

IBM Ciencia de Datos

Proyecto Capstone

Introducción: Problema comercial

La ciudad de Oakland está situada en el área de la bahía de San Francisco del estado de California, Estados Unidos, cuenta con una población de 440,646 habitantes con fecha del 2020, siendo la octava ciudad más poblada del estado, a nivel etnico cuenta con una gran diversidad entre asiáticos, afroamericanos, nativos del pacifico, árabes, indígenas, indios, hispanos y mestizos.

Dado eso, la ciudad proporciona una amplia diversidad de mundos gastronómicos brindando una alta oportunidad de inversión a pequeños y grandes empresarios de la industria panificadora, logrando establecer sus operaciones de forma estratégica y acertada, alcanzando poder iniciar el establecimiento de Panaderías subyacentes entre los distintos vecindarios

Con este proyecto analizaremos los distintos vecindarios de Oakland que presentan comercios de Panaderías como competidores directos e indirectos a la industria, para poder recomendar sobre qué categoría de vecindario será el más óptimo para iniciar operaciones.

El resultado del proyecto está dirigido a las empresas enfocadas a la venta de pan, pero puede expandirse al sector de repostería y alimentos relacionados, dado que las categorías de comercios considerados involucra a ambas partes como un factor decisivo para la ubicación final de la inversión.

Datos

Para lograr analizar los vecindarios de la ciudad de Oakland, California, obtenemos la información de la página de Local Wiki de vecindarios de Oakland, en la siguiente URL:

<https://localwiki.org/oakland/Neighborhoods>

Posteriormente, se utiliza la librería de Geopy para obtener las coordenadas geográficas del listado de vecindarios, almacenando dicha información para un mejor rendimiento en el archivo **df_vecindarios_oakland.csv** en el siguiente repositorio de Github.

https://github.com/i3nriq/Ciencia-de-Datos-Aplicada---Curso-Capstone/blob/master/Data_oakland/df_vecindarios_oakland.csv

La obtención de lugares utilizando la API de Foursquare es almacenado de igual manera resguardado en el archivo **oakland_venues.csv** dentro del mismo repositorio de Github.

https://github.com/i3nriq/Ciencia-de-Datos-Aplicada---Curso-Capstone/blob/master/Data_oakland/oakland_venues.csv

Cabe mencionar que ambos archivos **son de autoría propia**, este último archivo contiene un total de 2462 lugares, analizados con un radio de 500 metros a la redonda desde el punto geográfico de los vecindarios, los cuales filtraremos únicamente 16 categorías de comercios relacionados a la industria Panificadora y sus competidores directos e indirectos.

- Café
- Bakery
- Sandwich Place
- Burger Joint
- Convenience Store
- Donut Shop
- Market
- Latin American Restaurant
- Bagel Shop
- Supermarket
- Cupcake Shop
- Pet Café
- Street Food Gathering
- Food Truck
- American Restaurant
- Grocery Store

Obteniendo una matriz de ubicaciones geográficas por categoría entre vecindarios, que posteriormente serán analizados por medio del modelo K-means para poder determinar los patrones entre vecindarios.

Resumen

- Listado de vecindarios de Oakland, California.
- Coordenadas de los vecindarios.
- Lugares de moda en un radio de 500 metros.
- Categorías de lugares relacionados a la industria Panificadora.

Metodología

Con la finalidad de obtener un vecindario óptimo para la inversión, se requiere utilizar un proceso de codificación **one-hot**, creando de forma dinámica las variables por categorías de forma horizontal.

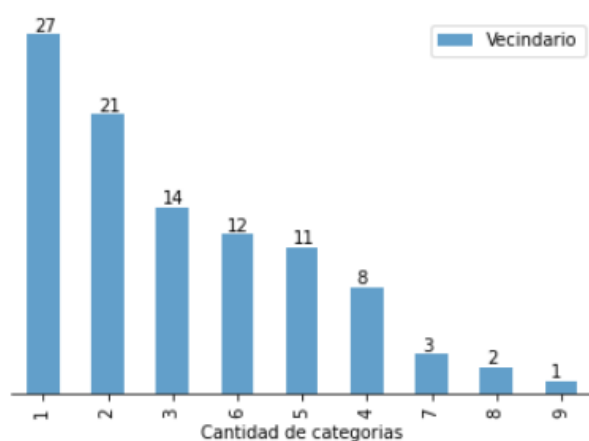
Vecindario	American Restaurant	Bagel Shop	Bakery	Burger Joint	Café	Convenience Store	Cupcake Shop	Donut Shop	Food Truck	Grocery Store	Latin American Restaurant	Market	Pet Café	Sandwich Place	Street Food Gathering
Adams Point	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Adams Point	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Adams Point	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Adams Point	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Adams Point	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0

Posteriormente se representa en porcentajes la frecuencia de las categorías dentro de un vecindario..

Vecindario	American Restaurant	Bagel Shop	Bakery	Burger Joint	Café	Convenience Store	Cupcake Shop	Donut Shop	Food Truck	Grocery Store	Latin American Restaurant	Market	Pet Café
Adams Point	0.0	0.0	0.000000	0.2	0.000000	0.20	0.0	0.000000	0.200000	0.400000	0.0	0.0	0.0
Chinatown	0.0	0.0	0.500000	0.0	0.250000	0.00	0.0	0.000000	0.000000	0.250000	0.0	0.0	0.0
City Center (Neighborhood)	0.0	0.0	0.500000	0.0	0.250000	0.00	0.0	0.000000	0.000000	0.250000	0.0	0.0	0.0
Civic Center	0.0	0.0	0.117647	0.0	0.411765	0.00	0.0	0.058824	0.235294	0.117647	0.0	0.0	0.0
Cleveland Heights	0.0	0.0	0.000000	0.0	0.250000	0.25	0.0	0.000000	0.250000	0.250000	0.0	0.0	0.0

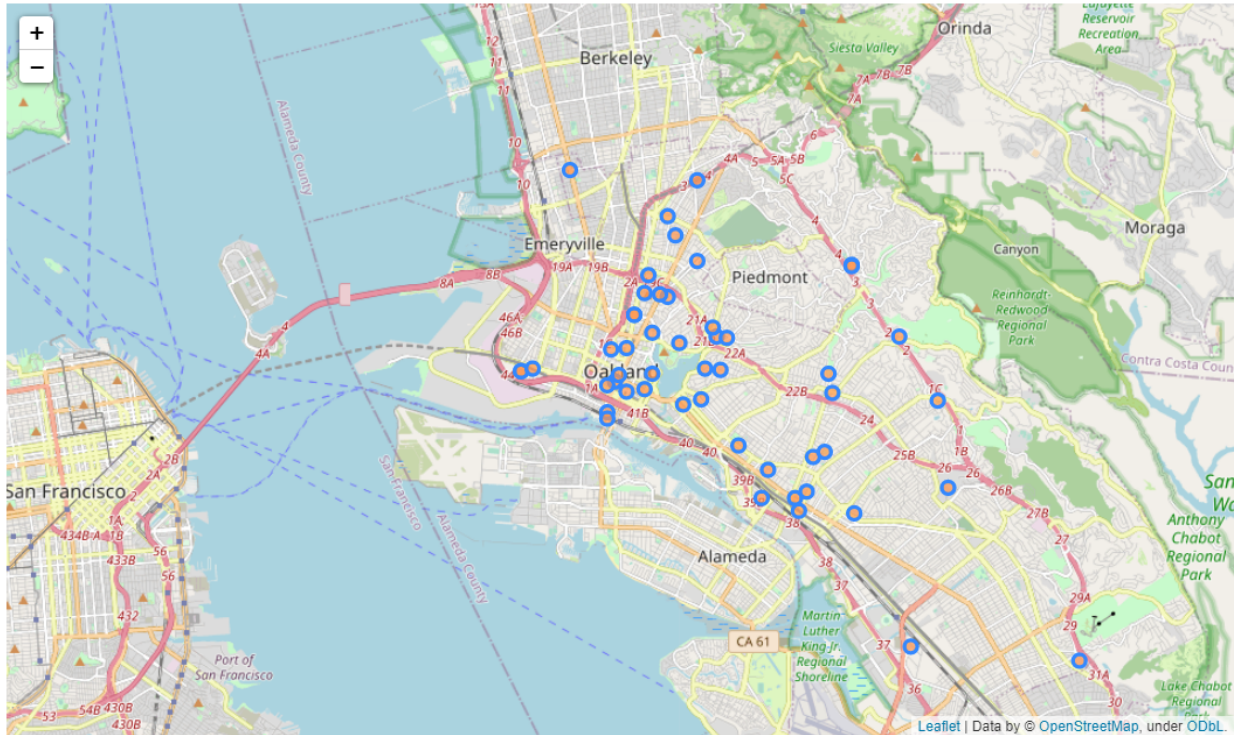
Como parte del análisis exploratorio se realiza una tabla de frecuencias de categorías por vecindarios, determinando descartando 48% de vecindarios, representando el 18% de lugares del conjuntos de datos, logramos establecer un ranqueo de 3 categorías por vecindario, dando un criterio más balanceado para el modelo k-means.

Frecuencia de categorias distintas por vecindarios



Dando como resultado la siguiente depuración de vecindarios de Oakland..

99 Vecindarios de Oakland	representando 410 lugares.
48 Vecindarios descartado	representando 73 lugares.
51 Vecindarios de Oakland a analizar	representando 337 lugares.



Mapa de vecindarios a analizar de la ciudad de Oakland, California.

Modelo K-means

En base a las frecuencias de las categorías más recurrentes en cada vecindarios, analizamos de forma descendente , logrando determinar nuestras variables que alimentara en modelo cluster, estableciendo 5 distintos clusters para ajustar el modelo, retornando las etiquetas de cluster sobre los vecindarios analizados, que nos ayudará a poder observar la distribución de las características, siendo clave en la toma de decisión sobre qué grupo de vecindarios se recomienda iniciar operaciones dentro de la ciudad de Oakland.

Vecindario	Latitud	Longitud	Etiquetas de cluster	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Adams Point	37.810403	-122.255903	3	Grocery Store	Burger Joint	Convenience Store
Chinatown	37.79993	-122.270239	0	Bakery	Café	Grocery Store
City Center (Neighborhood)	37.802165	-122.273128	0	Bakery	Café	Grocery Store
Civic Center	37.800584	-122.264735	4	Café	Food Truck	Bakery
Cleveland Heights	37.804546	-122.245325	1	Café	Convenience Store	Food Truck

IBM - Ciencia de Datos 2022

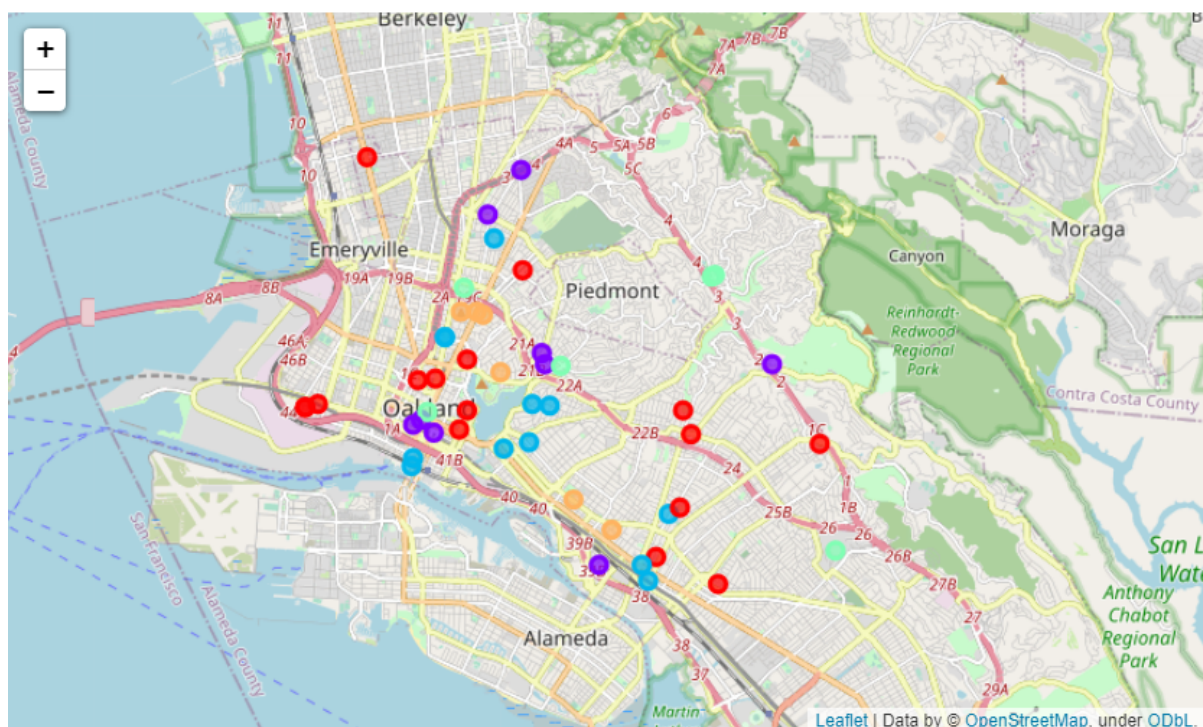
Proyecto Capstone - La Batalla de los Vecindarios

Resultado

Del modelo K-means , obtenemos la clasificación de los 51 vecindarios de Oakland, con una composición de clusters,teniendo como variables principales la ponderaciones descendente de las 3 categorías más relevantes por vecindario.

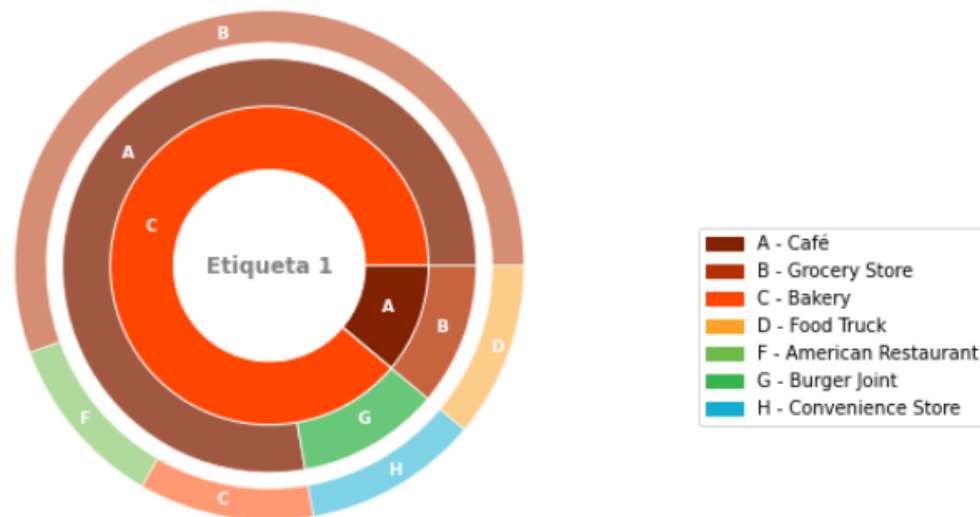


La representación de los cluster se visualiza a través de gráficas de pies de 3 niveles, brindando una mayor comprensión a la composición del vecindario. Dando mayor relevancia al anillo más cercano al centro como la variable más significativa dentro de la frecuencia de categorías más relevantes para la toma de decisiones.



Detalle de clusters

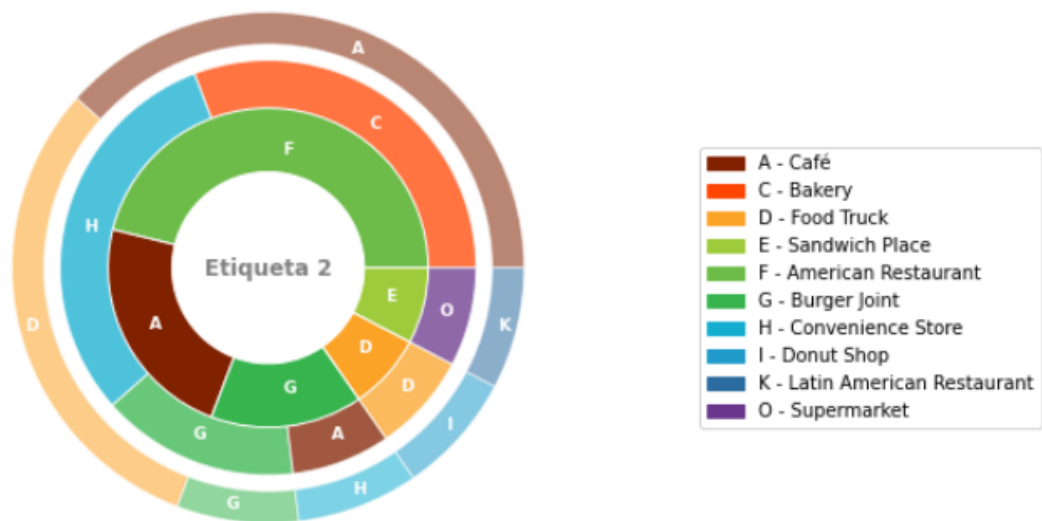
Etiqueta 1



Siendo un grupo de vecindarios predominado por las panaderías y tiendas de golosinas en sus dos primeras capas, la competencia más directa para el inversionista interesado.

Vecindario	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Chinatown	Bakery	Café	Grocery Store
City Center (Neighborhood)	Bakery	Café	Grocery Store
Grand Lake	Bakery	Café	American Restaurant
Jingletown	Bakery	Café	Convenience Store
Lynn	Bakery	Café	Grocery Store
Old Oakland	Bakery	Café	Grocery Store
Rockridge	Café	Grocery Store	Bakery
Temescal	Bakery	Café	Food Truck
Woodminster	Bakery	Burger Joint	Grocery Store

Etiqueta 2



La distribución de lugares se encuentra mayormente influenciada por la comida rápida y de fácil acceso al consumidor, brindando una mayor variedad y competidores clave en la industria.

Vecindario	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Cleveland Heights	Café	Convenience Store	Food Truck
Coliseum Industrial (Neighborhood)	American Restaurant	Burger Joint	Convenience Store
Fruitvale Station (neighborhood)	Burger Joint	Café	Food Truck
Haddon Hill	Café	Convenience Store	Food Truck
Ivy Hill	Café	Convenience Store	Donut Shop
Jack London Square	American Restaurant	Bakery	Café
Koreatown Northgate	American Restaurant	Bakery	Café
Merritt	Sandwich Place	Supermarket	Café
Northgate	American Restaurant	Bakery	Café
North Oakland	Burger Joint	Convenience Store	Food Truck
Patten	American Restaurant	Bakery	Burger Joint
Tuxedo	Food Truck	Burger Joint	Latin American Restaurant
Webster	American Restaurant	Food Truck	Café

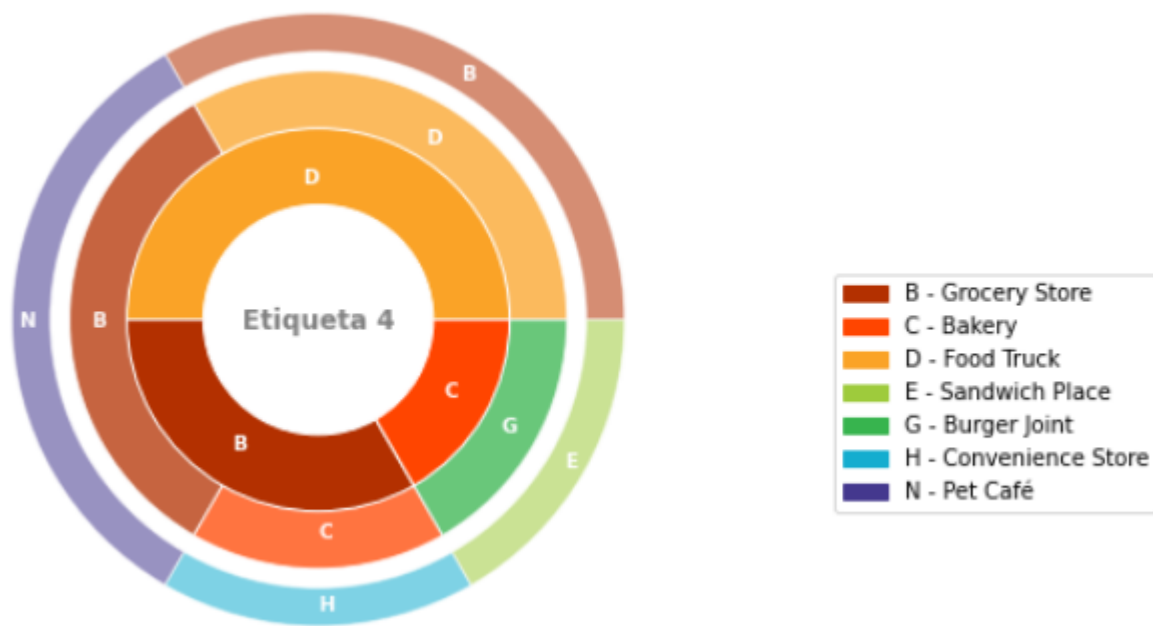
Etiqueta 3



Un grupo de vecindarios más retirados, acordonando las áreas residenciales de la ciudad de Oakland, compuestos en su mayor parte por panaderías y tiendas de Bagels.

Vecindario	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Downtown Oakland	Bakery	Bagel Shop	Café
Lakeshore	American Restaurant	Bakery	Café
Mills College	Bakery	Cupcake Shop	Sandwich Place
Montclair	Bagel Shop	Bakery	Sandwich Place
Montclair Village	Bagel Shop	Bakery	Sandwich Place
Mosswood	Bagel Shop	Bakery	Street Food Gathering

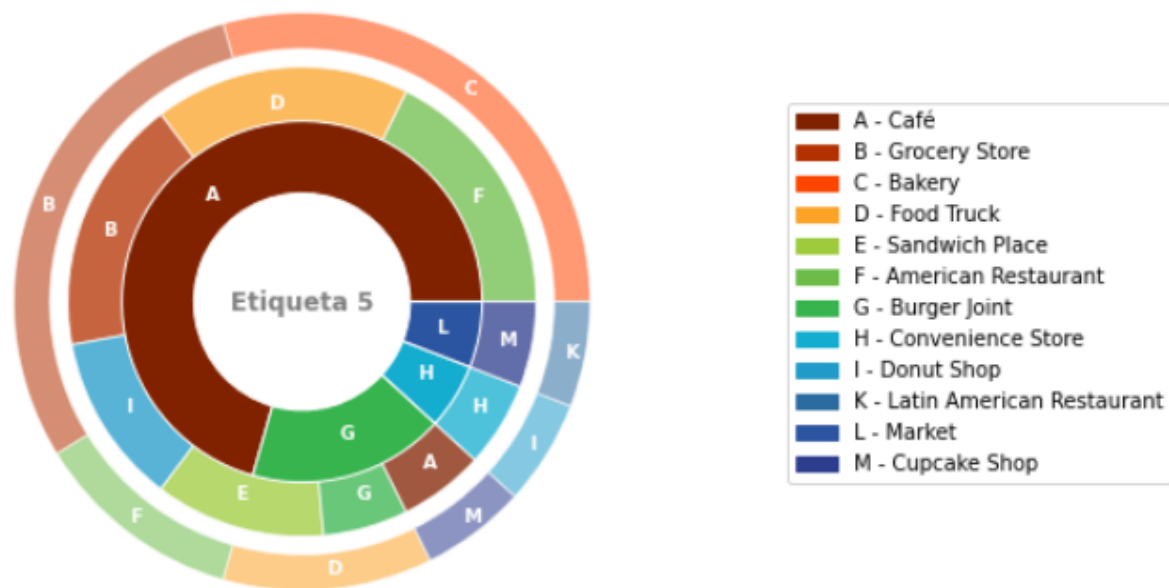
Etiqueta 4



La comida informal y tiendas de golosinas, predominan entre los distintos niveles del grupo de vecindarios. Siendo competidores que puedan satisfacer la necesidad de cualquier consumidor.

Vecindario	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Adams Point	Grocery Store	Burger Joint	Convenience Store
Glen Echo	Food Truck	Grocery Store	Pet Café
Hawthorne	Food Truck	Grocery Store	Pet Café
Oak Tree	Food Truck	Bakery	Grocery Store
Pill Hill	Bakery	Food Truck	Grocery Store
Rancho San Antonio (neighborhood)	Grocery Store	Food Truck	Sandwich Place

Etiqueta 5

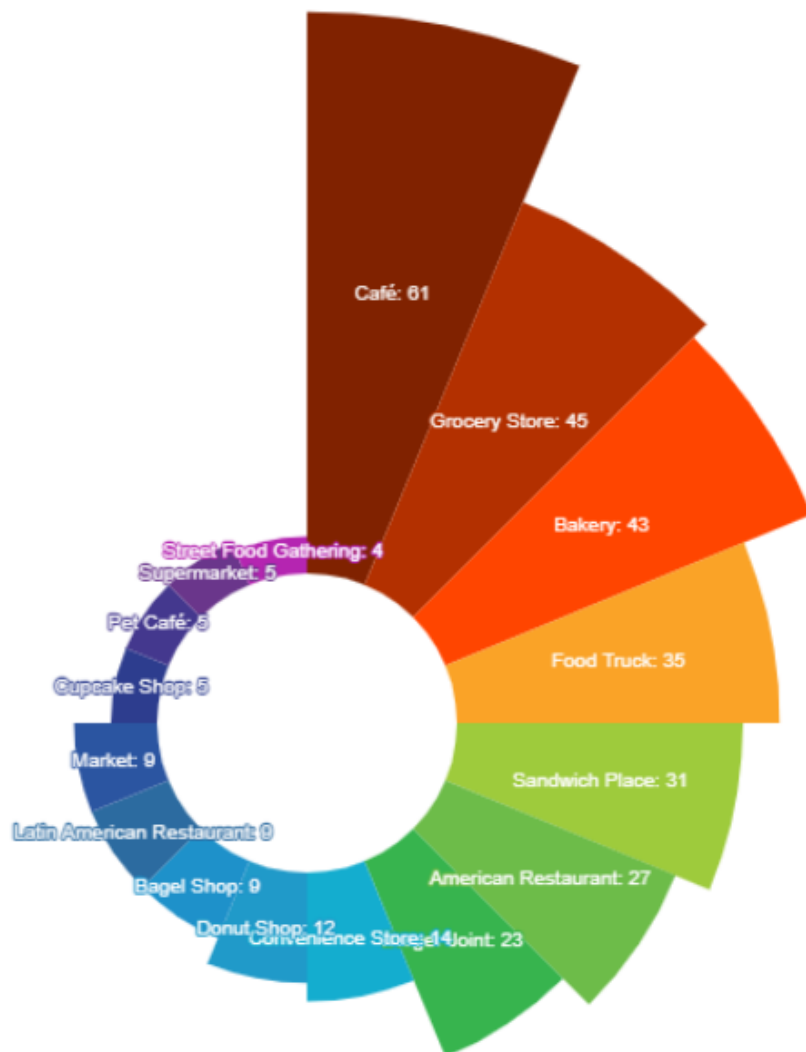


El grupo de vecindarios con mayor distribución de lugares por categoría, teniendo los restaurantes de comida latinoamericana Café y tiendas de golosinas la mayor parte de participación en este mercado.

Vecindario	1st Lugar mas comun	2nd Lugar mas comun	3rd Lugar mas comun
Civic Center	Café	Food Truck	Bakery
Dimond District	Café	Grocery Store	Bakery
Foothill Square (neighborhood)	Burger Joint	Convenience Store	Grocery Store
Fremont	Café	Donut Shop	Food Truck
Lakeside	Café	American Restaurant	Cupcake Shop
Lower Dimond	Café	Grocery Store	Bakery
Paradise Park	Café	Cupcake Shop	Donut Shop
Peralta Hacienda	Market	Burger Joint	Grocery Store
Piedmont Avenue	Convenience Store	American Restaurant	Bakery
Prescott	Café	Food Truck	Grocery Store
Redwood Heights	Burger Joint	Café	Grocery Store
San Pablo Gateway	Café	Sandwich Place	American Restaurant
South Prescott	Café	Food Truck	Grocery Store
St. Elizabeth	Burger Joint	Grocery Store	Latin American Restaurant
Uptown	Café	Sandwich Place	American Restaurant
Waverly	Café	American Restaurant	Bakery
West Oakland	Café	Donut Shop	Food Truck

Debate

Consideramos la frecuencia de categorías de la ciudad de Oakland en base a los 51 vecindarios y obtenemos el siguiente gráfico de rosa, siendo las Panaderías el 3er local con más penetración en dichos vecindarios en análisis.



En base al análisis de K-means y la distribución de frecuencias anterior, se recomienda al inversionista iniciar operaciones en el grupo de vecindarios de la **Etiqueta 5**, al estar geográficamente posicionado en puntos estratégicos mayor poblados de la ciudad de Oakland y no teniendo mayor participación por las Panaderías locales, brinda una ventaja competitiva para la industria Panificadora y derivados.

Conclusiones

En base al análisis realizado, su finalidad es comprender la composición de los clusters a una escala mayor, pudiendo ser las regulaciones urbanas en los comercios, la plusvalía y la distribución poblacional variables determinantes que puedan apoyar y brindar la ubicación definitiva dentro del cluster de vecindarios recomendados dentro de esta investigación para la ciudad de Oakland, el factor determinante en el éxito de las operaciones para el inversionista interesado.

En base al análisis de K-means y la distribución de frecuencias anterior, se recomienda al inversionista iniciar operaciones en el grupo de vecindarios de la Etiqueta 5, al estar geográficamente posicionado en puntos estratégicos mayor poblados de la ciudad de Oakland y no teniendo mayor participación por las Panaderías locales, brinda una ventaja competitiva para la industria Panificadora y derivados.