

Applied Data Science Capstone Project

Lucas Pereira Carini

IBM Data Science Professional Certificate, Coursera

1 Introduction

Everyone has heard the following phrase: 'This is so New York!!'. Or maybe saying it about another famous city in the heard. Maybe you heard it when someone just saw something that remembers this person of New York.

Of course, this can happen to almost everything or everyplace! However, there is some quite charming unique places in New York City, is not? But, actually, are there characteristics of New York that distinguish it from other cities? Like Toronto, perhaps?

We are going to try to answer that question by using some Foursquare data, such as restaurants, coffee places, etc! We are going to use some machine learning algorithms trying to extract the characteristics that defines it city in order to classify each Neighbourhood as belonging to Toronto or New York. If our models succeed to do it so, there are characteristics that defines each city, and, actually, it is correct to say 'This is so New York!!'. In the other hand, if our models fails to classify each Neighbourhood, we can say that both cities are very similar!

2 Data

We are going to use data from the past labs of the Applied Data Science Capstone Course. These data contain the coordinates of each Neighbourhood and the corresponding Borough. Then, we will use maps provided by the Folium package to illustrate these Neighbourhoods.

Moreover, we are going to use the Foursquare API to fetch data about venues, such as restaurants, museums and shops. These data will be the features of each sample to our model!

3 Methodology

4 Results and Discussion

5 Conclusion