Mini-project 2

Marc Baron

Goals

- **Use recommendation apps (Foursquare, Yelp) on local business to explore the quality and type of local businesses and restaurants**

Process

- Access the API of the respective sites and store the data
- Clean up the data so it is somewhat readable
- Store results as a database and retrieve for exploratory data analysis

Results - Foursquare

- **Both services provide basic information about local businesses (e.g. name, address, location)**
- **Foursquare has all types of businesses listed (e.g. tattoo parlours), as well as restaurants**
  - Top 5 most common restaurant types *(within 1km of Bloor GO Station), amount in brackets*:
    - Korean (8)
    - Italian (4)
    - Indian and Pakistani (3)
    - Ice Cream (3)
    - Desserts (3)

Results – Foursquare

- Top 5 most common businesses *(within 1km of Bloor GO Station); number in brackets*:
  - Bakery (5)
  - Bar (4)
  - Tattoo Parlour (3)
  - Park (2)
  - Laundry Services (2)

Results - Yelp

- **Yelp only provides data about restaurants (at least from my search), as well as a rating, in increments of 0.5**
  - Top 5 highest rated restaurants *(within 1km of Bloor GO Station)*:
    - La Bella Managua
    - Hodo Kwaja
    - Sugo
    - Famiglia Baldasarre
    - Ramen Isshin
      - (All had a rating of 4.5; although there were others)

Challenges

- Learning curve accessing and cleaning the data, and familiarization of software tools
  - About 5% of time was spent doing the Exploratory Data Analysis that was the focus of the mini-project
- Giving tables consistent properties so they can be joined and data from each API can be aggregated for a given business
  - ...which I finally got to, about 25 minutes before we began

Next Steps

- Connect Yelp's *Tips* API, which are mini-reviews about a business
  - (an example from a local brewery: 'Good brewery with nice sours and citrus forward IPAs. Friendly and helpful staff. Good space, but gets crowded.')
- Connect additional APIs to get more information (*e.g. Google Maps, Untappd)*
- Make more queries
  - Limit of 50 results per query
  - Slightly changing the start coordinates can produce completely unique results; certain parameters can also offset this (*i.e. the offset parameter)*
  - Making more queries will give more unique business results