A Simple Methodology for Coffeeshop Site Selection Using Geospatial Analysis

Mark Attwood

May 12, 2023

1 Motivation and Guiding Questions

We have a successful coffeeshop and are interested in choosing an additional location to grow our business. The goals for the new location are: 1) increase profit and revenue, 2) expand our brand, and 3) community-building. Community-building entails forming a diverse, safe and welcoming social environment where people from all walks of life can share ideas and conversation. Our initial model-building process began by asking a few basic questions:

- Where do our customers live?
- Where do our customers work?
- Does walkability matter?
- For how long will the median person travel in order to get their morning coffee and pastry?
- Do neighborhoods with similar customers have a comparable business within the median established travel time?

The above questions suggest an ideal model would be similar to the one proposed by David Huff [Huf63]:

$$P_{ij} = \frac{A_j^{\alpha} D_{ij}^{-\beta}}{\sum\limits_{k=1}^{n} A_k^{\alpha} D_{ik}^{-\beta}}$$

where:

- P_{ij} is the probability of customer i visiting store j.
- A_j is a measure of the attractiveness of store j.
- D_{ij} is the distance from the consumer's location, i, to store j.
- α is an attractiveness parameter.
- β is a distance decay parameter.
- n is the total number of stores, including store j.

However, this model demands a set of input data that we do not have access to: customer locations and individual-level demographics. In order to get this data we would need to pay for a professional survey of our customer base. Now what can we do?

2 A Simple Model Using ArcGIS Pro Tools and a Python Data-Scraping Script

Our approach is simple:

1. Map out the location of each competing coffeeshop along with their service radius (more on this later).

2. Plot these radii against available Census-block level demographic and market sales data.

3. Locate the gaps in these service areas and pick potential sites that will have access to thoroughfares as well as good coverage of high-demand demographics regions.

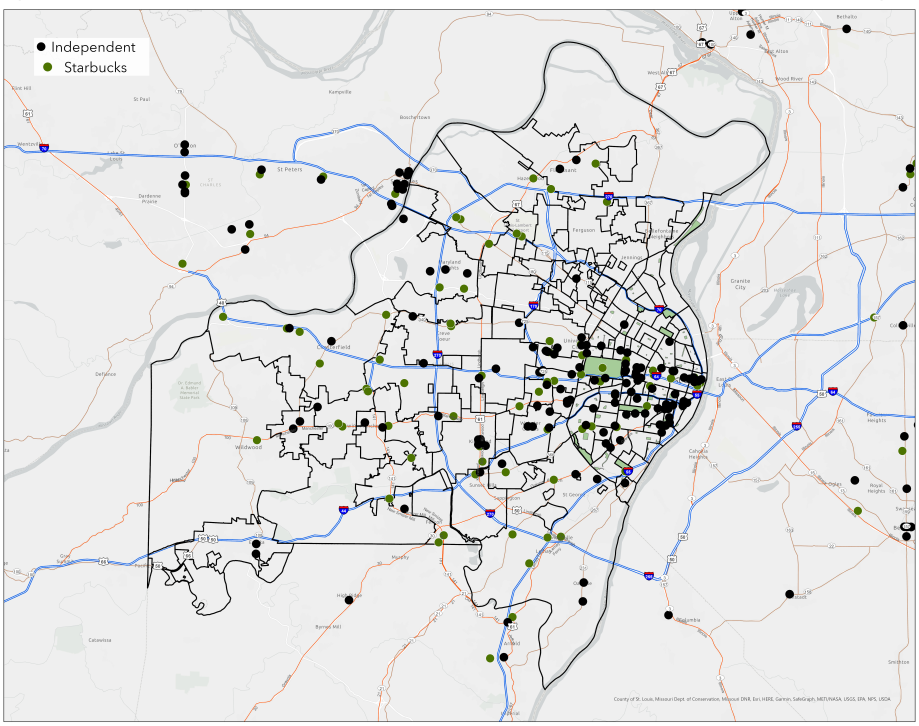
4. Find the service radii of these new sites and then derive the Census-block level value captured by each polygon as well as the various demographic features of interest.

5. Choose the site which has the highest expected value. (We will ignore locational costs for simplicity in this analysis).

2.1 Acquiring Up-To-Date Locational Data for Area Coffeeshops

Our first order of business is the get an up-to-date list of every relevant business that operates as a coffeeshop.

Figure 1: Locations of Independent Coffeeshops in the Saint Louis Region

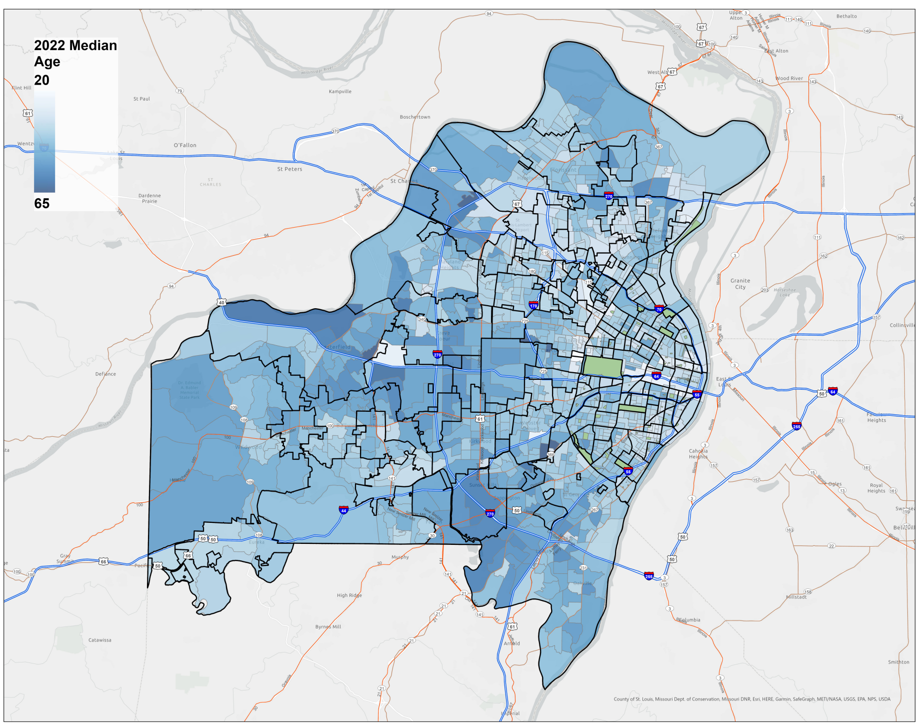


Yelp has an active community that is quick to update and create business profiles while also giving ratings which we can extract. Tied in with this data is the business address as well as an expense value that rates a business on a 1-4 scale from cheapest to most expensive. We were able to extract every business within a 50 mile radius of Saint Louis's centroid that contained "coffee" as one of the 3 category variables using Python and the Requests API (code is included in an appendix). This data was then winnowed down based on additional parameters such as whether the business is a chain and whether there is table service. As for chain stores, we retained the Starbucks locations for side-analysis and comparison (though this is omitted in this study). In general, we view Starbucks as an indicator of a healthy coffee market rather than as competition. Our results were then geocoded with ArcGIS Pro's geocoding tool. The results can be seen in Figure 1.

2.2 Building-Out the Demographic Map

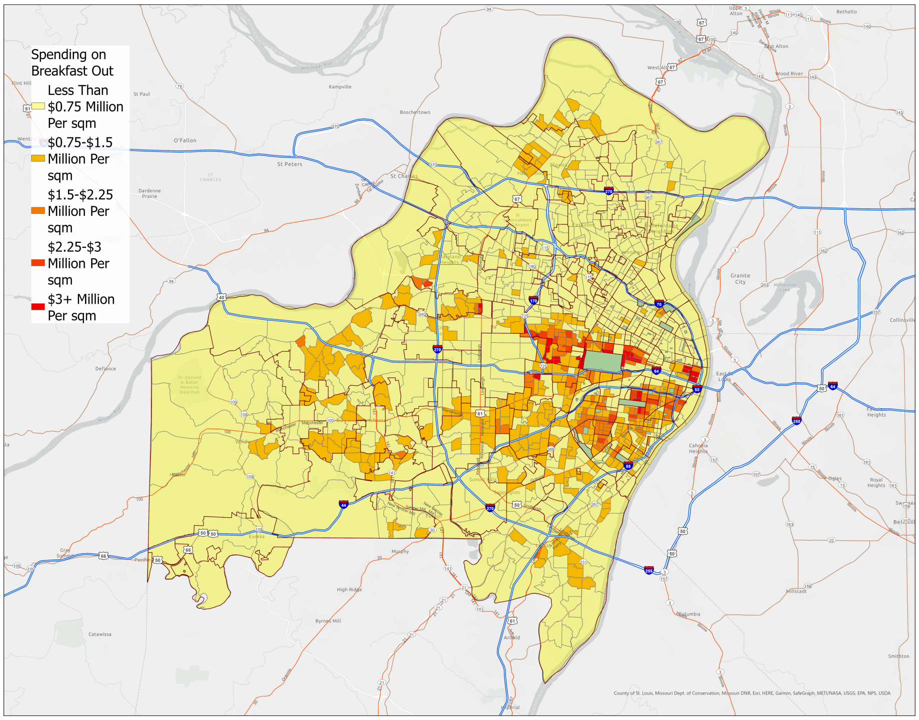
Esri (ArcGIS Pro's parent company) has a wealth of data available through a multitude of avenues. For ease of access we will be using ArcGIS Pro's enrich feature [dev]. The enrich feature breaks data down by polygon area and is, for the most part, natively collected on the Census block group level [Deva]. With this knowledge in hand, we can visit the US Census site [Bur] and download the Census block group shape files filtered for Saint Louis's FIPS codes. These polygons provide a natural canvas for the demographic features we will be using and require little guess work and fudging because they are the native aggregation areas of our data sets.

Figure 2: Median Age Within Census Block Groups



While several features were considered, there was a strong correlation between education, income, and spending on breakfast-out so we opted to keep matters simple and focus on the breakfast-out demographic feature. Additionally, our age demographic tends to be younger than our current surrounding area and we are looking for a median age of 35-40 for the population of a surround site. This number is anecdotally derived as the observed median of our customer base. Figures 2 (Age) and 3 (Breakfast) illustrate these layers as color gradient chloropleths. Of note is that we have normalized the expenditures on breakfast by polygon area (by square miles) in order to create a “spending density” pseudo-measure.

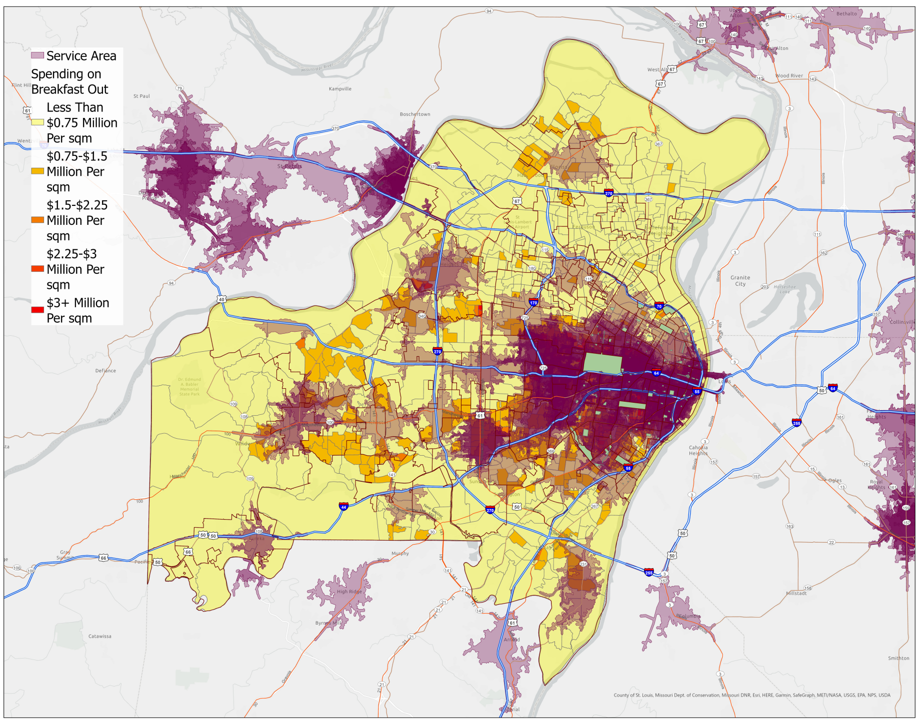
Figure 3: Spending on Breakfast-Out Per Square Miles Within Census Block Groups



2.3 Determining and Extrapolating Coffeeshop Service Areas

A 2022 survey of over 2000 random stratified participants conducted by consumer research firm Access Development [Devb] found that for regular small purchases consumers tend to travel no more than 6 minutes. Likewise, we will define our service area by a 5 minute travel radius. Next, we must consider how our consumers will get to our locations. For Saint Louis, a large dataset and deep analysis is not needed: everyone drives and very few areas about the region are pedestrian friendly. In contrast, if we were to run a similar analysis for San Francisco, we would look more toward walking radii and we would need to take access from public transportation under consideration. The Business Analyst Generate Drive Time Trade Area tool [Esr] handles all of the above listed modes of transportation. Figure 4 illustrates the service areas in purple, with deeper hues where there are more overlapping areas.

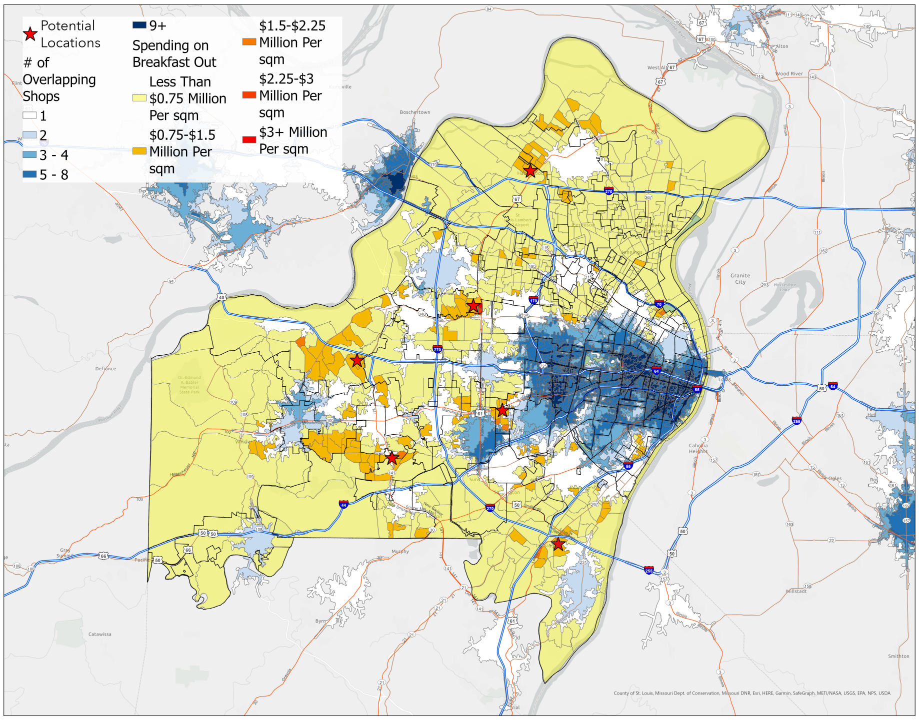
Figure 4: Independent Coffeeshop Market Coverage Overlaid on Breakfast-Out Expenses



2.4 Selection of Several Potential Sites

Our next order of business is to get a cleaner map of the market coverage of independent coffeeshops. The Count Overlapping Features (Analysis) tool makes quick work of this query (Figure 5).

Figure 5: Independent Coffeeshop Market Saturation Overlaid on Breakfast-Out Expenses



It is now visually clear where there is market saturation and where there are gaps in service. Furthermore, we can see where there is a demand for breakfast-out, where our age demographic is located, and pick sites near these demographic groups with good thoroughfare access. Sites were chosen by simple visual inspection. Our chosen potential sites and their service areas can be observed in Figures 6 (Breakfast) and 7 (Age).

Figure 6: Service Areas of Potential Sites Over Breakfast-Out Spending

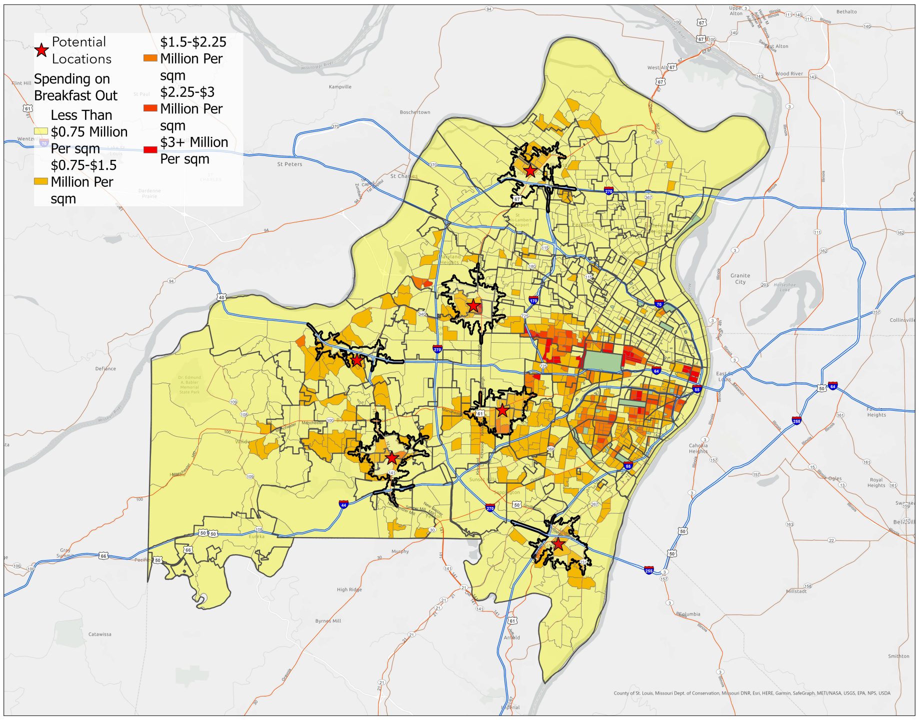
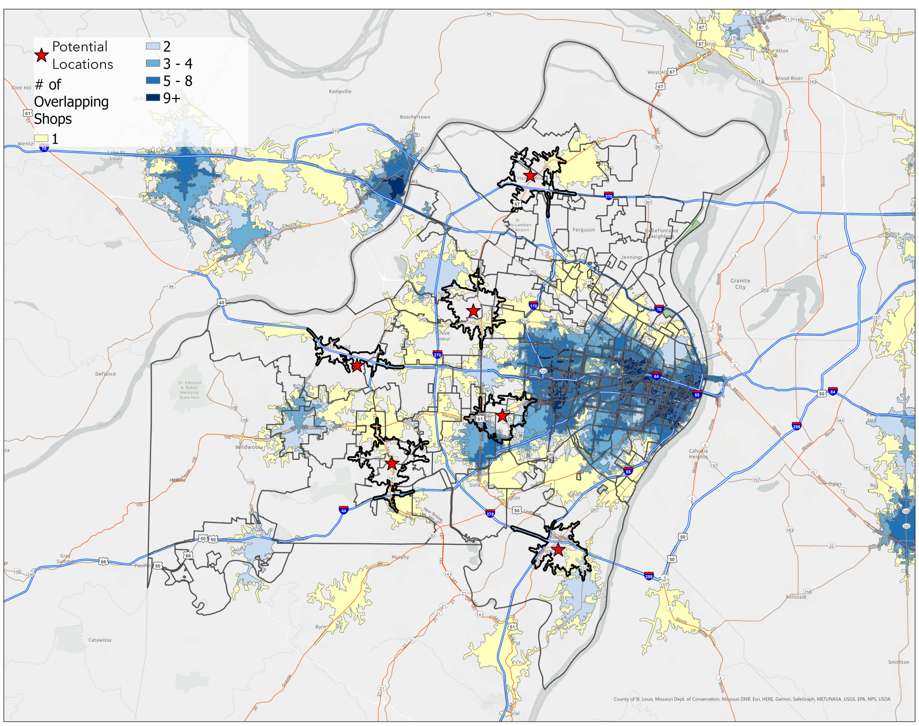


Figure 7: Service Areas of Potential Sites Against Competitor Service Areas



2.5 Analysis of Potential Site Demographics and Sales Potential

We are on the final step; we have our potential sites selected and probably have a good idea of which one will be the most likely to succeed. However, we can do better. By re-applying the enrich tool to our new polygons when can get more precise demographic values for comparison among our candidate sites. Figures 8 (Age) and 9 (Breakfast) illustrate visually the information describe in the following table.

Figure 8: Median Age Within Potential Site Market Areas

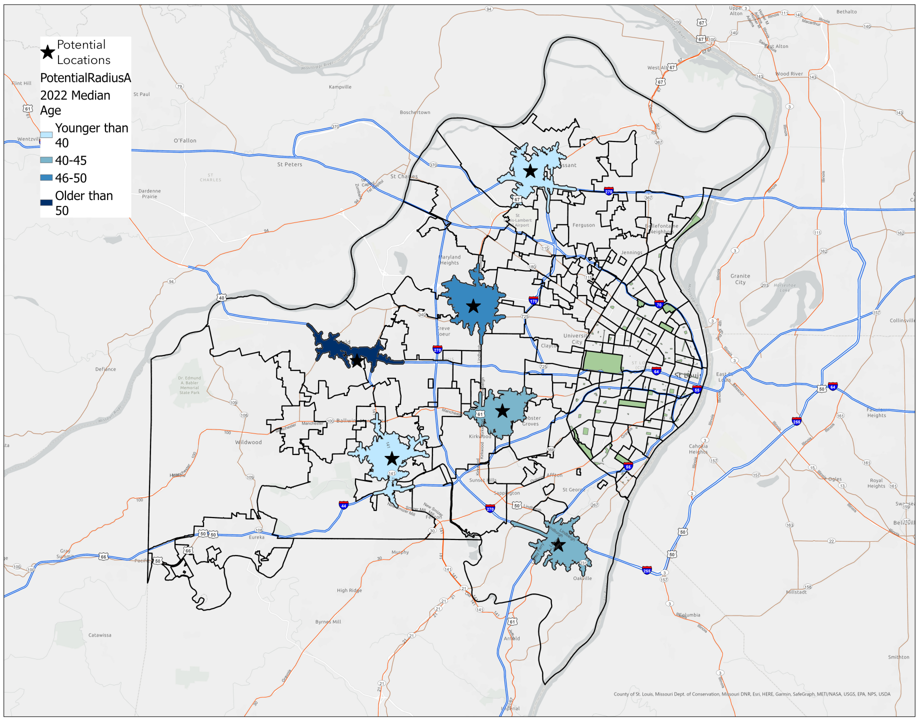
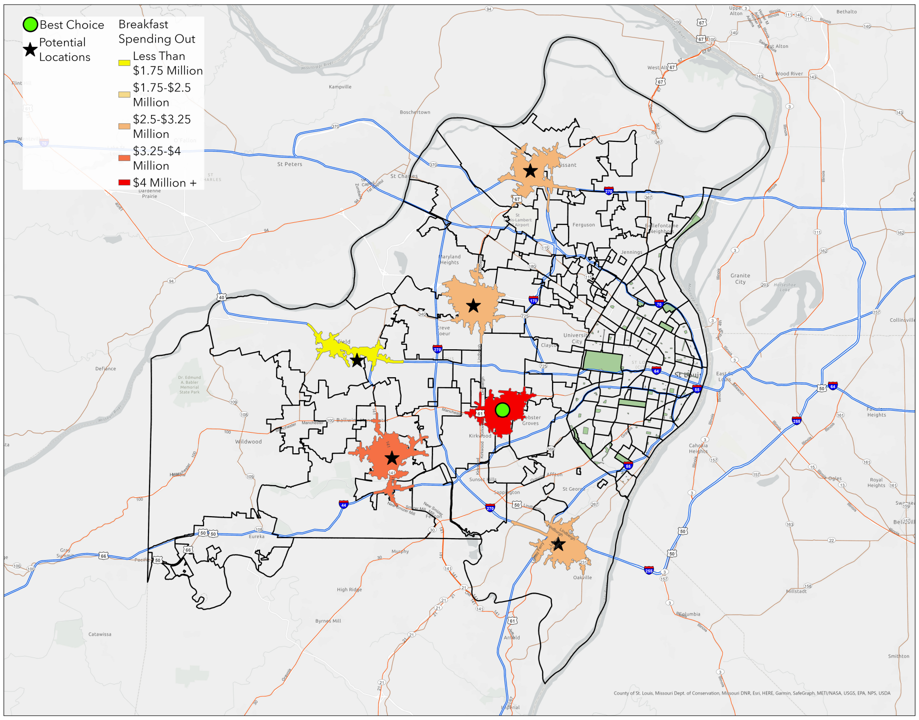


Figure 9: Spending on Breakfast-Out Within Potential Site Market Areas



Area	**Median Age**	**Breakfast-Out Spending**	**Latitude**	**Longitude**
Chesterfield	54	$1,008,338	38.640139	-90.528505
Creve Coeur	46	$3,128,031	38.682696	-90.412925
Florissant	39	$2,501,145	38.787744	-90.356072
Mehlville	43	$2,825,672	38.496855	-90.328981
Warson Woods	43	$4,990,433	38.601124	-90.384158
Valley Park	40	$3,690,201	38.564031	-90.493853

The Warson Wood site (green) is our clear front-runner with a much higher potential market value, though the median age is a bit high and may suppress our expectations a bit. Our next choice, should that location fall through, would be Valley Park (orange). Its age demographic matches our current customer base and it has the next highest potential market value.

3 Conclusions

In summary, with relatively little cost we have been able to find potential locations for a new business that will have a much higher success rate than one chosen arbitrarily. This analysis only scratches the surface of what can be achieved with geospatial analysis and the right datasets. Our model could be greatly improved with customer-level data on where they live and where they work. Transit paths could be mapped and added to our market area measure for much better estimates.

Yelp_Data

May 2, 2023

[1]:
```python
import requests
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from YelpAPIKey import get_key
from copy import deepcopy
pd.options.display.max_columns = 2000
```

[2]:
```python
#Define API Key, Define the Endpoint, Define Header
API_KEY = get_key()
ENDPOINT = 'https://api.yelp.com/v3/businesses/search'
HEADERS = {'Authorization': f'bearer {API_KEY}'}
```

[3]:
```python
# Define Parameters
PARAMETERS = {'term':'coffee',
              #'radius':25000,
              'limit': 50,
              'offset': 400,
              'location': 'Saint Louis'}
```

[4]:
```python
# Make a request to Yelp API
response = requests.get(url = ENDPOINT, params = PARAMETERS, headers = HEADERS)
```

[5]:
```python
# Convert JSON string to dictionary
business_data = response.json()
```

[6]:
```python
business_data.keys()
```

[6]:
```
dict_keys(['businesses', 'total', 'region'])
```

[7]:
```python
business_data.get('total')
```

[7]:
```
1400
```

[8]:
```python
names = []
for item in business_data.get('businesses'):
    names.append(item.get('name'))
```

```python
print(names)
```

```
['Smallcakes', 'First Watch', 'Baked By Ashley', 'Crafted', 'Fresh Thyme
Market', 'Midwest Pasta', 'The Great American Bagel Bakery', 'Shaw Market',
"Grandma's Cookies", "Michelle's Cafe", "Kaldi's Coffee at Mid Campus Center",
'Urban Chestnut Brewing Company', 'Cookie Factory Bakery and Cafe', 'Sweetie Cup
Thai Cafe', 'Nordstrom Ebar Artisan Coffee', 'The Bistro and Bar', 'Local
Harvest Grocery', "Fitz's SoCo", 'Green Earth', 'Grappa Growlers', 'Britt's
Bakehouse: A Gluten-Free Bakery', 'Einstein Bros. Bagels', 'Sugarfire Pie',
'Adriana's On The Hill', 'Trolley Track Cookie', 'The Weeping Willow Tea Room',
"Drew's Donut Stop", 'Big Sky Cafe', 'Dough Depot', 'Hot Box Cookies', "Dad's
Cookie Company", 'Riverside Diner', 'The Boardwalk Cafe', 'Strange Donuts',
"Scooter's Coffee", 'The Corner Butcher', 'The Brown Bag Bistro', 'Silver Spoon
Ice Cream & Sweets', "Kelly's Donuts", 'Wei Hong Bakery', "Letty Lou's Cafe",
'Kingside Diner', 'Starbucks', "DiGregorio's Market", 'Press Waffle',
"Bissinger's Handcrafted Chocolatier", 'The Schlafly Tap Room', "Jilly's Cupcake
Bar & Café", 'Olive Oil Marketplace', 'Balkan Store & Bakery']
```

[9]:
```python
business_data.get('businesses')[8]
```

[9]:

```
{'id': 'AWN5UYfF-nto6XihHc2dIA',
 'alias': 'grandmas-cookies-st-charles-2',
 'name': "Grandma's Cookies",
 'image_url':
'https://s3-media3.fl.yelpcdn.com/bphoto/VwCJJfrxtPZEhnuUBGLhrA/o.jpg',
 'is_closed': False,
 'url': 'https://www.yelp.com/biz/grandmas-cookies-st-charles-2?adjust_creative=
vb2A4YF95dm9CQOW13BC3w&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_searc
h&utm_source=vb2A4YF95dm9CQOW13BC3w',
 'review_count': 143,
 'categories': [{'alias': 'desserts', 'title': 'Desserts'}],
 'rating': 4.5,
 'coordinates': {'latitude': 38.77868, 'longitude': -90.48301},
 'transactions': ['delivery', 'pickup'],
 'price': '$',
 'location': {'address1': '401 S Main St',
  'address2': '',
  'address3': '',
  'city': 'St Charles',
  'zip_code': '63301',
  'country': 'US',
  'state': 'MO',
  'display_address': ['401 S Main St', 'St Charles, MO 63301']},
 'phone': '+16369470088',
 'display_phone': '(636) 947-0088',
 'distance': 23216.90161972317}
```

[ ]:

```

```

[10]:

```
businesses = []
for offset in np.arange(0, 1000, 50):
    PARAMS = {'term':'coffee',
              #'radius':25000,
              'limit': 50,
              'offset': offset,
              'location': 'Saint Louis'}
    response = requests.get(url = ENDPOINT, params = PARAMS, headers = HEADERS)
    data = response.json()
    for item in data.get('businesses'):
        businesses.append(item)
```

[11]:

```
type(businesses[1])
```

[11]: dict

[12]:

```
# maximum number of categories represented
nums = []
for item in businesses:
    nums.append(len(item.get('categories')))
print(max(nums))
```

3

[13]:

```
tester = deepcopy(businesses)
```

[14]:

```
for business in businesses:
    if len(business.get('categories')) == 3:
        business["cat_1"] = business.get('categories')[0].get('alias')
        business["cat_2"] = business.get('categories')[1].get('alias')
        business["cat_3"] = business.get('categories')[2].get('alias')
    if len(business.get('categories')) == 2:
        business["cat_1"] = business.get('categories')[0].get('alias')
        business["cat_2"] = business.get('categories')[1].get('alias')
        business["cat_3"] = 'None'
    if len(business.get('categories')) == 1:
        business["cat_1"] = business.get('categories')[0].get('alias')
        business["cat_2"] = 'None'
        business["cat_3"] = 'None'
businesses[1]
```

[14]:


```
{'id': '2xvewxEqwSkOK2G9Z2lLsA',
 'alias': 'la-finca-coffee-shop-st-louis',
 'name': 'La Finca Coffee Shop',
 'image_url':
```

'https://s3-media2.fl.yelpcdn.com/bphoto/NfflPKO8x5mdpmCiOa4riw/o.jpg',
 'is_closed': False,
 'url': 'https://www.yelp.com/biz/la-finca-coffee-shop-st-louis?adjust_creative=vb2A4YF95dm9CQOW13BC3w&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=vb2A4YF95dm9CQOW13BC3w',
 'review_count': 8,
 'categories': [{'alias': 'coffee', 'title': 'Coffee & Tea'}],
 'rating': 5.0,
 'coordinates': {'latitude': 38.6263785750439,
  'longitude': -90.26076674330857},
 'transactions': [],
 'location': {'address1': '4440 Manchester Ave',
  'address2': '',
  'address3': None,
  'city': 'St. Louis',
  'zip_code': '63110',
  'country': 'US',
  'state': 'MO',
  'display_address': ['4440 Manchester Ave', 'St. Louis, MO 63110']},
 'phone': '',
 'display_phone': '',
 'distance': 5488.11726188813,
 'cat_1': 'coffee',
 'cat_2': 'None',
 'cat_3': 'None'}

[15]:
```python
for business in businesses:
    business["street_address"] = business.get("location").get("address1")
    business["zip_code"] = business.get("location").get("zip_code")
businesses[1]
```

[15]:
{'id': '2xvewxEqwSkOK2G9Z2lLsA',
 'alias': 'la-finca-coffee-shop-st-louis',
 'name': 'La Finca Coffee Shop',
 'image_url':
'https://s3-media2.fl.yelpcdn.com/bphoto/NfflPKO8x5mdpmCiOa4riw/o.jpg',
 'is_closed': False,
 'url': 'https://www.yelp.com/biz/la-finca-coffee-shop-st-louis?adjust_creative=vb2A4YF95dm9CQOW13BC3w&utm_campaign=yelp_api_v3&utm_medium=api_v3_business_search&utm_source=vb2A4YF95dm9CQOW13BC3w',
 'review_count': 8,
 'categories': [{'alias': 'coffee', 'title': 'Coffee & Tea'}],
 'rating': 5.0,
 'coordinates': {'latitude': 38.6263785750439,
  'longitude': -90.26076674330857},
 'transactions': [],
 'location': {'address1': '4440 Manchester Ave',

```
 'address2': '',
 'address3': None,
 'city': 'St. Louis',
 'zip_code': '63110',
 'country': 'US',
 'state': 'MO',
 'display_address': ['4440 Manchester Ave', 'St. Louis, MO 63110']},
'phone': '',
'display_phone': '',
'distance': 5488.11726188813,
'cat_1': 'coffee',
'cat_2': 'None',
'cat_3': 'None',
'street_address': '4440 Manchester Ave',
'zip_code': '63110'}
```

[16]:
```
for business in businesses:
    business['latitude'] = business.get('coordinates').get('latitude')
    business['longitude'] = business.get('coordinates').get('longitude')
```

[ ]:

[17]:
```
df = pd.DataFrame.from_dict(businesses)
```

[18]:
```
df.head()
```

[18]:
```
                       id                                      alias  \
0  HUKrpwmcMlwPbDF5cz0VkQ  maypop-coffee-and-garden-shop-webster-groves
1  2xvewxEqwSk0K2G9Z21LsA               la-finca-coffee-shop-st-louis
2  JPTxxPgVUEkb9BJl06PAhw       la-cosecha-coffee-roasters-maplewood-2
3  8JRdV8M8b2bIZGH92gMQ0Q        coma-coffee-roasters-richmond-heights
4  hv-rci0u6KR75ALB_70dHw                blueprint-coffee-saint-louis

                            name  \
0  Maypop Coffee & Garden Shop
1         La Finca Coffee Shop
2   La Cosecha Coffee Roasters
3         Coma Coffee Roasters
4             Blueprint Coffee

                                            image_url  is_closed  \
0  https://s3-media4.fl.yelpcdn.com/bphoto/Knz83_…      False
1  https://s3-media2.fl.yelpcdn.com/bphoto/NfflPK…      False
2  https://s3-media1.fl.yelpcdn.com/bphoto/X2AkTN…      False
3  https://s3-media3.fl.yelpcdn.com/bphoto/iwpmdl…      False
4  https://s3-media1.fl.yelpcdn.com/bphoto/piqdTq…      False
```

```
                                               url  review_count  \
0  https://www.yelp.com/biz/maypop-coffee-and-gar…            82
1  https://www.yelp.com/biz/la-finca-coffee-shop-…             8
2  https://www.yelp.com/biz/la-cosecha-coffee-roa…            63
3  https://www.yelp.com/biz/coma-coffee-roasters-…           200
4  https://www.yelp.com/biz/blueprint-coffee-sain…           366

                                        categories  rating  \
0  [{'alias': 'coffee', 'title': 'Coffee & Tea'},…     5.0
1     [{'alias': 'coffee', 'title': 'Coffee & Tea'}]     5.0
2  [{'alias': 'coffeeroasteries', 'title': 'Coffe…     4.5
3  [{'alias': 'coffeeroasteries', 'title': 'Coffe…     4.5
4  [{'alias': 'coffeeroasteries', 'title': 'Coffe…     4.5

                                       coordinates       transactions  \
0     {'latitude': 38.60432, 'longitude': -90.33716}         [delivery]
1  {'latitude': 38.6263785750439, 'longitude': -9…                  []
2   {'latitude': 38.61261, 'longitude': -90.3192785}         [delivery]
3     {'latitude': 38.63644, 'longitude': -90.34447}  [delivery, pickup]
4  {'latitude': 38.6556838057126, 'longitude': -9…         [delivery]

                                          location         phone  \
0  {'address1': '803 Marshall Ave', 'address2': '…  +13147642140
1  {'address1': '4440 Manchester Ave', 'address2'…
2  {'address1': '7360 Manchester Rd', 'address2':…  +13149258880
3  {'address1': '1034 S Brentwood Blvd', 'address…  +13142501042
4  {'address1': '6225 Delmar Blvd', 'address2': '…  +13142666808

    display_phone     distance             cat_1      cat_2            cat_3  \
0  (314) 764-2140  1646.303448            coffee  gardening             None
1                  5488.117262            coffee       None             None
2  (314) 925-8880   194.610612  coffeeroasteries       None             None
3  (314) 250-1042  3332.371913  coffeeroasteries     coffee  breakfast_brunch
4  (314) 266-6808  5139.762623  coffeeroasteries       None             None

          street_address zip_code   latitude  longitude price
0        803 Marshall Ave    63119  38.604320 -90.337160   NaN
1     4440 Manchester Ave    63110  38.626379 -90.260767   NaN
2      7360 Manchester Rd    63143  38.612610 -90.319278     $
3  1034 S Brentwood Blvd    63117  38.636440 -90.344470    $$
4        6225 Delmar Blvd    63130  38.655684 -90.300594    $$
```

[19]:
```
df.columns
```

[19]:
```
Index(['id', 'alias', 'name', 'image_url', 'is_closed', 'url', 'review_count',
       'categories', 'rating', 'coordinates', 'transactions', 'location',
       'phone', 'display_phone', 'distance', 'cat_1', 'cat_2', 'cat_3',
```

```
       'street_address', 'zip_code', 'latitude', 'longitude', 'price'],
      dtype='object')
```

[20]:
```
drop_columns = ['id', 'alias', 'image_url', 'url', 'categories', 'coordinates',␣
 ↪'transactions', 'location',
       'phone', 'display_phone', 'is_closed']
df.drop(labels = drop_columns, axis=1, inplace=True)
```

[21]:
```
df.head()
```

[21]:
```
                         name  review_count  rating     distance  \
0  Maypop Coffee & Garden Shop            82     5.0  1646.303448
1         La Finca Coffee Shop             8     5.0  5488.117262
2   La Cosecha Coffee Roasters            63     4.5   194.610612
3         Coma Coffee Roasters           200     4.5  3332.371913
4             Blueprint Coffee           366     4.5  5139.762623

              cat_1      cat_2             cat_3         street_address  \
0            coffee  gardening              None       803 Marshall Ave
1            coffee       None              None     4440 Manchester Ave
2  coffeeroasteries       None              None      7360 Manchester Rd
3  coffeeroasteries     coffee  breakfast_brunch  1034 S Brentwood Blvd
4  coffeeroasteries       None              None        6225 Delmar Blvd

  zip_code   latitude  longitude price
0    63119  38.604320 -90.337160   NaN
1    63110  38.626379 -90.260767   NaN
2    63143  38.612610 -90.319278     $
3    63117  38.636440 -90.344470    $$
4    63130  38.655684 -90.300594    $$
```

[22]:
```
df_2 = df.loc[(df["cat_1"].str.match('coffee'))
       | (df["cat_2"].str.match('coffee'))
       | (df["cat_3"].str.match('coffee'))]
```

[23]:
```
df[df["review_count"]>50].rating.value_counts(normalize=False, sort=False,␣
 ↪dropna=False).sort_index(ascending=False)
```

[23]:
```
5.0     17
4.5    163
4.0    130
3.5     64
3.0     15
2.5     10
2.0      6
1.5      4
Name: rating, dtype: int64
```

```
[24]: df_2.tail(20)
```

```
[24]:
             name  review_count  rating      distance        cat_1  \
923     Starbucks            70     2.0  14163.963572       coffee
926     Starbucks             4     3.0  17121.099754       coffee
927    McDonald's            27     2.0  20473.931698      burgers
929     Starbucks            13     2.0  14842.153321       coffee
930     Starbucks             3     2.5  19745.970151       coffee
931      7-Eleven             8     2.5   5360.067858  convenience
938      7-Eleven             5     4.5  16650.373752  convenience
949      7-Eleven             2     4.5  10458.777180  convenience
955    McDonald's            33     2.5  26838.705991      hotdogs
957        Dunkin'           15     1.5  16059.031043       coffee
960    McDonald's            15     2.5   7193.372009      burgers
962      7-Eleven             3     3.5  14288.813274  convenience
964        Dunkin'           34     2.0  14900.770031       coffee
968    McDonald's            18     3.5   8323.616595      hotdogs
973    McDonald's             8     2.0   9381.710159      hotdogs
982    McDonald's            46     1.0   3264.637731      hotdogs
987  Krispy Kreme            47     2.5  31166.789278       donuts
988      7-Eleven             3     2.5   7244.737215  convenience
991        Dunkin'            4     3.0  19853.061106       coffee
992      7-Eleven             2     4.0  22223.480750  convenience

               cat_2   cat_3                     street_address zip_code  \
923             None    None           10701 Natural Bridge Rd    63145
926             None    None  1272 Town And Country Crossing Dr    63011
927          hotdogs  coffee                    300 Columbia Ctr    62236
929             None    None     10701 Lambert International Blvd    63145
930             None    None                15025 Manchester Rd    63011
931           coffee    None                 3160 Morganford Rd    63116
938  servicestations  coffee                6197 Lemay Ferry Rd    63129
949           coffee    None                 1193 Colonnade Ctr    63131
955          burgers  coffee                   24 Harvester Sq    63303
957           donuts    None                   1410 Big Bend Rd    63088
960          hotdogs  coffee               8127 Olive Street Rd    63130
962  servicestations  coffee                  13515 Big Bend Rd    63122
964           donuts    None            10701 Lambert Intl Blvd    63145
968          burgers  coffee               9406 Olive Street Rd    63132
973          burgers  coffee            4979 Natural Bridge Rd    63115
982          burgers  coffee                    7259 Watson Rd    63119
987           coffee    None            6021 Mid Rivers Mall Dr    63304
988           coffee    None                   8159 Olive Blvd    63130
991           donuts    None              8115 N Lindbergh Blvd    63031
992           coffee    None                 900 Shackleford Rd    63031

     latitude  longitude price
```

```
923  38.737026 -90.355081    $$
926  38.620260 -90.518156   NaN
927  38.446791 -90.218471     $
929  38.741518 -90.364688   NaN
930  38.594571 -90.547437   NaN
931  38.602260 -90.261043     $
938  38.465367 -90.357605     $
949  38.602643 -90.441066     $
955  38.746421 -90.578490     $
957  38.567516 -90.496973     $
960  38.673910 -90.346998     $
962  38.567687 -90.475481     $
964  38.742004 -90.364974     $
968  38.673646 -90.376450     $
973  38.676216 -90.250689     $
982  38.583090 -90.319115     $
987  38.747988 -90.635592    $$
988  38.674049 -90.348332     $
991  38.789802 -90.347145   NaN
992  38.810808 -90.352153     $
```

[25]:
```
df_2 = df_2[df_2['name']!="McDonald\'s"]
```

[26]:
```
df_2 = df_2 = df_2[df_2['name']!="7-Eleven"]
```

[27]:
```
df_2 = df_2 = df_2[df_2['name']!="Dunkin\'"]
```

[30]:
```
len(df_2)
```

[30]:
```
324
```

[32]:
```
df_2["cost"] = np.where(df_2["price"]=="$", 1,
                        np.where(df_2["price"]=="$$", 2,
                                np.where(df_2["price"]=="$$$", 3, np.nan)))
```

[33]:
```
df_2["cost"].value_counts(dropna=False)
```

[33]:
```
NaN    126
2.0    117
1.0     79
3.0      2
Name: cost, dtype: int64
```

[34]:
```
df_2.to_csv("yelp_stl_coffee_data.csv")
```

[ ]:
```

```

References

[Bur] United States Census Bureau. *Cartographic Boundary Files*. URL: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html. (accessed: 05.01.2023).

[Deva] ArcGIS Developer. *Data apportionment*. URL: https://developers.arcgis.com/rest/geoenrichment/api-reference/data-apportionment.htm#ESRI_SECTION1_206F094B25E540B3A01DB46E499EB08C. (accessed: 05.01.2023).

[Devb] Access Development. *The Impact of Poximity on Consumer Purchases: A National Study Conducted by Access Development*. URL: https://ww2.accessdevelopment.com/consumer_proximity_study. (accessed: 05.10.2023).

[dev] Esri data development. *2022/2027 Consumer Spending*. URL: https://storymaps.arcgis.com/stories/6618fb04c4e4419ea2caa041b54436c0. (accessed: 05.07.2023).

[Esr] Esri. *Generate Drive Time Trade Areas (Business Analyst)*. URL: https://pro.arcgis.com/en/pro-app/latest/tool-reference/business-analyst/generate-drive-time-trade-area.html. (accessed: 04.07.2023).

[Huf63] David L. Huff. “A Probabilistic Analysis of Shopping Center Trade Areas”. In: *Land Economics* 39.1 (1963), pp. 81–90.