Introduction

I used the street map data from my hometown Bellevue, Washington (https://www.openstreetmap.org/export#map=13/47.6065/-122.1647).

I sought to analyze the following:
- The number of food establishments
- The most proliferate establishment (fast food or resstaurant)
- The cuisines with the most number of establishments
- The number of zip codes contained in the map area
- The top contributors to the map

Problems Encountered in the Map

In analyzing the map, I encountered some problems:

- First, the road names lack consistency, which I had to clean in the following way:
  - Audit the map to inspect the titles of the roads
  - Create a dictionary with the titles mapped to a consistent road title, e.g. Ave. or ave -> "Avenue"
  - Write a function so that when I process the map to create a JSON file, I clean up these road names
- Second, I inspected the map tags to make sure the tags contained the write data or the right format that I would expect but I found that 5 tags contained "testing point" as its value, which was not what was expected
- Third, in analyzing the zip codes, some users inputted the values of multiple zip codes for elements that span multiple zip codes. This made the zip code field irregular and hence these lists as a value appeared as distinct from the zip codes contained in the list

Overview of the Data

Once in JSON format, the data becomes easier to read. Each document contains information about points on the map.
For instance, in the below document, we can see that this point represents a Burger King that user "goldfndr" created.
These documents then determine the points of interest or the information that goes on the map.


```
{
 "cuisine": "burger",
 "amenity": "fast_food",
 "name": "Burger King",
 "created": {
  "changeset": "814784",
  "user": "goldfndr",
  "version": "3",
  "uid": "7247",
  "timestamp": "2009-01-20T03:03:12Z"
 },
 "pos": [
  47.616762,
  -122.1833799
 ],
 "created_by": "Potlatch 0.10f",
 "source": "knowledge",
 "type": "node",
 "id": "333392511"
}
```


Analysis of the Data

I analyzed the map to look at the following information:

- The number of food establishments
- The most proliferate establishment (fast food or resstaurant)
- The cuisines with the most number of establishments
- The number of zip codes contained in the map area
- The top contributors to the map

My results are as follows:
- The number of food establishments
  67 Total fast food establishments
  246 Total restaurants
  313 Total number of food establishments

- The most proliferate establishment (fast food or resstaurant):
  'Subway', 'count': 10
  "McDonald's", 'count': 5
  'Red Robin', u'count': 4
  'Burger King', u'count': 3
  "Jimmy John's", u'count': 3}

- The cuisines with the most number of establishments:
  'burger', u'count': 18
  'mexican', u'count': 16
  'pizza', u'count': 14
  'sandwich', u'count': 13
  None, u'count': 148

- The number of zip codes contained in the map area:
  There are 15 zip codes:
    '98006',
    '98008',
    '98027',
    '98052',
    '98005',
    '98033',
    '98007',
    '98040',
    '98004',
    '98039',
    '98115',
    '98105',
    '98053',
    '98118',
    '98056',

- The top contributors to the map:
  'Glassman_Import', u'count': 55175
  'zephyr', u'count': 44502
  'sctrojan79-import', u'count': 32251
  'Glassman', u'count': 31432
  'STBrenden', u'count': 25690

Other Ideas About the Dataset

With more time and ability, I'd be curious to inspect the following:
- Information about the top contributors, e.g. are they knowledgeable about a certain area (e.g. zipcode 98006) or a specific type of information (e.g. highways or restaurants)
- Correlation between number of fast food establishments and schools
- The density of religious institutions and schools

With even more time, I would be interesting to compare some of these characteristics against the same characteristics in other map areas