



# Infrastructure and Analytics as Code

Dion Edge, Senior Cloud Data Architect  
2024-05-14

## Proposal for a Data Product Using Dataform, Looker, and BigQuery

The Data Product is a result of the Analytics Pipeline: the data model includes a semantic layer and data transformations both with version control, sophisticated analytics, practical ML, and data integrity. This solution is intended to be a blueprint for the customer MediaConfidant to build on Google Cloud Infrastructure and Analytics as Code means instructions are provided for provisioning the infrastructure, code snippets as examples, and thorough documentation for every step.

The modern data stack is a collection of tools that work together to help businesses collect, clean, analyze, and visualize data. Dataform, Looker, and BigQuery create a powerful analytical stack but are only part of the solution. This document outlines the necessary supporting infrastructure with the following objectives:

- Improved data insights: BigQuery's analytics engine can be used with Looker to produce more than just data.
- Improved data quality: Dataform can be used to clean and validate data before it is loaded into BigQuery.
- Increased data accessibility: Looker provides a user-friendly interface for users to access and explore data.
- Data integration: Dataform will be used to integrate data from multiple sources into BigQuery.

