



Comprehensive Market Analysis of US Fast Food Restaurants

**SHIBBIR AHMED ARIF
MS – DATA SCIENCE
MONTCLAIR STATE UNIVERSITY**



Report Outline

- ▶ Objective
- ▶ Data Collection
- ▶ Data Load
- ▶ Data Exploration
- ▶ Data Cleaning
- ▶ Data Analysis
- ▶ Map Visualization
- ▶ Aggregation Visualization
- ▶ Interactive Visualization
- ▶ Conclusion

Objective

- ▶ Choosing the fast food industry for market analysis.
- ▶ Focusing on the distribution of fast food restaurants across various states in the USA.
- ▶ Aiming to study the openness of the market for potential opportunities.
- ▶ Analyzing the potential benefits for new fast food stores entering the market.
- ▶ Identifying possible locations for new fast food restaurants based on the analysis.
- ▶ Considering the dataset to be representative of the broader fast food market.

Data Collection

- ▶ This is a dataset of 10,000 fast food restaurants provided by Business Database of Datafiniti.
- ▶ The dataset includes the restaurant's address, city, latitude and longitude coordinates, name, and more.
- ▶ Data is collected from 2014 to 2019.
- ▶ Data Source: <https://data.world/datafiniti/fast-food-restaurants-across-America>

Data Load

	id	dateAdded	dateUpdated	address	categories	primaryCategories	city	country
0	AWrSh_KgsVYJT2BJAzaH	2019-05-19T23:58:05Z	2019-05-19T23:58:05Z	2555 11th Avenue	Fast Food Restaurants,Hamburgers and Hot Dogs,...	Accommodation & Food Services	Greeley	US
1	AWEKIA-LIxWefVJwxG9B	2018-01-18T18:30:23Z	2019-05-19T23:45:05Z	2513 Highway 6 And 50	Restaurant,Mexican Restaurants,Fast Food Resta...	Accommodation & Food Services	Grand Junction	US
2	AWrSfAcYsVYJT2BJAzPt	2019-05-19T23:45:04Z	2019-05-19T23:45:04Z	1125 Patterson Road	Sandwich Shops,Fast Food Restaurants,Restauran...	Accommodation & Food Services	Grand Junction	US
3	AWrSa3NAQTFama1Xpkbz	2019-05-19T23:26:58Z	2019-05-19T23:26:58Z	3455 N Salida Court	Fast Food Restaurants,Mexican Restaurants,Rest...	Accommodation & Food Services	Aurora	US
4	AWrSaVGzZ4Yw-wtdgcaB	2019-05-19T23:24:38Z	2019-05-19T23:24:38Z	5225 E Colfax Avenue	Fast Food Restaurants,Mexican Restaurants,Rest...	Accommodation & Food Services	Denver	US

keys	latitude	longitude	name	postalCode	province	sourceURLs	websites
11thavenue/554191587	40.39629	-104.69699	Carl's Jr.	80631	CO	https://www.yellowpages.com/greeley-co/mip/car...	https://www.carlsjr.com/?utm_source=Yextandutm...
ay6and50/1550891556	39.08135	-108.58689	Del Taco	81505	CO	http://www.citysearch.com/profile/772076870/gr...	http://www.deltaco.com
ersonroad/-2137447852	39.09148	-108.55411	Which Wich	81506	CO	https://www.yellowpages.com/grand-junction-co/...	http://www.whichwich.com
salidacourt/1143321601	39.76369	-104.77671	Chipotle Mexican Grill	80011	CO	https://www.yellowpages.com/aurora-co/mip/chip...	http://www.chipotle.com
lifaxavenue/-864103396	39.74044	-104.92636	Taco Bell	80220	CO	https://www.yellowpages.com/denver-co/mip/taco...	https://locations.tacobell.com/co/denver/5225-...

Data Exploration

► Data Shape – (10000, 16)

Checking Missing Values

```
[8] df.isnull().sum()
```

```
id          0
dateAdded   0
dateUpdated 0
address      0
categories   0
..
name         0
postalCode   0
province     0
sourceURLs   0
websites     82
Length: 16, dtype: int64
```

As we can see that there are some missing values in the dataset

Checking Duplicate Values

```
[9] df.duplicated().any()
```

```
False
```

As we can see the output is False, meaning there is no duplicate data in the dataset

Data Cleaning

Remove Unnecessary Columns

```
[10] df.drop(columns=['id', 'dateAdded', 'dateUpdated', 'categories', 'primaryCategories', 'keys', 'sourceURLs', 'websites'], inplace=True)
```

```
# shape of dataset
print("Dimension after dropping columns:")
print(df.shape)

df.head()
```

Dimension after dropping columns:
(10000, 8)

	address	city	country	latitude	longitude	name	postalCode	province
0	2555 11th Avenue	Greeley	US	40.39629	-104.69699	Carl's Jr.	80631	CO
1	2513 Highway 6 And 50	Grand Junction	US	39.08135	-108.58689	Del Taco	81505	CO
2	1125 Patterson Road	Grand Junction	US	39.09148	-108.55411	Which Wich	81506	CO
3	3455 N Salida Court	Aurora	US	39.76369	-104.77671	Chipotle Mexican Grill	80011	CO
4	5225 E Colfax Avenue	Denver	US	39.74044	-104.92636	Taco Bell	80220	CO

Summary Statistics

Summary Statistics for Numerical Columns:

	latitude	longitude
count	10000.000000	10000.000000
mean	37.298215	-93.284547
std	5.566995	16.421424
min	19.639560	-159.464310
25%	33.850652	-98.342758
50%	37.490189	-86.677452
75%	40.808905	-81.803061
max	64.856880	-71.082180

Summary Statistics for Categorical Columns:

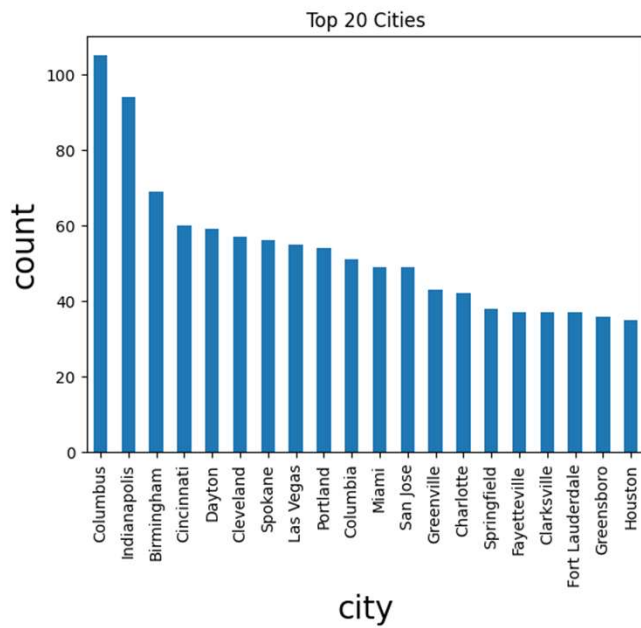
	address	city	country	name	postalCode	province
count	10000	10000	10000	10000	10000	10000
unique	9954	2810	1	1545	5310	47
top	2005 Kalia Rd	Columbus	US	McDonald's	37040	OH
freq	4	105	10000	765	17	922

In **Province** column, OH (Ohio) is the top value which shows occurrence 922 times.

In **City** column, Columbus is the top value which shows occurrence 105 times.

In **Name** column, McDonald's is the top value which shows occurrence 765 times.

Data Analysis



```
[15] rest_count_by_city = df['city'].value_counts()  
rest_count_by_city.describe()
```

```
count    2810.000000  
mean       3.558719  
std        6.260422  
min         1.000000  
25%         1.000000  
50%         2.000000  
75%         3.000000  
max        105.000000  
Name: count, dtype: float64
```

Many cities has only 1-3 restaurants opened and the mean number of restaurants opened is 3.55

Data Analysis – Cont.

How many cities have opened less than 4 fast food restaurants?

```
[16] print(len(rest_count_by_city[rest_count_by_city < 4]), "cities opened less than 4 fast food restaurants")
      print("They are", round(100*len(rest_count_by_city[rest_count_by_city < 4])/len(rest_count_by_city),2), "%"),
      print("of the total cities recorded in the dataset")
```

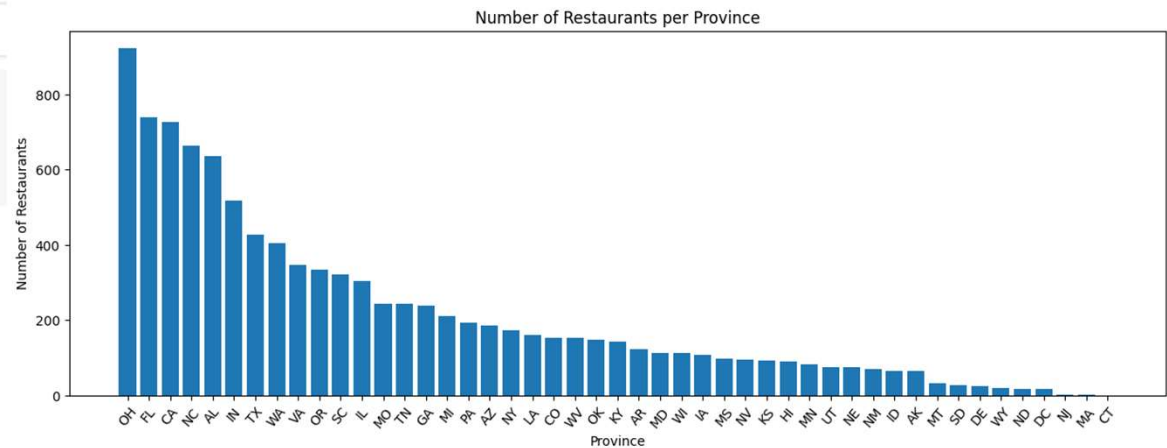
```
2121 cities opened less than 4 fast food restaurants
They are 75.48 %
of the total cities recorded in the dataset
```

1. The analysis of market openness and opportunities

▶ # Group dataset by province and count the number of restaurants

```
df1 = df.groupby('province').size().reset_index(name='restaurant_count')
df1 = df1.sort_values(by='restaurant_count', ascending=False)
df1
```

	province	restaurant_count
33	OH	922
9	FL	739
4	CA	727
26	NC	663
1	AL	635



Interactive Map Visualization 1

Finding out provinces with less than 100 and 50 restaurants

```
[19] states_less_100 = df1[df1['restaurant_count'] < 100]  
states_less_100
```

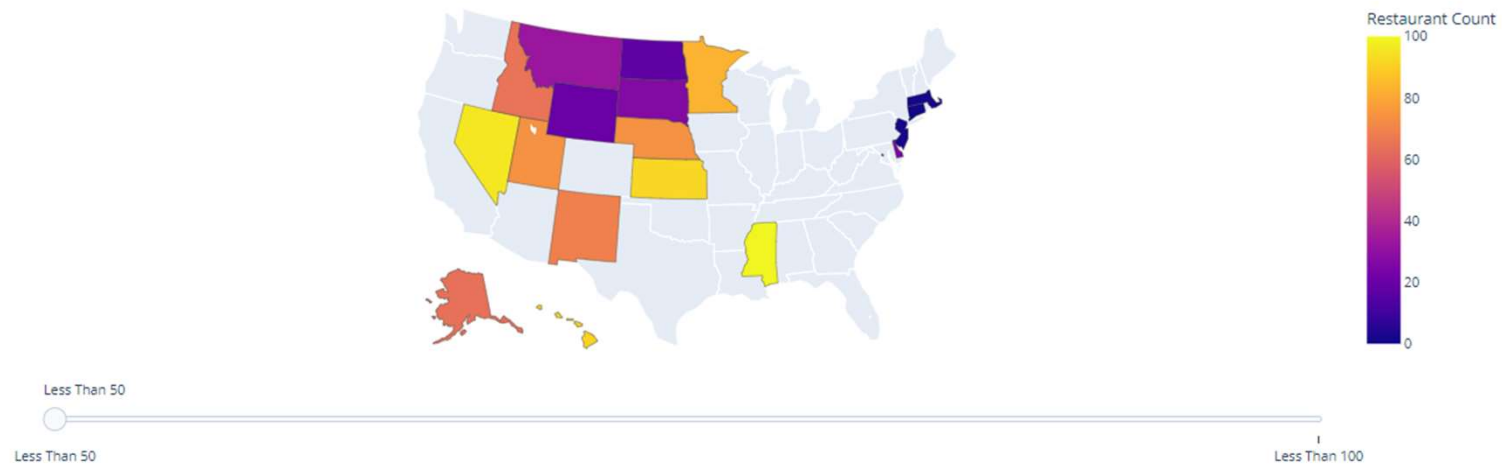
	province	restaurant_count
24	MS	99
31	NV	96
16	KS	92
11	HI	91
22	MN	83

```
states_less_50 = df1[df1['restaurant_count'] < 50]  
states_less_50
```

	province	restaurant_count
25	MT	33
38	SD	27
8	DE	25
46	WY	20
27	ND	18
7	DC	18
29	NJ	2
19	MA	2
6	CT	1

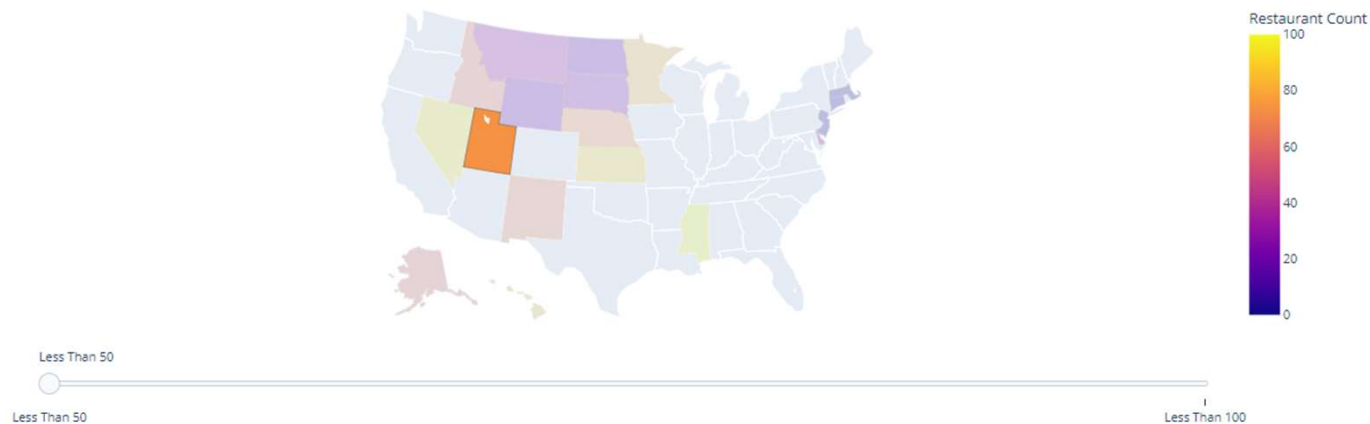
Map 1

Provinces with Less Than 50 or 100 Restaurants



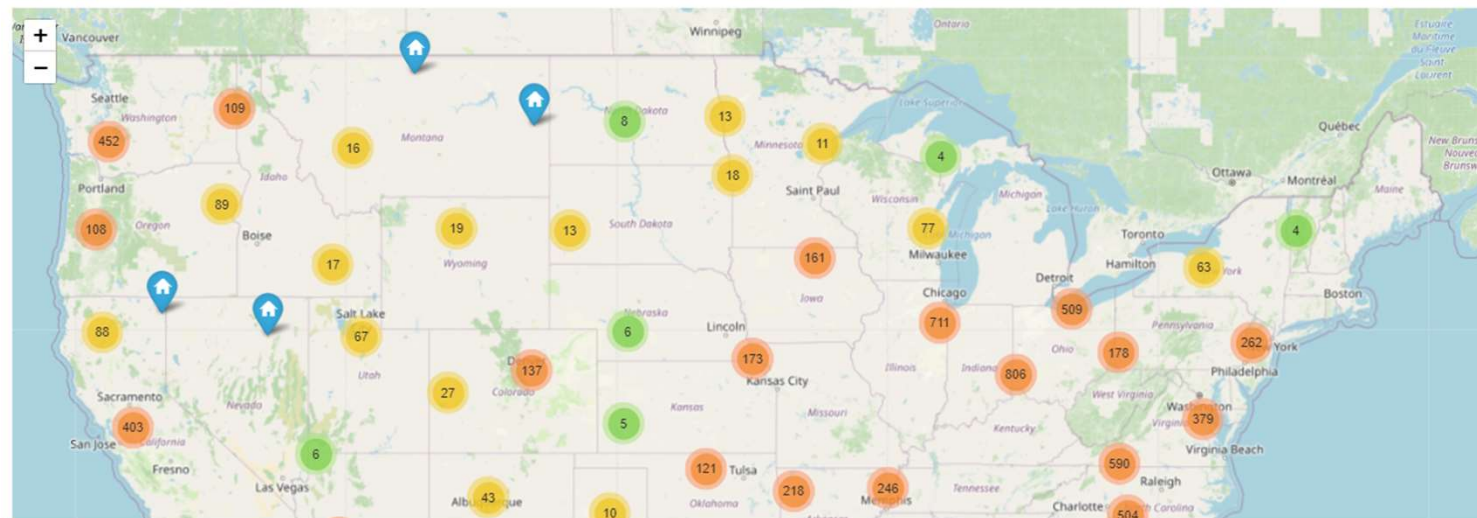
Map 1

Provinces with Less Than 50 or 100 Restaurants



Observation: Select the above areas where the amount of stores is less than 50 or 100 for the development of new fast food stores, having a large relatively unopened market and being given more opportunities, reducing competition with strong competitors in bustling areas.

Interactive Map Visualization 2



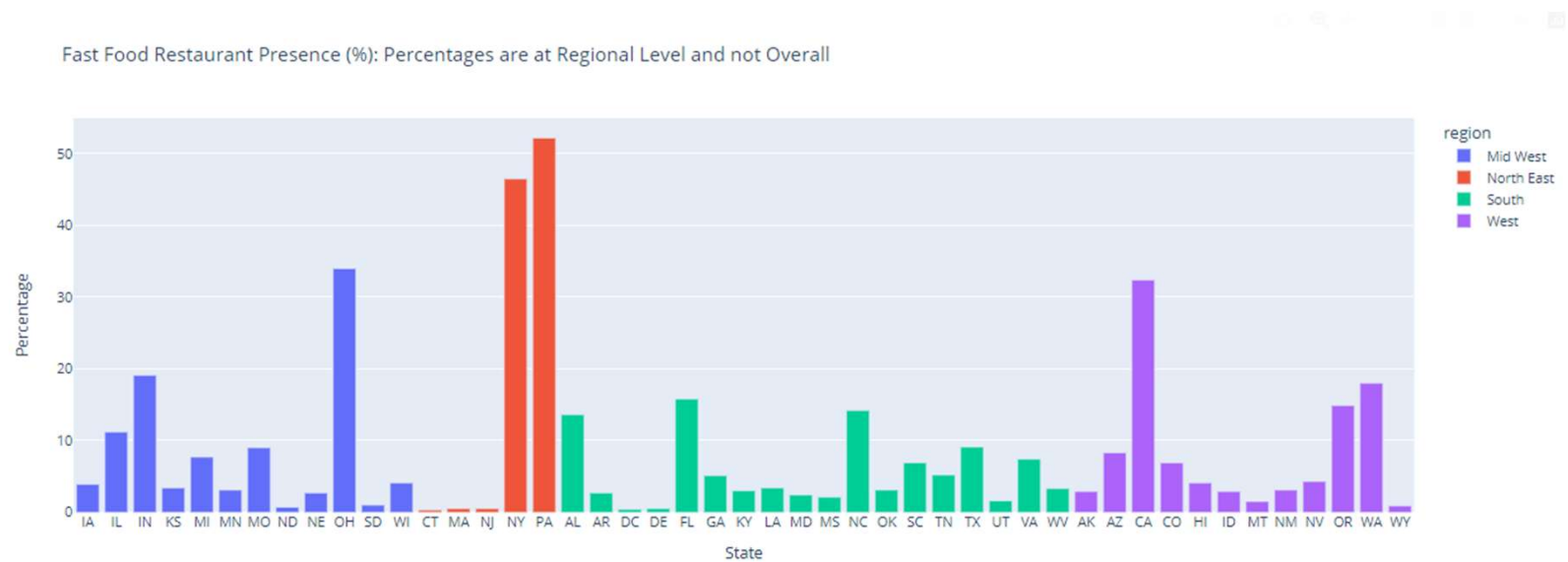
I see presence across all the states with Georgia, Indiana and Florida leading the way.

Continue

- Now lets split the entire market into 4 key regions

	address	city	country	latitude	longitude	name	postalCode	province	region
0	2555 11th Avenue	Greeley	US	40.39629	-104.69699	Carl's Jr.	80631	CO	West
1	2513 Highway 6 And 50	Grand Junction	US	39.08135	-108.58689	Del Taco	81505	CO	West
2	1125 Patterson Road	Grand Junction	US	39.09148	-108.55411	Which Wich	81506	CO	West
3	3455 N Salida Court	Aurora	US	39.76369	-104.77671	Chipotle Mexican Grill	80011	CO	West
4	5225 E Colfax Avenue	Denver	US	39.74044	-104.92636	Taco Bell	80220	CO	West

Interactive Aggregation Visualization 1



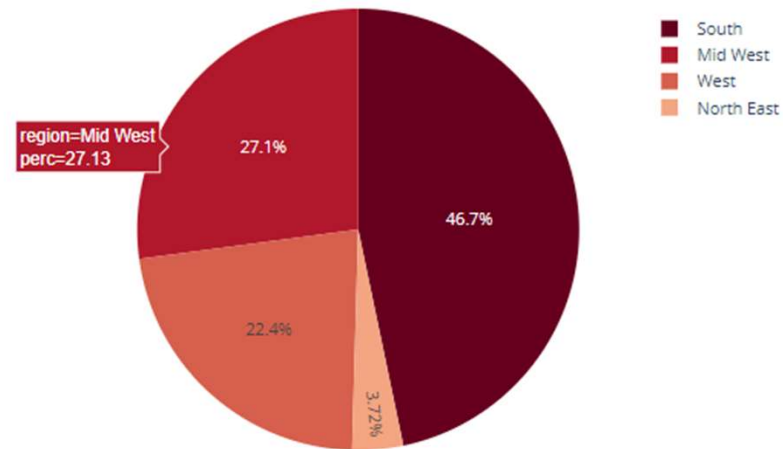
Continue

Observation

- ✓ In the West, California accounts for 32.4% of the total number of restaurants, and Washington and Oregon are still a long way off, accounting for about 18% and 14.9% respectively.
- ✓ Florida (15.8%) and North Carolina (14.2%) are the top two markets in the South, while Pennsylvania (52.2%) and New York (46.5%) lead the Northeast.
- ✓ In the Midwest, Ohio has the highest at about 34%.

Interactive Aggregation Visualization 2

Brand Presence by Region

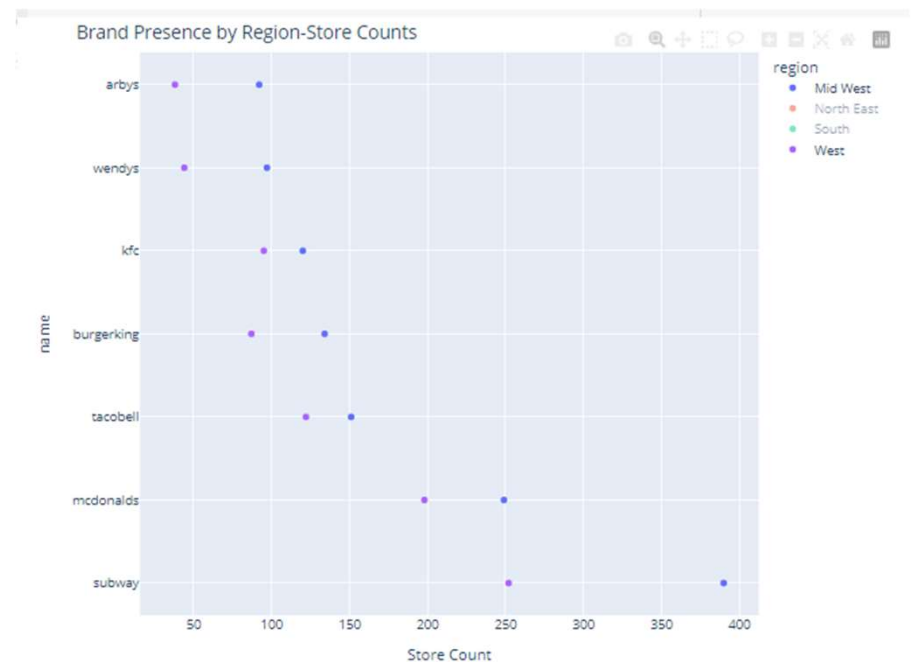
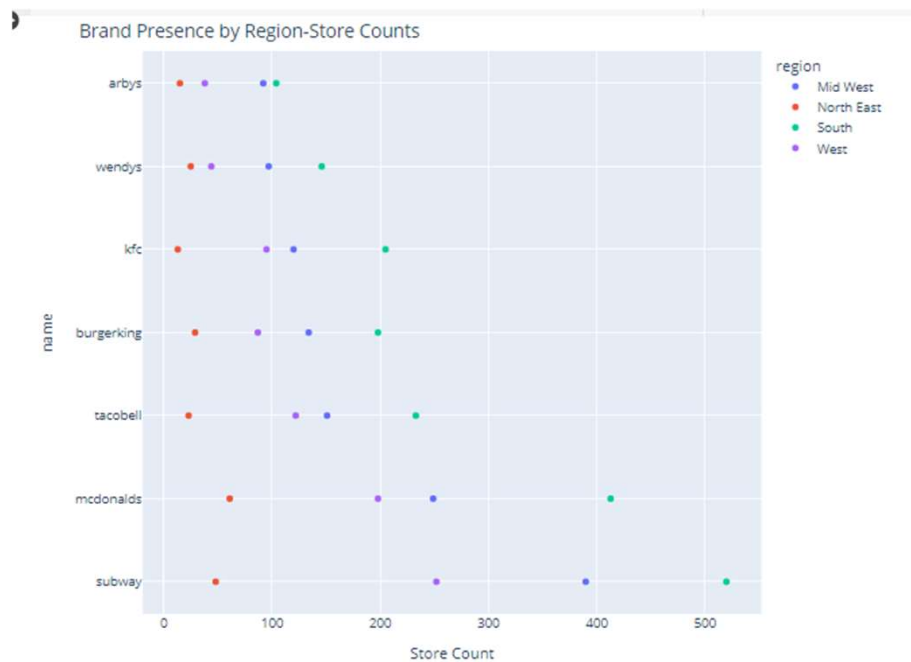


Continue

Observation

- ✓ With 46.7% of fast food restaurants located at South takes the top spot while North East contains only 3.72%.
- ✓ For Brands looking to enter the market or expand their footprint, South and Mid West are favorable.

Interactive Aggregation Visualization 3



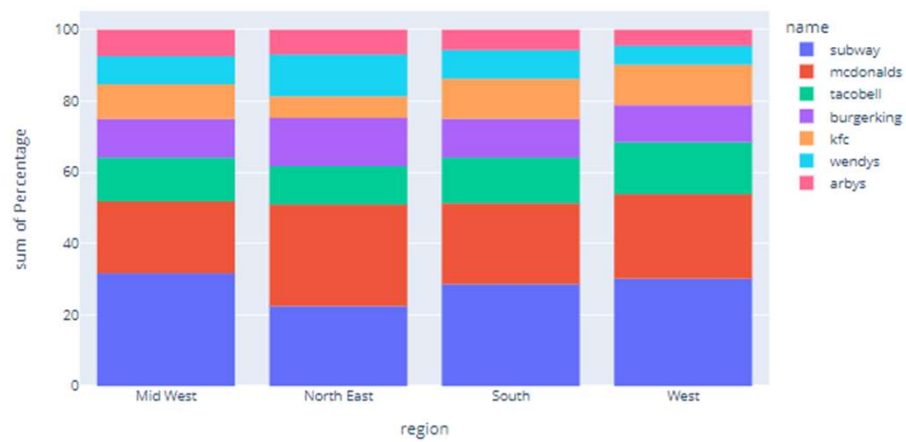
Continue

Observation

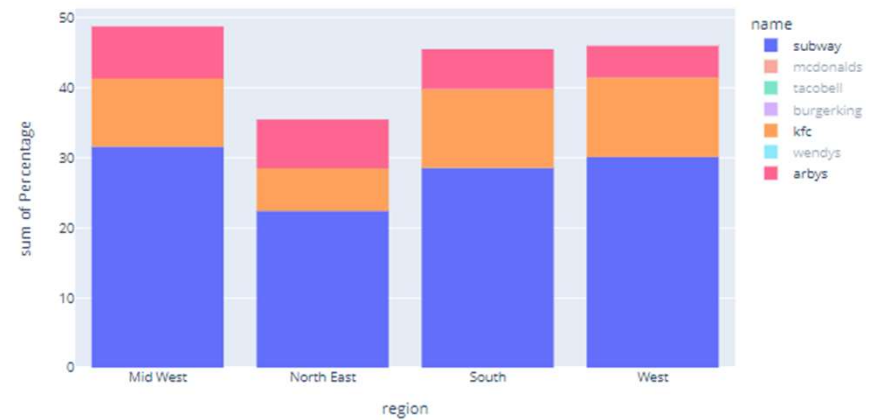
- ▶ Subway, McDonalds, Taco Bell, KFC, Burger King, Wendys and Arbys are the most popular Brands. Clearly all the Brands have big presence in South followed by Mid West.

Brand Presence by Region

Brand Presence by Region



Brand Presence by Region

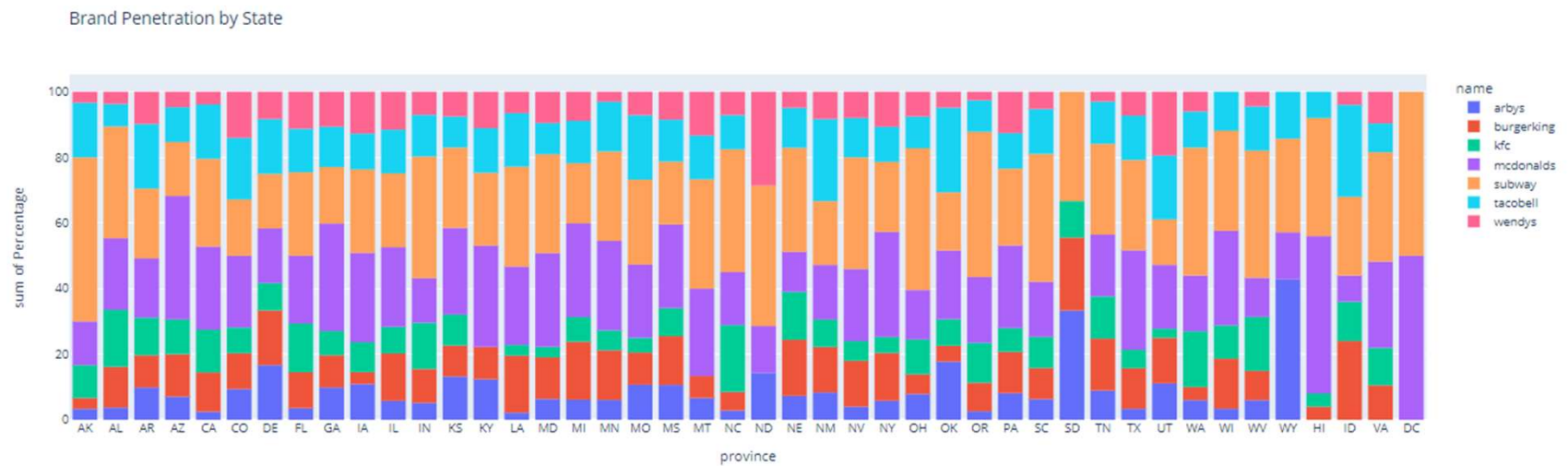


Continue

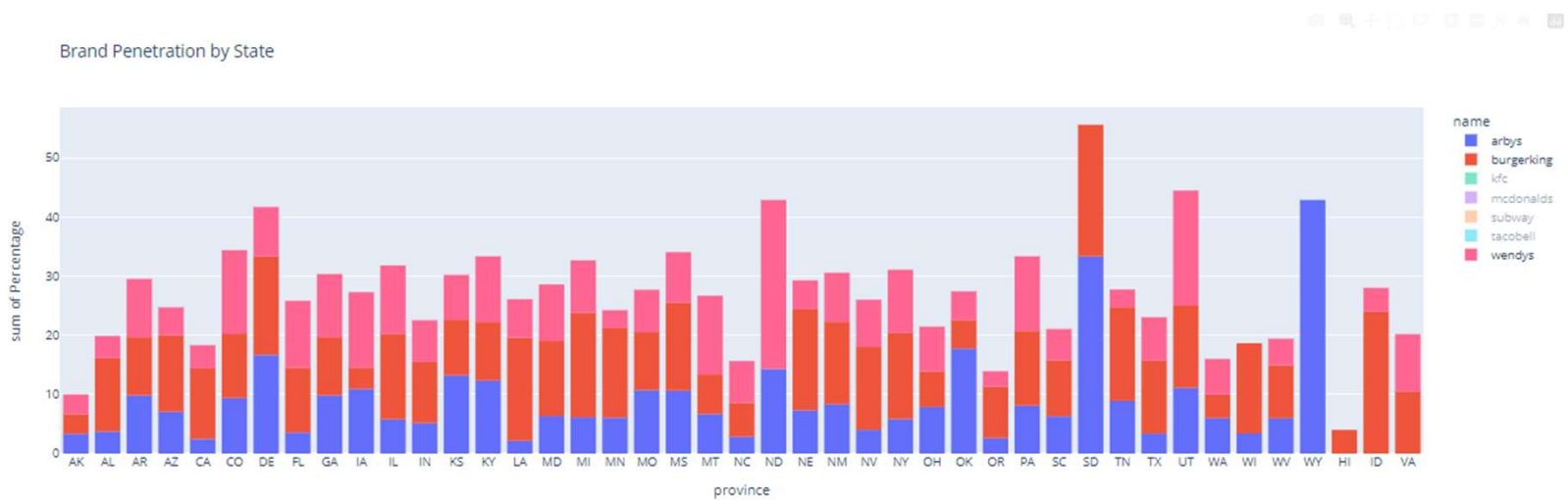
Observation

- ✓ Subway market penetration is above 28% in Mid West, South and West regions with highest of 31.6% in Mid West
- ✓ Burger King has higher penetration in North East after McDonalds and Subway.
- ✓ Taco Bell has concentrated more in West
- ✓ Arbys has least presence in West

Interactive Aggregation Visualization 4



Continue



Continue

Observation

- ✓ Analyzing by state I learn the top markets for each brand. I observe that they are present in most of the markets.
- ✓ In Washington, D.C. only 2 brands are present
- ✓ For McDonalds, Washington, D.C., Hawaii and Arizona are key markets.
- ✓ Wendys is present in all states except Wisconsin, South Dakota, Wyoming and Hawaii
- ✓ For Arbys, Wyoming and South Dakota are the key ones.
- ✓ KFC is absent in 4 states. They are - Kentucky, Montana, North Dakota and Wyoming



Conclusion

Market Landscape

- ▶ Most of the fast food restaurants are located in South and Mid West. Further in each of the four regions few states have higher market share than others by large margin.
- ▶ At National level, Ohio, Florida, California, North Carolina and Alabama are top states attracting most of the businesses

Continue

Brand Presence

- ▶ Top Brands have presence in almost all geographies with Subway having more stores than others by a considerable margin.
- ▶ This tells us that key growth strategy for Brands has been "Market Development" i.e. offer products/services in newer markets as a way to grow revenues Subway share is above 28% in all regions followed by McDonald's at 20-28%.
- ▶ Given the ubiquitous presence of brands, there is intense stiff competition in market. This creates added pressure on Brands to differentiate their product offerings
- ▶ In market characterized by high volume-low margin transactions mere survival may not be best long term strategy



Thank You!