



Big Data Overview: 311 Service Calls Incoming!



By Joey Tran, Garrett Tice, Brandon Ascencio,
Kyle Alfonso, and Michael Campos

Analyzing Service Calls in NYC Non-Emergencies

311 Service Requests from 2010 to Present

NOTE: This data does not present a full picture of 311 calls or service requests, in part because of ▸



🔍 Find in this Dataset

[More Views](#) [Filter](#) [Visualize](#) [Export](#) [Discuss](#) [Embed](#) [About](#)

Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor
42043549	03/24/2019 05:40:49 PM	03/25/2019 08:05:03 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42042821	03/24/2019 07:07:51 PM	03/25/2019 08:05:19 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42042652	03/24/2019 04:40:09 PM	03/25/2019 08:05:42 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42040531	03/24/2019 10:53:04 PM	03/25/2019 08:05:11 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42040351	03/24/2019 12:52:20 PM	03/24/2019 01:24:29 PM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42040316	03/24/2019 09:10:08 AM	03/24/2019 10:10:21 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42039548	03/24/2019 05:08:08 PM	03/25/2019 08:05:27 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42039439	03/24/2019 08:12:46 PM	03/25/2019 08:04:56 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42038939	03/24/2019 01:38:15 PM	03/24/2019 02:38:23 PM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42038935	03/24/2019 10:01:23 AM	03/24/2019 10:52:34 AM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	
42038696	03/24/2019 01:32:09 PM	03/24/2019 01:36:32 PM	DHS	Operations Unit - Department of Ho...	Homeless Person Assistance	

< Previous [Next](#) >

Showing 311 Service Requests 1 to 100 out of 24,507,90

Purpose and Benefit of 311 Service Calls

- 911 is for emergencies while 311 is for non-emergency service based calls. The types of complaints can vary and data.
- “Can be used in order to characterize the local context of the urban neighborhood and demonstrate moderate to strong correlations with wealth, education level, unemployment, racial structure of its population”(Wang).

Purpose of Big Data Research and Analysis

- The purpose of our analysis of 311 service calls is to survey the different types of non-emergency calls within the New York City Area between 2010 and 2020.
- Analyzing the data we want to discover patterns of most and least frequent types of calls in different boroughs based on most common complaints.

Proposed Analysis

- Determining most and least common service calls based on responding agency and the area of new york city (I.E. park borough manhattan)
- What is the Park Borough with most "pounding" and greatest loud noise complaints? Could this be considered more dangerous areas to live?

Proposed Analysis cont.

- Analyze the number of cases each agency responded to, and use the information in the reallocation of budgets.
- Analyzing timeframe 2010-2020 Complaint Types (looking for most common.)

Analytic Software

- Hadoop Hive - Querying data and creating tables
- Tableau - Visualizations and analysis



Specifications

- Cluster Version - 20.3.3-20
- Cluster Number of Nodes - 3
- Memory Size - 180 GB
- OCPUs - 12
- Storage - 957 GB



Data Source Urls

NYC OpenData

<https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9>

Data Size

Data size - 12.4 GB

Uploading the data set to the hdfs

- Downloaded the data(12.3 Gb)
- Compressed the data to a zip file (2.18 Gb)
- Uploaded zip file to an Amazon S3 bucket
- Used the S3 bucket to download to /dev/shm/ directory
- Unzipped the file, and made a /Services directory in the hdfs
- Moved the unzipped file to the Services Directory

Tables - Main Database Table

Data File Format: CSV

Table Name: Services

Total columns/attributes: 41

Data types: date, int, string

Number of rows(tentative): 24,515,013

```
0: jdbc:hive2://bigdai-nov-bdcsce-1:2181,bigdai> DESC services;
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| unique_key | string | |
| created_date | date | |
| closed_date | date | |
| agency | string | |
| agency_name | string | |
| complaint_type | string | |
| descriptor | string | |
| location_type | string | |
| incident_zip | string | |
| incident_address | string | |
| street_name | string | |
| cross_street_1 | string | |
| cross_street_2 | string | |
| intersection_street_1 | string | |
| intersection_street_2 | string | |
| address_type | string | |
| city | string | |
| landmark | string | |
| facility_type | string | |
| status | string | |
| due_date | date | |
| resolution_description | string | |
| resolution_action_updated_date | date | |
| community_board | string | |
| bbl | string | |
| borough | string | |
| x_coordinate_state_plane | int | |
| y_coordinate_state_plane | int | |
| open_data_channel_type | string | |
| park_facility_name | string | |
| park_borough | string | |
| vehicle_type | string | |
| taxi_company_borough | string | |
| taxi_pick_up_location | string | |
| bridge_highway_name | string | |
| bridge_highway_direction | string | |
| road_ramp | string | |
| bridge_highway_segment | string | |
| latitude | double | |
| longitude | double | |
| location | string | |
+-----+-----+-----+
41 rows selected (0.229 seconds)
```

Queries Run to test the database

- `Select * from Complaints limit 50;`
- `Select * from Complaints where agency_name LIKE 'New York City Police Department' limit 10;`
- `Select DISTINCT * from Services Limit 100;`
- `Select complaint_type, numberofcomplaints from Complaints where agency_name LIKE 'New York City Police Department' limit 10;`

Tables - Complaints Schema

- Agency_name - City govt agency responding to call
- Complaint_Type - Topic of incident or condition
- Numberofcomplaints - total number of complaints of each complaint type

```
0: jdbc:hive2://bigdai-nov-bdcsce-1:2181,bigd> DESC complaints;
+-----+-----+-----+
| col_name | data_type | comment |
+-----+-----+-----+
| agency_name | string | |
| complaint_type | string | |
| numberofcomplaints | int | |
+-----+-----+-----+
3 rows selected (0.371 seconds)
```

Tables - Complaints top 10 complaint types

Key takeaways

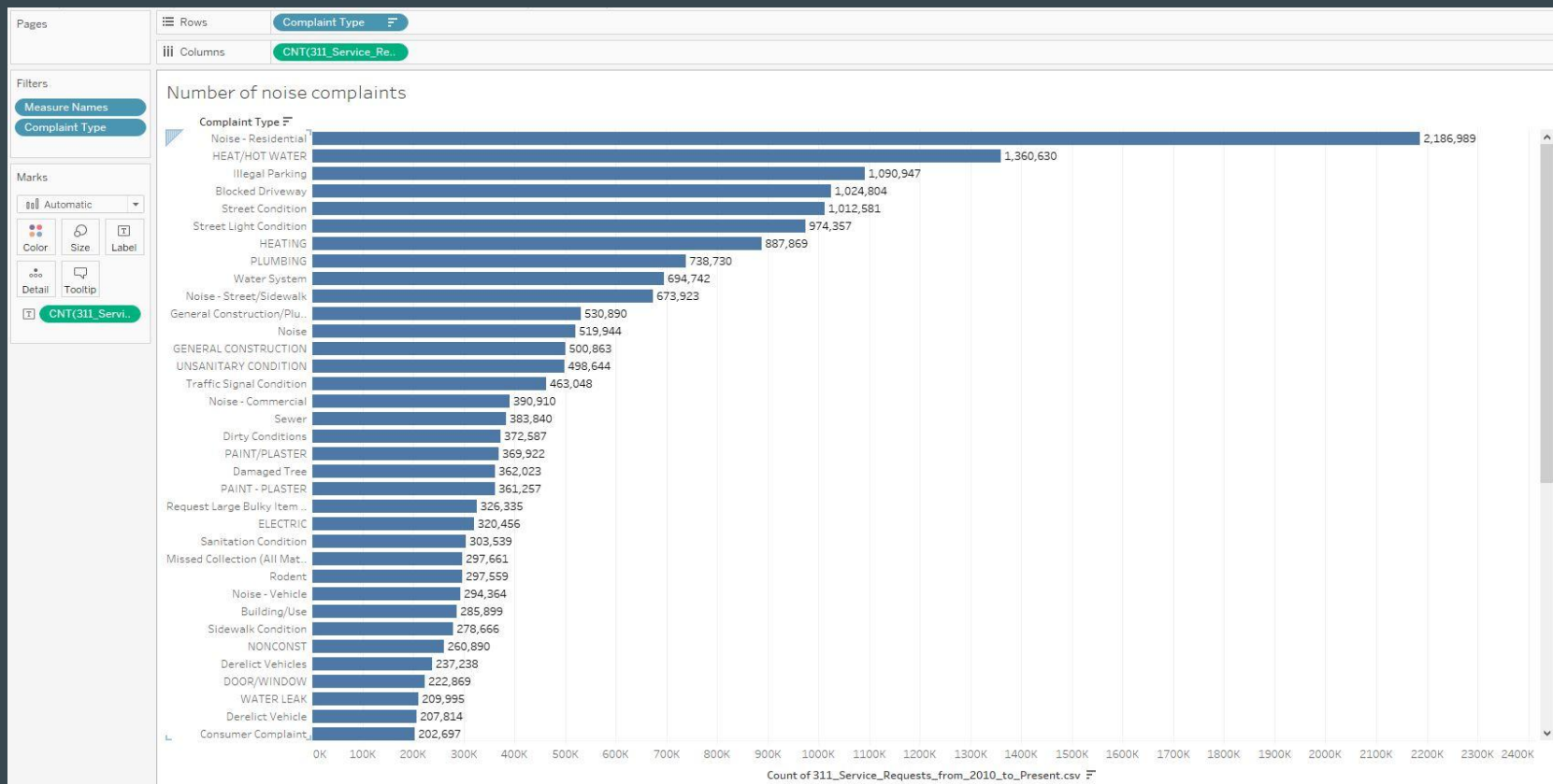
- Residential noise complaints - 2,189,115
- NYCPD appears 4 times

```
ERROR: Error while compiling statement: FAILED: RuntimeException: cannot make an array of type 'emp/nyc/1524250' out of type '0: jdbc:hive2://bigdai-nov-bdcscs-1:2181,bigd> select * from Complaints limit 10;
```

complaints.agency_name	complaints.complaint_type	complaints.numberofcomplaints
New York City Police Department	Noise - Residential	2189115
Department of Housing Preservation and Development	HEAT/HOT WATER	1362946
New York City Police Department	Illegal Parking	1091003
New York City Police Department	Blocked Driveway	1025839
Department of Transportation	Street Condition	1012885
Department of Transportation	Street Light Condition	974690
Department of Housing Preservation and Development	HEATING	887850
Department of Housing Preservation and Development	PLUMBING	739135
Department of Environmental Protection	Water System	694961
New York City Police Department	Noise - Street/Sidewalk	674609

```
10 rows selected (0.102 seconds)
```

Visualization of Results of 1



Number of Complaints handled by each Agency

- NYC-PD responded to the most complaints
- Housing Preservation was the second highest
- Both are about double the third highest

agency_name	totalcomplaints
New York City Police Department	6463414
Department of Housing Preservation and Development	6309625
Department of Transportation	3249152
Department of Environmental Protection	1868305
Department of Buildings	1250272
Department of Parks and Recreation	1118923
Department of Health and Mental Hygiene	615149
Department of Sanitation	563970
Taxi and Limousine Commission	271158
Department of Consumer Affairs	229061

10 rows selected (5.524 seconds)

Number of complaint types per Agency

- Health and mental hygiene responds to the most types of complaints
- NYPD was second despite responding to the most calls

agency_name	complaints_handled
Department of Health and Mental Hygiene	40
New York City Police Department	32
Department of Buildings	30
Department of Transportation	29
Department of Housing Preservation and Development	27
Department of Environmental Protection	18
Department for the Aging	18
Manhattan 02	16
BCC - Brooklyn South	16
BCC - Bronx	16

10 rows selected (5.525 seconds)

Total complaints and types by Agency

Which agency handles which complaint



CNT(311_Service_Req...

1 2,186,924

Tables - City Schema

City - city the complaint call was taken in

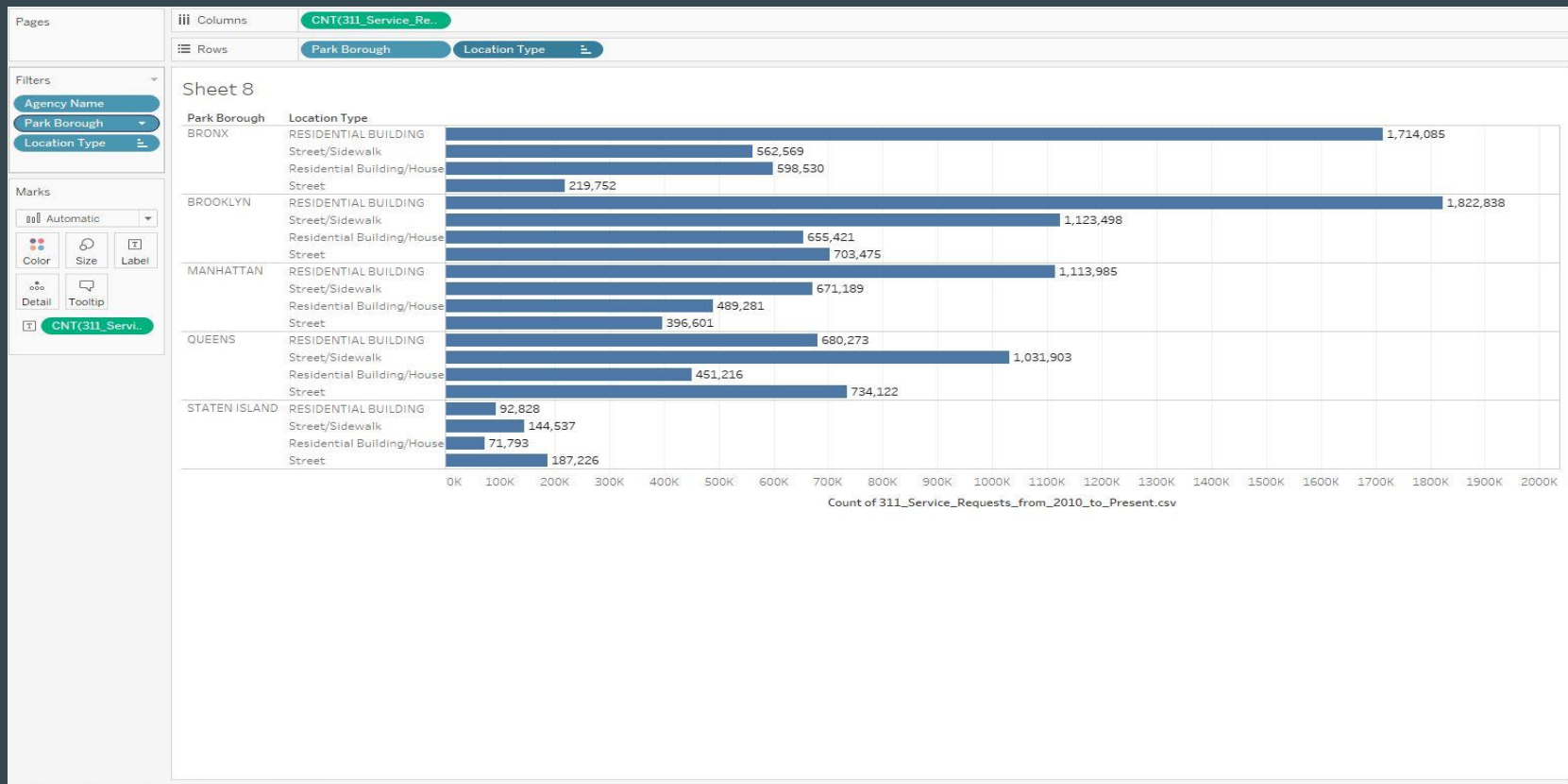
Complaint_Type - Topic of incident or condition

number - total # of complaints

city	complaint_type	number
BROOKLYN	Noise - Residential	626961
BRONX	Noise - Residential	587451
NEW YORK	Noise - Residential	468613
BRONX	HEAT/HOT WATER	444856
	Street Light Condition	413055
BROOKLYN	HEAT/HOT WATER	410288
BROOKLYN	Illegal Parking	384957
BROOKLYN	Blocked Driveway	375005
NEW YORK	HEAT/HOT WATER	307021
BROOKLYN	HEATING	280356

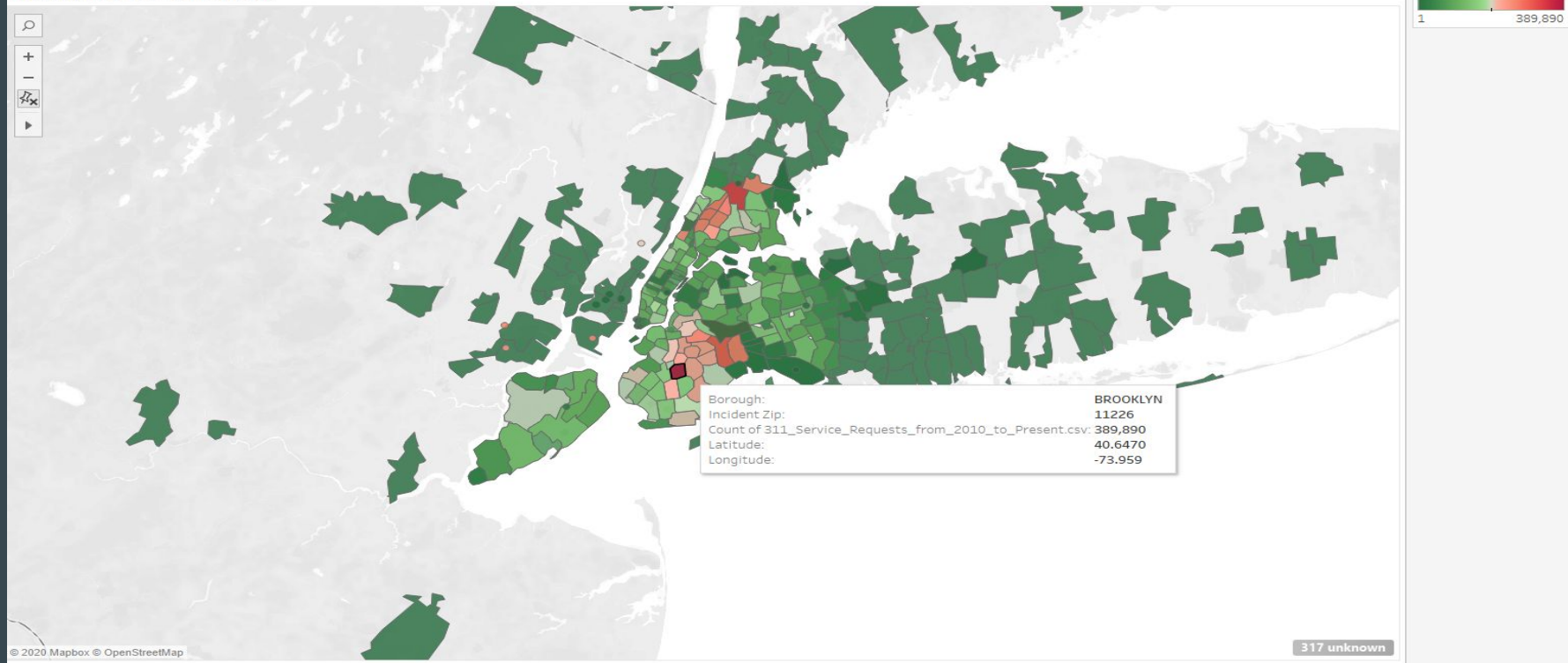
10 rows selected (86.168 seconds)

Residential Type by City

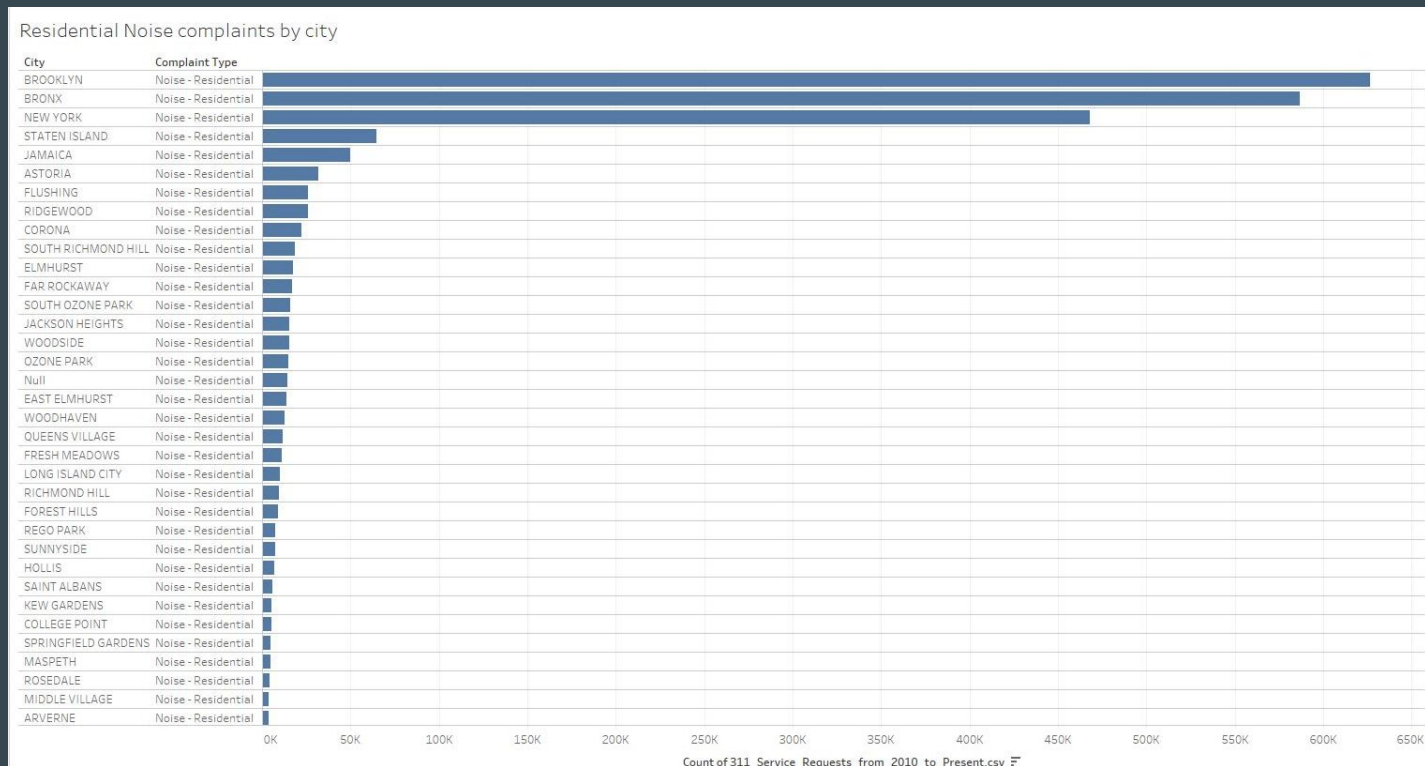


Heatmap of complaints by City

Heat map of 311 calls by zip



Residential Noise Complaints by City



Banging/Pounding Descriptor

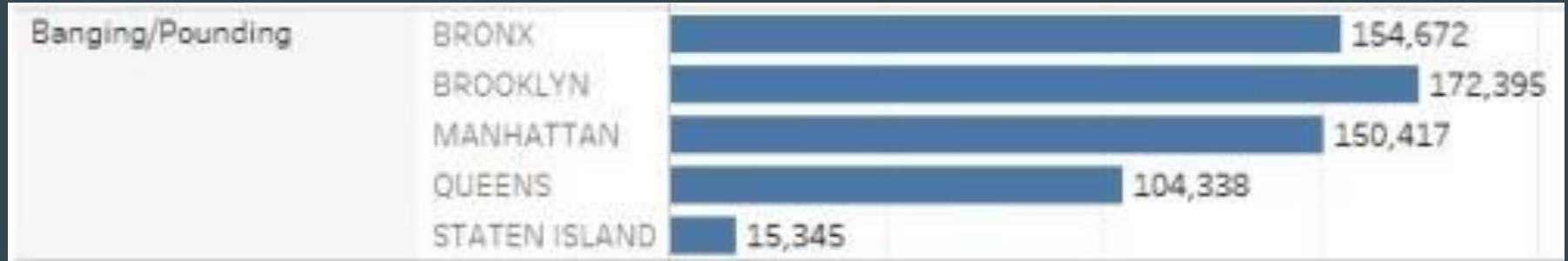
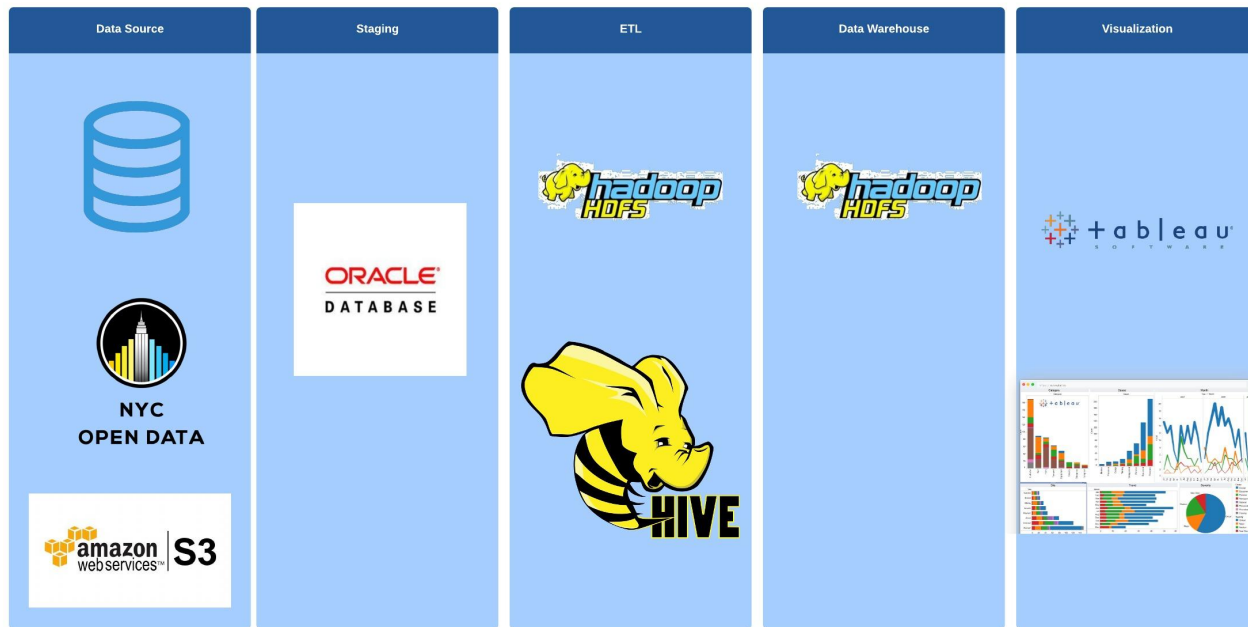


Figure 2. Boroughs with most common noise/related complaints (Pounding/Banging).

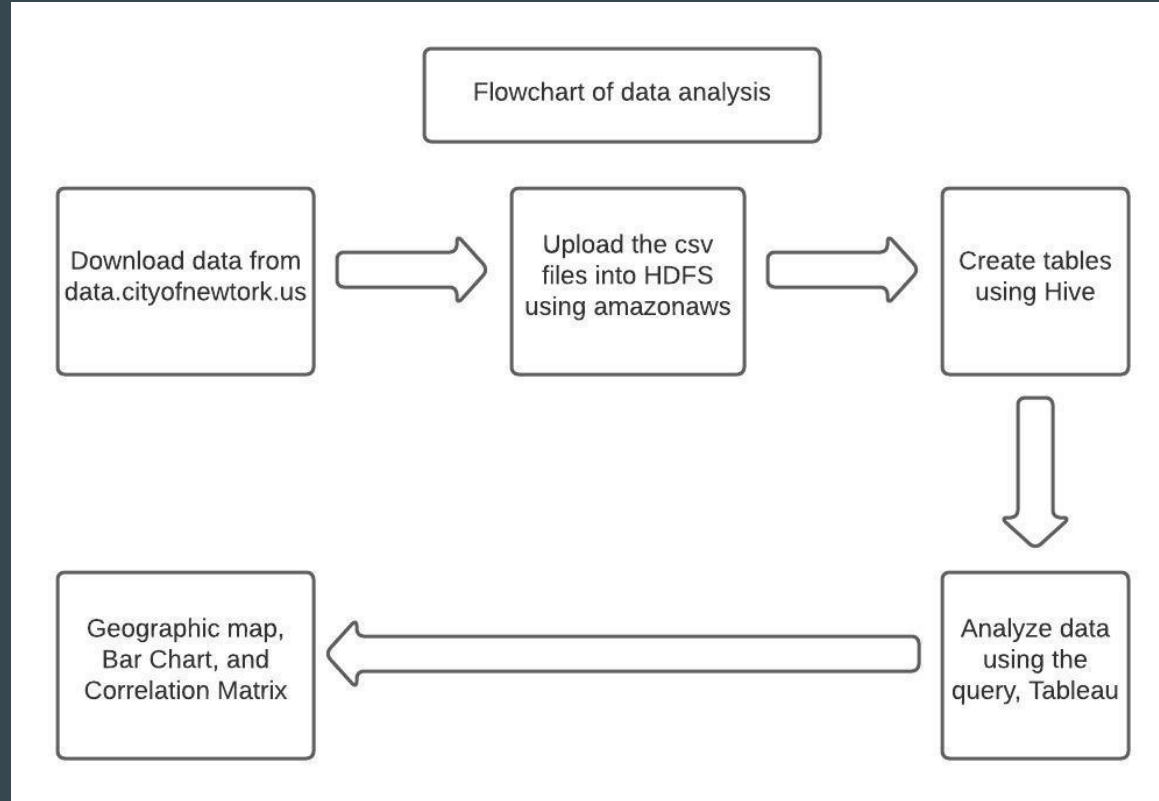
Architecture

Workflow Chart

Campos, Michael A | December 6, 2020



FlowChart



Findings

- Residential noise complaints was the most common complaint type
- Areas with the highest residential buildings have the highest complaints
- Brooklyn and Bronx had higher numbers of residential buildings resulting in the highest noise complaints

Findings Cont.

- NYPD and Housing Preservation were the agencies that handled the highest total of complaints received
- Together they account for just over half of the complaints/records received in the data set

GitHub Link



- <https://github.com/branasce/CIS4560Project>

Approval form the Professor



Woo, Jongwook

Mon 11/30/2020 11:28 AM



To: Campos, Michael A

Great. I look forward to seeing your presentation.

JW

