Hospitals Investment

1. INTRODUCTION

1.1 Background

The healthcare industry is an integration of sectors that provides goods and services to treat patients with curative, preventive, rehabilitative, and palliative care. The modern healthcare industry includes three essential branches, which are services, products, and finance, and may be divided into many sectors and categories and depends on the interdisciplinary teams of trained professionals to meet the health needs of individuals and populations. This industry is a strategic sector of a country, whereby this is one of the most important areas to do business.

1.2 Problem description

A Corporate group in the Health industry is studying the possibility of invest in the construction and management of 5 Hospitals. The Corporate wants to know where to allocate that investment, namely, in which countries, and in which cities.

1.3 Interest

This idea comes from personal curiosity to know more about the health industry a make a comparison of the healthcare system between countries.

2. DATA ACQUISITION AND CLEANING

2.1 Data sources

We need general information about the healthcare system of the countries, for example, population, GDP, number of hospitals, and any relevant information that helps us to make an informed decision.

It is not easy to find all the data needed in one place. Our primary sources of information will be the OCDE [1] and the World Health Organization (WHO) [2]. Any additional data we will retrieved from various sources, given that those datasets are inconsistent and incomplete.
There are dozens of even hundreds of variables to study the health system of a country. We are going to focus only on some relevant features to assess the suitability of a city to allocate the investment using only with data publicly available online.

We will build two datasets with information about the Countries, and the other with details of the Cities. The primary feature will be the population to decide which cities we will study. With that, we will choose the 20 most populated cities of countries members of the OCDE.

One feature that is difficult to obtain is the number of hospitals in each city, the OCDE and the WHO datasets does not provide this information. In this case, we are going to use the Foursquare API to obtain the number of hospitals given the geolocalization coordinates and a radius of search.

Once we collected and prepared the data, we are going to use clustering to determine if there is a cluster of cities more suitable for the allocation of the investment.

References

[1] https://stats.oecd.org/Index.aspx?DataSetCode=CITIES

[2] https://www.who.int/data/gho