**Exploring the level of trade/ market near to the beaches most popular beaches in Costa Rica.**

Introduction

In Costa Rica the Tourism is an important source of employment and generates indirectly a huge number of other jobs, it is an engine for the economy reaching up to 8.8% of the GDP and almost a 10% of the employment in the country. Then creating more attractive places to help the foreign people enjoy our country is a good idea, opening a business near a beach or in a community near to a beach could be a good business thanks the number of visits we have where 72% of people visit the beach or is looking for sun and beach.

Business Problem

Expanding a business is a great deal at any scale and the owners of one restaurant in one of the most popular beaches want to open a second one, it is natural that some questions surge as where is a good place to open? Which other beaches are like the beach where we are located? and how many restaurants exist there? how is the competition on those places?

Python and Foursquare are going to try answering these questions and generate new insights that will help the owners to take a decision based on data.

Data Section

Data Acquisition

To work with this project was needed to scrape the web and create a completely new dataset.

The pages used were [Playas Costa Rica](https://es.wikipedia.org/wiki/Anexo:Playas_de_Costa_Rica), [Go visit Costa Rica](https://www.govisitcostarica.co.cr/travelInfo/beaches/beaches-list.asp) and [Go Playa](https://goplaya.cr/index.php/search_controller/buscar?latitud=9.935607&longitud=-84.1833856&busqueda=todas&filtros%5B%5D=81&radio=350#sa) the first one for a list of beaches and the next to extract general features of each beach.

Data Cleaning

In this case is was required to do more steps working on matching the beaches because some of them had another name, an extra space, in general another format that did not help when building the dataset, for example playa Samara and playa Samara or Playa Mal País and Malpais.

Another issue was some beaches had different names or a beach in one site was two beaches in another site, example Playa Punta Leona (a great place to relax) was Playa Mantas and Playa Blanca.

Those cases where it was not possible or required analyze the case individually were skipped and them no features were added.

Data Analysis

that is some cases there were more beaches