Example Problem - What is the Population of Each Ward in Minneapolis?

List:
- the data you need
- the spatial functions you’ll use
- tabular operations you’ll apply

But don’t sketch out the workflow just yet (flowchart)
Tracts are polygons, contain population as an attribute recording the number of individuals in the polygon.
Tracts are polygons, contain population as an attribute recording the number of individuals in the polygon.
Data: Wards, with census tracts underneath

Census tracts: tabular data includes area, and number of people in the tract
Sketch out the steps (draw a flowchart) in your analysis to calculate the number of people per ward

e.g.,

- Tracts
- Subset
- Tracts that overlap wards

etc.,
Spatial selection: variations on subsetting the tract data
Spatial selection: variations on subsetting the tract data

Select By Location window:

- **Select features from one or more target layers based on their location in relation to the features in the source layer.**
- **Selection method:** select features from
- **Target layer(s):**
  - mnwards
  - MNtract2000
- **Only show selectable layers in this list**
- **Source layer:** mnwards
  - **Use selected features** (0 features selected)
- **Spatial selection method for target layer feature(s):** intersect the source layer feature
  - **Apply a search distance**
    - 30000.000000 Feet

Buttons:
- OK
- Apply
- Close
Even a partial overlap selects the polygon (with this selection variant)
Other variants of spatial selection

- Intersect the source layer feature
- Intersect (3D) the source layer feature
- Are within a distance of the source layer feature
- Are within a distance of (3D) the source layer feature
- Contain the source layer feature
- Completely contain the source layer feature
- Contain (Clementini) the source layer feature
- Are within the source layer feature
- Are completely within the source layer feature
- Are within (Clementini) the source layer feature
- Are identical to the source layer feature
- Touch the boundary of the source layer feature
- Share a line segment with the source layer feature
- Are crossed by the outline of the source layer feature
- Have their centroid in the source layer feature
Some variants give unexpected results
Overlay/intersect the ward polygons with tract polygons

Identify split tract pieces (new smaller polygons from the intersection) that are in each ward

Sum across the tract polygon IDs
Overlay/intersect the ward lines with tract lines

Intersect

Computes a geometric intersection of the input features. Features or portions of features which overlap in all layers and/or feature classes will be written to the output feature class.

INPUT

INTERSECT FEATURE

OUTPUT
Intersect Output
Intersect Output
What is the next step?
What is the next step?

Sum within wards to count population?
What is the next step?

Sum within wards to count population?

Does it matter if the ward and tract lines aren't co-occurring (one on top of the other)?
Partial overlaps, carry the attributes of the entire polygon (most of which is outside the wards)
After the intersection, the resulting polygons contain all the attributes and values for the input polygons, including the population value for the ENTIRE input polygon.
What to do?
What to do?

Calculate a density variable - add column(s) in the original data, and calculate number of people per unit area, e.g., people per square mile, or per square feet.

BE CAREFUL of precision specified for the variable, e.g.

0.002 persons per square mile over 1,000,000 square miles is 2000 people.

If you store density in a variable that only carries two decimal places, then the density rounds to 0.00, and when you multiply by area you get 1,000,000 * 0.00, which equals 0 people.
Original, whole-polygon population of 2194
More reasonable population estimate of 26
Sum by Wards

Summarize:

1. Select a field to summarize:
   - WARDS

2. Choose one or more summary statistics to include in the output table:
   - Owner_OCC
   - Renter_OCC
   - SQFT
   - Qt
   - popcalc
   - popcalc2
   - Sum
   - Standard Deviation

3. Specify output table:
   - Z:\Sum_Output_2.dbf

About summarizing data
Process Flowchart

Census Polygons → Add columns calculate population density → Density Polygons

Wards → Intersect

Census/Ward Combo → Table Operations: Calculate polygon area, mult. by popul. density for polygon count of people → Table Operations: Summarize across population values for polygons by Ward ID → Census/Ward Combo, with modified table

Table output:

<table>
<thead>
<tr>
<th>Ward</th>
<th>Popul.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30,122</td>
</tr>
<tr>
<td>2</td>
<td>28,256</td>
</tr>
<tr>
<td>etc.</td>
<td>.......</td>
</tr>
</tbody>
</table>
How can we improve the process?

What is our main assumption during intersection?

How do we minimize the likely violation of this assumption?
Smallest sampling units possible.

Here, blocks (because individuals not tracked)