# CATERPILLAR : Machines & Equipment

PURDUE UNIVERSITY | The Data Mine

Aarav Mangla | Roman Marchuk | Alexander Yang | Suhani Dadhich | Vraj Patel | Siddharth Bhasker | Anjali Pai

## 1. INTRODUCTION

- Caterpillar, a global leader in construction and mining equipment, operates across three primary segments: Construction Industries, Resource Industries, and Energy & Transportation, embodying excellence in innovation and reliability.
- Caterpillar's equipment seamlessly collaborates to accomplish diverse tasks, from excavators loading mining trucks to multiple engines powering marine vessels, showcasing a commitment to efficiency and performance in every endeavor.
- The equipment is electronically controlled using dynamic software that helps engineers understand equipment performance, efficiency, and usage patterns.
- A critical challenge arises when analyzing data from two or more assets – accurate time alignment is crucial for precise analysis of equipment that must work together.
- The project's goal is to devise methods and procedures for accurately time-aligning multiple assets and applying these techniques to real-world examples.

## 2. THE PROBLEM

- Accurately time-aligning data from multiple assets is crucial for precise analysis, especially for synchronized equipment, posing a significant challenge.
- The project seeks to develop methods for precise time alignment to enhance accuracy in real-world analysis scenarios down to the millisecond.

## 3. KEY WORDS

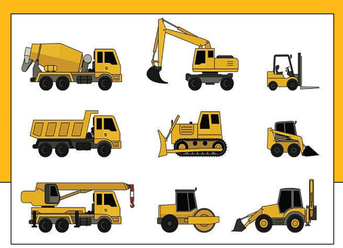
- Hydraulic Fracturing (Fracking): A method used to extract oil or gas from deep underground formations by injecting high-pressure fluid into the rock to create fractures, allowing the extraction of hydrocarbons.
- Machine: Refers to industrial equipment used in hydraulic fracturing operations.
- Total sum flow rate: The sum of the flow rates from all the machines involved in the process.

## 4. BACKGROUND

- Analyzed data frequencies (numerical data) with Python for pattern recognition and anomaly detection and discovering machine behavior.
- Used Tableau to create histograms and visualize datasets for clear conclusions and mastered data visualization.
- Learned to filter, group, and cluster data using Python-enhancing data analysis.
- Analyzed GPS and payload data to pinpoint machine failure origins to analyze root cause.
- Analyzed fracking data to match the time logs of requested flow rates with the actual flow rates recorded from individual trucks, then used Tableau and Python to create a comprehensive chart of total flow rates.

## 5. MOTIVATION

- Enhanced data integration for comprehensive understanding of equipment interactions.
- Empower engineers with accurate data alignment for deeper analysis.
- Drive operational efficiency through preventive maintenance and repair prioritization.
- Pave the way for future innovations by establishing a robust framework for multi-machine data analysis.



## 6. TOOL DEVELOPMENT

**Goal: create a tool that allow users to time align fracking data.**

Fracking trucks work together to meet the requested flow rate at any given moment. We can use this fact to align machine data even if their time series are inaccurate due to faulty behavior.

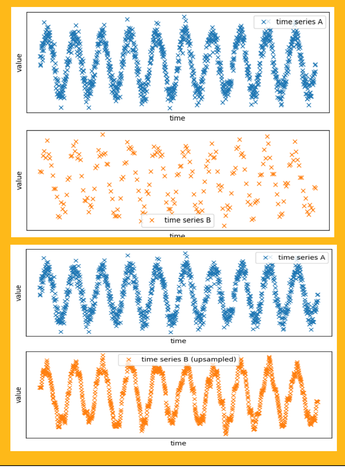
**Alignment Algorithm:**

1. Determine which machine to become the basis of our alignment.
2. Find a major gear shift to -1 that is earliest in time.
3. Align the rest of the machines by finding the first the first time the pump flow rate drops to 0.
4. Store the data in a csv file that contains the aligned machines and the total sum flow rate.


**Messy Real-World Data: Sampling rates**

- Involves increasing the rate of a data stream.
- Match the sampling rates of different signals to allow integration and compatibility.

- **Python's Pathway library**
  - Generate artificial data to time align the series.
  - Aggregate and interpolate close points from the lower-rate time series from the higher-rate series.




## 7. CONCLUSION

- This tool gives Caterpillar an automated solution to align the time values recorded by one machine or multiple machines working together.

- **Features of the tool**:
  - Automated time aligning based on different attributes like:
    - Engine Speed
    - Latitude
    - Longitude
  - The goodness of fit calculated on these attributes.

- **Applications of the tool**:
  - Easier analysis of data among multiple machines working together and precise time aligning.
- **Advantages of the tool**:
  - Reduces manual data analysis, produces accurate results and saves time by automating the process.

## 8. FUTURE GOALS

- Improved tool to identify outlying data points and filtering them.
- Accommodating more machines and incorporating their data.

## ACKNOWLEDGEMENTS

- We would like to thank our mentors, Jim Katter, Anthony Sanders, Josh Moon and everyone else at Caterpillar, for giving us this opportunity to learn.
- We also thank Cai Chen, Doug Crabill, Emily Hoeing, the rest of the Data Mine Staff and our TA, Saumya Verma for their guidance throughout this project.