



## **GROUP PROJECT No1**

### **YELP & GOOGLE MAPS - REVIEWS AND RECOMMENDATIONS**



Google Maps

## Context

"User opinion is very valuable data, which grows every day thanks to review platforms. Its analysis can be decisive for strategic planning. Yelp is a review platform for all types of businesses, restaurants, hotels, services among others. Users use the service and then upload their review according to the experience they have received. This information is very valuable for companies, since it helps them find out about the image that users have of the company's different locations. , being useful to measure the performance and usefulness of the premises, in addition to knowing in which aspects the service needs to be improved. In addition, Google has a review platform for all types of businesses, restaurants, hotels, services, among others, integrated into its service. location and maps, Google Maps. Users use the service and then upload their review based on their experience. Many users read reviews of the places they plan to go to make decisions about where to buy, eat, sleep, meet, etc. . This information is very valuable for companies, since it helps them find out about the image that users have of the company's different premises, being very useful to measure the performance and usefulness of the premises, in addition to identifying the aspects of the service to improve.""

## Role to develop

"As part of a data consultancy, we have been hired to carry out an analysis of the US market. Our client is part of a conglomerate of restaurant and related companies, and they want to have a detailed analysis of user opinion on Yelp and cross them with those from Google Maps about hotels, restaurants and other businesses related to tourism and leisure, using sentiment analysis, predict which business areas will grow (or decline) the most. In addition, they want to know where it is convenient to locate the new ones. local restaurants and similar, and they want to be able to have a restaurant recommendation system for the users of both platforms to give the user, for example, the possibility of being able to discover new flavors based on their previous experiences. They can change the type of commerce (not they need to be restaurants)."

## Job offer

Collect, purify and make information available: Creation of a database (DataWarehouse) from different sources, both provided by Henry and incorporated by you, running locally or hosted in cloud providers. The refined database must consider at least two different types of data extraction, for example: static data, calls to an API, scrapping, among others.

Significant report and analysis of the chosen line(s) of research: The analysis must consider the

relationships between variables and conclude, if it exists, a relationship between them, and the possible factors that cause said relationship in reality.

Training and putting into production an unsupervised or supervised classification machine learning model: The model must solve a problem and connect globally with the proposed objectives proposed as a project. .

## **Analysis and implementation ideas**

Improving marketing strategies: micro-segmented campaigns

Recommendation systems: About a particular place, they can be businesses such as restaurants, hotels, among others.

Additional data to cross: Stock market quotes, trends in social networks and media about expanding businesses.

## **Datasets and complementary sources**

"The data is extracted from the review platform Yelp and Google Maps, for the United States. It has information about the location of the businesses, their category, average scores, whether they are open or not, about the users, the reviews they made, how many reviews they made, how many votes those reviews have received, etc. The analysis can be complemented with additional datasets

that offer information that they consider relevant to the request (share values of the companies, additional geographic information of the premises, etc.). We strongly recommend the use of NLP to process review information."

### **Data sources**

[Data Dictionary](#)

Mandatory data sources:

[Dataset de Google Maps](#)

[Dataset the Yelp!](#)

# Disclaimer

On behalf of Henry's team, we would like to clarify and emphasize that the purposes of the proposed projects are exclusively pedagogical, with the aim of carrying out projects that simulate a work environment, in which various topics adjusted to reality are worked on. They do not necessarily reflect the philosophy and values of the organization. Furthermore, Henry does not encourage or recommend that students and/or anyone reading the repositories (and project submissions) take action based on data they may or may not have collected. All the information presented and results obtained in the projects should never be taken into account for real decision-making (especially in the areas of finance, health, politics, etc.).