Boston 311 Service Request
GCP Data Warehousing Project

Introduction

The Boston 311 Service Request system provides a centralized platform for handling non-emergency city service requests. This project focuses on preparing raw 311 request data from Boston for advanced analytics by building a robust ETL (Extract, Transform, Load) pipeline. The goal is to process historical and real-time service request data efficiently, enabling seamless reporting and analysis.

The pipeline processes millions of records and leverages **Apache Beam** for distributed data processing, **Google Cloud Dataflow** for execution, and **BigQuery** as the **Data Warehouse**. This ensures structured, high-performance analytics capabilities.

Links

Dataset: Boston 311 | Boston.gov

GitHub for Previous Implementation on Local: amey379/311_Request_Analysis

GitHub for GCP: amey379/GCPDataWarehousingProject

Tools

| Storage:	Development & System:
BigQuery – Data Warehouse for structured storage Google Cloud Storage (GCS) – Staging and intermediate data storage	*Apache Avro & Parquet – Data serialization and storage formats* *Java – Programming language for ETL processing* *GCP VM (Unix) – Environment setup & scripting*
Orchestration & Workflow Management:	**Visualization & Reporting:**
Google DataFlow – ETL pipelines and Orchestration *Apache Beam – ETL pipelines and Orchestration*	*Power BI – Dashboarding & data visualization*

Skills

Dimensional Modeling & Data Warehousing

BigQuery Optimization & Performance Tuning

ETL Development (Extract, Transform, Load)

Apache Beam & Google Dataflow

Dimensional Model
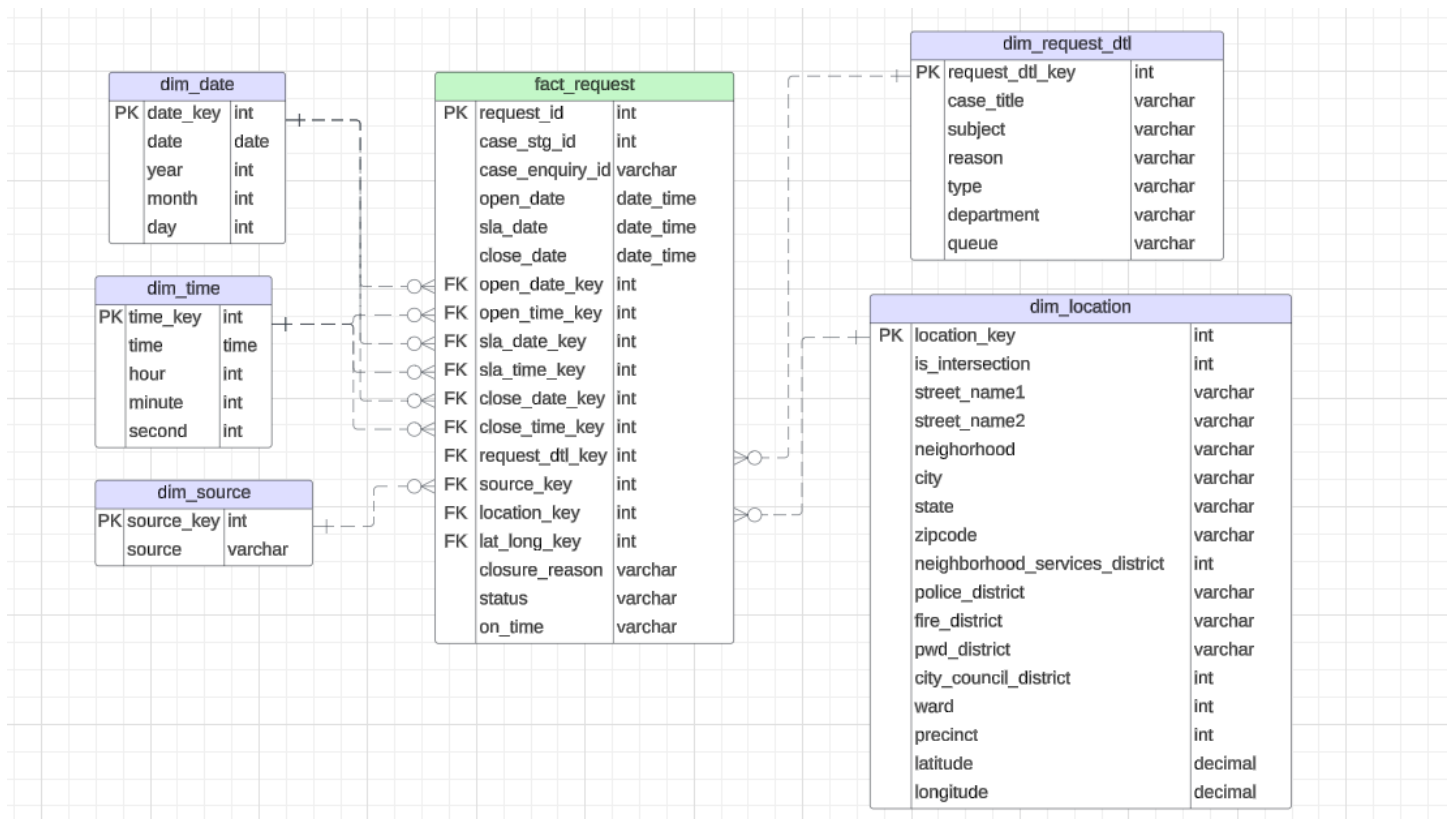
Fact Table:

- **Fact_311_Requests** – Stores key metrics like resolution time, case status, request type

Dimension Tables:

- **Dim_Date Dim_Time** – Date and time details for trend analysis

- **Dim_Location** – Geographic information of incidents

- **Dim_Request_Details** – Categories and types of 311 requests

- **Dim_Source** – Channel through which requests were received

Files:

Technical Implementation:

Data Ingestion Layer (Bronze Layer - Raw Data Storage)

- Source: Boston 311 Service Requests (Parquet in GCS)

- Bronze Layer: Store raw CSV files in Google Cloud Storage (GCS) without modifications

- Extract raw 311 data from GCS (Bronze Layer) and load it into Apache Beam (Dataflow) for processing

Data Processing & Transformation Layer (Silver Layer - Cleaned Data)

- Silver Layer: Apply cleaning, deduplication, and standardization in Apache Beam

- ETL Pipelines (Apache Beam - Dataflow) process and enrich data

- Handle missing values, data type inconsistencies, and data standardization

- Apply deduplication and ensure incremental processing

- Stored Data in Parquet Files.

Data Storage & Query Layer (Gold Layer - BigQuery Data Warehouse)

- Gold Layer: Load cleaned and structured data into BigQuery for analytics

- Convert raw data into structured dimension tables (dim_date, dim_time, dim_location, dim_source, dim_request_dtl)

- Optimize table schema with indexing and partitioning for query performance

- Store Fact Table (fact_311_requests) to support analytical queries

Performance Optimization

- Used **Apache Avro & Parquet** for optimized storage.

- Implemented **batch processing** for efficient large-scale data ingestion.

- Applied **partitioning and clustering** in BigQuery for improved query performance.

Output:

Stage_load

📌 **my-311-service-data-lake**

Location	Storage class	Public access	Protection
us-central1 (Iowa)	Standard	Subject to object ACLs	Soft Delete

OBJECTS　　CONFIGURATION　　PERMISSIONS　　PROTECTION　　LIFECYCLE　　OBSERVABILITY　　INVENTORY REPORTS　　OPERATIONS

Folder browser I<

Buckets > my-311-service-data-lake > silver 📄

CREATE FOLDER　　UPLOAD ▾　　TRANSFER DATA ▾　　OTHER SERVICES ▾

- 📦 my-311-service-data-lake ⋮
  - 📁 bronze/ ⋮
  - 📁 gold/ ⋮
  - 📁 silver/ ⋮
  - ▶ 📁 staging/ ⋮
  - ▶ 📁 temp/ ⋮

Filter by name prefix only ▾　　≡ Filter　Filter objects and folders　　　　Show Live objects only ▾

	Name	Size	Type	Created ❓	Storage class	
☐	📄 nul	0 B	application/octet-stream	Mar 15, 2025, 4:35:25 PM	Standard	⬇ ⋮
☐	📄 output-00000-of-00001.parquet	141.8 KB		Mar 16, 2025, 8:29:45 PM	Standard	⬇ ⋮
☐	📄 stage_boston_311_requests-0000...	8.4 MB		Mar 17, 2025, 3:37:15 AM	Standard	⬇ ⋮
☐	📄 stage_boston_311_requests-0000...	5.6 MB		Mar 17, 2025, 3:37:15 AM	Standard	⬇ ⋮
☐	📄 stage_boston_311_requests-0000...	12.6 MB		Mar 17, 2025, 3:37:15 AM	Standard	⬇ ⋮

Dim_date

▦ **dim_date**　　🔍 QUERY　　OPEN IN ▾　　+👤 SHARE　　📋 COPY　　⊞ SNAP

< 　SCHEMA　　DETAILS　　**PREVIEW**　　TABLE EXPLORER [PREVIEW]　　INSI

Row	date_key	date	year	month	day
1	20150101	2015-01-01	2015	1	1
2	20150102	2015-01-02	2015	1	2
3	20150103	2015-01-03	2015	1	3
4	20150104	2015-01-04	2015	1	4
5	20150105	2015-01-05	2015	1	5
6	20150106	2015-01-06	2015	1	6
7	20150107	2015-01-07	2015	1	7
8	20150108	2015-01-08	2015	1	8
9	20150109	2015-01-09	2015	1	9
10	20150110	2015-01-10	2015	1	10
11	20150111	2015-01-11	2015	1	11
12	20150112	2015-01-12	2015	1	12
13	20150113	2015-01-13	2015	1	13
14	20150114	2015-01-14	2015	1	14
15	20150115	2015-01-15	2015	1	15
16	20150116	2015-01-16	2015	1	16
17	20150117	2015-01-17	2015	1	17
18	20150118	2015-01-18	2015	1	18

Dim_time

⊞ **dim_time**	🔍 QUERY	OPEN IN ▾	👥 SHARE	📋 COPY	⊞ SNAPS

‹	SCHEMA	DETAILS	**PREVIEW**	TABLE EXPLORER PREVIEW	INSIGH

Row	time_key	time	hour	minute	second
1	0	00:00:00	0	0	0
2	1	00:00:01	0	0	1
3	2	00:00:02	0	0	2
4	3	00:00:03	0	0	3
5	4	00:00:04	0	0	4
6	5	00:00:05	0	0	5
7	6	00:00:06	0	0	6
8	7	00:00:07	0	0	7
9	8	00:00:08	0	0	8
10	9	00:00:09	0	0	9
11	10	00:00:10	0	0	10
12	11	00:00:11	0	0	11
13	12	00:00:12	0	0	12
14	13	00:00:13	0	0	13
15	14	00:00:14	0	0	14
16	15	00:00:15	0	0	15
17	16	00:00:16	0	0	16
18	17	00:00:17	0	0	17

Dim_location

⊞ **dim_location**	🔍 QUERY	OPEN IN ▾	👥 SHARE	📋 COPY	⊞ SNAPSHOT	🗑 DELETE	⬆ EXPORT

‹	SCHEMA	DETAILS	**PREVIEW**	TABLE EXPLORER PREVIEW	INSIGHTS	LINEAGE	DATA PROFILE	D

Row	location_key	is_intersection	street_name1	street_name2	neighborhood	city
1	-1	0	UNKNOWN	*null*	Not Available	Boston
2	772	1	1 CITY HALL PLZ	null	BOSTON	Boston
3	9	1	1 BEACON ST	null	BOSTON	Boston
4	240	1	8 BEACON ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
5	543	1	1P CHARLES ST	null	BOSTON	Boston
6	788	1	1B CHARLES ST	null	BOSTON	Boston
7	699	1	139 TREMONT ST	null	BOSTON	Boston
8	206	1	71 MOUNT VERNON ST	null	BEACON HILL	Boston
9	897	1	125-127 FULTON ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
10	711	1	130-132 FULTON ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
11	469	1	20 CLINTON ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
12	846	1	61-63 CHARTER ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
13	482	1	529P COMMERCIAL ST	null	BOSTON	Boston
14	251	1	88 BLACKSTONE ST	null	BOSTON	Boston
15	632	1	21 MILK ST	null	DOWNTOWN / FINANCIAL DIS…	Boston
16	933	1	1P POST OFFICE SQ	null	BOSTON	Boston
17	190	1	106-112 BEACH ST	null	DOWNTOWN / FINANCIAL DIS…	Boston

Dim_source

▦ **dim_source** 🔍 QUERY OPEN IN ▾

< SCHEMA DETAILS **PREVIEW** TAB

Row	source_key	source
1	-1	Not Available
2	1	CITIZENS CONNECT APP
3	2	CITY WORKER APP
4	3	CONSTITUENT CALL
5	4	SELF SERVICE
6	5	EMPLOYEE GENERATED

Dim_request_dtl

▦ **dim_request_dtl** 🔍 QUERY OPEN IN ▾ ＋⚊ SHARE ▭ COPY ⊞ SNAPSHOT 🗑 DELETE ⬆ ↻

< SCHEMA DETAILS **PREVIEW** TABLE EXPLORER `PREVIEW` INSIGHTS LINEAGE DATA PROFILE DATA

Row	request_dtl_key	case_title	subject	reason	type
1	79	ABANDONED BICYCLE	MAYOR'S 24 HOUR HOTLINE	ABANDONED BICYCLE	ABANDONED BICYCLE
2	123	GENERAL REQUEST	BOSTON WATER & SEWER CO...	ADMINISTRATIVE & GENERAL ...	BWSC GENERAL REQUEST
3	116	PARKS GENERAL REQUEST	PARKS & RECREATION DEPART...	ADMINISTRATIVE & GENERAL ...	PARKS GENERAL REQUEST
4	18	PUBLIC WORKS GENERAL REQ...	PUBLIC WORKS DEPARTMENT	ADMINISTRATIVE & GENERAL ...	PUBLIC WORKS GENERAL REQ...
5	85	PUBLIC WORKS GENERAL REQ...	PUBLIC WORKS DEPARTMENT	ADMINISTRATIVE & GENERAL ...	PUBLIC WORKS GENERAL REQ...
6	151	PUBLIC WORKS GENERAL REQ...	PUBLIC WORKS DEPARTMENT	ADMINISTRATIVE & GENERAL ...	PUBLIC WORKS GENERAL REQ...
7	228	PUBLIC WORKS GENERAL REQ...	PUBLIC WORKS DEPARTMENT	ADMINISTRATIVE & GENERAL ...	PUBLIC WORKS GENERAL REQ...
8	59	TRANSPORTATION GENERAL R...	TRANSPORTATION - TRAFFIC ...	ADMINISTRATIVE & GENERAL ...	TRANSPORTATION GENERAL R...
9	181	ANIMAL GENERIC REQUEST	ANIMAL CONTROL	ANIMAL ISSUES	ANIMAL GENERIC REQUEST
10	235	BUILDING INSPECTION REQUE...	INSPECTIONAL SERVICES	BUILDING	BUILDING INSPECTION REQUE...
11	127	CONTRACTORS COMPLAINT	INSPECTIONAL SERVICES	BUILDING	CONTRACTORS COMPLAINT
12	189	ELECTRICAL	INSPECTIONAL SERVICES	BUILDING	ELECTRICAL
13	198	ILLEGAL OCCUPANCY	INSPECTIONAL SERVICES	BUILDING	ILLEGAL OCCUPANCY
14	201	ILLEGAL ROOMING HOUSE	INSPECTIONAL SERVICES	BUILDING	ILLEGAL ROOMING HOUSE
15	193	UNSAFE/DANGEROUS CONDIT...	INSPECTIONAL SERVICES	BUILDING	UNSAFE DANGEROUS CONDITI...
16	104	WORK W/OUT PERMIT	INSPECTIONAL SERVICES	BUILDING	WORK W/OUT PERMIT
17	3	WORKING BEYOND HOURS	INSPECTIONAL SERVICES	BUILDING	WORKING BEYOND HOURS

Fact

	fact_request	🔍 QUERY	OPEN IN ▾	👥 SHARE	📋 COPY	⊞ SNAPSHOT	🗑 DELETE	⬆ EXPORT

SCHEMA	DETAILS	PREVIEW	TABLE EXPLORER PREVIEW	INSIGHTS	LINEAGE	DATA PROFILE

Row	request_id	ca	case_enquiry_id	open_date	sla_date	close_date
1	728		101005384842	1970-01-01	1970-01-01T07:05:51	null
2	875		101005654998	1970-01-01	1970-01-01T06:28:39	1970-01-01T07:32:29
3	521		101005633434	1970-01-01	1970-01-01T11:22:41	null
4	493		101005671525	1970-01-01	1970-01-01T09:34:56	1970-01-01T10:24:18
5	732		101005378921	1970-01-01	1970-01-01T09:49:56	null
6	432		101005351458	1970-01-01	1970-01-01T10:34:23	null
7	169		101005571512	1970-01-01	1970-01-01T11:44:20	1970-01-01T04:06:54
8	553		101005624340	1970-01-01	1970-01-01T15:56:54	1970-01-01T05:07:01

Data Visualization

Year 2022

277K
No of Requests

17.33%
Overdue Request %

13
Avg_Resolution_Time (Days)

82.67%
SLA Compliance Rate

YoY Analysis 2022-2023

neighborhood
All

13.22%
Yearly Requests Change %

Top 5 Reasons remains same
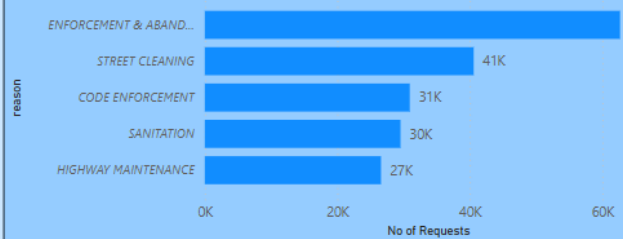
Year 2023

313K
No of Requests

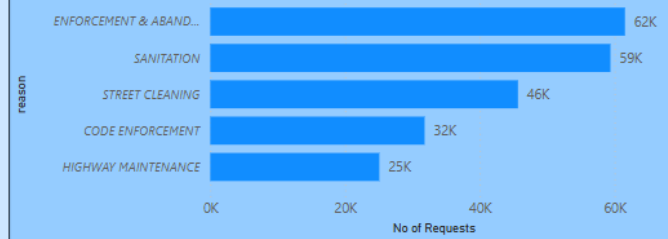**18.76%**
Overdue Request %

7
Avg_Resolution_Time (Days)

81.24%
SLA Compliance Rate

Top 5 Reasons (2022)

reason	No of Requests
ENFORCEMENT & ABAND...	63K
STREET CLEANING	41K
CODE ENFORCEMENT	31K
SANITATION	30K
HIGHWAY MAINTENANCE	27K

Top 5 Reasons (2023)

reason	No of Requests
ENFORCEMENT & ABAND...	62K
SANITATION	59K
STREET CLEANING	46K
CODE ENFORCEMENT	32K
HIGHWAY MAINTENANCE	25K

Day Period Analysis

Select all	Afternoon	Early Morning	Evening	Late Night	Morning	Night

Year: 2016 — 2023

Neighborhood: All

No of Requests by time (bins) and DayPeriod

DayPeriod ● Afternoon ● Early Morning ● Evening ● Late Night ● Morning ● Night

Values shown: 8K, 24K, 69K, 136K, 196K, 168K, 141K, 113K, 92K, 72K, 56K, 42K, 13K, 8K, 6K
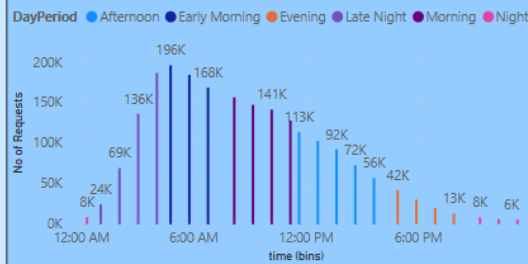
No of Requests by DayPeriod

105K (4.98%)
572K (27.17%)
415K (19.73%)
436K (20.73%)
548K (26.03%)

DayPeriod
● Morning
● Early Morning
● Afternoon
● Late Night
● Evening
● Night
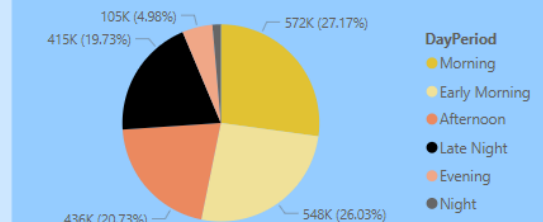
After Sunrise

DayPeriod	Rank	reason	TotalRequests
Afternoon	1	ENFORCEMENT & ABANDONED VEHICLES	126527
Afternoon	2	SANITATION	73478
Afternoon	3	STREET CLEANING	67936
Early Morning	1	STREET CLEANING	98724
Early Morning	2	SANITATION	92334
Early Morning	3	ENFORCEMENT & ABANDONED VEHICLES	91737
Morning	1	SANITATION	109636
Morning	2	ENFORCEMENT & ABANDONED VEHICLES	103626
Morning	3	STREET CLEANING	94973

"Street Cleaning" and "Enforcement & Abandoned Vehicles" remain prominent throughout the day.

Reason "Street Lights" is in top 3 during Evening and Night times.

After Sunset

DayPeriod	Rank	reason	TotalRequests
Evening	1	ENFORCEMENT & ABANDONED VEHICLES	42788
Evening	2	STREET CLEANING	14584
Evening	3	STREET LIGHTS	12586
Late Night	1	STREET CLEANING	96977
Late Night	2	ENFORCEMENT & ABANDONED VEHICLES	91231
Late Night	3	CODE ENFORCEMENT	54865
Night	1	STREET CLEANING	9272
Night	2	ENFORCEMENT & ABANDONED VEHICLES	8410
Night	3	STREET LIGHTS	4642

Weekday v/s Weekend

18.66%
Weekday Overdue Request %

16.42%
Weekend Overdue Request %

31
Weekday Avg_Resolution_Time (Days)

23
Weekend Avg_Resolution_Time (Days)

81.34%
Weekday SLA Compliance Rate

83.57%
Weekend SLA Compliance Rate

Year
2016   2023

Resolution time on Weekend > Weekday.

No of Requests by Weekday_Weekend

0M
(16.91%)

2M
(83.09%)

Weekday_Weekend
- Weekday
- Weekend

Avg Req Per day by Year and Weekday_Weekend

Weekday_Weeke... • Weekday • Weekend

Avg Req Per day

1000 / 800 / 600 / 400 / 200

Year: 2016, 2018, 2020, 2022

Top 5 Reasons on Weekend

No of Requ... 100K / 50K / 0K

ENFORCEMENT & ABANDONED VEHICLES	STREET CLEANING	CODE ENFORCEMENT	HIGHWAY MAINTENANCE	SANITATION
78K	62K	47K	30K	26K

Top 5 Reasons on Weekday

No of Reque... 0.4M / 0.2M / 0.0M

ENFORCEMENT & ABANDONED VEHICLES	SANITATION	STREET CLEANING	CODE ENFORCEMENT	HIGHWAY MAINTENANCE
0.31M	0.24M	0.22M	0.17M	0.17M

Geomapping

Year
2016   2023

Neighborhood
All

Average of latitude, Average of longitude and No of Requests by zipcode

Future Scope

- Delta Lakehouse Implementation
- Enhanced Airflow Orchestration
- Microservices & Real-Time Processing