

The Battle of Neighbourhoods

Final Project

Karl De Ruyck

A report presented in due course of the capstone project from the
IBM Data Science Professional Certificate.



coursera

by IBM at coursera.org
June 24, 2021

Contents

- 1 Introduction 2
 - 1.1 Background 2
 - 1.2 Business Problem 2
 - 1.3 Interest 3
- 2 Materials and Methods 4
 - 2.1 Data Acquisition 4
 - 2.1.1 State Capitals and GPS Co-Ordinates 4
 - 2.1.2 Popularity Contest 4
 - 2.2 Exploratory Analysis 5

Chapter I

Introduction

I.1 Background

The United States of America constitutes a federation of states, each of which has its own capital city. These state capitals are not always the most populous city in the state, but are usually considered the location of each state's legislative, judicial, and/or executive authority.

Considering the intimate relationship between a state's capital city and the interests of that state, it could be assumed that the state capital is representative of the state in even more ways than those previously mentioned. For example, it could be assumed that the culinary specialities of a state may be well represented in its capital city.

On the other hand, modern society is under strong influences toward global homogeneity. Consequently, it could also be assumed that all the capital cities of American states would be very similar in terms of popularly available cuisine.

I.2 Business Problem

Which assumption posited above is accurate? Can each state capital *differentiate themselves* by popular types of restaurant, or do people in state capitals *all prefer the same* types of food? Such an investigation could also produce mixed results, where *some* state capitals are similar to each other, while others stand apart.

1.3 Interest

Observing these trends would provide guidance to a variety of decisions:

- Restaurateurs may understand where to expand without needing to change their menus to suit the local palate.
- Tourists may be enticed to cities for a cuisine they wouldn't easily find elsewhere.
- Anthropologists may recognize patterns in human movement that link groups of state capitals with similar tastes.
- Alex Aklson may allow me to graduate from this course.

Consequently, it can be demonstrated that this is really essential research, and we cannot escape having to go through with it.

Chapter 2

Materials and Methods

2.1 Data Acquisition

2.1.1 State Capitals and GPS Co-Ordinates

A list of American states and respective capital cities was available on the Wikipedia website, and downloaded from the following URL:

```
https://en.wikipedia.org/wiki/List\_of\_capitals\_in\_the\_United\_States
```

Each state and the name of its capital city was extracted, assembling a basic data frame. The name of each capital city was then passed to Nominator for lookup on OpenStreetMaps, in order to obtain the GPS co-ordinates of the city. Latitudes and longitudes were added to each city in the data frame.

2.1.2 Popularity Contest

The location data for each city was passed to the Foursquare REST API with an ‘explore’ endpoint, along with keys indicating that the category ID should match ‘Food’ and that the results should be ranked by popularity. This query returned a JSON object including up to 10 of the most popular restaurants or eateries within a 2 km radius of the geographical centre of each capital city.

The JSON object was parsed to extract the name of each restaurant, its location, and the type of food served there. A new data frame was then created to list

each restaurant returned as popular venues in each capital city, and include the aforementioned variables alongside each listing.

2.2 Exploratory Analysis

By grouping restaurants according to the type of food served, it was possible to visualize the variety in each city. As shown in Fig 1 below, cities where many different cuisines were represented in their top 10 restaurants can be distinguished from cities where certain cuisines are predominant.

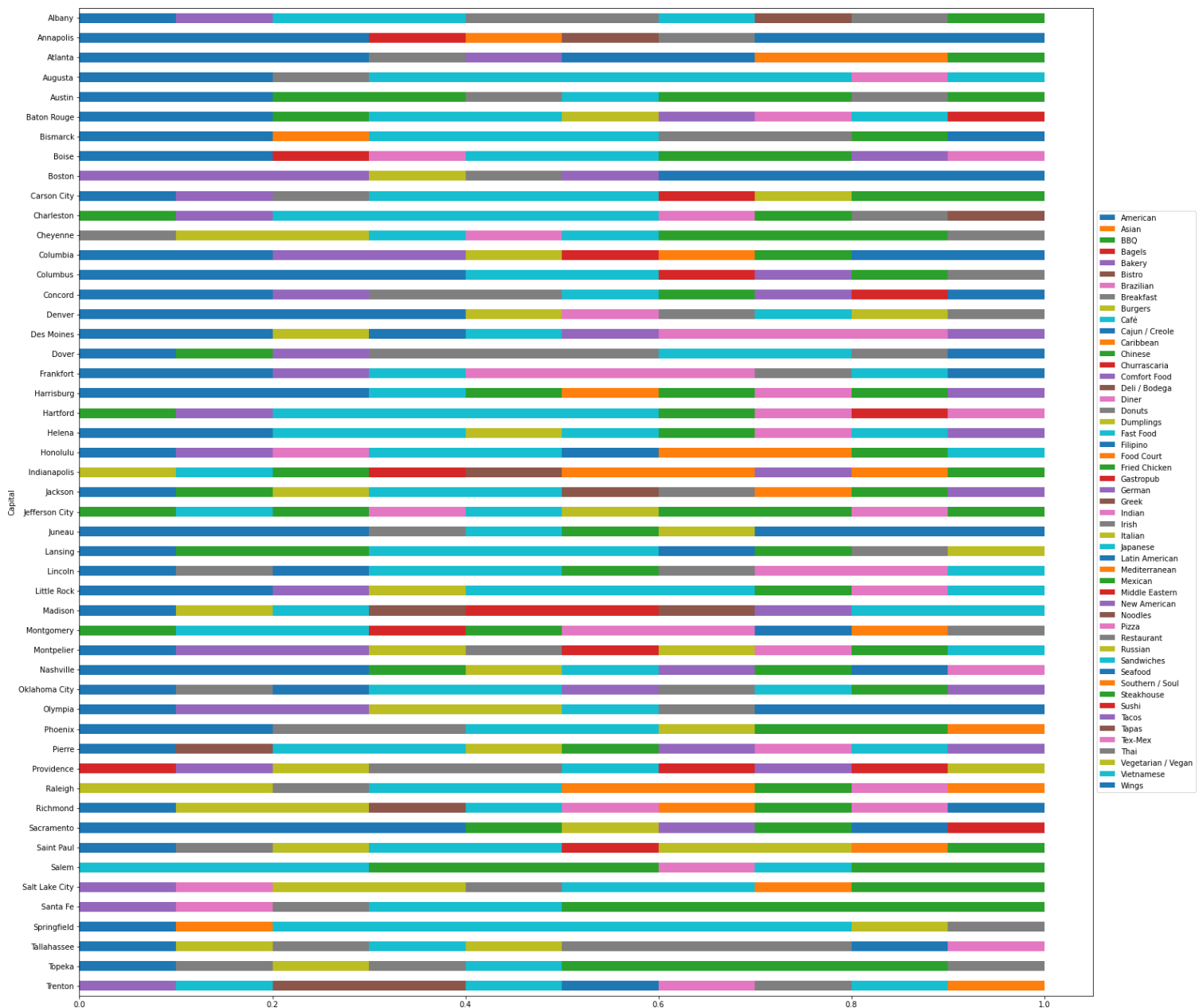


Figure 1: Stacked bar chart showing culinary heterogeneity in each capital city.