BUDT703_Project_0506_Design

Team 13 Members:

Anisha Bahl
Shikha Rani
Nikhil Nagesh
Malay Patel

Project Name:

TransitPulse KPI & Incident Monitoring System

Mission Statements:

To deliver an integrated analytics dataset and data model that combines WMATA ridership, operational performance, and crime data so WMATA decision makers can Identify station and line level priorities for service reliability and passenger safety.

Mission Objectives:

1. Which stations show the highest and lowest average weekday boardings over time, and how have these changed year over year?

2. Which lines have the worst on-time performance by month, and when did the largest drops occur?

3. Where and when do service incidents concentrate (by line and station), and what are the leading causes?

4. Which crime types and which stations show the highest incident rates when normalized by ridership (crime per million riders)?

5. Which stations are the highest priority for intervention when we combine ridership, incident counts, and crime rates (a composite intervention score)?

6. Is there a measurable correlation between ridership volume and incident/crime frequency at the station level (time lagged if needed)?

Business Terms:

- **MetroStation** - A WMATA rail station; unique stationId.

- **TransitLine** - A WMATA transit line (e.g., Red Line); unique lineId.

- **TimeLineReference** - A reporting period (monthly grain for this project) with dateId.
- **CrimeCategory** - A crime classification in the Ten-Year Crime Report; crimeTypeId.
- **StationSnapshot** - Station-level monthly snapshot (avg weekday boardings, entries, exits).

- **SystemRidership** - Monthly system mode ridership totals.

- **OnTimePerformance** - Monthly on time performance metrics by line.

- **ServiceIncident** - Monthly counts or aggregated incidents/delays attributed to line/station.

- **ReportedCrime** - Monthly counts of crimes at station by crime type.
- **StationLine** - (physical) bridge table linking stations and lines.
- **StationCrime** - (physical) bridge table linking stations and crime types for summary/metadata.

Facts, Attributes, Identifiers, & Cardinalities:

Entity: TimelineReference

- **Business Fact:** Time anchor that relates every monthly metric (ridership, performance, incidents, crimes).
- **Identifier (Primary Key):** dateId
- **Attributes:** year, month, quarter *(derived)*, periodStartDate *(derived)*, periodEndDate *(derived)*, periodType
- **Cardinalities (1:N):**

    Date (1) → Snapshot (N)

    Date (1) → Ridership (N)

    Date (1) → OnTime (N)

Date (1) → Incident (N)

Date (1) → Crime (N)

Entity: <u>MetroStation</u>

- **Business Fact:** Physical metro location that links ridership, incidents, and crimes.
- **Identifier (Primary Key):** stationId
- **Attributes:** stationCode, stationName, stationLat, stationLon, openedYear
  Cardinalities (1:N and M:N):

  Station (1) → Snapshot (N)

  Station (1) → Incident (N)

  Station (1) → Crime (N)

  Station (N) ↔ Line (N) M:N mapping stored in **StationLine** table

  Station (N) ↔ CrimeType (N) M:N mapping stored in **StationCrime** table

Entity: <u>TransitLine</u>

- **Business Fact:** Transit route used to relate on-time performance and incidents.
- **Identifier (Primary Key):** lineId
- **Attributes:** lineCode, lineName, mode, totalStations *(derived)*
- **Cardinalities (1:N and M:N):**

  Line (1) → OnTime (N)

  Line (1) → Incident (N)

  Line (N) ↔ Station (N) M:N mapping stored in **StationLine** table

Entity: <u>CrimeCategory</u>

- **Business Fact:** Classification that groups crime events and relates to stations for analysis.
- **Identifier (Primary Key):** crimeTypeId
- **Attributes:** crimeTypeCode, crimeTypeName, severity, categoryGroup

- **Cardinalities (1:N and M:N):**

    CrimeType (1) → Crime (N)

    CrimeType (N) ↔ Station (N) M:N mapping stored in **StationCrime**

Entity: <u>StationSnapshot</u>

- **Business Fact:** Monthly station snapshot that connects a Date to a Station for ridership activity.
- **Identifier (Primary Key):** snapshotId
- **Attributes:** avgWeekdayBoardings, totalEntries, totalExits, dateId, stationId
- **Cardinalities (1:N):**

    Date (1) → Snapshot (N)

    Station (1) → Snapshot (N)

Entity: <u>SystemRidership</u>

- **Business Fact:** Monthly system ridership by mode connected to Date.
- **Identifier (Primary Key):** ridershipId
- **Attributes:** mode, avgWeekdayRidership, totalRidership, percentChangePrevMonth *(derived)*, dateId
- **Cardinalities (1:N):**

    Date (1) → Ridership (N)

Entity: <u>OnTimePerformance</u>

- **Business Fact:** Monthly on-time performance that relates Date and Line.
- **Identifier (Primary Key):** onTimeId
- **Attributes:** scheduledTrips, onTimeTrips, onTimePct, avgDelayMinutes*(derived)*, dateId, lineId

- **Cardinalities (1:N):**

    Date (1) → OnTime (N)

    Line (1) → OnTime (N)

Entity: <u>ServiceIncident</u>

- **Business Fact:** Monthly service incidents linking Date and Line, and optionally Station.
- **Identifier (Primary Key):** incidentId
- **Attributes:** incidentCount, avgDelayMinutes*(derived)*, primaryCause, dateId, lineId, stationId
- **Cardinalities (1:N):**

    Date (1) → Incident (N)

    Line (1) → Incident (N)

    Station (1) → Incident (N)

Entity: <u>ReportedCrime</u>

- **Business Fact:** Monthly reported crimes linking Date, CrimeType, and optionally Station.
- **Identifier (Primary Key):** crimeId
- **Attributes:** incidentCount, arrestMadeFlag, crimeRatePerMillion *(derived)*, dateId, stationId , crimeTypeId
- **Cardinalities (1:N):**

    Date (1) → Crime (N)

    CrimeType (1) → Crime (N)

    Station (1) → Crime (N)

Entity: <u>StationLine</u> (bridge table)

- **Business Fact:** Physical implementation of the Station ↔ Line many-to-many.
- **Identifier (Primary Key):** stationLineId

- **Attributes:** stationId, lineId, openedYear
- **Cardinalities (1:N):**

  Station (1) → StationLine (N)

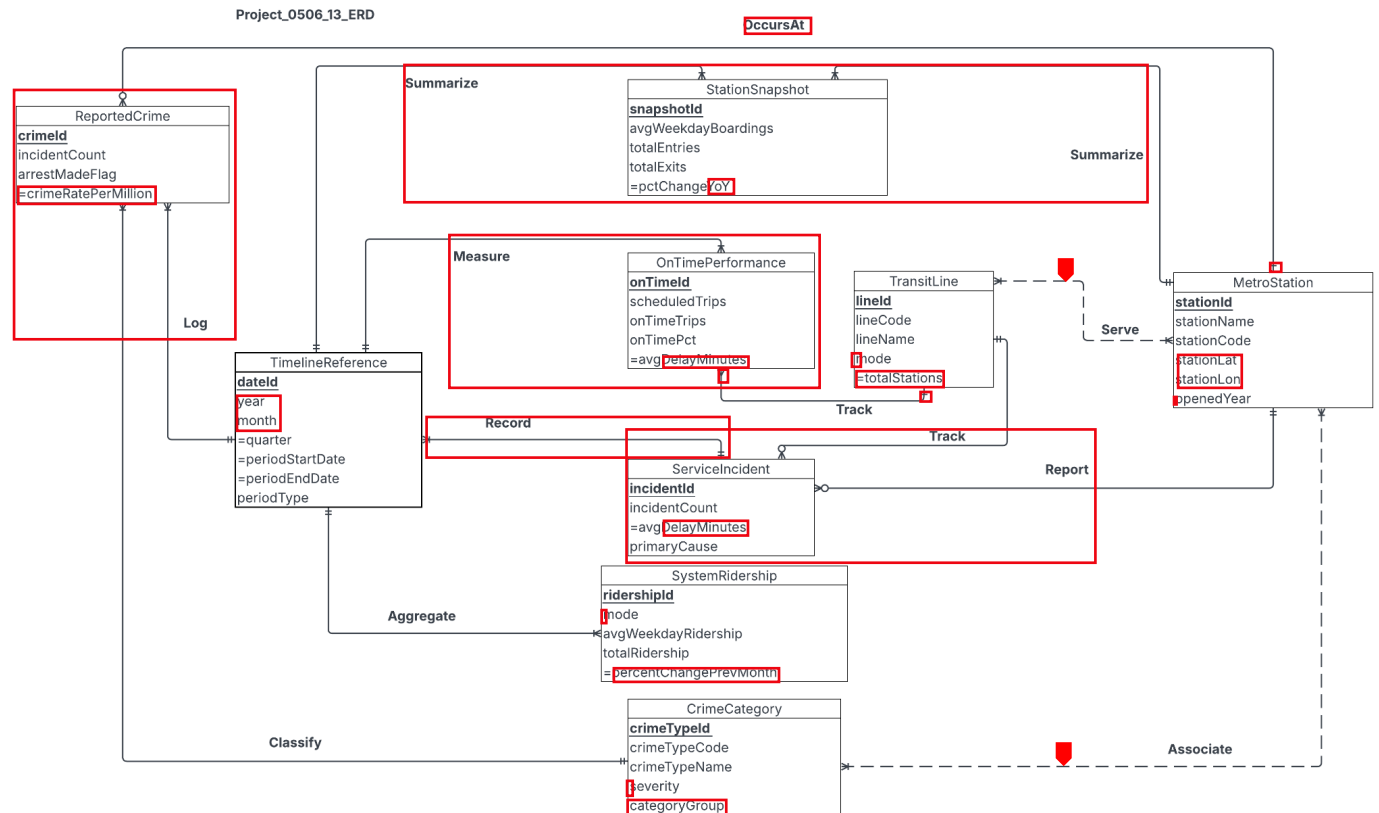 Line (1) → StationLine (N)

Entity: <u>StationCrime</u> (bridge table)

- **Business Fact:** Physical implementation of the Station ↔ CrimeType many-to-many

  (yearly summaries).
- **Identifier (Primary Key):** stationCrimeId
- **Attributes:** stationId, crimeTypeId, reportedYear, incidentCount
- **Cardinalities (1:N):**

  Station (1) → StationCrime (N)

  CrimeType (1) → StationCrime (N)

<u>Entity-Relationship Diagram:</u>

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)

Business Rules:

R1] When a TimelineReference row is deleted, it cannot be deleted if any StationSnapshot records reference it.

[R2] When TimelineReference information is changed, related StationSnapshot records cannot update automatically; the update is restricted because the dateId is referenced by foreign keys.

[R3] When a MetroStation row is deleted, it cannot be deleted if any StationSnapshot records reference it.

[R4] When MetroStation information is changed, related StationSnapshot records cannot update automatically because the stationId is referenced by foreign keys.

[R5] When a TimelineReference row is deleted, it cannot be deleted if any SystemRidership records reference it.

[R6] When TimelineReference information is changed, related SystemRidership records cannot update automatically due to foreign key restrictions.

[R7] When a TimelineReference row is deleted, it cannot be deleted if any OnTimePerformance records reference it.

[R8] When TimelineReference information is changed, related OnTimePerformance records cannot update automatically because SQL Server restricts FK key updates.

[R9] When a TransitLine row is deleted, it cannot be deleted if any OnTimePerformance records reference it.

[R10] When TransitLine information is changed, related OnTimePerformance records cannot update automatically due to FK constraints.

[R11] When a TimelineReference row is deleted, it cannot be deleted if any ServiceIncident records reference it.

[R12] When TimelineReference information is changed, related ServiceIncident records cannot update automatically because the referenced dateId cannot change.

[R13] When a TransitLine row is deleted, it cannot be deleted if any ServiceIncident records reference it.

[R14] When TransitLine information is changed, related ServiceIncident records cannot update automatically due to FK restrictions.

[R15] When a MetroStation row is deleted, it cannot be deleted if any ServiceIncident records

reference it; SQL Server prevents the delete to protect historical data.

[R16] When MetroStation information is changed, related ServiceIncident records cannot update automatically because FK key updates are restricted.

[R17] When a TimelineReference row is deleted, it cannot be deleted if any ReportedCrime records reference it.

[R18] When TimelineReference information is changed, related ReportedCrime records cannot update automatically.

[R19] When a MetroStation row is deleted, it cannot be deleted if any ReportedCrime records reference it.

[R20] When MetroStation information is changed, related ReportedCrime records cannot update automatically due to FK restrictions.

[R21] When a CrimeCategory row is deleted, it cannot be deleted if any ReportedCrime records reference it.

[R22] When CrimeCategory information is changed, related ReportedCrime records cannot update automatically.

[R23] When a MetroStation row is deleted, it cannot be deleted if any StationLine records reference it.

[R24] When MetroStation information is changed, related StationLine records cannot update automatically because FK updates are restricted.

[R25] When a TransitLine row is deleted, it cannot be deleted if any StationLine records reference it.

[R26] When TransitLine information is changed, related StationLine records cannot update automatically due to FK restrictions.

[R27] When a MetroStation row is deleted, it cannot be deleted if any StationCrime records reference it.

[R28] When MetroStation information is changed, related StationCrime records cannot update automatically.

[R29] When a CrimeCategory row is deleted, it cannot be deleted if any StationCrime records reference it.

[R30] When CrimeCategory information is changed, related StationCrime records cannot update automatically.

Referential Integrity:

Relation	Foreign Key	Base Relation	Primary Key	Business Rule	ON DELETE	Business Rule	ON UPDATE
StationSnapshot	dateId	TimelineReference	dateId	R1	NO ACTION	R2	CASCADE
StationSnapshot	stationId	MetroStation	stationId	R3	NO ACTION	R4	CASCADE
SystemRidership	dateId	TimelineReference	dateId	R5	NO ACTION	R6	CASCADE
OnTimePerformance	dateId	TimelineReference	dateId	R7	NO ACTION	R8	CASCADE
OnTimePerformance	lineId	TransitLine	lineId	R9	NO ACTION	R10	CASCADE
ServiceIncident	dateId	TimelineReference	dateId	R11	NO ACTION	R12	CASCADE
ServiceIncident	lineId	TransitLine	lineId	R13	NO ACTION	R14	CASCADE
ServiceIncident	stationId	MetroStation	stationId	R15	SET NULL	R16	CASCADE
CrimeCategory	dateId	TimelineReference	dateId	R17	NO ACTION	R18	CASCADE
CrimeCategory	stationId	MetroStation	stationId	R19	SET NULL	R20	CASCADE
CrimeCategory	crimeTypeId	CrimeCategory	crimeTypeId	R21	NO ACTION	R22	CASCADE
StationLine	stationId	MetroStation	stationId	R23	CASCADE	R24	CASCADE
StationLine	lineId	TransitLine	lineId	R25	CASCADE	R26	CASCADE
StationCrim e	stationId	MetroStation	stationId	R27	CASCADE	R28	CASCADE
StationCrim e	crimeTypeId	CrimeCategory	crimeTypeId	R29	NO ACTION	R30	CASCADE

Sample Data:

- TimelineReference: (dateId=1, year=2019, month=5, periodType='Month'')

- MetroStation: (stationId=101, stationCode='MC', stationName='Metro Center', stationLat=38.8983, stationLon=-77.0281, openedYear=1976)

- TransitLine: (lineId=1, lineCode='RD', lineName='Red Line', mode='Rail')
- CrimeCategory: (crimeTypeId=5, crimeTypeCode='T01', crimeTypeName='Theft/Larceny', severity='PartII', categoryGroup='Property')

- StationSnapshot: (snapshotId=1001, dateId=1, stationId=101, avgWeekdayBoardings=15234, totalEntries=15234, totalExits=14800)

- SystemRidership: (ridershipId=2001, dateId=1, mode='Rail', avgWeekdayRidership=450000, totalRidership=13500000)

- OnTimePerformance: (onTimeId=3001, dateId=1, lineId=1, scheduledTrips=12000, onTimeTrips=11000, onTimePct=91.7, avgDelayMinutes=2.1)

- ServiceIncident: (incidentId=4001, dateId=1, lineId=1, stationId=101, incidentCount=5, avgDelayMinutes=8.5, primaryCause='Signal')

- ReportedCrime: (crimeId=5001, dateId=1, stationId=101, crimeTypeId=5, incidentCount=2, arrestMadeFlag=true)

- StationLine: (stationLineId=6001, stationId=101, lineId=1, openedYear=1976)

- StationCrime: (stationCrimeId=7001, stationId=101, crimeTypeId=5, reportedYear=2019, incidentCount=2)