATA

A. Temporal – data that represents a state in time – for example, rain fall for a one day
    a. the world is constantly changing. A static map can only show one time period at once.
    b. Dynamic maps can allow a user to slide between different time periods to show changes.
    c. Multiple maps can be created to show the changes or somehow superimpose multiple time period pieces of data on a static map

Additional Information:

http://desktop.arcgis.com/en/arcmap/10.3/map/time/what-is-temporal-data.htm

http://wiki.gis.com/wiki/index.php/Temporal_GIS

208. KNOWLEDGE OF SPATIAL DATA STANDARDS, INCLUDING ISO, FGDC, AND OGC

A. Federal Geographic Data Committee (FGDC) - who, what, when, where, why, and how
    a. Include title, abstract and date, geographic extent and projection info, attribute label definitions, and domain values
B. Content Standard for Digital Geospatial Metadata (CSDGM)
    a. ISO 19115 - developed for documenting vector and point data and geospatial services (web-mapping, data catalogs, and data modeling applications)
    b. ISO 19115-2 - adds elements to describe imagery and gridded data as well as data collected using instruments (monitoring stations and measurement devices)
C. OGC – Open GIS Consortium
    a. Describes basic data model for holding geographic data
    b. These are data file types, such as KML

Additional Information:

https://www.fgdc.gov/metadata/geospatial-metadata-standards

https://www.fgdc.gov/metadata

https://en.wikipedia.org/wiki/Open_Geospatial_Consortium

CARTOGRAPHY AND VISUALIZATION

301. UNDERSTANDING OF GRAPHIC REPRESENTATION TECHNIQUES AND IMPLICATIONS

A. Thematic map is a type of map especially designed to show a particular theme connected with a specific geographic area
B. Choropleth - areas are shaded according to prearranged key, each shading or color type represents a range of values
C. Proportional Symbol - symbol drawn proportional in size to the size of the variable being represented
D. Isarithmic or Isopleth - lines of equal value are drawn (contour lines) or ranges of similar values are filled with similar colors or patterns
E. Dot - shows distribution of phenomena where values and locations are known - place a dot where the location of variable is
F. Dasymetric - alternative to choropleth - ancillary information is used to model internal distribution of the phenomenon
G. Multivariate displays - putting more than two sets of data on one map (i.e. single map shows population density and annual rainfall and cancer rates)
H. Web mapping - process of using maps delivered by GIS - web maps are both served and consumed

Additional Information:

https://en.wikipedia.org/wiki/Thematic_map

http://geographyfieldwork.com/DataPresentationMappingTechniques.htm

https://en.wikipedia.org/wiki/Web_mapping

302. UNDERSTANDING OF MAP DESIGN PRINCIPLES AND ESSENTIAL MAP ELEMENTS

A. Map layout elements - a title, map, legend, map scale, supporting media, north arrow, metadata (sources, currency of information, projection, copyright, authorship)
B. Symbols - represent things on a map
C. Map accuracy - difficult to assess, all maps show a selective view of reality - instead should ask if the map is appropriate for my purposes
D. Map scale - 1:100 - one inch represents 100 inches in the real world
    a. Large scale (more zoomed in) shows more detail than small scale (more zoomed out)
E. Symbolization variables - size, shape, orientation, pattern, hue, value
F. Quantitative –
    a. Size – the size of the point or the thickness of a line
    b. Value – the shade of the color such as dark red or light red
G. Qualitative –
    a. Shape – for points different symbol

   b. Pattern – lines having different styles such as dashed lines
   c. Hue – different colors, such as red and blue
H. Typography - the design of text, point size, line length, typefaces
I. Map Scales
   a. Verbal scale - expresses in words a relationship between a map distance and ground distance: one inch represents 16 miles
   b. Visual scale - graphic scale or bar scale
   c. Representative scale - representative fraction or ratio scale 1:24,000 - 1" = 24,000"
   d. Absolute scale - system of measurement that begins at a minimum or zero point and progresses in only one direction
   e. Relative scale (arbitrary) - begins at some point selected by a person and can progress in both directions
   f. Display vs Data - The data is built at a certain scale/accuracy but once the data is displayed in any other format that the one it was made for, the scale gets warped. Ex: a map made as 9"x10" that is then scaled down and printed in a newspaper.
   g. Large scale – small ratio between map units and ground units
      i. Depict small areas such as USGS topographic maps or neighborhoods
   h. Small scale – large ratio between map units and ground unit
      i. Depict large areas such as countries or continents

Additional Information:

http://wiki.gis.com/wiki/index.php/Cartographic_Design

http://www.esri.com/industries/k-12/education/~/media/Files/Pdfs/industries/k-12/pdfs/intrcart.pdf

http://wiki.gis.com/wiki/index.php/Typography

http://richardphillips.org.uk/maps/type.html

http://www.sfei.org/book/export/html/1321

https://en.wikipedia.org/wiki/Absolute_scale

https://www.perceptualedge.com/articles/visual_business_intelligence/geographical_data_visualization.pdf

303. UNDERSTANDING OF SURFACE INTERPRETATION AND REPRESENTATION

A. Be able to read an aerial map and interpret features
B. Be able to decide how a geographic feature should be represented in GIS
C. See Section 102

304. UNDERSTANDING OF 2D AND 3D VISUALIZATION

A. 3D mapping brings in z-value
   a. Typically, elevation data

B. Contour line - (aka isoline, isopleth, or isarithm) a function of two variables is a curve along which the function has a constant value - joins points of equal value on a line
C. Contour interval - difference in elevation between successive contour lines
D. Index contour – the contour that is thicker and typically labeled
E. Iso – equal – equal distances between lines
F. 3d mapping has also included building modeling

Additional Information:

https://en.wikipedia.org/wiki/Contour_line

https://gis.usc.edu/blog/4-uses-of-3d-gis/

DATA ACQUISITION

401. UNDERSTANDING OF DIGITIZATION AND OTHER MANUAL DATA COLLECTION AND CONVERSION METHODS

A. Primary data - collected specifically for the purpose of a researcher's particular study
B. Secondary data - collected for another purpose by someone other than the researcher
C. 5 types of measurement - physical measurement, observation of behavior, archives, explicit reports, computational modeling
    a. Physical Measurement - recording physical properties of the earth or its inhabitants - size, number, temperature, chemical makeup, moisture, etc.
    b. Observation of behavior - observable actions or activities of individuals or groups - not thoughts, feelings or motivations
    c. Archives - records that have been collected primarily for non-research purposes (secondary)
    d. Explicit reports - beliefs people express about things – survey
    e. Computational Modeling - models as simplified representations of portions of reality
D. Quantitative data - numerical values, measured on at least an ordinal level but could be on a metric level
E. Qualitative data - nonnumerical or numerical (nominal) values that have no quantitative meaning
F. Deceptive mapping - maps can be distorted for propaganda, military protection, or ignorance
G. Layer – mechanism to display geographic datasets
A. Data Transfer Standards
    a. Transfer - follow Spatial Data Transfer Standard (SDTS) - Federal Information Processing Standard (173)- robust way of transferring GIS data between computers with no information loss, including metadata
    b. Industry Standards - typically do not exchange topology, only graphic info; large number of format translators
    c. Open GIS Consortium (OGC) – non-profit, international, voluntary consensus standards organization - created GML or Geography Markup Language - XML based encoding standard
B. For information on degrees/minutes/seconds, see the Section on GPS

Additional Information:

http://www.geog.ucsb.edu/~kclarke/Geography128/Lecture07.pdf

http://dusk.geo.orst.edu/gis/Chapter9_notes.pdf

http://www.igic.org/resources/standards/map-scale-accuracy/

https://www.fgdc.gov/standards/projects/FGDC-standards-projects/SDTS

http://maps.unomaha.edu/Peterson/gis/notes/DataFormatExchange.html

402. KNOWLEDGE OF FIELD DATA COLLECTION

A. Field data collection types
    a. Remote Sensing - 3 resolutions - spatial, spectral (electromagnetic spectrum measured), temporal (repeat cycle)
    b. Ground survey
        i. Ground surveying based on the principle that the 3-D location of any point can be determined by measuring angles and distances from other known points
        ii. Expensive and time consuming
    c. GPS – using a GPS receiver to receive signals from GPS satellites to calculate the current position and time
    d. Inspection – data has already been geographically located and needs to be inspected
B. Field collection process
    a. Determine the result of field work
    b. Determine what needs to be collected, inspected, or surveyed
        i. Set up a field collection form
    c. Determine how it will be collected, i.e. pen and paper, mobile tablet, drone
    d. Begin field collection on a good representation of the entire dataset
    e. Review sample field collection and adjust the data being collected or the method of data collection
    f. Plan locations and timing for field work
    g. Start the field collection for all assets

403. KNOWLEDGE OF AUTOMATED DATA COLLECTION AND CONVERSION METHODS

A. Aerial imagery can be used to automatically update GIS data
B. Data can be automatically updated (or appended to existing data, such as a time series dataset) and instantly converted into a useable format
C. Most automation can be completed by writing scripts

Additional Information:

https://pdfs.semanticscholar.org/2f1b/a5017ffdaf8e5c62958abc8bb1a01a103311.pdf

https://www.esri.com/arcgis-blog/products/product/water/automated-data-acquisition-for-integrating-multiple-datasets-using-gis-applications/

404. KNOWLEDGE OF REMOTELY SENSED DATA SOURCES AND COLLECTION METHODS

A. Remote Sensing - 3 resolutions - spatial, spectral (electromagnetic spectrum measured), temporal (repeat cycle)
B. Aerial photography and satellite imagery
C. Passive sensors – gather radiation that is emitted from objects
    a. Photography, infrared, radiometers
D. Active sensors – emit energy and measure the amount of energy bounced back from objects

a. RADAR and LiDAR

Additional Information:

https://serc.carleton.edu/research_education/geopad/imagery_data.html

https://www.sciencedirect.com/topics/computer-science/remotely-sensed-data

https://en.wikipedia.org/wiki/Remote_sensing

405. KNOWLEDGE OF ACQUISITION, USE, AND LIMITATIONS OF CROWDSOURCED AND OPEN SOURCE DATA AND SERVICES

A. Understand the limitations of using another entity's collected GIS data
   a. GIS data gathered for one entity's project may not fully encompass the needs of a new or different project the data is being used for
B. Understand the limitations between a paid data source and a free to download data source
C. Understand the limitations of a data source completed by multiple individuals
   a. Open Street map – a huge collaboration of different contributors – due to the huge collaboration, it could be more accurate due to a lot of contributors, but it could also be inaccurate due to it relying on individuals

DATA MANIPULATION

501. UNDERSTANDING OF GEOREFERENCING, DATA FORMAT CONVERSION, AND DATA TRANSFORMATION

A. Georeferencing – associating a map (such as a pdf without spatial information) or image (such as an aerial image without spatial information) with spatial locations
B. Data format conversion – such as converting from a vector based dataset to a raster based dataset
C. Data transformation – converting from different coordinate systems
    a. Also, data transformation may refer to moving from one data structure to another data structure – such as aggregating road data may require attributes to be converted to a standard type (such as road type)

Additional Information:

https://desktop.arcgis.com/en/arcmap/latest/map/projections/geographic-transformation-methods.htm

https://en.wikipedia.org/wiki/Data_transformation

502. UNDERSTANDING OF SPATIAL DATA GENERALIZATION OPERATIONS AND METHODS

A. A map or GIS dataset does not need to contain an excess of detail for certain unimportant aspects
    a. Not every road needs to be shown on a weather map
B. Generalization can be achieved by
C. Generalization will cause data to be less accurate
D. Typically for smaller scale maps
E. Generalization can be achieved by
    a. Selection – only selecting certain features to be displayed
    b. Removing detail
    c. Simplification - smoothing out detailed and complicated features, such as rivers do not need every curve
    d. Dissolve or Merge – combining similar neighbor elements
    e. Aggregation – combining features into a new compsite feature
    f. Exaggeration – make features larger than they are

Additional Information:

https://www.gislounge.com/generalization-gis/

https://en.wikipedia.org/wiki/Cartographic_generalization

503. UNDERSTANDING OF SPATIAL FILE TYPES AND THEIR APPLICATIONS AND LIMITATIONS

A. Vector - a coordinate based data model that represents features such as points, lines, polygons
B. Raster - defines space as an array of equally sized cells arranged in rows and columns, single or multiple bands - each cell contains an attribute value
C. Vector Advantages
    a. Represent point, line, area very accurately
    b. More efficient than raster in storage
    c. Supports topology
    d. Interactive retrieval
    e. Enables map generalization
D. Vector Disadvantages
    a. Less intuitively understood
    b. Multiple vectors overlay is computationally intensive
    c. Display and plotting vectors can be expensive
E. Raster Advantages
    a. Easy to understand
    b. Good to represent surfaces
    c. Easy to input and output
    d. Easy to draw on a screen
    e. Analytical operations are easier
F. Raster Disadvantages
    a. Inefficient for storage
    b. Compression techniques not efficient with variable data
    c. Large cells could potential cause information loss
    d. Poor at representing points, lines, areas
    e. Each cell can be owned by only one feature
    f. Must include redundant or missing data

G. Vector Formats
    a. PostScript - page definition language to export or print a map
    b. Digital Exchange Format (DXF) - AutoCAD - no topology but lots of details
    c. Digital Line Graph (DLG) - distributed by the government and most GIS packages will import but extra manipulation needed
    d. TIGER - block level maps of every village, town, and city in US
    e. Shapefile - vector data format stores location, shape, and attributes
    f. Scalable Vector Graphics (SVG) - extension of the XML language
    g. ArcInfo Coverage - stores set of thematically associated data considered to be a unit
    h. ArcInfo Interchange File (.e00) - known as ArcGIS export file
    i. Geodatabase - object oriented data model represents features and attributes as objects
H. Raster Format
    a. Standard - rows and columns with a header information
    b. Tagged Image File Formats (TIFF) - associated with scanners

c. GEO-TIFF - puts latitude/longitude at edges of pixels
d. Graphic Interchange Format (GIF) - image files for sharp edges and few gradations of color
e. Joint Photograph Experts Group (JPEG) - variable-resolution compression system with both partial and full resolution recovery
f. Digital Elevation Model (DEM) - 30 meter elevation data 1:24000 7.5 minute quadrangle maps or 1:250,000 3 arc second digital terrain data
g. Band Interleaved by Pixel (BIP) or Band Interleaved by Line (BIL) - good at storing different brightness levels
h. RS Landsat - satellite imagery and BIL information are combined
i. 3 Types of Resolution
    i. Spatial Resolution – size of object that can be resolved and the most usual measure is the pixel size
    ii. Spectral resolution - parts of the electromagnetic spectrum that are measured
    iii. Temporal resolution - repeat cycle - frequency with which images are collected for the same area

I. Raster to Vector is not difficult based on pixel value
J. Vector to Raster is very difficult because pixels may distort the lines or exact point locations and would need to be re-digitized
K. File Types
    a. SHP - shapefile
        i. .shp - shape format - feature geometry itself
        ii. .shx - shape index format - positional index of the feature geometry to allow seeking forwards and backwards quickly
        iii. .dbf - attribute information
        iv. .prj - projection format
        v. .sbn & .sbx - spatial index
        vi. .shp.xml - geospatial metadata in XML format
        vii. .shp, .dbf, and .prj are required to properly display a shapefile and its attributes
    b. GDB - geodatabase
        i. .gdb - file geodatabase
        ii. .mdb - personal geodatabase based on microsoft access
    c. Coverage file - point, arc, node, route, route system, section, polygon, and region
    d. DGN - AutoCAD and MicroStation
    e. Txt - Text
    f. IMG - Image
    g. LiDAR - remote sensing technology that measures distance by illuminating a target with a laser and analyzing the reflected light
    h. Raster - .jpg, .tif, .gif
    i. Rendered file formats are raster formats

Additional Information:

http://gif.berkeley.edu/documents/GIS_Data_Formats.pdf

A. File creation, edit, management, back up data, keep track of files, organize files
B. Individual files - shapefiles, file gdb, personal gdb, tables, spreadsheets, CAD, rasters
C. Databases - direct connection to relational database management systems and big data databases - manage tables and feature classes in database
D. Geodatabase - stores GIS in central location for easy access
E. Cloud - store files in the cloud to be accessible anywhere
    a. Editing data can be multi or single user editing
F. Control of big data - visualize multiple different types
G. Integrate enterprise - data stored in big business systems to extend their analytical capabilities
H. Data rules and relationships - define relationships between datasets and set rules (domains and subtypes)
I. Manage metadata - describes content, quality, origin, and other characteristics of data
J. Secures data - flexibility and control over how GIS platform is deployed, maintained, secured, and used
K. Versioning - allows multiple editors to edit one database by creating "duplicates" of the base data - changes are recorded with addition and deletion tables - versions can be created or deleted - edits are isolated in that version until admin merges changes - edits can be posted to parent version - DEFAULT is the root version

Additional Information:

http://srnr.arizona.edu/rnr/rnr417/lab2p.pdf

http://www.webopedia.com/TERM/F/file_management_system.html

http://www.esri.com/products/arcgis-capabilities/data-management

http://www.esri.com/news/arcuser/0110/versioning101.html

504. UNDERSTANDING OF DATA INTEGRATION

A. Combining data from multiple sources into one unified view
B. Integration happens with Extract, Transform, Load methodology (ETL)
    a. Extract data from the current source
    b. Transform the data into the current format
        i. For example, if there are multiple road shape files and they have different attribute tables, making all these attribute tables uniform is transform step.
    c. Load the data into the final storage point

Additional Information:

https://en.wikipedia.org/wiki/Data_integration

http://wiki.gis.com/wiki/index.php/Data_integration

ANALYTICAL METHODS

601. UNDERSTANDING OF DATA SELECTION QUERIES AND VIEWS

A. Attribute – select features based on filters of attributes
    a. New Selection, Add to Selection, Remove from Selection, Subset Selection, Switch Selection, Clear Selection
B. Spatial -
    a. Intersect
    b. Within a distance
    c. Contains - features contain an input polygon (input polygon is selected)
    d. Completely contains - features must be completely in an input polygon (input polygon selected)
    e. Contains Clementini - features must be completely in the input polygon but if it's on the boundary, it will not be selected
    f. Within - features will be selected if inside a selecting polygon
    g. Completely within - features will be selected if completely within selecting polygon - no overlap
    h. Within Clementini - features will be selected and cannot be entirely on the boundary of the features
    i. Are Identical To - features are identical to input layer
    j. Boundary Touches - features will be selected if they have a boundary that touches a selecting features - must be completely inside or outside the polygon
    k. Share a Line Segment With - features selected if they share a line segment
    l. Crossed by the Outline of - Input features will be selected if they are crossed by the outline of a selecting feature
    m. Have their Center In - Features will be selected if their center falls within a selecting feature
    n. Contained by - Same as within
C. Location – Select features within a certain distance of a location
D. Understand basic SQL language, such as
    a. SELECT * FROM TABLE WHERE X=1
E. Spatial data selection functions
    a. Join - combine two attribute tables into one using a common key between tables
    b. Merge - combines multiple input datasets of the same data type into a single new output
    c. Append - combines datasets of same data type into an existing dataset
    d. Union - combines input features with another feature dataset
    e. Clip - extracts input features that overlay the clip features (keeps inputs attributes)
    f. Intersect - extracts features which overlap in all layers to new feature class (joins attribute tables)

Additional Information:

http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/geoprocessing/overlay_analysis.htm

602. UNDERSTANDING OF TECHNIQUES AND IMPLICATIONS OF DATA CLASSIFICATION

A. Classification - objects with similar symbols - up to 7 classes recommended but should stick to 5 - classes should be exhaustive (describe all possible values) and should not overlap (no value can fall into two classes)
   a. Equal range - equal distance between class breaks
   b. Quantiles - equal number of observations in each class
   c. Standard deviation - class breaks based on distance of standard deviation from the mean.
   d. Natural breaks - class breaks conform to gaps in data distribution
   e. Symbology - one layer can be symbolized by attribute

B. Implications of using different classification styles is that each will highlight different aspects of the trends in the dataset
   a. Choosing the wrong classification methodology could cause the data to hide important information or misinform the reader

Additional Information:

http://wiki.gis.com/wiki/index.php/Classification

https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/s10-03-data-classification.html

603. UNDERSTANDING OF ANALYTICAL OPERATIONS AND METHODS

A. Overlay analysis methodology (based on Boolean logic)
   a. Define problem
   b. Break problem into submodels
   c. Determine significant layers (some of these layers may need to be created)
   d. Reclassify or transform data within a layer

B. Spatial overlay - process of superimposing layers of geographic data that cover the same area to study the relationship between them

C. Overlay - two or more maps or layers are superimposed for showing relationships between features

D. Vector Overlay Tools
   a. Identity - Input features, split by overlay features
   b. Intersect - Only features common to all input layers
   c. Symmetrical Difference - Features common to either input layer or overlay, layer but not both
   d. Union - All input features
   e. Update - Input feature geometry replaced by update layer

E. Raster Overlay Tools

a. Zonal Statistics - Summarizes values in a raster layer by zones (categories) in another layer—for example, calculate the mean elevation for each vegetation category
b. Combine - Assigns a value to each cell in the output layer based on unique combinations of values from several input layers.
c. Single Output Map Algebra - Lets you combine multiple raster layers using an expression you enter—for example, you can add several ranked layers to create an overall ranking
d. Weighted Overlay - Automates the raster overlay process and lets you assign weights to each layer before adding (you can also specify equal influence to create an unweighted overlay)
e. Weighted Sum - Overlays several rasters multiplying each by their given weight and summing them together.

Additional Information:

http://resources.esri.com/help/9.3/arcgisdesktop/com/gp_toolref/geoprocessing/overlay_analysis.htm

http://resources.arcgis.com/EN/HELP/MAIN/10.1/index.html#/Understanding_overlay_analysis/009z0000000rs000000/

http://resources.arcgis.com/EN/HELP/MAIN/10.1/index.html#/How_Fuzzy_Membership_works/009z0000000rz000000/

http://resources.arcgis.com/EN/HELP/MAIN/10.1/index.html#/How_Fuzzy_Overlay_works/009z000000s0000000/

604. KNOWLEDGE OF MAP ALGEBRA

A. Map algebra – various functions performed on cells based on neighboring cells for raster datasets
B. Local operations - combine rasters that overlay each other
    a. Add/subtract/etc. the cells that are in the same location
C. Global operations – apply a formula to all cells
    a. Add/subtract/multiply/etc. all cells based on one value
    b. Find the distance from one cell to all cells
D. Focal operations – calculates a value based on all neighboring cells
    a. Find the average of all the cells around a certain cell
E. Zonal operations – computing a value based on cells in a particular zone (such as a watershed)

Additional Information:

https://en.wikipedia.org/wiki/Map_algebra

https://gisgeography.com/map-algebra-global-zonal-focal-local/

605. KNOWLEDGE OF DESCRIPTIVE AND SPATIAL STATISTICS

A. Descriptive statistics - discipline of quantitatively describing the main features of a collection of information
    a. Summarizes a sample to learn about the population
B. Summary statistics - used to summarize a set of observations
C. Coefficient of determination - R squared - number that indicates how well data fit a statistical model
    a. Fit to a line or curve
    b. 1 indicates the line fits perfectly with the data
    c. 0 indicates the line does not fit at all (data is random)

Additional Information:

https://en.wikipedia.org/wiki/Descriptive_statistics

https://en.wikipedia.org/wiki/Summary_statistics

https://en.wikipedia.org/wiki/Coefficient_of_determination

https://libraries.mit.edu/files/gis/spatialstatpresentation_iap2012.pdf

DATABASE DESIGN AND MANAGEMENT

701. UNDERSTANDING OF RELATIONSHIPS AMONG DATABASE OBJECTS

A. Schema - structure or design of the database or database object (table, view, index, stored procedure, trigger) - defines the tables, fields in each table, relationships between fields - a schema will include information on which fields have domains and what those domains are
B. Data dictionary - catalog or table containing information about the datasets stored in a database
C. Domain - the range of values for a particular metadata element
D. Attribute domain - enforces data integrity, identify what values are allowed in a field in a feature class
E. Coded value domain - attribute domain that defines a set of permissible values for an attribute in a geodatabase - it has a code and its equivalent value
F. Range domain - type of attribute domain that defines the range of permissible values for a numeric attribute
G. Spatial domain - allowable range for x, y coordinates and for m, z values
H. Tables - collection of related data held in structured format within a database, contains fields and rows
I. Views - result set of a stored query on the data - users can query - virtual table computed dynamically from data when the view is accessed
J. Sequences - ordered collection of objects in which repetitions are allowed (finite or infinite) number of elements is the length of the sequence
K. Synonyms - Alias or alternate name for a table, view, sequence or other object
L. Indexes - data structure that improves the speed of data retrieval operations in a database table
    a. Causes more storage space and additional writes
    b. Quickly locate data in the database
    c. Indexes can be on multiple columns
M. Clusters – Can either be
    d. Multiple servers share one storage
        i. This is typically used to handle user load balancing
    e. Databases distributed to different servers using replication
        i. This is typically used if you have multiple users utilizing the same data in different physical locations
        ii. There is a master database that the replica databases sync between
N. Database Links – data stored in a different database but accessible by to the database currently being accessed
O. Snapshot - state of a system at a particular point in time - can be a backup
P. Procedure - subroutine available to applications that access a relational database system (data validation, access control mechanisms)

Q. Trigger - procedural code automatically executed in response to certain events on a particular table or view in a database
R. Functions (subroutine) - sequence of program instructions that perform a specific task
S. Package - built from source with one of the available package management systems
T. Non-schema objects - users, roles, contexts, directory objects

Additional Information:

https://en.wikipedia.org/wiki/Database_schema

702. UNDERSTANDING OF DATABASE DESIGN

A. Database design - process of producing a detailed data model of a database
B. Design process -
    a. Conceptual schema - Determine where relationships and dependency is within the data.
    b. Logical Data Model - Arrange data in a logical structure that can be mapped into the storage objects supported by the database management system
    c. Physical database design
        i. Physical configuration of the database on the storage media
        ii. Detailed specification of data elements, data types, indexing options, and other parameters residing in the DBMS data dictionary
        iii. Modules, hardware, software
C. Field Types – the proper field type will secure data and make databases more efficient
    a. Short integer - between -32768 and 32768
    b. Long integer - between -2,147,483,648 and 2147483647
    c. Float - single-precision floating-point numbers
    d. Double- double-precision floating-point numbers
    e. Text - could be a coded value - assign to an integer through a domain
    f. Dates – a calendar date and sometimes a time is associated
    g. BLOBs - data stored as a long sequence of binary numbers - ArcGIS stores annotation and dimensions as BLOBs - images, multimedia, bits of code
    h. Object Identifiers - Unique IDs and FIDs
    i. Global Identifiers - Global ID and GUID - data types store registry style strings consisting of 36 characters enclosed in curly brackets
    j. Raster field types - raster can be stored within the geodatabase
    k. Geometry - point, line, polygon, multipoint, multipatch

Additional Information:

http://wiki.gis.com/wiki/index.php/Database_design

http://webhelp.esri.com/arcgisserver/9.3/java/index.htm#geodatabases/an_over-776141322.htm

703. KNOWLEDGE OF DATABASE MANAGEMENT AND ADMINISTRATION

A. Basic tasks
    a. Backup and recover databases - test backup and recovery plan, ensure backups done on schedule
    b. Database security - prevent hackers, security models, tasks - authentication, authorization, auditing (making sure the right people have the right access)
    c. Storage and capacity planning - disk storage is needed and monitor disk space and watch growth trends
    d. Performance monitoring and tuning - identify bottlenecks, tuning (indexing, queries on speed of return, right monitoring tools, capacity of server hardware)
    e. Troubleshooting - quickly ascertain problem and correct it
    f. Other
        i. High availability and ETL functions
        ii. Data extraction, transformation, and loading
B. Archiving - captures, manages, and analyzes data changes
    a. most often done with geodatabases
C. Retrieval – extracting data from a backup due to data loss or data corruption

Additional Information:

https://desktop.arcgis.com/en/desktop/latest/manage-data/gdbs-in-sql-server/geodatabase-administrator-sqlserver.htm

http://www.wiki.gis.com/wiki/index.php/GIS_Data_Administration

704. KNOWLEDGE OF DATA SECURITY

A. Data owner - user who creates tables, feature classes owns those datasets.
B. User access - database must verify the user accounts that connect to it.
    a. Administrator – Full control of the database, can read, create, update, delete features,
        i. Can create and delete feature classes, tables
    b. Editor – can read, update, create, and delete features
    c. Reader – can only view data
    d. Creator – can create additional feature classes, tables, as well as read, update, create, and delete features
C. Authentication - database checks the list of users to make sure a user is allowed to make a connection
    a. Operating System (OS) authentication
    b. Database Authentication
D. Groups - grant users based on their common functions.
E. Public role - right granted to anyone connected to database

Additional Information:

https://desktop.arcgis.com/en/desktop/latest/manage-data/gdbs-in-sql-server/user-accounts-groups.htm http://www.techrepublic.com/blog/the-enterprise-cloud/what-does-a-dba-do-all-day/

APPLICATION DEVELOPMENT

801. KNOWLEDGE OF DATA TRANSFER PROTOCOLS

A. Knowledge of data exchange procedures - transfer constructs (based on data models): (1) logical constructs solely pertaining to this standard, (2) constructs relating to the implementation method, and (3) constructs solely pertaining to the transfer media.
   a. File based transfer - data is in a structured file format.
   b. Application Programming Interface (API) - data is accessed and exchanged as needed between software systems
   c. Web services - data is accessed and exchanged over networks and the internet between software components, using http and other web based protocols.

B. Transfer of data between a client server and an end user

C. SSL – secure sockets layer

D. TLS – transport layer security (latest version of SSL)

E. SSL and TSL are used to encrypt data

F. Network transport protocols
   a. Communication packet is constructed at different intervals
      i. Transmission Control Protocol – header package for the data at the transport layer
      ii. Internet Protocol – header is added at the internet layer
      iii. Media Access Control (MAC) address – added at the physical network layer
      iv. Transferred from the host to the receiver

G. Network File Services (NFS)

H. Common Internet File System (CIFS)

I. HTTP Protocols – Hypertext Transfer Protocol – standard web transmission protocol – a way for delivering map images or map data to web browsers

Additional Information:

https://en.wikipedia.org/wiki/Comparison_of_file_transfer_protocols

https://en.wikipedia.org/wiki/File_Transfer_Protocol

https://www.sciencedirect.com/topics/computer-science/data-transfer-protocol

http://wiki.gis.com/wiki/index.php/Network_Communications

802. KNOWLEDGE OF CODING, SCRIPTING, AND MODELING BASICS

A. Scripting – interpreted by the computer rather than compiled
   a. Used to manipulate, customize, and automate existing software

B. Object oriented programming (OOP) - programming paradigm based on concept of "objects" which are data structures that contain data in the form of fields (aka attributes) and code in the form of procedures (aka methods) - most common are class based

C. Extensibility - system design principle where the implementation takes future growth into consideration - level of effort to extend the system and implement the extension
D. Query expressions – select a subset of features or records
E. Scripting basics
    a. Expression – most basic programming instruction
        i. Contain values and operators that can reduce to a single value
    b. Variables – value that can change depending on the program or information passed to the program
        i. i.e. x = 2
        ii. x is the variable, the value 2 is assigned
    c. Iterations – For loops, do while loops, do until loops
        i. Repeating a process to generate an outcome
        ii. For x in range(0, 3):
        iii. In Python, this would take the variable x from 0 to 2, running it 3 times
    d. Condition statements
        i. If – a conditional expression that is either true or false
        ii. Else – combined with an if statement and if that statement is false, defaults to the else condition
        iii. Elif – check if a different condition from the first "if" is true
F. Query Basics
    a. The following expression is a basic SQL query expression.
    b. SELECT * FROM Table WHERE "EXRPESSION"
        i. SELECT – determines what the operation will be
        ii. * - indicates the fields that should be returned
            1. * denotes that all fields should be returned
        iii. FROM Table – indicates which table or feature class the data should be selected from
        iv. WHERE "EXPRESSION" – is a filter on the table or feature to indicate what records should be returned
    c. Esri's ArcGIS Desktop and QGIS typically only request users to only fill out the "EXPRESSION" portion
        i. Such as "Street = 'MAIN'

Additional Information:

https://en.wikipedia.org/wiki/Object-oriented_programming

https://en.wikipedia.org/wiki/Extensibility

https://automatetheboringstuff.com/chapter1/

http://desktop.arcgis.com/en/arcmap/10.3/map/working-with-layers/building-a-query-expression.htm

803. AWARENESS OF BASIC APPLICATION DEVELOPMENT METHODS

A. Types of application development methods

a. Agile – developing software in mini-increments of the new functionality – i.e. 2 week sprints
    i. Scrum, crystal, extreme programming (XP) and feature-driven development are types of Agile
b. DevOps – organizational changes that enhances collaboration between departments responsible for different segments of the development life cycle (development, quality assurance, and operations)
c. Waterfall – Rigid linear model that contains sequential phases (requirements, design, implementation, verification, maintenance) – each phase must be complete before moving to the next phase
d. Rapid application development (RAD) – condensed development process that produces high quality system with low investment costs
e. Spiral development – combination of waterfall and RAD – combine the advantages of top-down and bottom-up concepts

Additional Information:

https://www.synopsys.com/blogs/software-security/top-4-software-development-methodologies/

https://en.wikipedia.org/wiki/Software_development_process

SYSTEMS DESIGN AND MANAGEMENT

901. KNOWLEDGE OF SYSTEMS ARCHITECTURE AND DESIGN, INCLUDING VARIOUS GIS SOFTWARES, PLATFORMS, AND ENVIRONMENTS

A. Architecture design
   a. Requirements Phase - user needs assessment and workflow loads analysis (baseline and peak traffic)
   b. Design Phase - Infrastructure requirements, network communication capacity, hardware and software procurement, software development and data acquisition must be identified
   c. Construction Phase - system procurement, data acquisition and database design, authorization for application design and development, prototype testing
   d. Implementation Phase - Initial deployment and operational testing, final system delivery, user training, system maintenance operations
   e. Capacity Planning Tool (CPT) - developed as a framework to promote successful GIS system design and implementation

B. Enterprise environments
   a. Enterprise GIS Environment - broad spectrum of technology integration of enterprise technologies connected by local area networks, wide area networks, internet communications
   b. Enterprise technologies - database servers, storage area networks, windows terminal servers, web servers, map servers, desktop clients

C. Platforms
   a. Desktop - individual user on a computer, make maps, data analysis, data creation
   b. Server - bring GIS into hands of everyone in organization, allows access to web GIS, control of GIS data on your own infrastructure, control over how GIS platform is deployed, maintained, secured and used
   c. Hosted (cloud) - ability to discover, use, make, and share maps with any device anywhere, anytime - access other users maps and data - connect more people outside of organization and share latest maps, data, and ideas
   d. Enterprise GIS - integrated through entire organization so that a large number of users can manage, share, and use spatial data and related information to address a variety of needs, including data creation, modification, visualization, analysis, dissemination
   e. Enterprise GIS can utilize both hosted (cloud) and server but if organizational data is not stored in the cloud, only data is accessed from the cloud, then it's not quite an enterprise system

D. Software
   a. Software runs on a variety of hardware types
      i. Centralized servers to desktop computers
   b. Software may rely on DBMS type, OS type
   c. System infrastructure - hardware, software, and communication network
      i. Required information products and spatial and non-spatial data resources

ii. Essential spatial analysis, display, and reporting functions, needed data management resources, anticipated number of end users within the department
d. Virtualization – creation of a virtual machine that acts like a real computer with an operating system

E. Types of GIS Software – listed the most commonly known GIS software
    a. GRASS – developed by US Army Corps of Engineers – open source
    b. Esri – ArcMap, ArcGIS Pro
    c. QGIS
    d. MapInfo
    e. Smallworld

Additional Information:

http://lbd.udc.es/Repository/Thesis/432677332J.pdf

http://www.wiki.gis.com/wiki/index.php/System_Design_Strategies

http://www.wiki.gis.com/wiki/index.php/System_Implementation

http://www.esri.com/software/arcgis/arcgis-for-desktop

http://www.esri.com/software/arcgis/arcgisserver

http://www.esri.com/software/arcgis/arcgisonline

https://en.wikipedia.org/wiki/Enterprise_GIS

http://www.essex-countynj.org/GIS.pdf

http://www.esri.com/library/whitepapers/pdfs/enterprise-gis-localgov.pdf

https://en.wikipedia.org/wiki/Virtualization

https://en.wikipedia.org/wiki/List_of_geographic_information_systems_software

902. KNOWLEDGE OF SYSTEMS AND APPLICATION SECURITY

A. Application security - Making applications more secure by finding, fixing, and enhancing the security
B. Application security is important because it will keep the data safe
C. Continuous deployment and integration cause security fixes to be deployed constantly (daily or hourly)
D. Security testing
    a. Static testing – analyzes code at fixed points
    b. Dynamic testing – analyzes running code and can simulate an attack
    c. Interactive testing – combines static and dynamic testing
    d. Mobile testing – designed to examine how an attacker can attack mobile apps

E. Methods to prevent unauthorized access to data and metadata
    a. Define who has access
    b. Other software to enforce these policies
    c. Identify management systems to check the credentials of users
    d. Data authenticity verification

Additional Information:

https://www.csoonline.com/article/3315700/what-is-application-security-a-process-and-tools-for-securing-software.html

https://en.wikipedia.org/wiki/Application_security

https://www.gislounge.com/gis-security-applications/

https://gis.usc.edu/blog/gis-and-cybersecurity/

903. AWARENESS OF TRENDS IN GEOSPATIAL TECHNOLOGY

A. Be aware of where GIS is heading
    a. Read blogs
    b. Attend conferences
    c. Attend webinars
    d. Join local and online GIS groups
    e. Job postings for skills required
    f. Review LinkedIn profiles of similar GIS professionals
    g. Check out https://www.reddit.com/r/gis/
B. Some examples
    a. Artificial Intelligence in GIS data creation - https://www.esri.com/en-us/artificial-intelligence
    b. Machine learning in GIS - https://www.esri.com/en-us/arcgis/products/spatial-analytics-data-science/capabilities/machine-learning-ai
    c. IoT – Internet of Things - https://en.wikipedia.org/wiki/Internet_of_things
    d. Drones
    e. Mobile field data collection

https://www.esri.com/about/newsroom/arcnews/five-tech-trends-driving-new-geospatial-development/

PROFESSIONAL PRACTICE

1001. UNDERSTANDING OF APPROPRIATE INTERPRETATION OF WORK-RELATED POLICIES AND PROCEDURES

Each company has different policies and procedures for editing, storing, and archiving data, and those policies should be understood and followed.

1002. UNDERSTANDING OF ETHICS RELATED TO TECHNICAL GIS WORK

Read GISCI's code of ethics: https://www.gisci.org/Ethics/RulesofConduct.aspx

URISA's code of ethics (similar to GISCI's): https://www.urisa.org/about-us/gis-code-of-ethics/

Here are their four primary categories:

I. Obligations to Society

II. Obligations to Employers and Funders

III. Obligations to Colleagues and the Profession

IV. Obligations to Individuals in Society

Additional Information:

https://www.gislounge.com/need-ethics-gis/

1003. KNOWLEDGE OF MANAGING, DOCUMENTING, AND COMMUNICATING GIS WORK

There is a significant amount of GIS data being created and modified on constant basis. Following procedures to manage, document, and communicate this work is extremely important. A few ideas are to:

1004. AWARENESS OF HOW GIS IS USED ACROSS OTHER PROFESSIONS

GIS skills and jobs are vastly varying between professions, companies, and industries.

A. Read other LinkedIn profiles to get a sense of what skills others have in GIS
B. Attend conferences and webinars
C. Join a local group and discuss with other GIS professionals what tools they use and what kind of work they perform
D. Read blogs and social networking

1005. AWARENESS OF GIS-RELATED PROFESSIONAL ORGANIZATIONS AND CERTIFICATION

For individuals in the United States here are a few professional organizations:

URISA, GISCI, & American Association of Geographers (AAG)

https://www.gislounge.com/organizations/

There are many types of Professional GIS Certifications (outside of academic GIS certifications)

https://www.directionsmag.com/article/1822

MISCELLANEOUS

GPS

A. 24 satellites
B. Orbit earth twice a day
C. Revolution every 12 hours
D. Altitude of about 12,000 miles
E. Started by US Department of Defense in the 1970's for military
F. Space segment - NAVigation Satellite Timing and Ranging (NAVSTAR) constellation
    a. GPS satellites which transmit signals on two phase modulated frequencies
    b. Transmit a navigation message that contains orbital data for computing the positions of all satellites
G. Standard Positioning Service - signal broadcast for civilian use
    a. Horizontal location - 3 satellites are required
    b. Vertical position - min 4 satellites are required
H. Calculate distance by measuring the time interval between the transmission and reception of a satellite signal
I. Trilateration - used to determine position of the GPS receiver
J. Accuracy dependent on type of GPS receiver, field techniques, post processing of data, error from various sources
K. 3 types of GPS receivers - Recreational Grade, Mapping Grade, or Survey or High Accuracy Grade
L. GPS errors
    a. Multipath - errors caused by reflected GPS signals arriving at the GPS receiver - nearby structures and other reflective surfaces
    b. Atmosphere - GPS signals can experience delays when traveling through the atmosphere - Common atmospheric conditions can affect GPS signals such as tropospheric delays and ionospheric delays
    c. Distance from Base Station - differential correction will increase the quality of the data, accuracy is degraded slightly as the distance from the base station increases
    d. Selective Availability - intentional degradation of the GPS signals by the department of defense (DOD) to limit accuracy for non-U.S. military and government users - currently turned off, but can turn it back on whenever
M. Noise - error is the distortion of the satellite signal prior to reaching the GPS receiver and or additional signal piggy backing onto the GPS satellite signal
N. PDOP - Position Dilution of Precision - collect data when there is an optimum satellite availability (four or more) and when satellites are in an appropriate configuration to produce an acceptable (lower) PDOP value (higher PDOP values are bad)
    a. PDOP values - set to 6 or less. Higher levels will be less reliable data
O. Signal to Noise Ratio (SNR) mask - set the value of the SNR mask higher to help minimize noise error - user manufacturer recommendations

P. Elevation Mask - set it to 15 degrees - default angle to minimize the amount of atmosphere through which the satellite signal must travel

Q. Data Collection Rate (sync rate) - recommended to collect point data at 1-second interval - collect polygon and line data at a 5 second interval - collect point data at the same data collection interval as the base station

R. Datum - GPS receivers are designed to collect GPS positions relative to the WGS84 datum - can designate what datum to be used

S. GPS coordinates
    a. Latitude/Longitude - Degrees/Minutes/Seconds (DMS) 43° 5′ 20″
    b. Latitude/Longitude - Decimal Degrees (DD) 43.088889°
    c. Converting Degrees/Minutes/Seconds to Decimal Degrees
        i. There are 60 minutes in one degree
        ii. There are 60 seconds in one minute
            1. There are 3600 seconds in one degree
        iii. Latitude in the southern hemisphere is negative
        iv. Longitude in the western hemisphere is negative
    d. Latitude/Longitude - Degrees and decimal minutes - 43° 5.33333′
    e. UTM 18 - (4740283N, 434057E)
    f. State Plane - US feet - (312608N, 313525E)
    g. US National Grid - (18T WN 7125315437)

Additional Information:

http://support.esri.com/en/technical-article/000007058