Lab 4: Vectors Part 2 (Geoprocessing Overlays, Proximity, and Uncertainty)

Introduction
Vectors represent geographic features using points, lines, and polygons. This is in contrast to rasters, which represent data using pixels. In this lab, you will evaluate the interaction between built infrastructure and extreme hydrological hazards in Ellicott City. Raw spatial data (like FEMA flood maps) is often disjointed or overly segmented due to administrative mapping boundaries. You must pre-process this data using aggregation tools before executing overlays.

Learning Goals

- By the end of this lab you will be able to:
- Simplify complex administrative datasets using Dissolve and Multipart to Singlepart.
- Calculate continuous spatial relationships using the Near tool.
- Execute vector overlays (Intersect, Buffer) and define the geometric modifications produced by each.
- Enforce geodatabase topology rules to prevent overlaps and invalid intersections.
- Quantify the impact of positional uncertainty on spatial models by independently designing a sensitivity analysis.

Datasets Provided

- Lab04_EllicottCity_Buildings_Damage.shp (polygon): A comprehensive building footprint dataset.
- Lab04_EllicottCity_Floodplains.shp (polygon): FEMA modeled flood zones.
- Lab04_Patapsco_River_Streams.shp (polyline): USGS stream centerlines for the watershed.
- Maryland Six Inch Imagery - Cached Map Service (https://mdgeodata.md.gov/imagery/rest/services/SixInch/SixInchImagery/MapServer)

Part A: Project Setup and Topology Quality Control
Before executing spatial overlays, you must verify that the input geometry is valid. Overlapping building footprints or self-intersecting polygons will artificially inflate structural counts and area calculations. It’s good practice to ensure data validity first to save yourself from potential future headaches.

1. Initialize Project

- Create the lab04_lastname directory structure (raw, working, outputs, docs).
- Create a new ArcGIS Pro project and a default geodatabase named lab04.gdb.
- Import the provided shapefiles. Ensure the Map frame is set to NAD 1983 StatePlane Maryland (Meters).

2. Establish Topology Rules
Topology describes the spatial relationships between geographic features (e.g., connectivity, adjacency, and containment). It allows the system to understand how points, lines, and polygons relate to one another, ensuring data integrity and enabling advanced spatial analysis. You can define rules for topology in your project to ensure data integrity:

- In the Catalog pane, right-click lab04.gdb → New → Feature Dataset. Name it Urban_Topology and assign the Maryland State Plane CRS.
- Import EllicottCity_Buildings.shp into this Feature Dataset.
- Right-click the Urban_Topology dataset → New → Topology.
- Add the rule: EllicottCity_Buildings Must Not Overlap.
- Validate the topology. Note the presence of topological violations for your final analysis. Do not attempt to fix all errors manually.

Part B: Proximity and Exposure Overlays
Binary risk models assume a structure is either completely safe or completely at risk. You will calculate the continuous distance to the physical hazard, followed by an exact areal exposure overlay.

3. Continuous Risk (Proximity)
- Execute the Near tool.
- Input Features: Ellicott City Buildings.
- Near Features: Patapsco River Streams.
- Open the attribute table for your buildings. The tool has added a NEAR_DIST field, calculating the shortest straight-line distance (in meters) from each building to the stream centerline.

4. Isolate the 100-Year Floodplain
- Use Select By Attributes on the FEMA floodplain to select the 100-year flood zones.
- Export this selection to your geodatabase as FEMA_100yr.

5. Intersect Buildings and Floodplains
- Execute the Intersect tool.
- Input Features: Ellicot City Buildings and FEMA_100yr.
- Output: Buildings_Intersect_100yr.
- Note that building polygons that cross the floodplain boundary are physically cut. The resulting layer only contains the portion of the building footprint situated inside the flood zone.

Part C: Uncertainty and Sensitivity Analysis (Do It Yourself)
Spatial boundaries are mathematical models, not absolute physical barriers. A line on a map implies a precision that hydraulic models do not possess. You will model a systemic +/- 15 meter positional uncertainty in the FEMA boundary to evaluate how sensitive Ellicott City's infrastructure exposure is to mapping errors.

6. Generate Uncertainty Envelopes
- Use the Buffer tool to create two new layers from the FEMA_100yr layer.
- High Exposure Scenario: Buffer by +15 meters. Dissolve all outputs into a single feature.
- Low Exposure Scenario: Buffer by -15 meters. Dissolve all outputs into a single feature.

7. Calculate Structural Exposure (Area)

You must quantify the total square meters of building footprints exposed to water in all three scenarios.

- Task: Calculate the total building area (in square meters) intersecting the baseline 100-year floodplain. (You started this in Part B).
- Task: Calculate the total building area intersecting the -15m low exposure scenario.
- Task: Calculate the total building area intersecting the +15m high exposure scenario.

8. Evaluate Proximity vs. Zones

- Task: Using the NEAR_DIST values calculated in Part C, determine the average distance to the stream for buildings inside the 100-year floodplain versus buildings outside the 100-year floodplain.

9. Final Documentation

In your docs/ folder, create a plain text file named Lab04_Results.txt. Address the following:

- Explain how the topological errors (overlapping buildings) you observed in Step 2 mathematically alter the total exposed area calculations you performed in Step 5.
- Report the total exposed building area for the Low (-15m), Baseline, and High (+15m) scenarios.
- Based on the variance between those three area calculations, is the building exposure model in Ellicott City highly sensitive or relatively insensitive to a 15-meter horizontal boundary shift? Formulate a physical reason for this sensitivity (or lack thereof) based on the topography of a steep bedrock valley system like the Tiber Branch.
- Report the average distance to the stream for structures inside versus outside the flood zone. Identify the geometric limitation of using straight-line distance (NEAR_DIST) to predict flood risk in a sinuous valley.