Coursera Capstone

IBM Applied Data Science Capstone

Opening a cricket stadium in India

By: Mihir M Kestur

Introduction

Cricket is a sport that's watched and cherished by many in the world. The craze for the sport, in the Indian subcontinent, is arguably unparalleled. The country hosts a variety of international tournaments ranging from the shortest format of the game to the longest 5-day test match series'. It also hosts a number of domestic tournaments, amongst which the Indian premier league (IPL) is a world-class, highly anticipated, annual festival. As per https://en.wikipedia.org/wiki/Indian_Premier_League, the Indian cricket industry is easily a billion dollar industry thus having a major positive impact on the economy. Therefore it is easy to realize the importance of having sufficient cricket stadiums to cater to the demand the sport creates in the country.

Business problem

The objective of this capstone project is to analyse and select, using data science methodologies and machine learning techniques like clustering, the best location to open a new cricket stadium in the country of India.

Data Requirements

Some parameters that affect the selection of the city are proximity/existence of an airport, restaurants and other attractive venues for the players and spectators. It is also important to exclude the cities in which the cricket stadiums already exist.

- List of top 100 cities in the country web scrapped from
  https://www.nriol.com/india-statistics/biggest-cities-india.asp
- Cities in which cricket stadiums already existing scrapped from
  https://en.wikipedia.org/wiki/List_of_international_cricket_grounds_in_India
- Airport in the city scrapped from
  https://www.goibibo.com/airports/airports-in-india/
- Restaurants and other eateries in the city obtained from foursquare API.