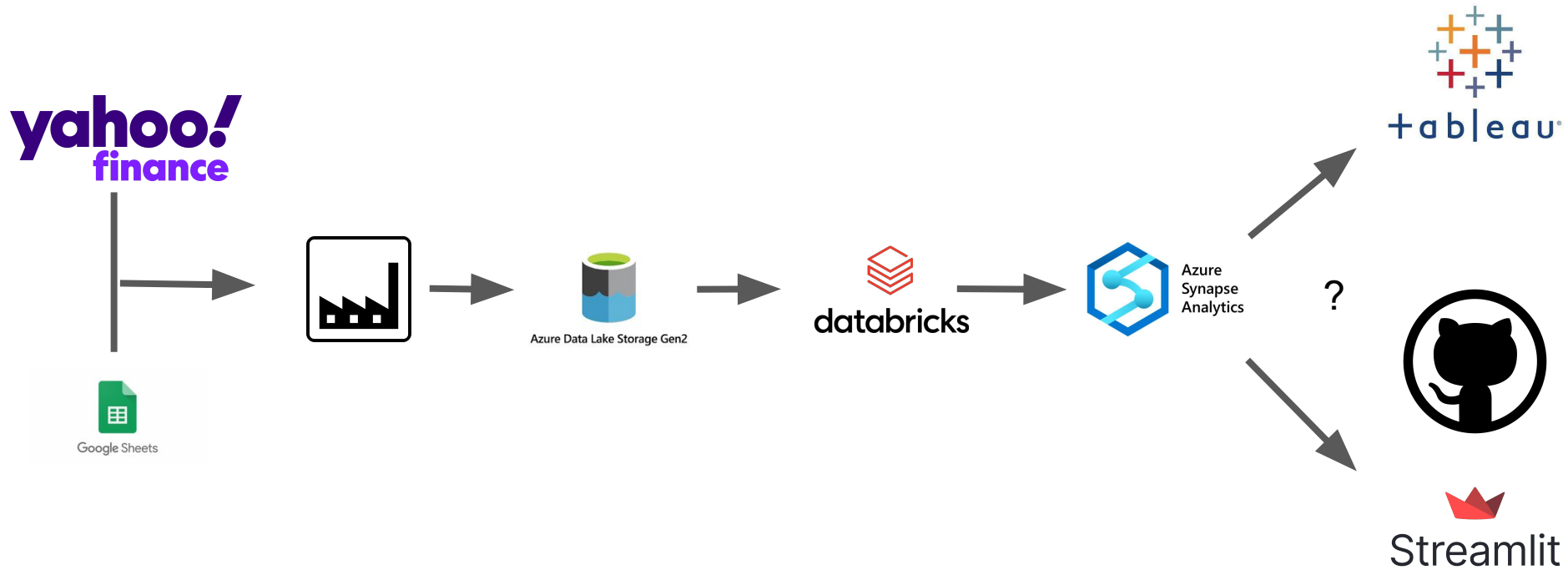


# Pipeline DAG



# Project Goals

- Utilize personal investment data to have a public facing dashboard for accountability and analysis of portfolio performance
- Gain cloud data infrastructure experience (ETL Pipeline)
- Realistic and purposeful data modeling
- Minimize costs. Free to maintain if possible, otherwise look to refactor later with services to lower upkeep costs

# Tools and Use Cases

- CI/CD
  - GitHub Actions?
- Source Data
  - Personal stock transactions
    - Google Sheets API
  - Current/Historical stock prices
    - Yahoo Finance API
- Orchestration Tooling
  - Azure Data Factory
- Object storage of raw data
  - Azure Data Lake Gen2

# Tools and Use Cases

- Transformation
  - Azure Databricks with spark
- OLAP Datawarehouse
  - SQL in Azure Synapse Analytics
- Interactive Dashboard
  - Tableau or Streamlit/GitHub Pages