



The Battle of the Neighborhoods

APPLIED DATA SCIENCE CAPSTONE BY IBM/COURSERA

Rubén Ureña | | 02/09/2020

The Problem

Hello, my name is Rubén and I am currently living in Barcelona. I usually drive to go to work among other tasks. The problem I am about to study is the one below.

- How do I select the safest route for me and my passengers?
- Is it the shortest route the only variable to consider?

I see this problem relevant for anyone living or ever visiting this beautiful and touristic city.

The problem fits with the requirements of the Project since we will use clustering, map visualization and therefore the Foursquare location of the city neighbourhoods.

The Data Section

Based on definition of our problem, factors that will influence our decision are:

- * The total number of accidents happened in each street during the year.
- * The most common neighborhoods in each district.

List of accidents handled by the local police in the city of Barcelona. Incorporates the number of injuries by severity, the number of vehicles and the point of impact.

https://www.kaggle.com/xvivancos/barcelona-data-sets?select=accidents_2017.csv

Data sets from the Open Data BCN portal, the Ajuntament de Barcelona's open data service.

Open Data BCN, a project that was born in 2010, implementing the portal in 2011, has evolved and is now part of the Barcelona Ciutat Digital strategy, fostering a pluralistic digital economy and developing a new model of urban innovation based on the transformation and digital innovation of the public sector and the implication among companies, administrations, the academic world, organizations, communities and people, with a clear public and citizen leadership.

The Dataset

The set contains the following variables:

Id: A unique number assigned to each accident by the police.

District Name: Name of the district where the accident occurred.

Neighborhood Name: Name of the neighbourhood where the accident occurred.

Street: Name of the street where the accident occurred.

Weekday: Day of the week when the accident occurred.

Month: Month of the week when the accident occurred.

Day: Number of the day of the month when the accident occurred.

Hour: Time in hours when the accident occurred.

Part of the day: Part of the day when the accident occurred.

Mild injuries: Mild injuries at the accident occurred.

Serious injuries: Serious injuries at the accident occurred.

Victims: Number of victims at the accident occurred.

Vehicles involved: Number of vehicles involved at the accident occurred.

	Id	District Name	Neighborhood Name	Street	Weekday	Month	Day	Hour	Part of the day	Mild injuries	Serious injuries	Victims	Vehicles involved
0	2017S008429	Unknown	Unknown	Número 27 ...	Friday	October	13	8	Morning	2	0	2	2
1	2017S007316	Unknown	Unknown	Número 3 Zona Franca / Número 50 Zona Franca ...	Friday	September	1	13	Morning	2	0	2	2
2	2017S010210	Unknown	Unknown	Litoral (Besòs) ...	Friday	December	8	21	Afternoon	5	0	5	2
3	2017S006364	Unknown	Unknown	Número 3 Zona Franca ...	Friday	July	21	2	Night	1	0	1	2
4	2017S004615	Sant Martí	el Camp de l'Arpa del Clot	Las Navas de Tolosa ...	Thursday	May	25	14	Afternoon	1	0	1	3