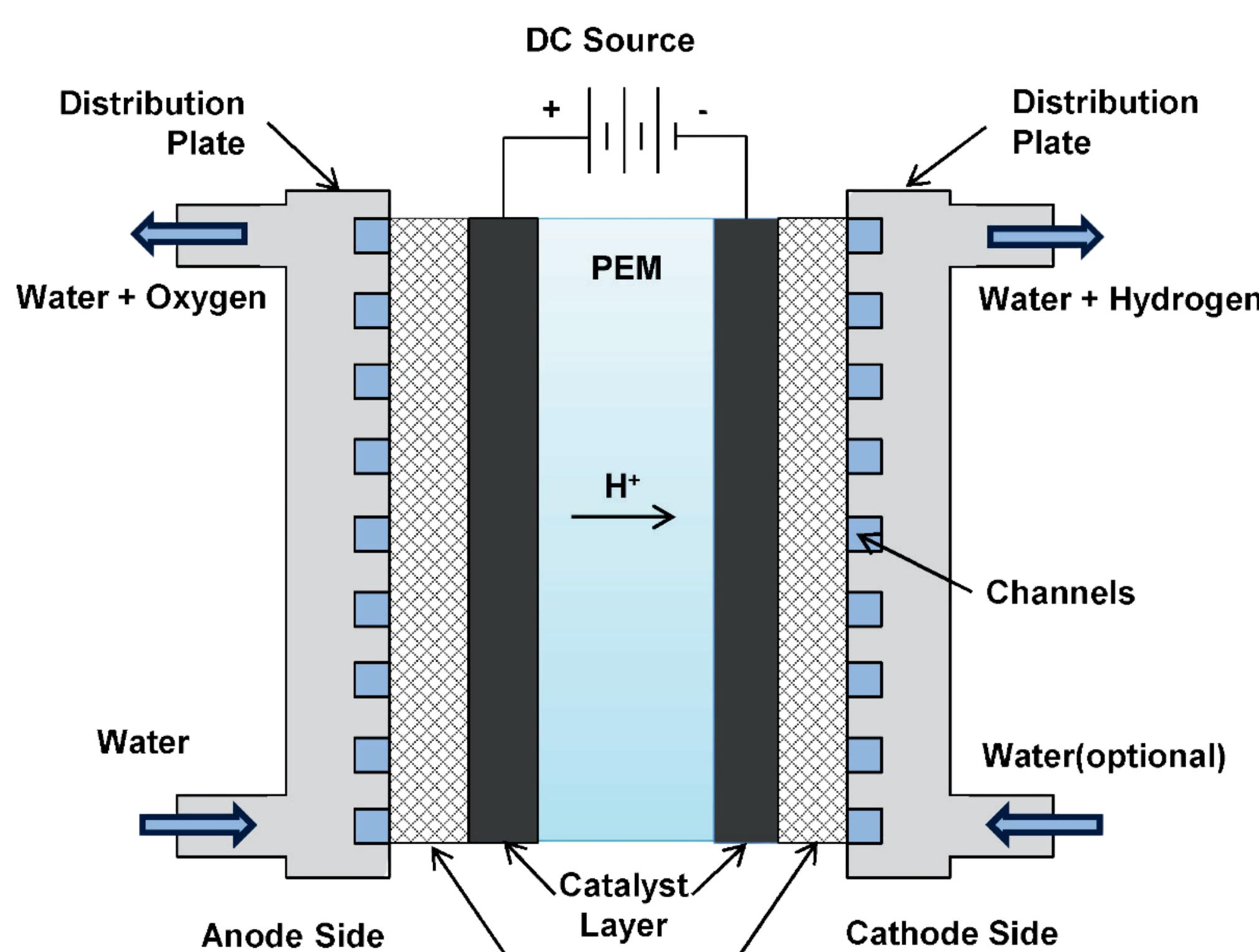


VOLTAGE TO VALUE: ENGINEERING AN ANALYTICS PIPELINE FOR ELECTRODE DATA

Annabel Hiiu, Anett Sandberg, Heili Aavola
Course: Data Engineering (2024)



QUESTIONS THAT CAN BE ANSWERED BASED ON THE DATA

Which material performs better at a specific current density?

How does the data from different instruments compare for the same electrode material?

Which material degraded faster?

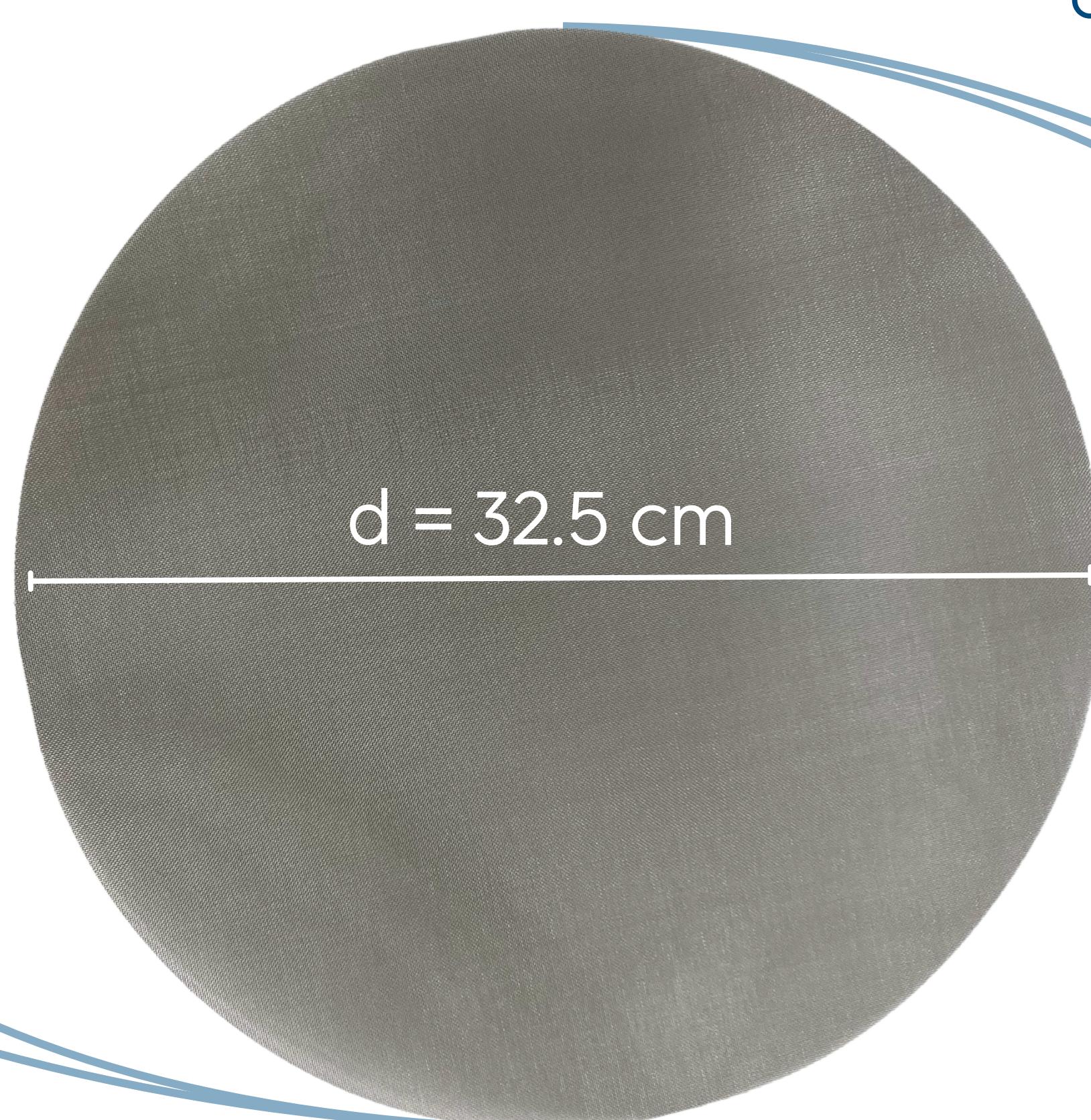
How does the volume of hydrogen produced compare between the two electrode materials?

BACKGROUND

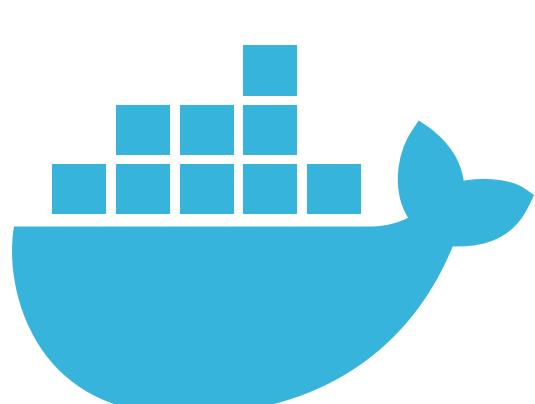
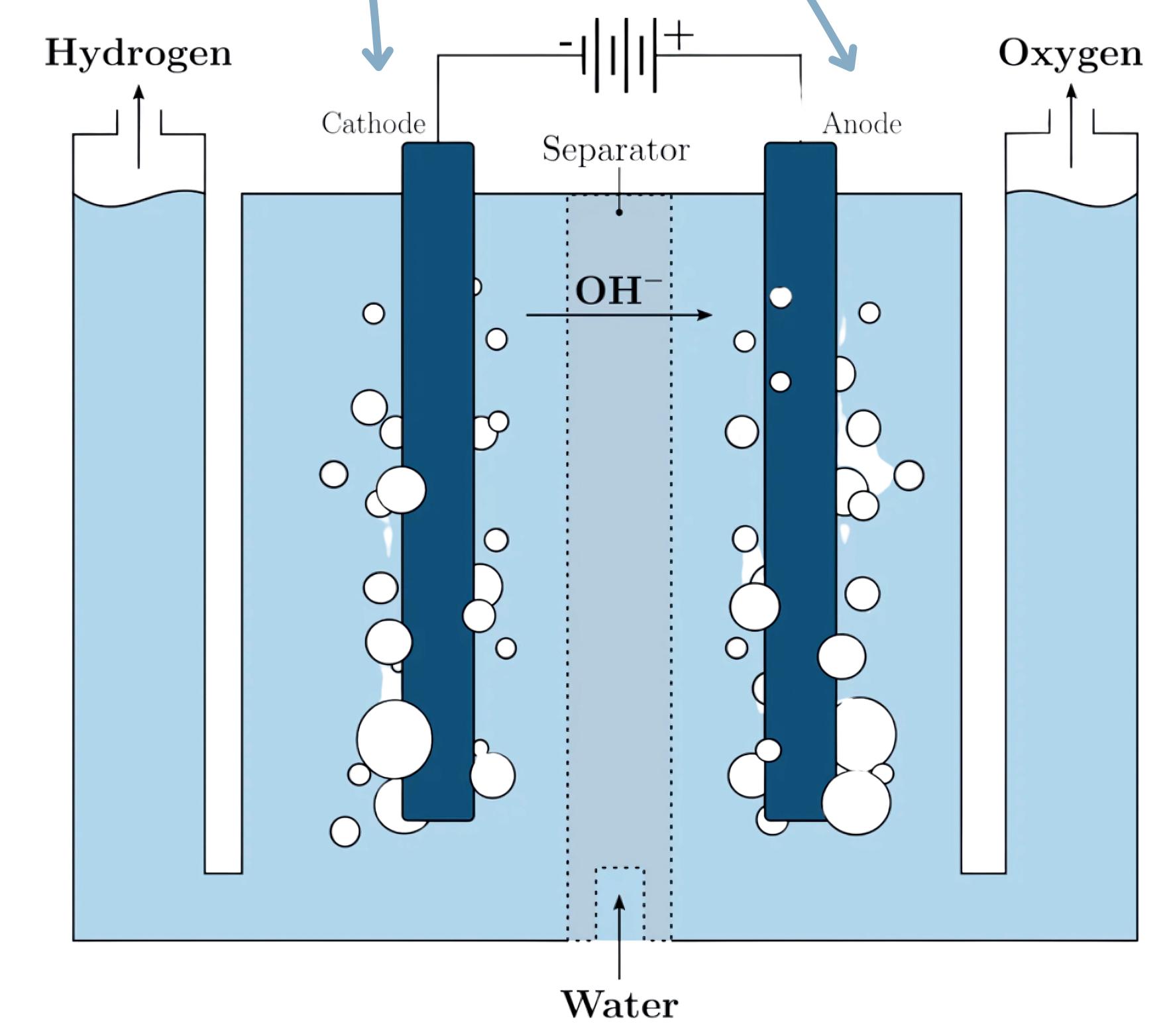
The data for this project comes from the R&D team at OÜ Stargate Hydrogen Solutions. Stargate Hydrogen manufactures electrolyzers, and the data, primarily focusing on different electrodes and their variations, is generated by lab personnel and machines during experiments.

OBJECTIVE

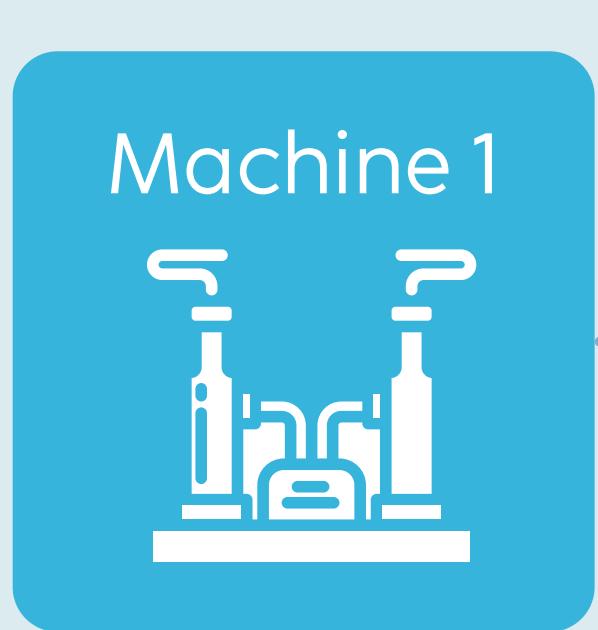
The main objective of the data engineering (DE) project is to take measurement data from two different chemical analysis machines (datasets) and integrate them into a database. Afterwards, visualizations of the data will be used to draw conclusions and make decisions based on the plot.



ELECTRODE(S)



DATA GENERATION



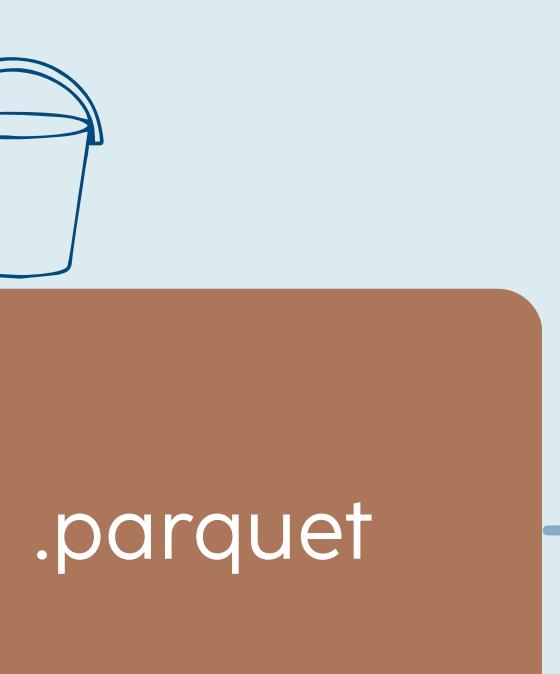
EXTRACT

LOAD

ICEBERG

DuckDB

TASK 1 Apache Airflow



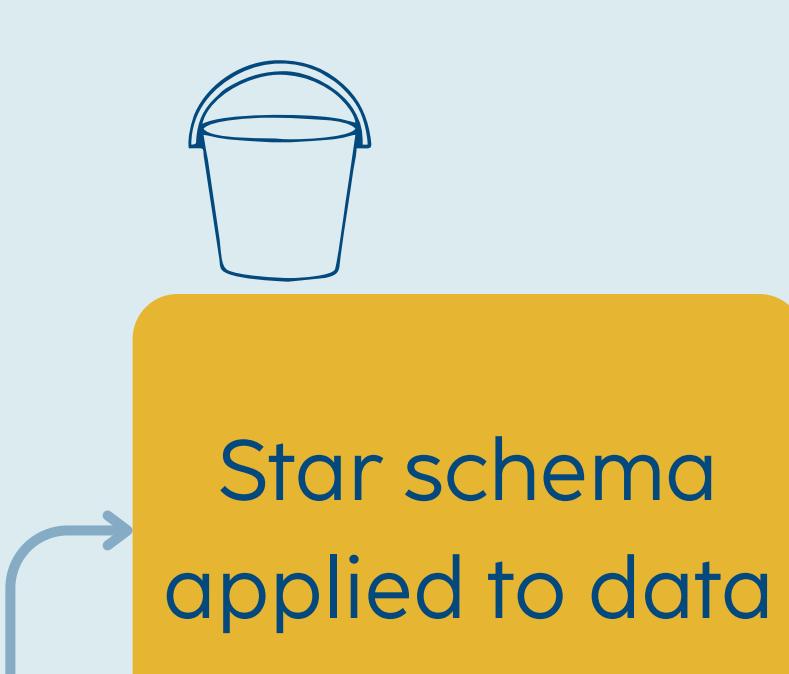
.parquet

TASK 2 Apache Airflow



Iceberg formed .parquet

TASK 3 Apache Airflow



Star schema applied to data



Streamlit

DuckDB

dbt

TRANSFORM

DATA VISUALIZATION