Contents

1. Inti	roduction	
1.1.	Background	
	Problem	
	Target Audience	
	ita acquisition and cleaning	
	Description of Data	

Comparing the neighborhoods of the two New York City and Toronto

Denny Thomas

27th November 2020

1. Introduction

1.1. Background

Homo sapiens evolved from Africa and spread all over world over the time of millions of years. Earlier men were hunter gatherers slowly they start living in societies and by the introduction of agriculture and farming humans create villages, then slowly into different countries with fixed boundaries. Today we are in world having different countries which is having different cities of different cultures as per the demography.

Now migration going on with high pace as compared to the earlier times with the advancement of science technology and advanced logistics facilities. Earlier humans were migrated in search of food, climatic condition etc. In Present scenario humans' needs are changed, the main governing factors are, living conditions, facility, social security, opportunities etc. There for it is advantageous to analysis how similar and dissimilar are different cities in world.

1.2. Problem

Compare the neighborhoods of the two cities, New York City and Toronto and determine how similar or dissimilar based on the various factors such as climate, facilities, living conditions. The project aims to decided how similar are there and cluster the similar neighborhood based on the similar venues.

1.3. Target Audience

Developed for the people who wants to know the different Neighborhoods similarity around the Toronto and New York Cities. Based on this one can take the business decision, sales/marketing strategies for these regions.

2. Data acquisition and cleaning

Neighborhood details and climatic conditions can be found in the Wikipedia and foursquare location data is used to explore or compare neighborhoods.

2.1. Description of Data

 Wikipedia: - Used for identifying the neighborhoods of the cities and analyzing the demographic and climatic data, this data will be used for clustering the neighbor hood