JAT's Capstone Project
Data Section
"Dinner and a Date"

Jerry Ackaret

Data to Use in JAT's Capstone Project

- Primarily, only Foursquare data is necessary to accomplish this project, with the exception of determining Portland, OR latitude and longitude, using geopy.geocoders.
- The Foursquare data to be used is in the Explore and Details Endpoints:
  - Explore Endpoint
    - Limit the Fun Venue results to Museums, Art Galleries, Concert Halls, Historic Sites, Stadiums, Zoos
    - Near Portland, OR (there is a 'near' function inside of Foursquare)
  - Example Output:


```
'reasonName': 'globalInteractionReason'}]},
'venue': {'id': '43ee1bd9f964a5205f2f1fe3',
 'name': 'Portland Art Museum',
 'location': {'address': '1219 SW Park Ave',
  'lat': 45.51622988093649,
  'lng': -122.68359661102295,
  'labeledLatLngs': [{'label': 'display',
    'lat': 45.51622988093649,
    'lng': -122.68359661102295}],
  'postalCode': '97205',
  'cc': 'US',
  'city': 'Portland',
  'state': 'OR',
  'country': 'United States',
  'formattedAddress': ['1219 SW Park Ave',
   'Portland, OR 97205',
   'United States']},
 'categories': [{'id': '4bf58dd8d48988d18f941735',
```

Data to Use in JAT's Capstone Project

- Explore continued
  - Obviously, Foursquare returns far more information than is needed, so I will scrub the data down to Name, Category, Latitude, Longitude, Address, and Venue ID.
    - The Latitude and Longitude will be needed for mapping and the Venue ID will be needed to obtain details of the Fun Venue than the explore endpoint provides.
- Data Contraction: If there are any venues that fall outside of the core part of Portland, I will drop those for being outside of the scope of the project.
- Data Expansion: Continuing to use the Explore Endpoint, find the Food Venues that are around each of the Fun Venues, within the radius that is chosen by the user (set early in the program).
  - The Food Venue Name, ID, Latitude, Longitude, Distance to Fun Venue, Category, and Address will be retrieved, and combined with the appropriate Fun Venue.
  - Example of output:

Fun_Venue	Fun_Address	Fun_Latitude	Fun_Longitude	Fun_ID	Venue	Venue_ID	Venue_Latitude	Venue_Longitude
Portland Art Museum	1219 SW Park Ave	45.51623	-122.683597	43ee1bd9f964a5205f2f1fe3	Higgins Restaurant & Bar	41ddd100f964a520c51e1fe3	45.515464	-122.682074
Portland Art Museum	1219 SW Park Ave	45.51623	-122.683597	43ee1bd9f964a5205f2f1fe3	Behind The Museum Café	4ee11c6abe7b7e4d131503ff	45.516492	-122.684151
Portland Art Museum	1219 SW Park Ave	45.51623	-122.683597	43ee1bd9f964a5205f2f1fe3	Raven & Rose	50960315e4b0e34d634b48f5	45.514794	-122.682321
Portland Art Museum	1219 SW Park Ave	45.51623	-122.683597	43ee1bd9f964a5205f2f1fe3	Southpark Seafood & Oyster Bar	40b13b00f964a520fef51ee3	45.517909	-122.681596
Portland Art Museum	1219 SW Park Ave	45.51623	-122.683597	43ee1bd9f964a5205f2f1fe3	Addy's Sandwich Bar	4a8eec36f964a5202b1320e3	45.518534	-122.682720

Data to Use in JAT's Capstone Project

- Data Contraction: I do not want food venues that might not meet the level of a Dinner and a Date, and so will drop some categories:
  - 'Food Truck','Donut Shop','Bakery','Bagel Shop','Coffee Shop','Dessert Shop','Fast Food Restaurant', 'Food Court','Food Stand','Juice Bar','Snack Place'
- Data Contraction: Further reduce the Food Venues to the closet ones. The number is set by the user at the beginning of the program.

- Detail Endpoint:
  - The explore endpoint does not have phone numbers or ratings, and so I will pull just the Venue IDs from Fun and Food, retrieve the phone numbers and ratings where they exist, and combine this data back into the master list using merges.
- Output:
  - Automatically, a spreadsheet will be created that has the appropriate columns for the 'Donation Solicitor' to use in order to first contact the Fun Venue, and if they agree to provide tickets, then find an associated Food Venue. Since there can be as many results as the solicitor wishes, the odds of finding Food Venue to combine with the Fun Venue is high.
    - I will drop the Latitudes, Longitudes, and ID numbers from the spreadsheet as unimportant data for its function. Another dataframe will still be there that includes this data for use in mapping.
  - In addition, clustering will be used to find what types of food venues cluster into what regions of the city. If one assumes that the best of a breed tend to cluster, this will further enabling the Solicitor to decide on the best Food Venues from which to solicit donations.