TransitPulse KPI & Incident Monitoring System

Team 13: Malay Patel, Shikha Rani, Anisha Bahl, Nikhil Nagesh

Background: Audience & Data Sources

Who are our users?

Our system serves diverse stakeholders across WMATA's operational ecosystem:

WMATA executives — Strategic decision-makers requiring high-level insights

Safety and security analysts — Teams monitoring crime patterns and passenger safety

Rail operations & maintenance management — Staff optimizing service delivery

Urban mobility & policy researchers — Professionals studying transit trends

Data sources used

We integrate multiple datasets to create a comprehensive view of system performance:

- Metro station boarding records
- System-wide mode ridership data
- Crime incident database
- Service incident logs
- On-time performance datasets

Mission Statement

> To deliver a unified analytics dataset and model that links ridership, reliability, and crime activity to empower WMATA decisions on safety and operational performance.

By combining disparate data sources into a single analytical framework, we enable evidence-based decision-making that improves passenger experience, enhances safety protocols, and optimizes operational efficiency across the entire WMATA network.

Mission Objectives

We aim to answer critical questions:

Ridership Patterns

Which stations have the highest and lowest boardings over time?

Service Reliability

Which lines show the worst on-time performance?

Incident Clusters

Where do service incidents cluster and why?

Crime Analysis

Which crime types are most frequent and at which stations?

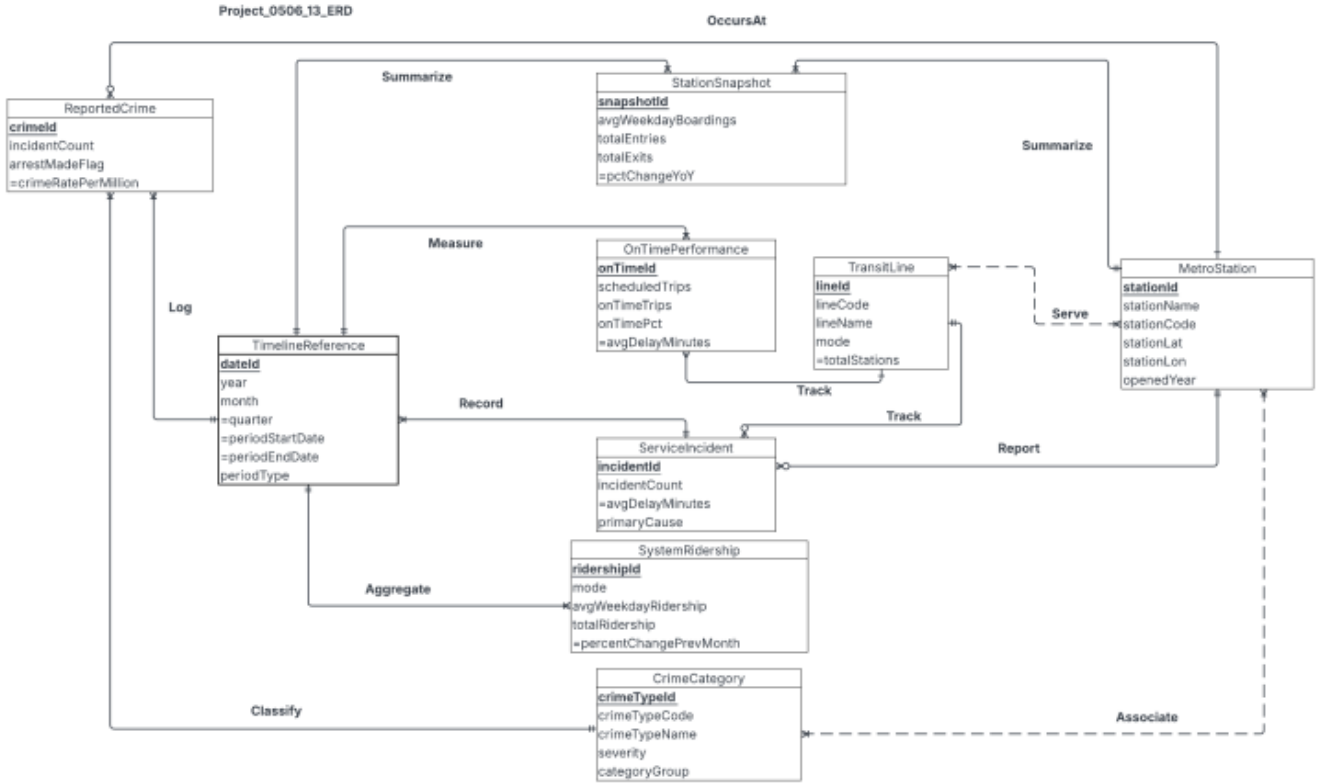
Risk Assessment

Which stations require intervention when combining ridership and crime activity?

Correlation Study

Is there correlation between ridership and crime frequency?

Conceptual Database Design (ER Diagram)



Logical Database Design (Relational Schema)

TimelineReference(**dateId**, year, month, periodType)

MetroStation(**stationId**, stationName, stationCode, stationLat, stationLon, openedYear)

TransitLine(**lineId**, lineCode, lineName, mode)

CrimeCategory(**crimeTypeId**, crimeTypeCode, crimeTypeName, severity, categoryGroup)

StationSnapshot(**snapshotId**, avgWeekdayBoardings, totalEntries, totalExits, *dateId, stationId*)

SystemRidership(**ridershipId**, mode, avgWeekdayRidership, totalRidership, *dateId*)

OnTimePerformance(**onTimeId**, scheduledTrips, onTimeTrips, onTimePct, *dateId, lineId*)

ServiceIncident(**incidentId**, incidentCount, primaryCause, *dateId, stationId, lineId*)

ReportedCrime(**crimeId**, incidentCount, arrestMadeFlag, *dateId, stationId, crimeTypeId*)

StationLine(**stationLineId**, *stationId*, *lineId*, openedYear)

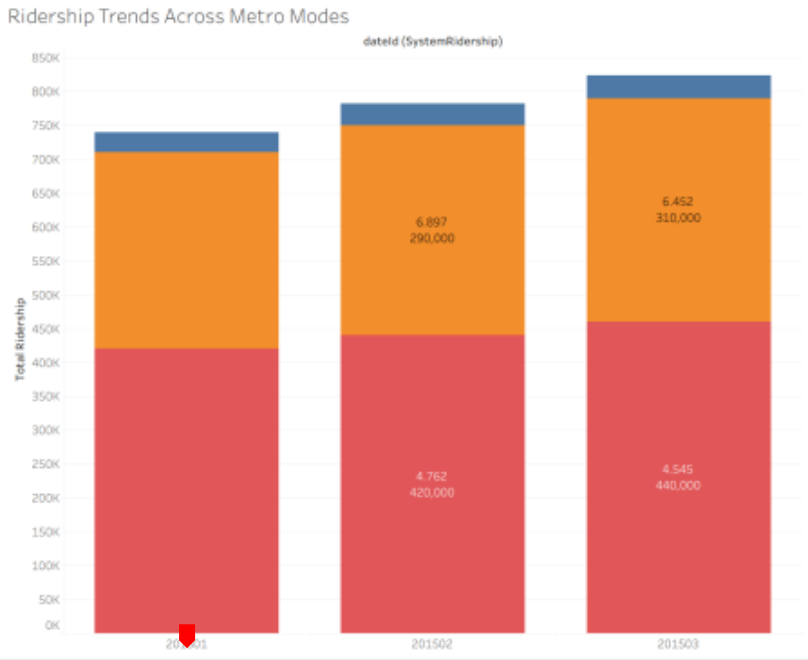
StationCrime(**stationCrimeId**, *stationId*, *crimeTypeId*, reportedYear, incidentCount)

Physical Design: Example SQL CREATE TABLE

```
CREATE TABLE Team13.StationSnapshot(
 snapshotId CHAR(10) PRIMARY KEY,
 dateId CHAR(8) NOT NULL REFERENCES Team13.TimelineReference(dateId),
 stationId CHAR(8) NOT NULL REFERENCES Team13.MetroStation(stationId),
 avgWeekdayBoardings INT,
 totalEntries INT,
 totalExits INT
);
GO
```

Business Transaction 4

Query: How has WMATA Ridership changed month-to-month across Metrorail, Metrobus, MetroAccess?



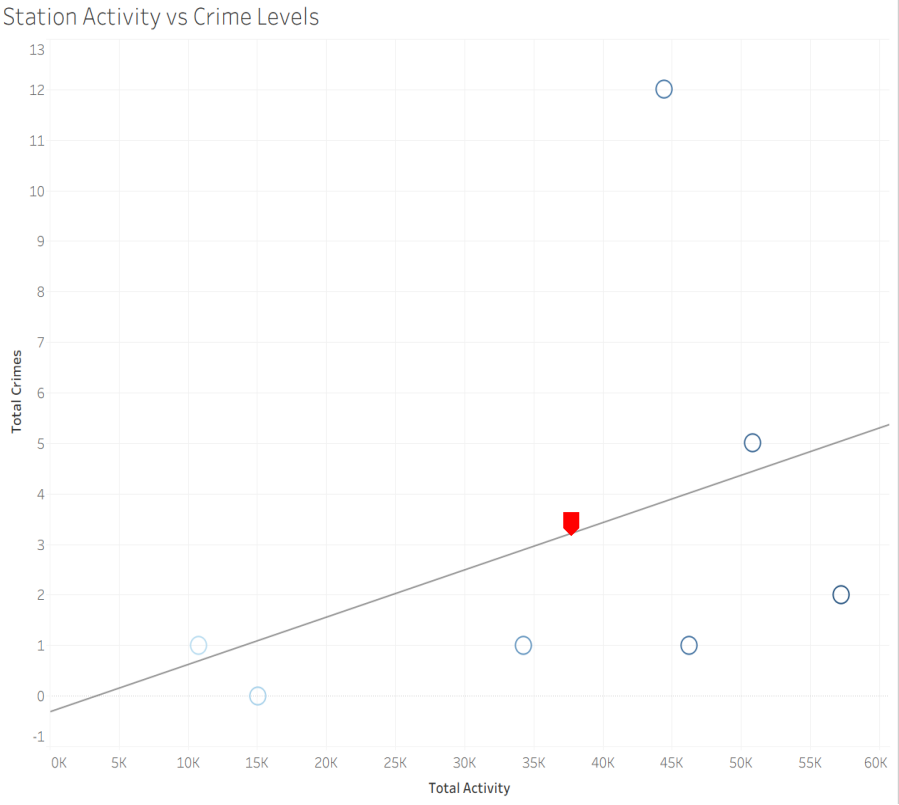

```sql
--4. How has WMATA ridership changed month-to-month across Metrorail, Metrobus, and MetroAccess?
SELECT
    R.mode,
    R.dateId,
    R.totalRidership,
    LAG(R.totalRidership) OVER (PARTITION BY R.mode ORDER BY R.dateId) AS prevMonthRidership,
    ROUND(
        100.0 *
        (R.totalRidership - LAG(R.totalRidership) OVER (PARTITION BY R.mode ORDER BY R.dateId))
        / NULLIF(LAG(R.totalRidership) OVER (PARTITION BY R.mode ORDER BY R.dateId), 0),
        2
    ) AS percentChangePrevMonth
FROM Team13.SystemRidership R
ORDER BY R.mode, R.dateId;
```

Results | Messages

	mode	dateId	totalRidership	prevMonthRidership	percentChangePrevMonth
1	MetroAccess	201501	30000	NULL	NULL
2	MetroAccess	201502	32000	30000	6.670000000000
3	MetroAccess	201503	34000	32000	6.250000000000
4	Metrobus	201501	290000	NULL	NULL
5	Metrobus	201502	310000	290000	6.900000000000
6	Metrobus	201503	330000	310000	6.450000000000
7	Metrorail	201501	420000	NULL	NULL
8	Metrorail	201502	440000	420000	4.760000000000
9	Metrorail	201503	460000	440000	4.550000000000

Business Transaction 6

Query: How does station activity(entries/exits) relate to crime levels across WMATA stations?




```sql
-- 6. How does station activity (entries/exits) relate to crime levels across WMATA stations?
SELECT
    SS.stationId,
    ST.stationName,
    SS.totalEntries + SS.totalExits AS totalActivity,
    COALESCE(SUM(SC.incidentCount), 0) AS totalCrimes
FROM Team13.StationSnapshot SS
JOIN Team13.MetroStation ST
     ON SS.stationId = ST.stationId
LEFT JOIN Team13.StationCrime SC
     ON SS.stationId = SC.stationId
GROUP BY
    SS.stationId,
    ST.stationName,
    SS.totalEntries,
    SS.totalExits
ORDER BY totalCrimes DESC;
```

	stationId	stationName	totalActivity	totalCrimes
1	ST16	Gallery Pl-Chinatown	44458	12
2	ST15	Metro Center	50858	5
3	ST18	Union Station	57268	2
4	ST19	Rhode Island Ave-Brentwood	10766	1
5	ST13	Dupont Circle	34264	1
6	ST14	Farragut North	46250	1
7	ST17	Judiciary Square	15046	0
8	ST20	Brookland-CUA	11566	0
9	ST21	Fort Totten	16594	0

Key Insights and Conclusion

Ridership fluctuations correlate with operational incidents

Analysis reveals a direct relationship between passenger volume changes and the frequency of service disruptions.

High-traffic stations report proportionally higher crime rates

Stations with increased foot traffic are often hotbeds for various types of criminal activity, requiring focused security measures.

Certain lines experience recurring causes of delays

Specific metro lines consistently suffer from particular operational issues, indicating a need for targeted maintenance and infrastructure upgrades.

Metrics provide targeted intervention recommendations

The system generates clear, data-backed suggestions for optimizing resource allocation, improving safety, and enhancing overall service reliability.

Future Road Map: Enhancing TransitPulse

Our vision for TransitPulse extends beyond current capabilities, focusing on integrating advanced data, real-time analytics, and predictive intelligence to further optimize transit operations and public safety.

Extended Data Horizon

Unlock deeper insights into long-term trends, seasonal patterns, and historical performance by integrating a decade of operational data. Add 10 year data.

Real-time Operational Insights

Incorporate live data feeds from sensors and operational systems to provide immediate updates on incidents, ridership, and performance.

AI-Powered Predictive Analytics

Implement machine learning models to forecast crime hotspots, predict service delays, and anticipate maintenance needs with greater accuracy.

Dynamic Timetable Optimization

Develop algorithms to automatically adjust timetables based on predicted disruptions, ridership demand, and incident likelihood for smoother operations.


Q&A