Final Project Proposal—GEO265

Ben Stabley

October 31, 2022

1 Introduction

In Portland, Oregon there are approximately 10,000 to 12,000 vehicle accidents each year[7]. Although many *human factors*, such as intoxication and distractions, contribute to vehicle accidents, this project will not examine human factors. Determining "high risk" sections of road has historically been done using both recorded accident incidents and modeling based on physical road characteristics [1]. In this GIS-based approach, I hope to combine both methods to rate roads in Portland, OR based on the likelihood that a vehicle will have an accident there.

The research questions I would like to investigate are:

1. Can a "risk score" for vehicular accidents be assigned to roads in Portland? Where are some of the highest risk segments of roads?
2. How do "winter" conditions affect the assigned dry surface risk scores?

2 Data

Most of the data to be used comes from the Portland Maps Open Data website[6] and includes information such as speed limits, street width, traffic volume and speed, and potential roadside hazards. Additional traffic accident data may come from the Portland Police Bureau's database of dispatched calls [5]. Analysis for winter risk assessment will be performed possibly using raster DEM for the Portland area, and possibly obtained from Portland Maps or the Oregon Department of Geology and Mineral Industries (DOGAMI)[4]. Miscellaneous other data for use as a base map or to enhance the visual appeal of map products may be obtained from any of the above sources in addition to TIGER/Line[3].

Specific data sets that have already been obtained, and attributes that are of interest are as follows, in **name:feature** format.

streets: line The complete catalog of Portland area streets and highways as center lines. Attributes of interest are: `LOCALID` which might be able to be used to join across datasets, though unlikely; `DIRECTION` which encodes travel direction; `TYPE` which encodes street class from "Freeway" to "Forest Service Road"; and `FULL_NAME` used to identify the street and possibly join to other datasets.

speed limits: line Basically a copy of *streets* but with attributes `SpeedLimit` from 10-55 mph.

traffic speed, traffic volume: point Two datasets of points across the city's road network where speed and volume of vehicle traffic were sampled for at least a day. Traffic speed provides `PctOverPosted` which is the percent of total observations (in `ADTVolume`) that were over the posted speed limit, as well as the speeds of the 50th, 70th, 85th, and 90th percentiles. Traffic volume provides `ADTVolume` which tells how many vehicles were observed.

vehicle class: point Similar to the above, it records observations of traffic volume with breakdown by class in `PctCars` and `PctTrucks`.

high crash network: line Has information for the top 30 most dangerous streets in Portland, with Y/N fields indicating if the street is hazardous for autos, bicyclists, and pedestrians.

sign mounts, street lights: point These datasets show locations of signage (eg stop sign, no parking sign) and street lights. They are included as potential *roadside hazards* which may cause an accident to be worse if struck.

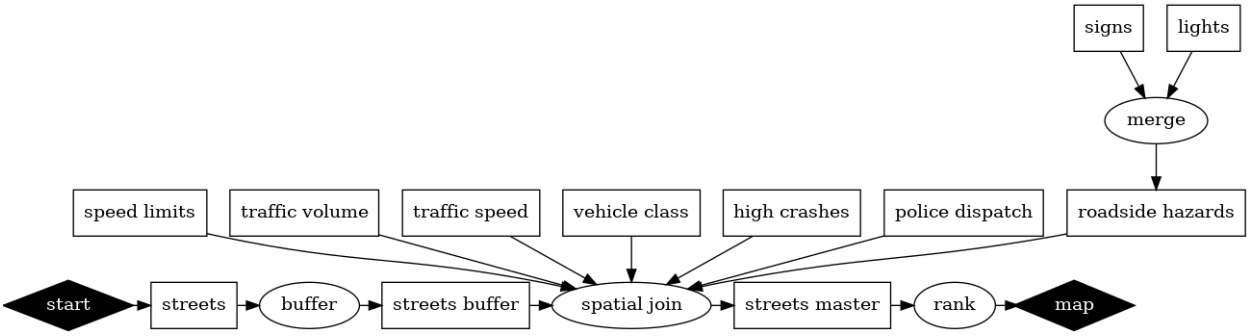

Figure 1: Approximate flowchart of basic processing tasks. No doubt things will change.

snow and ice routes: line This dataset describes plow and deicing plans for the city for winter conditions. The main attribute of interest is `Priority` which assigns an A to D priority for snow removal. Some streets in Portland, such as Powell Blvd., are state maintained and the responsibility of ODOT to plow and deice, and are not included in this dataset.

police dispatch: point This contains locations for police dispatch and non-emergency calls for 2012–2022. I would use only the "traffic accident" subset of these points and perform some sort of density calculation before applying it to road segments. I do not yet know if I will use this data or not.

3 Process

I will "do GIS" to the data as shown in Figure 1. The overall plan is to attach data from many sources to the primary *streets* layer, then use it to calculate an risk score for segments of the street. Because very few of the datasets have identifiers in common (such as a primary key), I think I will have to rely heavily on spatial joins. In order to intersect lines and points, I will first use the *buffer* operation to turn streets into wider polygons.

For the *roadside hazards* and *police dispatch* datasets, I may first combine each into a single feature class, possibly using *buffer* and *dissolve* to combine nearby features, then associate a count with each new feature using spatial join and aggregating. Another possibility is to *tesselate* and then perform a spatial join in order to calculate a density for each feature which can then be applied to road segments that intersect using, once again, a spatial join.

Winter risk scores will get additional input. For this, I was thinking of using the DEM data in order to calculate slope. I can then use some tool to sample the DEM along each line feature of *streets* and assign a slope and elevation to each street. The motivation for this is that steep streets are more hazardous when covered with snow and ice, and that higher elevation streets have increased chance of having snow and ice.

I will also use the *snow and ice routes* data to modify the winter risk scores, with the idea that roads which are better plowed and deiced will be safer to drive.

In order to combine all these factors into a final "risk score", I will determine (hopefully motivated by research) integer "importance ranks" for each factor. I will use the following formula to calculate weights in the interval $(0, 1)$ for each factor[2].

$$w_i = \frac{n - r_i + 1}{\sum_{k=1}^{n}(n - r_k + 1)} \tag{1}$$

Where w_i is the weight for the i^{th} factor and n is the total number of factors.

4 Map products

The two main map products I plan to make will show the final "risk score" for normal conditions and for winter conditions. Because there are so many roads which may be difficult to show on a single map at a

city-wide scale, I may have to make some kind of series of maps to decrease the "clutter" on each. This may involve focusing on areas of interest, or focusing on major classes of streets.

I will also make a few maps to illustrate the input datasets. This will probably require some sort of generalization, such as displaying the density of data instead of all the individual measurements.

References

[1] Ezra Hauer. "Identification of Sites with Promise". In: *Transportation Research Record* 1542 (Jan. 1996), pp. 54–60. DOI: 10.3141/1542-09.

[2] Paul Bolstad. *GIS Fundamentals. A First Text on Geographic Information Systems*. 6th ed. XanEdu, 2019. ISBN: 978-1-59399-552-2.

[3] US Bureau of the Census. *TIGER/Line Shapefile*. 2022. URL: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2020.html (visited on 10/30/2022).

[4] State of Oregon Department of Geology and Mineral Industries. *Lidar*. 2022. URL: https://www.oregongeology.org/lidar/ (visited on 10/30/2022).

[5] Portland Police. *Police Dispatched Calls Dashboard*. 2022. URL: https://www.portland.gov/police/open-data/police-dispatched-calls (visited on 10/30/2022).

[6] *PortlandMaps Open Data*. 2022. URL: https://gis-pdx.opendata.arcgis.com/ (visited on 10/30/2022).

[7] Portland Dept. of Transportation. *How Crash Data Works*. Jan. 1, 2022. URL: https://www.portland.gov/transportation/vision-zero/crash-data (visited on 10/30/2022).