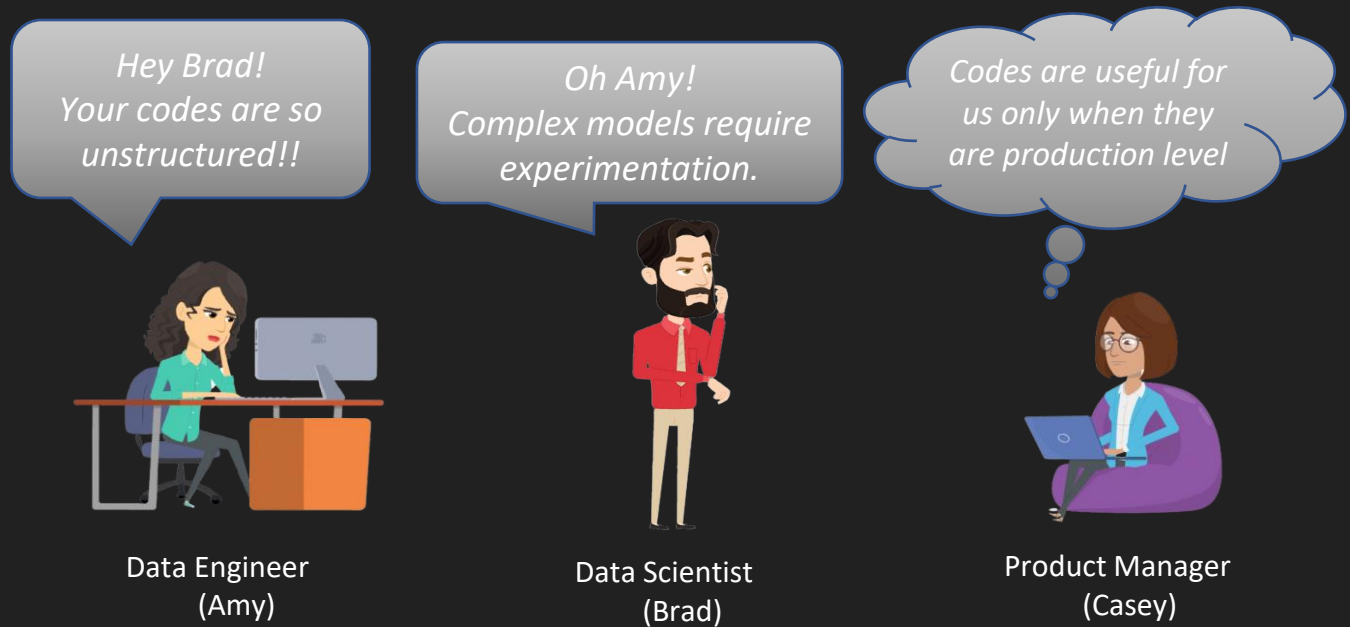


Kedro: Platform Agnostic Data Pipeline

A development workflow tool that allows to create portable data pipelines

Challenges in Data Analytics:



Kedro

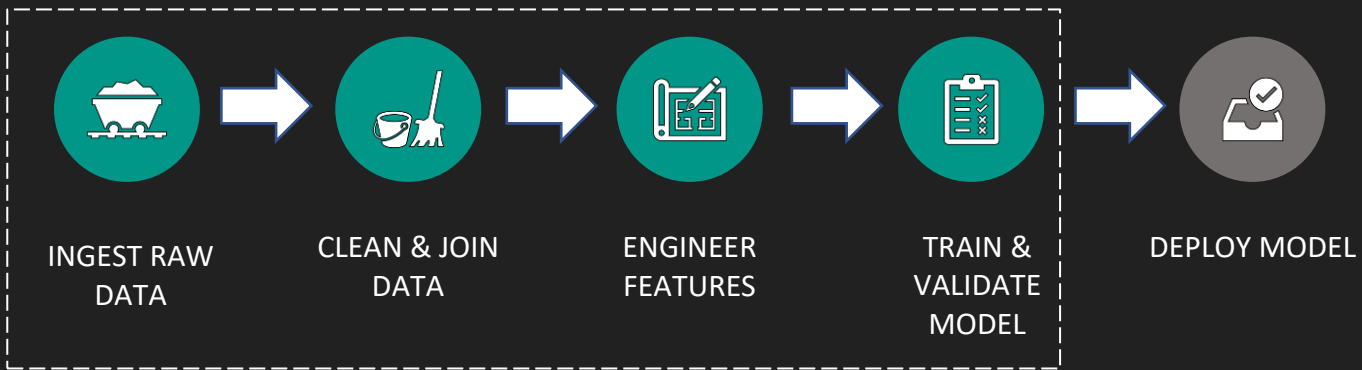
An open-source Python framework for creating reproducible, maintainable and modular data science codes. It implements software engineering best practices like modularity, separation of concerns and versioning

Features :



- **Platform Agnostic:** Migration and deployment across different platforms
- **Finding the balance:** Deployment of ML models applying best software engineering principles to codes

Integration in Data Modeling:



- Kedro fits in the ecosystem across the steps in the workflow of data modeling
- Builds pipelines that automates the redundant data science steps

Comparison across competitor packages :

Features	Airflow	Gokart	Luigi	Kedro
Parallel Execution of pipelines	★	★	★	★
Visualization of pipelines	★★	★	★	★
Built-in various data (file/database) operation (read/write) wrappers			★	★★
Modularity, reusability, testability of data operation			★	★★

★ Good ★ ★ Better

Kedro
Users:



Submitted by:
Brandon P, Pranav C, Pritika K, Sumit K, Tanay D

GitHub Repository:
shorturl.at/yGMRY

