

Analysis of the amenities around airports in the United States

A data science point of view

Introduction

1. Description of the problem
2. Data collection
3. Methodology
4. Results

Description of the problem

- Analyze the amenities available around the busiest airports in the US
- Dependencies of such amenities on several factors: geographical location? Passenger traffic?
- Business opportunities in amenities that are possibly lacking at some airport locations?

Data collection: several Wikipedia pages

Rank (2020) ↕	Airports (large hubs) ↕	IATA Code ↕	Major city served ↕	State ↕	2020** ↕	2019 ^[3] ↕	2018 ^[4] ↕	2017 ^[5] ↕	2016 ^[6] ↕	2015 ^[7] ↕	2014 ^[8] ↕	2013 ^[9] ↕	2012 ^[10] ↕	2011 ^[11] ↕	2010 ^[12] ↕	2009 ^[13] ↕
1	Hartsfield-Jackson International Airport	ATL	Atlanta	GA	42,918,685	53,505,795	51,866,464	50,251,964	50,501,858	49,340,732	46,604,273	45,308,407	45,798,809	44,414,121	43,130,585	42,280,868
2	Dallas/Fort Worth International Airport	DFW	Dallas	TX	39,364,990	35,778,573	32,800,721	31,816,933	31,283,579	31,589,832	30,766,940	29,038,128	28,022,877	27,518,358	27,100,656	26,663,984
3	Denver International Airport	DEN	Denver	CO	33,741,129	33,592,945	31,363,573	29,809,097	28,267,394	26,280,043	26,000,591	25,496,885	25,799,832	25,667,499	25,241,962	24,013,669
4	O'Hare International Airport	ORD	Chicago	IL	30,860,251	40,871,223	39,874,879	38,593,028	37,589,899	36,305,668	33,686,811	32,317,835	32,171,743	31,892,301	32,171,831	31,135,732
5	Los Angeles International Airport	LAX	Los Angeles	CA	28,779,527	42,939,104	42,626,783	41,232,432	39,636,042	36,351,226	34,314,197	32,425,892	31,326,268	30,528,737	28,857,755	27,439,897
6	Charlotte Douglas International Airport	CLT	Charlotte	NC	27,205,082	24,199,688	22,283,574	22,011,251	21,511,880	21,913,166	21,542,277	21,346,601	20,032,426	19,022,535	18,629,181	18,165,476
7	George Bush Intercontinental Airport	IAH	Houston	TX	24,690,222	21,905,309	21,157,948	19,603,731	20,062,072	20,595,874	19,772,054	18,952,840	19,038,958	19,306,660	19,528,631	19,290,239
8	McCarran International Airport	LAS	Las Vegas	NV	22,201,479	24,728,361	23,655,285	23,364,393	22,833,267	21,824,231	20,551,016	19,946,179	19,941,173	19,854,759	18,996,738	19,445,952
9	Phoenix Sky Harbor International Airport	PHX	Phoenix	AZ	21,978,708	22,433,552	21,623,842	21,185,458	20,896,265	21,351,445	20,344,867	19,525,109	19,556,189	19,750,306	18,907,171	18,559,647
10	Orlando International Airport	MCO	Orlando	FL	21,617,803	24,562,271	23,184,634	21,565,448	20,283,541	18,759,938	17,278,608	16,884,524	17,159,425	17,250,415	17,017,491	16,371,016

Page containing all the major airports

Data collection: several Wikipedia pages

Rank (2020) ↕	Airports (large hubs) ↕	IATA Code ↕	Major city served ↕	State ↕	2020** ↕	2019 ^[3] ↕	2018 ^[4] ↕	2017 ^[5] ↕	2016 ^[6] ↕	2015 ^[7] ↕	2014 ^[8] ↕	2013 ^[9] ↕	2012 ^[10] ↕	2011 ^[11] ↕	2010 ^[12] ↕	2009 ^[13] ↕
1	Hartsfield-Jackson International Airport	ATL	Atlanta	GA	42,918,685	53,505,795	51,866,464	50,251,964	50,501,858	49,340,732	46,604,273	45,308,407	45,798,809	44,414,121	43,130,585	42,280,868
2	Dallas/Fort Worth International Airport	DFW	Dallas	TX	39,364,990	35,778,573	32,800,721	31,816,933	31,283,579	31,589,832	30,766,940	29,038,128	28,022,877	27,518,358	27,100,656	26,663,984
3	Denver International Airport	DEN	Denver	CO	33,741,129	33,592,945	31,363,573	29,809,097	28,267,394	26,280,043	26,000,591	25,496,885	25,799,832	25,667,499	25,241,962	24,013,669
4	O'Hare International Airport	ORD	Chicago	IL	30,860,251	40,871,223	39,874,879	38,593,028	37,589,899	36,305,668	33,686,811	32,317,835	32,171,743	31,892,301	32,171,831	31,135,732
5	Los Angeles International Airport	LAX	Los Angeles	CA	28,779,527	42,939,104	42,626,783	41,232,432	39,636,042	36,351,226	34,314,197	32,425,892	31,326,268	30,528,737	28,857,755	27,439,897
6	Charlotte Douglas International Airport	CLT	Charlotte	NC	27,205,082	24,199,688	22,283,574	22,011,251	21,511,880	21,913,166	21,542,277	21,346,601	20,032,426	19,022,535	18,629,181	18,165,476
7	George Bush Intercontinental Airport	IAH	Houston	TX	24,690,222	21,905,309	21,157,948	19,603,731	20,062,072	20,595,874	19,772,054	18,952,840	19,038,958	19,306,660	19,528,631	19,290,239
8	McCarran International Airport	LAS	Las Vegas	NV	22,201,479	24,728,361	23,655,285	23,364,393	22,833,267	21,824,231	20,551,016	19,946,179	19,941,173	19,854,759	18,996,738	19,445,952
9	Phoenix Sky Harbor International Airport	PHX	Phoenix	AZ	21,978,708	22,433,552	21,623,842	21,185,458	20,896,265	21,351,445	20,344,867	19,525,109	19,556,189	19,750,306	18,907,171	18,559,647
10	Orlando International Airport	MCO	Orlando	FL	21,617,803	24,562,271	23,184,634	21,565,448	20,283,541	18,759,938	17,278,608	16,884,524	17,159,425	17,250,415	17,017,491	16,371,016

Page containing all the major airports

Hartsfield–Jackson Atlanta International Airport

From Wikipedia, the free encyclopedia

(Redirected from Hartsfield-Jackson International Airport)

Coordinates: 33°38′12″N 084°25′41″W﻿ / ﻿33.63667°N 84.42806°W﻿ / 33.63667; -84.42806

"Atlanta Airport" redirects here. For the airport in Idaho, see [Atlanta Airport \(Idaho\)](#). For the airport in Atlanta, Texas, see [Hall-Miller Municipal Airport](#).

Hartsfield–Jackson Atlanta International Airport (IATA: **ATL**, ICAO: **KATL**, FAA LID: **ATL**), also known as **Atlanta Hartsfield–Jackson International Airport**, **Atlanta Airport**, **Hartsfield**, **Hartsfield–Jackson** or by its airport code **ATL** is the primary [international airport](#) serving [Atlanta](#), [Georgia](#), [United States](#). The airport is located 7 mi (11 km) south of the [Downtown Atlanta](#) district. It is named after former Atlanta mayors [William B. Hartsfield](#) and [Maynard Jackson](#).^[2] ATL covers 4,700 acres (1,902 ha) of land and has five parallel runways.^{[3][2]} It was the world's busiest airport by passenger traffic from 1998 to 2019, losing its title in 2020 due to the effect of the [COVID-19 pandemic](#) in the United States and being surpassed by [Guangzhou Baiyun International Airport](#) as of a result.^[4]

Hartsfield–Jackson is the primary [hub](#) of [Delta Air Lines](#), and is a focus city for low-cost carriers [Frontier Airlines](#) and [Southwest Airlines](#). With just over 1,000 flights a day to 225 domestic and international destinations, the Delta hub is the world's largest airline hub.^{[5][6]} In addition to hosting Delta's corporate headquarters, Hartsfield–Jackson is also the home of [Delta's](#)

Hartsfield–Jackson Atlanta International Airport

 Hartsfield-Jackson Atlanta International Airport



For each airport, page containing coordinates

Data collection: several Wikipedia pages

Rank (2020) ↕	Airports (large hubs) ↕	IATA Code ↕	Major city served ↕	State ↕	2020** ↕	2019 ^[3] ↕	2018 ^[4] ↕	2017 ^[5] ↕	2016 ^[6] ↕	2015 ^[7] ↕	2014 ^[8] ↕	2013 ^[9] ↕	2012 ^[10] ↕	2011 ^[11] ↕	2010 ^[12] ↕	2009 ^[13] ↕
1	Hartsfield-Jackson International Airport	ATL	Atlanta	GA	42,918,685	53,505,795	51,866,464	50,251,964	50,501,858	49,340,732	46,604,273	45,308,407	45,798,809	44,414,121	43,130,585	42,280,868
2	Dallas/Fort Worth International Airport	DFW	Dallas	TX	39,364,990	35,778,573	32,800,721	31,816,933	31,283,579	31,589,832	30,766,940	29,038,128	28,022,877	27,518,358	27,100,656	26,663,984
3	Denver International Airport	DEN	Denver	CO	33,741,129	33,592,945	31,363,573	29,809,097	28,267,394	26,280,043	26,000,591	25,496,885	25,799,832	25,667,499	25,241,962	24,013,669
4	O'Hare International Airport	ORD	Chicago	IL	30,860,251	40,871,223	39,874,879	38,593,028	37,589,899	36,305,668	33,686,811	32,317,835	32,171,743	31,892,301	32,171,831	31,135,732
5	Los Angeles International Airport	LAX	Los Angeles	CA	28,779,527	42,939,104	42,626,783	41,232,432	39,636,042	36,351,226	34,314,197	32,425,892	31,326,268	30,528,737	28,857,755	27,439,897
6	Charlotte Douglas International Airport	CLT	Charlotte	NC	27,205,082	24,199,688	22,283,574	22,011,251	21,511,880	21,913,166	21,542,277	21,346,601	20,032,426	19,022,535	18,629,181	18,165,476
7	George Bush Intercontinental Airport	IAH	Houston	TX	24,690,222	21,905,309	21,157,948	19,603,731	20,062,072	20,595,874	19,772,054	18,952,840	19,038,958	19,306,660	19,528,631	19,290,239
8	McCarran International Airport	LAS	Las Vegas	NV	22,201,479	24,728,361	23,655,285	23,364,393	22,833,267	21,824,231	20,551,016	19,946,179	19,941,173	19,854,759	18,996,738	19,445,952
9	Phoenix Sky Harbor International Airport	PHX	Phoenix	AZ	21,978,708	22,433,552	21,623,842	21,185,458	20,896,265	21,351,445	20,344,867	19,525,109	19,556,189	19,750,306	18,907,171	18,559,647
10	Orlando International Airport	MCO	Orlando	FL	21,617,803	24,562,271	23,184,634	21,565,448	20,283,541	18,759,938	17,278,608	16,884,524	17,159,425	17,250,415	17,017,491	16,371,016

Page containing all the major airports

Hartsfield–Jackson Atlanta International Airport

From Wikipedia, the free encyclopedia
(Redirected from Hartsfield-Jackson International Airport)

"Atlanta Airport" redirects here. For the airport in Idaho, see Atlanta Airport (Idaho). For the airport in Atlanta, Texas, see Hall-Miller Municipal Airport.

Hartsfield–Jackson Atlanta International Airport (IATA: **ATL**, ICAO: **KATL**, FAA LID: **ATL**), also known as **Atlanta Hartsfield–Jackson International Airport**, **Atlanta Airport**, **Hartsfield**, **Hartsfield–Jackson** or by its airport code **ATL** is the primary [international airport](#) serving [Atlanta](#), [Georgia](#), [United States](#). The airport is located 7 mi (11 km) south of the [Downtown Atlanta](#) district. It is named after former Atlanta mayors [William B. Hartsfield](#) and [Maynard Jackson](#).^[2] ATL covers 4,700 acres (1,902 ha) of land and has five parallel runways.^{[3][2]} It was the world's busiest airport by passenger traffic from 1998 to 2019, losing its title in 2020 due to the effect of the [COVID-19 pandemic](#) in the United States and being surpassed by [Guangzhou Baiyun International Airport](#) as of a result.^[4]

Hartsfield–Jackson is the primary [hub](#) of [Delta Air Lines](#), and is a focus city for low-cost carriers [Frontier Airlines](#) and [Southwest Airlines](#). With just over 1,000 flights a day to 225 domestic and international destinations, the Delta hub is the world's largest airline hub.^{[5][6]} In addition to hosting Delta's corporate headquarters, Hartsfield–Jackson is also the home of [Delta's](#)

Coordinates: 33°38′12″N 084°25′41″W﻿ / ﻿33.63667°N 84.42806°W﻿ / 33.63667; -84.42806

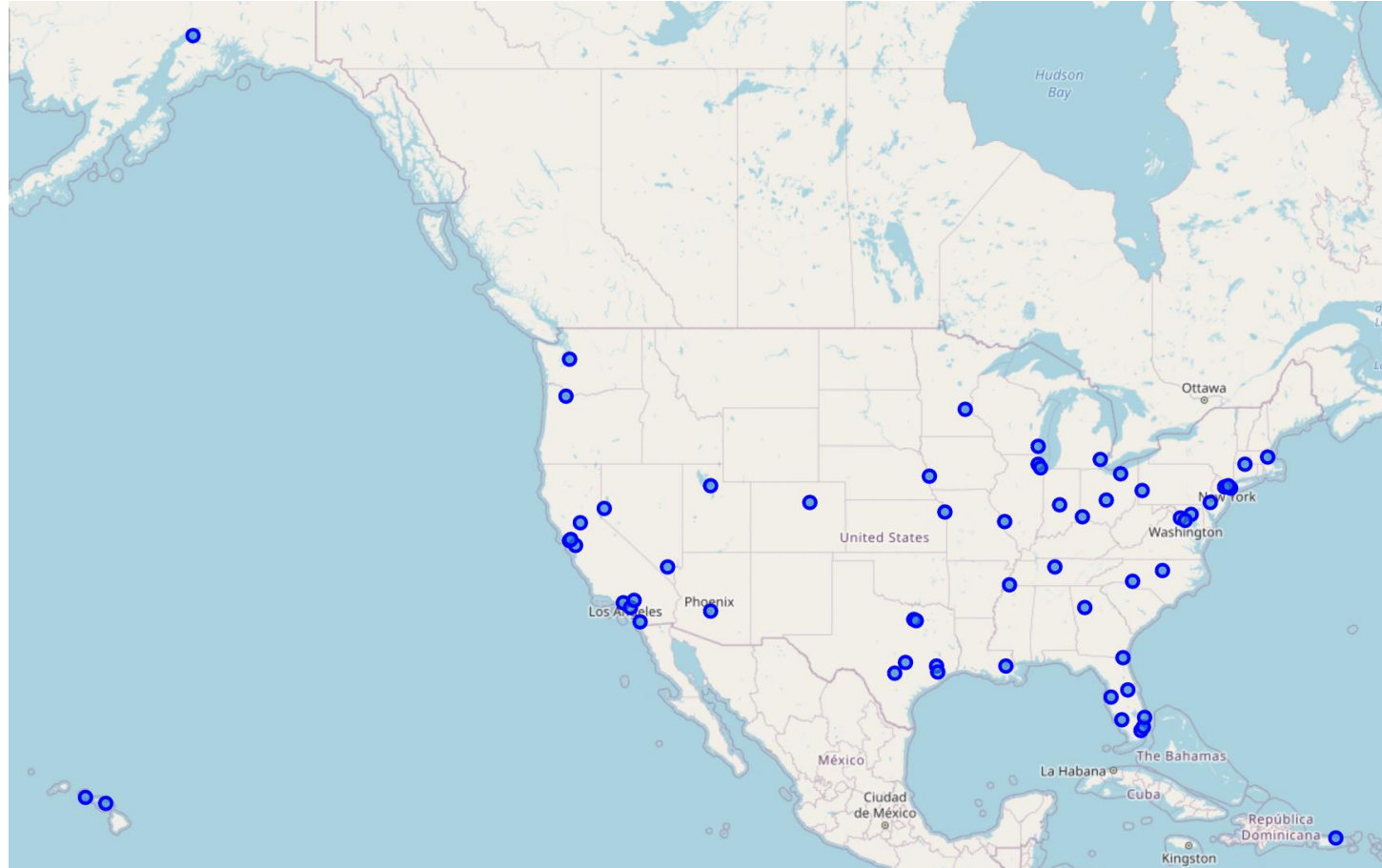


For each airport, page containing coordinates

Data collection: Airport location

	Airport name	Code	City	State	Passengers	Latitude	Longitude
0	Hartsfield-Jackson International Airport	ATL	Atlanta	GA	42,918,685	33.63667	-84.42806
1	Dallas/Fort Worth International Airport	DFW	Dallas & Ft. Worth	TX	39,364,990	32.89694	-97.03806
2	Denver International Airport	DEN	Denver	CO	33,741,129	39.86167	-104.67306
3	O'Hare International Airport	ORD	Chicago	IL	30,860,251	41.97861	-87.90472
4	Los Angeles International Airport	LAX	Los Angeles	CA	28,779,527	33.94250	-118.40806
5	Charlotte Douglas International Airport	CLT	Charlotte	NC	27,205,082	35.21389	-80.94306
6	McCarran International Airport	LAS	Las Vegas	NV	22,201,479	36.08000	-115.15222
7	Phoenix Sky Harbor International Airport	PHX	Phoenix	AZ	21,978,708	33.43417	-112.01167
8	Orlando International Airport	MCO	Orlando	FL	21,617,803	28.42944	-81.30889
9	Seattle-Tacoma International Airport	SEA	Seattle	WA	20,061,507	47.44889	-122.30944

Data collection: map of airports



Data collection: call foursquare

	Airport name	Venue name	Venue type
0	Hartsfield-Jackson International Airport	Chicken & Beer	Fried Chicken Joint
1	Hartsfield-Jackson International Airport	One Flew South	Sushi Restaurant
2	Hartsfield-Jackson International Airport	Delta Sky Club	Airport Lounge
3	Hartsfield-Jackson International Airport	The Sky Deck	Roof Deck
4	Hartsfield-Jackson International Airport	Delta Sky Club	Airport Lounge
...
5992	Reno-Tahoe International Airport	The Daily Bagel	Café
5993	Reno-Tahoe International Airport	Bistro 7	American Restaurant
5994	Reno-Tahoe International Airport	Grand Sierra Resort and Casino	Resort
5995	Reno-Tahoe International Airport	The Shore Riverfront Dining	American Restaurant
5996	Reno-Tahoe International Airport	The Brewer's Cabinet	Pub

Determine venues around each airport: radius = 10km

Methodology

1. Determine venue proportion
2. Unexpected extra pre-processing
3. Cluster airports according to surrounding venues
4. Cluster airports according to surrounding restaurants

Methodology: determine venue proportion

- C1: venue type with highest proportion
- C10: venue type with lowest proportion

	Airport name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0	Austin–Bergstrom International Airport	Coffee Shop	Food Truck	Brewery	Pizza Place	Convenience Store	Hotel	Music Venue	BBQ Joint	Trail	Racetrack
1	Baltimore/Washington International Airport	Hotel	Sandwich Place	Seafood Restaurant	American Restaurant	Fast Food Restaurant	Toy / Game Store	Pizza Place	Mexican Restaurant	Convenience Store	Asian Restaurant
2	Bradley International Airport	Hotel	American Restaurant	Sandwich Place	Italian Restaurant	Donut Shop	Pizza Place	Breakfast Spot	Pharmacy	Asian Restaurant	Liquor Store
3	Charlotte Douglas International Airport	Brewery	Bar	BBQ Joint	Mexican Restaurant	Grocery Store	Restaurant	Deli / Bodega	American Restaurant	Sushi Restaurant	Bakery
4	Chicago Midway International Airport	Pizza Place	Mexican Restaurant	Ice Cream Shop	Grocery Store	Brewery	Breakfast Spot	Liquor Store	Hot Dog Joint	American Restaurant	Bakery

Unfortunately, similar venues fall into separate categories and may bias the analysis
For example, all the different restaurants will have much more weight than other types of venues

Methodology: extra pre-processing

- Group all restaurants into one category: 'Restaurant'
- Group all bars into one category: 'Bar'
- Repeat the same process for gyms, ice cream places, grocery stores and all venues inside airport terminals
- **Note that we will DROP all venues inside airport terminals**

Methodology: extra pre-processing

	Airport name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0	Austin–Bergstrom International Airport	Restaurant	Bar	Coffee Place	Convenience Store	Hotel	Racetrack	Trail	Music Venue	Garden Center	Discount Store
1	Baltimore/Washington International Airport	Restaurant	Hotel	Bar	Coffee Place	Convenience Store	Toy / Game Store	Warehouse Store	Scenic Lookout	Liquor Store	Ice Cream
2	Bradley International Airport	Restaurant	Hotel	Bar	Coffee Place	Gym	Pharmacy	Groceries	Convenience Store	Liquor Store	Museum
3	Charlotte Douglas International Airport	Restaurant	Bar	Groceries	Gym	Bakery	Clothing Store	Park	Gift Shop	Convenience Store	Miscellaneous Shop
4	Chicago Midway International Airport	Restaurant	Bar	Ice Cream	Groceries	Coffee Place	Liquor Store	Bakery	Clothing Store	Military Base	Soccer Stadium

Sample of input dataset after grouping similar categories

Methodology: extra pre-processing

	Airport name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0	Austin–Bergstrom International Airport	Restaurant	Bar	Coffee Place	Convenience Store	Hotel	Racetrack	Trail	Music Venue	Garden Center	Discount Store
1	Baltimore/Washington International Airport	Restaurant	Hotel	Bar	Coffee Place	Convenience Store	Toy / Game Store	Warehouse Store	Scenic Lookout	Liquor Store	Ice Cream
2	Bradley International Airport	Restaurant	Hotel	Bar	Coffee Place	Gym	Pharmacy	Groceries	Convenience Store	Liquor Store	Museum

‘Restaurant’ is the most present category around all airport, so it has no specificity for clustering purposes

We drop the ‘Restaurant’ category, and will study it separately later

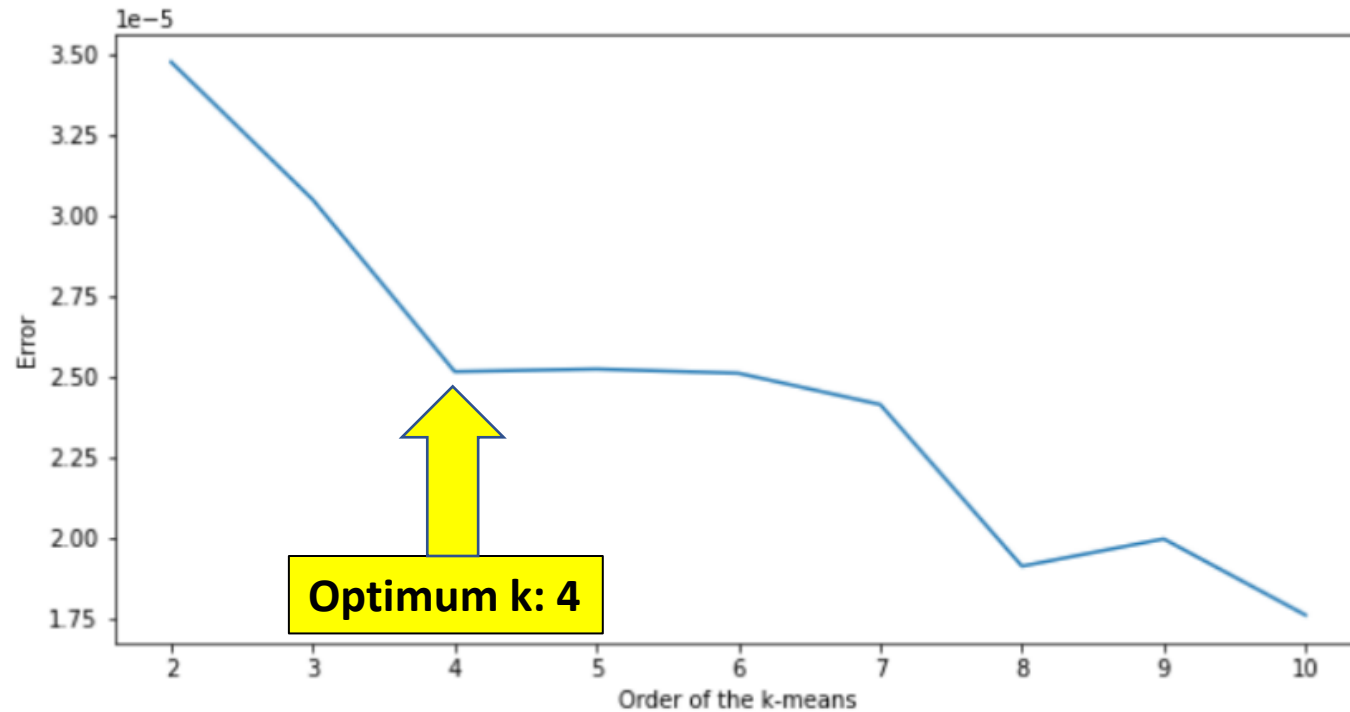
Methodology: extra pre-processing

	Airport name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0	Austin–Bergstrom International Airport	Bar	Coffee Place	Convenience Store	Hotel	Racetrack	Trail	Music Venue	Ice Cream	Discount Store	Garden Center
1	Baltimore/Washington International Airport	Hotel	Bar	Coffee Place	Toy / Game Store	Convenience Store	Ice Cream	Liquor Store	Scenic Lookout	Warehouse Store	Pet Store
2	Bradley International Airport	Hotel	Bar	Coffee Place	Gym	Pharmacy	Groceries	Convenience Store	Liquor Store	Museum	Pet Store
3	Charlotte Douglas International Airport	Bar	Groceries	Gym	Bakery	Clothing Store	Park	Gift Shop	Convenience Store	Miscellaneous Shop	Coffee Place
4	Chicago Midway International Airport	Bar	Ice Cream	Groceries	Coffee Place	Liquor Store	Bakery	Clothing Store	Military Base	Soccer Stadium	Other Great Outdoors

Sample of final input dataset: after dropping the 'Restaurant' category

Methodology: cluster airport venues

- Omitting the restaurants for now, we test k-means for $k=2, \dots, 10$



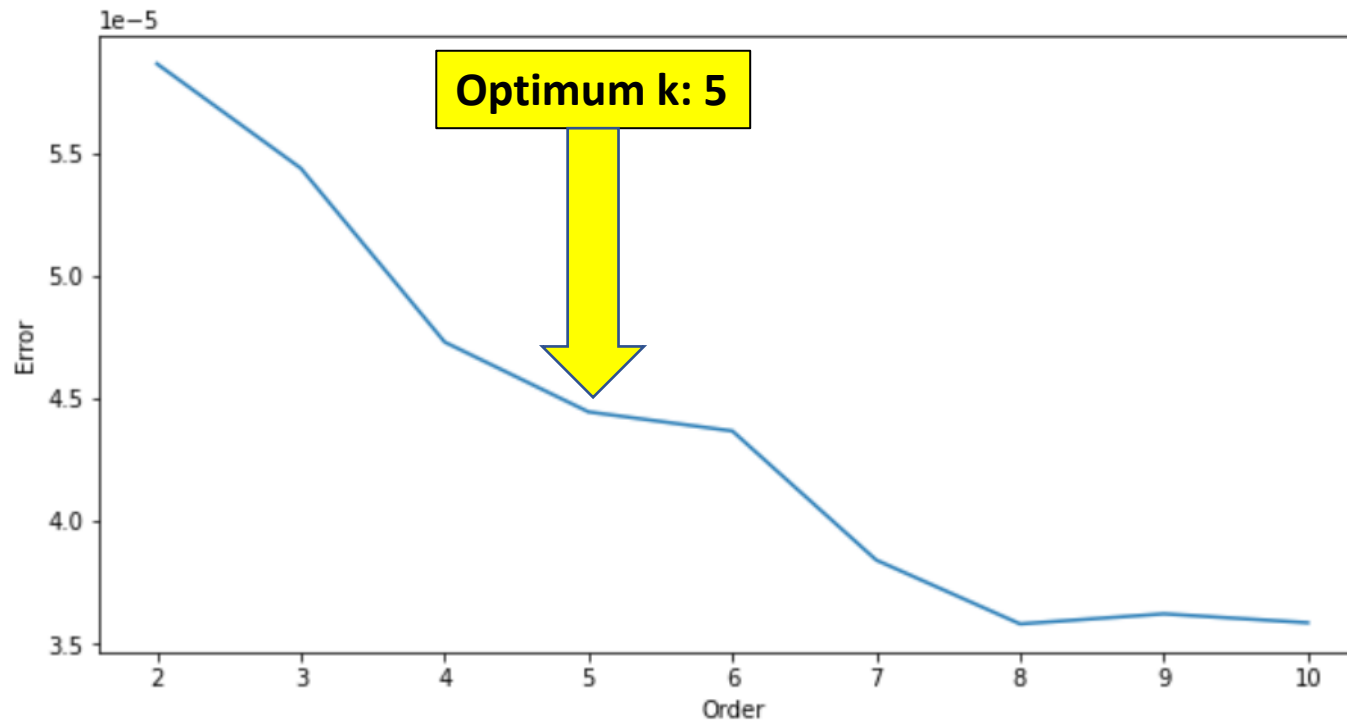
Methodology: cluster airport restaurants only

	Airport name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0	Austin–Bergstrom International Airport	Food Truck	Pizza Place	BBQ Joint	Italian Restaurant	Restaurant	Taco Place	Burger Joint	Vietnamese Restaurant	Asian Restaurant	Bagel Shop
1	Baltimore/Washington International Airport	Seafood Restaurant	Sandwich Place	American Restaurant	Fast Food Restaurant	Mexican Restaurant	Pizza Place	Asian Restaurant	Italian Restaurant	Snack Place	Sushi Restaurant
2	Bradley International Airport	American Restaurant	Italian Restaurant	Sandwich Place	Breakfast Spot	Pizza Place	Asian Restaurant	Diner	Bagel Shop	Wings Joint	Vietnamese Restaurant
3	Charlotte Douglas International Airport	Mexican Restaurant	BBQ Joint	Restaurant	Greek Restaurant	American Restaurant	Pizza Place	Asian Restaurant	Sushi Restaurant	Southern / Soul Food Restaurant	Mediterranean Restaurant
4	Chicago Midway International Airport	Pizza Place	Mexican Restaurant	Breakfast Spot	American Restaurant	Hot Dog Joint	BBQ Joint	Middle Eastern Restaurant	Taco Place	Italian Restaurant	Diner

Sample of input dataset containing all types of restaurants only

Methodology: cluster airport restaurants only

- Considering restaurants only, we test k-means for $k=2, \dots, 10$



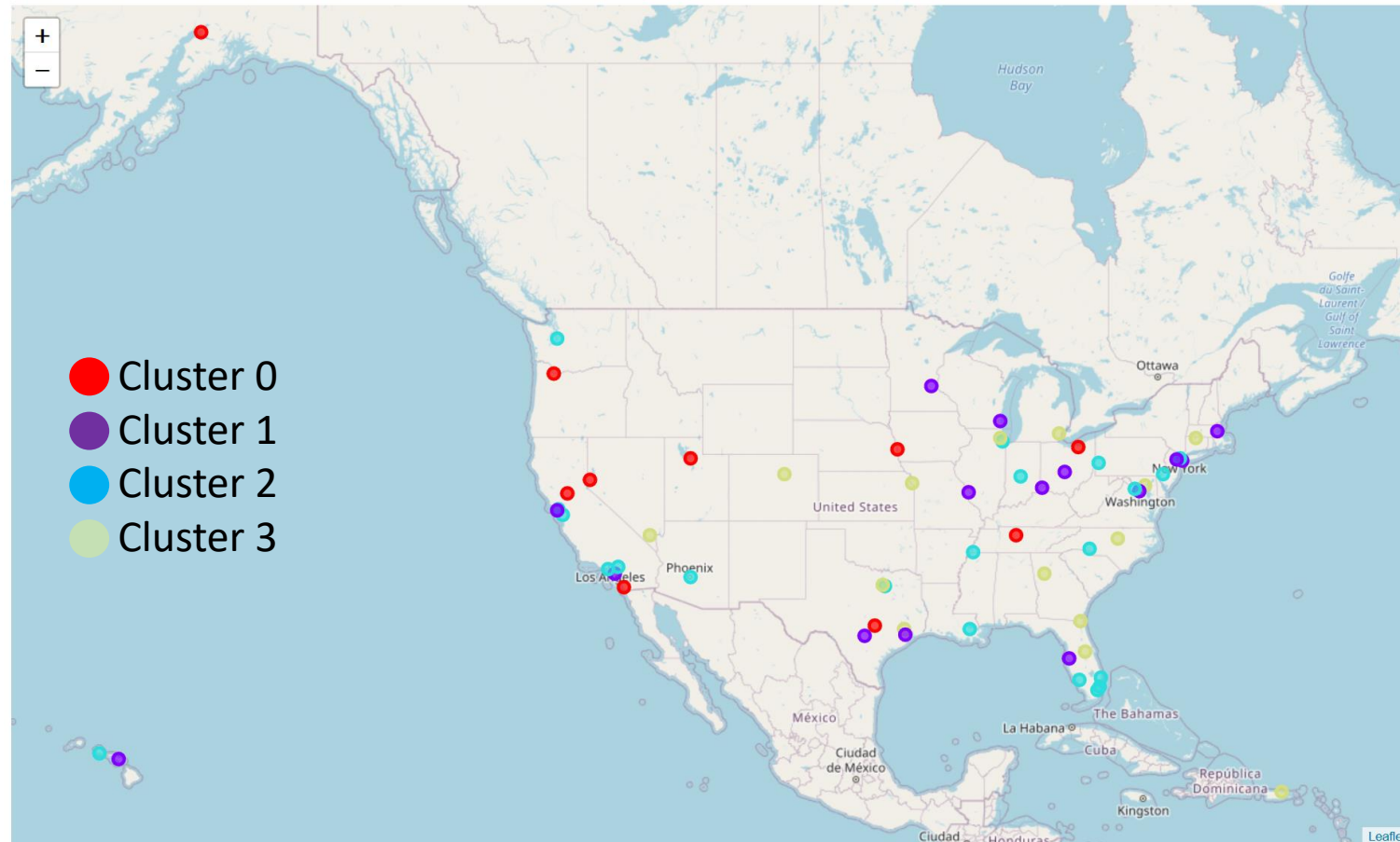
Results

1. All venues except restaurants
2. Restaurants only

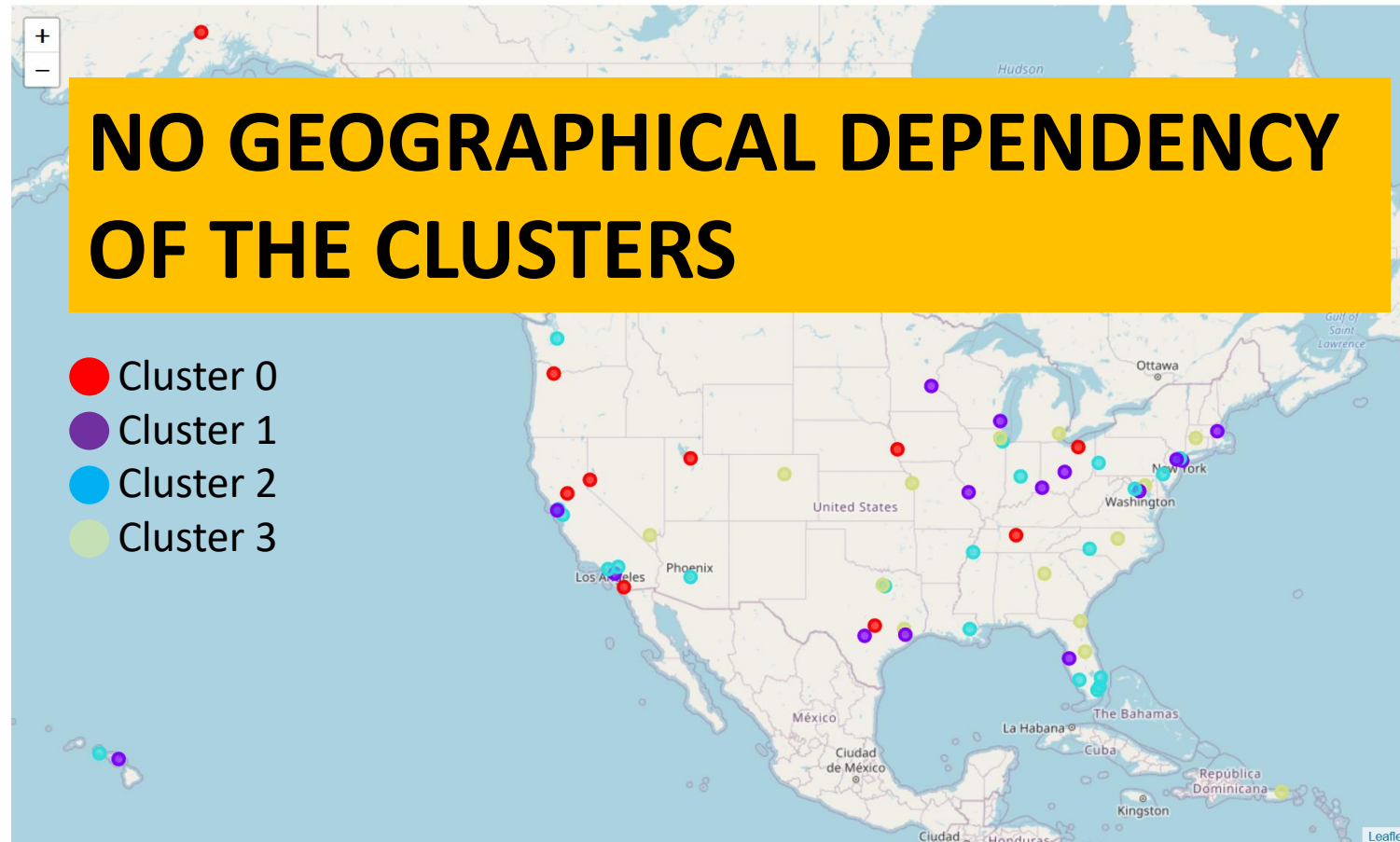
Cluster venues, no restaurants, 4-means

- Cluster 0: zoos, parks, art venues such as music venues and theaters, and no liquor stores
- Cluster 1: grocery stores, parks, gyms, beaches, gardens and trails
- Cluster 2: gyms, parks, beaches, home stores, some golf courses and museums
- Cluster 3: cluster with most hotels, cosmetic shops, convenience stores and rental car agencies

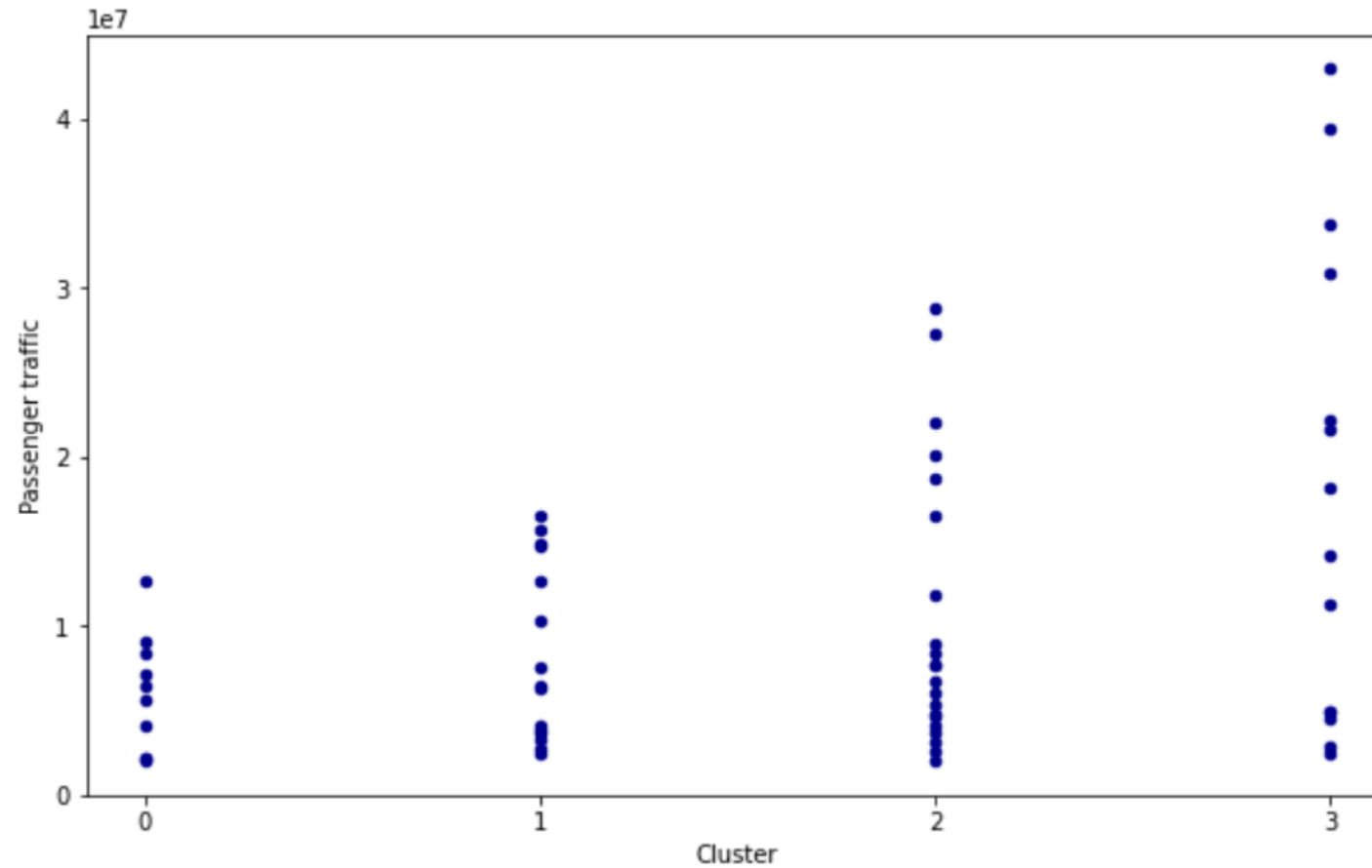
Cluster venues, no restaurants, 4-means



Cluster venues, no restaurants, 4-means



Cluster vs. passenger traffic



Cluster vs. passenger traffic

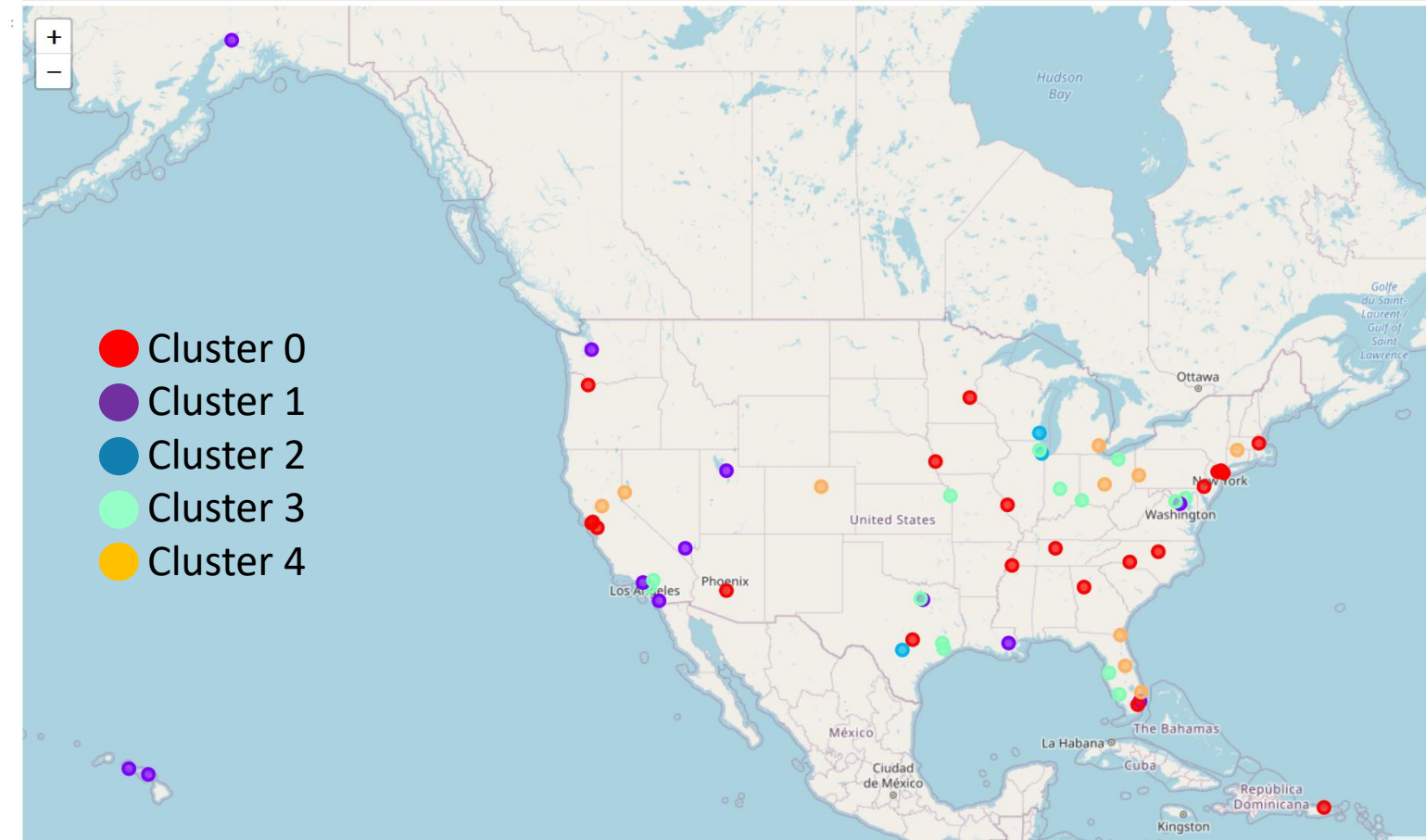
DEPENDENCY OF THE CLUSTERS ON PASSENGER TRAFFIC: HIGHER CLUSTERS INCREASINGLY INCLUDE AIRPORTS WITH MORE AND MORE TRAFFIC



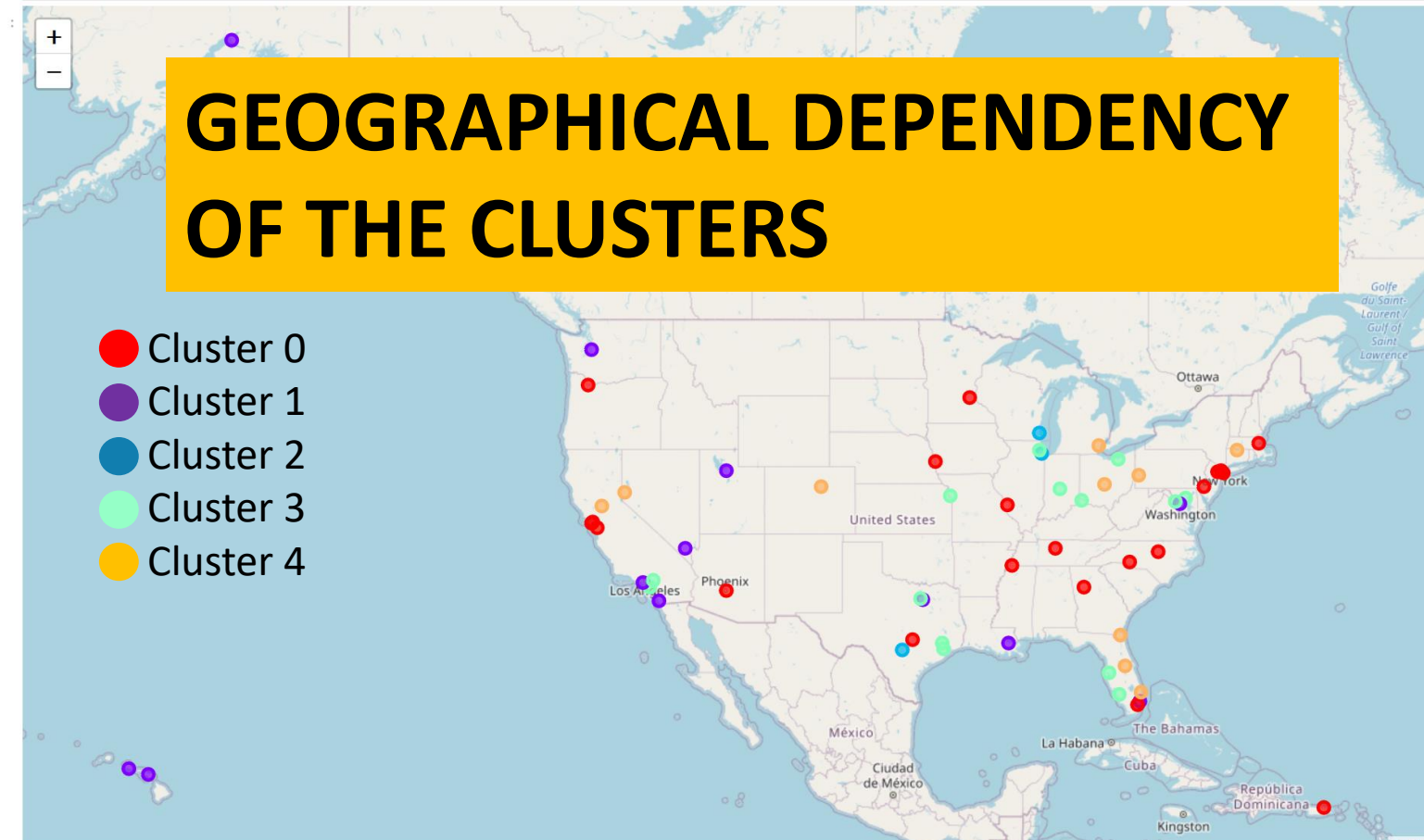
Cluster restaurants only, 5-means

- Cluster 0: Italian restaurants, BBQs and fast-food, but less than average Asian restaurants;
- Cluster 1: seafood and American restaurants;
- Cluster 2 Italian and American restaurants;
- Cluster 3 pizza places, fast food, American, seafood restaurants and steak houses;
- Cluster 4: American restaurants, fast food, Italian and Asian restaurants.

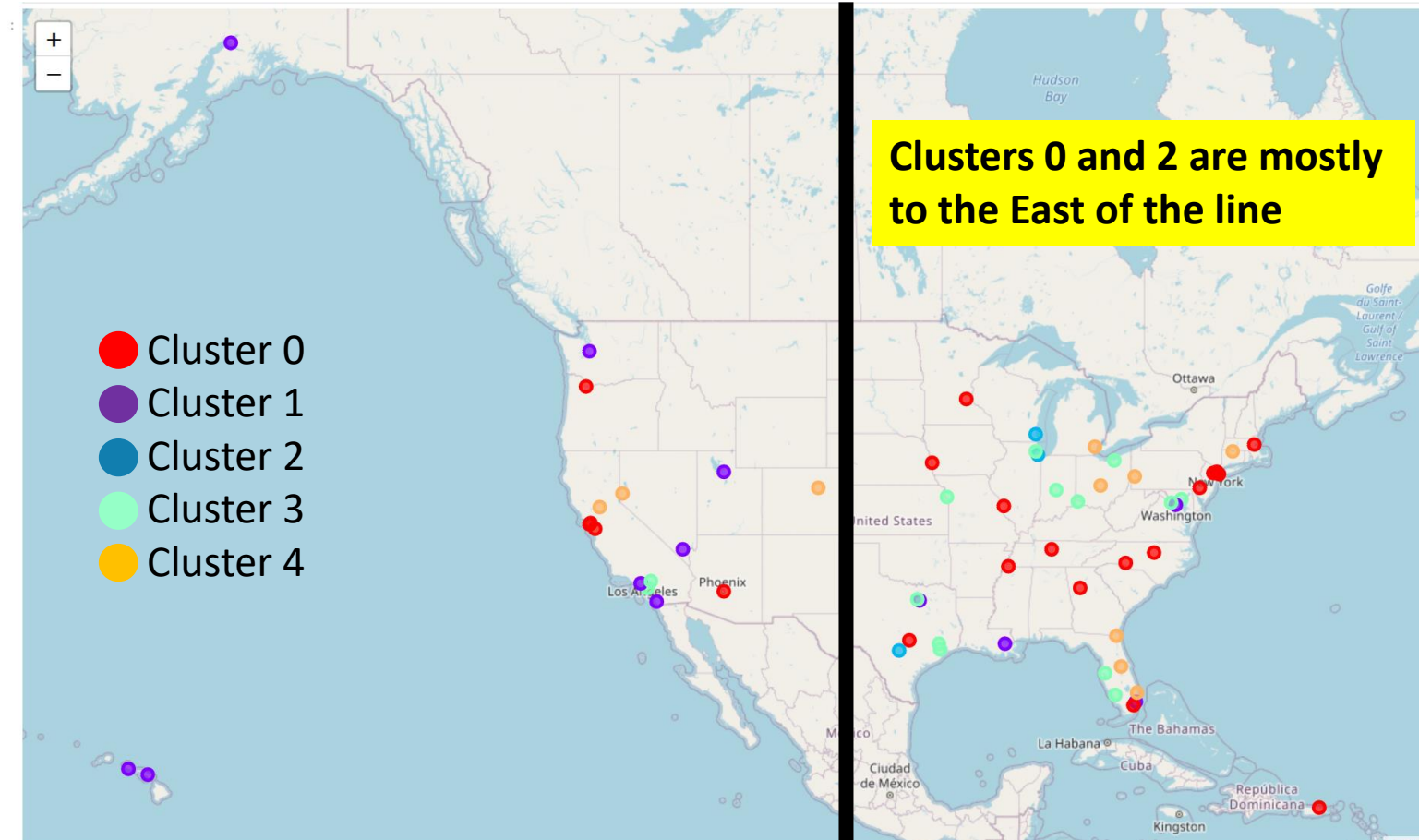
Cluster restaurants only, 5-means



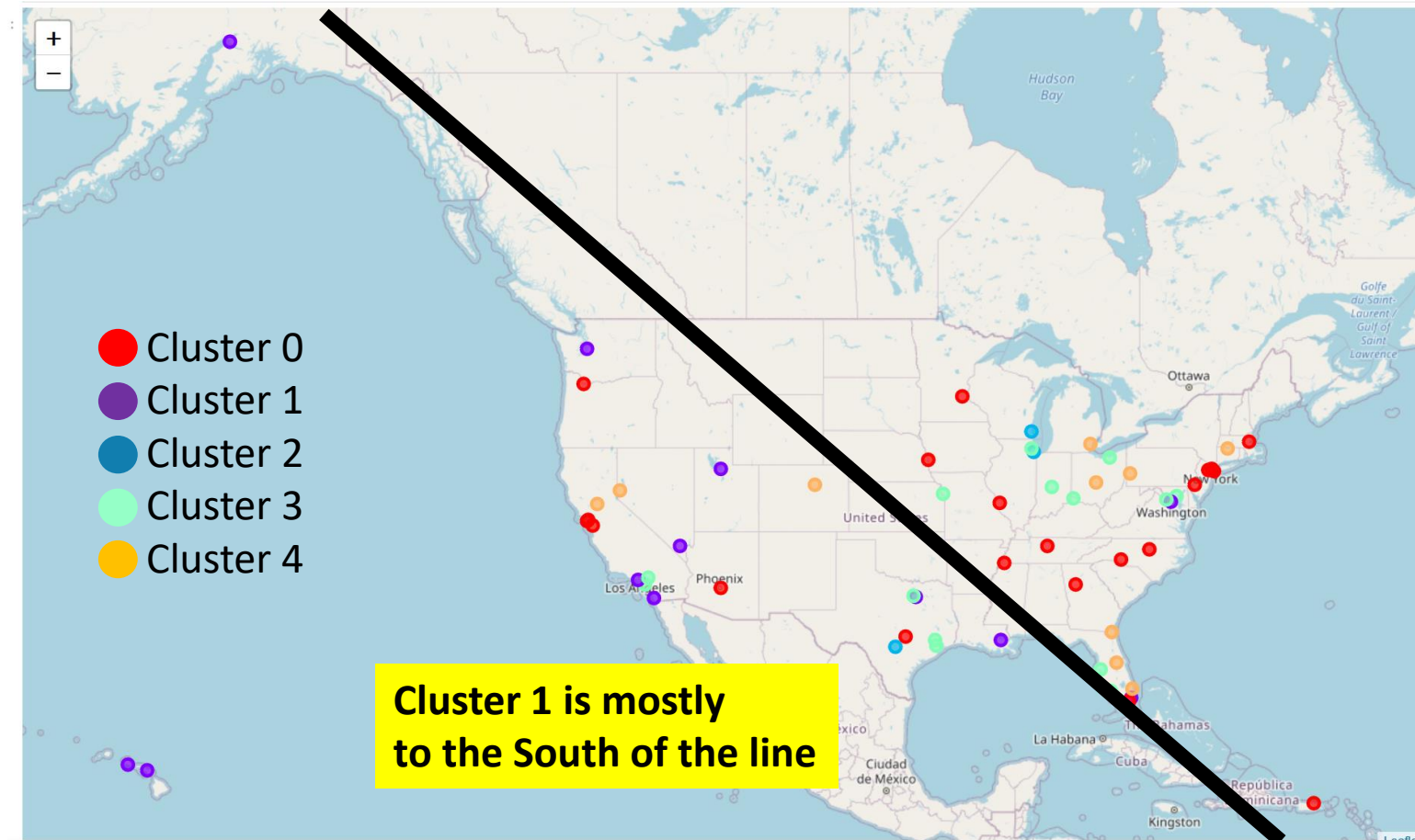
Cluster restaurants only, 5-means



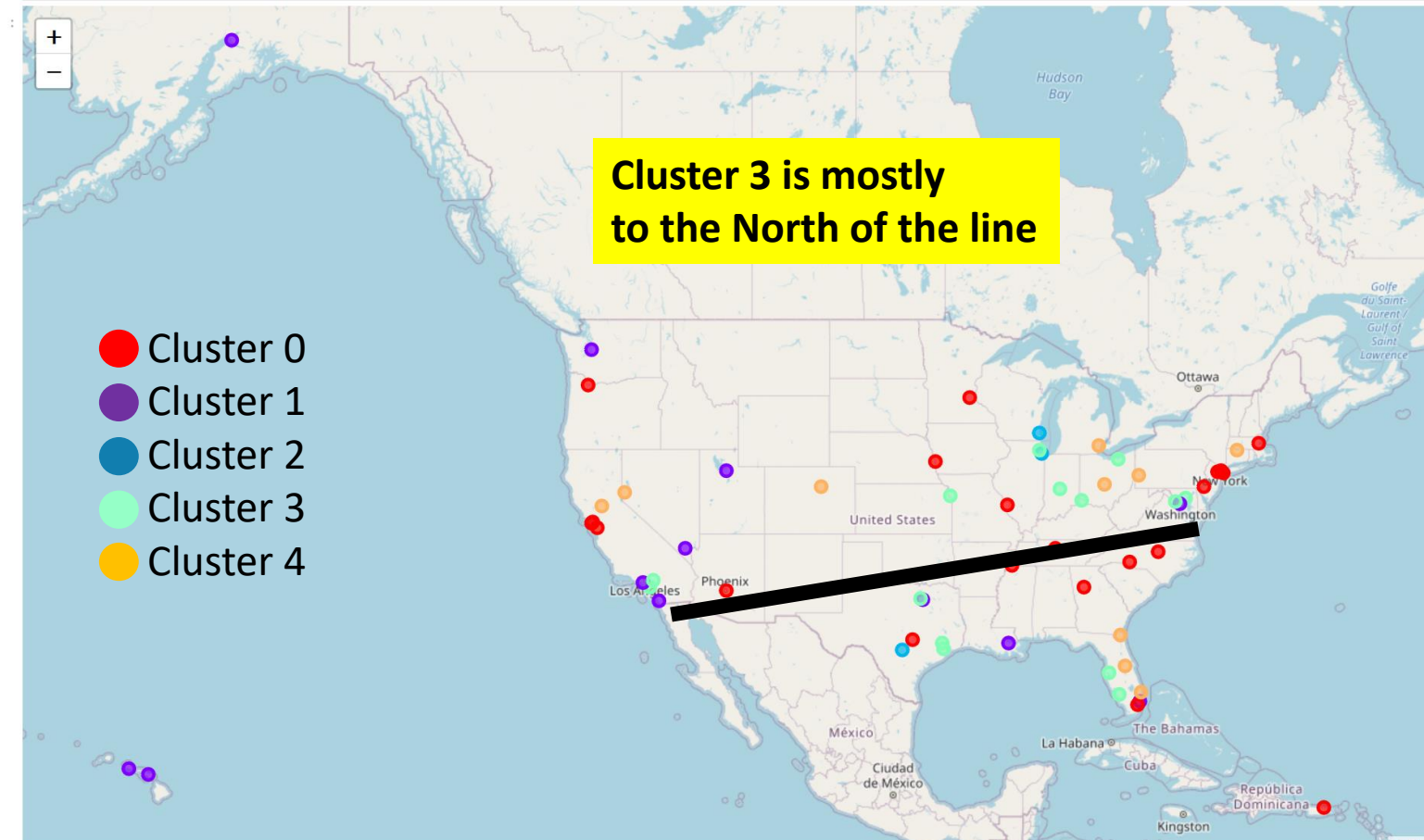
Cluster restaurants only, 5-means



Cluster restaurants only, 5-means

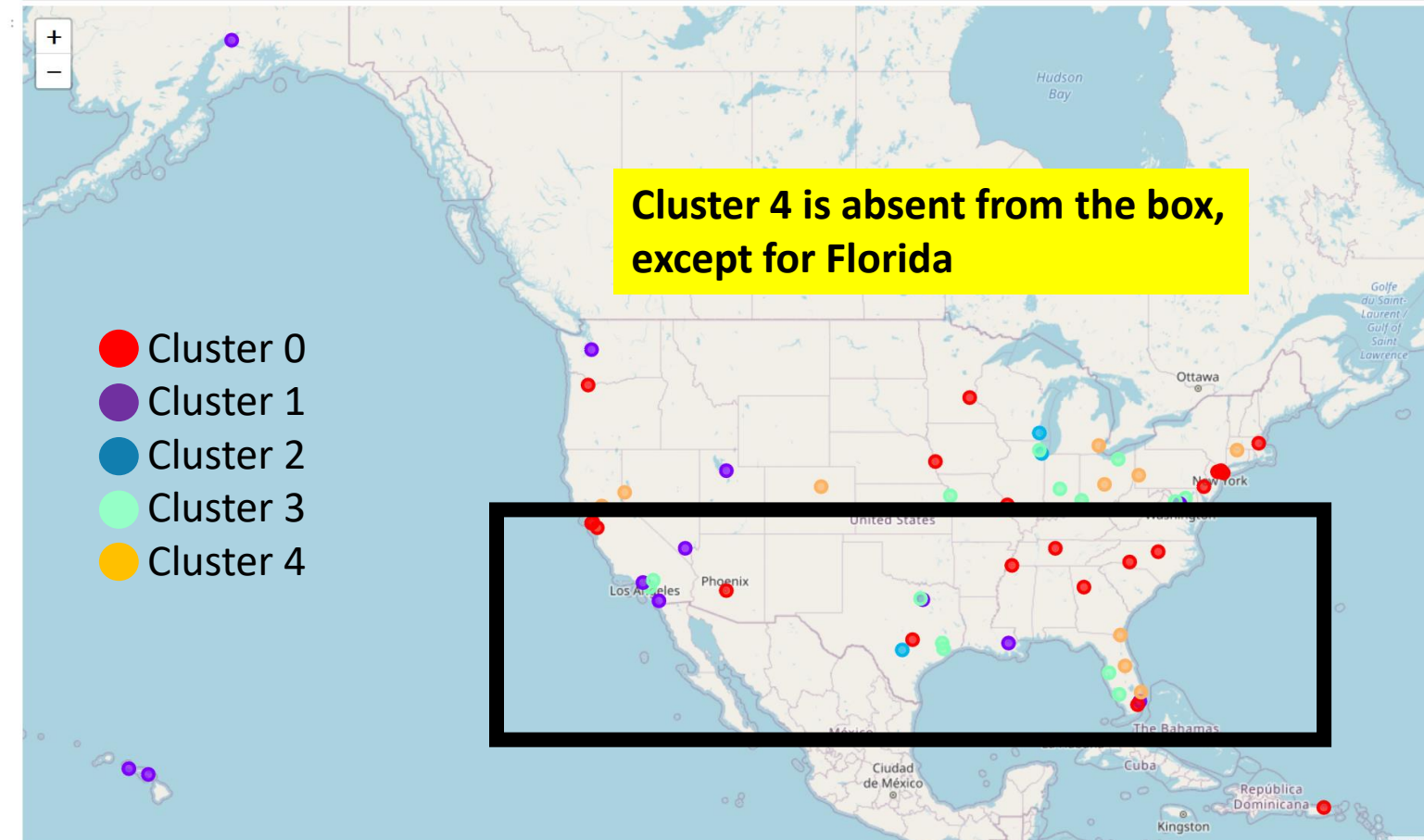


Cluster restaurants only, 5-means



Cluster restaurants only, 5-means

1



Conclusion

- We were able to cluster airports using their venue distribution, and their restaurant type distribution
- All venues except restaurants:
 - ✓ Distribution rather homogeneous geographically
 - ✓ Cluster 0 is lacking liquor stores but has zoos and art venues
 - ✓ Increasing cluster numbers include airports with more and more passenger traffic
- Business opportunity?
 - ✓ Airports that want to grow might need an increase of the proportion of hotels
 - ✓ Liquor stores might have room to be opened in cluster 0

Conclusion

- Restaurant clustering is geographically dependent
- Business opportunity: is it that the restaurants are already adapted to the local population, or some missing restaurant types need to be opened in order to fill the gap?
- Possible extension to this work: study the distribution of a one particular venue category, such as grocery stores and identify possible business opportunities at airports where this type of venue is missing