Computer Science Apprenticeship by the Faculty of Engineering

Ice Cream Site Selection – Report

Supervised by : Dr.Eehab Hamzi Hijazi

Prepared by:

- **Ne’meh Abu Issa 12114863**
- **Mayar Basheer 12112956**

1. Introduction

The goal of this project is to identify the best locations for opening new ice cream shops in Pennsylvania.
The analysis integrates demographic, agricultural, crime, and accessibility factors to determine optimal candidate cities.

The work was completed using PostgreSQL + PostGIS for spatial processing and QGIS for map visualization.

2. Dataset Description

2.1 Counties

Polygon layer representing Pennsylvania counties.

Key attributes:

no_farms87 – agricultural activity indicator

age_18_64 – working-age population

pop_sqmile – population density

2.2 Cities

Point dataset representing cities within the state.

Key attributes:

population

total_crim – total crime index

university (1 = contains university, 0 = no university)

2.3 Interstates

Polyline dataset representing major interstate highways.

2.4 Recreation Areas

Polygon dataset representing parks and recreational lands.

3. Methodology

The selection process is based on a set of spatial and non-spatial filters applied sequentially.

3.1 County–City Spatial Join

Cities were linked to their counties using the spatial function ST_Contains.

```
SELECT c.gid AS county_gid, c.name AS county_name, c.geom AS county_geom,
       c.no_farms87, c.age_18_64, c.pop_sqmile,
       ci.gid AS city_gid, ci.name AS city_name,
       ci.population, ci.total_crim, ci.university, ci.geom AS city_geom
FROM counties c
JOIN cities ci
ON ST_Contains(c.geom, ci.geom);
```

3.2 Identifying “Good Counties”

Counties were filtered based on the following criteria:

no_farms87 > 400

age_18_64 > 20,000

pop_sqmile > 50

```
SELECT *
FROM county_city_join
WHERE no_farms87 > 400
  AND age_18_64 > 20000
  AND pop_sqmile > 50;
```

3.3 Selecting Cities within Good Counties

Cities located inside the selected counties were extracted from the previous result.

3.4 Distance to Interstates

Distances between cities and interstate highways were computed using ST_Distance in meters:

```
SELECT city_name,
       ST_Distance(c.city_geom::geography, i.geom::geography) AS dist_meters
FROM good_counties c
CROSS JOIN interstates i
ORDER BY dist_meters;
```

The goal is to identify cities that are reasonably close to highways to ensure accessibility.

3.5 Final Candidate Selection

By evaluating the distance results and identifying a clear distance break (jump), the four closest cities to interstates were selected as the final candidates.

4. Final Candidate Cities

The four selected cities are:

City Name	Population	Crime Index	University	County	Farms	Distance to Interstate (m)
Nittanytown	85,000	1500	Yes	Center	817	~3032 m
Huntstown	7,680	120	Yes	Taft	1043	~4551 m
Victoria Falls	26,580	210	No	Raccoon	628	~4887 m
Cumberland	20,005	500	No	King	635	~5037 m

These cities satisfy:

adequate population

manageable crime levels

presence of universities (for two cities)

strong agricultural surroundings

close proximity to major interstates

5. Visualization

Final maps were created in QGIS, including:

Counties

Good Counties (highlighted)

Cities

Final Candidate Cities (yellow points)

Interstates overlay

6. Conclusion

After applying demographic, spatial, and accessibility filters, four cities emerged as the most suitable locations for new ice cream shops.
These cities show strong business potential due to population characteristics, proximity to highways, and favorable county-level indicators.

The workflow demonstrates effective use of PostGIS spatial functions and QGIS visualization to support location-based decision-making.