Ideas for a successful business for Seattle, USA and Vancouver, Canada

[DOCUMENT SUBTITLE]
ANASTASIYA BUKHALA

Business problem and background

Both Seattle and Vancouver are one of major cities in USA, financial, cultural and seaport centers. They lie geographically close to each other, have approximately the same population size and climate. The goal of this research is to predict top five best start/up ideas for these cities. Startup success depends on many factors, especially if a startup is a venue, like restaurant, café, coffee shop, etc. One of the most important factors is location. So, in this project we will analyze neighborhoods of both Seattle and Vancouver to find the most common and trending places in these cities. The results of the research will give us a general idea of a startup and its location in each city mentioned before.

Datasets used in research

Dataset for neighborhoods in Seattle will be scraped using BeautifulSoup package from https://www.zipdatamaps.com/zipcodes-seattle-wa

Dataset for Vancouver will be scraped from http://zip-code.en.mapawi.com/canada/4/british-columbia/1/2/bc/vancouver-ubc-/v6t/278/

Methodology

K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. Typically, unsupervised algorithms make inferences from datasets using only input vectors without referring to known, or labelled, outcomes. In this project we will use k-means clustering to define similar groups of neighborhoods for further analysis.