GPRM 230
A (very) quick look at Statistical and Geospatial data

Alex Guindon
GIS and Data Services Librarian
Concordia University
alex.guindon@concordia.ca

What we will cover today

- Definition of Microdata
- Access to Microdata files
- Microdata
  - <odesi> (for Statistics Canada microdata)
  - QICSS (to access Statistics Canada's master files)
- Definition of Geospatial Data (GIS data)
- A quick overview of GIS data sources

STATISTICAL MICRODATA

What is microdata?

Microdata consists of the data directly observed or collected from a specific unit of observation (individual cases, answers from one individual, from one respondent).

Aggregate data	Microdata
◦ Every cell represents *added up* data (or frequencies), i.e. how many people living in Quebec identify themselves as South Asian, Chinese, Black, etc.	◦ Each row represents the answers of an *individual respondent*
◦ These stats can be expressed in absolute numbers or percentages	◦ Often the numbers are codes, not actual values.
◦ Beyond basic sums and percentages, no serious statistical analysis can be done with aggregated data	◦ Thus you need the documentation to understand the data
	◦ Raw data is what should be used to conduct in-depth statistical analysis.

Aggregate data vs. Microdata

Aggregate Data (2006 Census)

Microdata (unaggregated)

AN OVERVIEW OF STATISTICS CANADA MICRODATA

Microdata from Statistics Canada

1. ***Public Use Microdata Files* (PUMF)**

- They are *samples* of the complete datasets
- Preserve the confidentiality of respondents

2. ***Master Files***

- Complete datasets
- Only available from the QICSS data centre

ACCESS TO DATASETS

How to access PUMFs

- PUMFs are made available to students and faculty through the <odesi> database
- Some files may not be in <odesi>
  - Ask the data librarian (that's me!) to download them for you

<ODESI> DEMO

Example of products

GEOSPATIAL DATA

A definition…

Geospatial data, also known as GIS data, is any numeric data that contains explicit geographic positioning information.

Can be accessed and manipulated by *Geographic Information System* software (such as ArcGIS) to produce:

- detailed maps
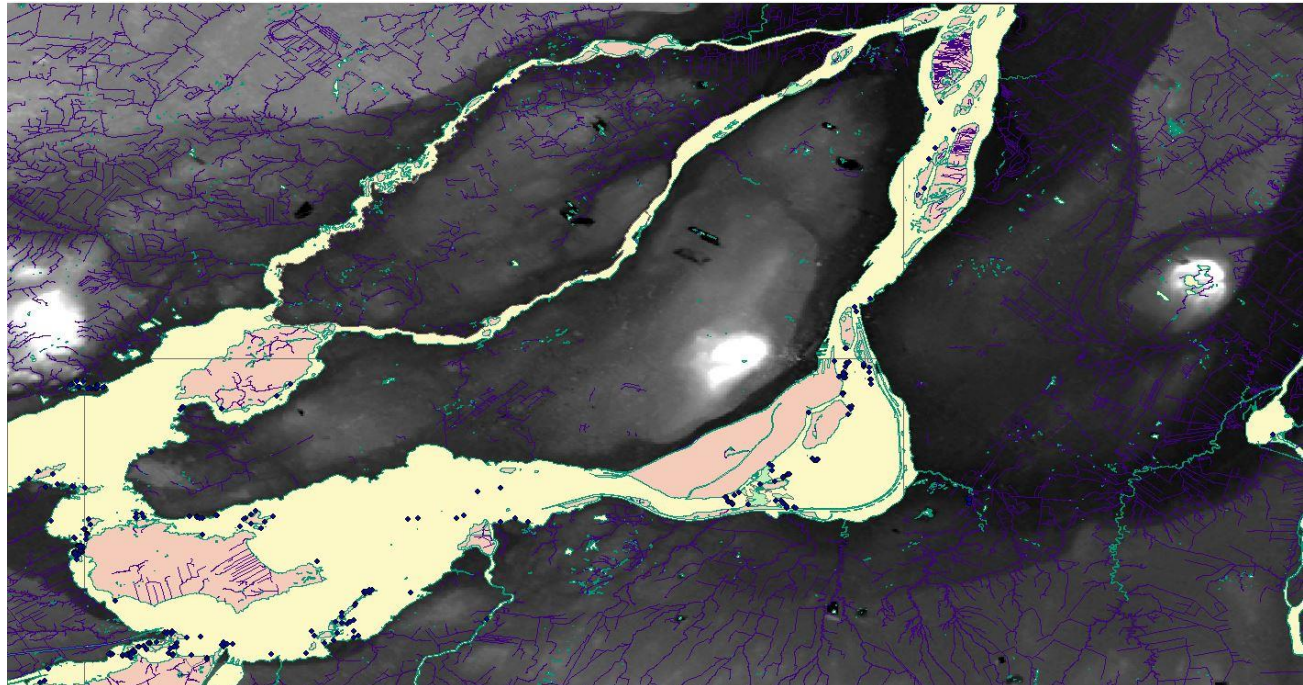- sophisticated spatial analysis

Topographic data – e.g. CanVec

Include features such as: parks, buildings, street networks, hydrography, administrative boundaries, points of interest, etc.

LiDAR and DEM

- "LIDAR (Light Detection and Ranging) is a remote sensing technology that collects 3-dimensional point clouds of the Earth's surface. These elevations and can be either ground or reflective surface elevations."
- "A Digital Elevation Model (DEM) is a digital representation of ground surface topology or relief"

Orthophotos

Orthophotos are digital images based on aerial photographs.

How to obtain the data from our collections?

1. Identify the specific area (product tiles) you want using the indexes available from the GIS Research Guide (if applicable)

2. Contact the GIS/Data Librarian to obtain the data: alex.guindon@Concordia.ca

N.B. If are confused by Step 1, just contact me ☺