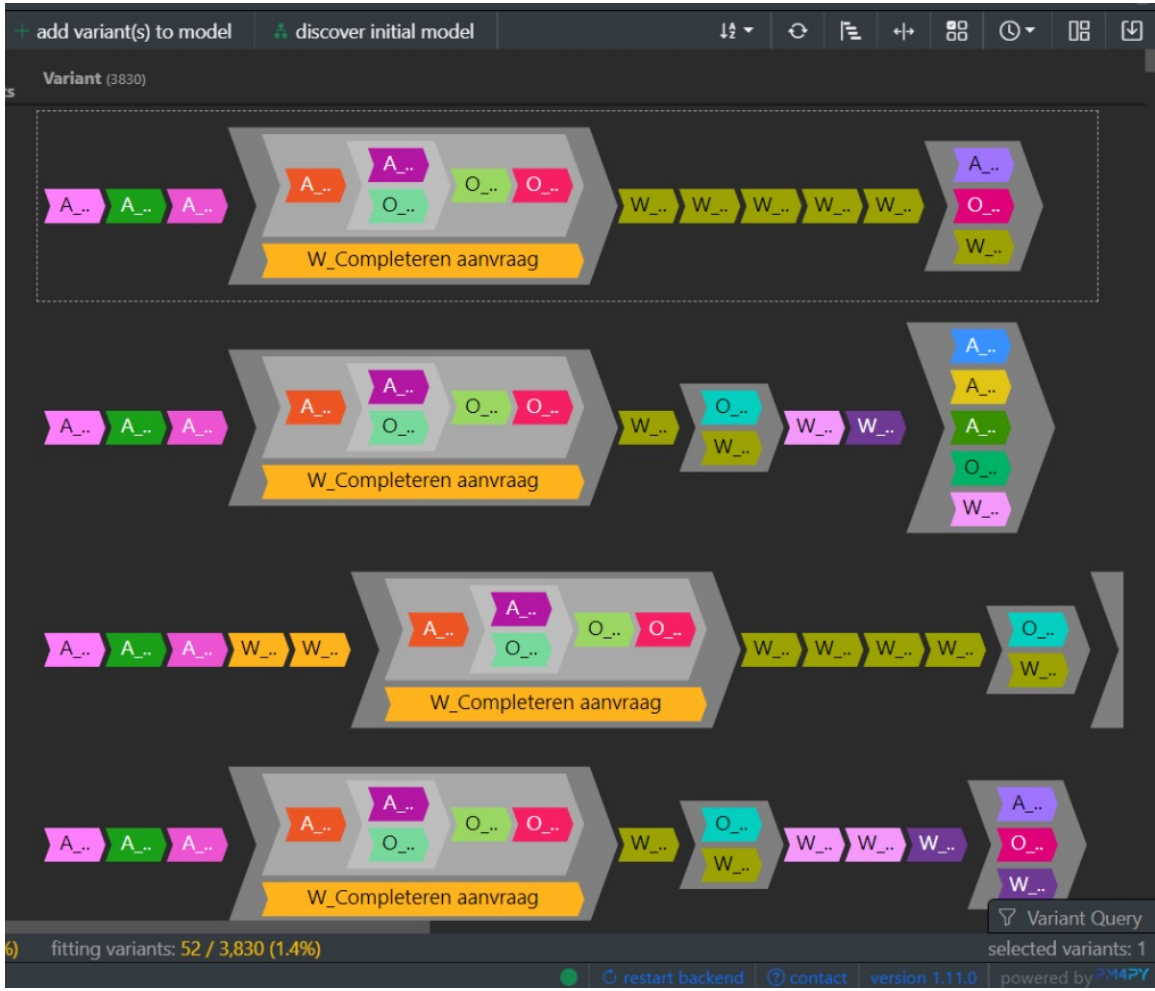
Team Project
Clustering Event Data in Process Mining

UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

- **Process mining** is a data-driven approach that uses techniques and algorithms to **discover, analyze, and improve business processes** based on event data
  → widely used across industries

- Event data and individual process executions (traces) are the key input for process mining techniques

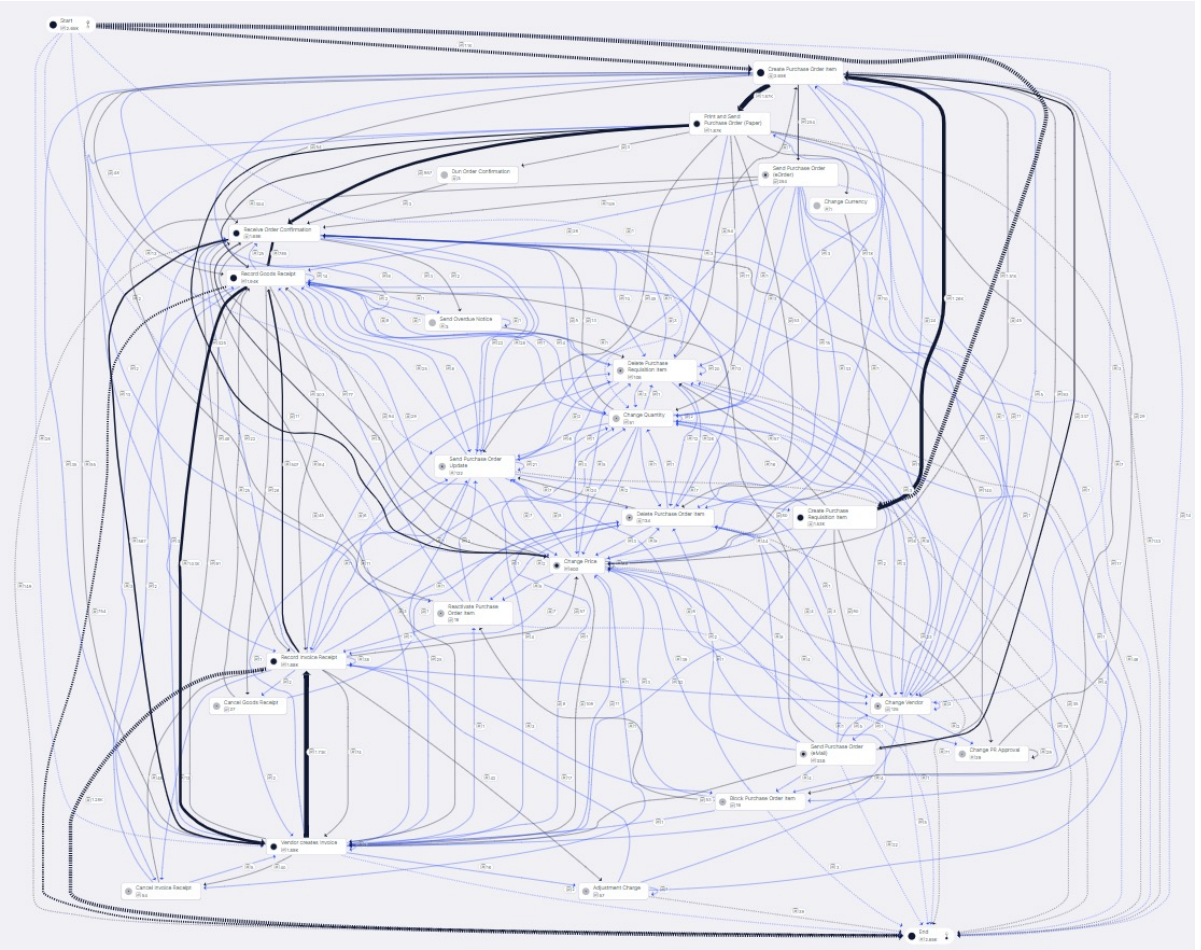


Team Project
Clustering Event Data in Process Mining

UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

Why clustering?

→ grouping similar traces so you can discover cleaner process models per group instead of one tangled "spaghetti" model over everything



Team Project
Clustering Event Data in Process Mining

UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

- **Topic & Goal**
  - Implementing (Python) diverse existing clustering approaches for event data in process mining (e.g. ActiTraC, k-Traceoid, …)
  - Performing benchmarking of the approaches
  - Writing a joint scientific paper (with the goal of a joint publication!)
- Instructor: **Prof. Dr. Daniel Schuster**
- Duration: **6 months** | Credits: **12 ECTS**
- **Hybrid setting (mostly online)**
  - In presence kick-off & final presentations (90min each)
  - Weekly status updates/consultation via Zoom (30min)

Team Project
Clustering Event Data in Process Mining

UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

- **Prerequisites**
  - Fundamental knowledge about process mining (beneficial)
  - Good programming skills in Python (required)

- **Learning objectives**
  - Learn to work in a student team
  - Learn to implement clustering algorithms in Python
  - Learn how to conduct scientific experiments via real-life data
  - Learn to write & contribute to a scientific paper