

Project 4: Jen and Barry's Site Selection in PostGIS

Frame the Question

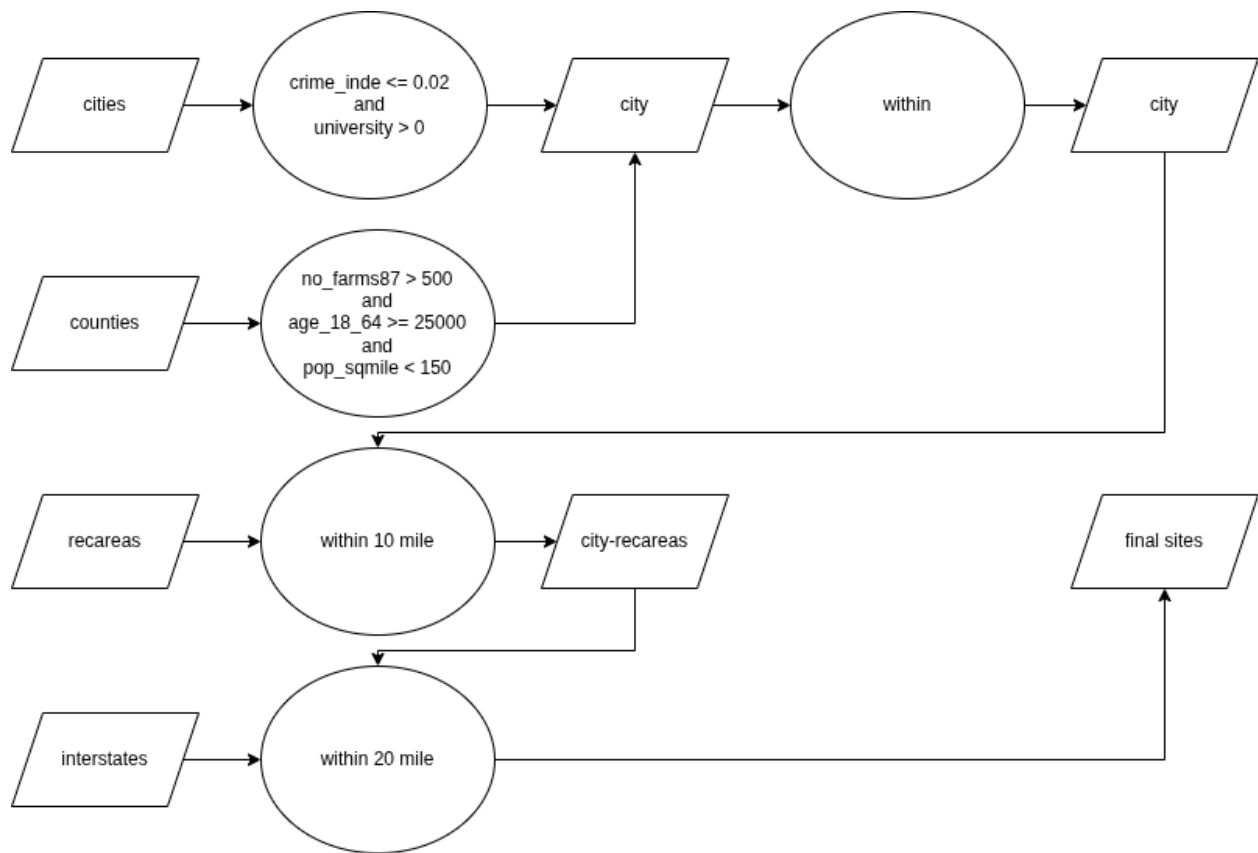
The main question for this project is to find the best possible sites to open a new ice cream business in Pennsylvania using. locate cities that meet both county and city level criteria while also being close to recreation areas and interstates. The analysis focuses on combining spatial and attribute conditions such as crime rate, university presence, farm counts, and proximity to major infrastructure. The process involves identifying which areas provide a safe environment, access to resources, and suitable business conditions for growth and sustainability.

Understand the Data

The data used in this project includes four main layers which are cities, counties, recreation areas, and interstates. Each layer contains spatial geometry and attribute data that describe population, crime index, farm numbers, and other key indicators. The cities and counties datasets provide demographic and economic information, while the recreation areas and interstates datasets help evaluate accessibility and location quality. All datasets are stored in geographic coordinates NAD27 and must be transformed into a projected coordinate system suitable for distance measurements. Understanding these datasets ensures accurate and reliable results.

Choose a Method

The PostGIS spatial functions used in this project include ST_Within to find cities located within selected counties and ST_DWithin to calculate proximity between cities, recreation areas, and interstates. Logical and spatial filters are applied to meet the given selection criteria, combining attribute conditions such as low crime rate, high number of farms, and population density. The method involves writing SQL queries that use joins, transformations, and distance evaluations to progressively narrow down the list of suitable sites. This structured method allows automation of the selection process and ensures reproducible spatial analysis.



Process the Data

The data processing begins by selecting counties with more than 500 farms, a population between 18 and 64 of at least 25000, and population density less than 150 per square mile. Cities within these counties are then filtered based on a crime index less than or equal to 0.02 and the presence of at least one university. Using the ST_Within function, these filtered cities are identified as being inside the selected counties. Next, ST_DWithin is used to find cities located within 10 miles of recreation areas and within 20 miles of interstates after transforming geometries to Pennsylvania State Plane coordinates. This step results in a refined set of final candidate sites.

Result

The final result is a list of distinct cities that satisfy all given conditions for Jen and Barry's business location. These cities represent safe, accessible, and economically suitable areas with strong agricultural support and nearby universities. The SQL query successfully automates the process, producing a shortlist of ideal sites ready for evaluation. The final query combines all spatial and attribute filters, yielding four cities in Pennsylvania for the new ice cream shops.