

# 1. Identifying Fast Food Restaurants in NYC

Data set: <https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8?fbclid=IwAR3CjTb7mf5K1oNklcONeYh6fdprYXwrMS4a6S71r8y9BD4NvBoQzPKY24>

```
In [135]: import pandas as pd
import os
import re

In [136]: os.getcwd()

Out[136]: 'C:\Users\l\l\l\Desktop\Spec_Graphics\ld7_supermarkets_workspace'
```

## Read in and view data

Data dictionary available at <https://data.cityofnewyork.us/Health/DOHMH-New-York-City-Restaurant-Inspection-Results/43nn-pn8?fbclid=IwAR3CjTb7mf5K1oNklcONeYh6fdprYXwrMS4a6S71r8y9BD4NvBoQzPKY24>

Explanation of columns: CAMIS: This is a unique identifier for the entity (restaurant): 10-digit integer, static per restaurant permit DBA: "doing business as" (storefront name) BORO: borough: 1 = MANHATTAN • 2 = BRONX • 3 = BROOKLYN • 4 = QUEENS • 5 = STATEN ISLAND BUILDING: building number STREET: street name ZIPCODE: zipcode PHONE: Phone number provided by restaurant owner/manager CUISINE DESCRIPTION: Optional field provided by provided by restaurant owner/manager INSPECTION DATE: NOTE: Inspection dates of 1/1/1900 mean an establishment has not yet had an inspection **COUNCIL DISTRICT: Important, we will use this to locate only those in D7**

**Drive link to DOHMH New York City Restaurant Inspection Results.csv:** <https://drive.google.com/file/d/1kHRSendYmawV-XcccdHk4mfKw68dF5k/view?usp=sharing>

```
In [137]: restaurants_df = pd.read_csv("../DOHMH_New_York_City_Restaurant_Inspection_Results.csv")
restaurants_df
restaurants_df.shape

Out[137]: (381502, 26)
```

## Select only District 7 Restaurants

Leaves us with 6220 of the original 380k NYC restaurants

```
In [138]: d7_restaurants_df = restaurants_df.loc[restaurants_df["Council District"] == 7]
#d7_restaurants_df
#d7_restaurants_df.shape
```

## What kind of DBA names do we have? Cuisine Descriptions?

(will help us filter out fast food)

```
In [139]: d7_restaurants_df["DBA"]

Out[139]: 81      JOHN'S FRIED CHICKEN
92      BIRCH COFFEE
175     AWASH ETHIOPIAN RESTAURANT
224     THE FAMOUS JIMBO'S HAMBURGER PALACE
272      AWADH
...
381295   HUNAN CHEN'S KITCHEN
381409   DEMITASSE COFFEE & TEA
381435   GEISHA ASIAN CUISINE & LOUNGE
381436   GEISHA ASIAN CUISINE & LOUNGE
381442   STROKOS GOURMET DELI
Name: DBA, Length: 6220, dtype: object

In [140]: d7_restaurants_df["CUISINE DESCRIPTION"]

Out[140]: 81      Chicken
92      Coffee/Tea
175     Ethiopian
224     Hamburgers
272     Indian
...
381295   Chinese
381409   Coffee/Tea
381435   Asian/Asian Fusion
381436   Asian/Asian Fusion
381442   Sandwiches
Name: CUISINE DESCRIPTION, Length: 6220, dtype: object

Top 50 US fast-food restaurant chains by 2020 sales: https://www.qsrmagazine.com/content/qsrf50-2021-top-50-chart

Cross-referenced with https://www.restaurantbusinessonline.com/top-50-chains?year=2021&page=1#data-table
```

```
In [141]: top_50_restaurant_chains = ["McDonald's", "Starbucks", "Chick-fil-A", "Taco Bell", "Wendy's", "Burger King", "Subway", "Domino's", "Chipotle Mexican Grill", "Sonic Drive-In", "Panera Bread", "Pizza Hut", "HFC", "Popeyes Louisiana Kitchen", "Arby's", "Little Caesars", "Dairy Queen", "Panda Express", "Jack in the Box", "Papa John's", "Whataburger", "Wingstop", "Zaxby's", "Bardde's", "Culver's", "Jimmy John's", "Raising Cane's Chicken Fingers", "Five Guys", "Jersey Mike's Subs", "Carl's Jr.", "Bojangles", "In-N-Out Burger", "El Pollo Loco", "Del Taco", "Checkers Drive-In Restaurants", "Firehouse Subs", "Krispy Kreme", "Papa Murphy's Pizza", "Steak 'n Shake", "Qdoba", "Marco's Pizza", "Church's Chicken", "Elm Hortons", "Tropical Smoothie Cafe", "Freddy's Frozen Custard and Steakburgers", "McAlister's Deli", "Baskin Robbins", "White Castle", "Moe's"]

FAST_FOOD_KEYWORDS = list()

for chain in top_50_restaurant_chains:
    FAST_FOOD_KEYWORDS.append(chain.lower())

#FAST_FOOD_KEYWORDS
len(FAST_FOOD_KEYWORDS)

Out[141]: 50
```

```
In [142]: def isFastFood(business_name):
    if type(business_name) != str:
        return False # in case of empty entries
    business_name = business_name.lower()
    businesses_list = business_name.split(",") # some locations have multiple names listed
    for keyword in FAST_FOOD_KEYWORDS:
        for business in businesses_list:
            if business in keyword:
                return True
    return False
```

```
In [143]: ## Example:
## isFastFood("MCDONALD'S")
```

## Determine if each D7 restaurant is a fast-food place

We'll mostly focus on filtering out big national fast-food chains

```
In [144]: is_fast_food_list = list()
dba_list = d7_restaurants_df["DBA"].tolist()
#dba_list
#type(dba_list[0])

In [145]: for dba in dba_list:
    is_fast_food_list.append(isFastFood(dba))

In [146]: #is_fast_food_list

In [147]: d7_restaurants_fast_df = d7_restaurants_df.assign(fast_food = is_fast_food_list)
#d7_restaurants_fast_df

In [148]: ## Keep only fast food restaurants and remove unnecessary columns
d7_restaurants_fast_df = d7_restaurants_fast_df.loc[d7_restaurants_fast_df["fast_food"] == True]
d7_restaurants_fast_df = d7_restaurants_fast_df.set_index("CAMIS")

d7_restaurants_fast_df = d7_restaurants_fast_df.drop_duplicates()
d7_restaurants_fast_df = d7_restaurants_fast_df[~d7_restaurants_fast_df.index.duplicated(keep='first')]
d7_restaurants_fast_df = d7_restaurants_fast_df.drop(columns=["BORO", "PHONE", "CUISINE DESCRIPTION", "INSPECTION DATE", "VIOLATION CODE", "VIOLATION DESCRIPTION", "CRITICAL FLAG", "SCORE", "GRADE", "GRADE DATE", "INSPECTION TYPE", "RECORD DATE", "Community Board", "Census Tract", "BIN", "BBL", "NTA", "ACTION", "Council District", "fast_food"])

In [149]: d7_restaurants_fast_df

Out[149]:
```

<

## Save output to csv

```
In [150]: d7_restaurants_fast_df.to_csv("fast_food_in_d7.csv")
```

## 2. Identifying and Categorizing Convenience Stores v. Grocery Stores in NYC

Get dataframe of all food retailers in D7 (narrowed down from <https://data.ny.gov/Economic-Development/Retail-Food-Stores-Map/p2dn-xhaw> using R to superimpose D7 geographical boundaries)

Raw retail food stores data (all NY): <https://drive.google.com/file/d/1HQaKZQHdRmCigJ56Pi0h-PrpFTxHytCTC/view?usp=sharing>

**food\_retailers\_in\_d7.csv:** [https://drive.google.com/file/d/1S\\_j6oIdUWcljeh3cvX7U7P6ETaAdk/view?usp=sharing](https://drive.google.com/file/d/1S_j6oIdUWcljeh3cvX7U7P6ETaAdk/view?usp=sharing)

```
In [151]: food_retailers_d7_df = pd.read_csv("../food_retailers_in_d7.csv")
# food_retailers_d7_df
# food_retailers_d7_df.shape
```

Get dataframe of all grocery/fresh produce stores in D7 (manually selected by examining business names)

**grocery\_stores\_in\_d7.csv:** <https://drive.google.com/file/d/1hhFlG9nsJnbxFPqg70xCFc408/view?usp=sharing>

```
In [152]: grocery_stores_d7_df = pd.read_csv("../grocery_stores_in_d7.csv")
grocery_stores_d7_df.rename(columns={})
grocery_stores_d7_df

Out[152]:
```

	County	License.Number	Operation.Type	Establishment.Type	Entity.Name	DBA.Name	Street.Number	Street.Name	City	State	...
0	New York	622503	Store	JAC	3550 FOOD CORP	KEY FOOD SUPERMARKET	3550	BROADWAY	NEW YORK	NY	...
1	New York	739952	Store	JAC	MANHATTEN AVENUE COMMUNITY CAFE I	MANHATTEN AVE COMMUNITY	51	WEST 106TH STREET	NEW YORK	NY	...
2	New York	725281	Store	A	ABC MART NC	ABC MART	3505	BROADWAY	NEW YORK	NY	...
3	New York	629213	Store	JAC	DDY MINI MARKET & DELI CORP	DDY MINI MARKET & DELI	955	AMSTERDAM AVE	NEW YORK	NY	...
4	New York	742768	Store	JAC	HB DELI GROCERY CORP	HB DELI GROCERY	3787A	BROADWAY	NEW YORK	NY	...
...	...	...	...	...	...	...	...	...	...	...	...

106	New York	731926	Store	JAC	DYLAN'S DELICIOUS DELI GROCERY COR	DYLAN'S DELICIOUS DELI GROCERY	900	RIVERSIDE DRIVE	NEW YORK	NY	...
107	New York	624394	Store	JAC	MILANO MARKET INC	MILANO MARKET	2892	BROADWAY	NEW YORK	NY	...
108	New York	627088	Store	JAC	91 CATHERINE GROCERY CORP	91 CATHERINE GROCERY	91	HAMILTON PL	NEW YORK	NY	...
109	New York	621015	Store	JAC	3011 GROCERY INC	C TOWN	3320	BROADWAY	NEW YORK	NY	...
110	New York	708796	Store	JAC	KING MEAT CORP	SHOP FAIR	3871	BROADWAY	NEW YORK	NY	...

111 rows x 22 columns

Set "grocery" column in food\_retailers\_d7\_df to "true"/"false" based on whether it appears in our manual selection of grocery stores listed in grocery\_stores\_d7\_df.

```
In [153]: # Get list of all grocery stores
grocery_stores_list = grocery_stores_d7_df["License.Number"].to_list()
#grocery_stores_list

In [154]: # Create column indicating if each store is a grocery store; all "A" establishments are NON-grocery stores
food_retailers_list = food_retailers_d7_df["License.Number"].to_list()
establishment_type_list = food_retailers_d7_df["Establishment.Type"].to_list()
is_grocery = list()

for i in range(len(food_retailers_list)):
    if (food_retailers_list[i] in grocery_stores_list) and (establishment_type_list[i] != "A"):
        is_grocery.append(True)
    else:
        is_grocery.append(False)
#is_grocery
```

```
In [155]: # Add new column to dataframe
food_retailers_d7_df["is_grocery"] = is_grocery

In [156]: food_retailers_d7_df
```

	County	License.Number	Operation.Type	Establishment.Type	Entity.Name	DBA.Name	Street.Number	Street.Name	Address.Line.2	A
0	New York	627413	Store	A	DOLLAR VARIETY STORE INC	DOLLAR VARIETY STORE	762	AMSTERDAM AVE		NaN
1	New York	717156	Store	JAC	GALAXY GOURMET INC	GALAXY GOURMET	3778	BROADWAY		NaN
2	New York	622503	Store	JAC	3550 FOOD CORP	KEY FOOD SUPERMARKET	3550	BROADWAY		NaN
3	New York	739966	Store	JAC	FINEST COLUMBIA DELI INC	FINEST COLUMBIA DELI	945	AMSTERDAM AVE		NaN
4	New York	628824	Store	JAC	SANTIAGO DELI FOOD CORP	SANTIAGO DELI FOOD	3415	BROADWAY		NaN
...	...	...	...	...	...	...	...	...	...	...

227	New York	741584	Store	JAC	NEW DELI 1 CORP	NEW DELI 1	3660	BROADWAY		NaN
228	New York	748699	Store	A	UPPER MANHATTAN PHARMA INC	UPPER MANHATTAN PHARMACY	1728	AMSTERDAM AVE		NaN
229	New York	626558	Store	A	DUANE READE ETAL PIRS	DUANE READE 14225	2864	BROADWAY		NaN
...	...	...	...	...	...	...	...	...	...	...

230	New York	721538	Store	JAC	JOES GOURMET DELI INC	JOES GOURMET DELI	920	COLUMBUS AVENUE		NaN
231	New York	627132	Store	A	1967 99 CENTS & DISCOUNT INC	99 CENTS & DISCOUNT	1967	AMSTERDAM		NaN

232 rows x 18 columns

**Note:** All non-grocery stores ("grocery" = False) are pharmacies, tobacco/liquor stores, delis, or other convenience stores that do not sell a wide enough range of fresh produce to be considered grocery stores.

```
In [157]: # See the grocery stores
food_retailers_d7_df.loc[food_retailers_d7_df["is_grocery"] == True]

Out[157]:
```

	County	License.Number	Operation.Type	Establishment.Type	Entity.Name	DBA.Name	Street.Number	Street.Name	Address.Line.2	A
2	New York	622503	Store	JAC	3550 FOOD CORP	KEY FOOD SUPERMARKET	3550	BROADWAY		NaN
5	New York	739952	Store	JAC	MANHATTEN AVENUE COMMUNITY CAFE I	MANHATTEN AVE COMMUNITY	51	WEST 106TH STREET		NaN
7	New York	629213	Store	JAC	DDY MINI MARKET & DELI CORP	DDY MINI MARKET & DELI	955	AMSTERDAM AVE		NaN
10	New York	742768	Store	JAC	HB DELI GROCERY CORP	HB DELI GROCERY	3787A	BROADWAY		NaN
12	New York	740014	Store	JAC	NY HARVEST TRADING INC	EVERFRESH SUSHI	2840	BROADWAY		NaN
...	...	...	...	...	...	...	...	...	...	...

215	New York	731926	Store	JAC	DYLAN'S DELICIOUS DELI GROCERY COR	DYLAN'S DELICIOUS DELI G	900	RIVERSIDE DRIVE		NaN
216	New York	624394	Store	JAC	MILANO MARKET INC	MILANO MARKET	2892	BROADWAY		NaN
...	...	...	...	...	...	...	...	...	...	...
227	New York	741584	Store	JAC	NEW DELI 1 CORP	NEW DELI 1	3660	BROADWAY		NaN
228	New York	748699	Store	A	UPPER MANHATTAN PHARMA INC	UPPER MANHATTAN PHARMACY	1728	AMSTERDAM AVE		NaN
229	New York	626558	Store	A	DUANE READE ETAL PIRS	DUANE READE 14225	2864	BROADWAY		NaN
230	New York	721538	Store	JAC	JOES GOURMET DELI INC	JOES GOURMET DELI	920	COLUMBUS AVENUE		NaN
231	New York	627132	Store	A	1967 99 CENTS & DISCOUNT INC	99 CENTS & DISCOUNT	1967	AMSTERDAM		NaN

141 rows x 18 columns

```
In [158]: # Save convenience stores
food_retailers_d7_df.loc[food_retailers_d7_df["is_grocery"] == False]

Out[158]:
```

	County	License.Number	Operation.Type	Establishment.Type	Entity.Name	DBA.Name	Street.Number	Street.Name	Address.Line.2	Ad
0	New York	627413	Store	A	DOLLAR VARIETY STORE INC	DOLLAR VARIETY STORE	762	AMSTERDAM AVE		NaN
1	New York	717156	Store	JAC	GALAXY GOURMET INC	GALAXY GOURMET	3778	BROADWAY		NaN
3	New York	739966	Store	JAC	FINEST COLUMBIA DELI INC	FINEST COLUMBIA DELI	945	AMSTERDAM AVE		NaN
4	New York	628824	Store	JAC	SANTIAGO DELI FOOD CORP	SANTIAGO DELI FOOD	3415	BROADWAY		NaN
6	New York	725281	Store	A	ABC MART NC	ABC MART	3505	BROADWAY		NaN
...	...	...	...	...	...	...	...	...	...	...
227	New York	741584	Store	JAC	NEW DELI 1 CORP	NEW DELI 1	3660	BROADWAY		NaN
228	New York	748699	Store	A	UPPER MANHATTAN PHARMA INC	UPPER MANHATTAN PHARMACY	1728	AMSTERDAM AVE		NaN
229	New York	626558	Store	A	DUANE READE ETAL PIRS	DUANE READE 14225	2864	BROADWAY		NaN
230	New York	721538	Store	JAC	JOES GOURMET DELI INC	JOES GOURMET DELI	920	COLUMBUS AVENUE		NaN
231	New York	627132	Store	A	1967 99 CENTS & DISCOUNT INC	99 CENTS & DISCOUNT	1967	AMSTERDAM		NaN

262 rows x 5 columns

The remaining formatting will require some manual spot-checking. Download so we can do just that:

```
In [167]: # Save csv
combined_df.to_csv("combined_food_before_manual_cleaning.csv")
```

## 3. Formatting for Mapbox

Combined fast food, convenience store, and grocery store data:

<https://drive.google.com/file/d/1NeefnY3ibDGXV8DlIXqkn77IptmZB6/view?usp=sharing>

```
In [160]: combined_df = pd.read_csv("../combined_food.csv")
combined_df

Out[160]:
```

	DBA	Address	latitude	longitude	retailer_type
0	DOLLAR VARIETY STORE	762 AMSTERDAM AVE	40.794936	-73.969728	convenience
1					