# Detecting Patterns in Train Position Data of Trains in Shunting Yards

Analysis of Arrival Time Distributions and Delays

Author: Amanda Krudde
Supervisors: Mathijs de Weerdt, Issa Hanou
CSE3000: Research Project

ProRail TUDelft

## Introduction

- Shunting yards for parking, maintenance and cleaning of trains
- Train Unit Shunting Problem [1]
- Detect patterns -> identify recurring issues
- Gap in analysis of arrival times of trains
  - Schedule additional staff
  - Reduce delays in future
- Hypothesis: patterns will be found

**Figure 1:** Area Amersfoort with shunting yard (red)

## Dataset

- Realization data of NS trains
- GPS points of each train unit
- 7 shunting yards in NL
- Period of 10 months
- Hosted in Azure environment

## Background

- Clustering algorithms
  - Hierarchical Clustering
  - DBSCAN
- Related work
  - New York City's subway system [2,3]

## Research Question

What patterns can be identified in arrival time distributions of trains in shunting yards using train position data?

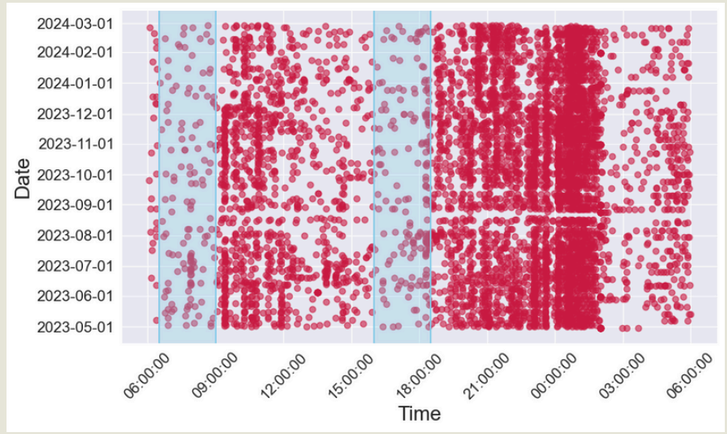


**Figure 2a:** Scatterplot of the arrival times

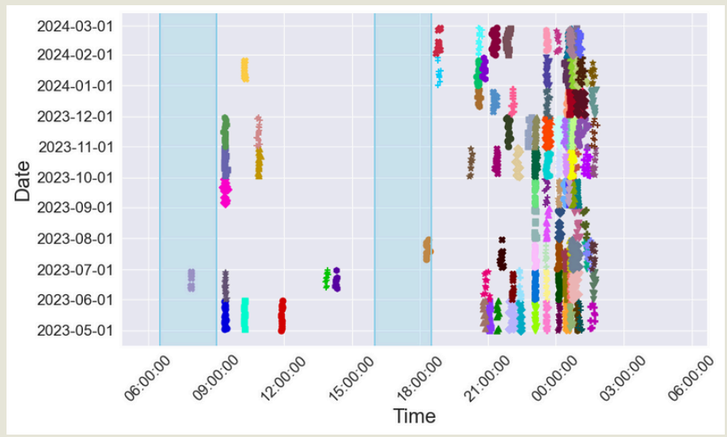


**Figure 2b:** Identified train series (using DBSCAN)
Peak hours are highlighted in blue.

## Methodology & Results

1. **Data Processing**
2. **Grouping similar train instances**
   - Cluster train units as trains
   - Train series across days
3. **Line graph visualization**
4. **Heatmap visualization**
5. **Detect patterns**

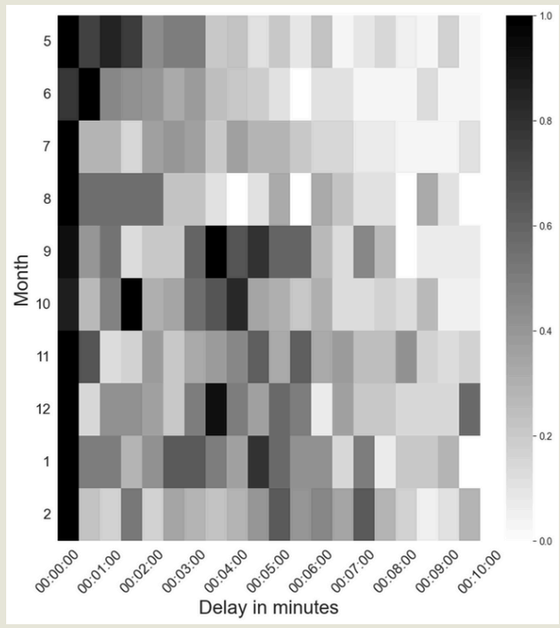


**Figure 4:** Heatmap visualization of delays

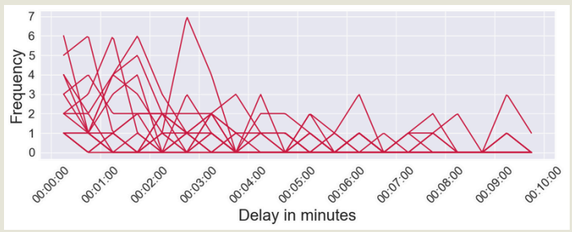


**Figure 3a:** Line graph visualization (DBSCAN)

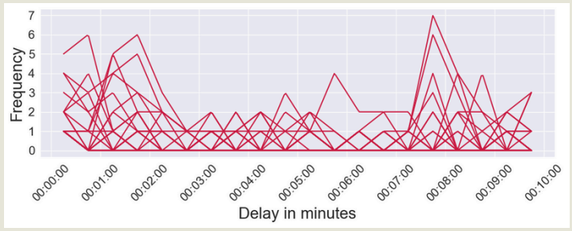


**Figure 3b:** Line graph visualization (created algorithm)

## Observations

**Arrival time distributions**

- Busy arriving moments
- Still trains during peak hour
- Train series after peak hours
- Less predictable outside busy moments
- No series during vacation breaks

**Delays**

- Less delay in summer
- More delay in fall & winter

## Conclusion

- Align staff and equipment at busy moments
- Assign recurring trains to certain tracks
- Additional maintenance during fall/winter to avoid delays
- Hypothesis aligns with outcomes

## Future Work

- Expanding dataset
- Deep learning
- Broader context

## Related Literature

[1] R. Freling, R. M. Lentink, L. G. Kroon, and D. Huisman. Shunting of passenger train units in a railway station. In *Transportation Science*, 39(2):261–272, 2005.
[2] A. Halvorsen, D. Jefferson, T. Stasko, and A. Reddy. Algorithm for tracing train delays to incident causes. In *Transportation Research Record*, volume 2674, pages 264–273, 2020. Publisher: SAGE Publications Inc.
[3] S. Lehmann, A. Reddy, C. Samsundar, and T. Huynh. Automated train identification and train position monitoring at new york city transit. In *Transportation Research Record*, volume 2674, pages 843–854, 2020. Publisher: SAGE Publications Inc.