Regional Committee of United Nations Global Geospatial Information Management for Africa

Workshop on Integration of Geospatial and Statistical Information

An Overview of the Concept of Geocoding

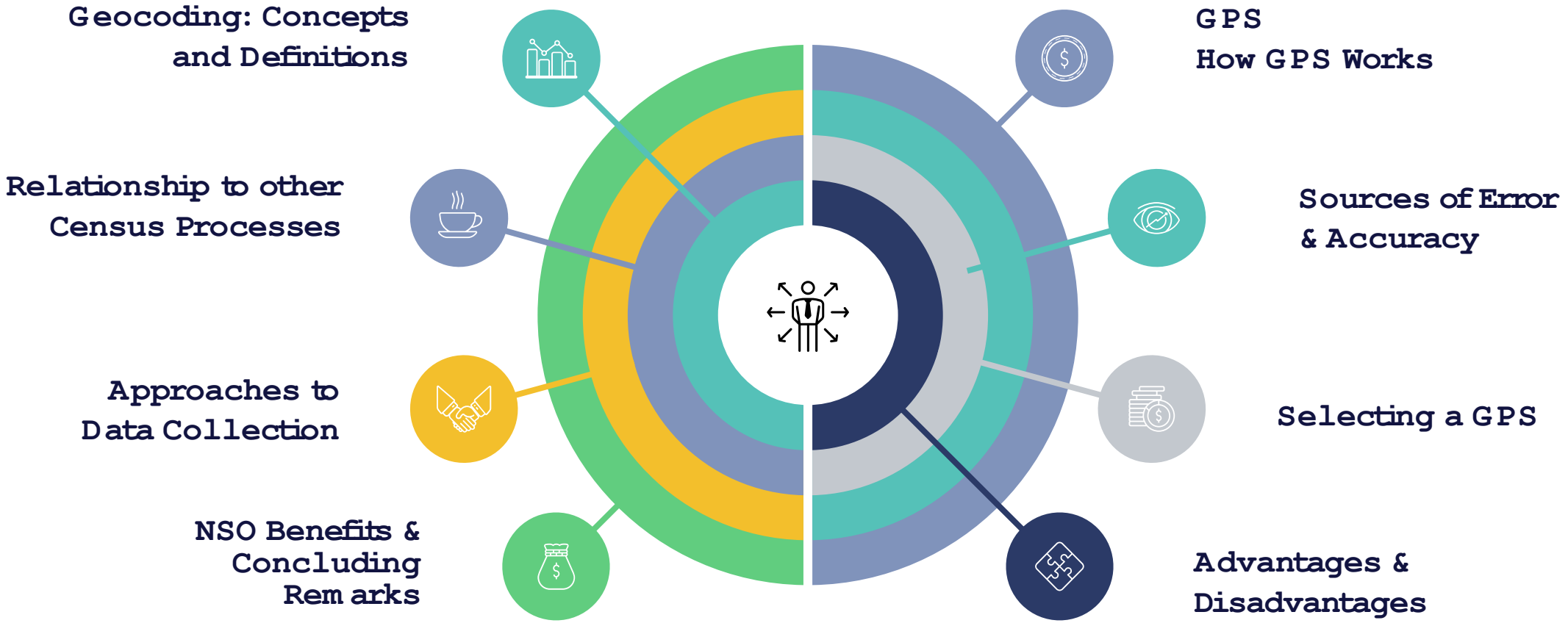
United Nations Economic Commission for Africa

Presented by:
Ayenika Godheart
ECA (Geospatial Information Management Section)

UN-GGIM: AFRICA
UNITED NATIONS INITIATIVE ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT

Presentation Summary

An Overview of the Concept of Geocoding in Population and Housing Censuses



Geocoding: Concepts and Definitions

Definitions

A Conceptual

B Operational

C Geocoding

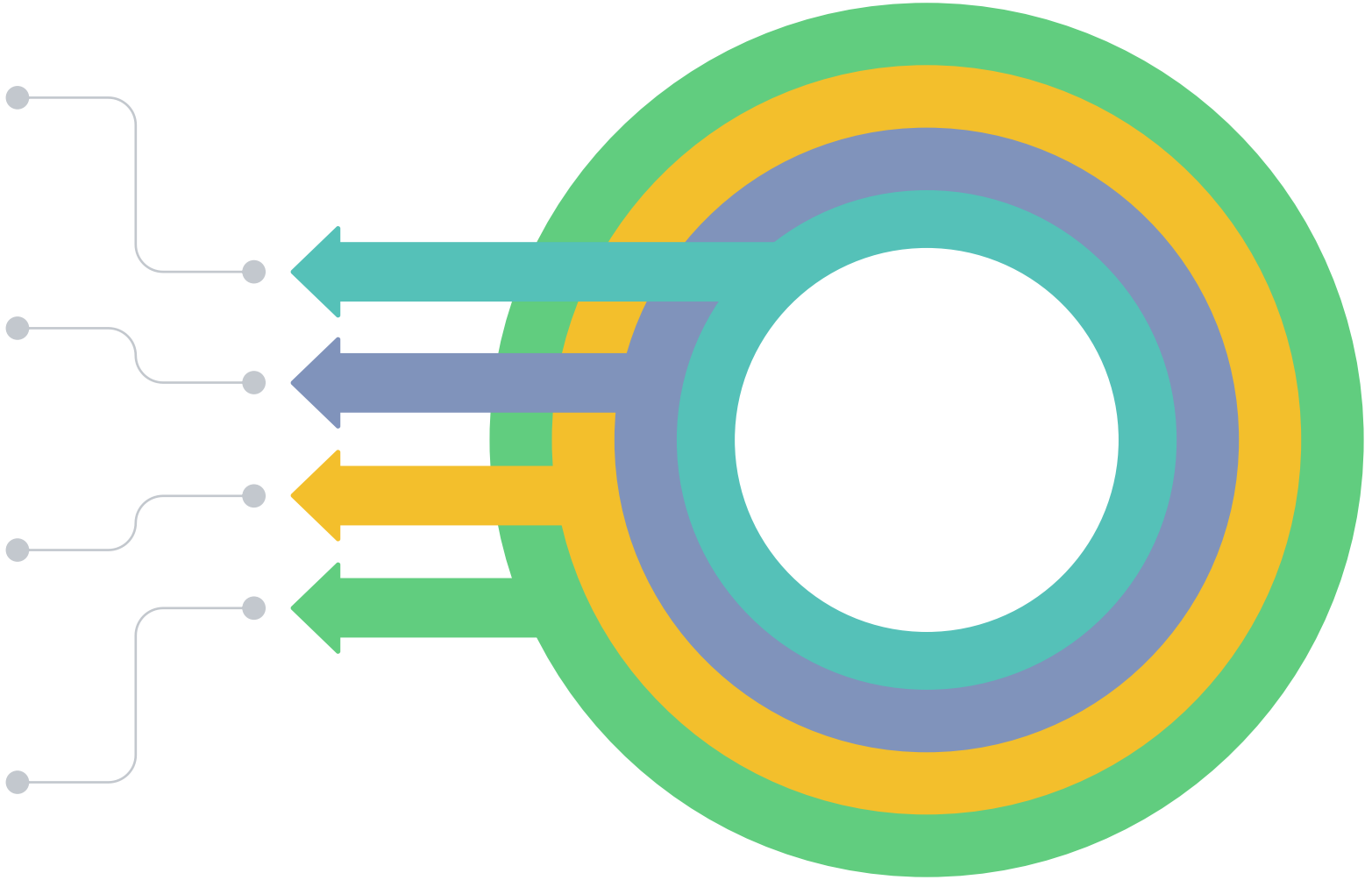
D Georeferencing

Geocoding: Concepts and Definitions

Geocoding can be broadly defined as the assignm code to a geographic location. Usually however, Geocoding refers to a m ore specific assignm geographic coordinates (latitude, Longitude) to an individual address.

The purpose of this section is to introduce geocoding concepts relevant for census m apping and the different approaches to related data collection.

Reference: UN Report of the Expert Group Meeting on Contem porary Practices in Census Mapping and Use of Geographical Inform ation System s (2007)



Definition of Geocoding

Conceptual - 2 situations

The more general process of assigning geographic codes to features in a digital database.

A GIS function

A GIS function that determines a point location based on an address. It could generally be expected that such point locations will be relatively precise (eg +/-2m) in accuracy and will be based upon use of GPS technology.

Operational

A Geocoding is the computer oriented process which converts information about a unit from which statistical information is collected into a set of coordinates describing the geographic position of that unit

B

C

D

Source UNSD, Workshop on Census Cartography and Management, Lusaka, Zambia, 08-12 October 2007

Definition of Geocoding (cont.)

Operational Elements

A Collecting precise data at the level of point locations (or very low geographic level such as a city block) and assigning codes for use in EAs, Spatial Analysis and dissemination.

B Coding the centroid, building corners, or building point of entry coordinates for a unit such as a block of land, building or dwelling

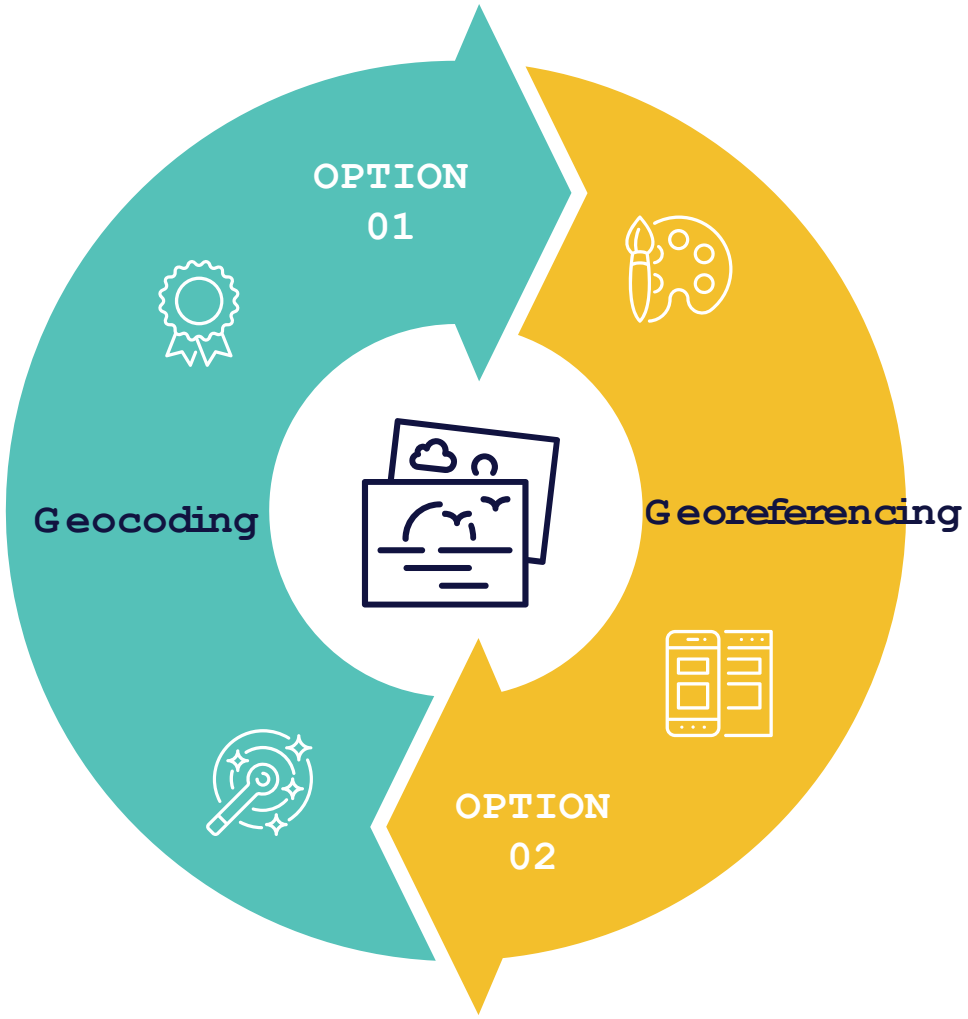
C Coordinates must contain latitude and longitude or standardized x and y points for gridded interpolation. A Z or Zed coordinate may represent altitude or elevation

D Codes cover each geographic unit and have a combinational relationship to distinguish different units (Enumeration Areas/Blocks)

Geocoding vs. Georeferencing

Geocoding

A GIS operation for converting street addresses into spatial data that can be displayed as features on a map



Georeferencing

Aligning geographic data to a known coordinate system so it can be analyzed, viewed, and queried with other geographic data

Census Enumeration & the Geocoding System

Delineation irrespective of the existence of address

Ability to apply a geocode to any geographic areal unit/Dwelling Frames

Flexible Coding Scheme

Ability to incorporate future administrative divisions

Pre-enumeration geocoding critical

links between GIS boundaries and tabular census data

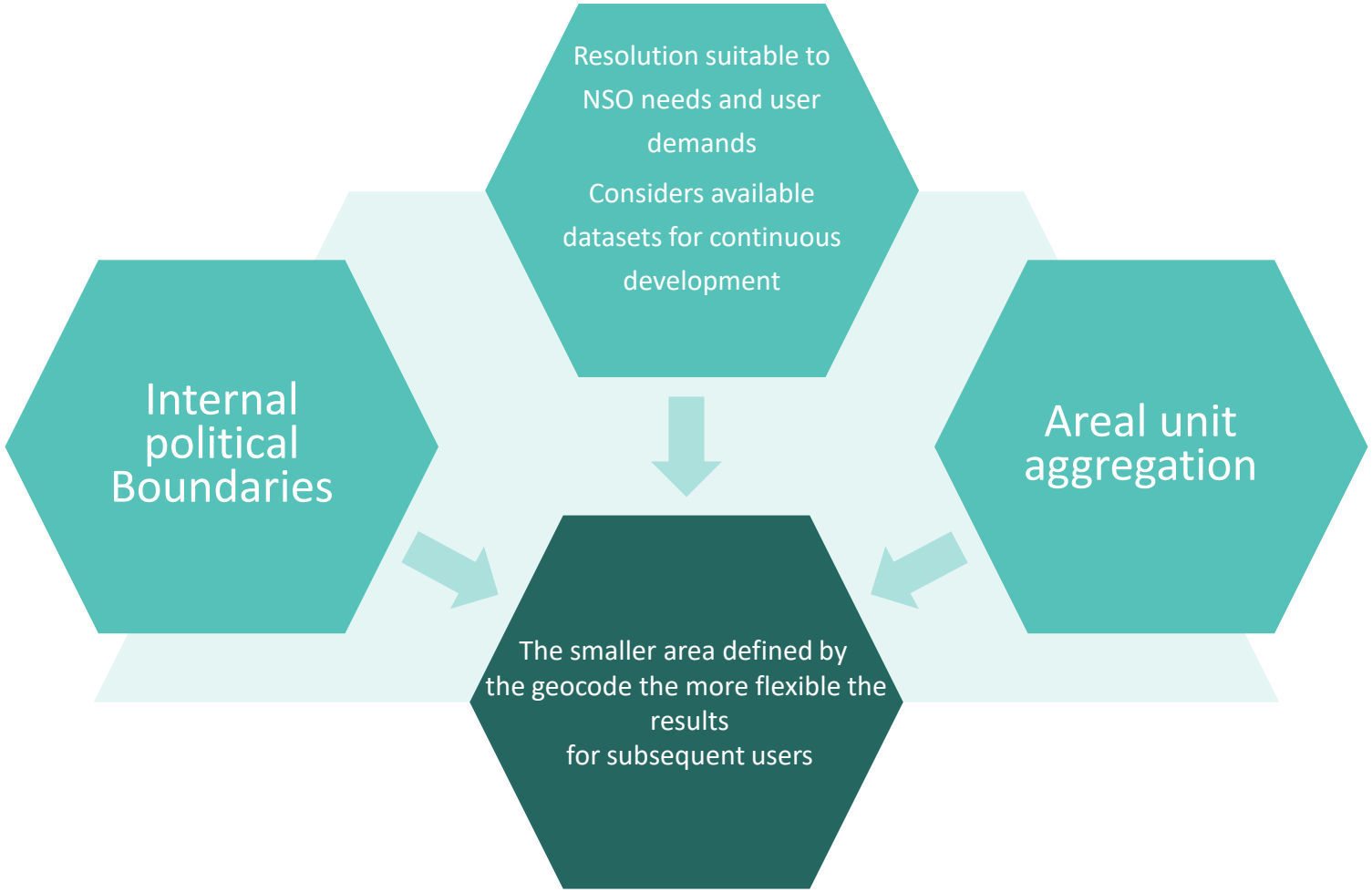
Geocoding During Census Enumeration

Ability to use mobile devices and geocode enumeration

Analysis of Geocoded data

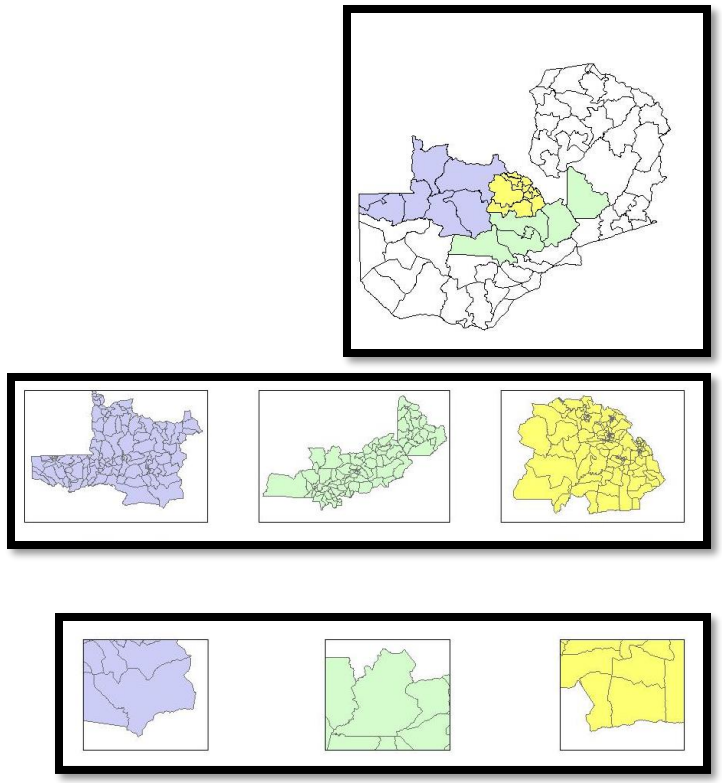
Ability to use geospatial techniques and analysis geocoded data

Census Hierarchies



Census Hierarchies

Given Country

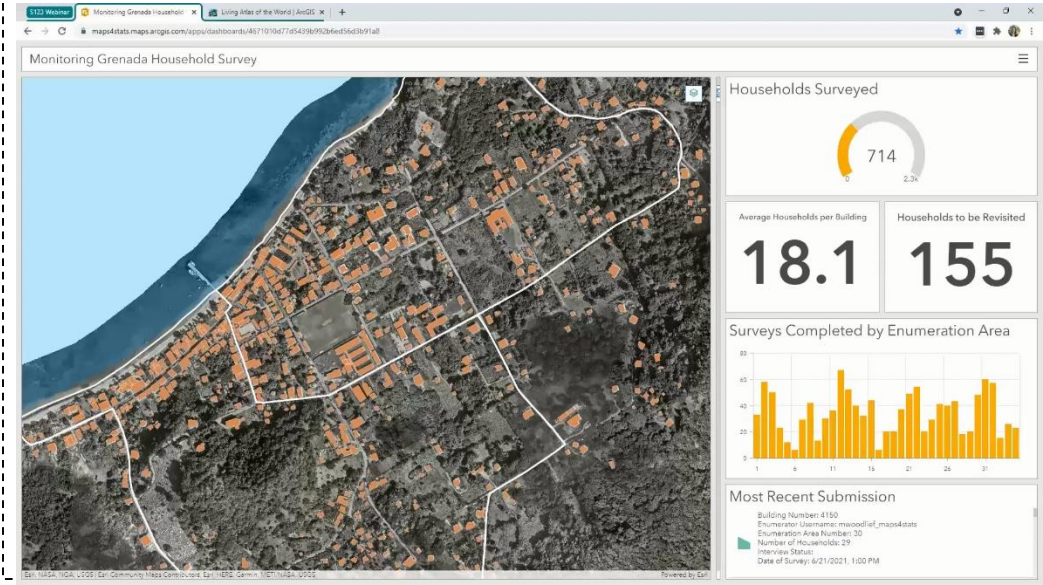


Country

Province

District

Locality

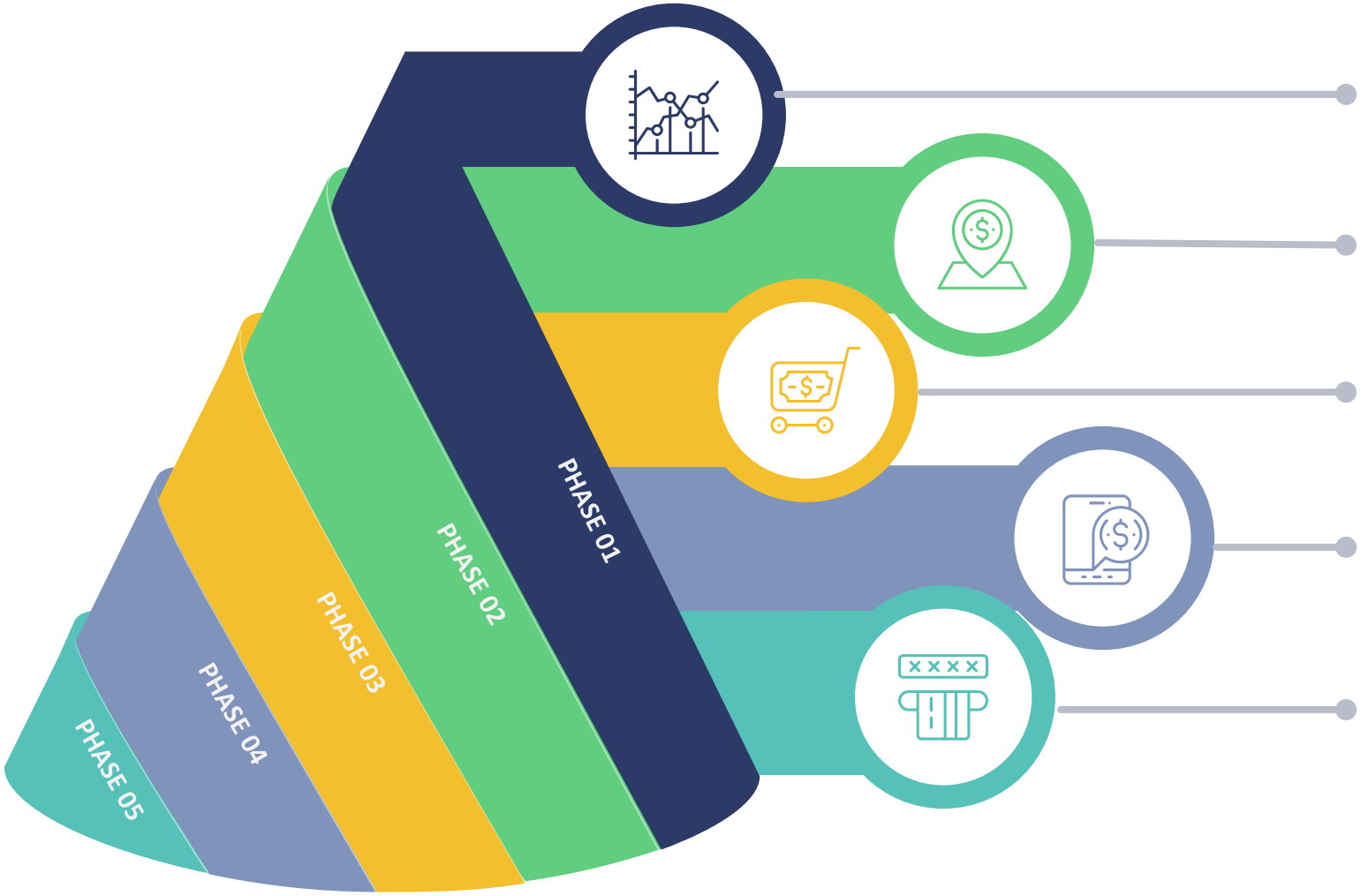


Enumeration Areas

Blocks Building Dwelling

Geocoding Classifications

Disaggregation into Spatial Entities or Civil Divisions and Compatibility



1st
Region, Province

2nd
District, Municipality

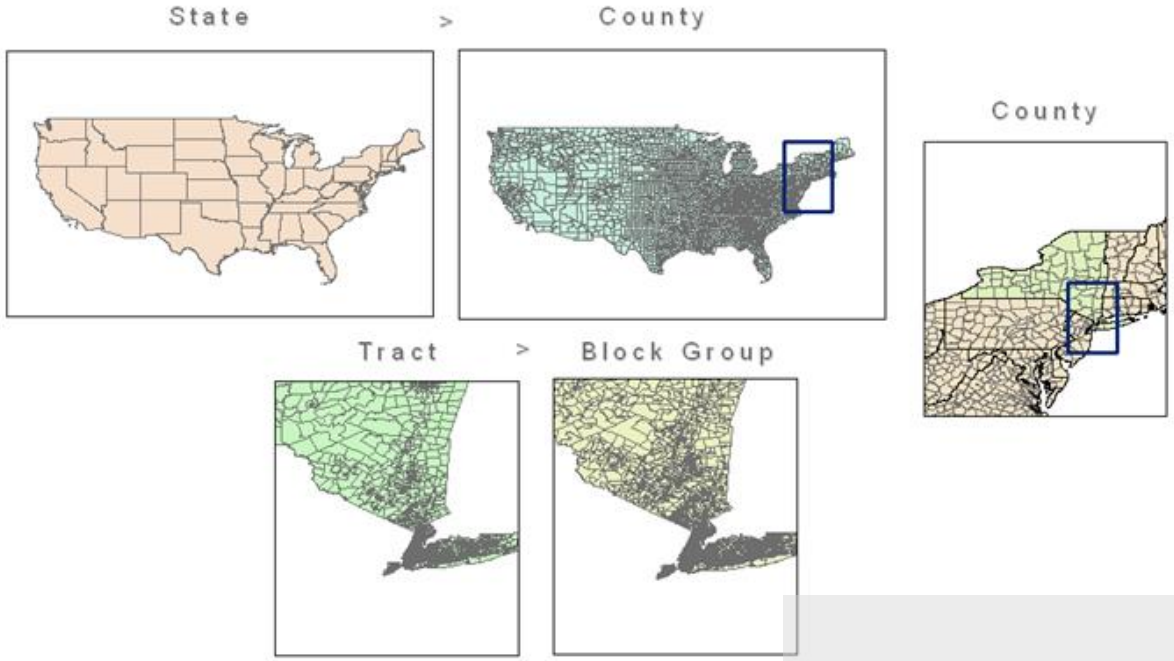
3rd
Town/Village

4th
Quarter/Locality

5th
Household, Dwelling

- Resultant geocoded units placed within a set of Latitude and Longitudinal boundaries
- Initial creation of Civil Divisions through digitizing or segmentation/pixel based-approaches
- Low to Zero levels of sampling through the accurate placing of coded units, but flexible

Coding Scheme



250131402013

Digits 1-2 = State code

Digits 3-5 = County Code

Digits 6-11 = Census Tract Code

Digit 12 = Blockgroup code

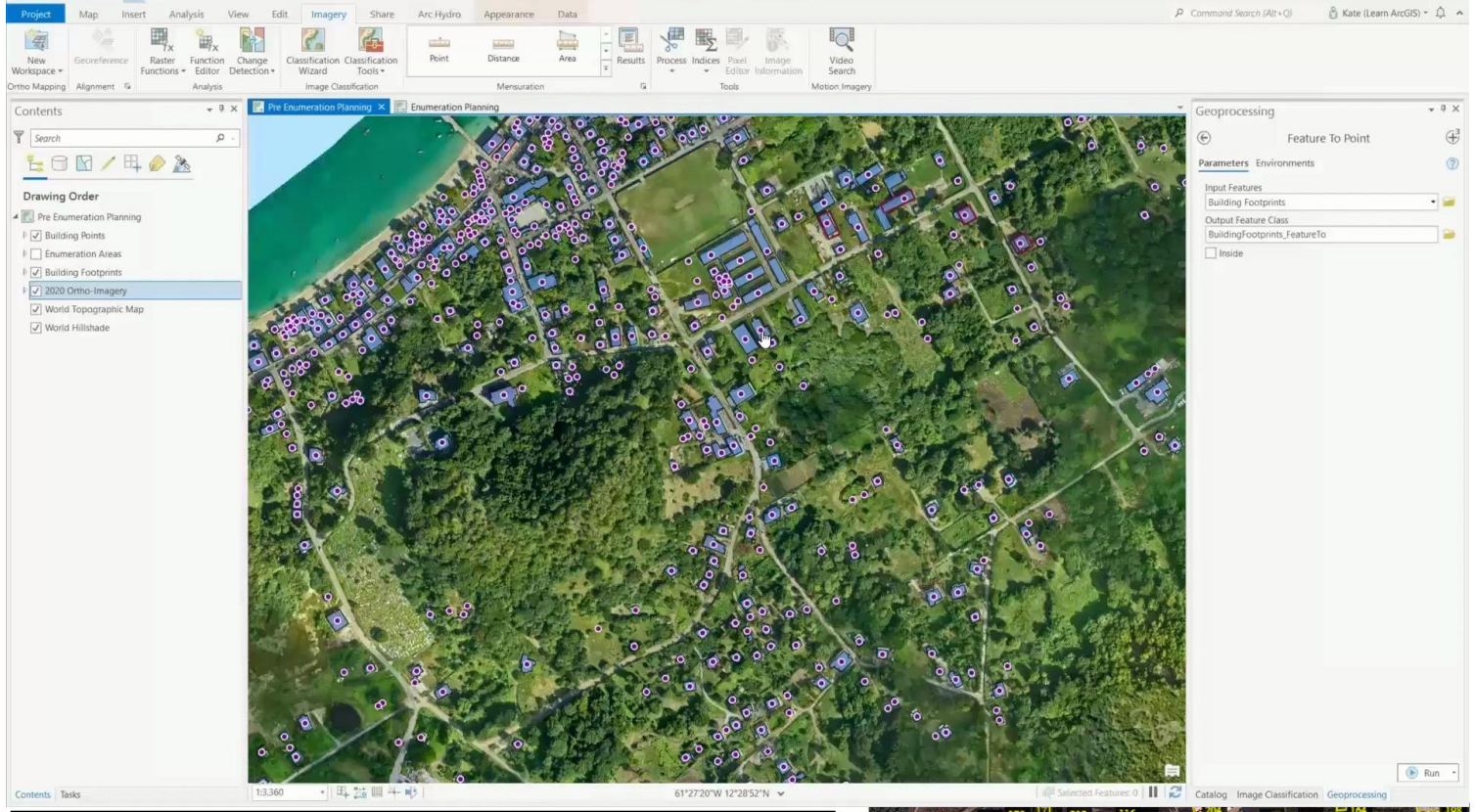
Data Collection Methods

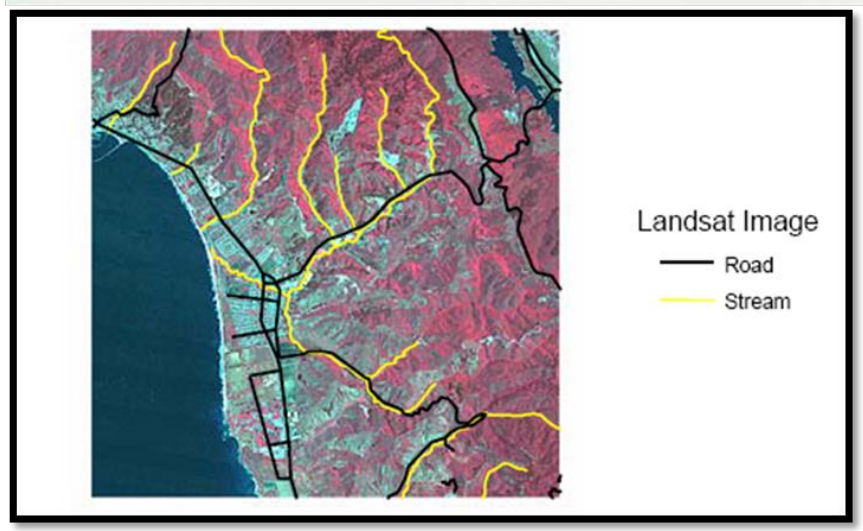
Two main methods:

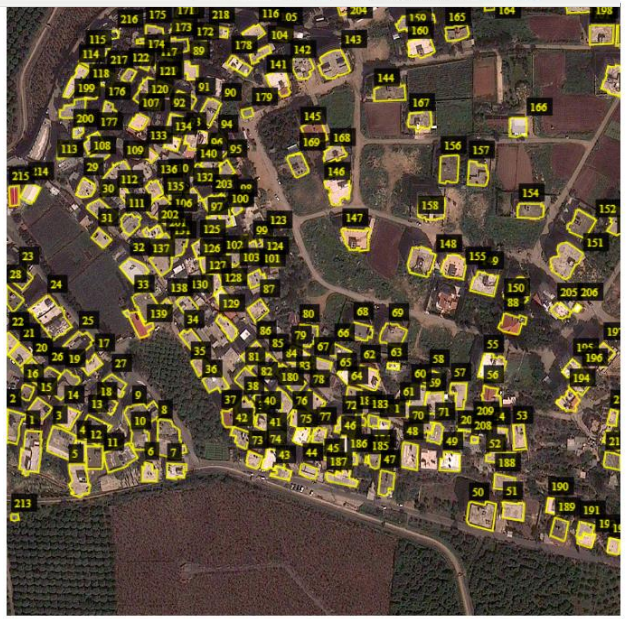
- Direct Collection Approach
- Matching Approach

Direct Collection Approach

- Digitizing from available topographic maps
- Direct collection using field techniques (ex.GPS)



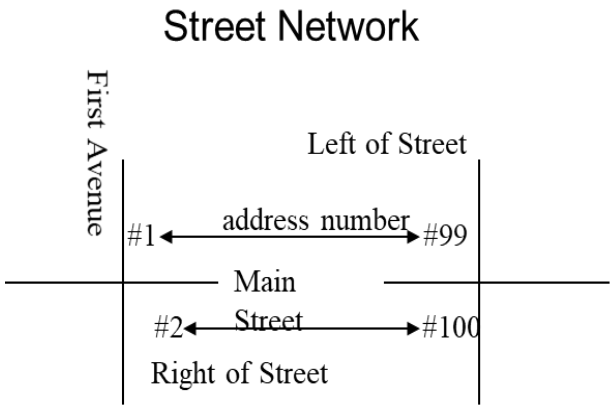


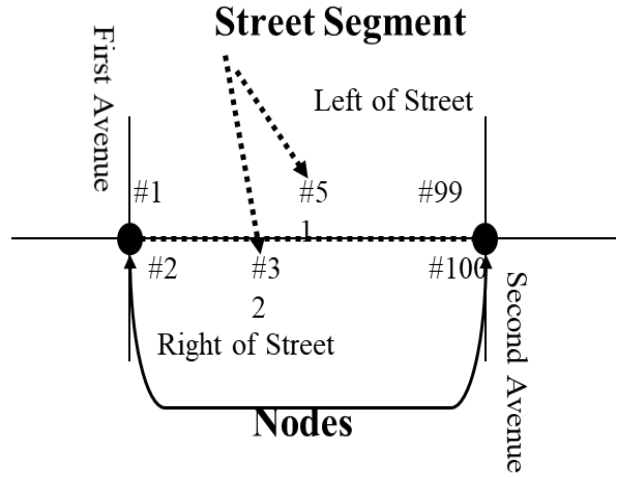


Matching approach

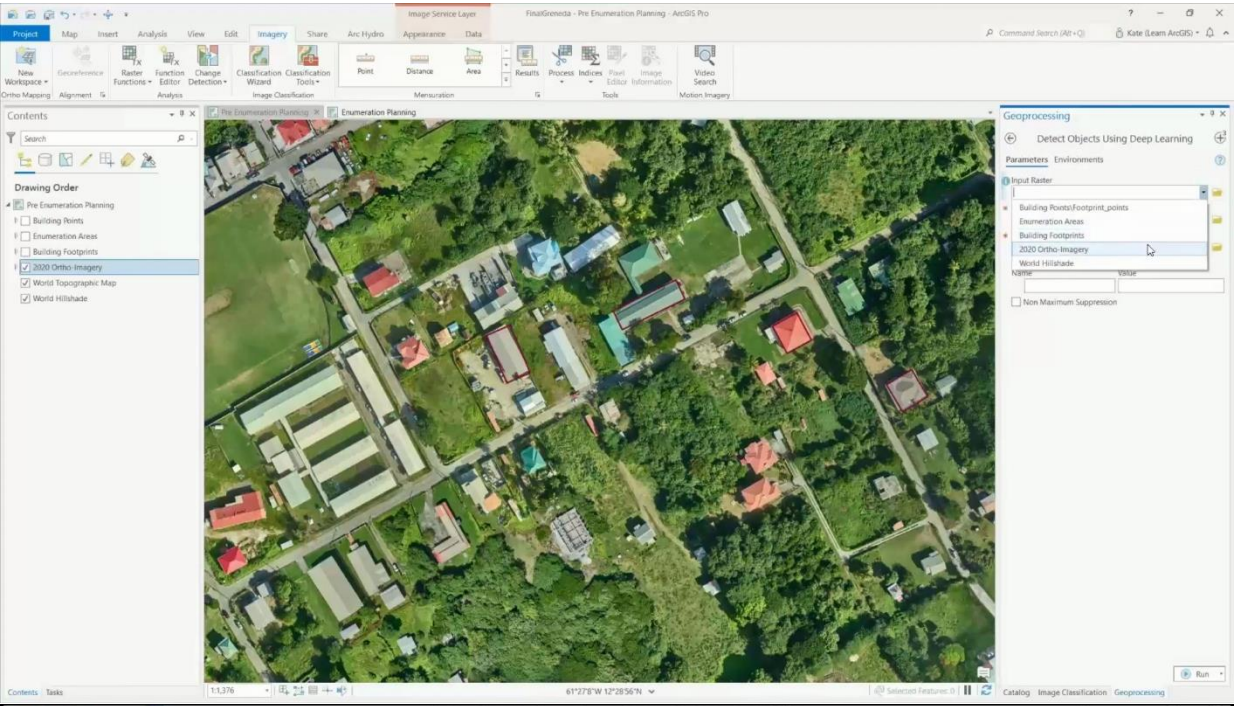
Using an Address locator database and street network database in a GI

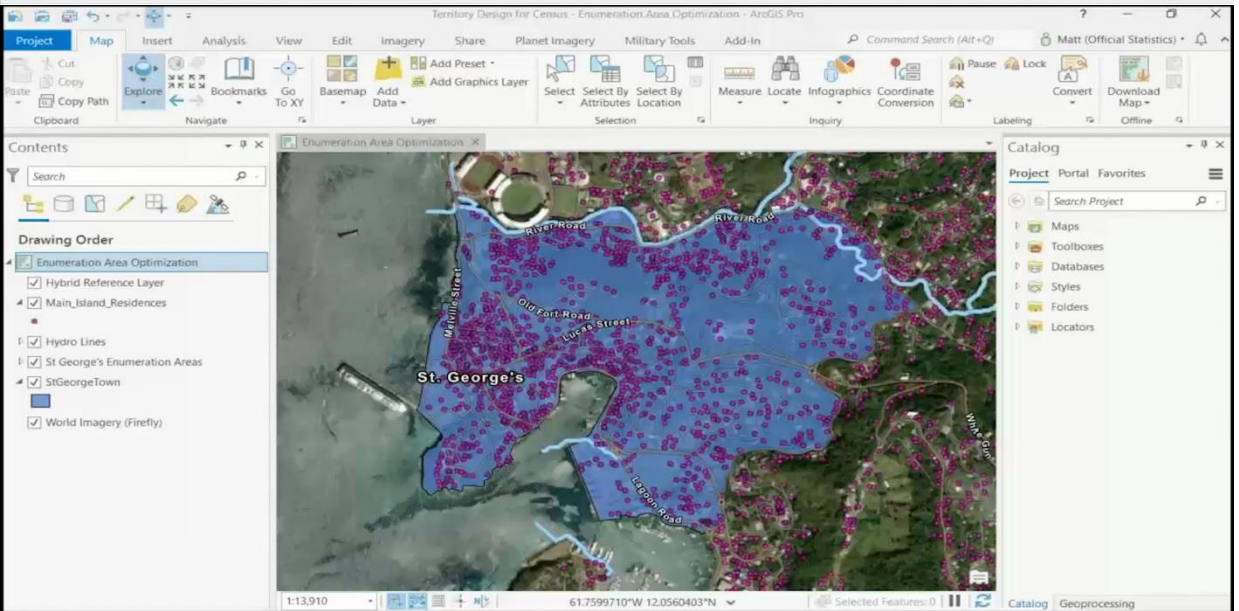
Joining an address database to an existing spatial database for the a of interest





Data Maintenance





- Cleaning Addresses
  - Retaining only the key address elements
    - Establish a Matchcode (indicator of which address elements will determine the geocode)

Record	Street Address	City	State	ZIPcode	Latitude	Longitude	Areakey	MatchCode
1	344 East 63rd	New York	NY	10023	40.47	73.58	3502508100	**AS0**

  - Eliminating extraneous characters
  - Standardizing Spelling

Global Positioning Systems (GPS)

Identify

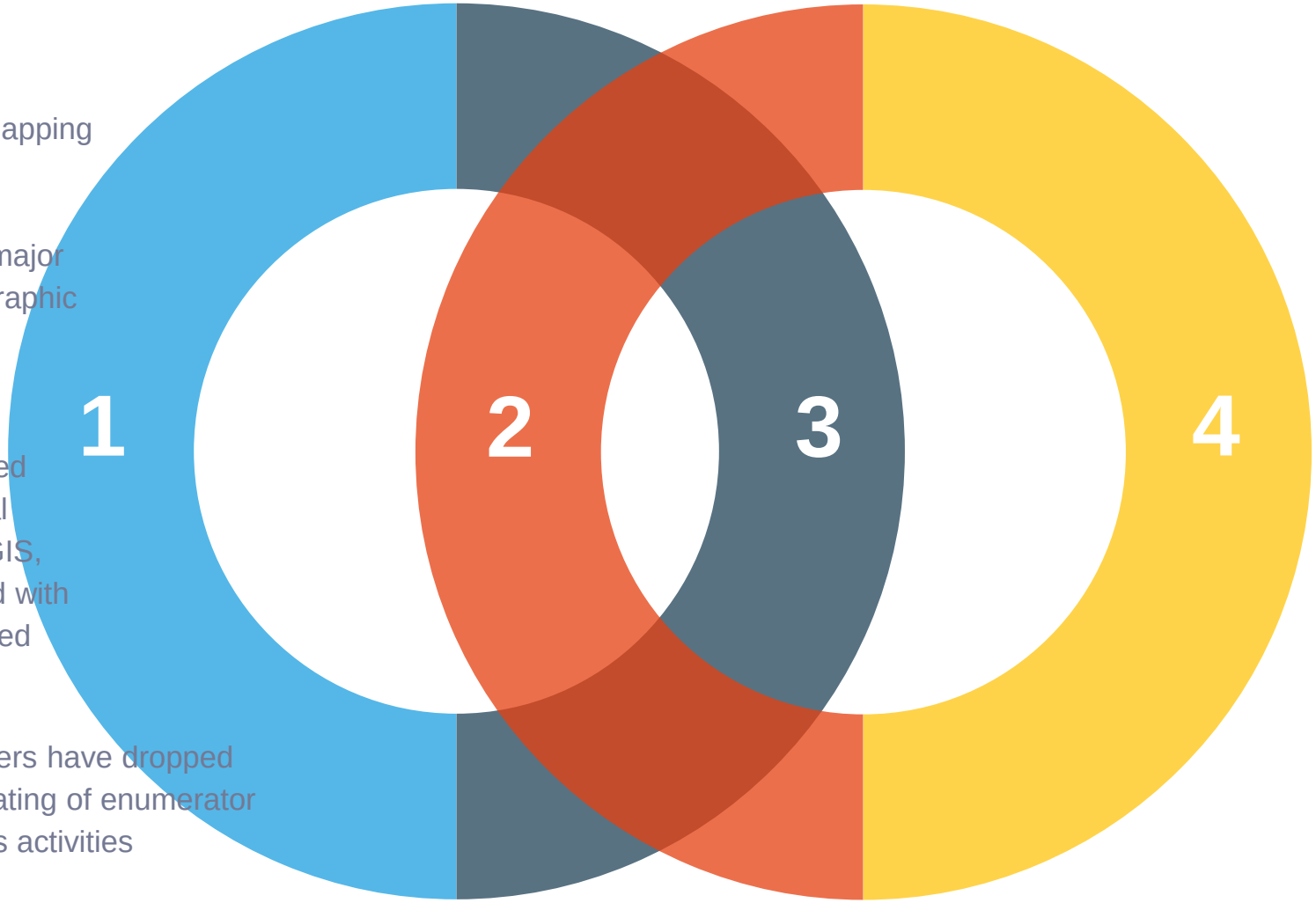
Technology has revolutionized field mapping in recent years

GPS has become a major tool in census cartographic applications

Coordinates can be downloaded or entered manually into a digital mapping system or GIS, and can be combined with existing, georeferenced information

Improve

Prices of GPS receivers have dropped
Preparation and updating of enumerator (EA) maps for census activities



Innovation

GPS methods have been integrated in many applications

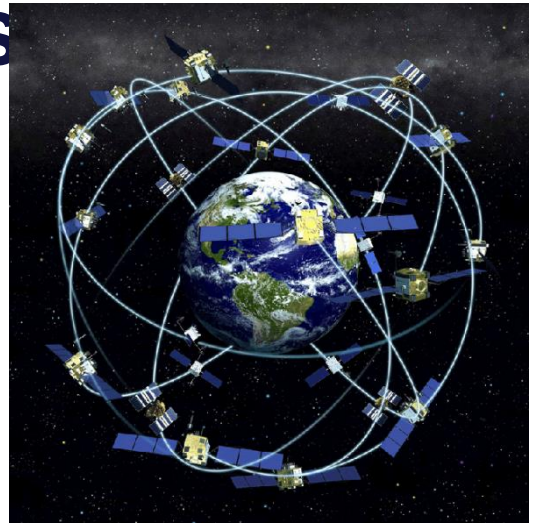
User groups are widespread (utilities management, surveying and navigation)

Location of point features such as service facilities or village centers

Management

GPS has contributed and advanced to improve field research in areas such as biology, forestry, geology, epidemiology and population studies

How GPS Works

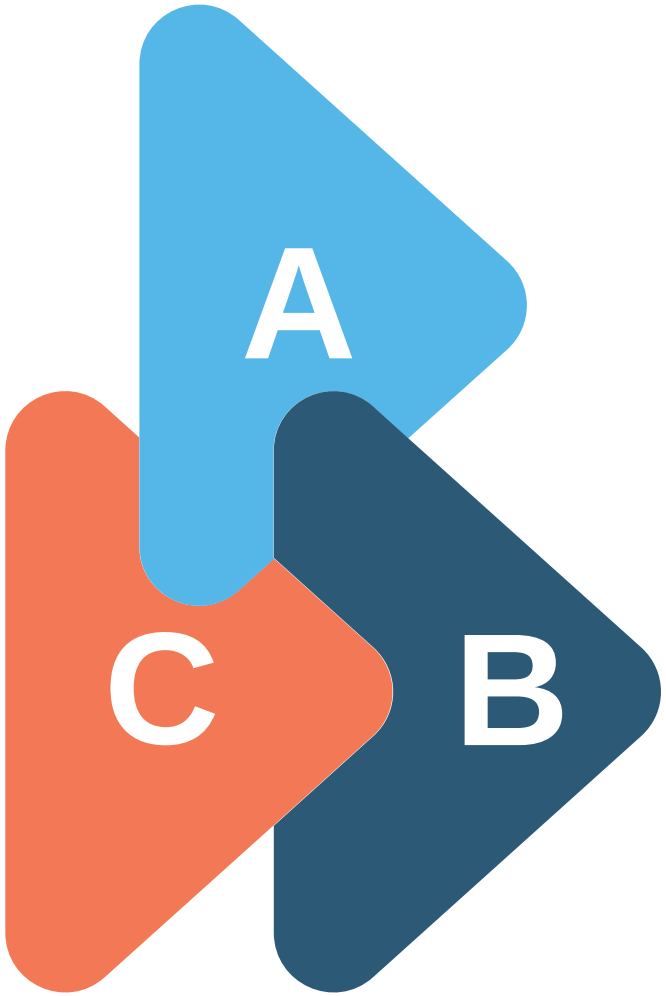


Identify

GPS receivers collect the signals transmitted from more than 24 satellites—21 active satellites and three spares. The system is called NAVSTAR, and is maintained by the U.S. Department of Defense

Improve

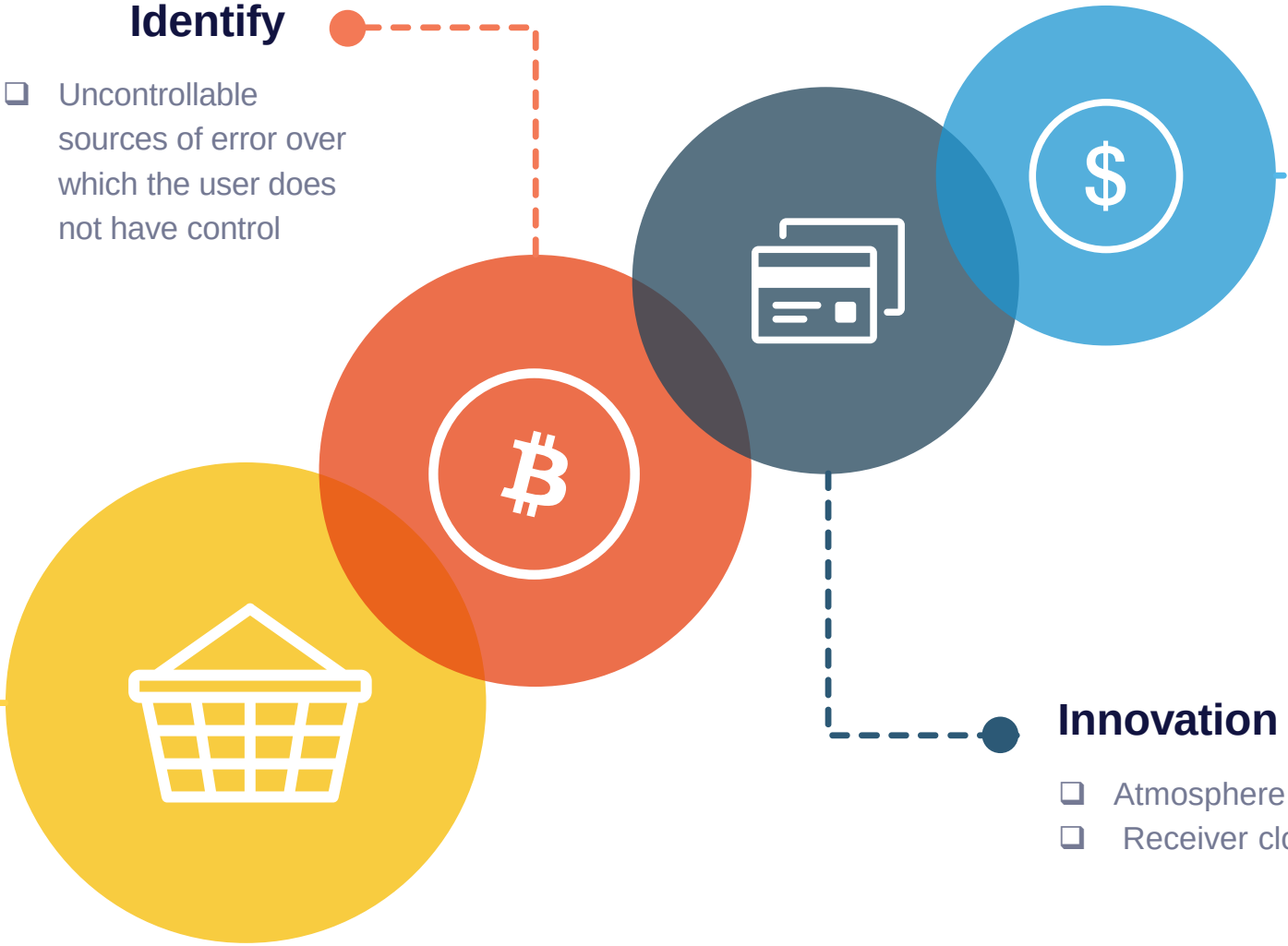
The satellites are circling the earth in six orbital planes at an altitude of approximately 20,000 km. At any given time five to eight GPS satellites are within the “field of view” of a user on the earth’s surface



Management

The position on the earth’s surface is determined by measuring the distance from several satellites

Sources of GPS Signal Errors



Identify

- Uncontrollable sources of error over which the user does not have control

Improve

- Good visibility and bad visibility of satellites due to obstacles

Management

- Signal multipath
- Orbital errors

Innovation

- Atmosphere delays
- Receiver clock errors

GPS Accuracy

Identify

- Inexpensive GPS receivers
- In dense urban settings, the possible error of standard GPS (standard ~15m up to 100 meters) may not be sufficient

Management

- Within 15 to 100 meters for civilian applications

Innovation

- Differential GPS reduces error further

Improve

- Accuracy of about 3-10m can be achieved with quite affordable hardware and shorter observation times.

Client

- More expensive systems and longer data collection for each coordinate reading can yield sub-meter accuracy.
- Differential GPS can be used for cross-checking GPS readings with other data sources: published maps, aerial photographs and

Summary: Advantages and Disadvantages of GPS

Advantages

- Fairly inexpensive, easy-to-use field data collection
- Modern units require very little training for proper use
- Collected data can be read directly into GIS databases minimizing intermediate data entry or data conversion steps
- Worldwide availability
- Sufficient accuracy for many census mapping applications—high accuracy achievable with differential correction

Disadvantages

- Signal may be obstructed in dense urban or wooded areas
- Standard GPS accuracy may require differential techniques
- Differential GPS is more expensive, requires more time in field data collection and more complex post-processing to obtain more accurate information
- A very large number of GPS units may be required for only a short period of data collection.
- Privacy Issues

ECA

COVID-19 RESPONSE

THANKS