Understanding the trends in Pizza restaurants and the pizza they sell.

Ashish Jain (MT 18052)
Sarosh Hasan (MT 18084)
Shubham Gupta (MT 18055)

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Motivation

The Motivation behind the problem is that suppose you have a business with multiple outlets and one would like to analyze the trends in the sale or in which region what type of food people like the most. This dashboard can help in regulating staff among the outlets and maximize the profit as per fulfilling the demands of customers.

Data Source

- Dataset is collected from Datafiniti's Business Database dated between January 2018 and May 2019.
- This dataset is a list of 10,000 samples. Total no of features are 24.

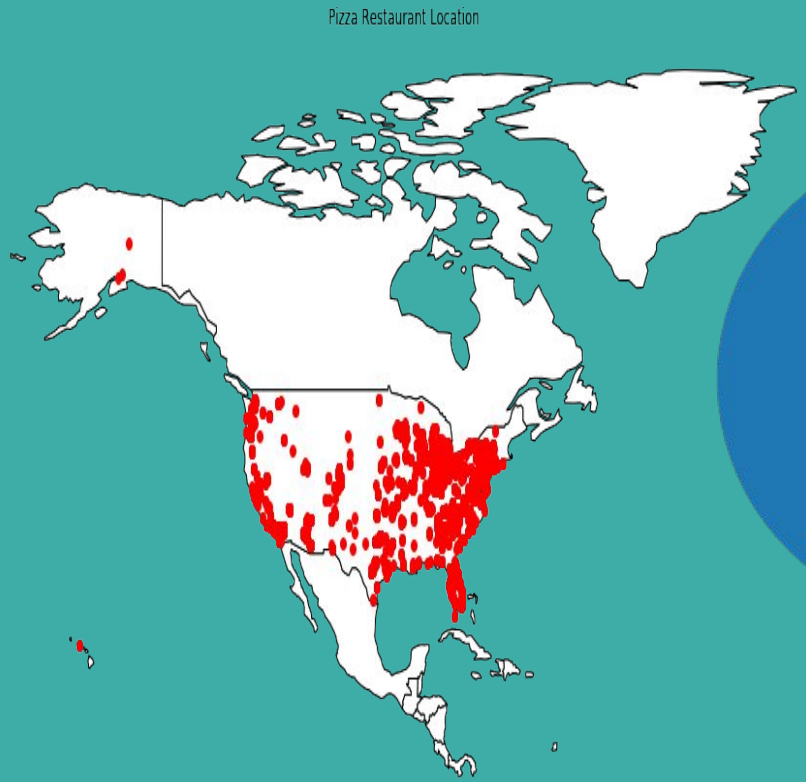
keys	latitude	...	menus.currency	menus.dateSeen	menus.description	menus.name	name	postalCode	priceRangeCurrency	priceRangeMin	p
91616	34.832300	...	USD	2018-05-01T04:25:37.197Z,2018-04-16T04:36:02.3...	NaN	Cheese Pizza	Shotgun Dans Pizza	72120	USD	0	
22936	33.509266	...	USD	2018-03-03T02:38:06.381Z,2018-01-18T20:18:10.0...	NaN	Pizza Cookie	Sauce Pizza Wine	85012	USD	0	
97122	39.144883	...	USD	2018-04-10T07:58:34.585Z,2018-04-21T05:43:21.4...	a saucelessampcomma double cheese pizza with a...	Pizza Blanca	Mios Pizzeria	45209	USD	0	
63116	42.516669	...	USD	2016-10-20T21:50:02Z,2016-03-29T05:08:59Z	NaN	Small Pizza	Hungry Howies Pizza	48071	USD	25	
65359	39.286630	...	USD	2016-03-31T02:34:04Z	NaN	Pizza Sub	Spartan Pizzeria	21224	USD	0	

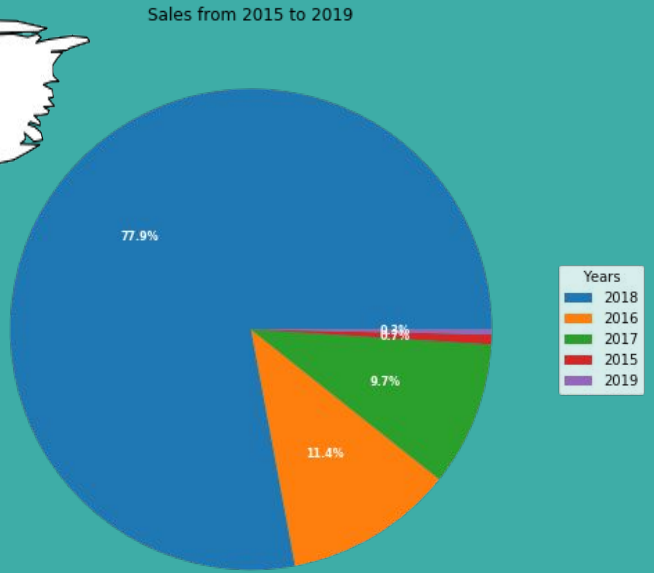
Data Analysis

- The dataset includes the category, name, address, city, state, menu information, price range, and more for each pizza restaurant in US.


```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 24 columns):
id                      10000 non-null object
dateAdded               10000 non-null object
dateUpdated             10000 non-null object
address                 10000 non-null object
categories              10000 non-null object
primaryCategories       10000 non-null object
city                    10000 non-null object
country                 10000 non-null object
keys                    10000 non-null object
latitude                10000 non-null float64
longitude               10000 non-null float64
menus.amountMax         10000 non-null float64
menus.amountMin         10000 non-null float64
menus.currency          10000 non-null object
menus.dateSeen          10000 non-null object
menus.description       3718 non-null object
menus.name              10000 non-null object
name                    10000 non-null object
postalCode              9996 non-null object
priceRangeCurrency      10000 non-null object
priceRangeMin           10000 non-null int64
priceRangeMax           10000 non-null int64
province                10000 non-null object
geometry                10000 non-null object
dtypes: float64(4), int64(2), object(18)
```

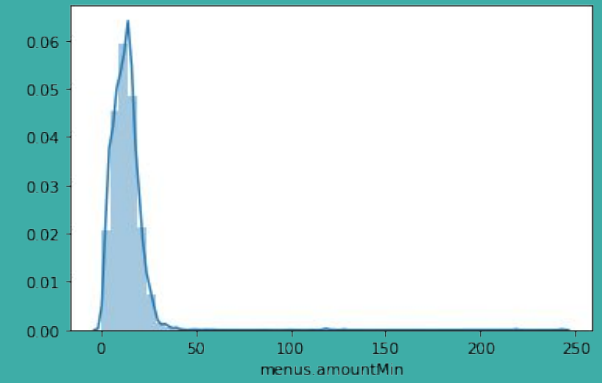
Data Analysis

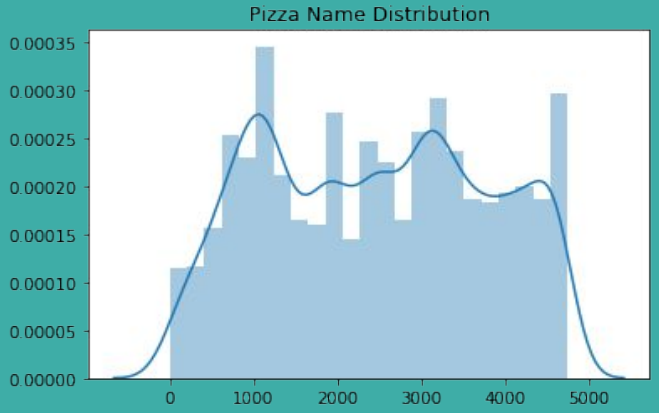


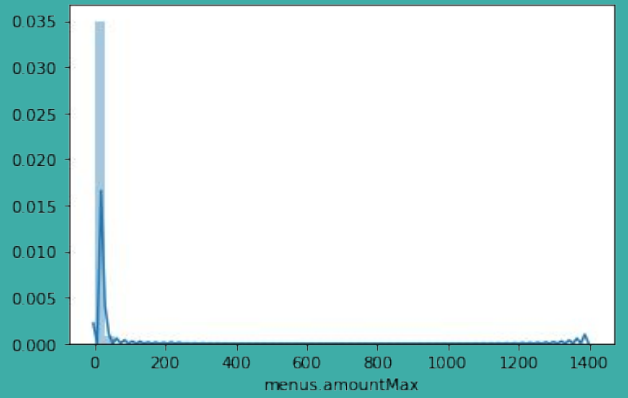


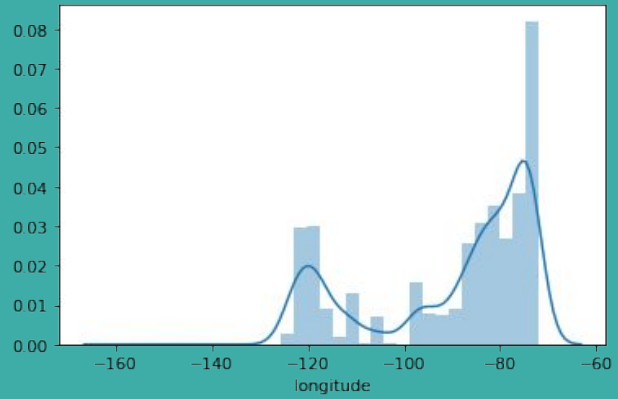
	latitude	longitude
0	34.832300	-92.183800
1	33.509266	-112.073044
2	39.144883	-84.432685
3	42.516669	-83.106630
4	39.286630	-76.566984
5	39.286630	-76.566984
6	37.875496	-122.260345
7	37.875496	-122.260345
8	30.537097	-84.215156
9	30.537097	-84.215156
10	42.912878	-85.566939
11	42.912878	-85.566939
12	42.912878	-85.566939
13	42.912878	-85.566939
14	34.137502	-117.865331

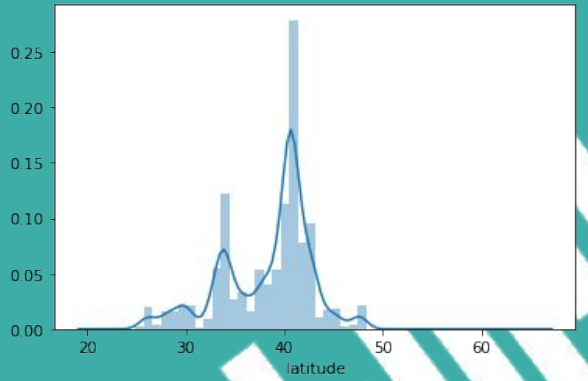
Data Distribution





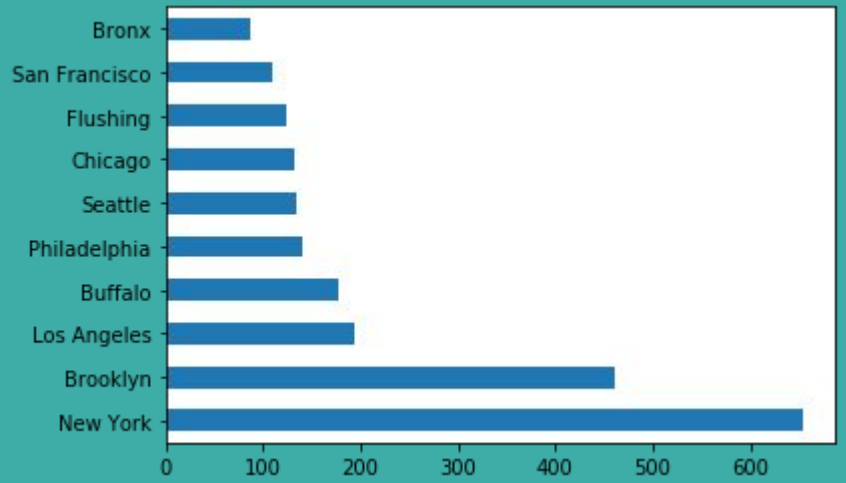


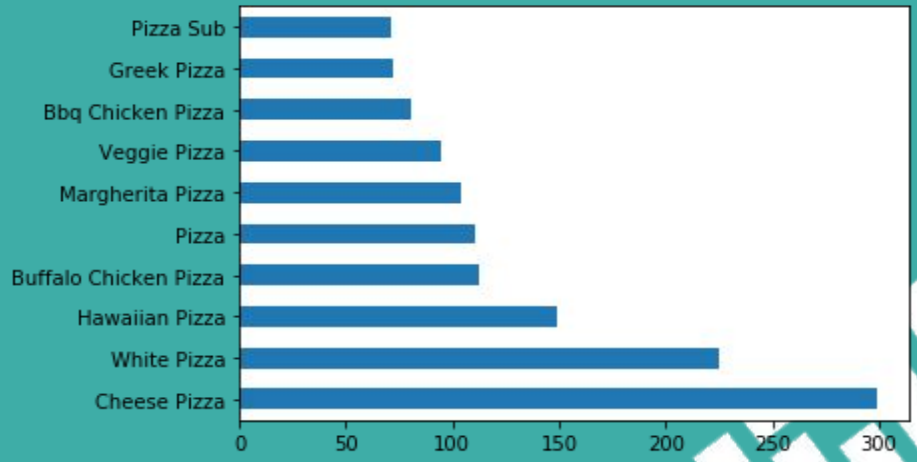




Inferences Drawn

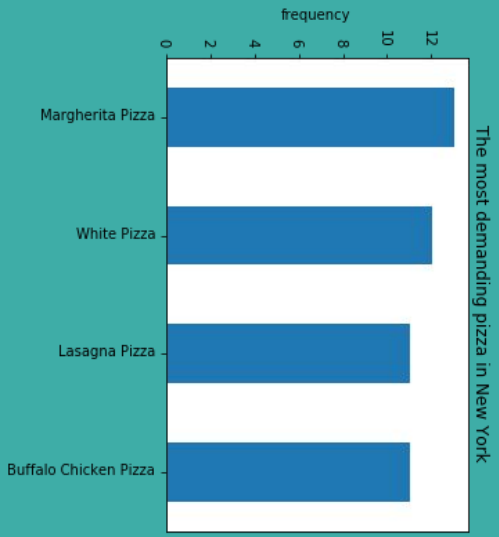
- Which are the top 10 pizzas in North America?
- Which are the top 10 cities with the most pizza restaurant ?

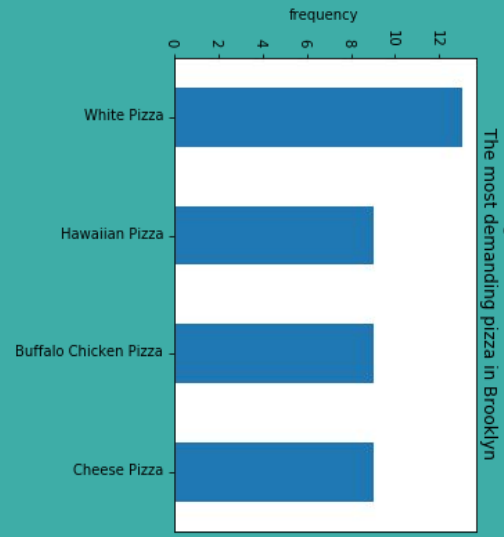


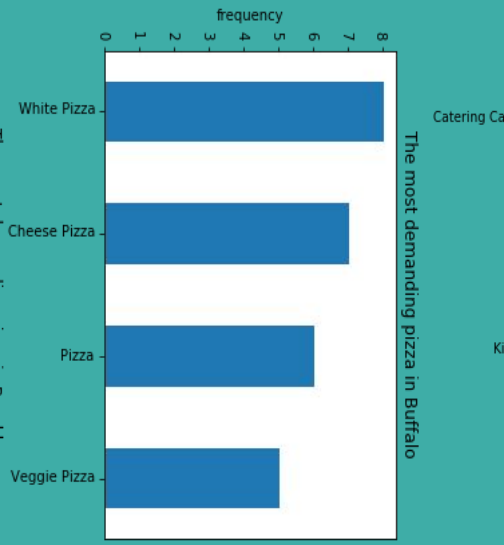


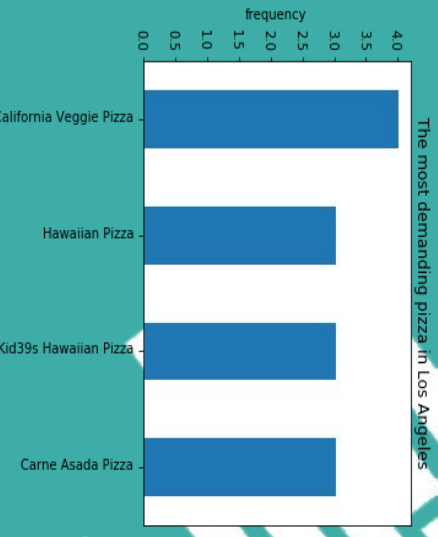
Inferences Drawn

- Which are the most demanding pizza in top 4 cities ?









Inferences Drawn

- Which are the cheapest and the most expensive pizza and its pizza restaurant?
  - We can infer that most expensive pizza i.e "Taco Pizza" is offered by "Rocco's" for 1395.0.
  - We can infer that cheapest pizza are offered by 3 restaurant, namely Fratellis Pizzeria,DiAngelos and Stacia's Gourmet Pizza and Pasta.

	name	menus.name	menus.amountMax
9337	Rocco's	Taco Pizza	1395.0

	name	menus.name	menus.amountMin
804	Fratellis Pizzeria	Pizza By the Slice	0.25
2777	DiAngelos	6" Pizza Sub	0.25
2778	DiAngelos	French Bread Pizza	0.25
7827	Stacia's Gourmet Pizza and Pasta	Garlic Herb Pizza Crust	0.25

Work To Do

- Clustering based on geography (Hotspot).
- Making a Dashboard which will show the plots of the data such as making hotspots of pizza restaurants.
- Tackle Problems of setup a new restaurant, so that the sell of restaurant maximizes.

References

- https://datafiniti.co/products/business-data/
- https://towardsdatascience.com/probability-distributions-in-data-science-cce6e64873a7
- Han, Jiawei, Micheline Kamber, and Anthony KH Tung. "Spatial clustering methods in data mining." Geographic data mining and knowledge discovery (2001): 188-217.
- Shu, Y., et al. "Clustering of hyperspectral image based on spatial-spectral Chinese restaurant process mixture model." Guang pu xue yu guang pu fen xi= Guang pu 36.4 (2016): 1158-1162