

Designing algorithms for searching for optimal/twin points of sale in expansion strategies for geomarketing tools

1. Problem statement

1.1. Introduction

Businesses such as Mercadona, Makro, PC City etc., which are looking for expanding their sales networks, have to develop strategies for searching for new points of sale. Preferably, these points should be located in areas with the largest market potential or in areas with characteristics similar to areas where their best point of sales are currently situated.

As this decision involves millions of euros in investments, it is very important to collect the necessary information and to have appropriate tools to make the final decision. An error in the decision could be catastrophic for the company.

We propose to develop a mathematical algorithm that would allow the automation of this decision-making. The algorithm will be incorporated into Geomarketing tools. This algorithm and tools will allow companies to reduce the number of errors and increase the decision efficacy.

1.2. Problem to solve

The objective of this problem is to design an algorithm that will allow the following:

- 1) For a particular point of sale of a chain store to be chosen and the characteristics that make it successful to be highlighted. To identify the main features of a successful shop, those which make it a good point of sale.
- 2) Identify, in a given area, a similar point of sale to its best selling point, based on the features found in (1)

The company will provide the following input information:

- a) The location of the best point of sale
- b) The area of influence of the point of sale (household at n minutes walking or driving time)

- c) List of factors that the commercial chain take into account when opening a new point of sale

2. Description of the work

2.1. Previous elements

To solve the problem, a database by sections census with socio-demographic information will be provided. The database provided will have both categorical and continuous variables (variables dictionary will be provided). A database with distances between sections census (walking and driving distances) will be also provided.

2.2. Work to be done

- a) To determine the best features for a given point of sale.

The business question, needing to be answered, could be like this: *"Suppose that we have a mobile phone shop in Plaza de Manuel Becerra which has the best sales in Madrid. Which are the socio-demographic features of this store which determine its high sales?"*

So the work to be done would be: *"Define an algorithm for a set of predefined variables, for a determined point of sale located at the specified coordinates (census) and an area of influence of a certain size, look for the variables that stand out compared to other point of sale "*

- b) Establish the criteria of likeness

This work consists of building a statistical measure of similarity (business similarity) between two points of sale, depending on the selected variables in (a)

- c) Find the most similar point of sale

The business question to be answered in this section is: *"I want to find the point of sale that is the most similar to my shop in Manuel Becerra in Sevilla".*

In order to find the solution, we must build an optimization algorithm to locate the most similar store to my reference store, in terms of the variables that score highly on that point of sale. The efficiency of the algorithm will be assessed.

Appendix: definitions and considerations

- **Sections census:** The INE has divided the country into 36,000 units between 1,000-2,000 inhabitants. These correspond to the groups of people in polling stations for the elections. This will be the minimum unit of study and points of sale should be located at this level. Each section census is associated to a pair of geographic coordinates.
- **Area of influence:** Each point of sale is characterized by the area where its costumers come from. The catchment area is characterized by all those who live within a certain distance (either travelled on foot or by car)
- **Walking Distance:** Time taken to walk from point X to point Y (Bearing in mind the traffic)
- **Driving distance:** Time taken to go from point X to point Y by car (the road and traffic directions have to be taken into account)