



Utility Data Collection, Monitoring, and Analytics System



Sponsor: Tupl

Students: Zihao Tao, Sixiang He, Ryan Linden

Faculty Advisor: Baosen Zhang Industry Advisor: Xiang Chen

Problem Statement

Utility companies have vast amounts of data that may contain untapped potential. With better methods of monitoring utilities data, these companies may greatly benefit by increasing their operational efficiency. Our group built a data analytics application for water meters, which can be used to visualize large sets of utilities data and identify potential problems.

Technologies & Workflow

Raw data:
This CSV data file contains 1,418,682 rows for 42610 different meters

Spark: Parallel Data Processing Engine

We converted the CSV file to a Spark DataFrame, processed the data, and then saved the new resulting CSV files

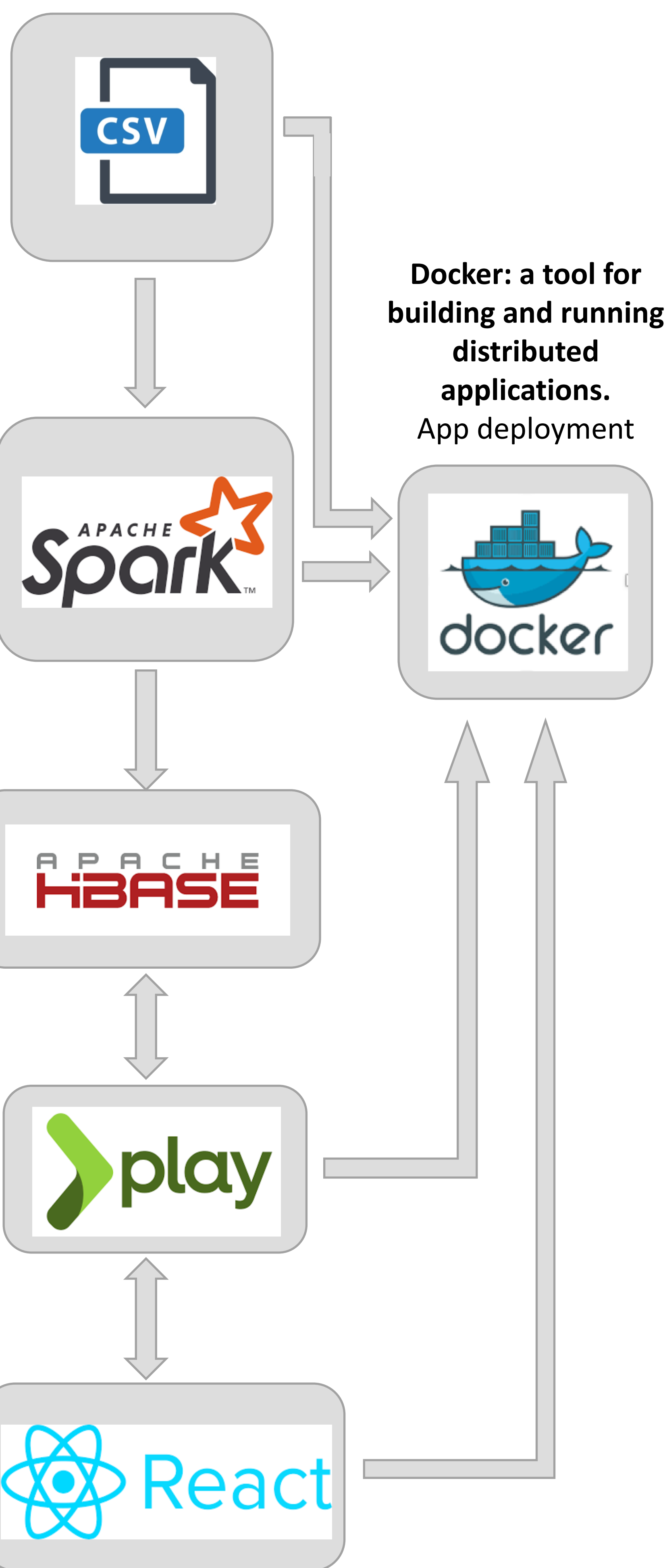
Hbase: Distributed File System
Each resulting CSV file is stored as an HBase table

Play: Web App Framework

1. Reads data from HBase and returns the data in JSON format
2. Updates HBase tables based on user input from the UI.

Ant Design:
A design system that supports ReactJs. Provides advanced UI components that can be used in React

React: A JavaScript library for building user interfaces in web applications.
Used for our UI and interfacing with the Play Framework to access and display HBase data.



Results & Functionality

Our water utility data analysis application has four main pages:
Overview, Meters, Problems, and Machine Learning

Overview Page:

The overview page accumulates data across all of the water meters and displays relevant statistics.

Each box in the box plot represents a specific day, which allows the user to easily visualize daily trends over a specified date range.

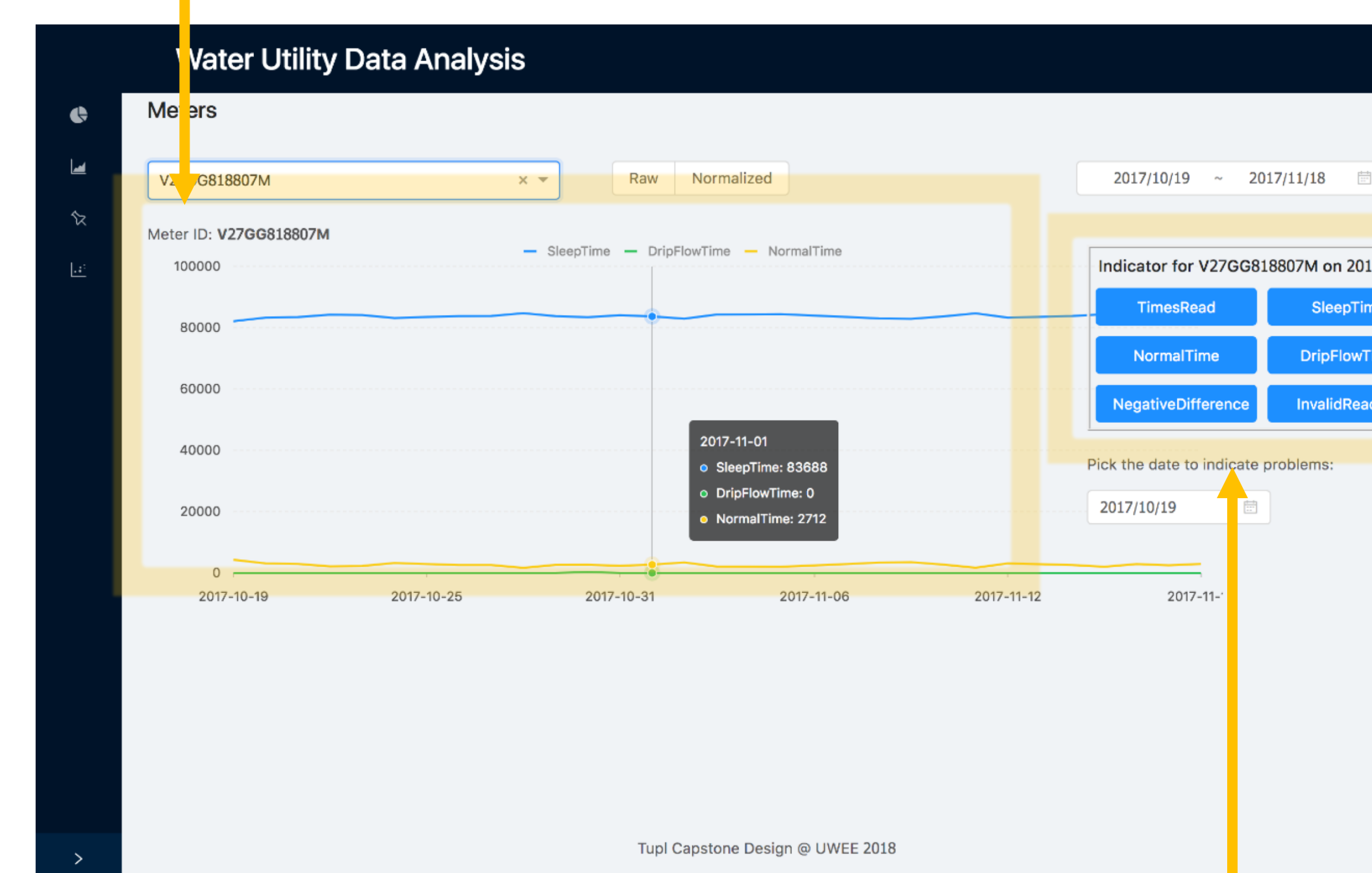


The pie graphs displays the water usage statistics for a single date.

Meters Page:

The Meters page displays information for an individual water meter.

Water usage is displayed in a line plot so the trends of the meter can be easily viewed.



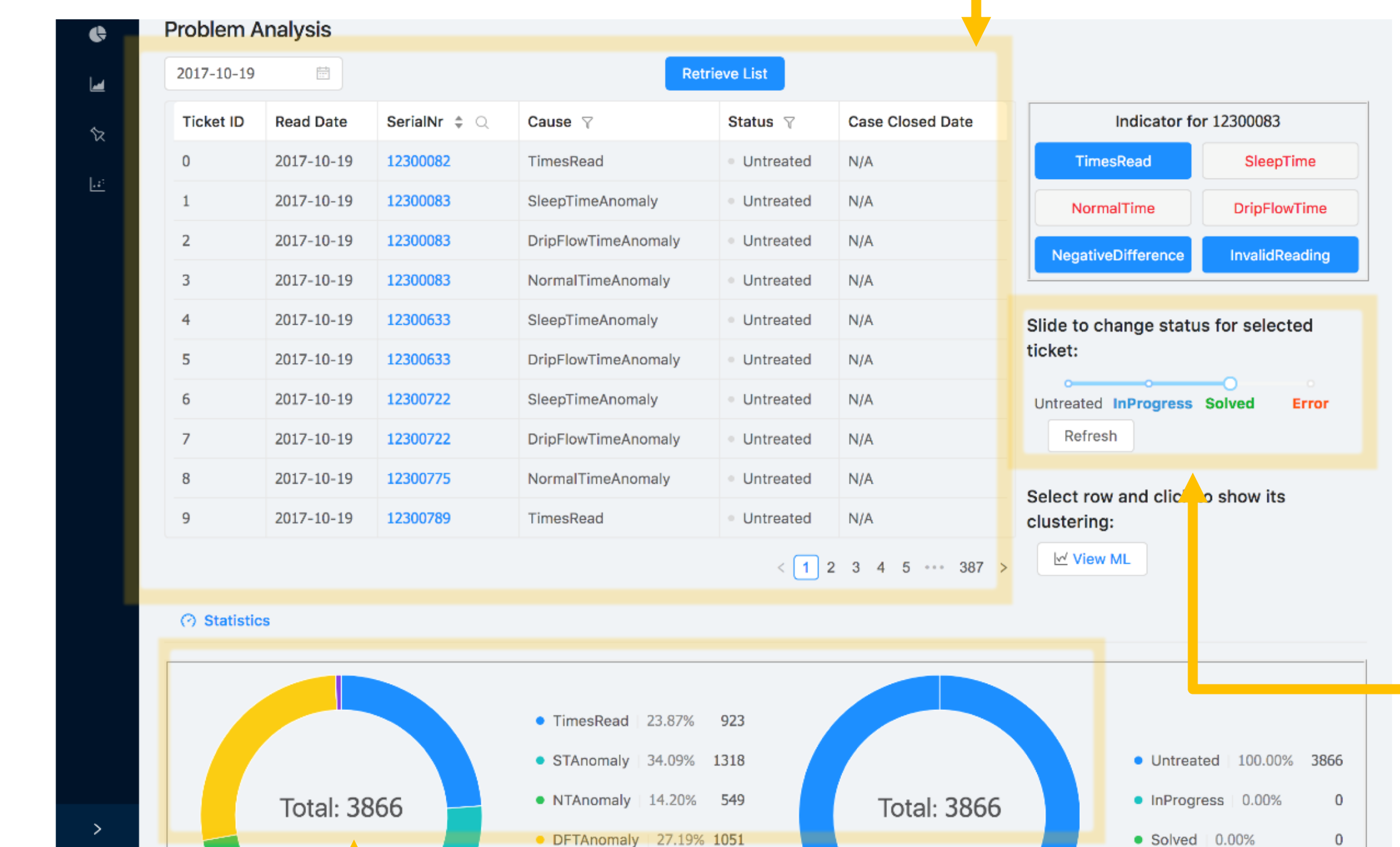
The problems associated with the current meter is displayed via the indicators on the right side of the page.

Results & Functionality Cont.

Problems Page:

Problems associated with specific water meters are displayed on this page.

All detected problems for the water meters are displayed in this list.



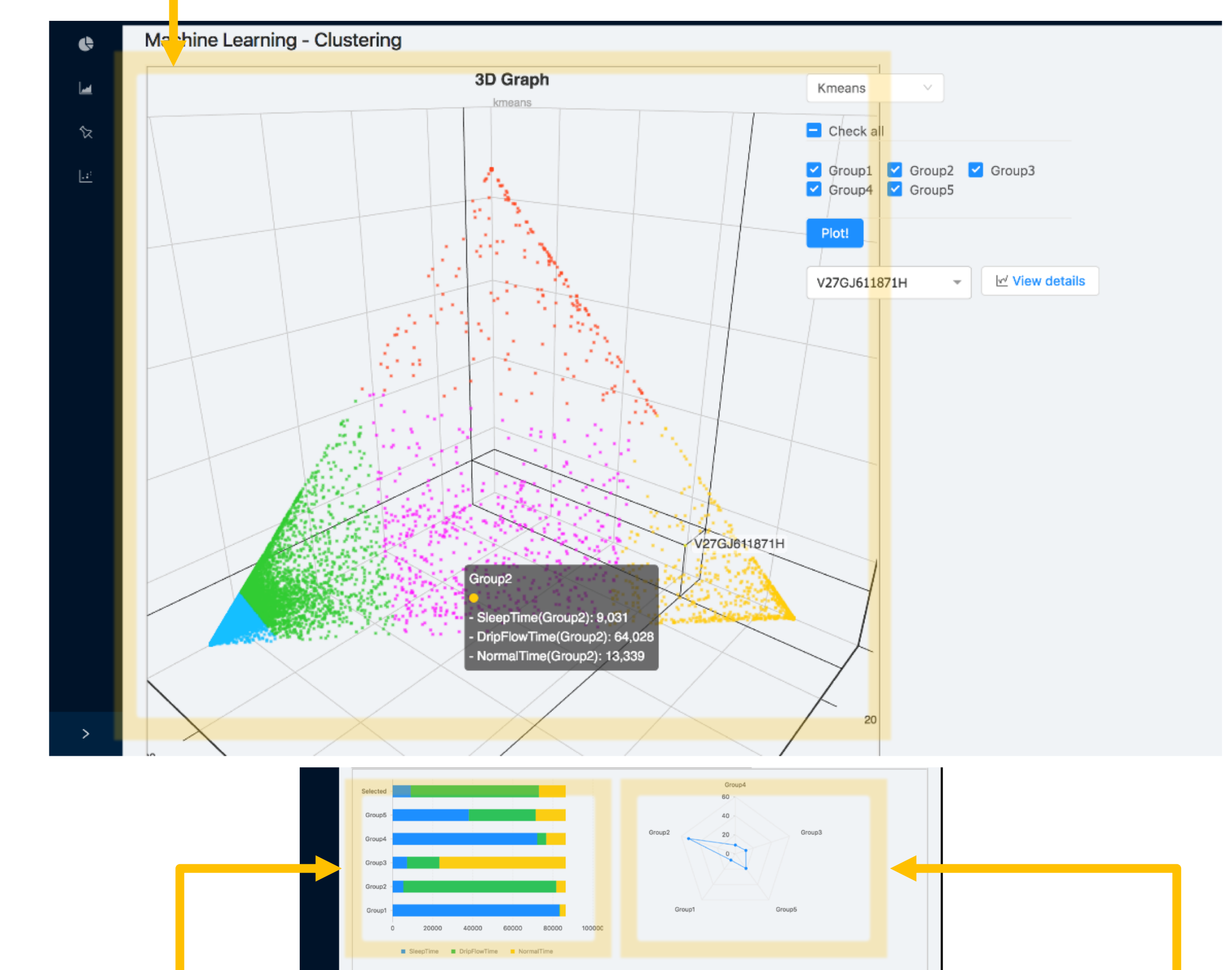
The pie charts help visualize how often each problem occurs on a given date or date range.

A ticket system is used to keep track of each individual problem and its progress.

Machine Learning Page:

The ML page shows the results of clustering the water meters. This helps to identify trends among the different water meters.

The 3D plot shows each individual water meters and its associated color group.



The bar graph displays the cluster centers for each group.

The spider plot shows how strong the relationship is between a meter and the different groups.