

BUDT703 - Database Management Systems

Project - Design

Created By: Palak Kishore, Ashna Sandoze, Maitreyee Tiwari, Manisha Ahuja

Brand Name: TerpTrak Solutions

Mission Statement: To analyze Amtrak records to identify meaningful correlations between key factors and trends over time

Mission Objectives:

1. To analyze the top and bottom ten stations with the highest and lowest ridership growth, and assess how these growth trends correlate with their total procurement and budget allocations.
2. To identify states where ridership exceeded rewards participation in 2023 and analyze the differences in ridership and rewards growth between 2022 and 2023.
3. To identify the bottom fifteen routes with the lowest on-time performance (OTP), while considering other relevant factors to develop strategies for improving overall route efficiency and performance.
4. To evaluate the stations that have experienced a drop in ridership count compared to the usual trends, and analyze the correlation with other factors to identify potential causes and areas for improvement.

Entities and Attributes:

There are seven entities each with a set of attributes and relationships with one another. Below is the breakdown of each entity and its attributes:

1. State

The state entity will contain the state details which have stations.

Attributes:

- stateId: Unique identifier for each state (primary key).
- stateName: Name of the state.

Relationships:

- Acts as a central entity connected to Employee, Reward, and Station.
- Represents where employees reside, where members receive rewards, and where the stations are located.

2. Station

The station entity will contain the Amtrak-related station details.

Attributes:

- stationCode: Unique code for the station (primary key).
- stationName: Name of the station.
- stationURL: URL for the station's webpage.

Relationships:

- Connected to Ridership and State, showing station location and ridership statistics.

- Linked to Route , showing which routes stop at the station.
- Connected to Budget and Procurement, showing budget allocation and procurement spending for the station.

3. Budget

The budget entity consists of budget details associated with certain stations and certain years.

Attributes:

- budgetId: Unique identifier for the budget (primary key).
- budgetYear: Year of the budget.
- budgetType: Type of the budget (e.g., design, construction, deployment).
- budgetAmount: Amount allocated in the budget.

Relationships:

- Connected to Station, representing budget allocations for particular stations.

4. Ridership

The ridership entity contains the station usage details. Ridership is defined as the total number of boarding and alightings at a station.

Attributes:

- ridershipYear: Year for which ridership data is recorded.(Primary Key combined with stationCode)
- ridershipCount: Number of riders for the given year.

Relationships:

- Connected to Station, indicating that ridership data is associated with specific stations.

5. Route

This entity represents Amtrak train routes.

Attributes:

- routeId: Unique identifier for each route (Primary key)
- routeName: Name of the route
- routeType: Type of the route
- routeFrequency: Frequency of route's service

Relationships:

- Connected to Station, showing which stations are stops for the route
- Connected to OTP, showing on-time performance percent for each route

6. OTP (On-Time Performance):

Attribute:

- routeOTPYear: Year of OTP tracked (Primary Key along with routeId)
- routeOTP: On time performance for the associated route in a particular year. It provides us insight about the timelines of the route.

Relationships:

- Connected to Route, showing performance metrics for each route per year.

7. Reward

The reward entity consists of Amtrak's AGR rewards details.

Attributes:

- rewardId: Unique identifier for the reward (primary key).
- rewardYear: Year of the reward count.
- rewardCount: Number of customers registered in the AGR reward program.

Relationships:

- Connected to State, indicating rewards associated with members in specific states.

8. Procurement

The procurement entity consists of Amtrak's procurement details.

Attributes:

- procurementId: Unique identifier for procurement data (Primary key).
- procurementYear: Year of procurement.
- procurementAmount: Amount procured by the associated station.

Relationships:

- Connected to Station, showing procurements for specific stations.

9. Employee

The employee entity contains a certain state's residing employees details.

Attributes:

- employeeYear: Year associated with employee data (Primary key along with stateId).
- employeeCount: The number of employees.
- employeeSalarySum: The total salary of all employees for the given year.

Relationships:

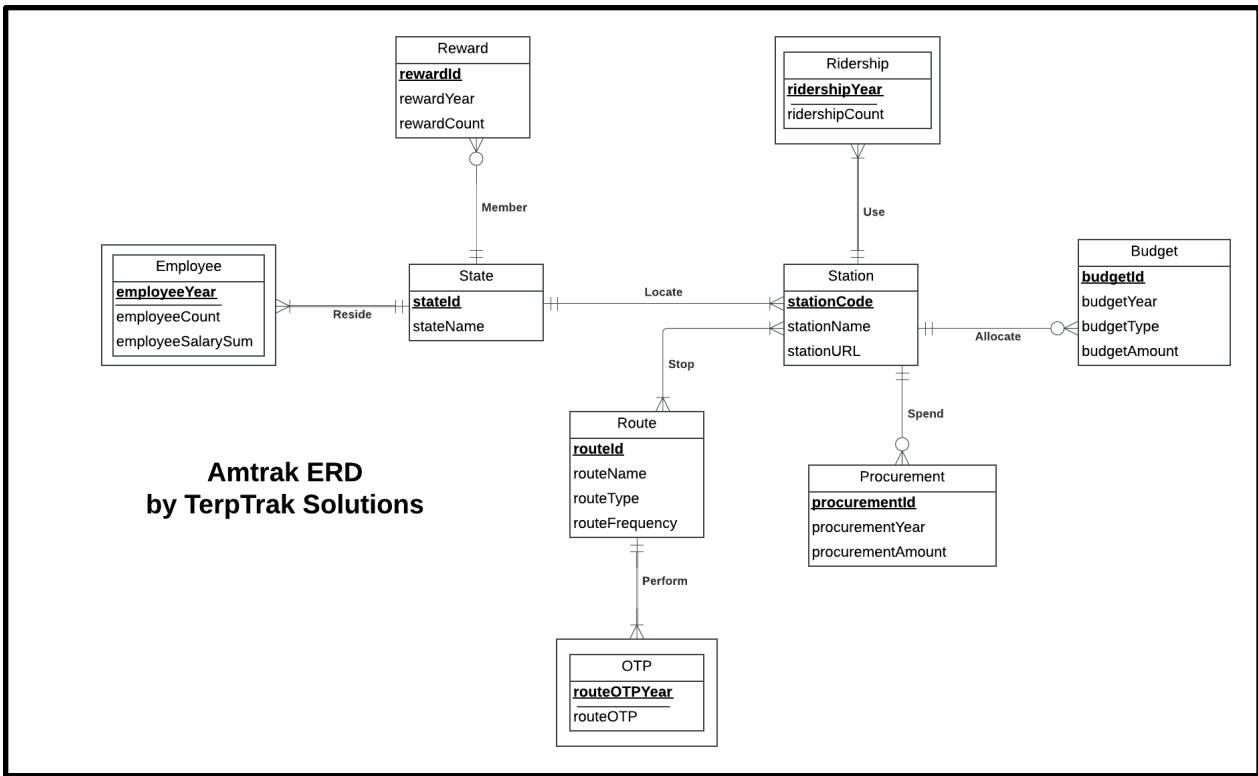
- Connected to State, indicating that employees reside in a particular state.

Assumptions:

- All states in our database have to have at least one station and a state can have multiple stations located in the same state.
- A station has to be located in a state and can only be located in a single state.
- A station might not have a budget allocated but can also have multiple budgets allocated of different types or in different financial years.
- A budget has to be allocated to a station and a budget can only be allocated to a single station.
- A station has to have at least one ridership count as ridership count is the station usage count and a station can have multiple ridership count records each for different financial years.

- A ridership record has to be linked with one station as it is the usage of a single station and the ridership record at most can be linked to only one station.
- Ridership records are not independent of the corresponding station as the ridership entity contains the stations riders count.
- A station might not have spent any procurement amount and a station could have spent different procurement amounts in multiple financial years.
- Certain procurement amount has to be spent by at least one station and at most can be spent by only one station.
- A state might not have any members in their rewards program in any given year and a state could also have multiple reward member counts in different financial years.
- Each rewards member count record has to belong to at least one state and the rewards member count record must at most belong to one state.
- Employees' details are not independent of the state they reside in. Therefore, each employee's yearly information is identified by a combination of state ID and a financial year.
- All states in our database have at least one financial year employee information and can have multiple financial year employee records.
- An employee details record has to be linked to at least and at most a single state as it contains the information of employees residing in the particular state.
- A particular station has to be a stop in at least one route and a station can be a stop in multiple routes.
- A particular route must have at least one station as a stop and a route can have multiple station stops.
- OTP details are not independent of Route as the record describes the on time performance of the corresponding route. Therefore, each OTP record is identified with a combination of OTP year and route ID.
- All routes in our database have at least one OTP record and can have multiple OTP records as it is recorded yearly.
- An OTP record has to be linked with at least and at most a single route as it contains the information of on time performance per year per route.

ER Diagram: [Link](#)



Relational schema

State(stateId, stateName)

Station(stationCode, stationName, stationURL, *stateId*)

Procurement(procurementId, procurementYear, procurementAmount, *stationCode*)

Ridership (stationCode, ridershipYear, ridershipCount)

Budget (budgetId, budgetYear, budgetType, budgetAmount, *stationCode*)

Route (routeId, routeName, routeType, routeFrequency)

Stop(routeId, stationCode)

Employee (stateId, employeeYear, employeeCount, employeeSalarySum)

Reward (rewardId, rewardYear, rewardCount, *stateId*)

OTP (routeId, routeOTPYear, routeOTP)

Business Rules:

R1: When a state's information is deleted from or changed in the database, all station records referencing that state should remain unchanged.

R2: When a station is removed from the system then the amount of procurement for that station should be set as null.

R3: Any changes to a station's information should result in an automatic update of the amount of procurement related to that station.

R4: When a station is removed from the system then the ridership records for that station should also be deleted.

R5: When a station's information is changed in the database, then all the ridership records related to that station should be changed accordingly.

R6: When a station is deleted from the system then the budget information for that station should be set as null.

R7: When the budget allocation for a station is changed in the database, the corresponding budget information should be changed accordingly.

R8: When a route is deleted from the system then all stop records related to that route should be deleted from the database.

R9: When a route's information is changed in the database then the corresponding stop information should be updated accordingly.

R10: All the stops related to a station should be deleted if that station is deleted from the database.

R11: If a station's information is changed or updated then the stops related to that station should also be changed.

R12: When a state's information is deleted from or changed in the database, all employee records referencing that state should remain unchanged.

R13: When we change or update a state's information then the employees associated with that state should be changed or reassigned.

R14: When a state is removed from the system then the number of people who have guest rewards should be set as null.

R15: When a state's information is changed or updated in the database then the number of people having guest rewards in that state should also be changed accordingly.

R16: When a Routes is removed from the system then the route's corresponding OTP record should be removed from the database.

R17: When a routes information is updated in the database then the corresponding OTP record would be updated.

Referential Integrity Actions:

Relation	Foreign Key	Base Relation	Primary Key	Business Rule	ON DELETE	Business Rule	ON UPDATE
Station	stateId	State	stateId	R1	NO ACTION	R1	NO ACTION
Procurement	stationCode	Station	stationCode	R2	SET NULL	R3	CASCADE
Ridership	stationCode	Station	stationCode	R4	CASCADE	R5	CASCADE

Budget	stationCode	Station	stationCode	R6	SET NULL	R7	CASCADE
Stop	routeId	Route	routeId	R8	CASCADE	R9	CASCADE
Stop	stationCode	Station	stationCode	R10	CASCADE	R11	CASCADE
Employee	stateId	State	stateId	R12	NO ACTION	R13	CASCADE
Reward	stateId	State	stateId	R14	SET NULL	R15	CASCADE
OTP	routeId	Route	routeId	R16	CASCADE	R17	CASCADE

Sample Records:State(stateId, stateName)

stateId	stateName
AL	Alabama
FL	Florida
MD	Maryland
NY	New York

Station(stationCode, stationName, stationURL, *stateId*)

stationCode	stationName	stationURL	stateId
BHM	Birmingham	https://www.greatamericanstations.com/stations/birmingham-al-bhm/	AL
HOL	Hollywood	https://www.greatamericanstations.com/stations/hollywood-fl-hol/	FL
RKV	Rockville	https://www.greatamericanstations.com/stations/rockville-md-rkv/	MD
ROM	Rome	https://www.greatamericanstations.com/stations/rome-ny-rom/	NY

Budget (budgetId, budgetYear, budgetType, budgetAmount, *stationCode*)

budgetId	budgetYear	budgetType	budgetAmount	stationCode
001	2023	Design	643	BHM
002	2025	Design	58	ABE
003	2025	Construction	4287	WLO

Ridership (stationCode, ridershipYear, ridershipCount)

stationCode	ridershipYear	ridershipCount
ROM	2022	8496
RKV	2022	4058
RKV	2023	3560

Route (routeId, routeName, routeType, routeFrequency)

routeId	routeName	routeType	routeFrequency
001	Crescent	Long Distance	Daily
002	Acela	Northeast Corridor	Daily
003	Vermonter	State Supported	Daily

OTP(routeId, routeOTPYear, routeOTP)

routeId	routeOTPYear	routeOTP
001	2021	0.54
002	2022	0.82
003	2023	0.70

Stop(routeId, stationCode)

routeId	stationCode
001	WAS
002	LOR
003	CHI
005	SJC

Reward (rewardId, rewardYear, rewardCount, *stateId*)

rewardId	rewardYear	rewardCount	stateId
001	2021	36,069	AL
002	2021	402,806	FL
003	2021	400,131	MD
004	2021	1,107,245	NY

Procurement(procurementId, procurementYear, procurementAmount, *stationCode*)

procurementId	procurementYear	procurementAmount	stationCode
001	2021	4015429	BHM
002	2021	18781522	RKV
003	2022	26497387	RKV

Employee (stateId, employeeYear, employeeCount, employeeSalarySum)

stateId	employeeYear	employeeCount	employeeSalarySum
AL	2021	13	\$1,292,695
FL	2021	594	\$52,083,356
MD	2021	2123	\$213,125,686
NY	2021	1449	\$129,777,100

OTP (routeId, routeOTPYear, routeOTP)

routeId	routeOTPYear	routeOTP
001	2021	53.0%
002	2021	28.0%
003	2023	21.0%