Data Analytics

**BigQuery**: BigQuery is a fully-managed data analysis service that enables businesses to analyze Big Data. It features highly scalable data storage that accommodates up to hundreds of terabytes, the ability to perform ad hoc queries on multi-terabyte datasets, and the ability to share data insights via the web.

**Cloud Composer**: Cloud Composer is a managed workflow orchestration service that can be used to author, schedule, and monitor pipelines that span across clouds and on-premises data centers. Cloud Composer allows you to use Apache Airflow without the hassle of creating and managing complex Airflow infrastructure.

**Cloud Data Fusion**: Cloud Data Fusion is a fully-managed, cloud native, enterprise data integration service for quickly building and managing data pipelines. Cloud Data Fusion provides a graphical interface to help increase time efficiency and reduce complexity and allows business users, developers, and data scientists to easily and reliably build scalable data integration solutions to cleanse, prepare, blend, transfer, and transform data without having to wrestle with infrastructure.

**Cloud Life Sciences (formerly Google Genomics)**: Cloud Life Sciences provides services and tools for managing, processing, and transforming life sciences data.

**Data Catalog**: Data Catalog is a fully-managed and scalable metadata management service that empowers organizations to quickly discover, manage, and understand their data in Google Cloud. It offers a central data catalog across certain Google Cloud Services that allows organizations to have a unified view of their data assets.

**\*Data Studio**: Data Studio is a data visualization and business intelligence product. It enables customers to connect to their data stored in other systems, create reports and dashboards using that data, and share them throughout their organization.

**Dataplex**: Dataplex is an intelligent data fabric that helps customers unify distributed data and automate management and governance across that data to power analytics at scale.

**Dataflow**: Dataflow is a fully-managed service for strongly consistent, parallel data-processing pipelines. It provides an SDK for Java with composable primitives for building data-processing pipelines for batch or continuous processing. This service manages the life cycle of Compute Engine resources of the processing pipeline(s). It also provides a monitoring user interface for understanding pipeline health.

**Datalab**: Datalab is an interactive tool for exploration, transformation, analysis and visualization of your data on Google Cloud Platform. It runs in your cloud project and enables you to write code to use other Big Data and storage services using a rich set of Google-authored and third party libraries.

**Dataproc**: Dataproc is a fast, easy to use, managed Spark and Hadoop service for distributed data processing. It provides management, integration, and development tools for unlocking the power of rich open source data processing tools. With Dataproc, you can create Spark/Hadoop clusters sized for your workloads precisely when you need them.

**Dataproc Metastore**: Dataproc Metastore provides a fully-managed metastore service that simplifies technical metadata management and is based on a fully-featured Apache Hive metastore. Dataproc Metastore can be used as a metadata storage service component for data lakes built on open source processing frameworks like Apache Hadoop, Apache Spark, Apache Hive, Presto, and others.

**Datastream**: Datastream is a serverless change data capture (CDC) and replication service that enables data synchronization across heterogeneous databases, storage systems, and applications with minimal latency.

**Google Earth Engine:**Google Earth Engine is a platform for global-scale analysis and visualization of geospatial datasets. Google Earth Engine can be used with custom datasets, or with any of the publicly available satellite imagery hosted (and ingested on a regular basis) by Earth Engine Data Catalog.

**Pub/Sub**: Pub/Sub is designed to provide reliable, many-to-many, asynchronous messaging between applications. Publisher applications can send messages to a "topic" and other applications can subscribe to that topic to receive the messages. By decoupling senders and receivers, Pub/Sub allows developers to communicate between independently written applications.