

Diagrams of data flows representing ETL, Profiling and Analytics

Anuj Bora & Gowtham Goli

Data Source 1 : 311 Service Request Data set

ETL of 311 Service request Dataset

...	Violation Date	Agency	Complaint Type	Borough
...	3/29/2016	NYPD	Street/Side walk	Queens
...	3/25/2016	DOHMH	Food Poisoning	Brooklyn
...	3/10/2016	DOT	Street Condition	Manhattan
...	2/18/2016	DOHMH	Rodent	Manhattan
...	2/10/2016	NYPD	Street/Side walk	Brooklyn
...	2/7/2016	DOHMH	Indoor Air Quality	Unspecified
...	2/5/2016	DOHMH	Rodent	Bronnx

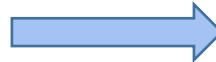


Date	Agency	Complaint Type	Borough
3/25/2016	DOHMH	Food Poisoning	Brooklyn
2/18/2016	DOHMH	Rodent	Manhattan
2/7/2016	DOHMH	Indoor Air Quality	Unspecified
2/5/2016	DOHMH	Rodent	Bronnx

Profiling of 311 Service request Dataset

Date	Agency	Complaint Type	Borough
3/25/2016	DOHMH	Food Poisoning	Brooklyn
2/18/2016	DOHMH	Rodent	Manhattan
2/7/2016	DOHMH	Indoor Air Quality	Unspecified
2/5/2016	DOHMH	Rodent	Bronnx

PROFILE



Date	Agency	Complaint Type	Borough
3/25/2016	DOHMH	Food Poisoning	Brooklyn
2/18/2016	DOHMH	Rodent	Manhattan
2/5/2016	DOHMH	Rodent	Bronx

Analytics of 311 Service request Dataset

Date	Agency	Complaint Type	Borough
3/25/2016	DOHMH	Food Poisoning	Brooklyn
2/18/2016	DOHMH	Rodent	Manhattan
2/5/2016	DOHMH	Rodent	Bronx

ANALYTICS



- In March 2016, the primary health concerns were mostly caused due to
 - Food poisoning = 1
 - Indoor air quality = 0
 - Rodents = 0
- In February 2016, the primary health concerns were mostly caused due to
 - Food poisoning = 0
 - Indoor air quality = 0
 - Rodents = 2

Data Source 2 : Water Quality Complaints Data set

ETL of Water Quality complaints Dataset

...	Created Date	Borough
...	3/29/2016	Queens
...	3/25/2016	Brooklyn
...	3/10/2016	Manhattan
...	2/18/2016	Manhattan
...	2/10/2016	Brooklyn
...	2/7/2016	Unspecified
...	2/5/2016	Bronnx

ETL


Created Date	Borough
3/29/2016	Queens
3/25/2016	Brooklyn
3/10/2016	Manhattan
2/18/2016	Manhattan
2/10/2016	Brooklyn
2/7/2016	Unspecified
2/5/2016	Bronnx

Profiling of Water Quality complaints Dataset

Created Date	Borough
3/29/2016	Queens
3/25/2016	Brooklyn
3/10/2016	Manhattan
2/18/2016	Manhattan
2/10/2016	Brooklyn
2/7/2016	Unspecified
2/5/2016	Bronnx

PROFILE



Created Date	Borough
3/29/2016	Queens
3/25/2016	Brooklyn
3/10/2016	Manhattan
2/18/2016	Manhattan
2/10/2016	Brooklyn
2/5/2016	Bronx

Analytics of Water Quality complaints Dataset

Created Date	Borough
3/29/2016	Queens
3/25/2016	Brooklyn
3/10/2016	Manhattan
2/18/2016	Manhattan
2/10/2016	Brooklyn
2/5/2016	Bronx

ANALYTICS



- In March 2016, the total number of water quality complaints are 3, distributed as follows
 - Queens = 1
 - Manhattan = 1
 - Brooklyn = 1
- In February 2016, the total number of water quality complaints are 3, distributed as follows
 - Manhattan = 1
 - Brooklyn = 1
 - Bronx = 1

Data Source 3 : OATH ECB Hearings Case Status Data set

ETL of OATH ECB Hearings Case Status Dataset

...	Violation Date	Issuing Agency	Violation Location (Borough)
...	3/29/2016	DOS - ENFORCEMENT AGENTS	Manhattan
...	3/25/2016	SANITATION RECYCLING	Brooklyn
...	3/10/2016	DOS - ENFORCEMENT AGENTS	Brooklyn
...	2/18/2016	SANITATION RECYCLING	Manhattan
...	2/10/2016	DOS - ENFORCEMENT AGENTS	Queens
...	2/7/2016	SANITATION RECYCLING	Bronx
...	2/5/2016	SANITATION RECYCLING	Staten Is

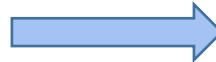
ETL


Violation Date	Issuing Agency	Violation Location (Borough)
3/29/2016	DOS - ENFORCEMENT AGENTS	Manhattan
3/25/2016	SANITATION RECYCLING	Brooklyn
3/10/2016	DOS - ENFORCEMENT AGENTS	Brooklyn
2/18/2016	SANITATION RECYCLING	Manhattan
2/10/2016	DOS - ENFORCEMENT AGENTS	Queens
2/7/2016	SANITATION RECYCLING	Bronx
2/5/2016	SANITATION RECYCLING	Staten Is

Profiling of OATH ECB Hearings Case Status Dataset

Violation Date	Issuing Agency	Violation Location (Borough)
3/29/2016	DOS - ENFORCEMENT AGENTS	Manhattan
3/25/2016	SANITATION RECYCLING	Brooklyn
3/10/2016	DOS - ENFORCEMENT AGENTS	Brooklyn
2/18/2016	SANITATION RECYCLING	Manhattan
2/10/2016	DOS - ENFORCEMENT AGENTS	Queens
2/7/2016	SANITATION RECYCLING	Bronx
2/5/2016	SANITATION RECYCLING	Staten Is

PROFILE



Violation Date	Issuing Agency	Violation Location (Borough)
3/25/2016	SANITATION RECYCLING	Brooklyn
2/18/2016	SANITATION RECYCLING	Manhattan
2/7/2016	SANITATION RECYCLING	Bronx
2/5/2016	SANITATION RECYCLING	Staten Is

Analytics of OATH ECB Hearings Case Status Dataset

Violation Date	Issuing Agency	Violation Location (Borough)
3/25/2016	SANITATION RECYCLING	Brooklyn
2/18/2016	SANITATION RECYCLING	Manhattan
2/7/2016	SANITATION RECYCLING	Bronx
2/5/2016	SANITATION RECYCLING	Staten Is

Analytics



- In March 2016,
 - Number of Sanitation Recycling Complaints = 1
- In February 2016,
 - Number of Sanitation Recycling Complaints = 3

Data Source 4 : Air Quality Data set

ETL of Air Quality Dataset

...	geo entity name	year description	data valuemessage
...	Manhattan	2005	126.1
...	Queens	2005	17.5
...	Greenpoint	2006	56.3
...	Bronx	2009	136.1
...	Brooklyn	2006	171.2
...	Jamaica	2009	224.8.



geo entity name	year description	data valuemessage
Manhattan	2005	126.1
Queens	2005	17.5
Greenpoint	2006	56.3
Bronx	2009	136.1
Brooklyn	2006	171.2
Jamaica	2009	224.8.

Profiling of Air Quality Dataset

geo entity name	year description	data valuemessage
Manhattan	2005	126.1
Queens	2005	17.5
Greenpoint	2006	56.3
Bronx	2009	136.1
Brooklyn	2006	171.2
Jamaica	2009	224.8.

PROFILE



geo entity name	year description	data valuemessage
Manhattan	2005	126.1
Queens	2005	17.5
Bronx	2009	136.1
Brooklyn	2006	171.2

Analytics of Air Quality Dataset

geo entity name	year description	data valuemessage
Manhattan	2005	126.1
Queens	2005	17.5
Bronx	2009	136.1
Brooklyn	2006	171.2

ANALYTICS



- In 2005,
 - Air Quality in Manhattan : Unhealthy for sensitive groups
 - Air Quality in Queens : *Good*
- In 2006,
 - Air Quality in Brooklyn : Unhealthy
- In 2009,
 - Air Quality in Bronx : Unhealthy for sensitive groups