



GROWTHTRAK

YingXin Ou
Truong Hien Vy Cai
Megan O'Brien
I-Tso Chen

December 6, 2024



TABLE OF CONTENTS

01

BACKGROUND

02

INTRODUCTION

03

DATABASE DESIGN

CONCEPTUAL / LOGICAL / PHYSICAL

04

BUSINESS TRANSACTIONS

05

RECOMMENDATIONS



1.BACKGROUND

► WHO USERS ARE (AMTRAK RIDERS)

Amtrak riders are typically a diverse group of individuals traveling for various reasons. We can break them down into three categories:

- Business Travelers
- Commuters
- Tourists and Leisure Travelers

► WHAT DATA/SOURCES IT INCLUDES?

- States
- Stations
- Route
- Ridership
- On-Time Performance
- Guest Rewards

2. INTRODUCTION

Mission:

To improve customer growth by analyzing and presenting data-driven improvements for ridership throughout stations for a reliable travel experience and customer rewards

Mission objectives:

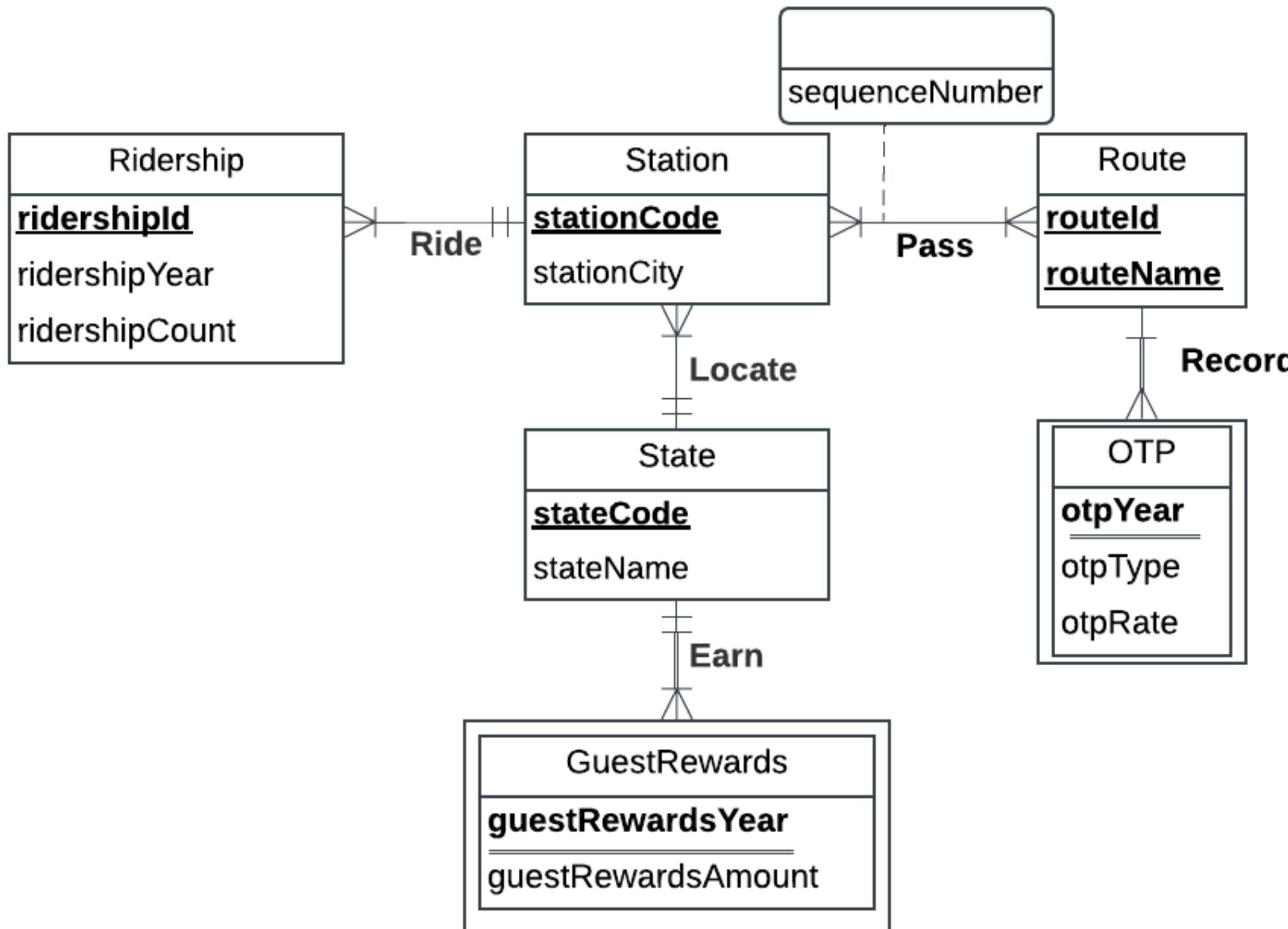
- Identify the **Distribution of Ridership** and growth rate of the past 3 years of each state.
- Highlight **Top 1% stations with low ridership but high growth rate** in the past 3 years to prioritize resources for faster growth.
- To find the **Top 10% Routes** with the **highest on-time performance** in the past 3 years.
- To identify **States with Low Guest Rewards participation**.



3. DATABASE DESIGN

CONCEPTUAL / LOGICAL / PHYSICAL

BUDT703_Project_0506_03_ERD



Relational schema:

State(stateCode, stateName)

Station(stationCode, stationCity, stateCode)

Route(routId, routeName, stationCode)

Ridership(ridershipId, ridershipYear, ridershipCount, stationCode)

OTP(routId, routeName, otpYear, otpType, otpRate)

GuestRewards(stateCode,guestRewardsYear, guestRewardsAmount)

Pass(routeName,stationCode, sequenceNumber)

```

CREATE TABLE [Ridership] (
    ridershipId CHAR(6) NOT NULL,
    ridershipYear SMALLINT,
    ridershipCount INT,
    stationCode CHAR(3) NOT NULL,
    CONSTRAINT pk_Ridership_ridershipId PRIMARY KEY(ridershipId),
    CONSTRAINT fk_Ridership_stationCode FOREIGN KEY (stationCode)
        REFERENCES [Station] (stationCode) ON DELETE NO ACTION ON UPDATE CASCADE
);
  
```

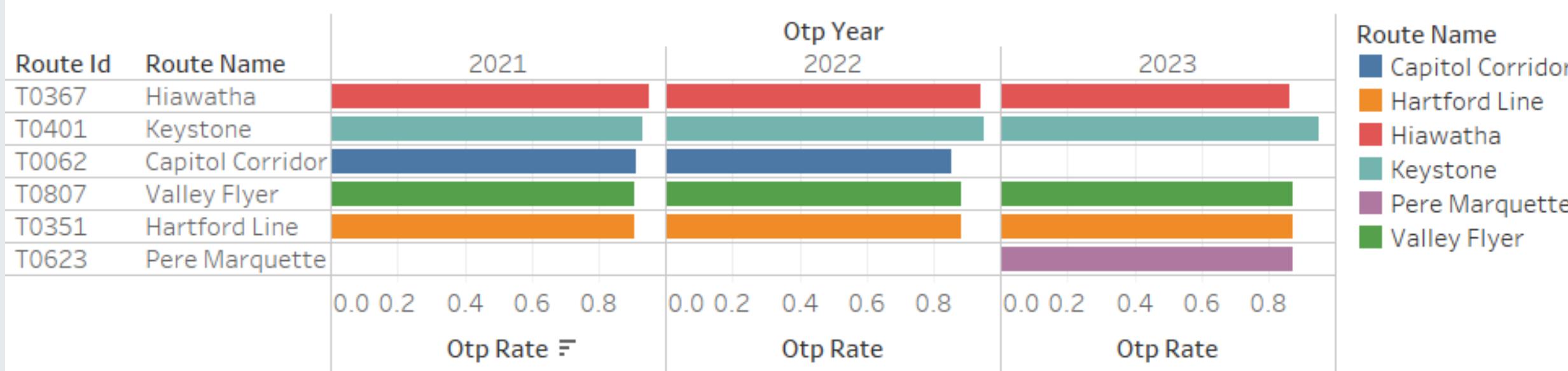
4.1. BUSINESS TRANSACTIONS #1

What are the top 10% routes with highest on-time performance?

```
WITH RankedRoutes AS (
    SELECT
        otpYear,
        routId,
        routeName,
        otpType,
        otpRate,
        NTILE(10) OVER (PARTITION BY otpYear ORDER BY otpRate DESC)
    AS RankGroup
    FROM OTP
)
SELECT
    otpYear AS 'Otp Year',
    routId AS 'Route Id',
    routeName AS 'Route Name',
    otpRate AS 'Otp Rate'
FROM RankedRoutes
WHERE RankGroup = 1
ORDER BY otpYear, otpRate DESC;
```

	Otp Year	Route Id	Route Name	Otp Rate
1	2021	T0367	Hiawatha	0.948
2	2021	T0401	Keystone	0.931
3	2021	T0062	Capitol Corridor	0.912
4	2021	T0351	Hartford Line	0.904
5	2021	T0807	Valley Flyer	0.904
6	2022	T0401	Keystone	0.950
7	2022	T0367	Hiawatha	0.940
8	2022	T0351	Hartford Line	0.880
9	2022	T0807	Valley Flyer	0.880
10	2022	T0062	Capitol Corridor	0.850
11	2023	T0401	Keystone	0.950
12	2023	T0351	Hartford Line	0.870
13	2023	T0623	Pere Marquette	0.870
14	2023	T0807	Valley Flyer	0.870
15	2023	T0367	Hiawatha	0.860

What are the top 10% route with highest On-time Performance?



4.2. BUSINESS TRANSACTIONS #2

What are the stations with the most and least ridership growth over the last 3 years?

```
--TOP
WITH a AS (
    SELECT r.stationCode,
        SUM(CASE WHEN ridershipYear = '2021' THEN ridershipCount ELSE 0 END) AS 'Ridership 2021',
        SUM(CASE WHEN ridershipYear = '2023' THEN ridershipCount ELSE 0 END) AS 'Ridership 2023'
    FROM Ridership r
    GROUP BY r.stationCode
)
SELECT TOP 5
    a.stationCode AS 'Station Code',
    s.stateName AS 'Station Name',
    "Ridership 2021" AS 'Total Ridership 2021',
    "Ridership 2023" AS 'Total Ridership 2023',
    ROUND(
        CASE
            WHEN "Ridership 2021" = 0 THEN 0
            ELSE (CAST("Ridership 2023" AS FLOAT) / CAST("Ridership 2021" AS FLOAT) - 1) * 100
        END,
        2
    ) AS 'Growth Rate %'
FROM a
JOIN Station st ON a.stationCode = st.stationCode
JOIN State s ON st.stateCode = s.stateCode
ORDER BY 'Growth Rate %' DESC;
```

```
(--BOTTOM
ASC; )
```

	Station Code	Station Name	Total Ridership 2021	Total Ridership 2023	Growth Rate %
1	SKT	Califomia	1	11714	1171300
2	MPK	Califomia	1050	23091	2099.14
3	MKA	Wisconsin	37580	509107	1254.73
4	CWH	Pennsylvania	234	2190	835.9
5	HMI	Indiana	445	3950	787.64

	Station Code	Station Name	Total Ridership 2021	Total Ridership 2023	Growth Rate %
1	BBY	Massachusetts	322928	0	-100
2	SND	Texas	153	0	-100
3	KKI	Illinois	25495	13944	-45.31
4	MKE	Wisconsin	198389	110938	-44.08
5	ESM	Montana	1740	1009	-42.01

Which stations have the most and least growth rate over the past 3 years?



4.3. BUSINESS TRANSACTIONS #3

Which stations fall within the top 1% for high growth rates but have low ridership over the past three years?

WITH RankedStations AS (

```
    SELECT
        r.stationCode AS [Station Code],
        st.stateName AS [State Name],
        SUM(CASE WHEN r.ridershipYear = 2021 THEN r.ridershipCount ELSE 0 END) AS [Total Ridership in 2021],
        SUM(CASE WHEN r.ridershipYear = 2023 THEN r.ridershipCount ELSE 0 END) AS [Total Ridership in 2023],
        CAST(ROUND(((SUM(CASE WHEN r.ridershipYear = 2023 THEN r.ridershipCount ELSE 0 END) -
            SUM(CASE WHEN r.ridershipYear = 2021 THEN r.ridershipCount ELSE 0 END)) * 1.0 /
            NULLIF(SUM(CASE WHEN r.ridershipYear = 2021 THEN r.ridershipCount ELSE 0 END), 0)) * 100, 2
        ) AS DECIMAL(10, 2)) AS [Growth Rate %],
        PERCENT_RANK() OVER (ORDER BY
            CAST(ROUND(((SUM(CASE WHEN r.ridershipYear = 2023 THEN r.ridershipCount ELSE 0 END) -
                SUM(CASE WHEN r.ridershipYear = 2021 THEN r.ridershipCount ELSE 0 END)) * 1.0 /
                NULLIF(SUM(CASE WHEN r.ridershipYear = 2021 THEN r.ridershipCount ELSE 0 END), 0)) * 100, 2
            ) AS DECIMAL(10, 2)) DESC) AS [GrowthPercentile]
```

FROM

Ridership r

JOIN Station s ON r.stationCode = s.stationCode

JOIN [State] st ON s.stateCode = st.stateCode

GROUP BY

r.stationCode, st.stateName

HAVING

SUM(CASE WHEN r.ridershipYear = 2023 THEN r.ridershipCount ELSE 0 END) <

```
    SELECT
```

```
        AVG(a.[Total Ridership in 2023])
```

FROM (

```
    SELECT
```

```
        SUM(CASE WHEN r.ridershipYear = 2023
            THEN r.ridershipCount ELSE 0 END) AS [Total Ridership in 2023]
```

FROM

Ridership r

GROUP BY r.stationCode) a))

SELECT

```
[Station Code],
[State Name],
[Total Ridership in 2021],
[Total Ridership in 2023],
[Growth Rate %]
```

FROM

RankedStations

WHERE

```
[GrowthPercentile] <= 0.01
```

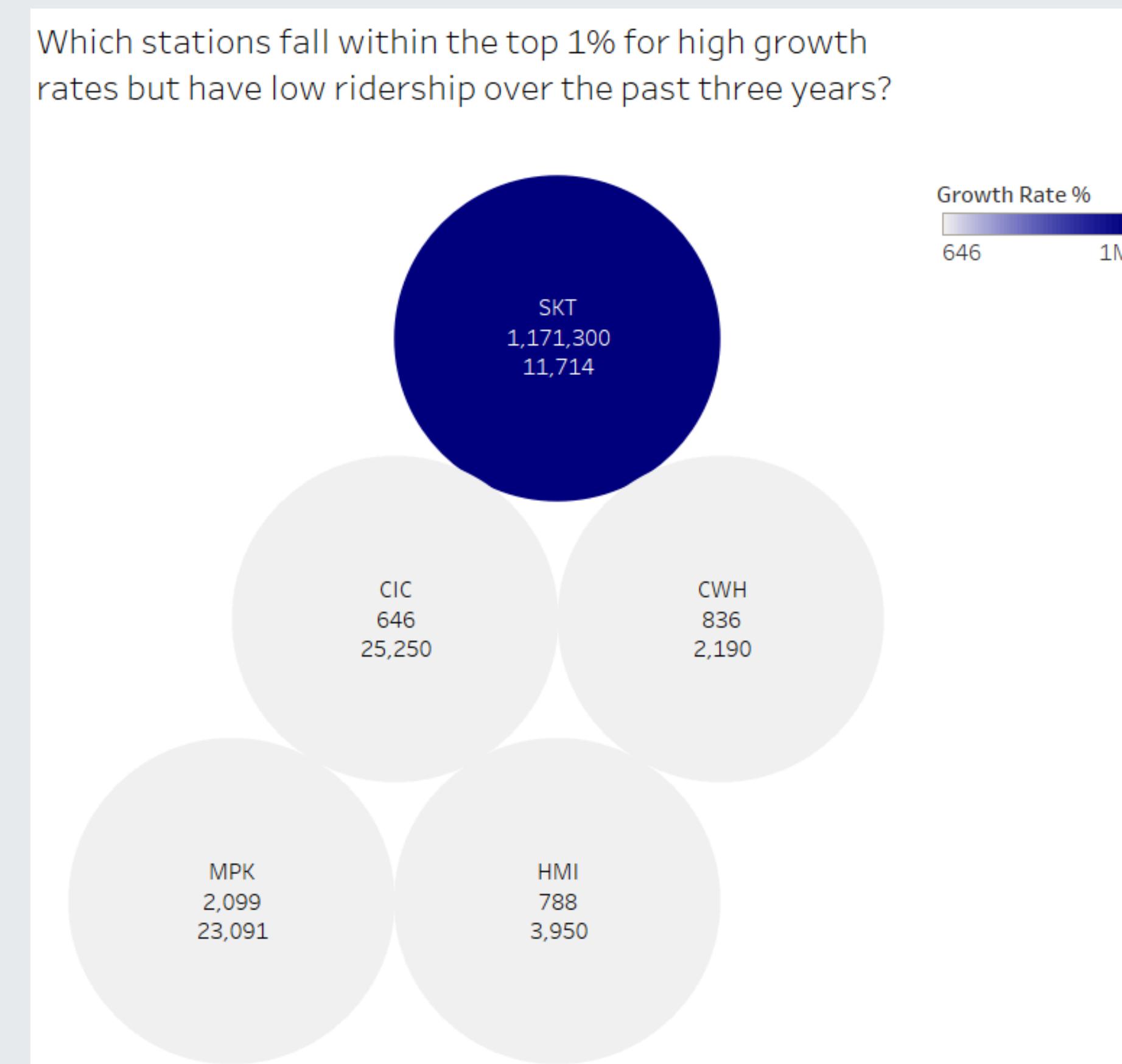
ORDER BY

```
[Growth Rate %] DESC,
[Station Code];
```

	Station Code	State Name	Total Ridership in 2021	Total Ridership in 2023	Growth Rate %
1	SKT	California	1	11714	1171300.00
2	MPK	California	1050	23091	2099.14
3	CWH	Pennsylvania	234	2190	835.90
4	HMI	Indiana	445	3950	787.64
5	CIC	California	3383	25250	646.38

4.3. BUSINESS TRANSACTIONS #3

Which stations fall within the top 1% for high growth rates but have low ridership over the past three years?



4.4. BUSINESS TRANSACTIONS #4

What are the states with the lowest guest rewards participation?

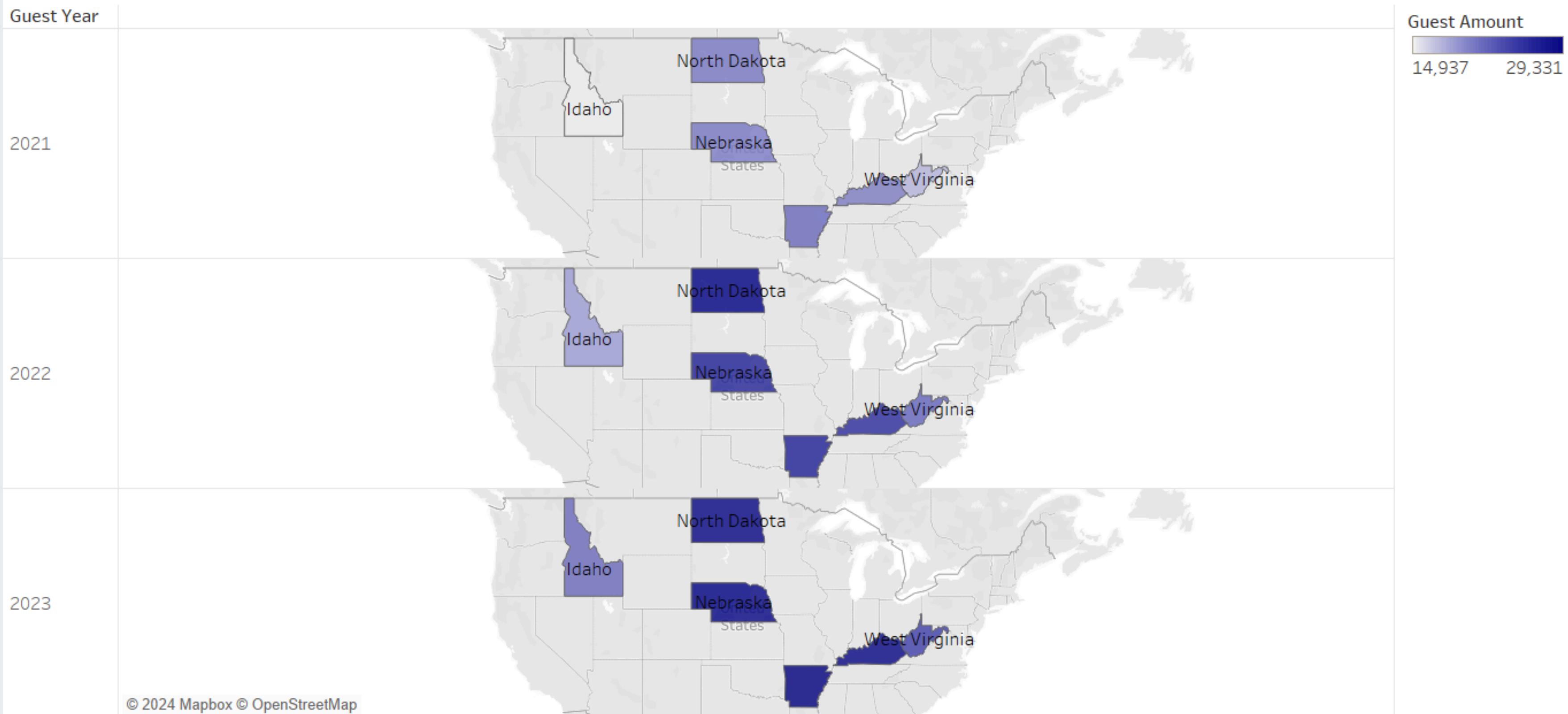
```
SELECT
    s.stateName AS 'State Name',
    a.guestYear AS 'Guest Year',
    a.guestAmount AS 'Guest Amount'
FROM (
    SELECT
        r.stateCode,
        r.guestYear,
        r.guestAmount,
        RANK() OVER (PARTITION BY r.guestYear ORDER BY r.guestAmount ASC) AS row
    FROM GuestRewards r
) AS a
JOIN State s ON a.stateCode = s.stateCode -- Join the State table to get the state name
WHERE a.row IN (
    SELECT row
    FROM (
        SELECT
            RANK() OVER (PARTITION BY r.guestYear ORDER BY r.guestAmount ASC) AS row
        FROM GuestRewards r
    ) AS r
    WHERE r.row <= 6
)
ORDER BY a.guestYear, a.guestAmount;
```

	State Name	Guest Year	Guest Amount
1	Idaho	2021	14937
2	West Virginia	2021	17360
3	Kentucky	2021	20363
4	Nebraska	2021	20399
5	North Dakota	2021	20529
6	Arkansas	2021	21101
7	Idaho	2022	18750
8	West Virginia	2022	21584
9	Kentucky	2022	25390
10	Nebraska	2022	25984
11	Arkansas	2022	26133
12	North Dakota	2022	28962
13	Idaho	2023	21088
14	West Virginia	2023	23924
15	Kentucky	2023	28394
16	North Dakota	2023	28677
17	Nebraska	2023	28866
18	Arkansas	2023	29331

4.4. BUSINESS TRANSACTIONS #4

What are the states with the lowest guest rewards participation?

Which states have the lowest guest rewards participation?



5. RECOMMENDATIONS

► Focus on On-Time Performance Improvement

- **Actions:**
 - **Identify Root Causes of OTP Issues:** Break down delays by type (e.g., equipment failures, crew shortages, track maintenance, weather).
 - **Optimize Operational Processes:** Collect input from train operators, station staff, and maintenance teams to understand operational challenges, providing adequate training to handle unexpected disruptions.
 - **Invest in Technology and Tools:** Implement predictive analytics to foresee and mitigate potential bottlenecks, such as traffic congestion or mechanical failures. Improve communication with passengers by providing real-time updates on delays and alternative solutions.

► Boost Growth at High-Potential Stations

- **Target Areas:** California (SKT, MPK, CIC), Pennsylvania (CWH), Indiana (HMI).
- **Actions:**
 - **Align schedules:** Adjust train schedules to better align with travel patterns.
 - **Upgrade facilities:** Modernize stations with amenities like Wi-Fi, charging ports, and food courts.

5. RECOMMENDATIONS

► **Grow Guest Rewards Participation**

- **Underperforming Areas:** Idaho, West Virginia, Kentucky, Nebraska, Arkansas.
- **Tailor engagement strategies:**
 - Offer discounts for young people, like students with ability to purchase tickets.
 - Launch localized marketing campaigns in targeted areas, like on campus-campaigns.
 - Offer incentives like bonus rewards for first-time sign-ups.
 - Partner with local businesses to enhance visibility and appeal.
 - Explore Barriers to Participation:
 - Address issues like awareness gaps or accessibility challenges.



THANK YOU

*FOR ATTENDING THE PRESENTATION
AND YOUR ENGAGEMENT*