

# Data Science Project | Where to live in Bogota, Colombia

## INTRODUCTION

### THE BACKGROUND



Bogota is the largest city, and also the capital of Colombia, and according to [worldpopulationreview.com](http://worldpopulationreview.com) has about 11 million people living in the metropolitan area.

Bogota is located in the center of Colombia, on a high plateau known as the Bogotá savanna, part of the Altiplano Cundiboyacense located in the Eastern Cordillera of the Andes. It is the third-highest capital in South America and in the world after Quito and La Paz, at an average of 2,640 meters (8,660 ft) above sea level. Subdivided into 20 localities, Bogotá has an area of 1,587 square kilometers (613 square miles) and a relatively cool climate that is constant through the year.

Bogotá stands out for its economic strength and associated financial maturity, its attractiveness to global companies and the quality of human capital. It is the financial and commercial heart of Colombia, with the most business activity of any city in the country. The capital hosts the main financial market in Colombia and the Andean natural region and is the leading destination for new foreign direct investment projects coming into Latin America and Colombia. It has the highest nominal GDP in the country, responsible for almost a quarter of the nation's total (24.7%).

The city's airport, El Dorado International Airport, named after the mythical El Dorado, handles the largest cargo volume in Latin America, and is third in number of people. Bogotá is home to the largest number of universities and research centers in the country, and is an important cultural center, with many theaters, libraries and museums.

Colombia is located at the north side of South America, and due to its economic stability compared to the rest of the Latin America, has become the hub for the rest of the region.

## **THE PROBLEM**

A lot of companies have located their Latin America headquarters in Bogota, which makes Bogota a place for a lot of expats.

Usually with a vast city like Bogota, expats don't know which will be the best place to live, sometimes they receive advice from local people, but don't have a data driven decision.

## **TARGET AUDIENCE**

This project has as target audience, people that want to live in Bogota, and take a data driven decision before moving to certain neighborhood.

## **DATA COLLECTION**

Before starting data mining, is important to have clarity of what kind of data will be needed to take the decision.

According to Twin Cities habitat for humanity web site, people first look for neighborhood safety, especially if you have a child, as will give greater peace of mind.

Food, medications, and toiletries are necessities for daily life. If convenience is important, you may want to consider choosing a location that's close to where you will be running a majority of your errands.

It is still important to check out the outdoor amenities a neighborhood has to offer, as sidewalks, parks, walking trails and dog parks.

Where you work in relation to where you live is an important factor to consider, especially if your goal is to have a short commute. If driving yourself isn't an option, or one you prefer to avoid, look into transportation options.

## **BOGOTA ZIP CODES**

To find the zip codes of neighborhood in Bogota, the postal code website was used, as it contains a csv file with all the required info:

[http://visor.codigopostal.gov.co/472/visor/Codigos\\_Postales\\_Nacionales.csv](http://visor.codigopostal.gov.co/472/visor/Codigos_Postales_Nacionales.csv).

As the name of the link indicates, the file contains all info for Colombia, and need to be filtered to keep only Bogota data.

## **NEIGHBORHOODS SAFETY**

A report from a local government web site was found, where information need to be scrapped, as it was published as a pdf file.

[https://scj.gov.co/sites/default/files/documentos\\_oaiee/Boletin\\_2020\\_05\\_Reporte\\_bogota\\_2020\\_05.pdf](https://scj.gov.co/sites/default/files/documentos_oaiee/Boletin_2020_05_Reporte_bogota_2020_05.pdf)

## **VENUES**

As source of venues data, foursquare will be used, taking advantage of the zip codes information.

For work information I find that is really dependable of where the office will be located at, for that reason, the scope of the project will be to show where will be the best places to live, using the classification named above.