

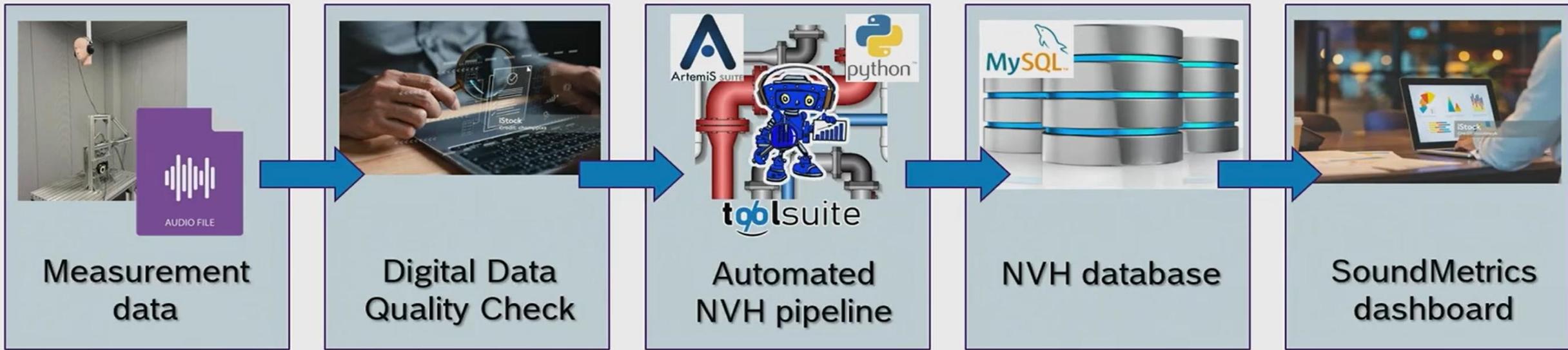
Master Thesis

Ideas discussion

Current Architecture

Data engineering pipeline for KPI processing established

Foundation for all AI use cases is a strong data base



- Continuously growing amount of data must be considered
- Find corrupted measurement files for revision
- Highly automated data processing
- Useable for Data Scientists w/o special NVH licences
- Interactive NVH reports based on PowerBI

Output Reports



BOSCH



**Sound
Metrics**

NVH Report

ERPA Report

NVH Drive Unit Comparison



BOSCH Order Spectrum

MIN DATE

04/12/2023

MAX DATE

04/12/2023

LAST REFRESH

13/11/2025 09:42:35

Reference Line

Show B-sample

Reference Sample



Generation

Sample

Load

ERPA

Condition

Part Number

AddInfo

BDU342

B-sample

LT

All

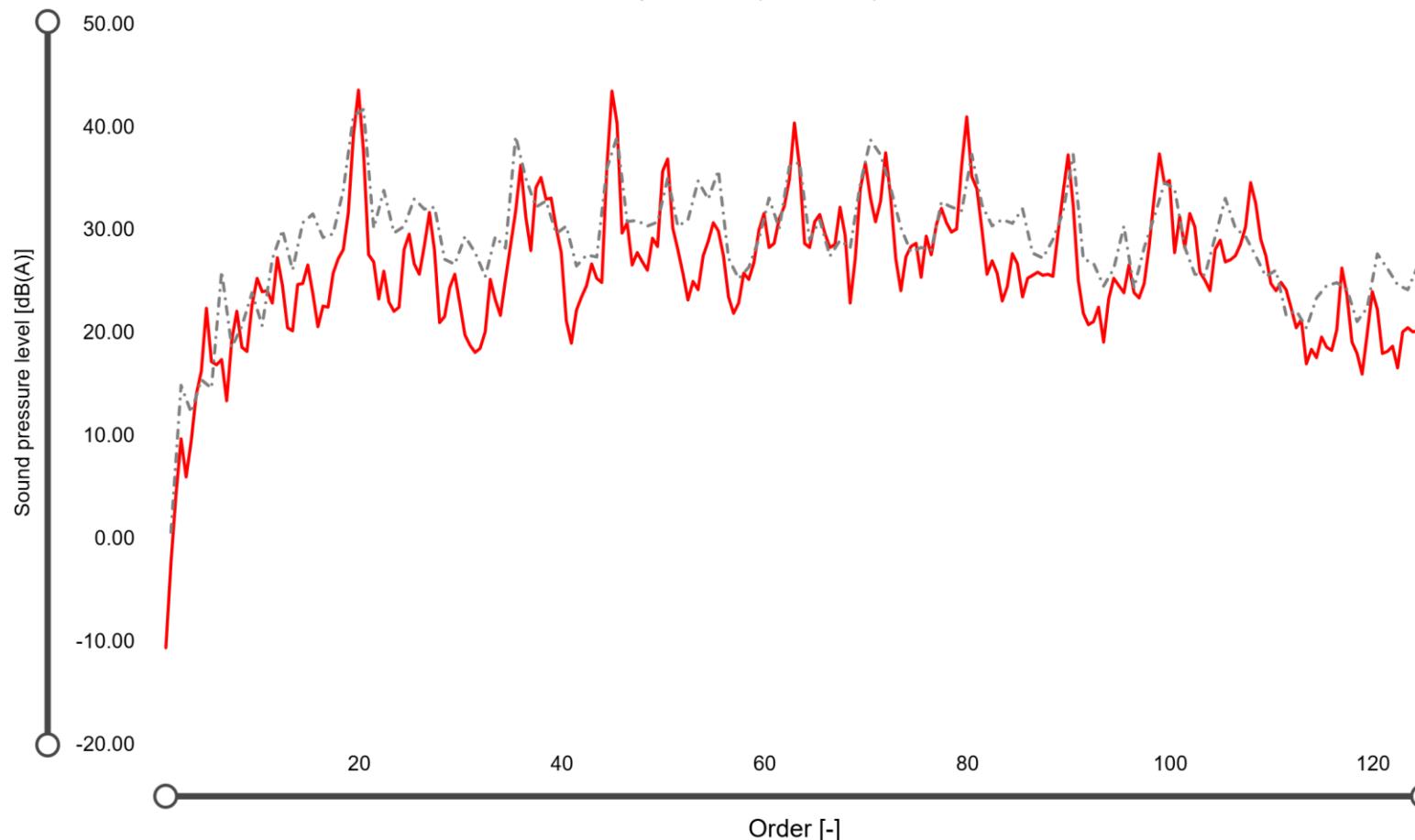
All

EB1110000J-30-37602-0004-Z2-300-00-0000-BDU3420

All

Reset

Order spectrum - LT (cadence 115)



Generation-Condition-Sample-PartNumber-AddInfo

BDU342-S08-B-sample-EB1110000J-30-
37602-0004-Z2-300-00-0000-BDU3420-250
Zyklen -20_+85



BOSCH ERPA Report

MIN DATE

14/07/2025

MAX DATE

28/08/2025

LAST REFRESH

06/10/2025 17:17:25

NEW (1)

S14 (2)

Reference Line

Reference Sample

Show

Series

▼

Generation

Sample

Erpa

Condition

PartNumber-AddInfo-condition

BDU38

B-sample

2682

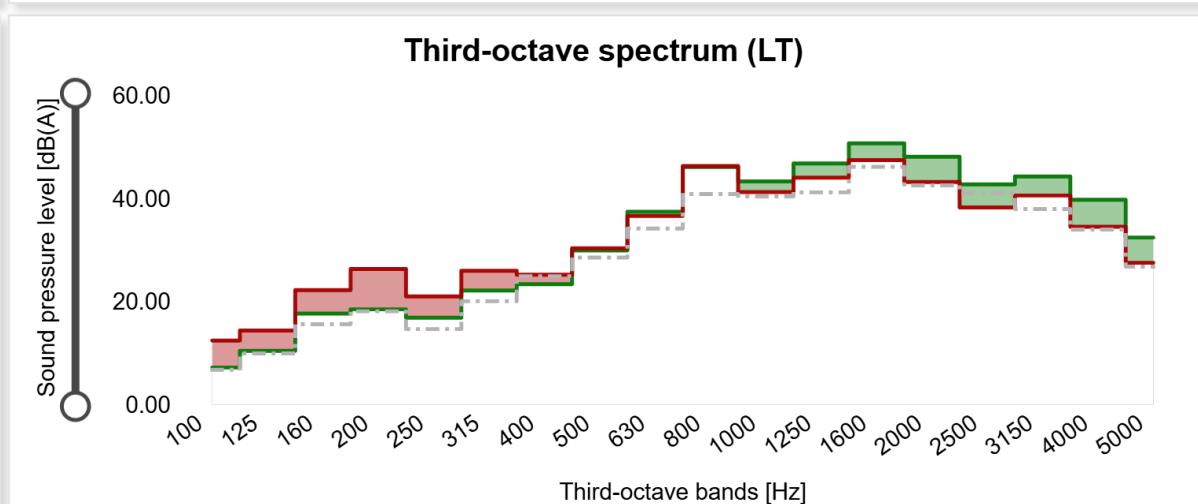
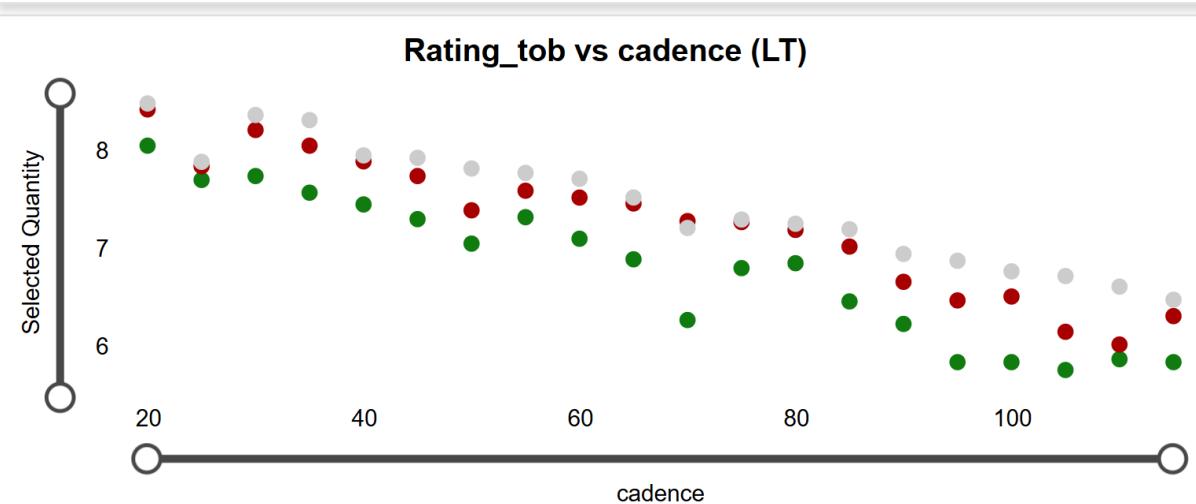
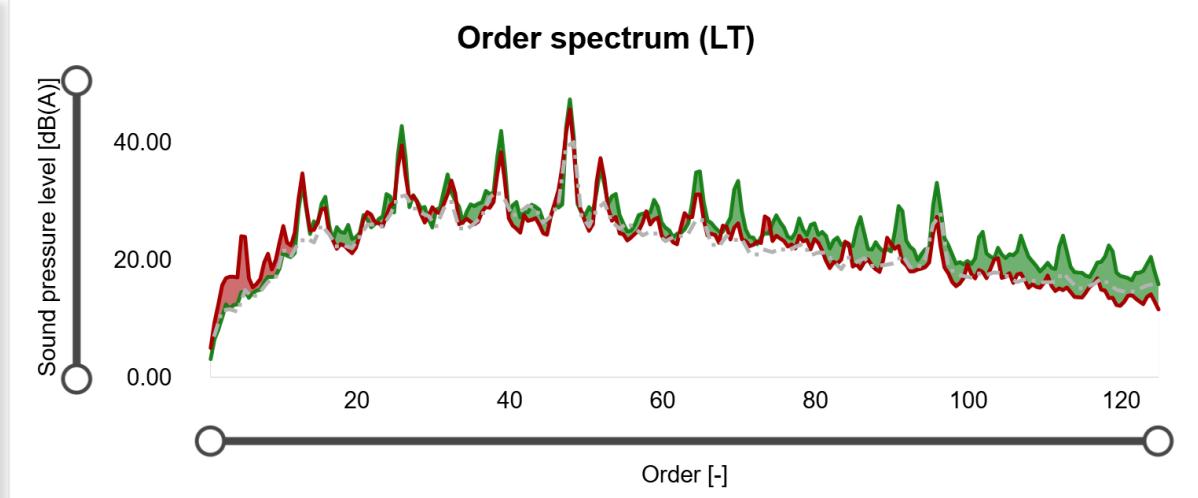
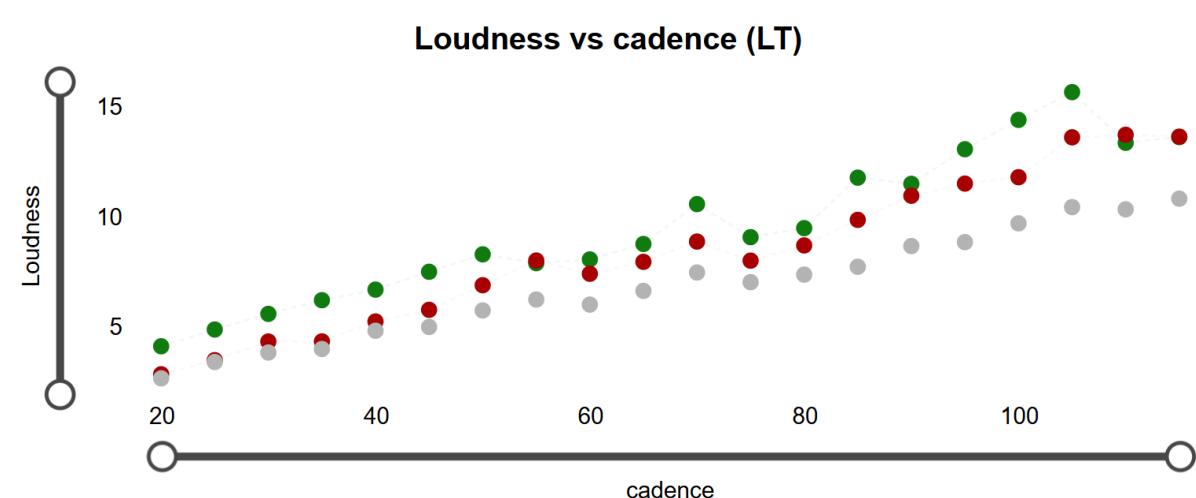
Multiple s...

All

Load

Performance Index

Reset





BOSCH Drive Units Comparison

MIN DATE

13/11/2023

MAX DATE

07/11/2025

LAST REFRESH

12/11/2025 14:18:26



Performance Index

Loudness

Cadence

All



BDU334-Series-LT--NEW



Filter Drive Unit 2

Generation

Sample

Load

ERPA

Condition

Reset

Part Number

AddInfo

All

All

Generation

Sample

Load

ERPA

Condition

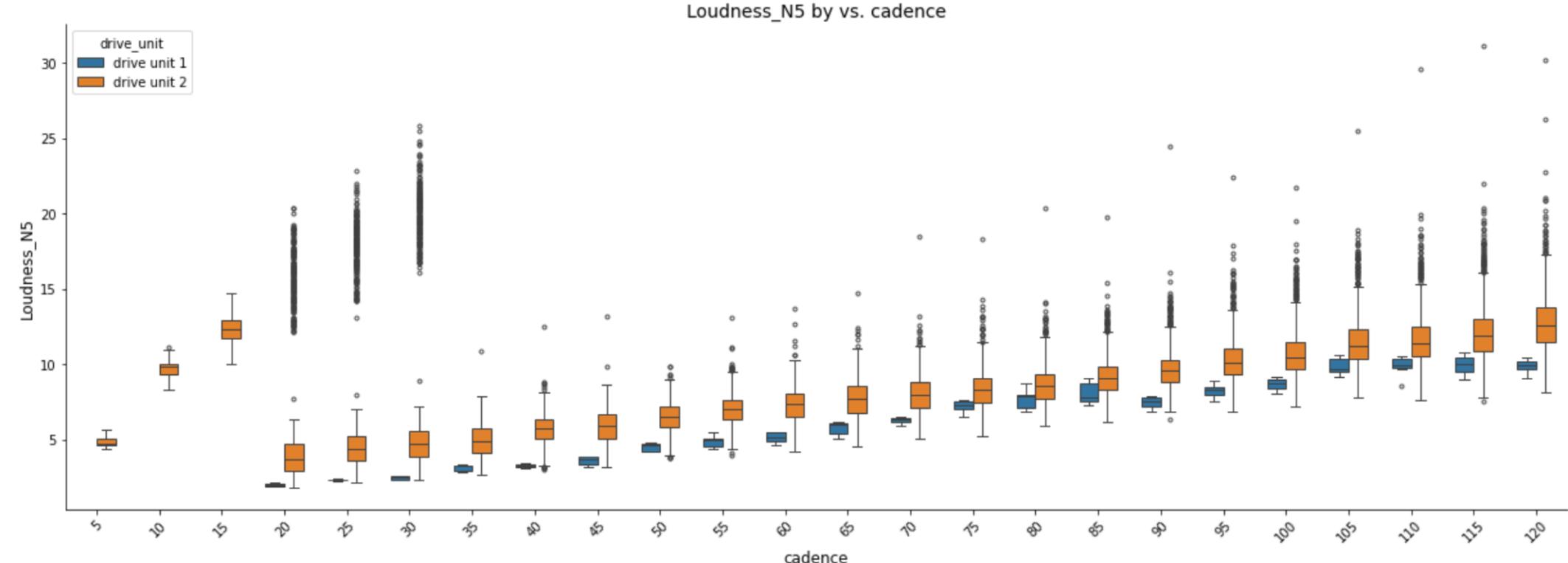
Reset

Part Number

AddInfo

All

All



Possible topics

Topic 1 : Automate ETL pipeline

- **Currently :** Batch Processing (ETL pipeline executed every wednesday)
- **Idea :** Automate the process and use spark streaming to process data in real time

Possible topics

Topic 2 : Integration of Power BI and Python

- **Currently :** There are some limitations in Python visuals in Power BI for example no tool tips and visual quality
- **Idea :** Develop a workflow to take filter values from Power BI and then display in Python Plotly

Possible topics

Topic 3 : Outlier detections using Machine Learning Models

- **Currently :** There is no work done on this topic
- **Idea:**
Develop and implement machine learning–based techniques to automatically detect anomalies in the dataset, improving accuracy and reducing manual effort.

Thank You !