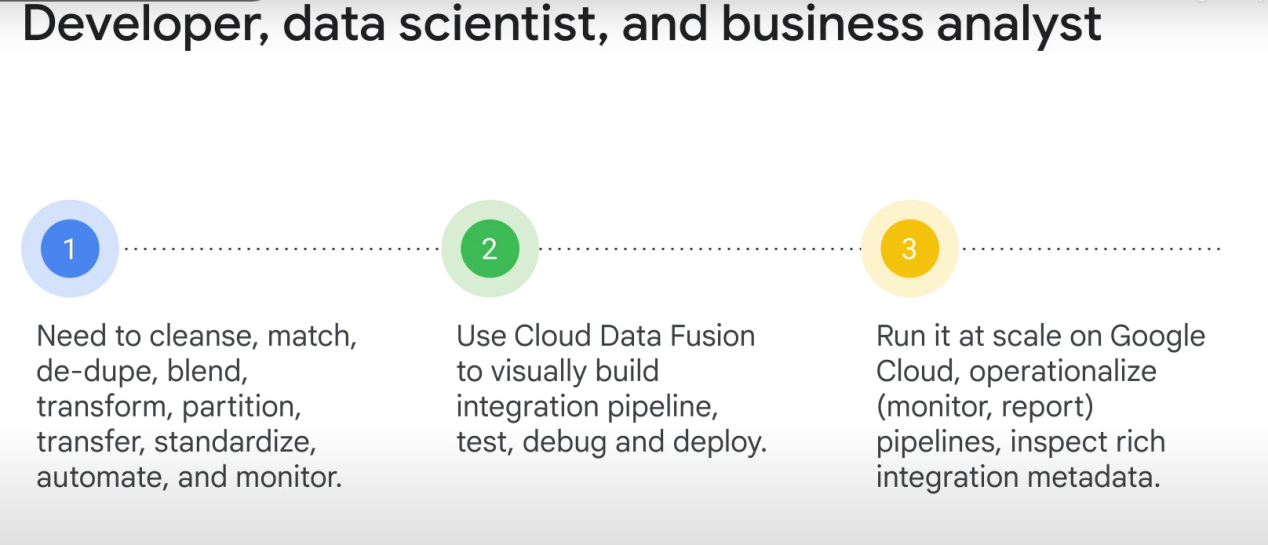
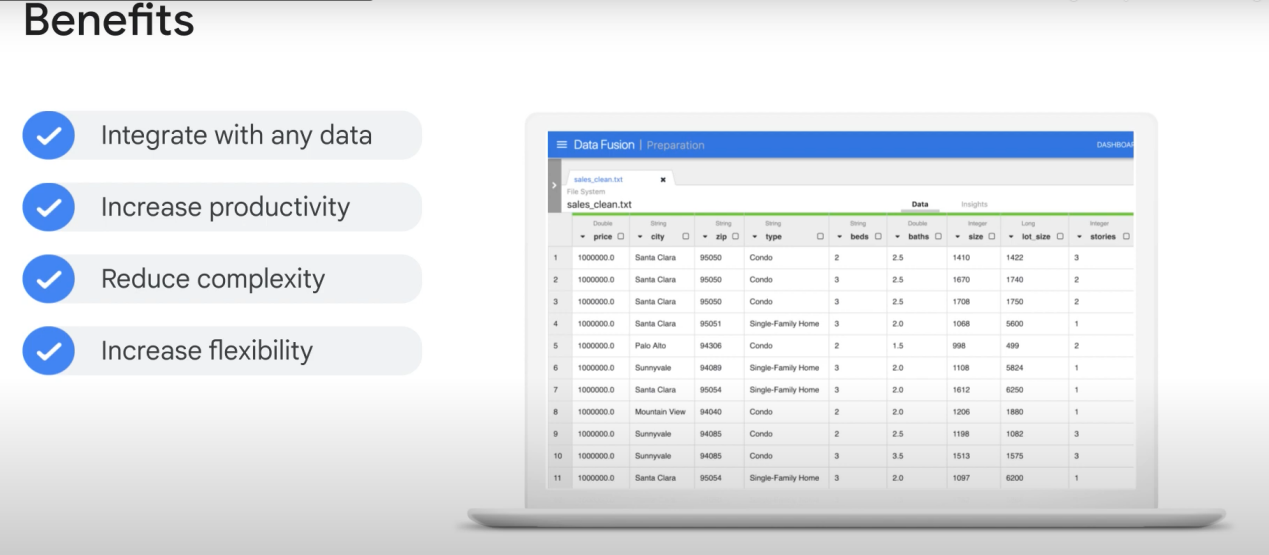


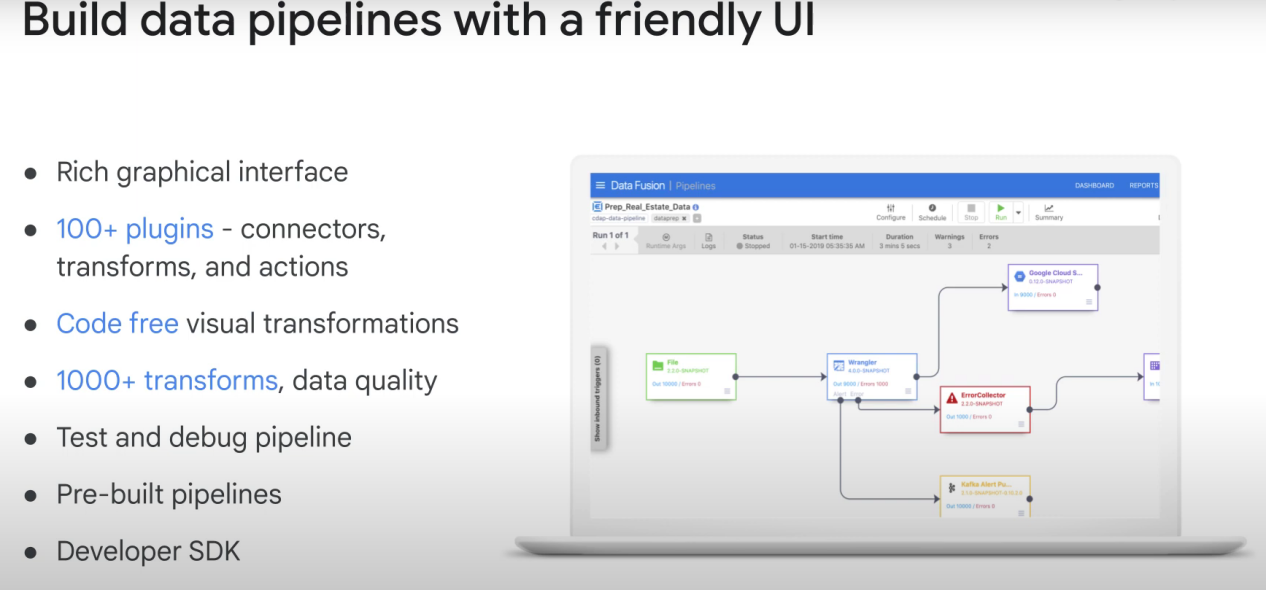
What does it do?

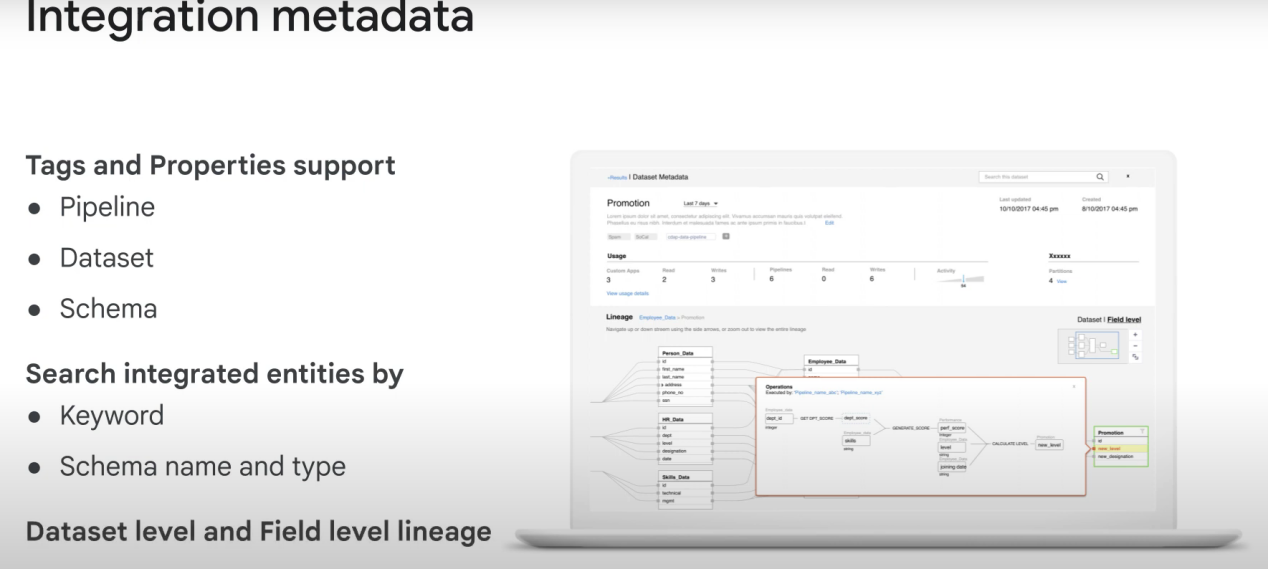


Benefits:



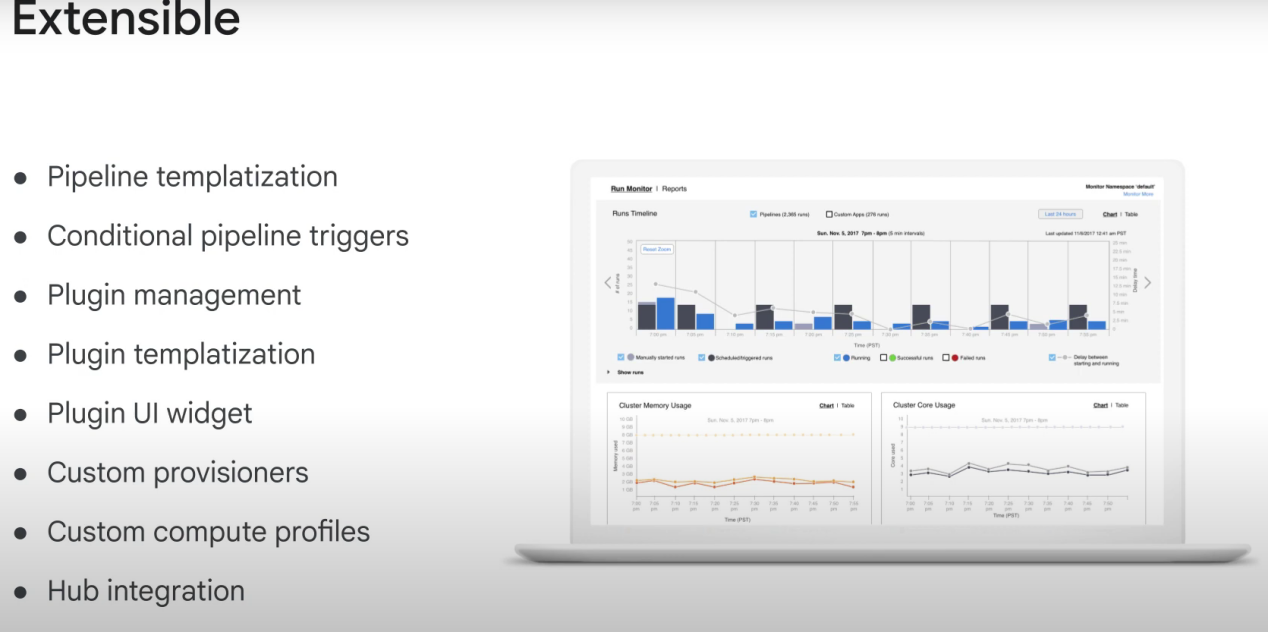
\*Merge data from different sources and create a common view

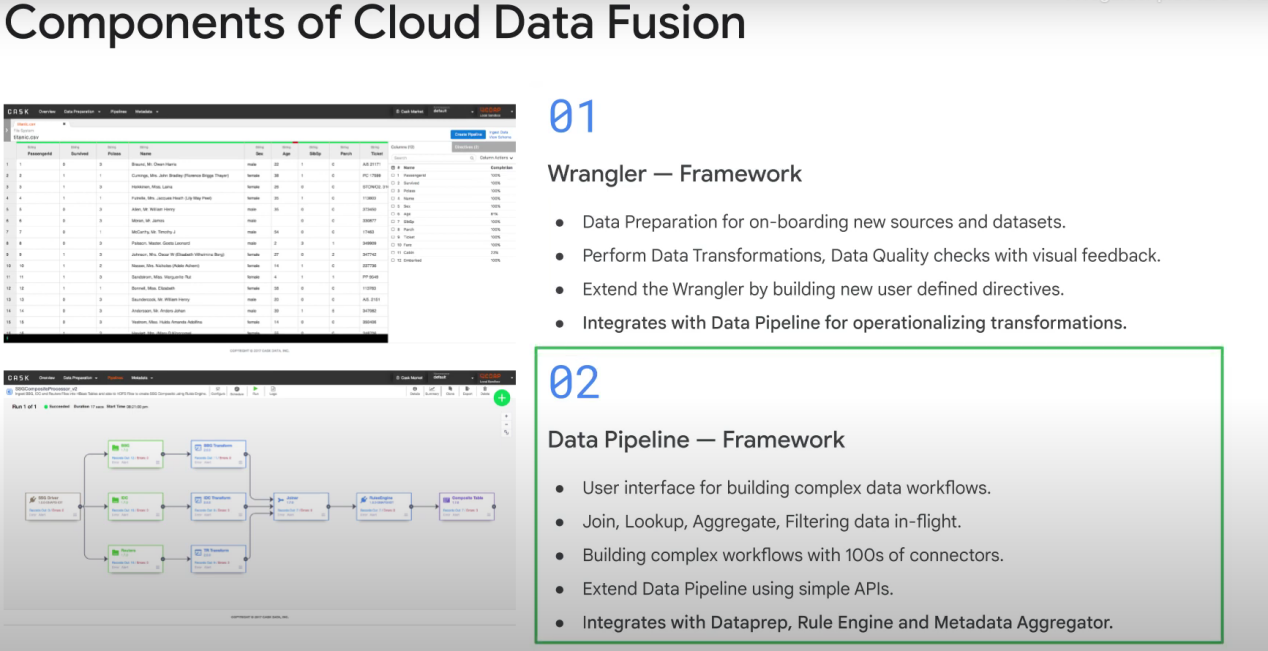




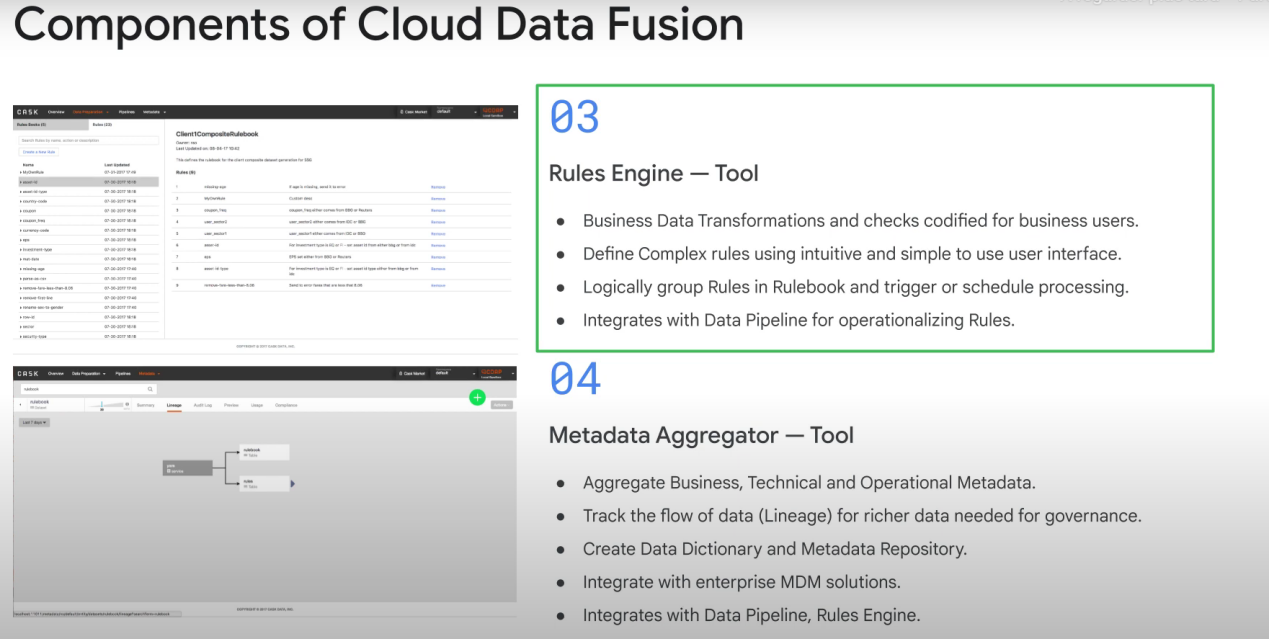
**LINEAGE!**

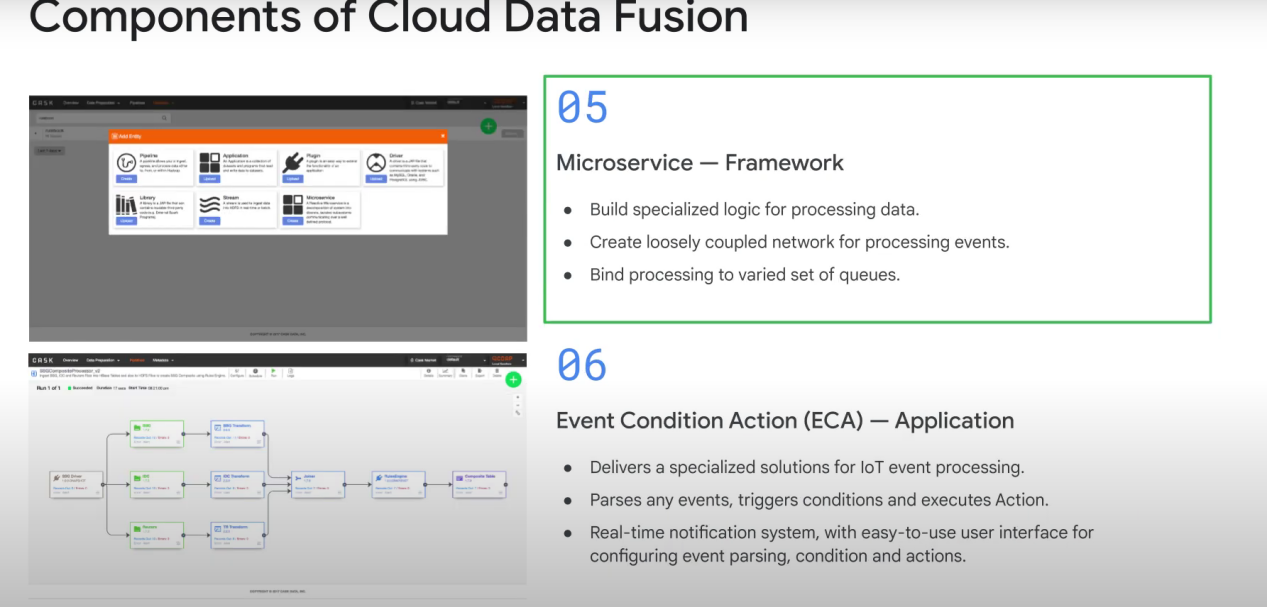
Cloud Fusion is extensible:





Data fusion (particulary wrangler I think) allows you to do everything from an easy-to-use UI.





Note:

When you run a pipeline, **Cloud Data Fusion provisions an ephemeral Cloud Dataproc cluster, runs the pipeline, and then tears down the cluster.** This could take a few minutes. You can observe the status of the pipeline transition from Provisioning to Starting and from Starting to Running to Succeeded during this time.