Commercial Analytics on AWS

Customer Challenges:

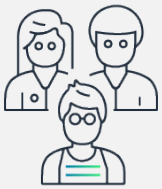
Healthcare and Life science organizations need to derive information from operational datasets while limiting the burden of data transformation on staff members. Applying analytics across operational data can provide business insights to improve profitability while serving the organization's mission. Though storing business operation data is key to these initiatives, transforming and securely sharing relevant data can still be a challenge.

AWS Health for Data:

AWS makes it easier for customers to share, search, and derive predictive insights at scale across organizational boundaries, allowing them to identify new opportunities to better run their businesses. Customers can use AWS Glue DataBrew to clean and normalize data using automated data preparation tasks to save time. Once the data is prepared, customers can soon use Amazon DataZone to unlock data across the organization securely and integrate into analytics tools such as Amazon Elastic Map Reduce with a simple to use portal.

Top Benefits:

Automate data preparation tasks



Unlock data across the organization

Built-in governance of data

Unify business intelligence

Featured AWS Services:

AWS Glue

Serverless data integration to discover, prepare, and integrate data at any scale

Amazon Elastic Map Reduce

Easily run petabyte-scale analytics faster

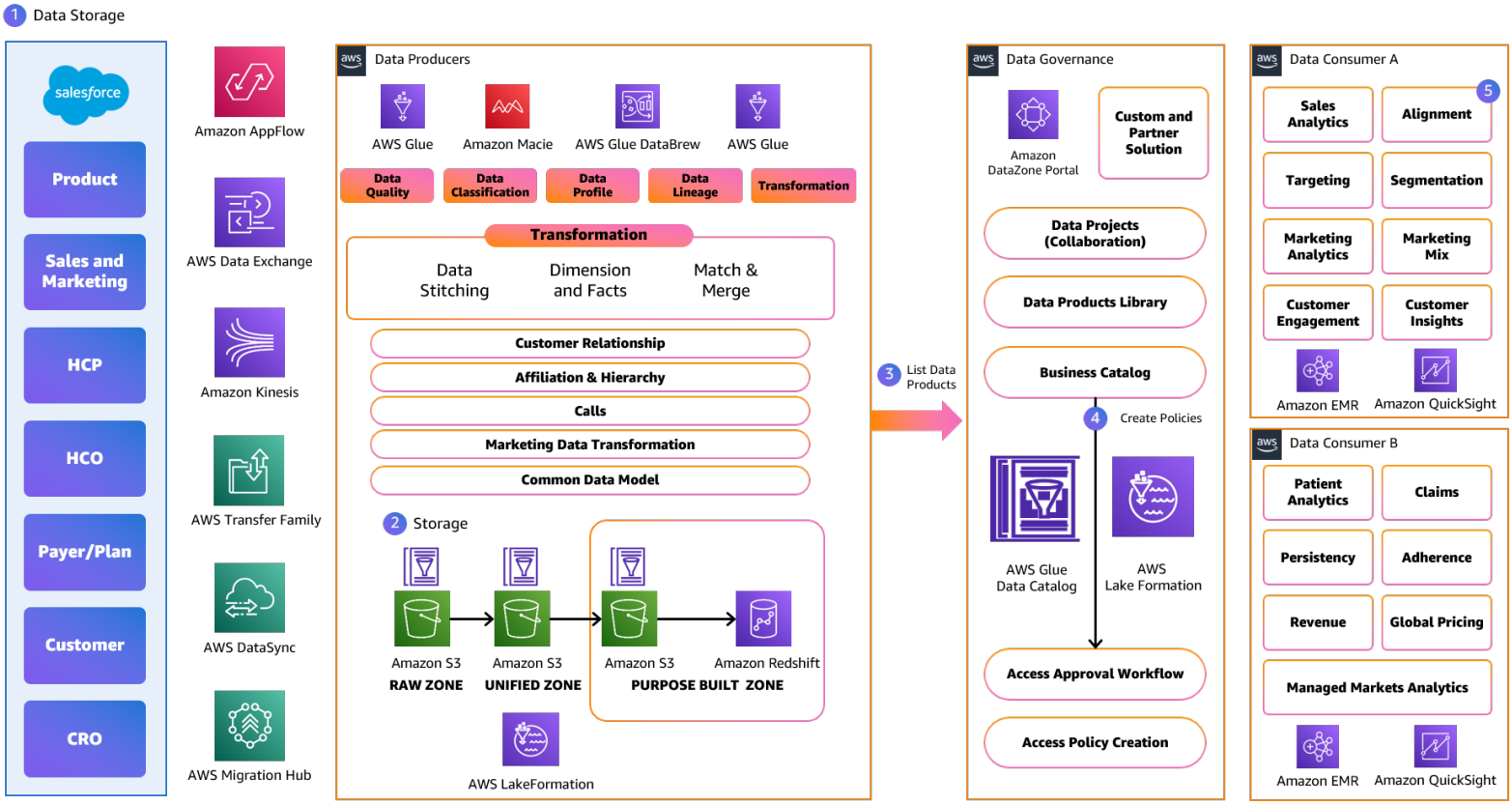
Amazon DataZone

Unlock data across organizational boundaries with built-in governance

Amazon QuickSight

Unified business intelligence at hyperscale

Commercial Analytics Architecture on AWS



1/ Diverse set of data sources are ingested using appropriate data ingestion services.

2/ Ingested data is stored in raw Amazon S3 layer and then cataloged using AWS Glue data catalog. Using fit-for-purpose data processing services, the ingested data is cleaned, curated, classified, transformed and stored in a purpose-built storage that can be published as a data product.

3/ Data products are published to the enterprise business catalog on Amazon DataZone along with metadata (schema/data quality metrics/lineage etc.). The schema is also shared with the central data catalog running on AWS Glue. AWS Lake Formation has a feature for sharing data catalogs across multiple accounts that can leveraged here.

4/ Amazon DataZone self-serve workflow accepts consumer request for data products. AWS Lake Formation shared database and tags are utilized for fine grained access controls.

5/ Data consumers access data to run and build analytics based on commercial KPIs.

Learn more about how AWS capabilities are used in Life Sciences:
https://aws.amazon.com/health/life-sciences/solutions/