

**Massachusetts Bay
Transportation Authority**

Schedule Deficiency

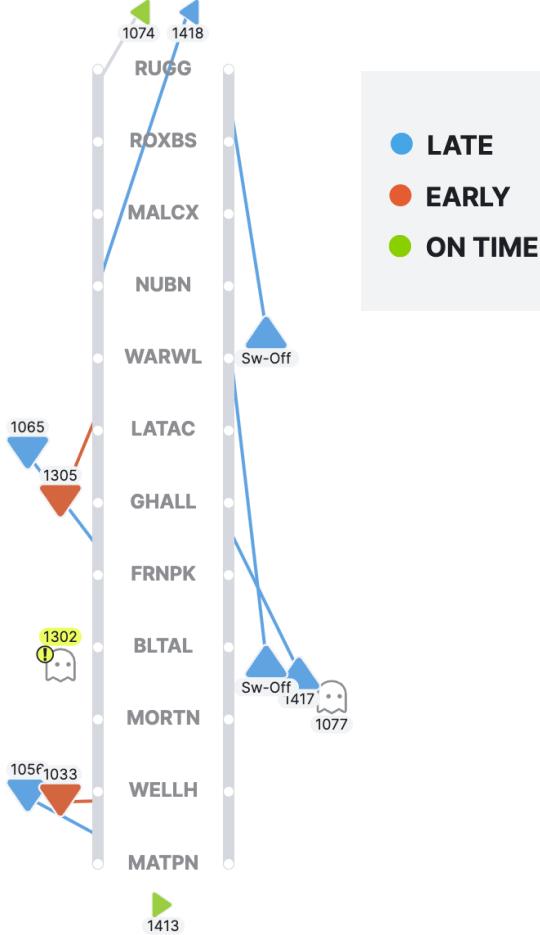
Final Presentation
08/29/25

Nail Bashan
Service Planning Intern

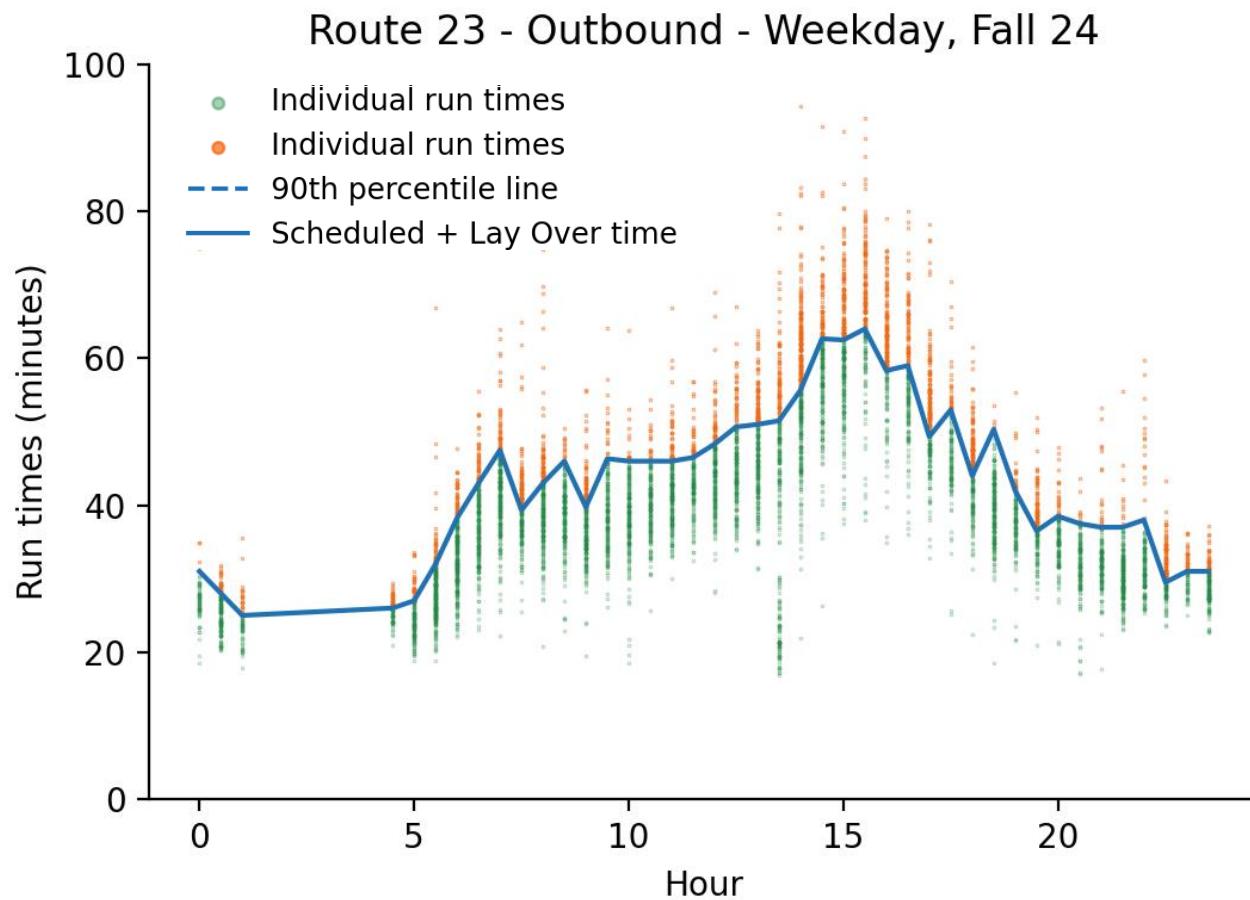
bashan.n@northeastern.edu

Schedule deficiency

28

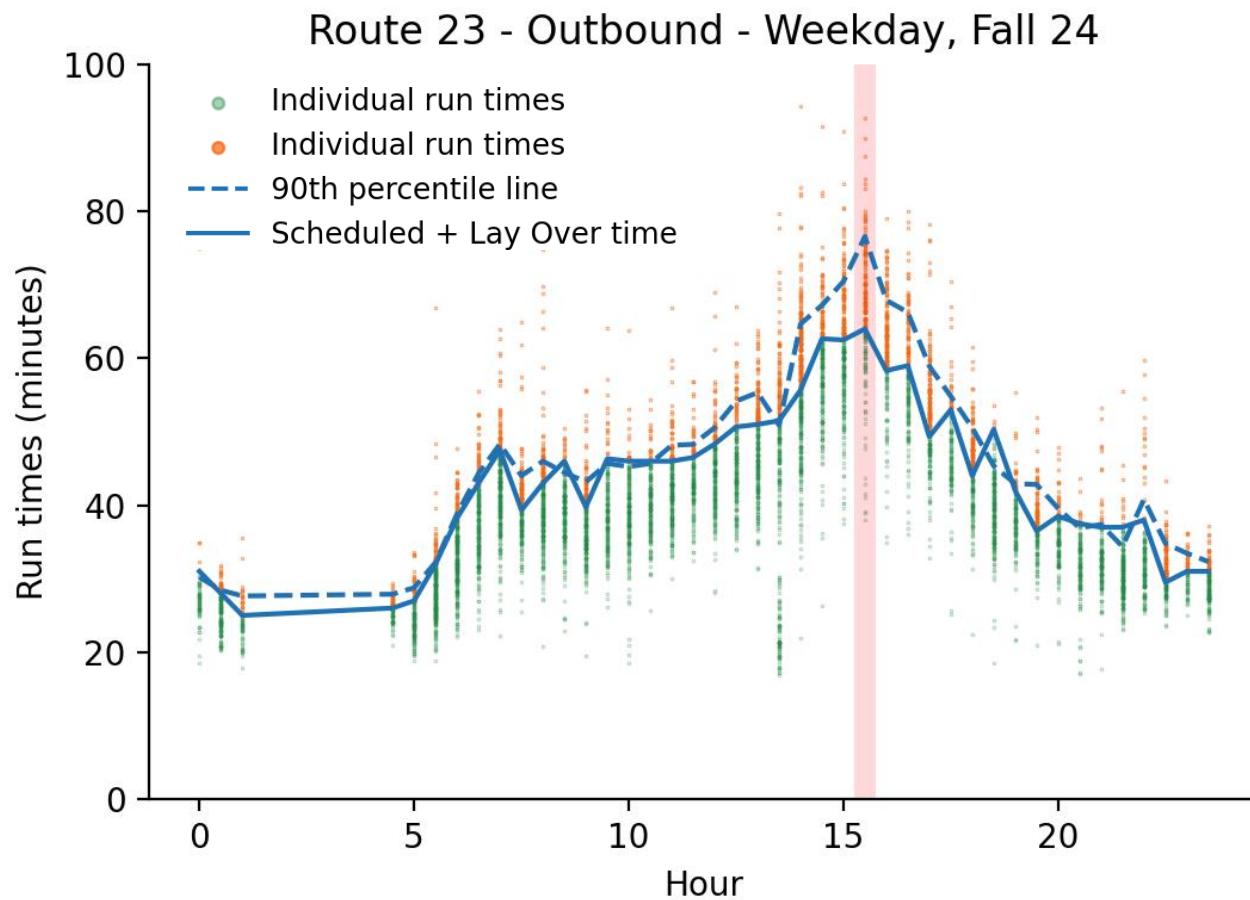


Schedule deficiency



Schedule deficiency is the total time by which buses typically run longer than the timetable allows.

Schedule deficiency



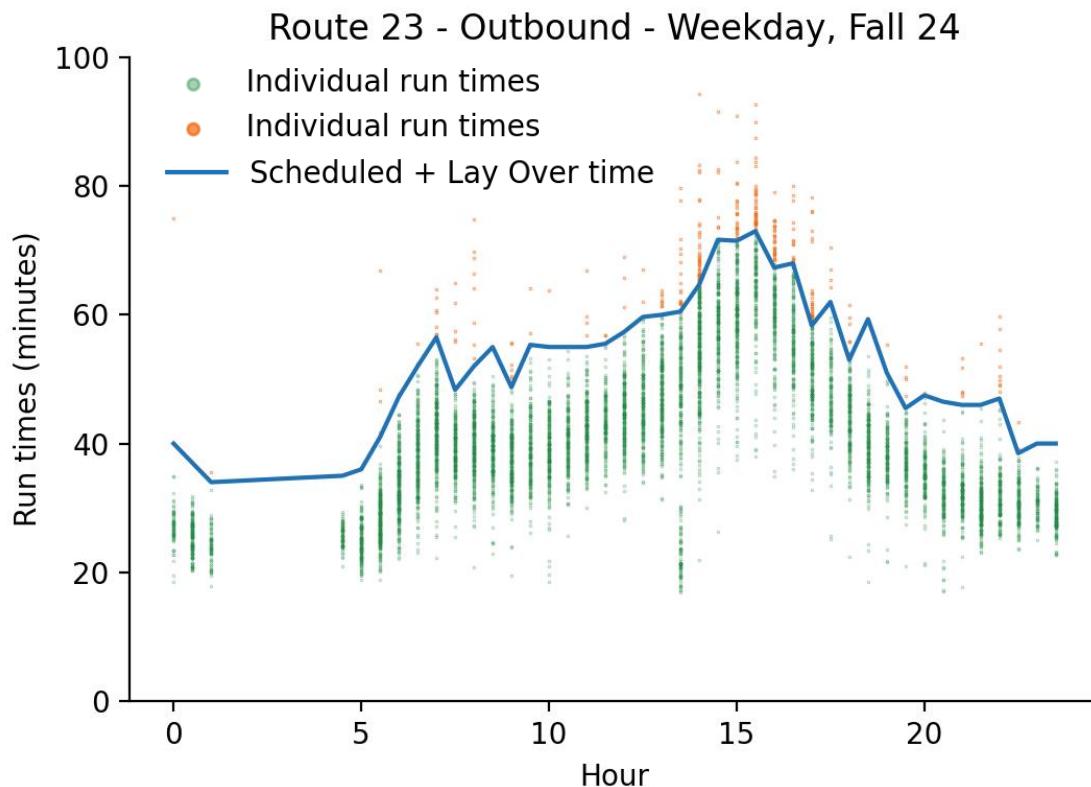
Schedule deficiency is the total time by which buses typically run longer than the timetable allows.

Goal is 90% of the busses will run less than this scheduled + lay over time (ideal schedule)

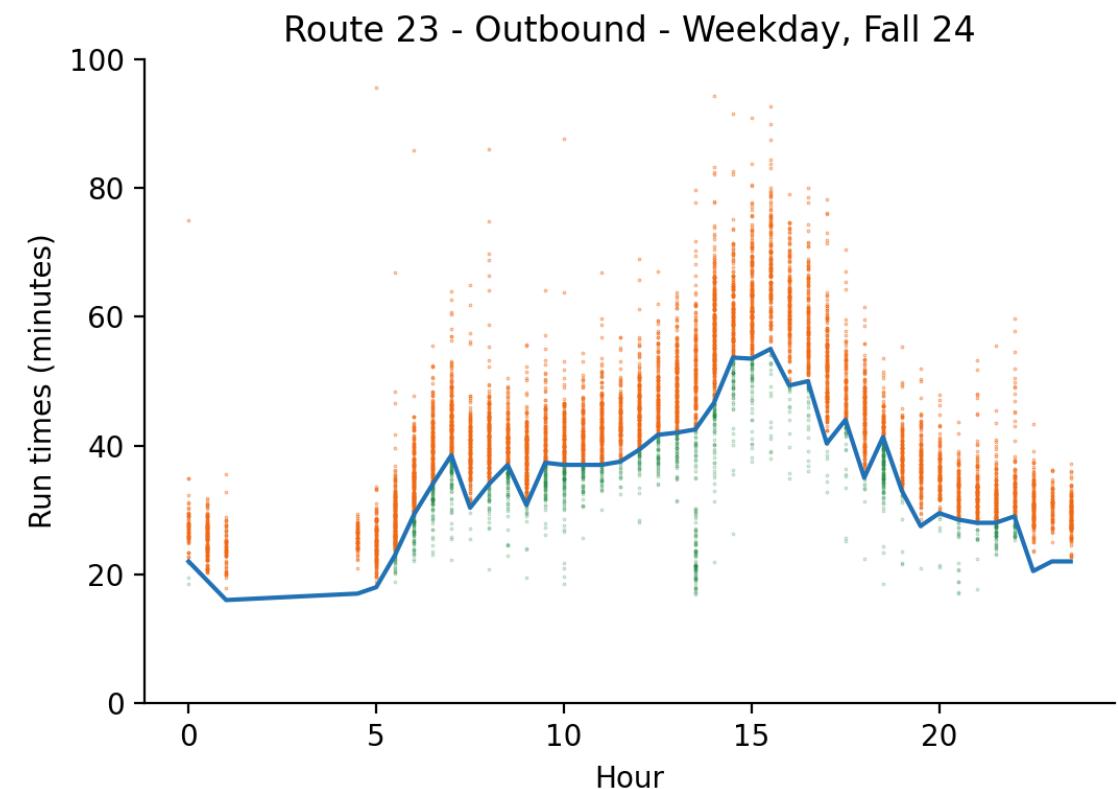
Impacts include decreased service quality for **riders**, less recovery time for **drivers**, and inefficient use of our **resources**

Schedule deficiency

What is the solution?



Zero deficiency but inefficient use of our resources



Zero surplus but no recovery time

Challenges and motivation

Schedule Deficiency should be

**1.Calculated
better**

**2.Understood
better**



**3.Corrected
better**

- **Better resource management**
- **Higher on-time performance**
- **Reasonable layover times**



Challenges and motivation

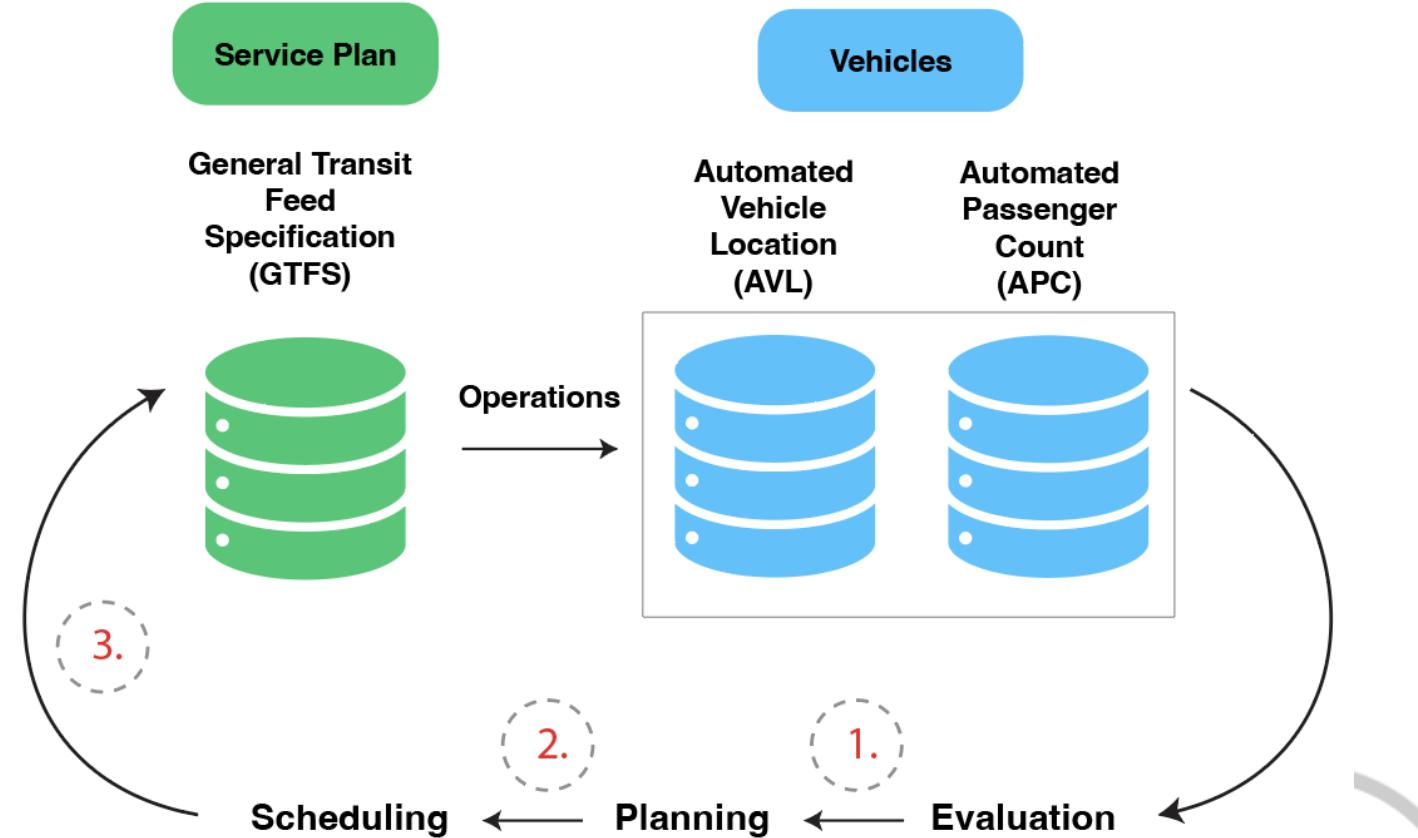
**Schedule Deficiency
should be**

**1.Calculated
better**

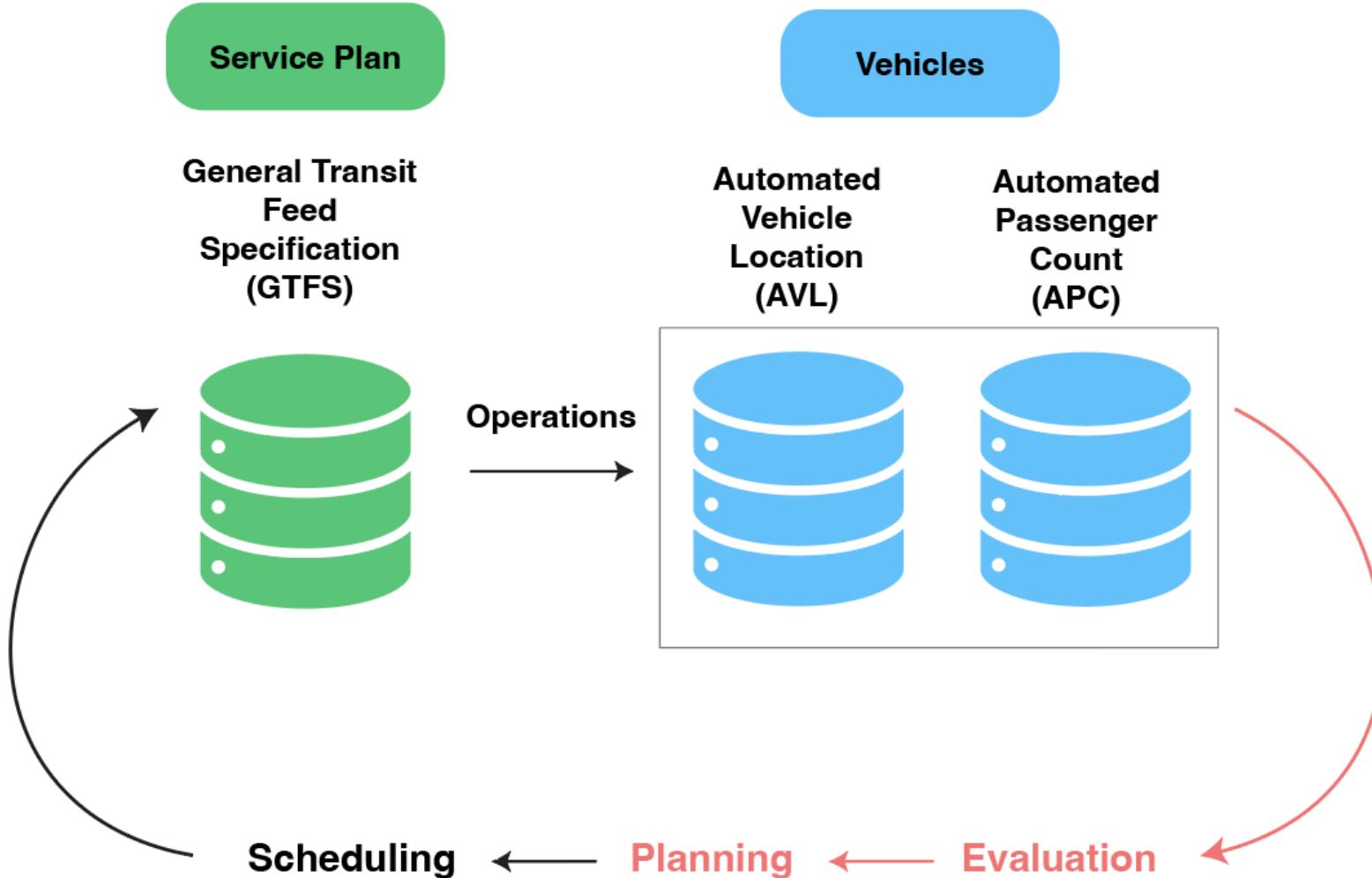
**2.Understood
better**

+

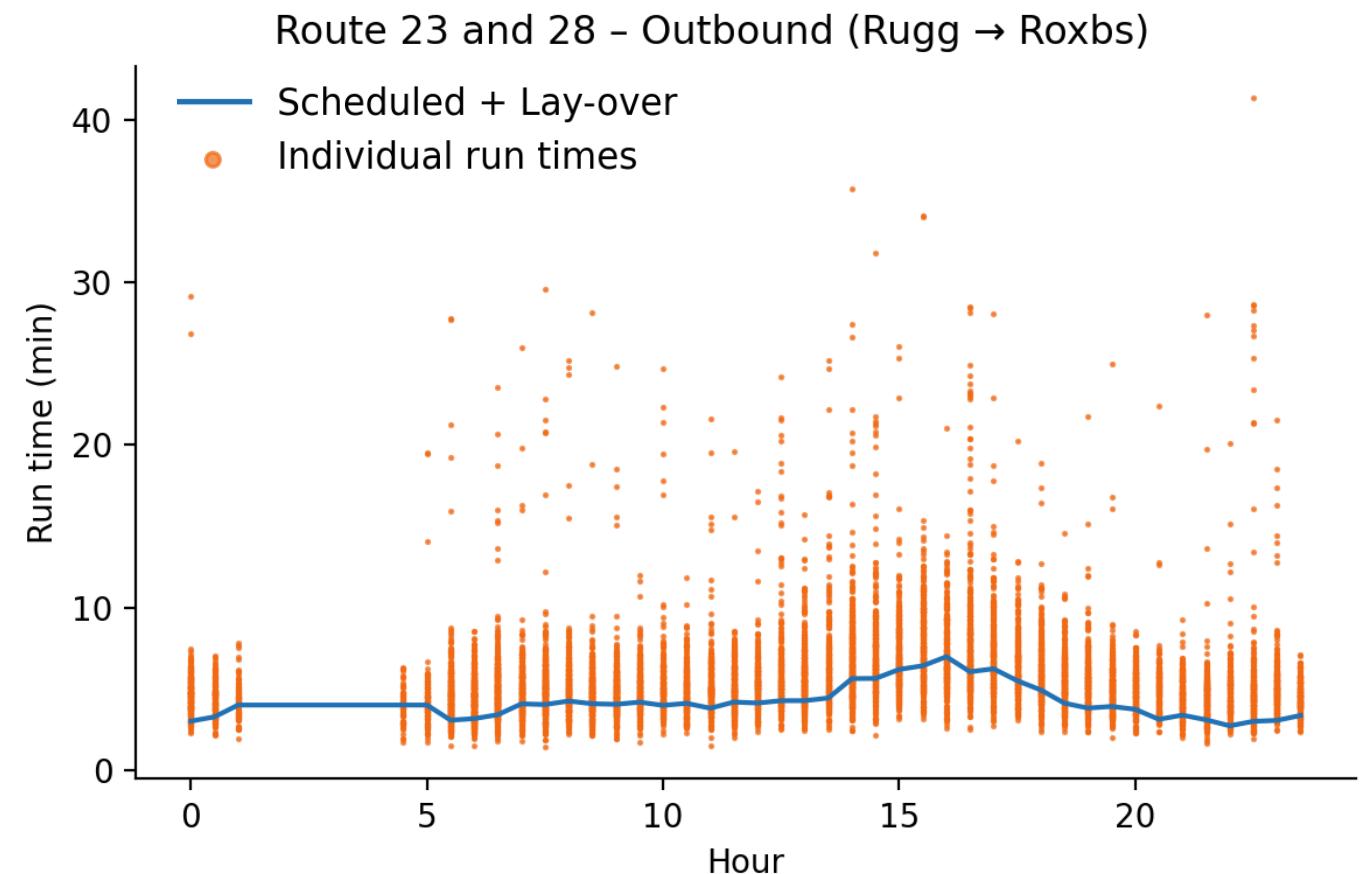
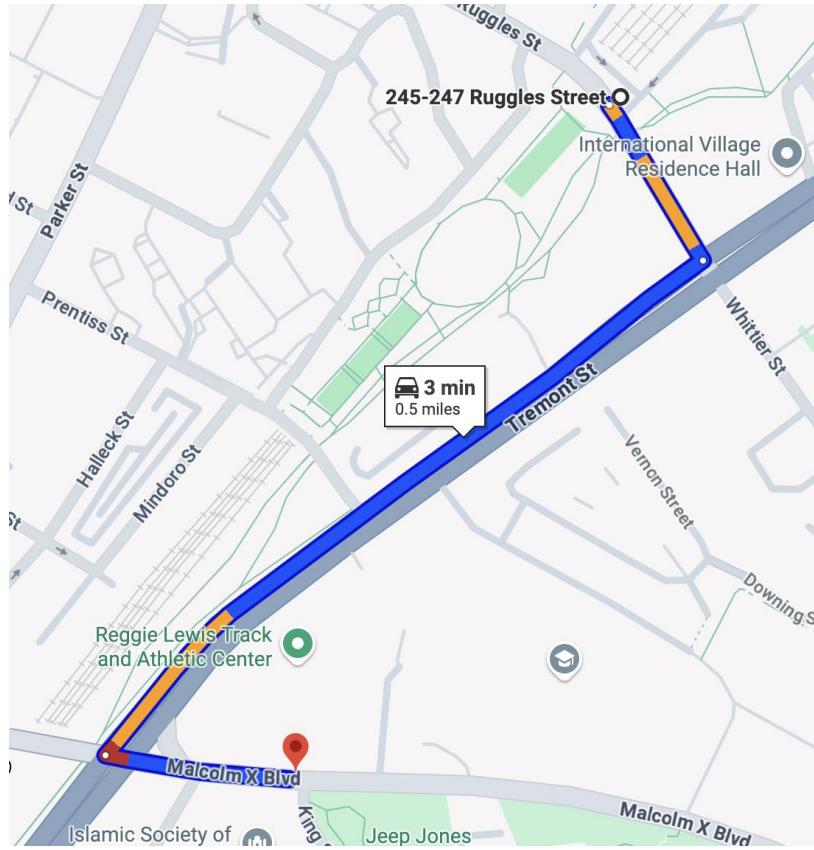
**3.Corrected
better**



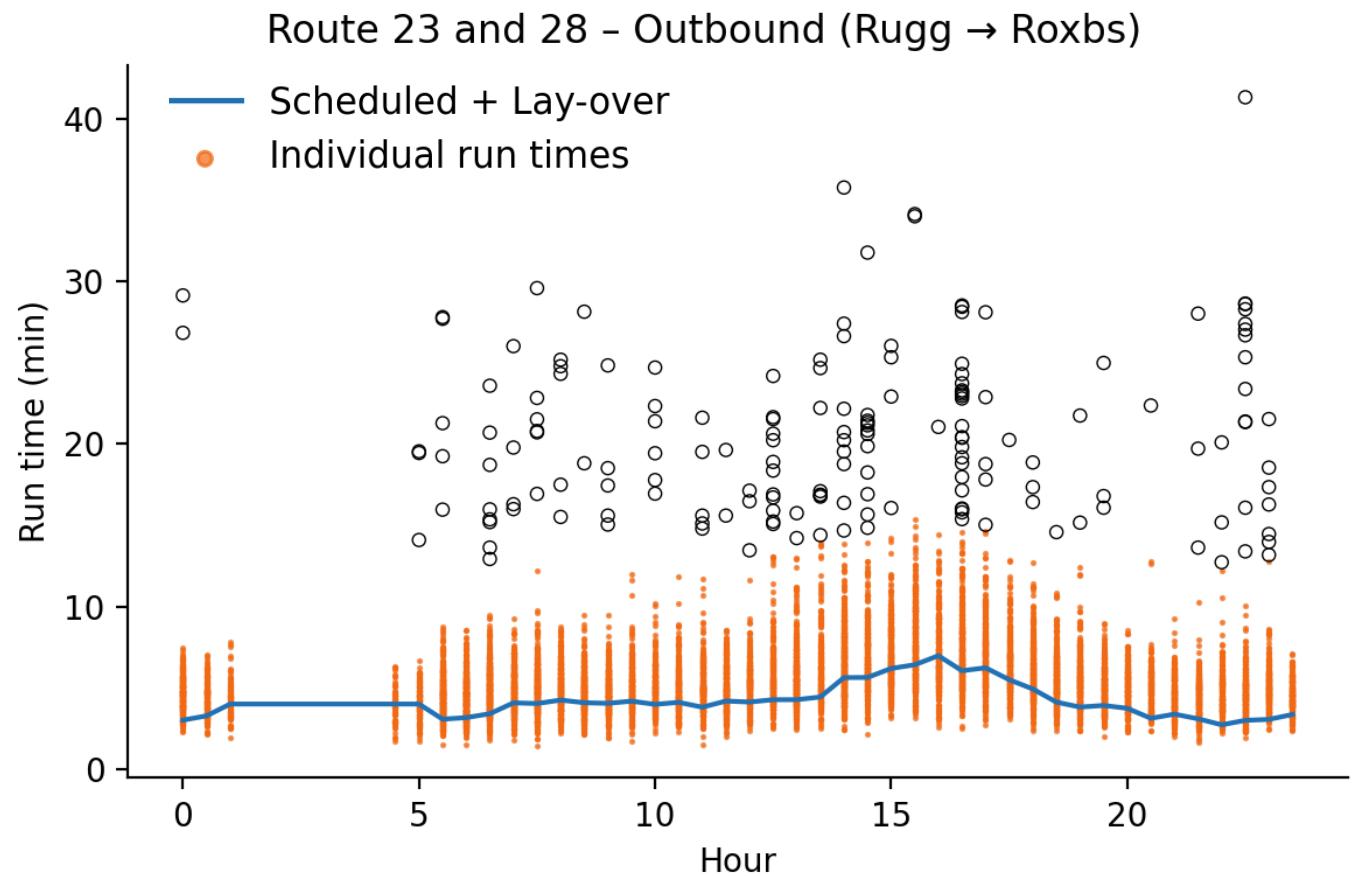
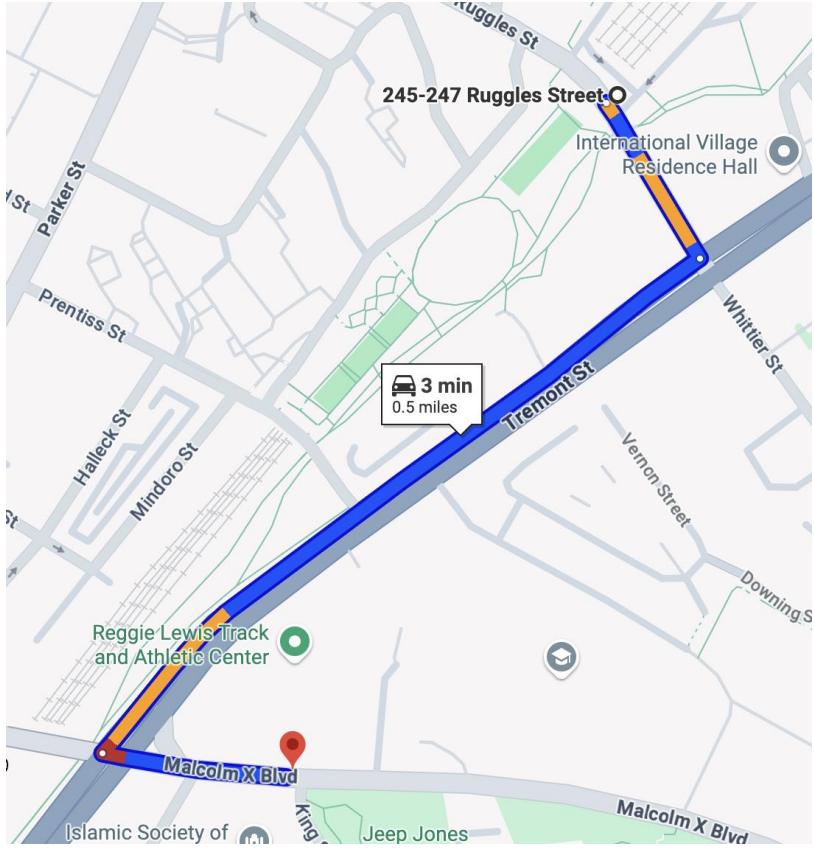
1. Calculate better



Calculate better



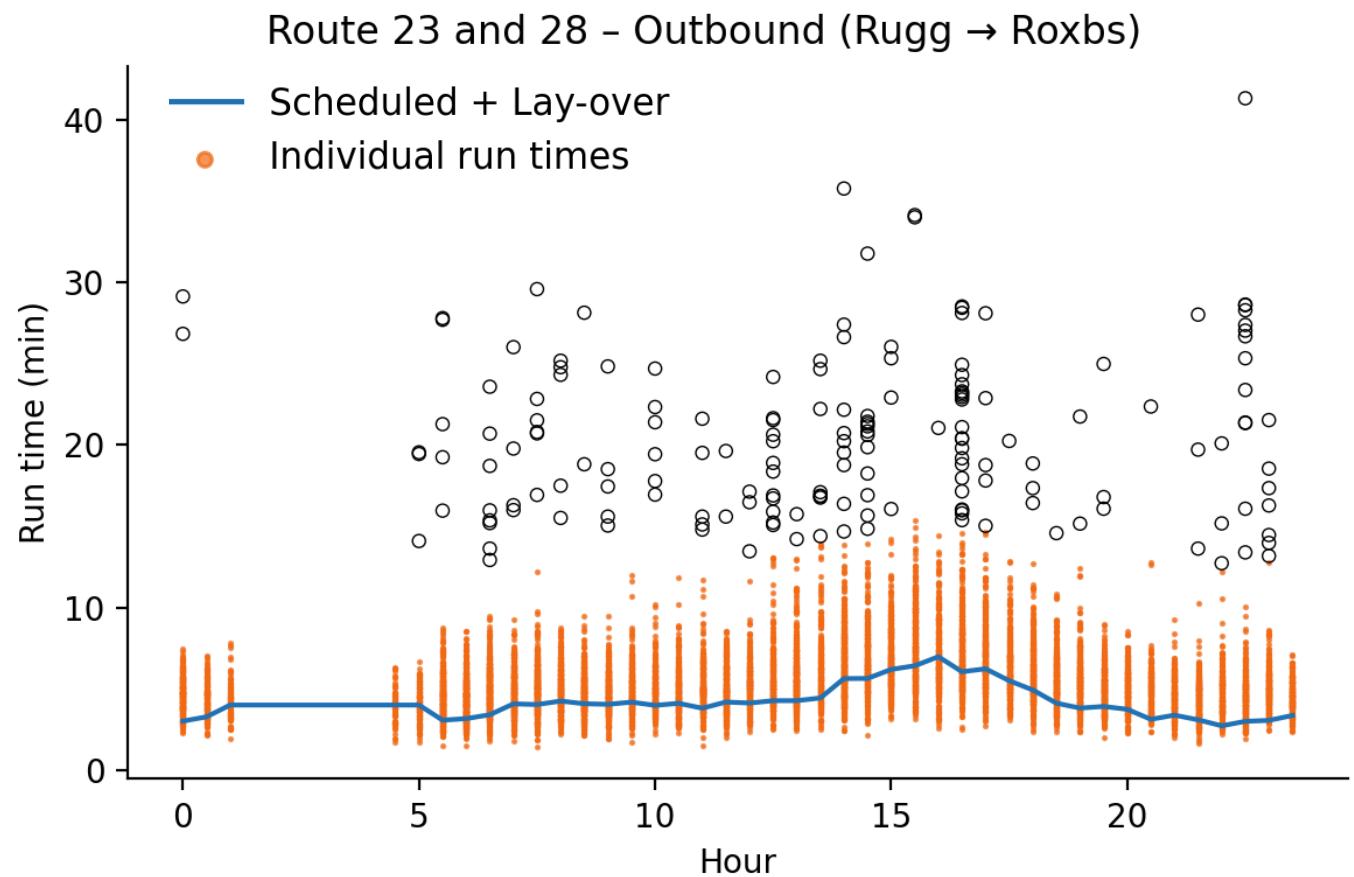
Calculate better



Calculate better

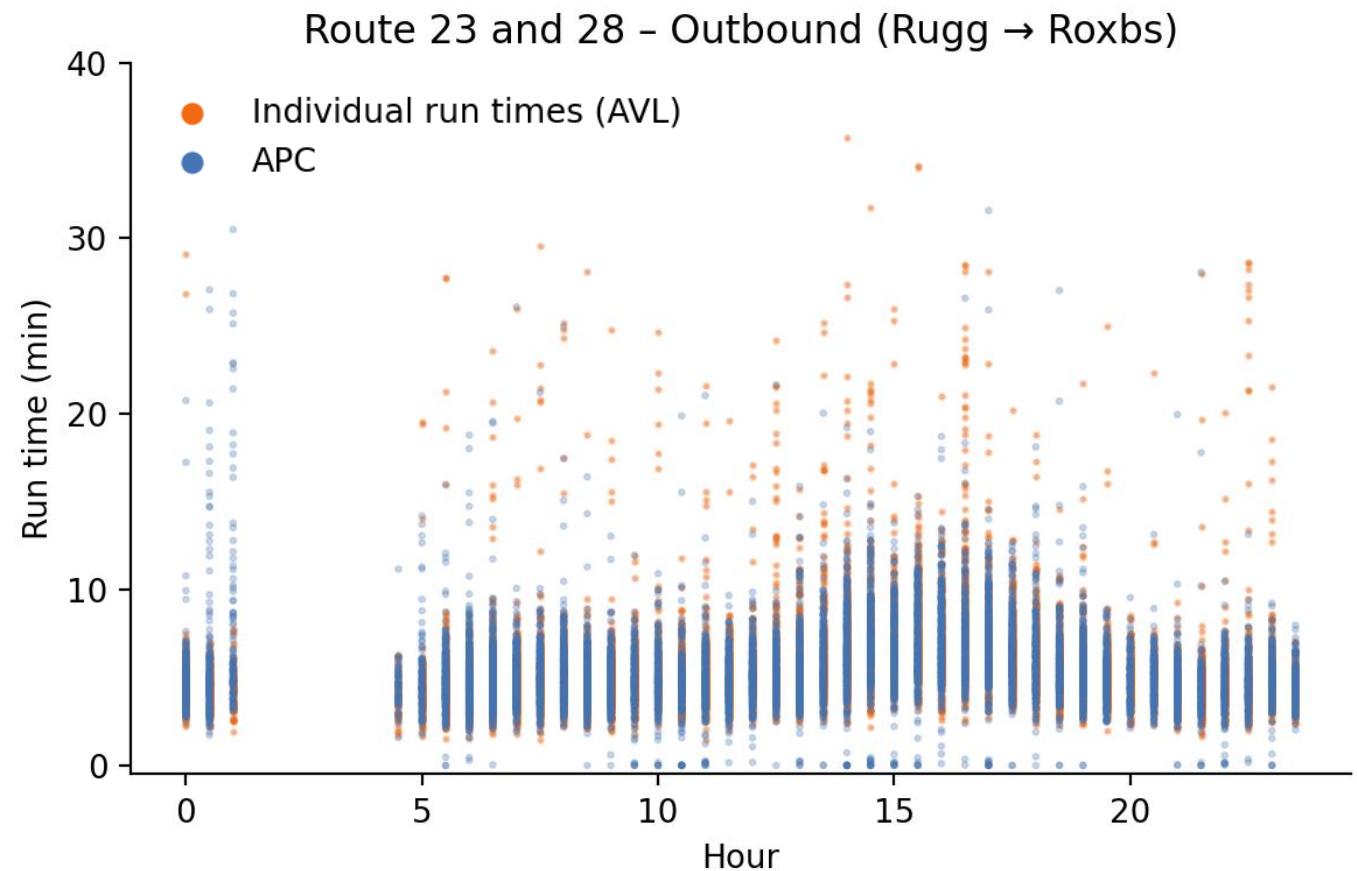
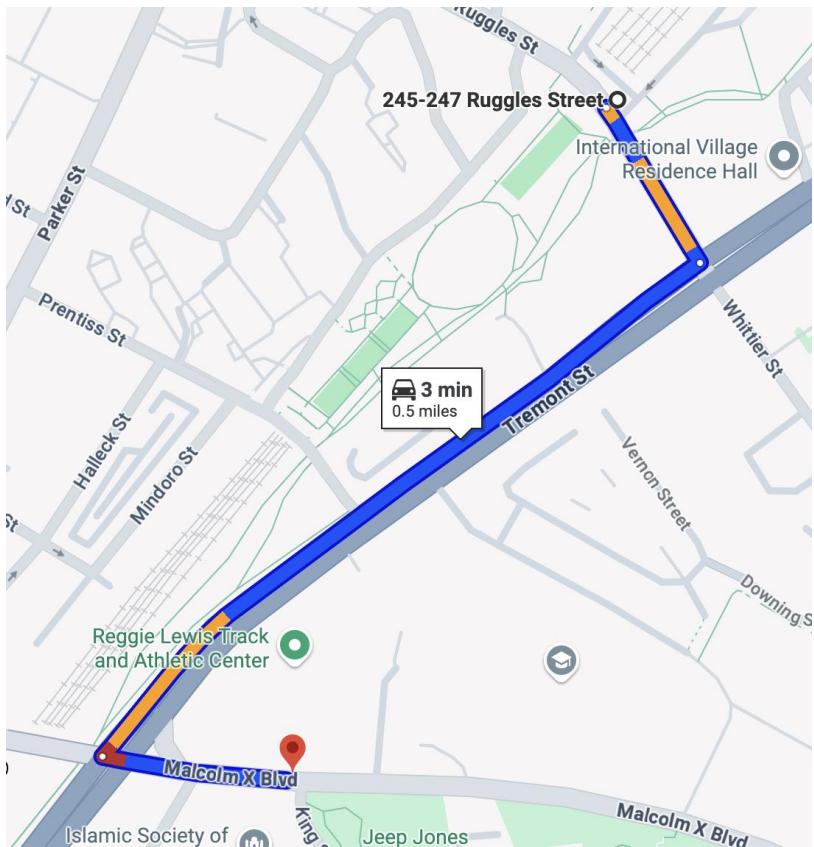
We are **scheduling** and **evaluating** with run times that are or not accurate

Data cleaning should be standardized

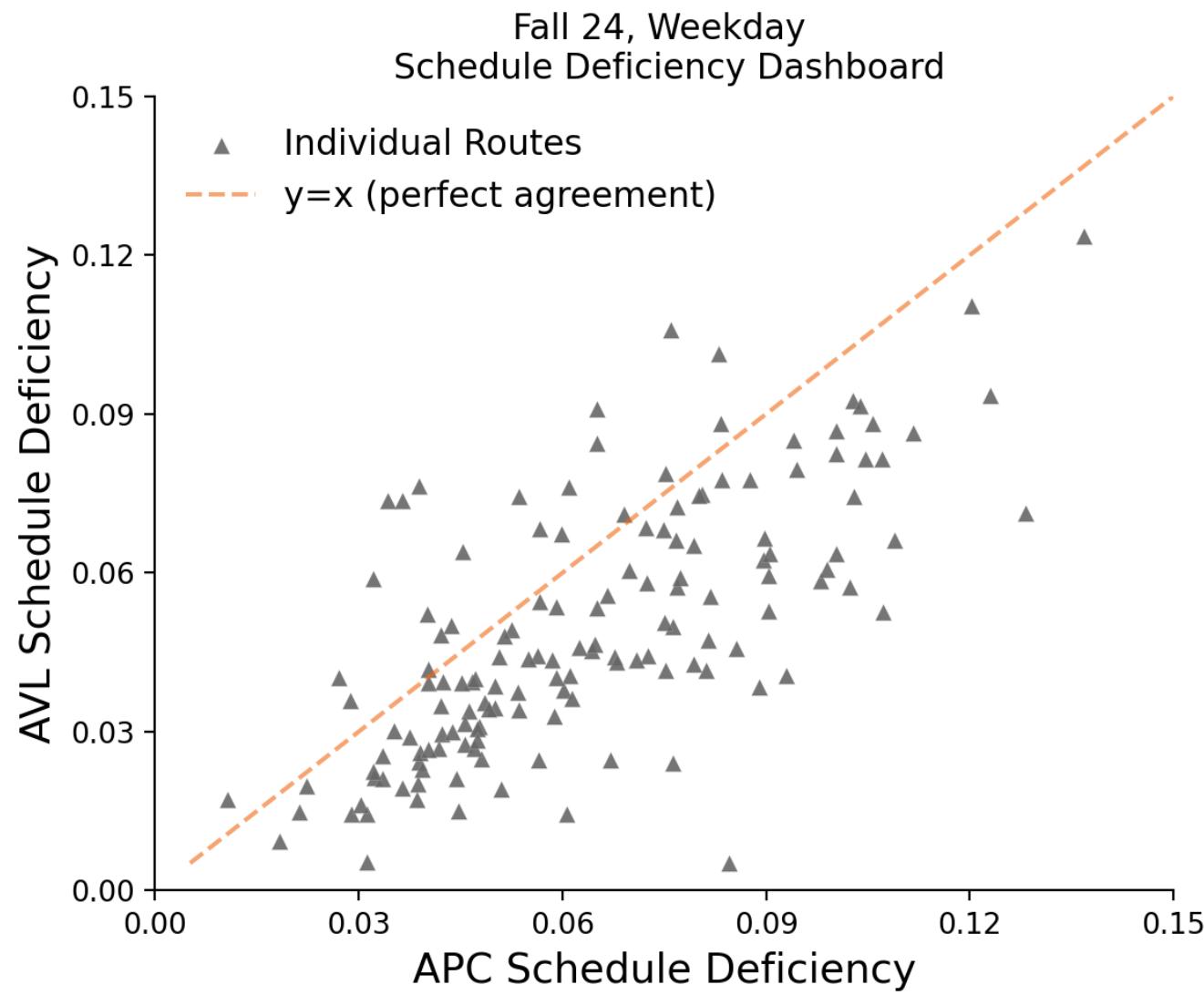


Calculate better

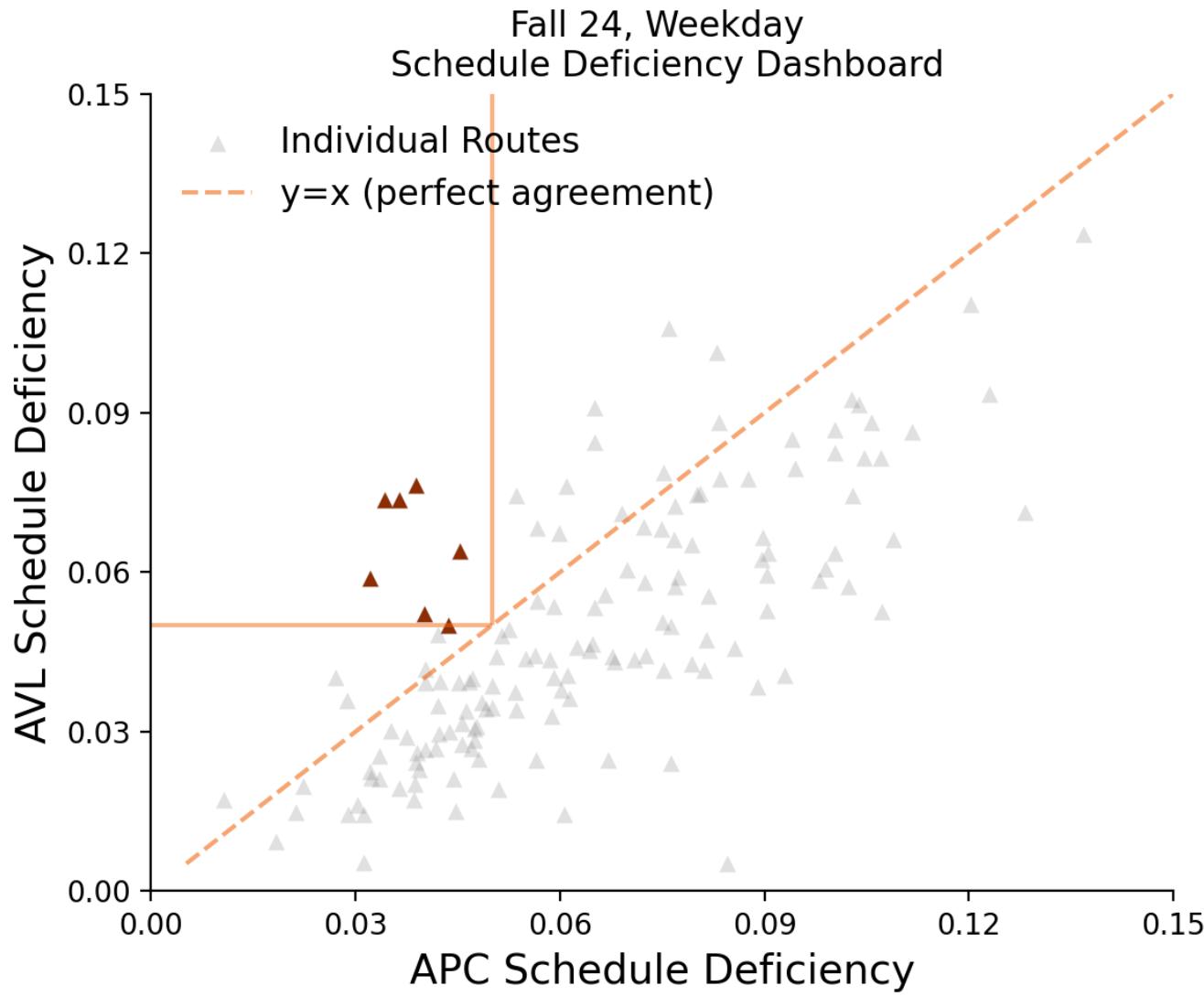
AVL-APC Mismatch



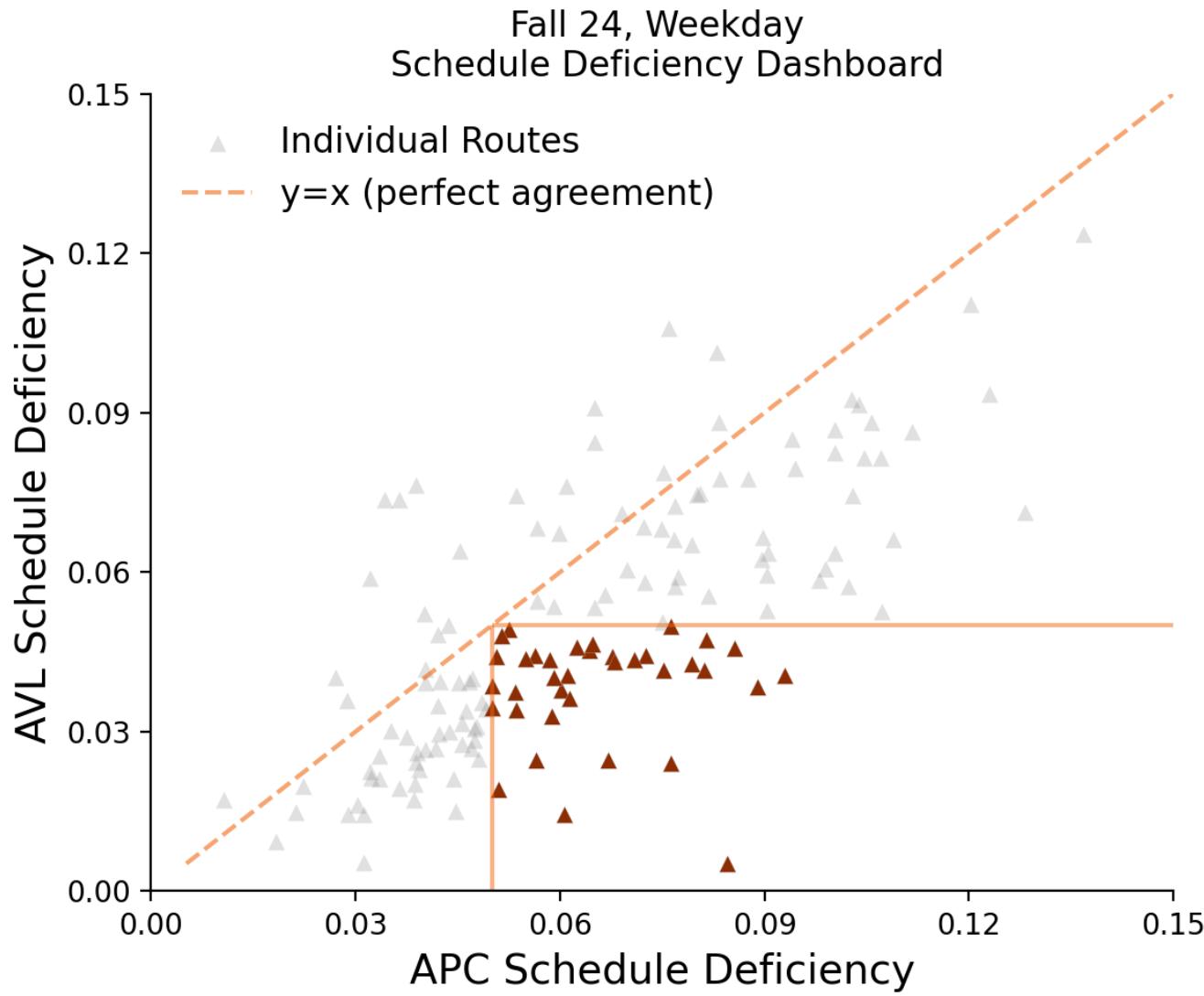
Calculate better



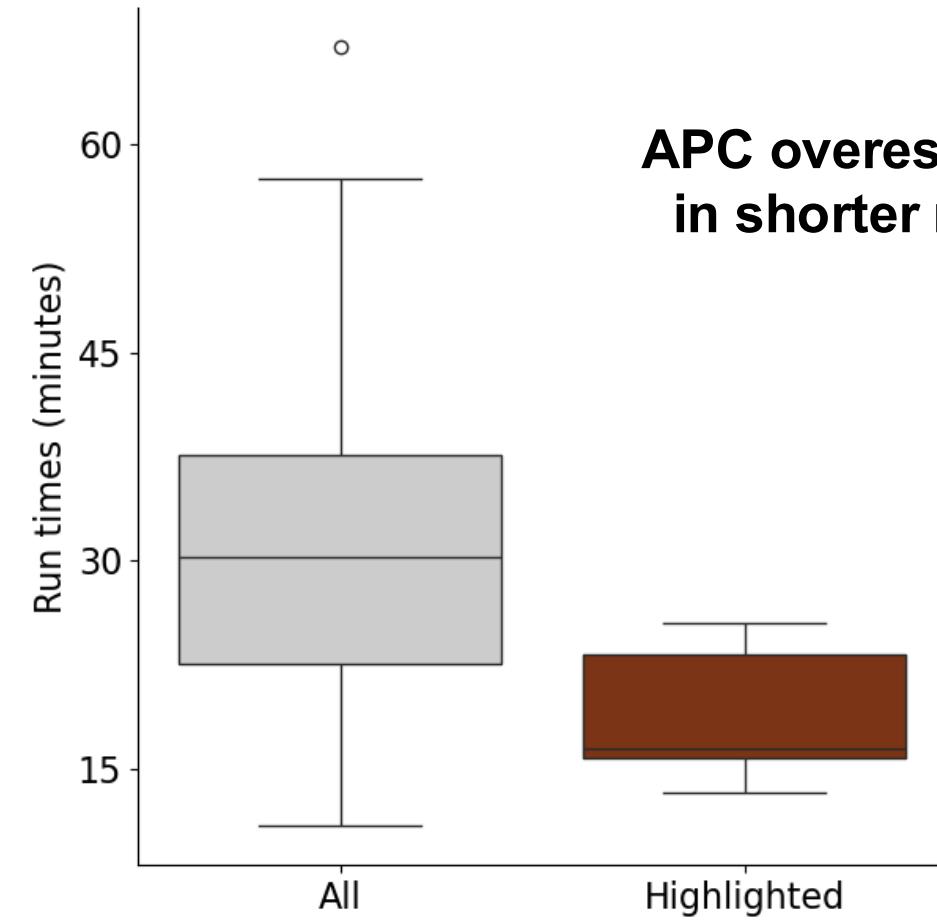
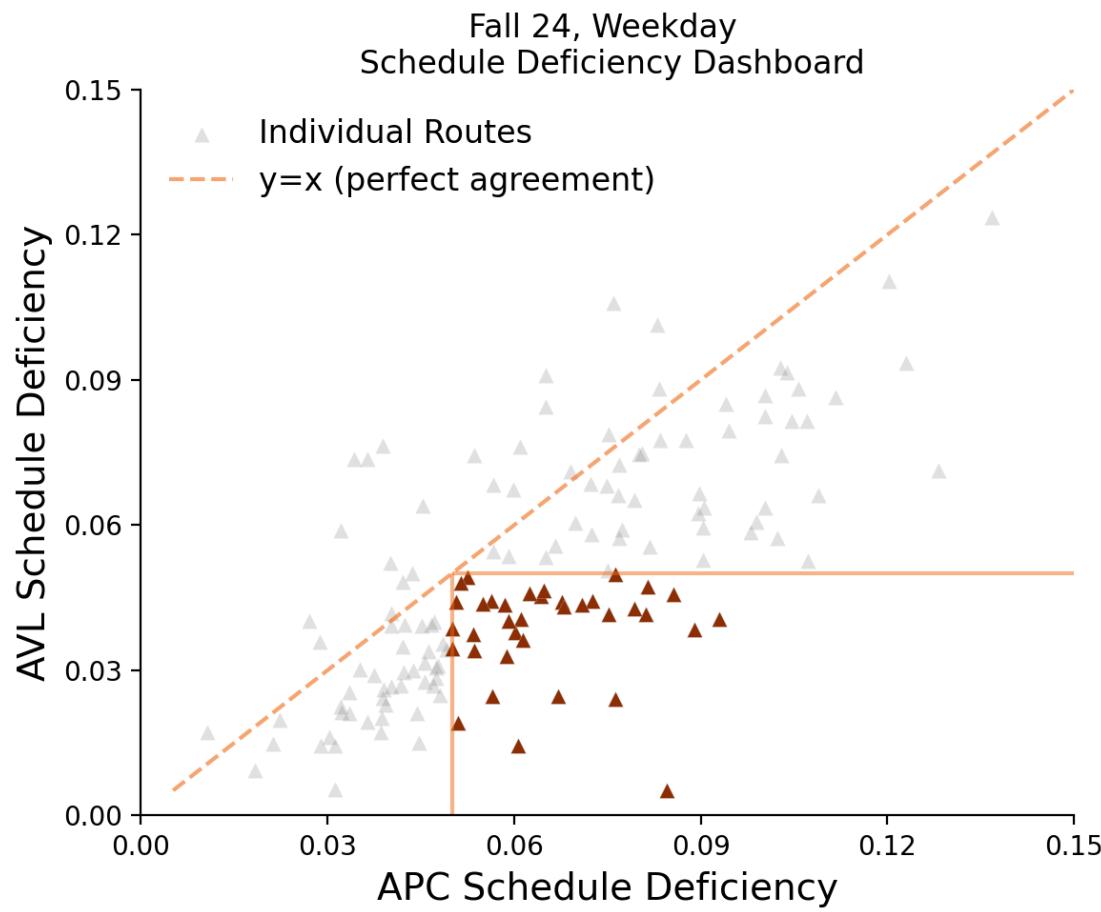
Calculate better



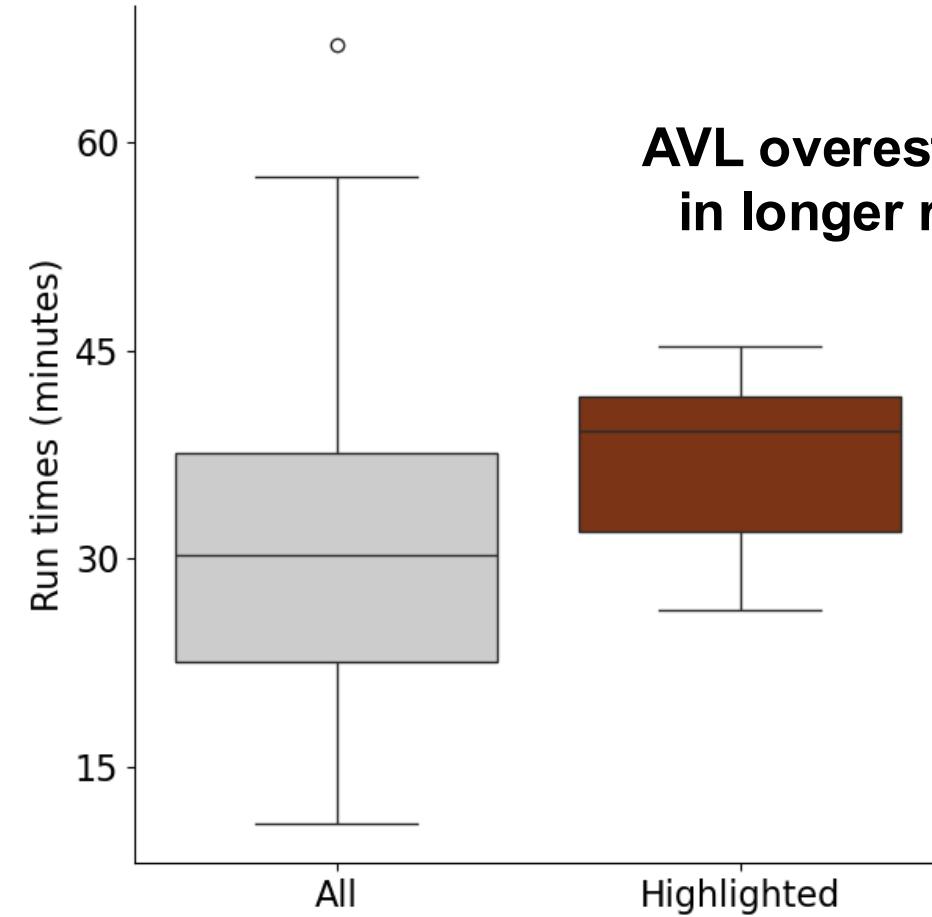
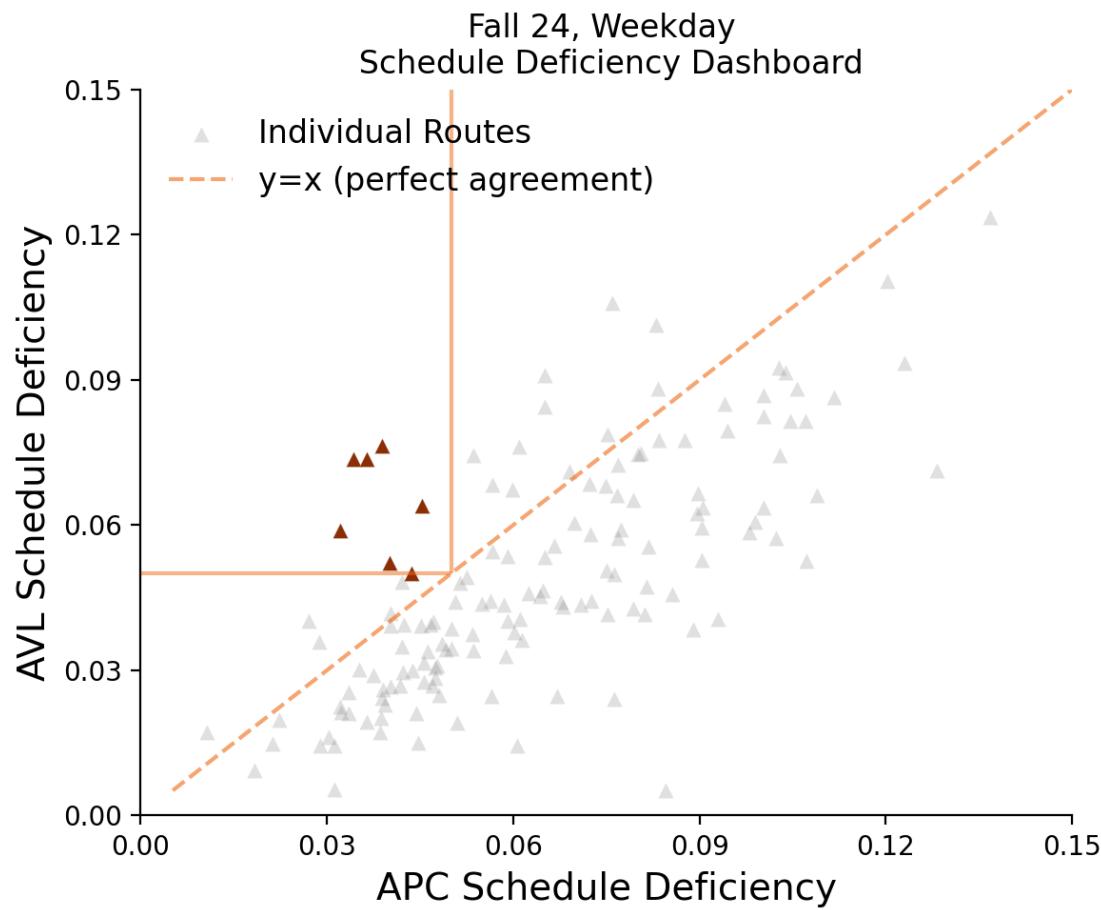
Calculate better



Calculate better



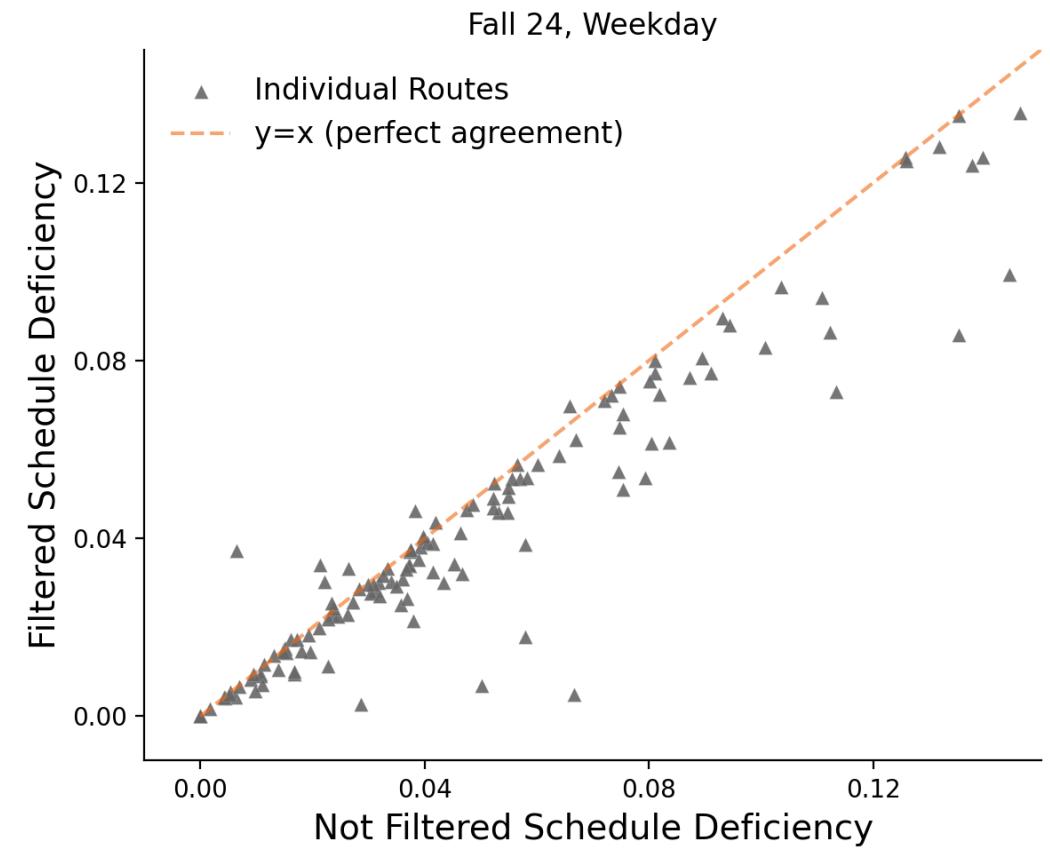
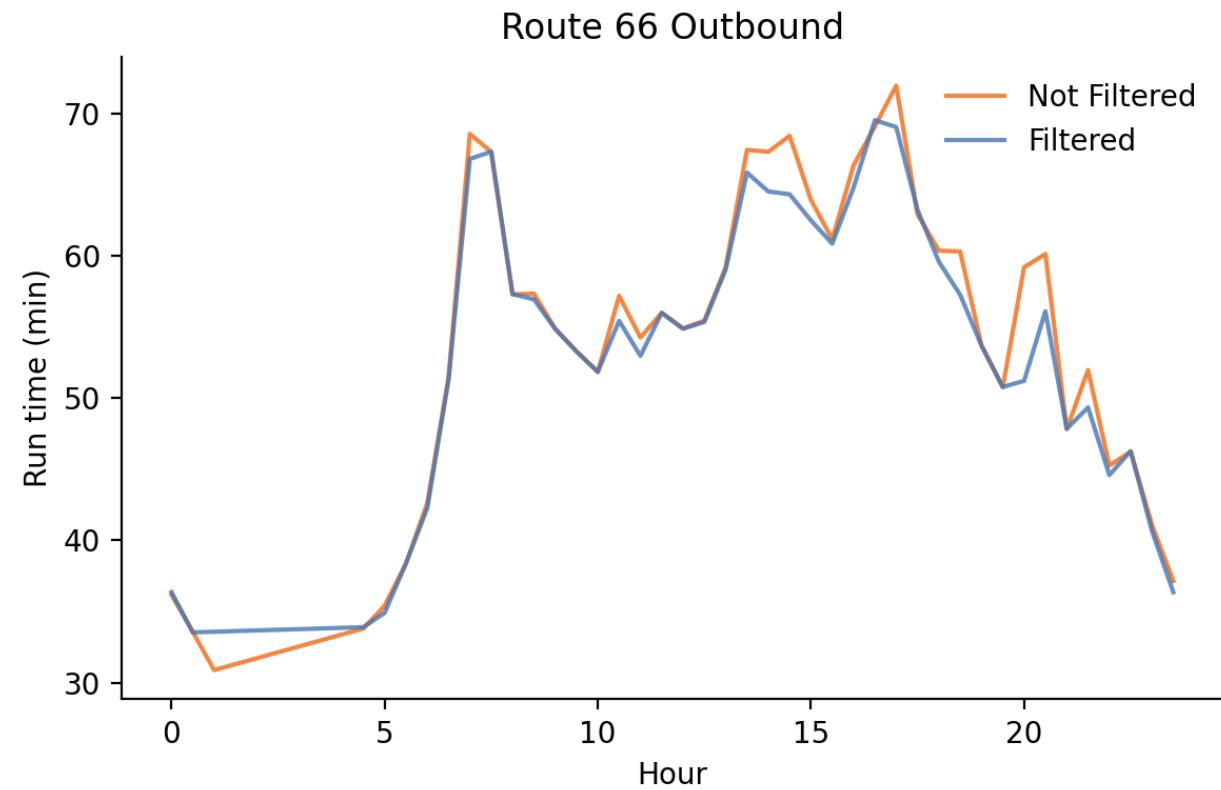
Calculate better



Calculate better

Methodology Differences

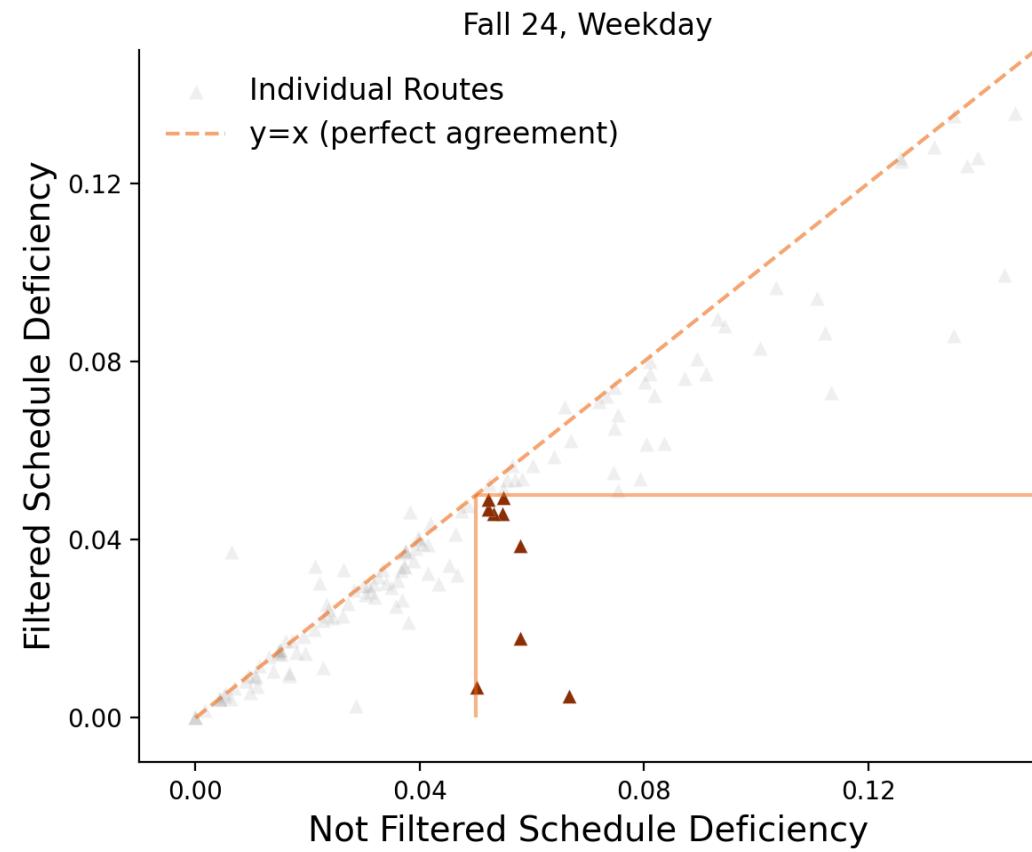
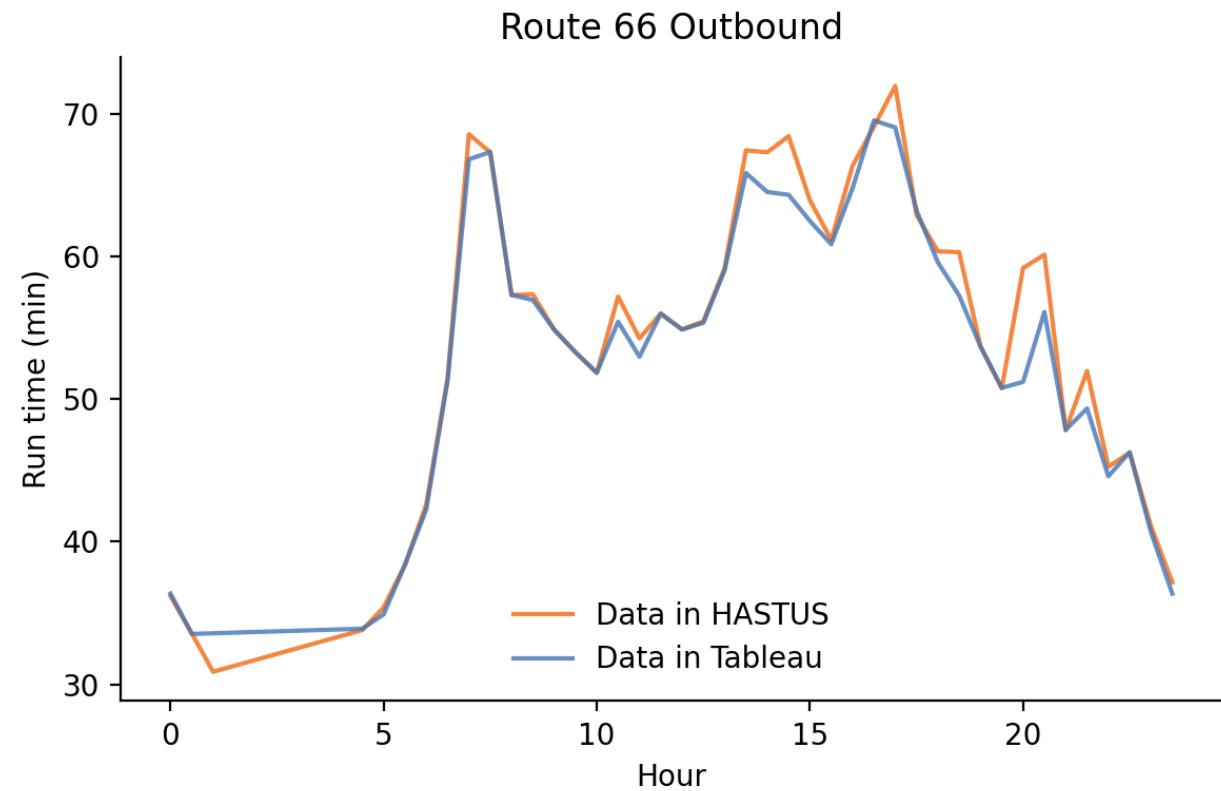
While calculating **we filter trips** started 15 minutes late or 10 minutes early



Calculate better

Methodology Differences

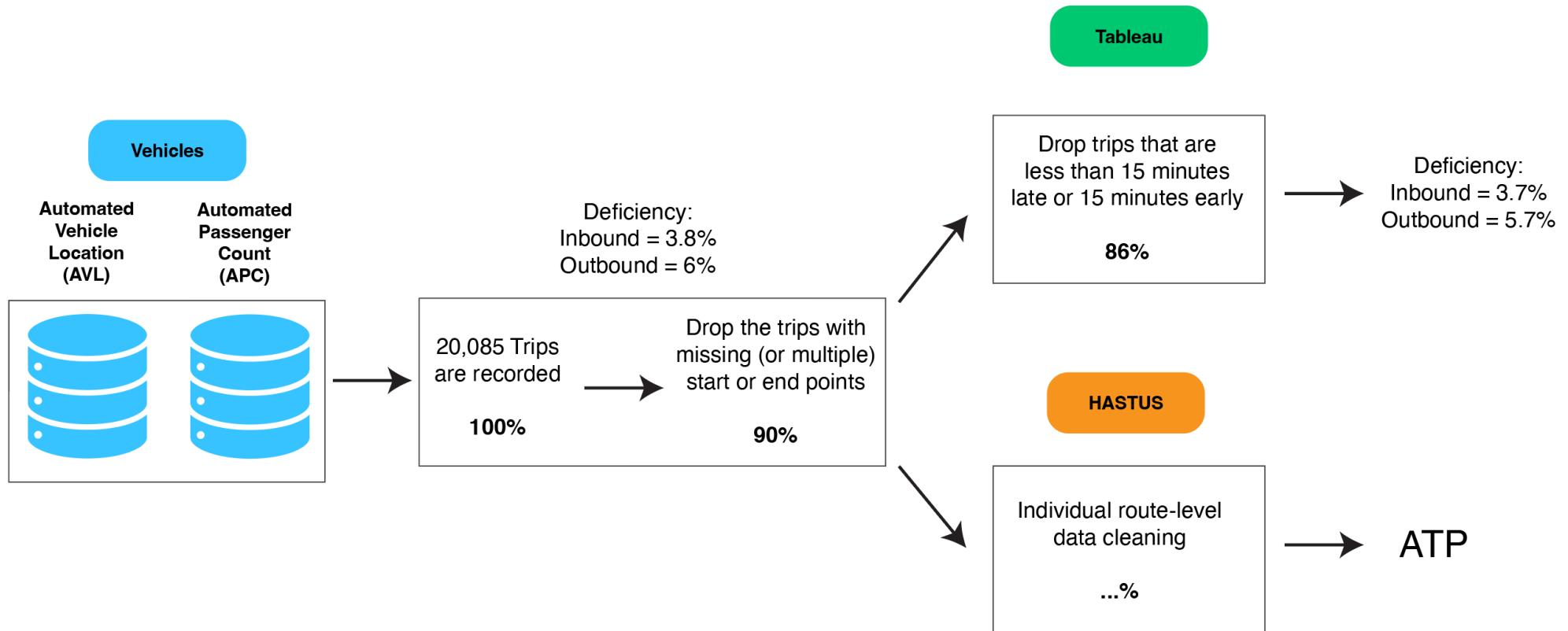
While calculating **we filter trips** started 15 minutes late or 10 minutes early



Calculate better

Data we evaluate = Data we schedule

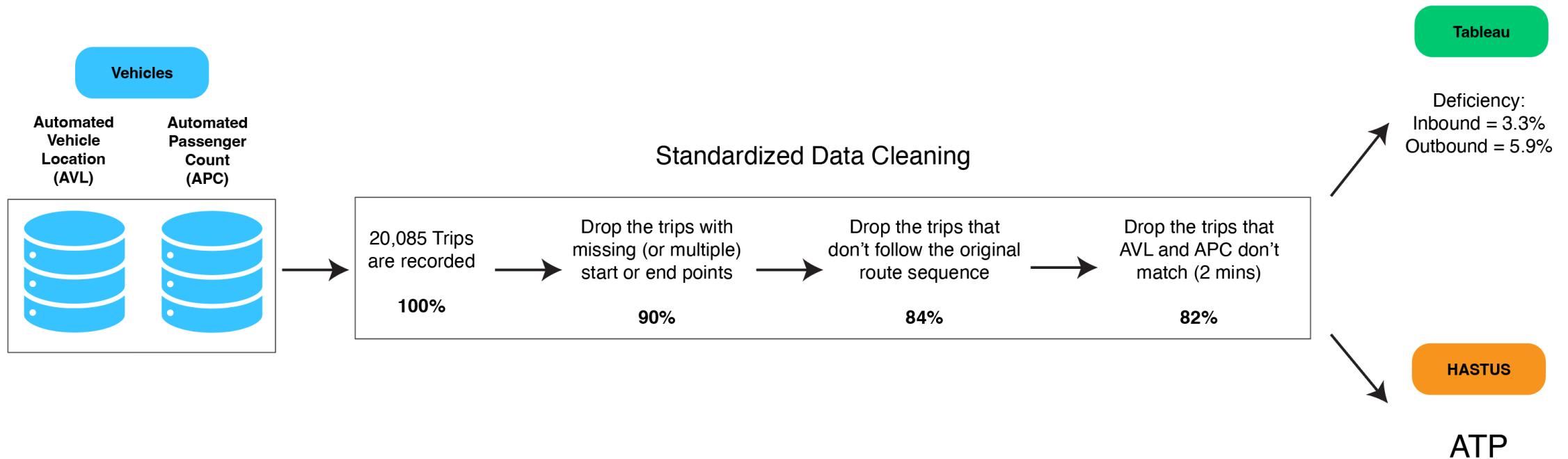
Route 22 – Fall 2024 Example (Existing)



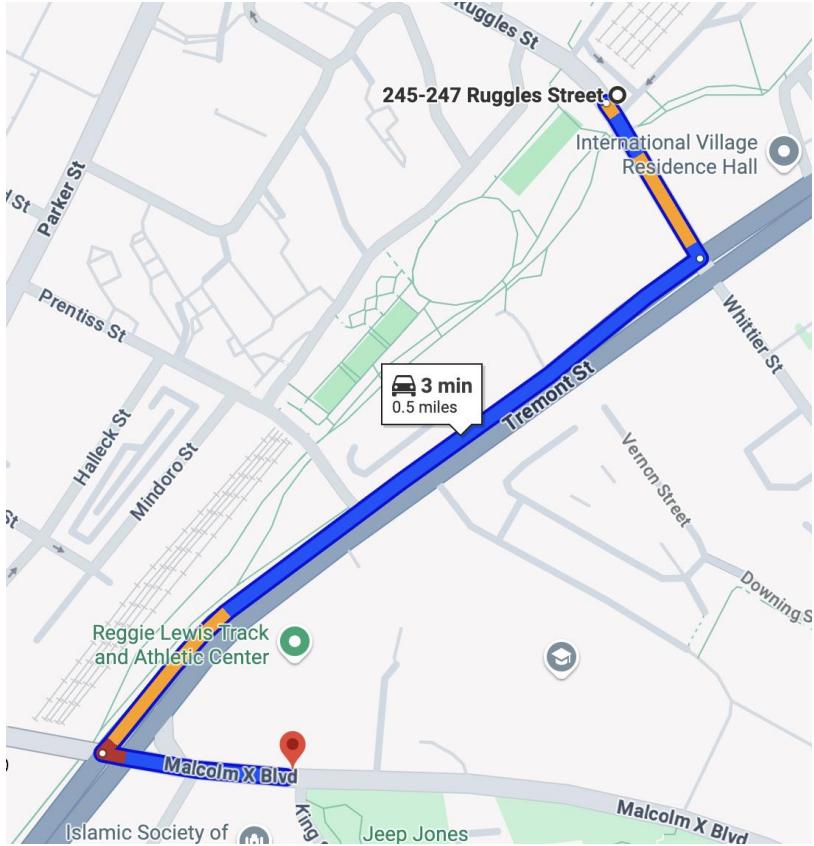
Calculate better

Data we evaluate = Data we schedule

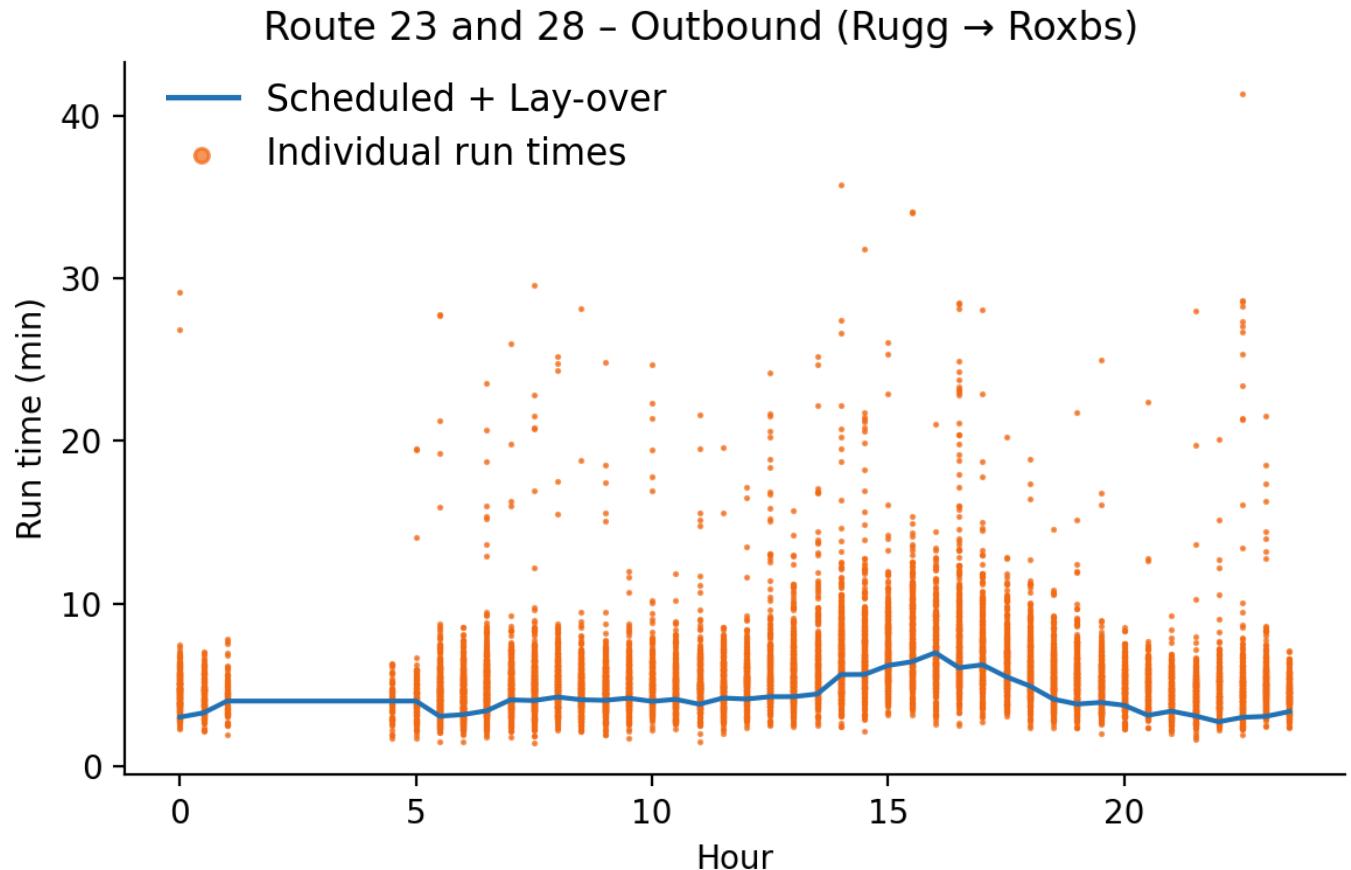
Route 22 – Fall 2024 Example (Proposed)



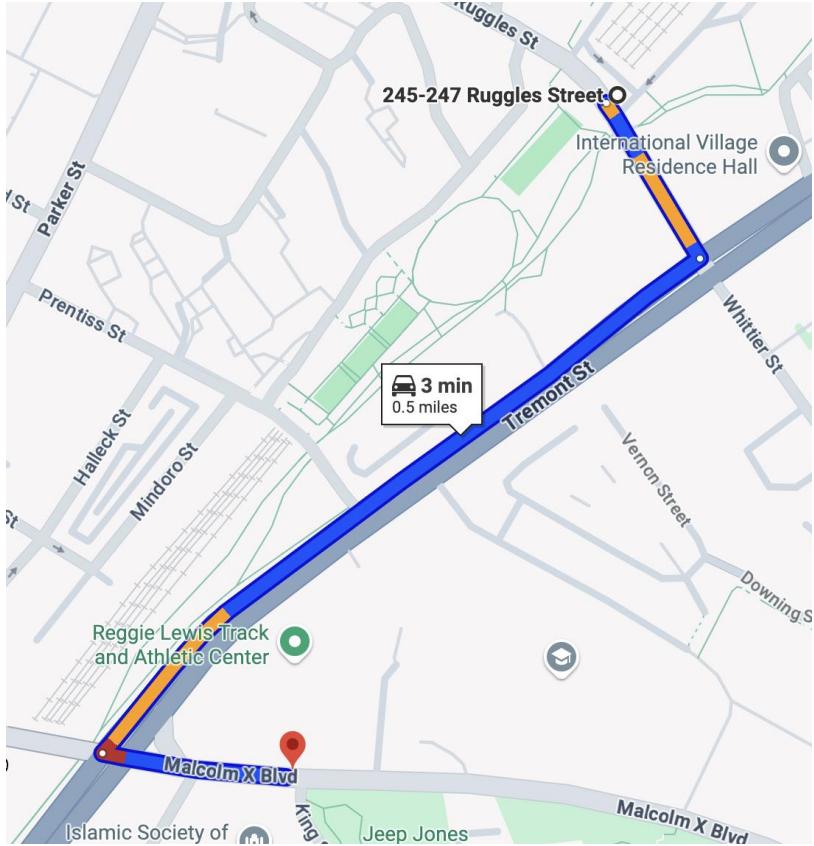
Calculate better



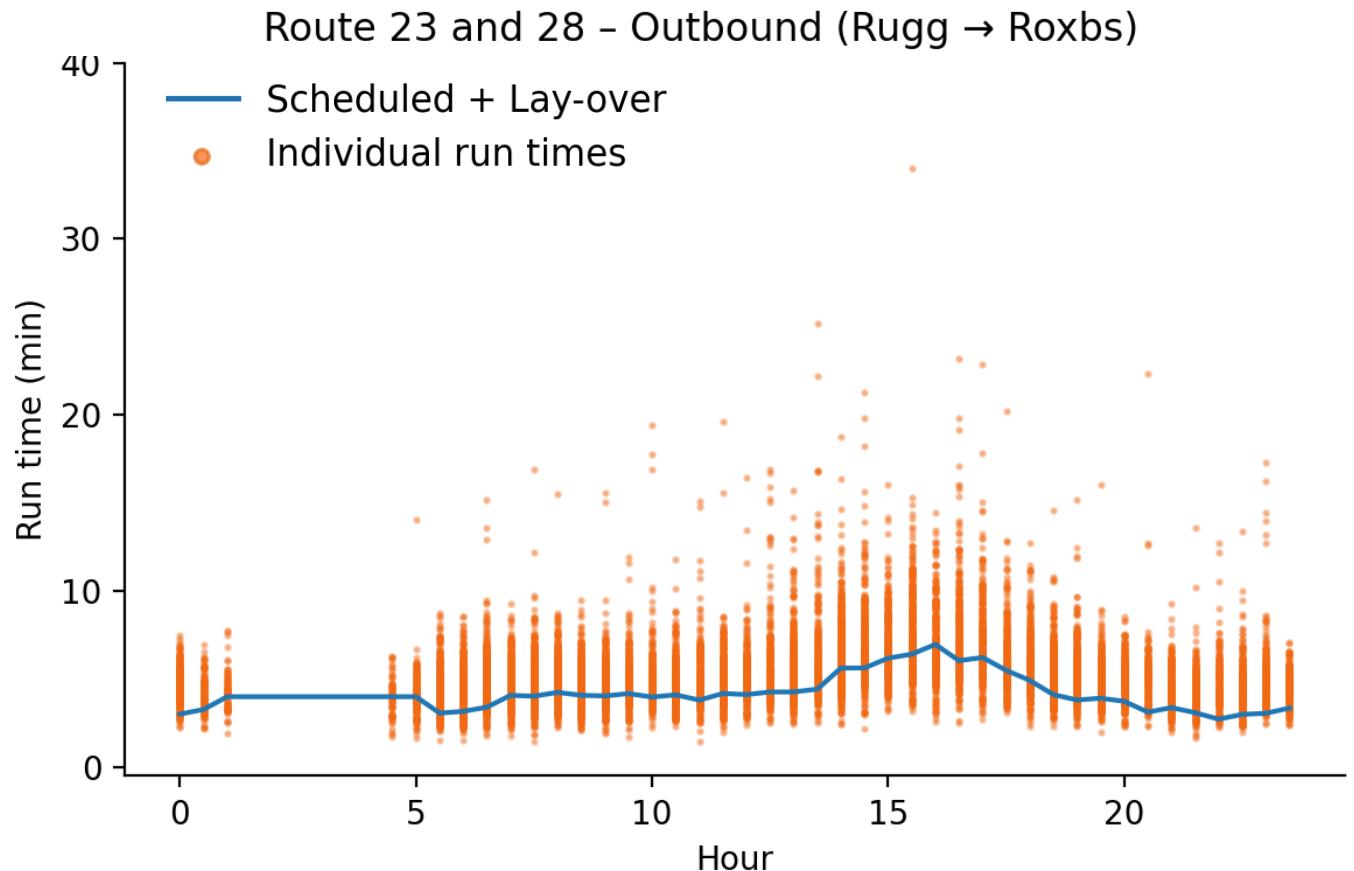
Before Cleaning (HASTUS)



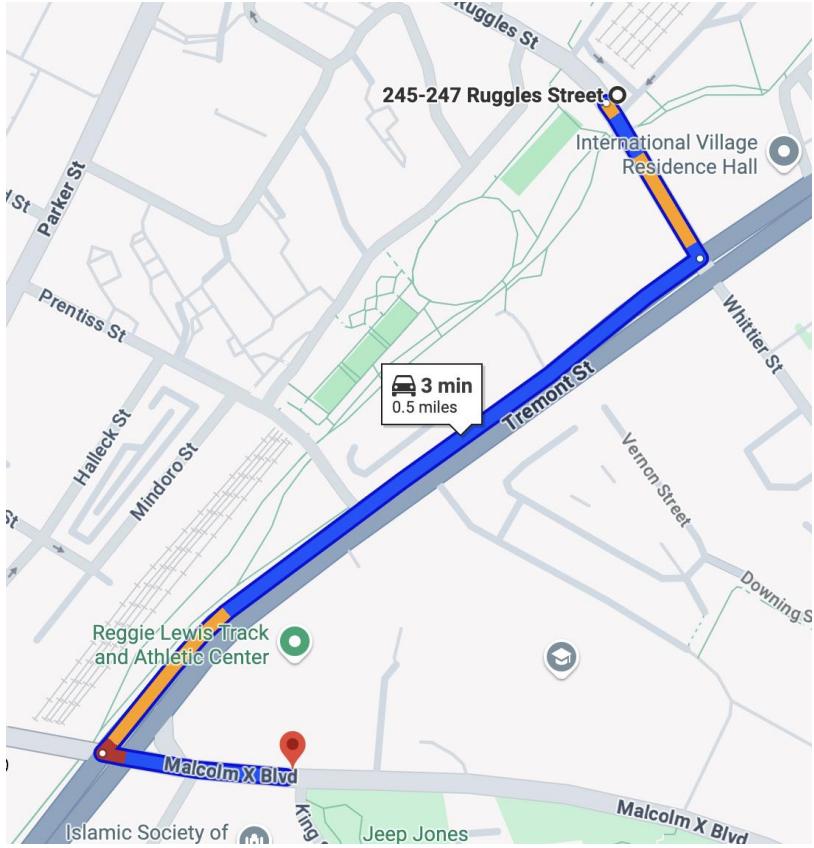
Calculate better



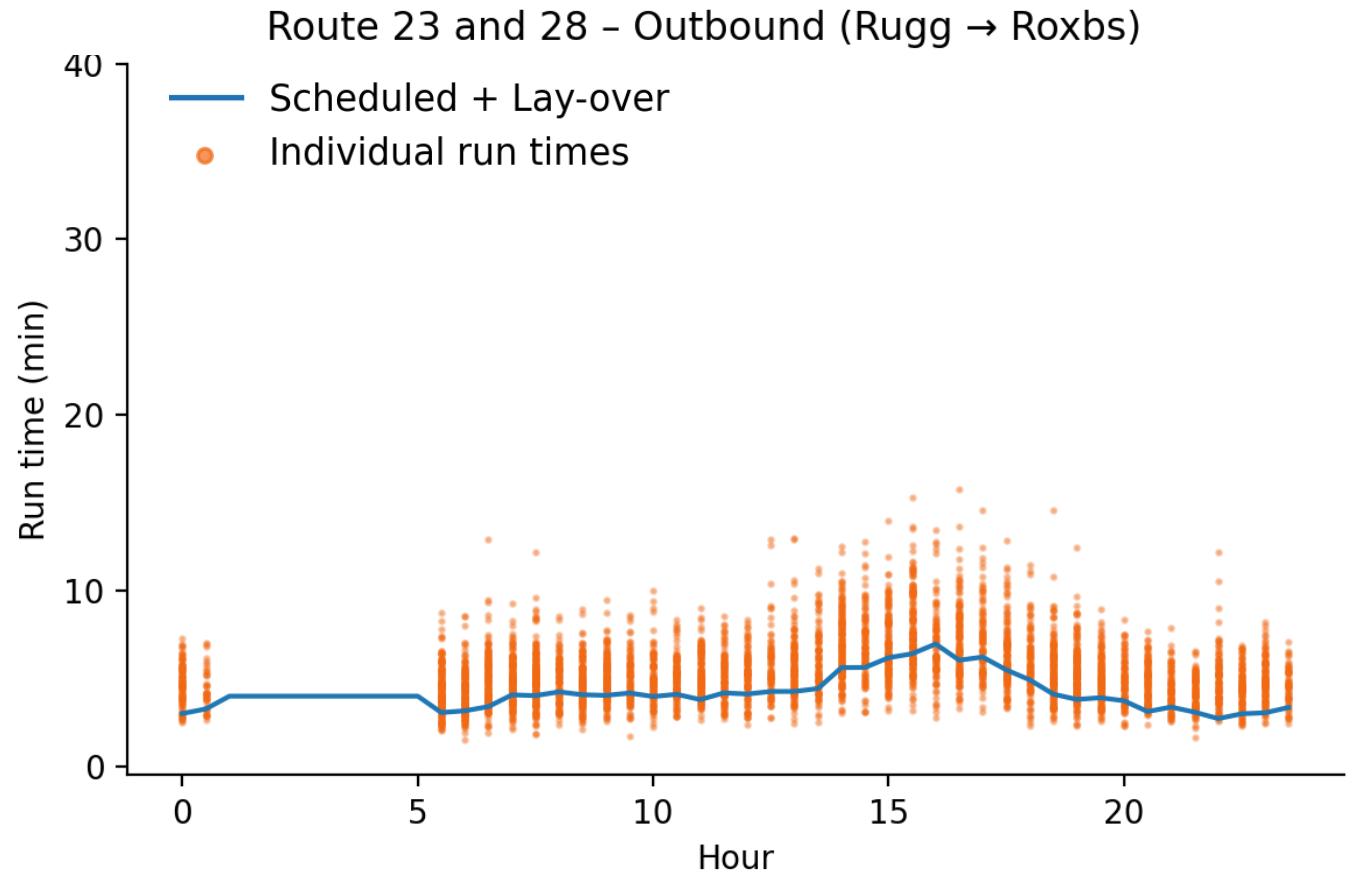
Current Cleaning (Tableau)



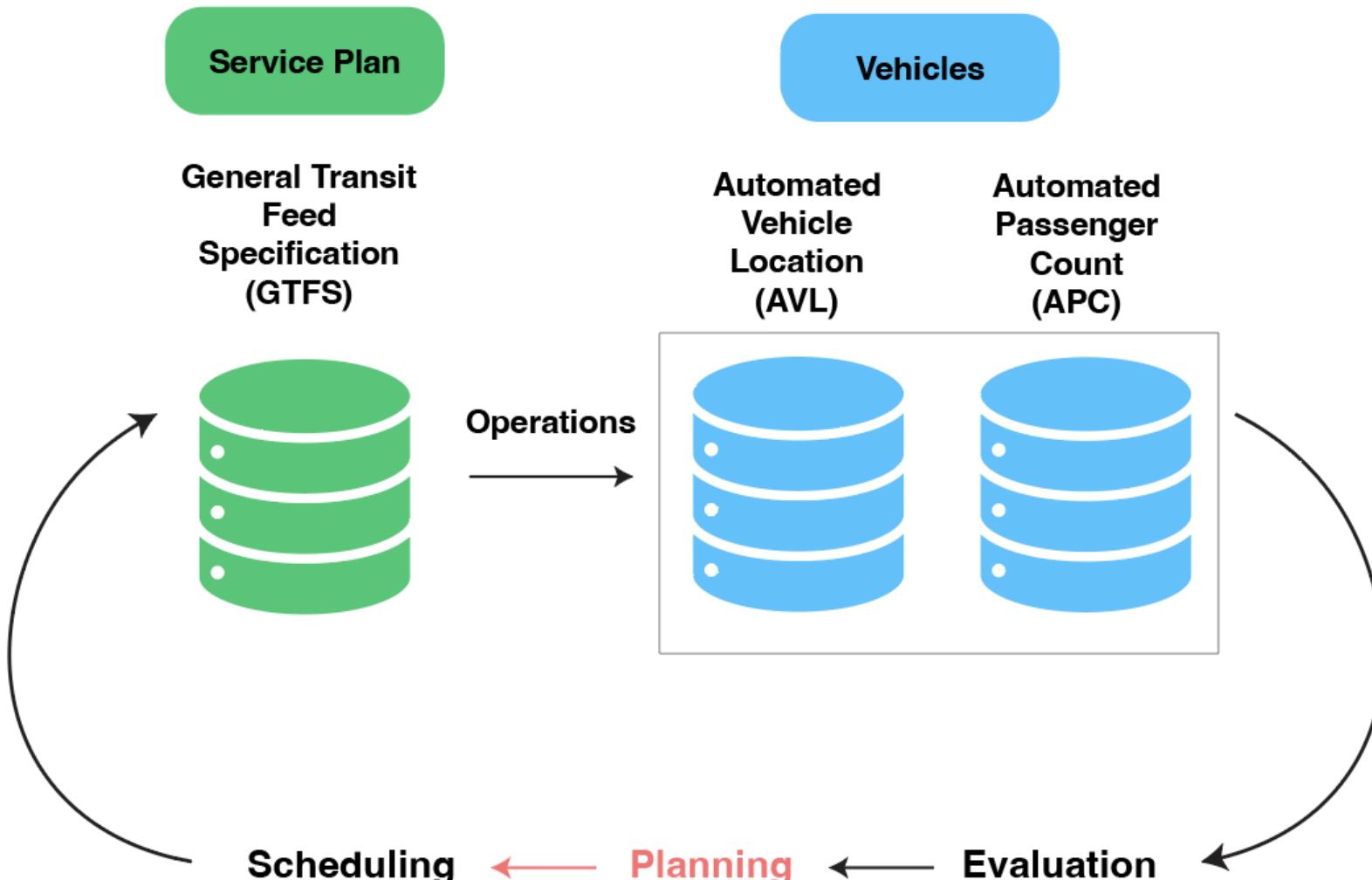
Calculate better



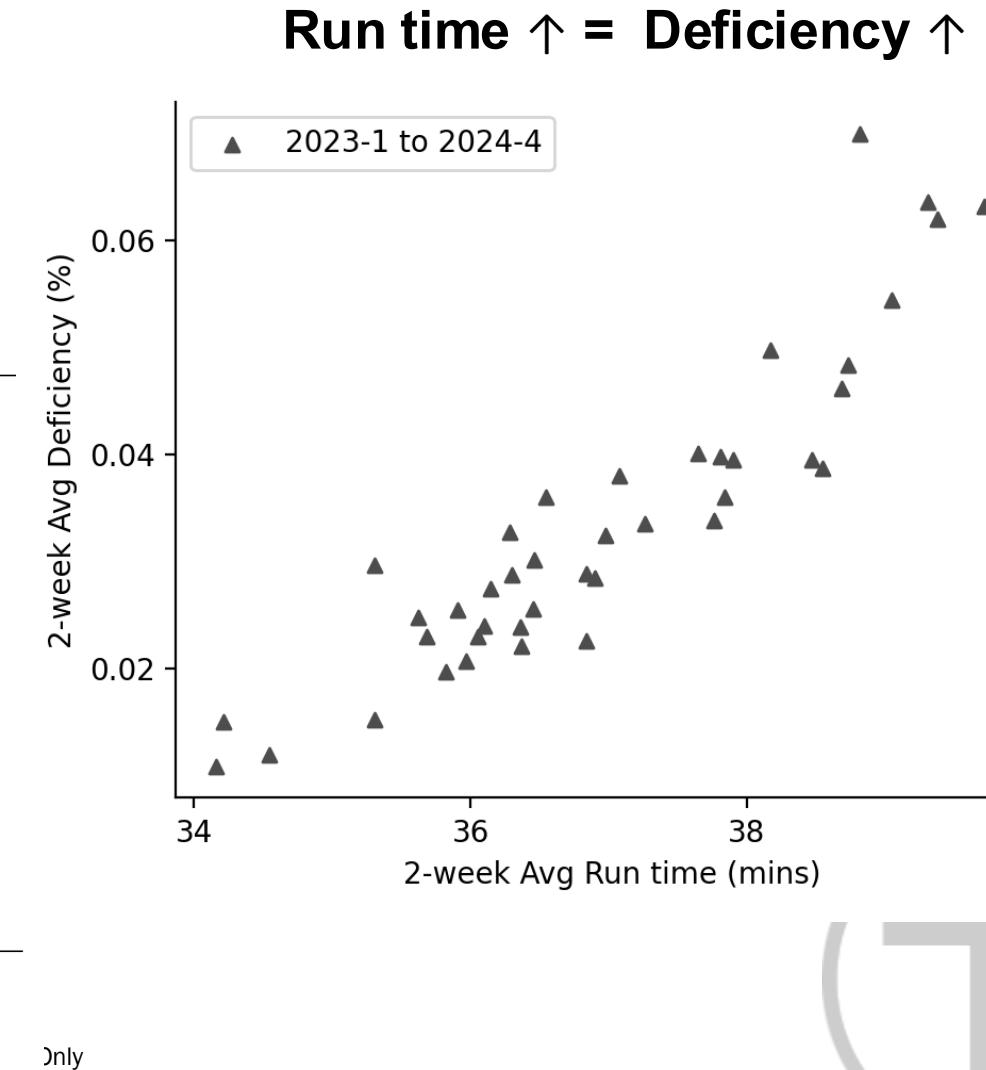
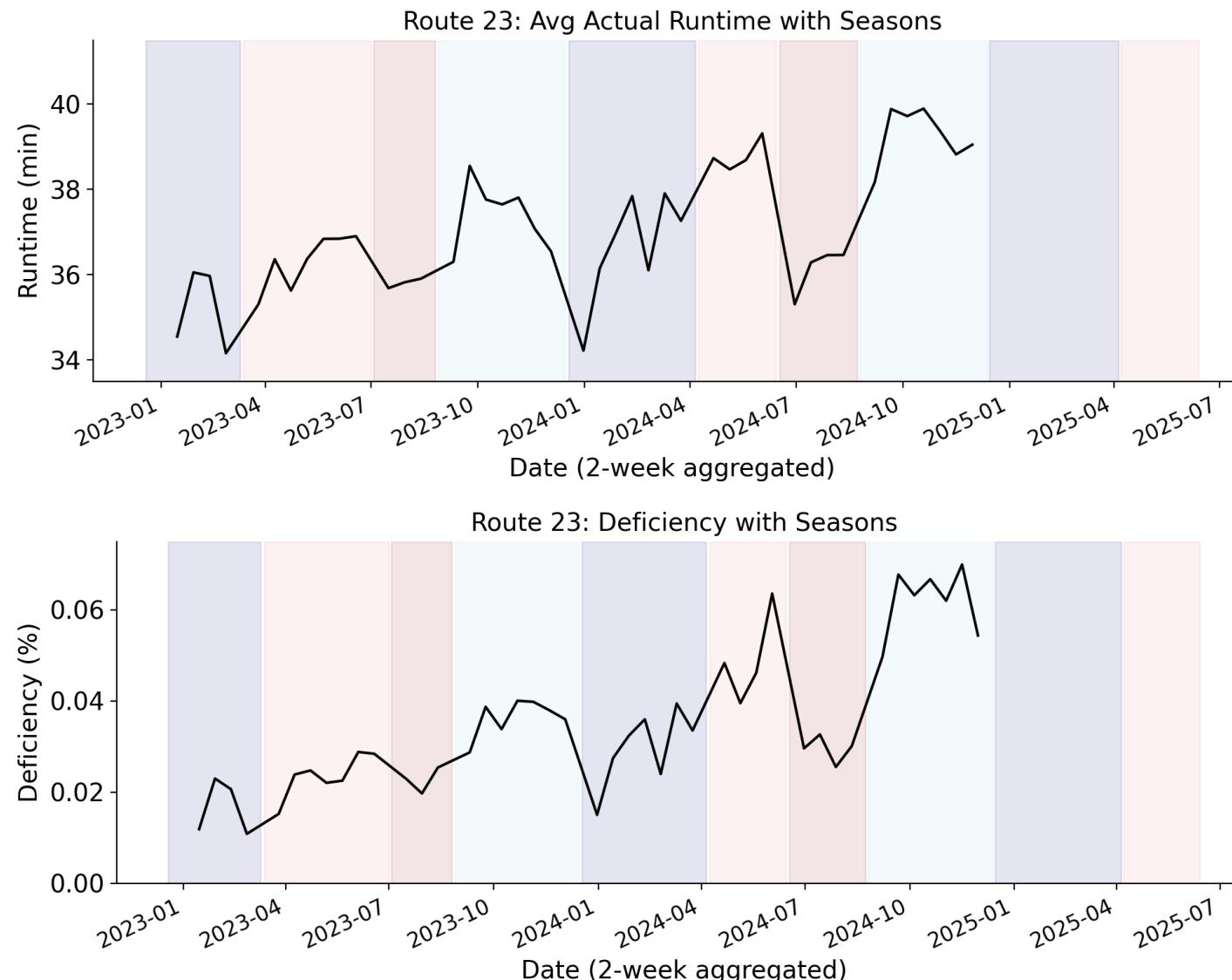
After Cleaning (Proposed)



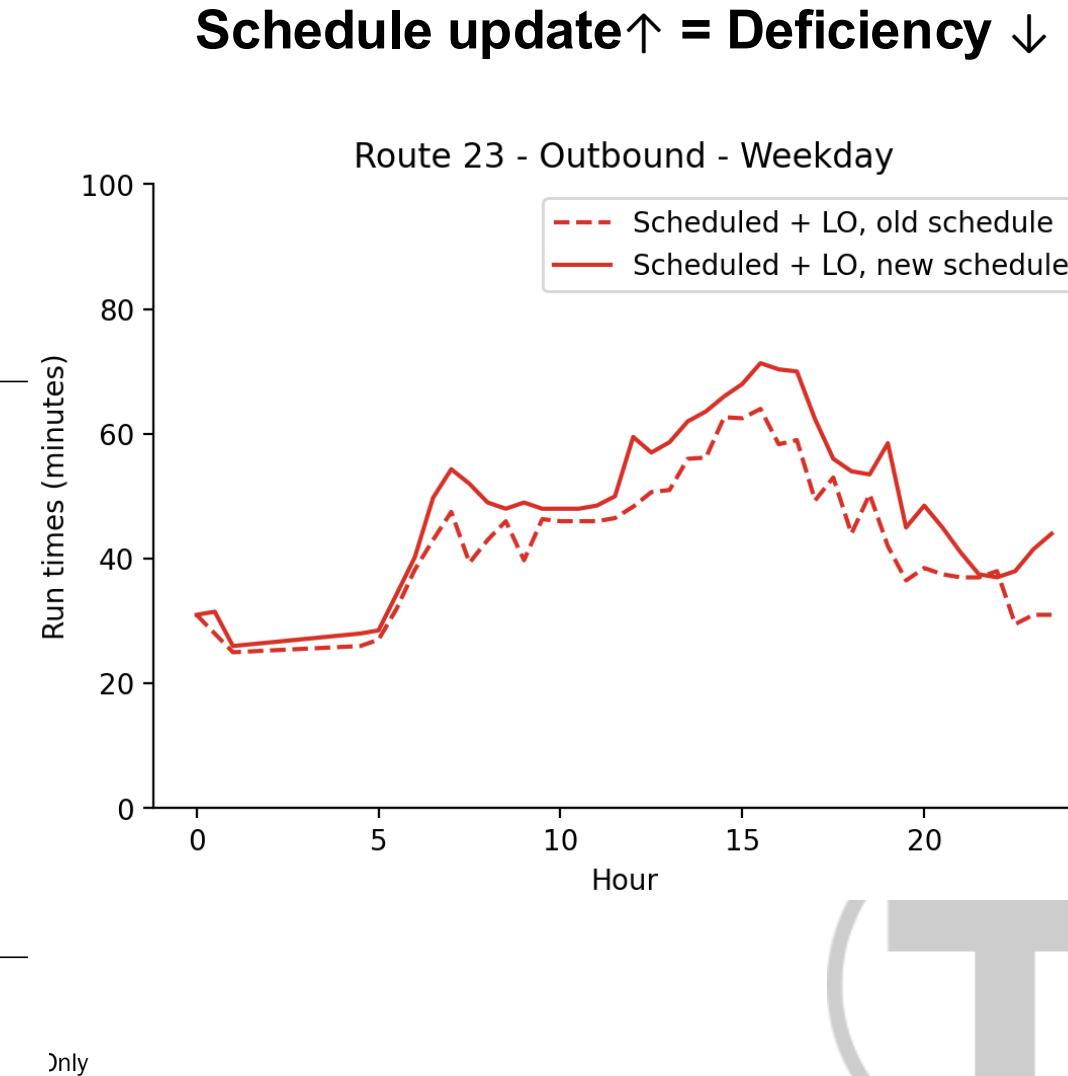
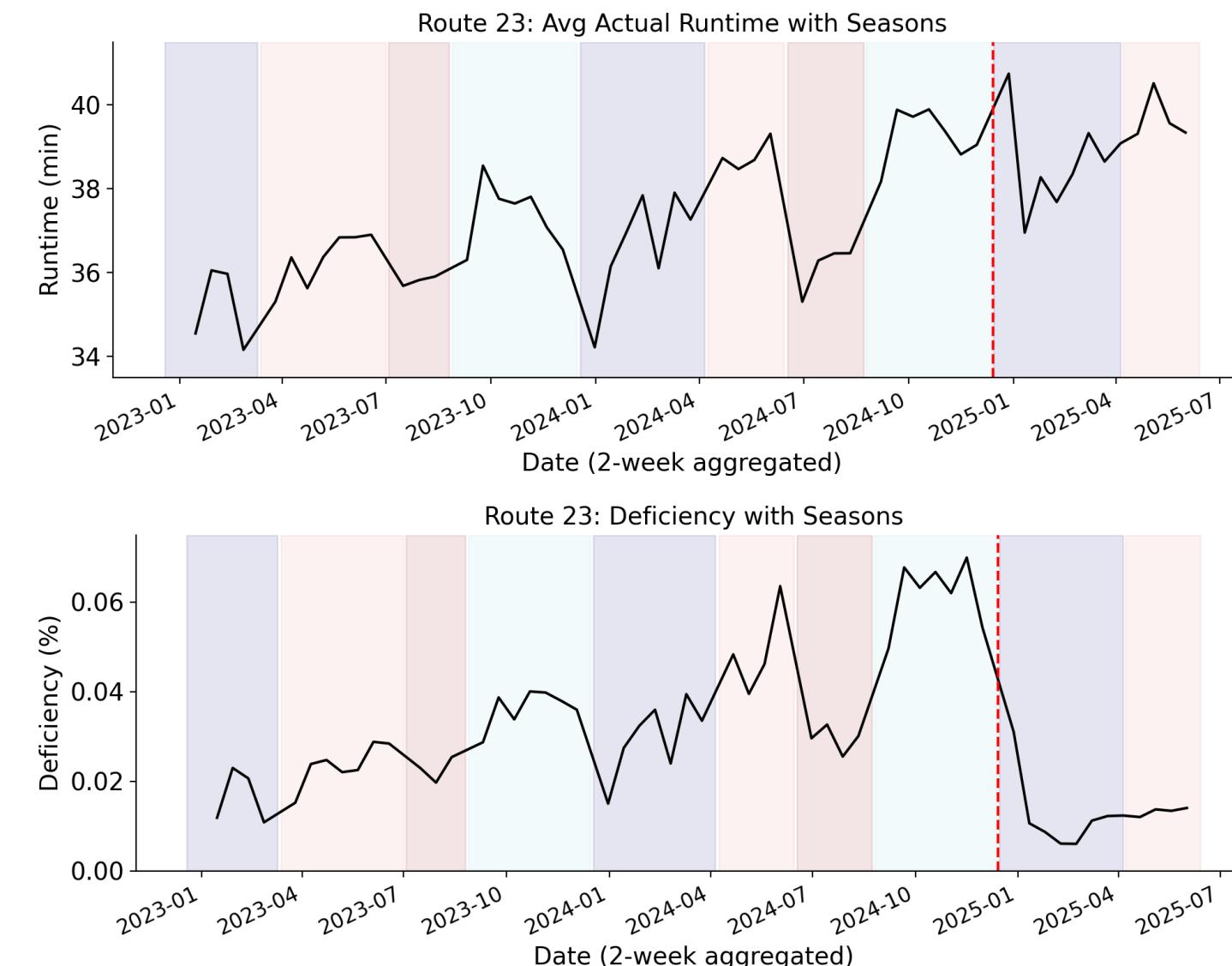
2. Understand better



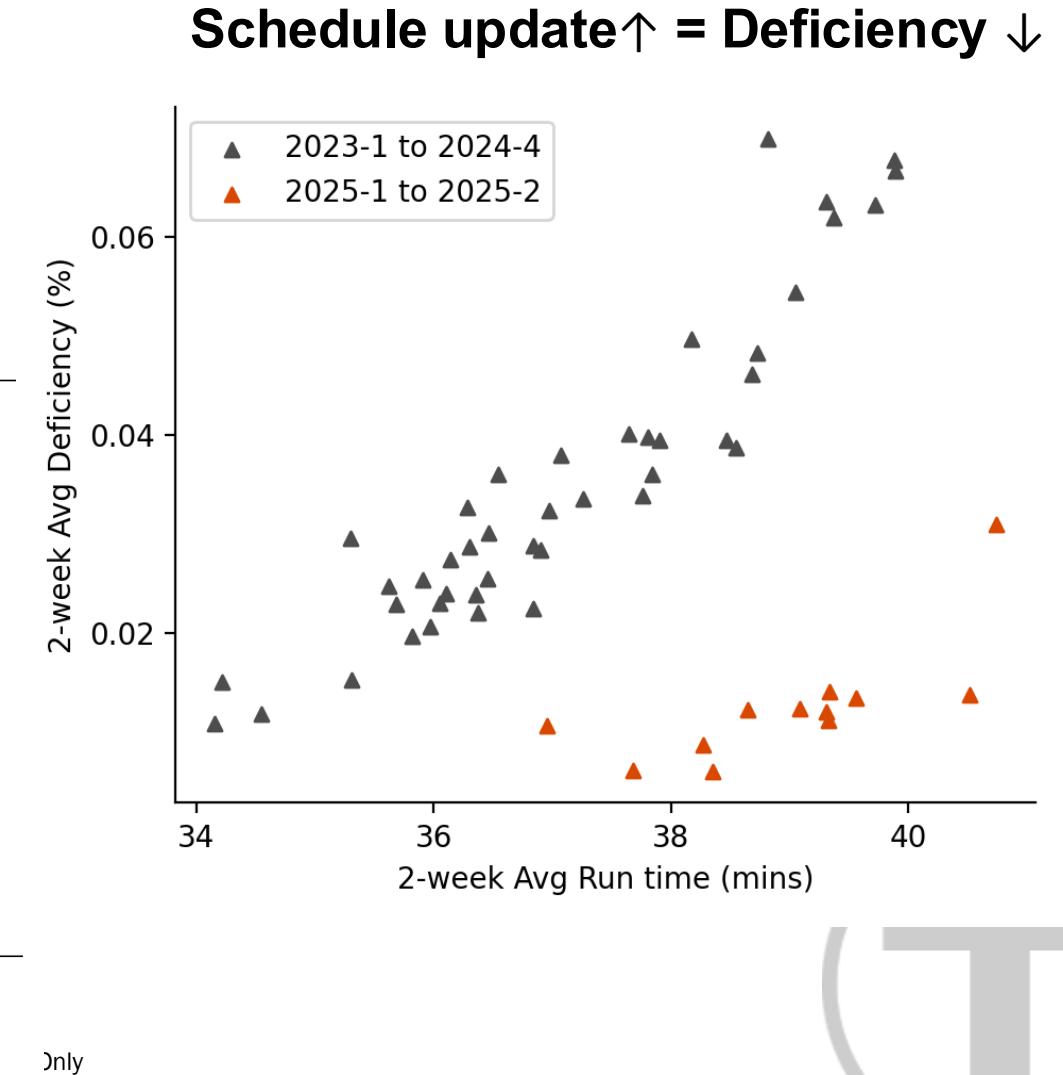
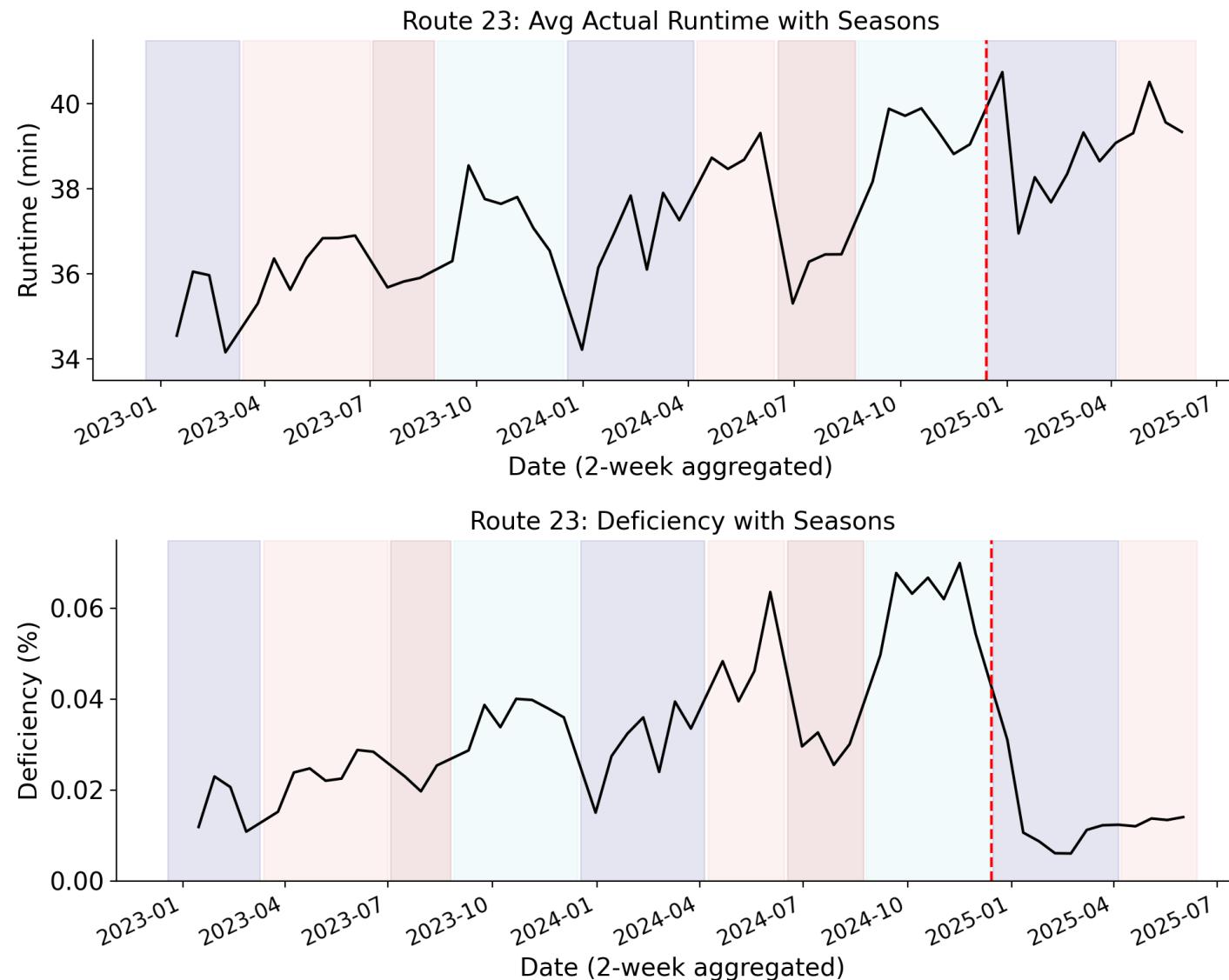
Understand better



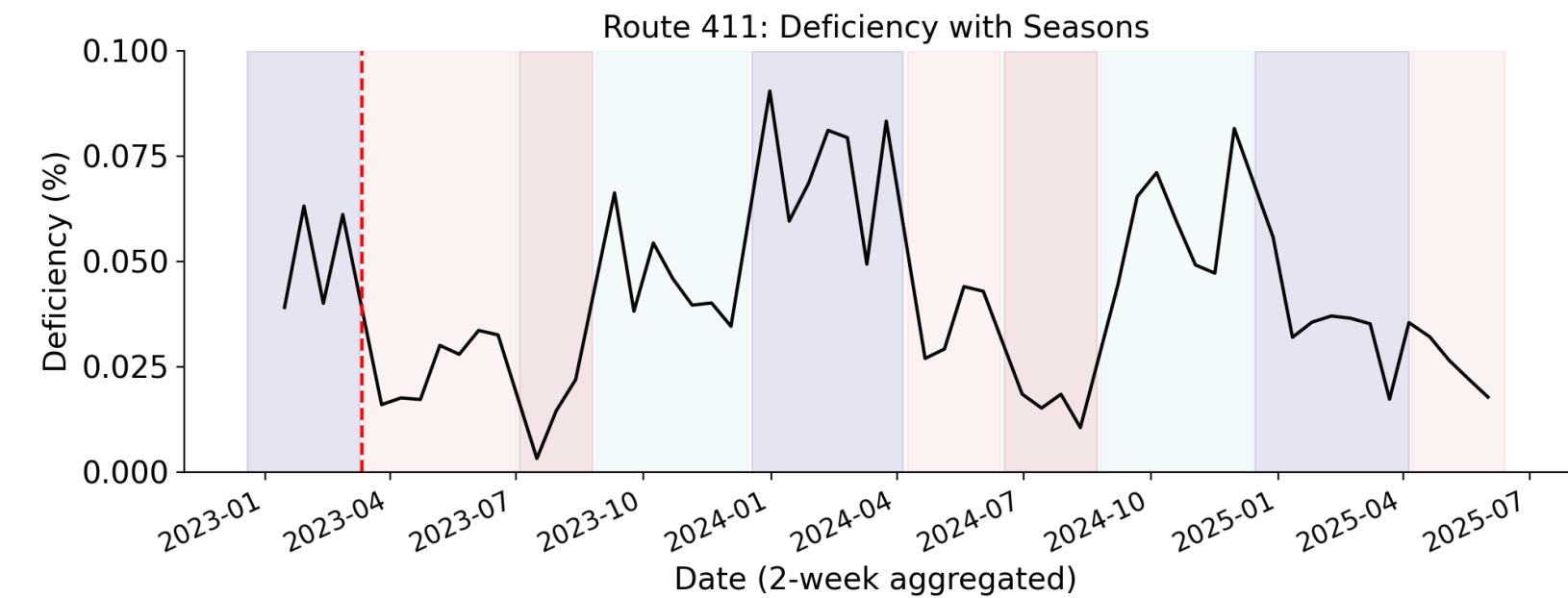
Understand better



Understand better



Understand better



**When do we actually
need to flag a
route for
schedule update?**

Understand better

Deficiency Priority Tool

A tool to **progress the deficiency** changes. Is it getting worse? Or seasonal?

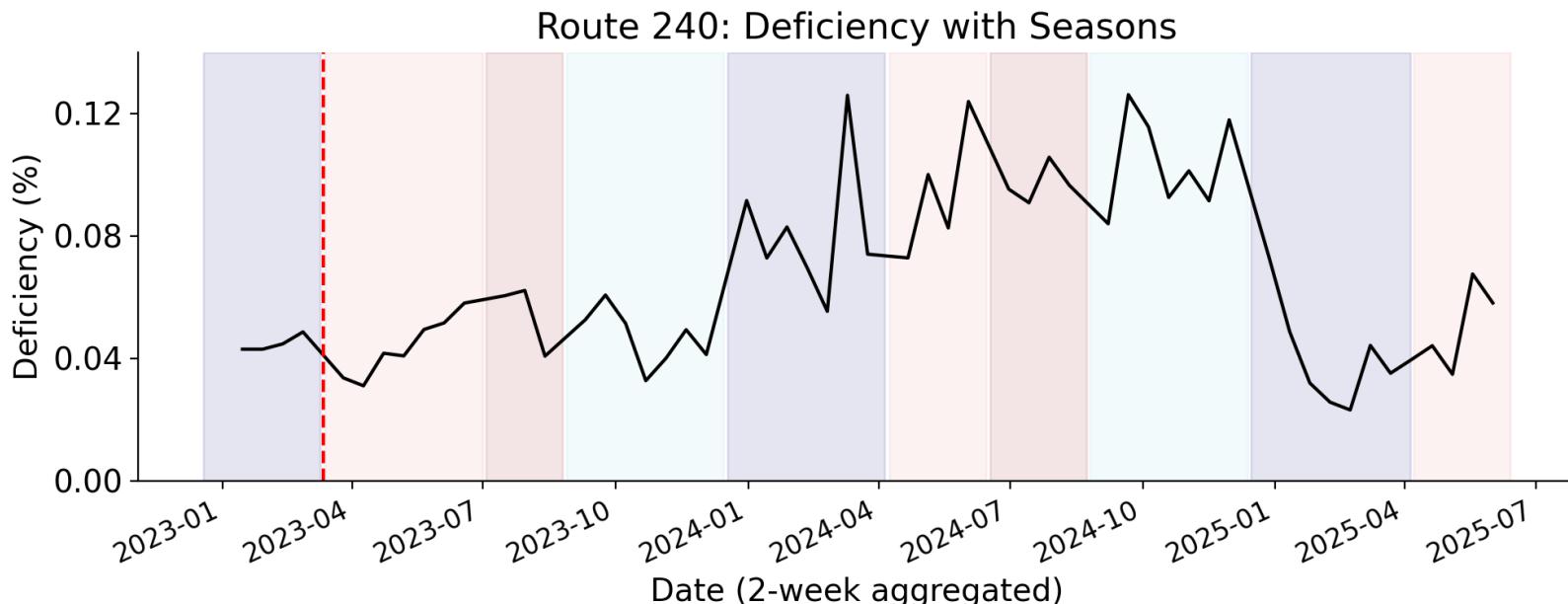
route/ season	Winter 24 - Winter 23	Spring 24 - Spring 23	Summer 24 - Summer 23	Fall 24 - Fall 23
1	0.10%	3.00%	2.90%	-2.30%
4	0.80%	5.60%	-1.20%	-1.60%
7	2.30%	-1.60%	-3.30%	0.40%
8	3.00%	-1.30%	-2.30%	-1.50%
9	4.80%	1.20%	1.50%	-2.40%
10	5.80%	3.30%	1.40%	-1.60%
11	1.30%	-2.60%	-1.70%	-0.10%
15	0.60%	1.20%	0.10%	2.10%
16	0.20%	0.80%	0.60%	-1.40%

Understand better

Deficiency Priority Tool

→

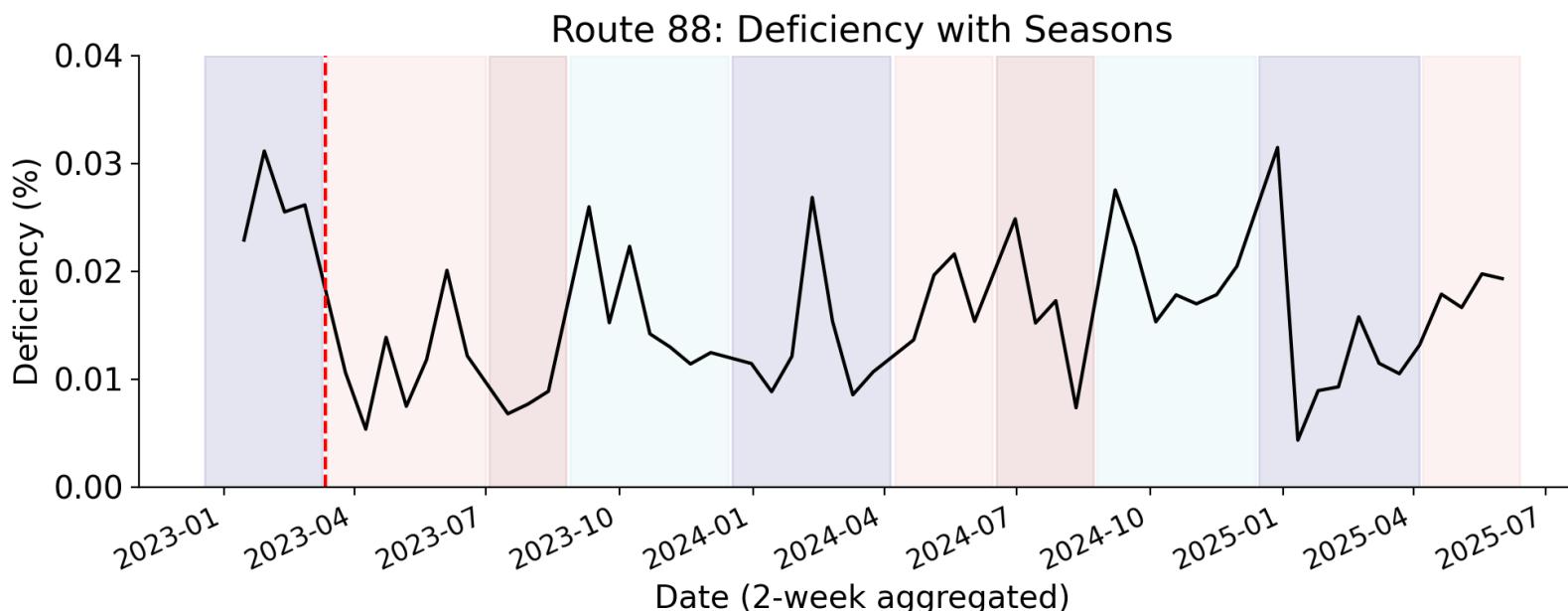
route/ season	Winter 24 - Winter 23	Spring 24 - Spring 23	Summer 24 - Summer 23	Fall 24 - Fall 23
240	3.70%	5.10%	4.30%	5.70%
101	4.30%	5.20%	2.40%	2.40%
83	1.70%	2.10%	2.60%	3.60%
37	1.60%	0.60%	1.40%	5.70%



Understand better

Deficiency Priority Tool

route/ season	Winter 24 - Winter 23	Spring 24 - Spring 23	Summer 24 - Summer 23	Fall 24 - Fall 23
211	0.10%	0.70%	-0.90%	0.80%
88	-1.30%	0.60%	0.80%	0.30%
16	0.20%	0.80%	0.60%	-1.40%
51	-0.70%	0.50%	0.10%	0.30%



Understand better

route	2024-3 vs 2023-3	2024-4 vs 2023-4	2025-1 vs 2024-1	2025-2 vs 2024-2
90	1.5%	2.6%	1.2%	7.7%
37	1.5%	6.7%	2.1%	1.4%
41	3.3%	2.6%	2.2%	2.3%
24	0.9%	1.6%	1.2%	2.0%

→

Data updated through: Rating Day Type Garage Route Percentile Layover Layover Used is defined as the Actual Layover, where if zero minutes, is the layover attached to the following deadhead.

6/14/2025 Spring 2025 Weekday (All) 90 90th Actual Layover

Schedule Summary

	Scheduled Trips	Trips Excluded (W or No Data)	Total Run Time	Total Run + Min Layover Hours	Total Run + Actual Layover
Inbound	19	0	10.0	11.7	10.0
Outbound	21	0	11.3	13.4	14.1

Deficiency Summary

	Deficient Trip Count	Surplus Trip Count	Total Deficit Hours	Deficiency % of Service	Total Surplus Hours	Surplus % of Service
Inbound	18	1	1.7	16.9%	0.0	0.4%
Outbound	6	15	0.2	1.7%	1.4	9.8%

Key

- Run + Layover
- Scheduled Run Time
- Median Trip Run Time
- Upper Percentile Run Time

Outbound

Inbound

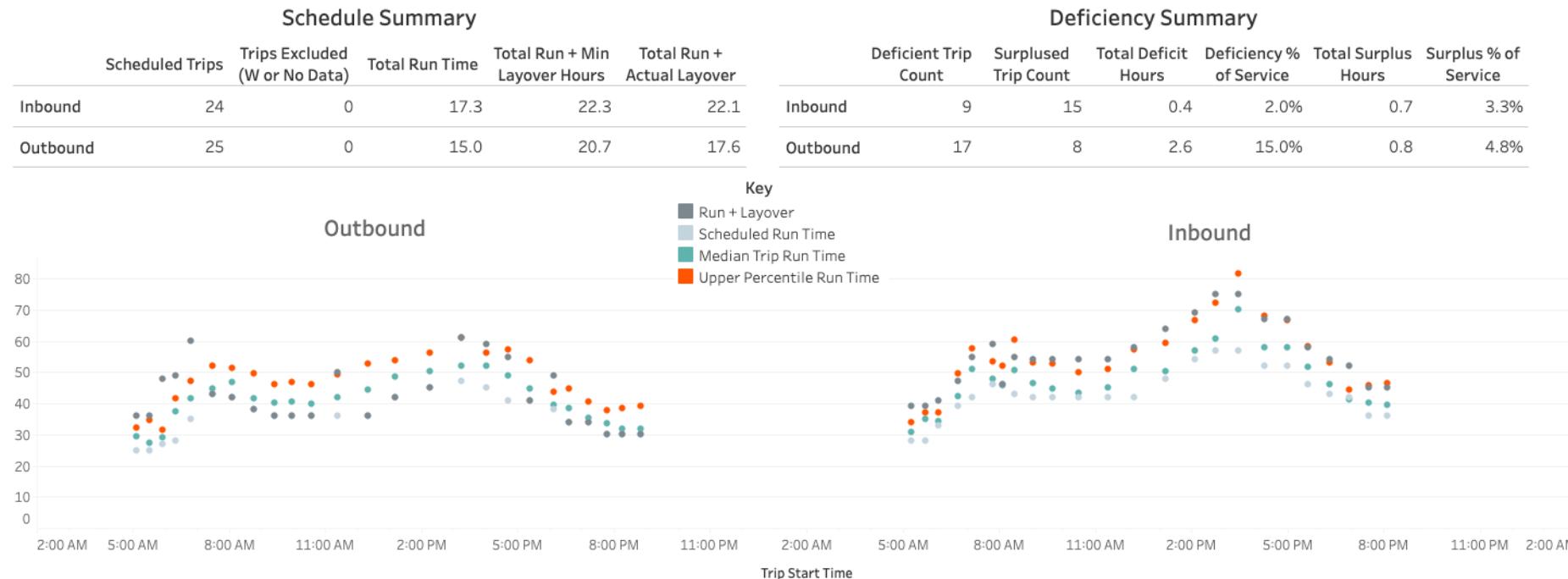
Understand better

route	2024-3 vs 2023-3	2024-4 vs 2023-4	2025-1 vs 2024-1	2025-2 vs 2024-2
90	1.5%	2.6%	1.2%	7.7%
37	1.5%	6.7%	2.1%	1.4%
41	3.3%	2.6%	2.2%	2.3%
24	0.9%	1.6%	1.2%	2.0%



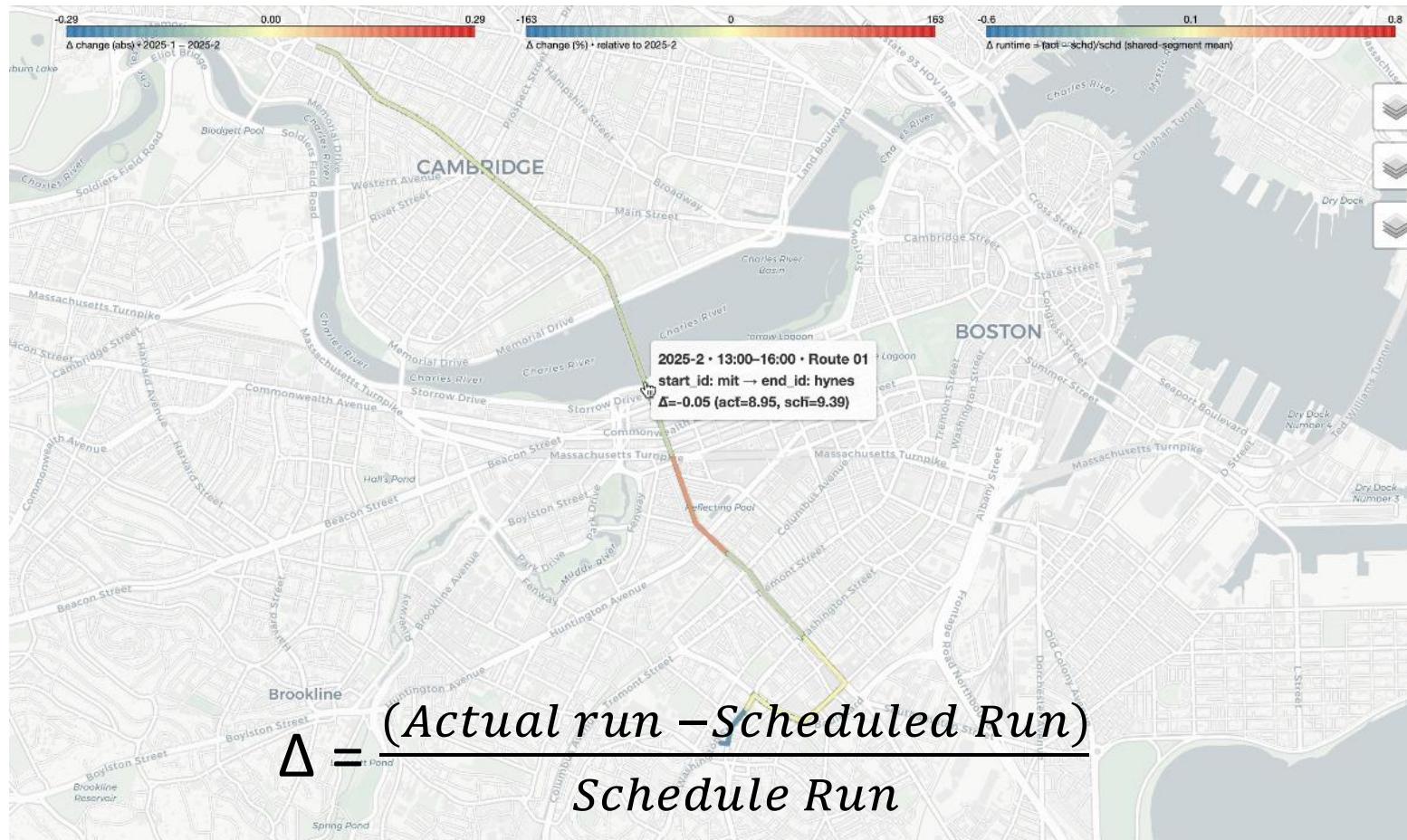
Data updated through: Rating Day Type Garage Route Percentile Layover Layover Used is defined as the Actual Layover, where if zero minutes, is the layover attached to the following deadhead.

6/14/2025 Spring 2025 Weekday (All) 41 90th Actual Layover



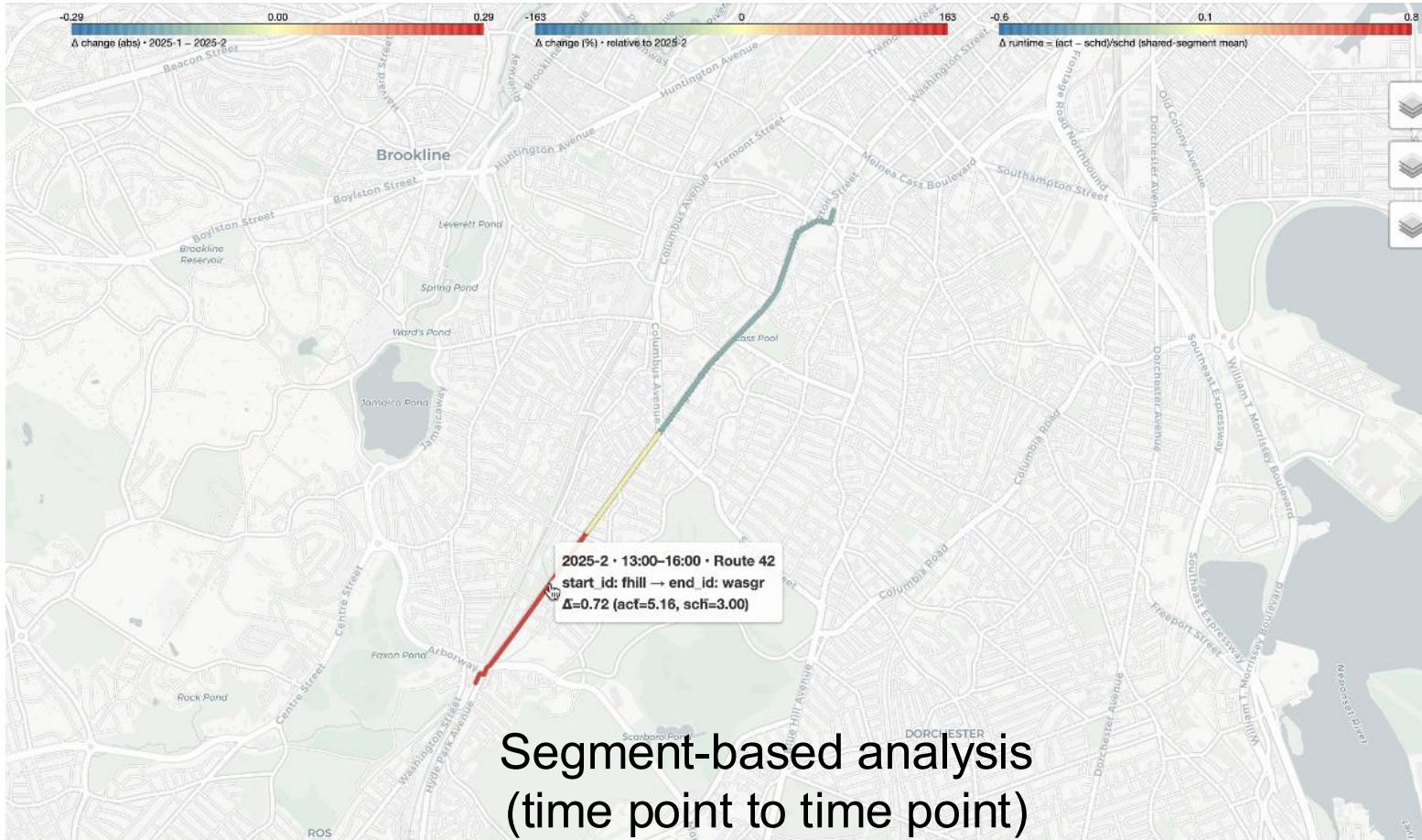
Understand better

Corridor Analysis Tool



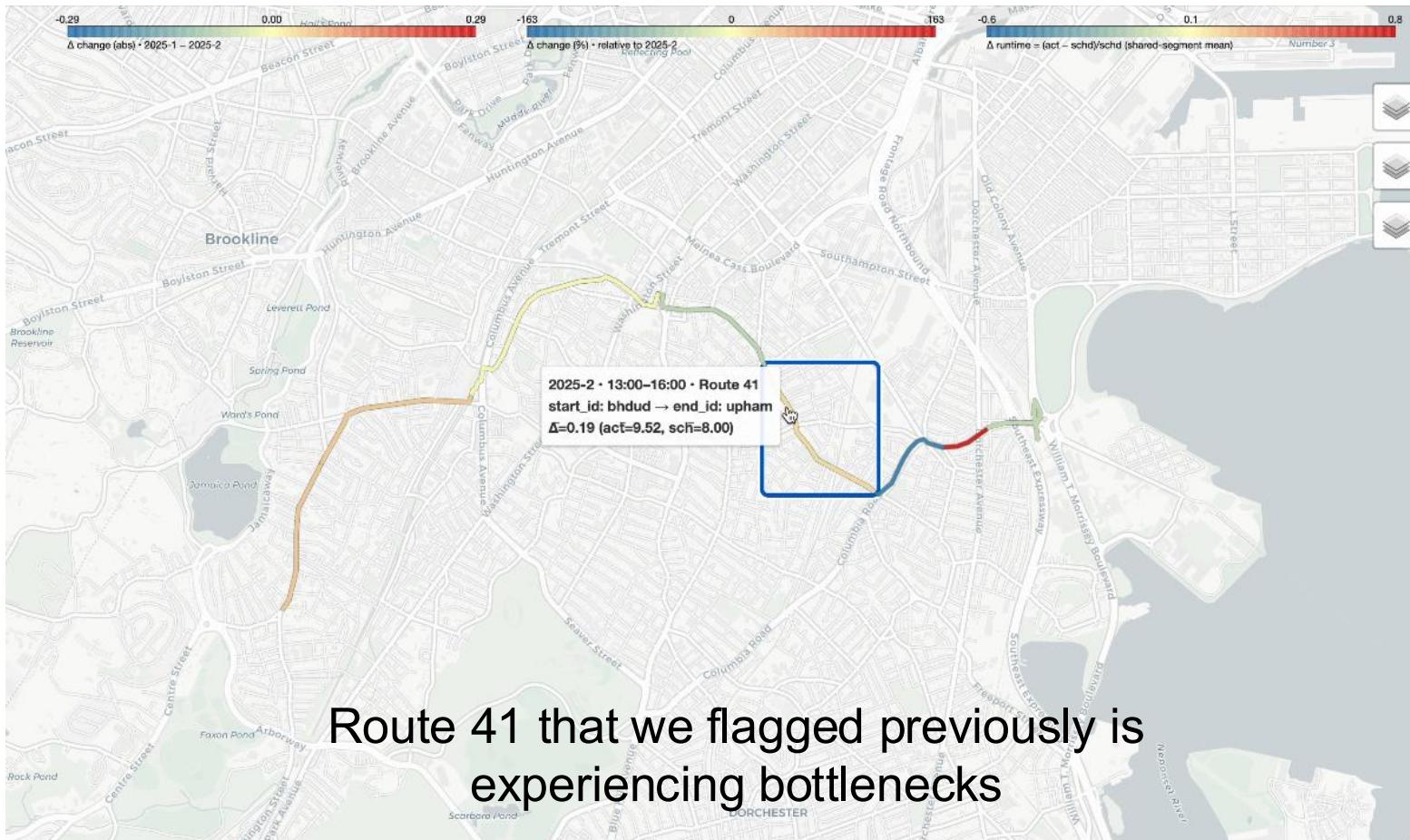
Understand better

Corridor Analysis Tool



Understand better

Corridor Analysis Tool

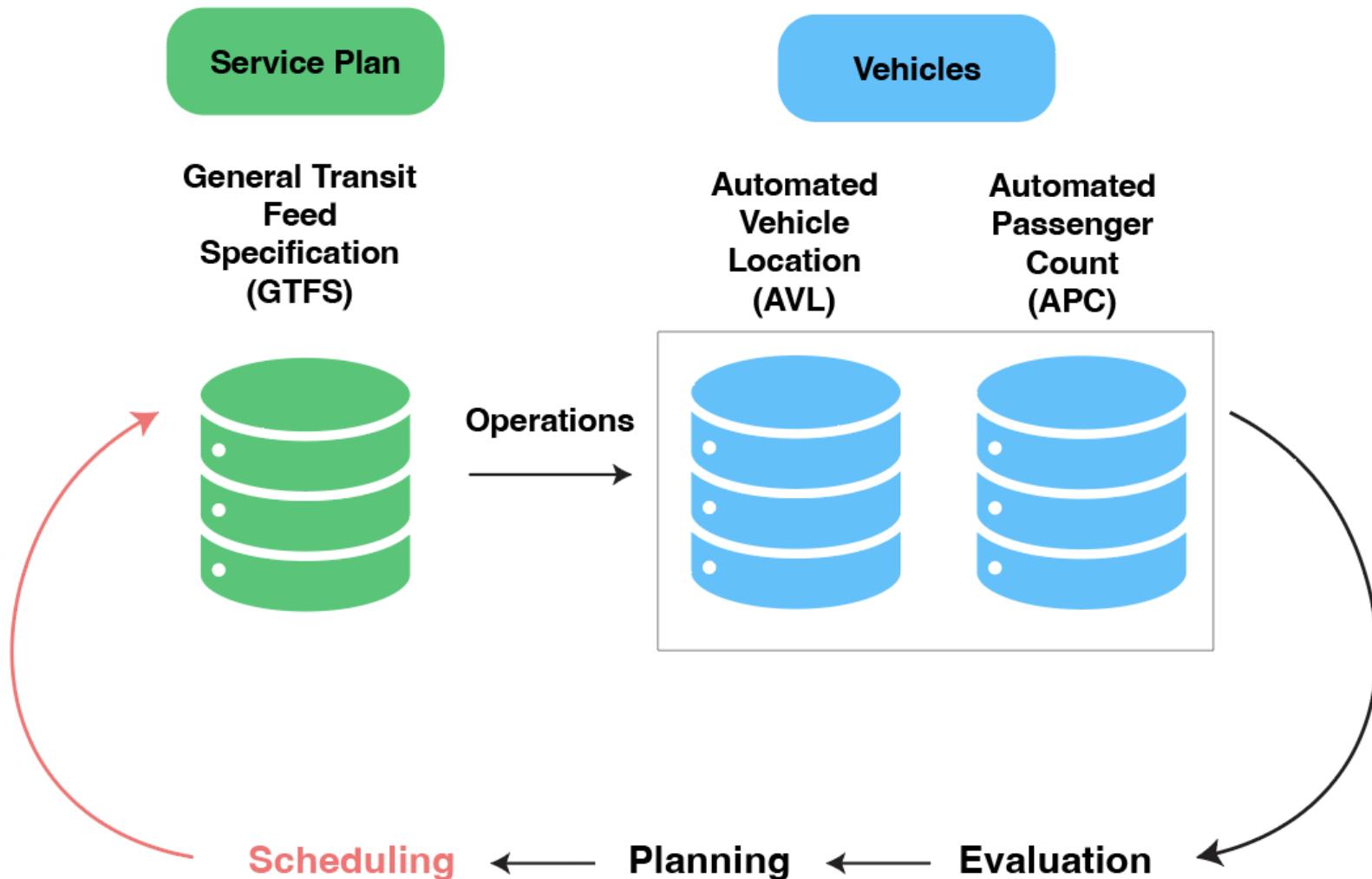


Understand better

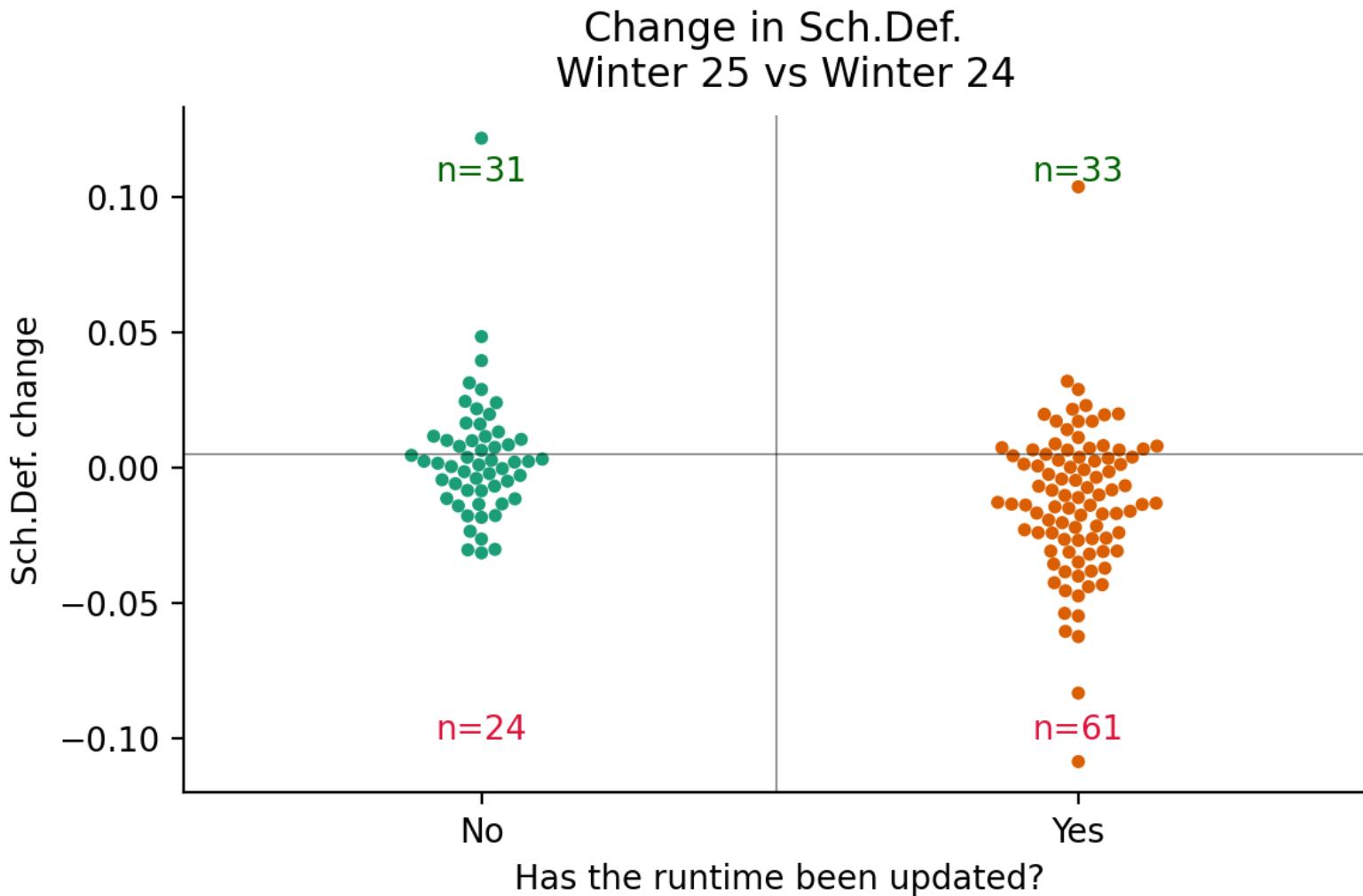
Corridor Analysis Tool



3. Correct better



Correct better

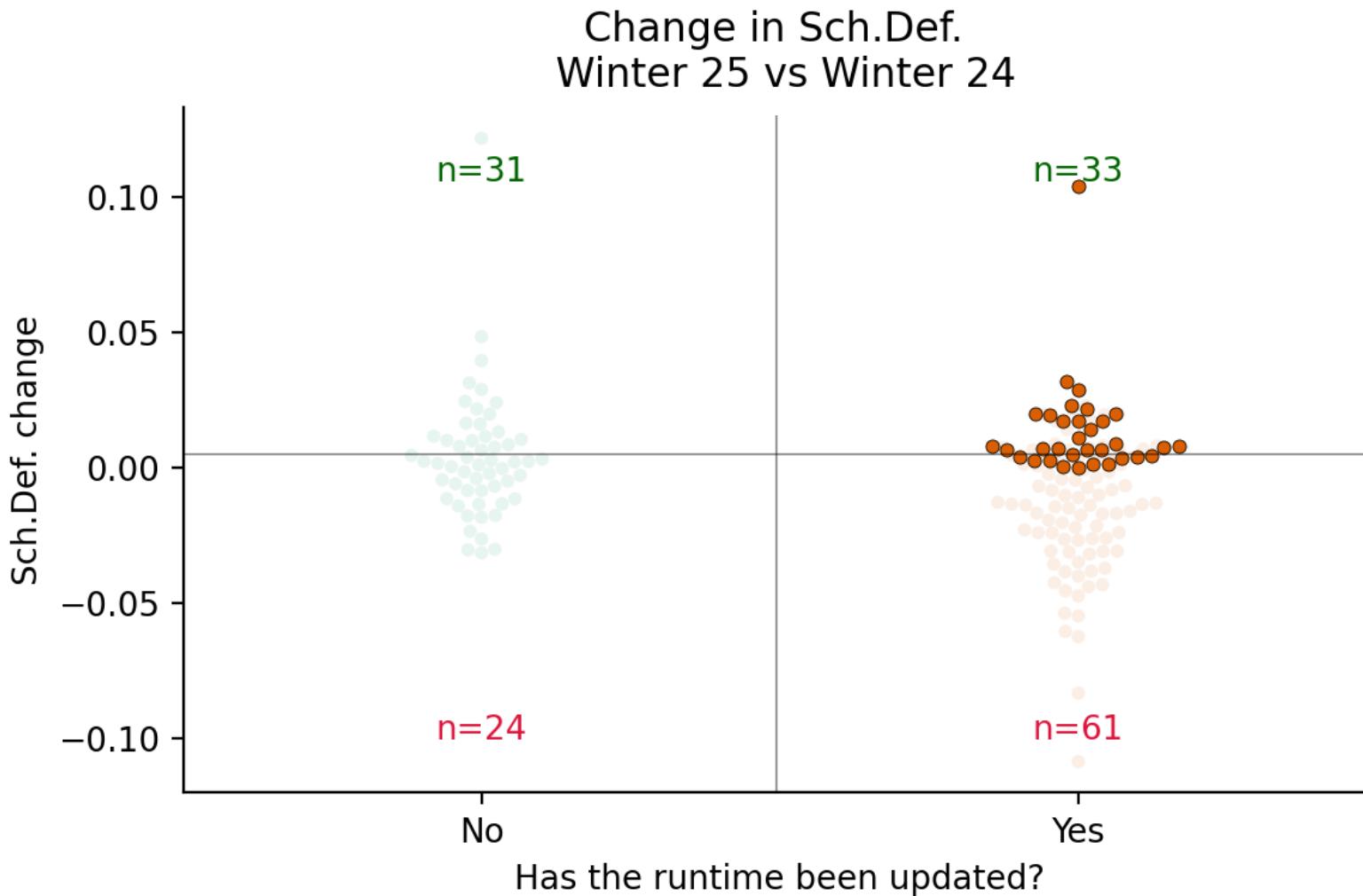


How many **periods** do we need?

What is the **time frame** for analysis?

How do we **clean** data?

Correct better

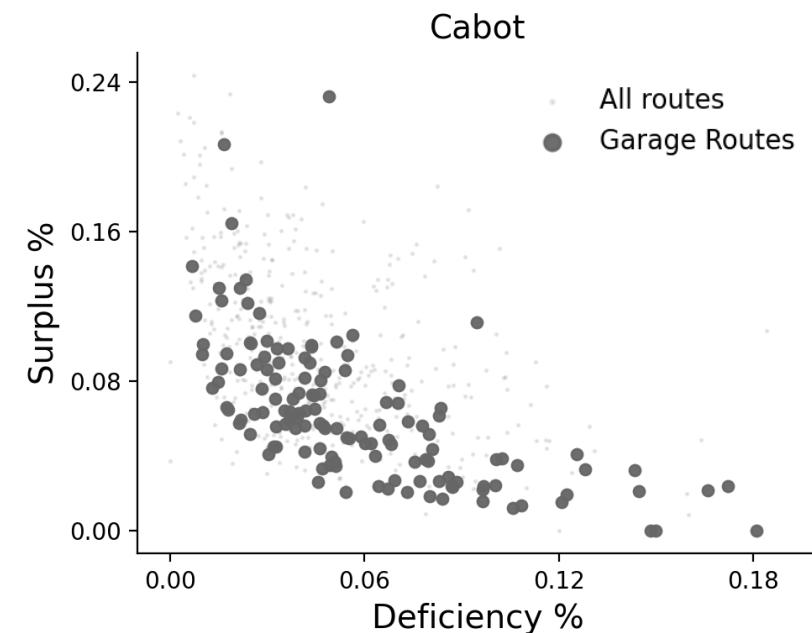
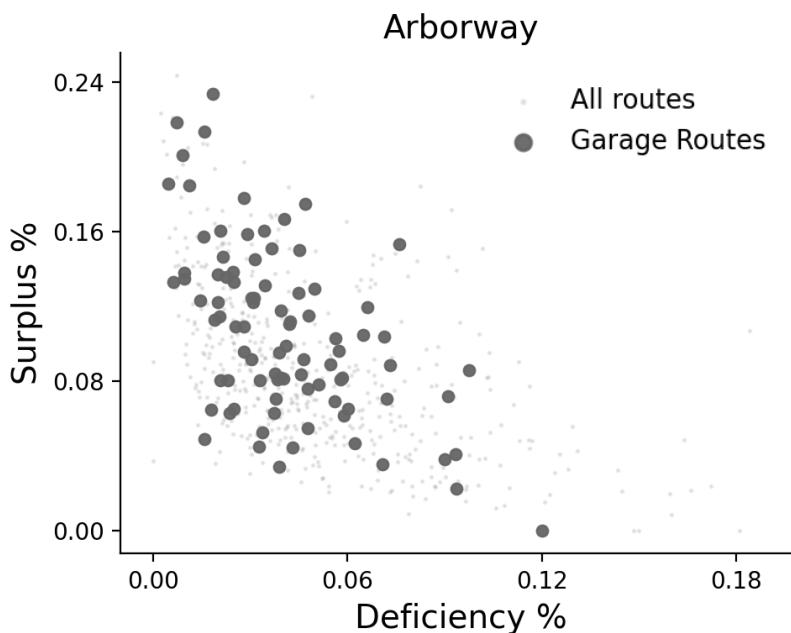
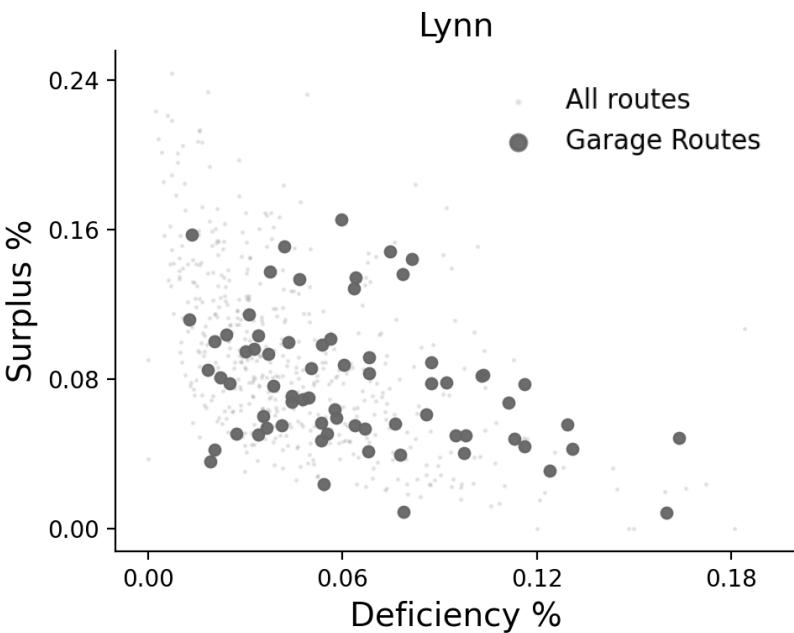


How many **periods** do we need?

What is the **time frame** for analysis?

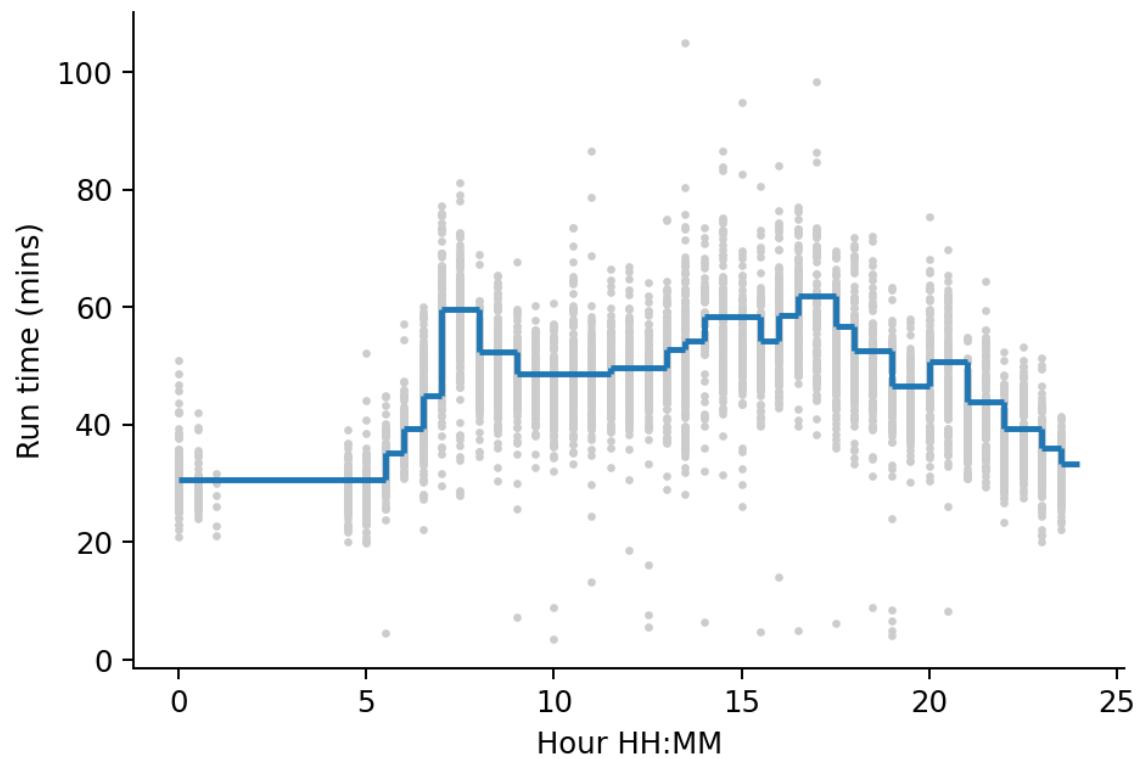
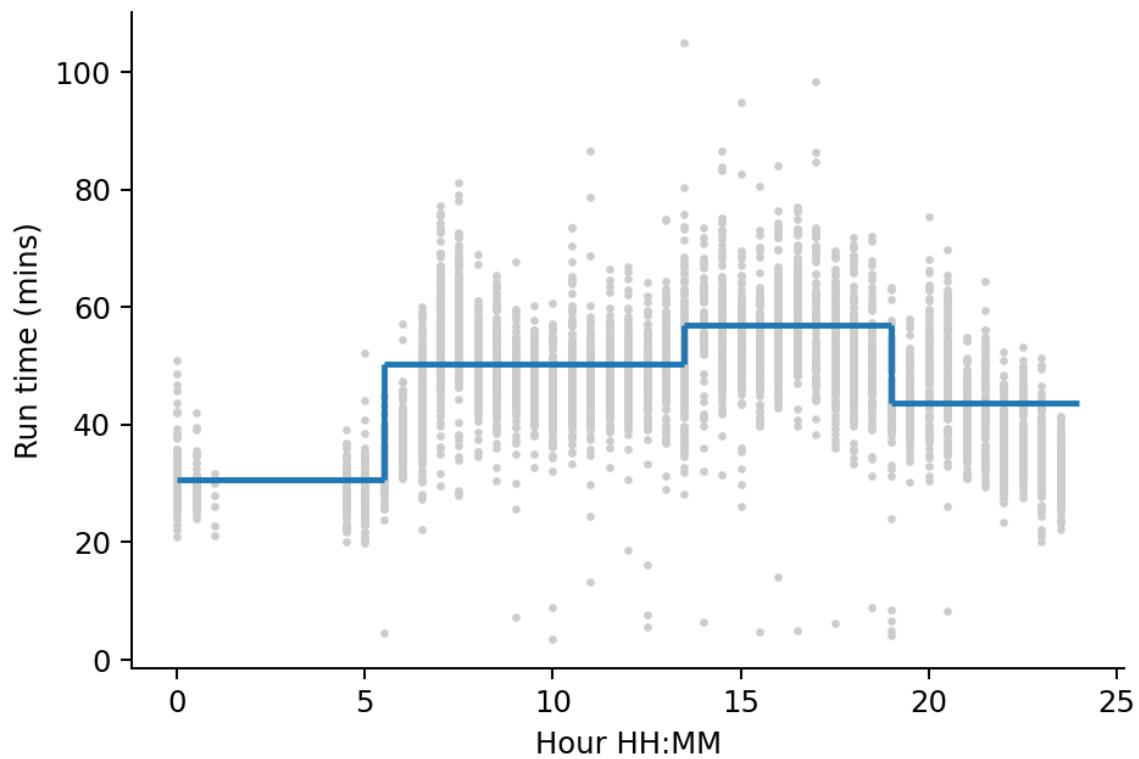
How do we **clean** data?

Correct better



We need informed choices for run time updates

Correct better

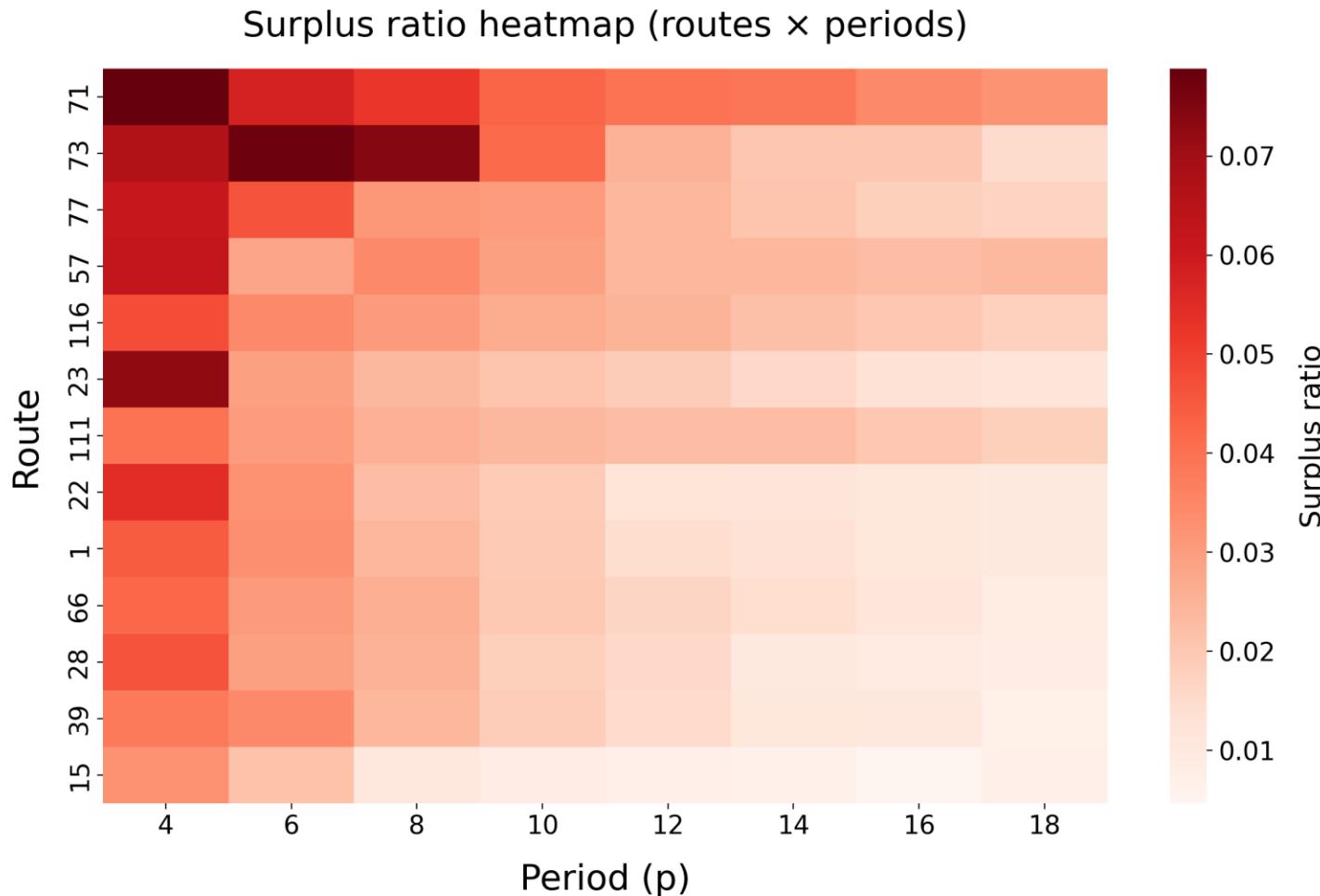


We need informed choices for run time updates

Correct better

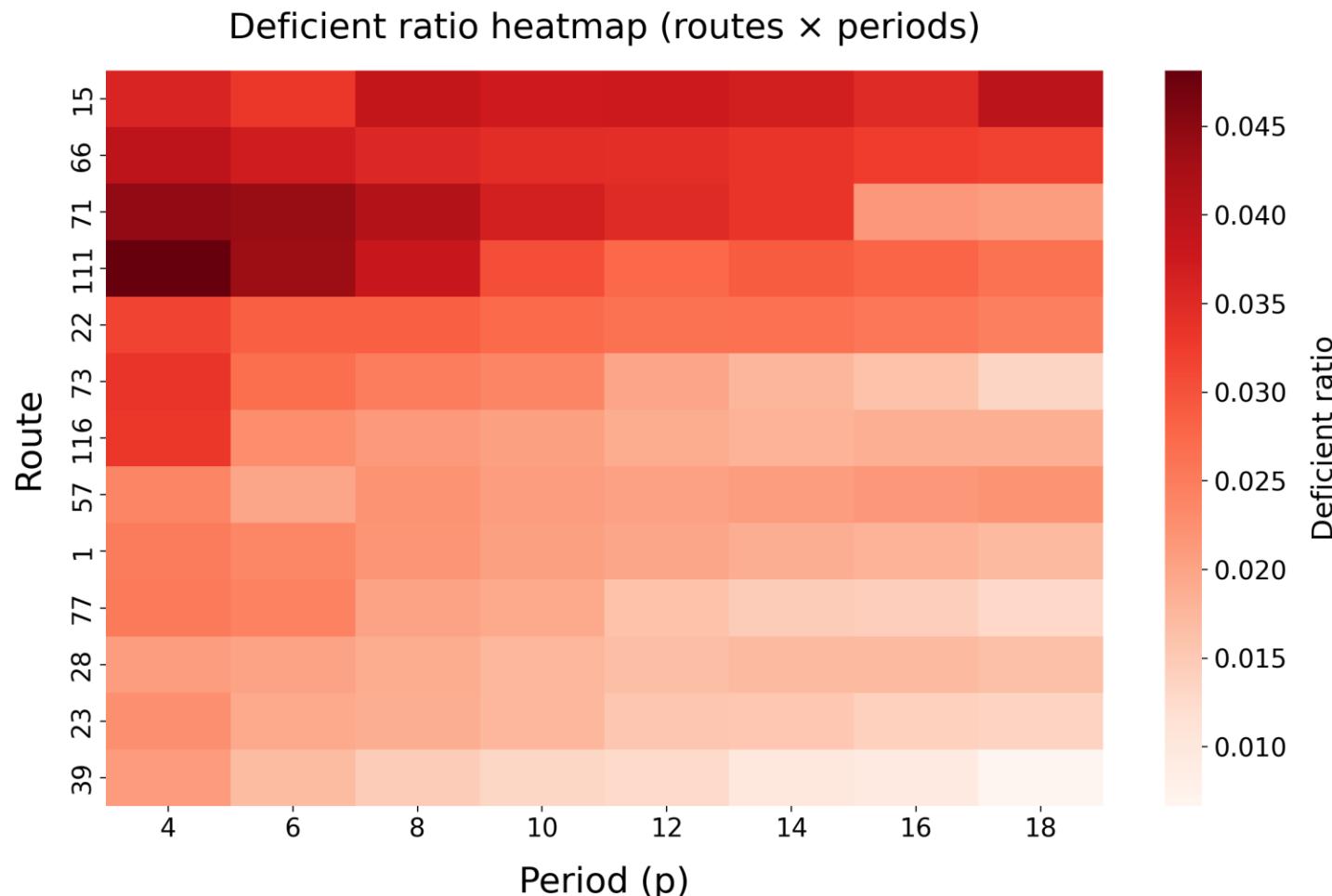
Period recommendation tool

+6 periods can decrease surplus



Correct better

Period recommendation tool



Correct better

Schedule thumbs-up tool

```
table = batch_compare(  
    sd_all,  
    combined,  
    route='23',  
    direction='Outbound',           # 'Outbound' / 'Inbound'  
    dow='Weekday',                 # 'Sunday' / 'Weekday' / 'Saturday'  
    schedule_season='2025-1',       # e.g., '2024-4'  
    schedule_type='N',             # School or No school (NS)  
    runtime_groups=runtime_groups,  
    rt_stat='quantile',            # 'quantile'|'median'|'mean'  
    q=0.9,                         # 90th vs 95th quantile  
    use_top_pair=True              # keep only most common Start-End pair  
)  
  
table[['runtime_label','deficiency_ratio','surplus_ratio']]
```

	runtime_label	deficiency_ratio	surplus_ratio
0	2025-1	0.005219	-0.071819
1	2024-4	0.008231	-0.068137
2	2024-1 to 2025-4	0.001986	-0.085233
3	2023-1 and 2024-1	0.000687	-0.130683

Compare proposed schedule update with previous run time scenarios to make sure our schedule works in long term

Correct better

Schedule thumbs-up tool

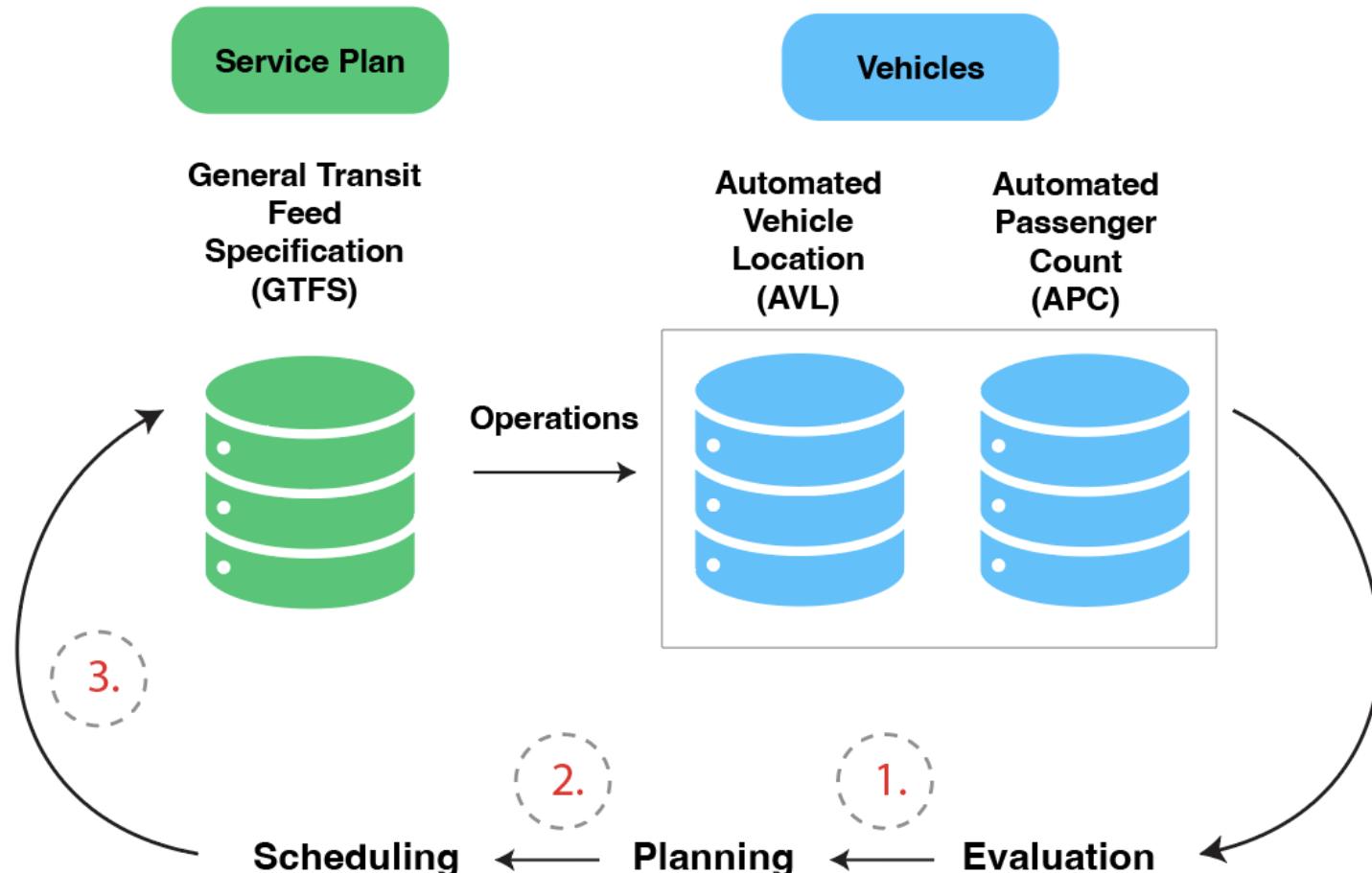
```
runtime_groups = [
    ("2024-2", ("2024-2",)),
    ("2024-1", ("2024-1",)),
    ("2023-2 to 2024-1", ("2023-2", "2023-3", "2023-4", "2024-1",)),
    ("2023-2", ("2023-2",)),
]
table2 = batch_compare(
    sd_all, combined,
    route='34', direction='Outbound', dow='Weekday',
    schedule_season='2024-2', schedule_type='N',
    runtime_groups=runtime_groups,
    rt_stat='quantile', q=0.9, use_top_pair=True
)

table2[['runtime_label','deficiency_ratio','surplus_ratio']].iloc[1:]

  runtime_label  deficiency_ratio  surplus_ratio
1      2024-1           0.0       -0.255768
2  2023-2 to 2024-1           0.0       -0.249994
3      2023-2           0.0       -0.254986
```

Detect potential seasonal mismatch **before** publishing schedules

Moving Forward



1. Calculate Better

- Understanding data limitations
- Sensitivity analysis for different methodological approaches
- Data cleaning standardization

2. Understand Better

- Deficiency priority tool
- Corridor analysis tool

3. Correct Better

- Period recommendation tool
- Schedule thumbs-up tool
- ATP standardization

Moving Forward

- Data collection and emerging methodologies bring down technological barriers
 - Comparison of different approaches
- But the institutional and human factors barriers remain
 - Culture of schedule adherence
- More important than adding methods
 - Develop mechanisms to safeguard against blind spots in methods



Moving Forward

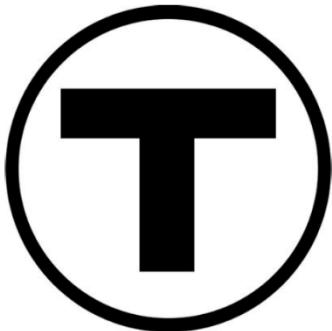
The screenshot shows a GitHub repository page for 'schedule_deficiency'. The repository is public and has 1 branch and 0 tags. The commit history shows the following activity:

File	Commit Message	Time
LICENSE	Initial commit	last week
README.md	Create README.md	now
corridor_vis_tool.ipynb	Add files via upload	1 minute ago
data_process.ipynb	Add files via upload	1 minute ago
hastus_tool.ipynb	Add files via upload	1 minute ago
presentation_figs.ipynb	Add files via upload	1 minute ago
schedule_thumbsup.ipynb	Add files via upload	1 minute ago

The README file contains the following text:

```
Created for the Schedule Deficiency Project at MBTA - Massachusetts Bay Transportation Authority  
Operations Planning, Scheduling, & Strategy (OPSS)  
Nail Bashan - Intern bashan.n@northeastern.edu
```

https://github.com/nnbashan/schedule_deficiency



**Massachusetts Bay
Transportation Authority**

Schedule Deficiency

Final Presentation
08/29/25

Nail Bashan
bashan.n@northeastern.edu