Building a data pipeline is not a trivial task, but the reward of owning your own data and processing is huge for both operational and business outcomes.

Our engineers are experienced in developing data pipelines, both within an existing environment or using the services provided by major cloud platform vendors. By plumbing together the most complicated data systems, we can provide data directly to the hands of business users.

IMPLEMENTATION
We leverage our deep expertise across big data technologies and common programming languages, and work alongside your team to develop a complete end-to-end pipeline, maintain data integrity, and perform data transformations and cleansing.

ETL / ELT PRINCIPLES
There are five key principles to developing a successful data pipeline resulting in optimized time to data, shortened recovery times, data quality and consistency, and data lineage and audit trail.

PARTITION DATA
Optimize underlying operations and segment data with security scheme in mind.

TIMELINES
Eliminate any bottlenecks in processing, favoring micro-service architecture where possible, and allowing SLAs to be met and exceeded.

COMPLETENESS
Ensure validations exist throughout the process to verify metrics such as: expected count in, equals expected count out.

DETERMINISTIC AND IDEMPOTENT
Avoid undesirable outcomes like multiple copies of the same data in your environments. Ensure that results are expected to be the same run over run.

METADATA MANAGEMENT
Track and maintain a log of all job runtimes, data transformation tracking, and statistics.