



NEIGHBORHOODS IN ALMAY, KAZAKHSTAN

DEVELOPMENT OF TOURISM

FOR COURSERA APPLIED DATA SCIENCE CAPSTONE PROJECT

Name: Yerdos Kulnazarov | Applied data science capstone | Date: 05/16/2021

Table of contents

No	Topic	Page number
1	Introduction	3
2	Data collection <ul style="list-style-type: none">- Neighborhoods data- Geographical coordinates- Venue data	4

INTRODUCTION

Tourism has always been a booming sector across the globe. No matter which country you live in, you can always come across a group of people, big or small, who always like to visit places. I being an adventurer myself can acknowledge this fact as to how tourism plays a salient role for a traveler/explorer. Tourism is not only an important aspect of a country's economy but also for its global standing.

Why Tourism is important to any country?

The tourism industry is important for the benefits it brings and due to its role as a commercial activity that creates demand and growth for many more industries. Tourism not only contributes to more economic activities but also generates more employment, revenues, and play a significant role in development.

- i. Tourism activity creates demand.
- ii. Tourism industry value chain meets & spreads demand across industries & boosts more economic activities.
- iii. Tourism induces more consumption.

Business Problem

All the benefits of tourism tend to reflect on the employment opportunity which it gives to the people of that country. The objective of this project is to analyze the tourist places of a given state in India and try to recommend the best location where they can open a restaurant or lodging to make the best use of the opportunity.

The target audience for this project includes people who are interested in opening a restaurant, lodging, transport services, or any other similar businesses which fall within the tourism industry. This also recommends travelers' tourist venues to be visited in a given state of a country.

Project object identification

Almaty, formerly known as Alma-Ata and Verny, is the largest city in Kazakhstan, with a population of about 2,000,000 people, about 11% of the country's total population, and more than 2.7 million in its built-up area that encompasses Talgar, Boraldai, Otegen Batyr and many other suburbs. It served as capital of the Kazakh Soviet Socialist Republic and later independent Kazakhstan from 1929 to 1997. In

1997, the government relocated the capital to Akmola (renamed Astana in 1998, later renamed Nur-Sultan in 2019) in the north of the country.

Almaty is still the major commercial and cultural centre of Kazakhstan, as well as its most populous and most cosmopolitan city. The city is located in the mountainous area of southern Kazakhstan near the border with Kyrgyzstan in the foothills of the Trans-Ili Alatau at an elevation of 700–900 m (2,300–3,000 feet), where the Large and Small Almatinka rivers run into the plain.

It's a project about learning to read a city's visual landscape, and my tool is systematic classification. My core idea is to systematize Almaty attractive places to give hint to potential tourists and business starters.

DATA COLLECTION

To tackle the above mentioned problem, we need to have the dataset that contains

- i. All the districts of a particular country.
- ii. Latitude and longitudes of all the districts.

Neighborhoods Data

The Wikipedia page <https://en.wikipedia.org/wiki/Almaty> is the major source of data that is being used to obtain all the districts of Almaty. Data for attractive places found from page <http://www.walkingalmaty.com>.

The data is read into a pandas data frame using the `read_html()` method. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be scraped using the `read_html()` method of pandas.

Geographical Coordinates

The geographical coordinates for Almaty data has been obtained from the GeoPy library in python. This data is relevant for plotting the map of Almaty using the Folium library in python. The geocoder library in python has been used to obtain latitude and longitude data for various neighborhoods in Almaty. The coordinates of all neighborhoods in Almaty are used to check the accuracy of coordinates given on Wikipedia and replace them in our data frame if the absolute difference is more than 0.001. These coordinates are then further used for plotting using the Folium library in python.

Venue Data

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all neighborhoods in Almaty and is used to study the popular venues of different neighborhood.