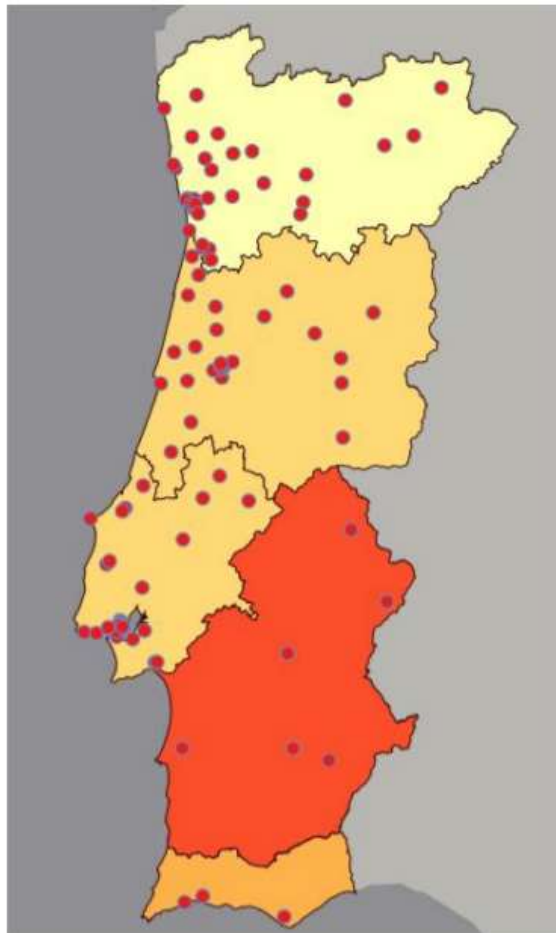


# **Capstone Project - The Battle of Neighborhoods (Week 2)**

# Introduction

I am a data scientist in a IT company in Portugal. My client, the Ministry of Portuguese Health, has approached our company recently seeking for our help. The Ministry of Health wants to open a restaurant for employees of hospitals, health centers and for the IT team that supports health services. They want to know the best district, in order to cover the largest number of people and what get more money to start this idea. Our team attempted to fulfill his need by the power of data science.

Regarding the status of existing institutions, it appears that the vast majority (77%) are public corporate entities (E.P.E.), endowed with administrative, financial and patrimonial autonomy and are governed by the legal regime applicable to public corporate entities. In the group of EPE, there are 8 Local Health Units (ULS). There are specialized hospital institutions, belonging to the SNS network, the three Oncology Institutes, located in the north, center and south of the country, two psychiatric institutions (Hospital de Magalhães Lemos, EPE and Centro Hospitalar e Psiquiatrico de Lisboa), and two Centers of Physical Medicine and Rehabilitation (Rovisco Pais Physical Medicine and Rehabilitation Center and Southern Physical Medicine and Rehabilitation Center, PPP)



# Business Problem

The business problem to be solved using data science and Foursquare's location data was to identify, for each district of Portugal, where to choose to establish a restaurant for employees. To better understand their needs, employees were asked the following question:

"What are your preferences about the location of this new restaurant to be established?"

And the response from the target audience, who was my client, was:

"I want a place close to where hospital staff work. However, I expect to face less competition and, ideally, candidates shouldn't have many restaurants."

After further communication and clarification, the following criteria were defined:

The proposed location should be one of the top 5 districts in Portugal with more restaurants in general; At the same time, the proposed location must have the lowest percentage of restaurants.

## DATA

The data that was used came from:

Number of hospital workers in each area of Portugal from  
: <https://transparencia.sns.gov.pt/explore/dataset/trabalhadores-por-grupo-profissional/table/?flg=pt&disjunctive.regiao&disjunctive.instituicao&sort=periodo> Four square location data to find out the numbers and the kinds of restaurants and hospitals in various districts in Portugal.

# DATA ANALISYS