



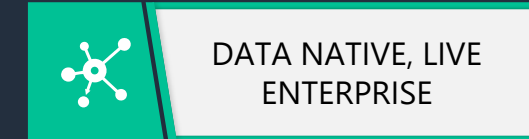
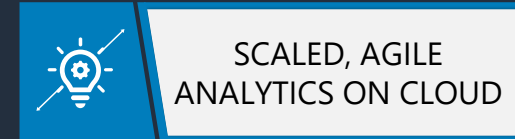
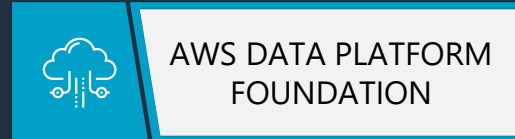
Streaming Analytics Patterns

Modernization Week

Infosys

Introduction to Data & Analytics @ Infosys

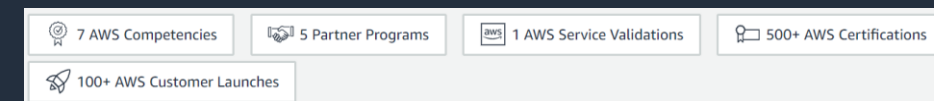
Service Offerings



APN Credentials



APN Partner since 2011
3rd Global SI Partner to achieve **AWS Data and Analytics Competency**



Analyst Recognition



Positioned as a **Leader** in the **Gartner MQ - Data and Analytics Services 2021**

Patterns of Streaming Use Cases



Streaming Data into Lake

- Fashion retailer streams data to data lake for analytics



Real Time Sales Decisions

- Premium retailer uses streamed sales data for real time decisions

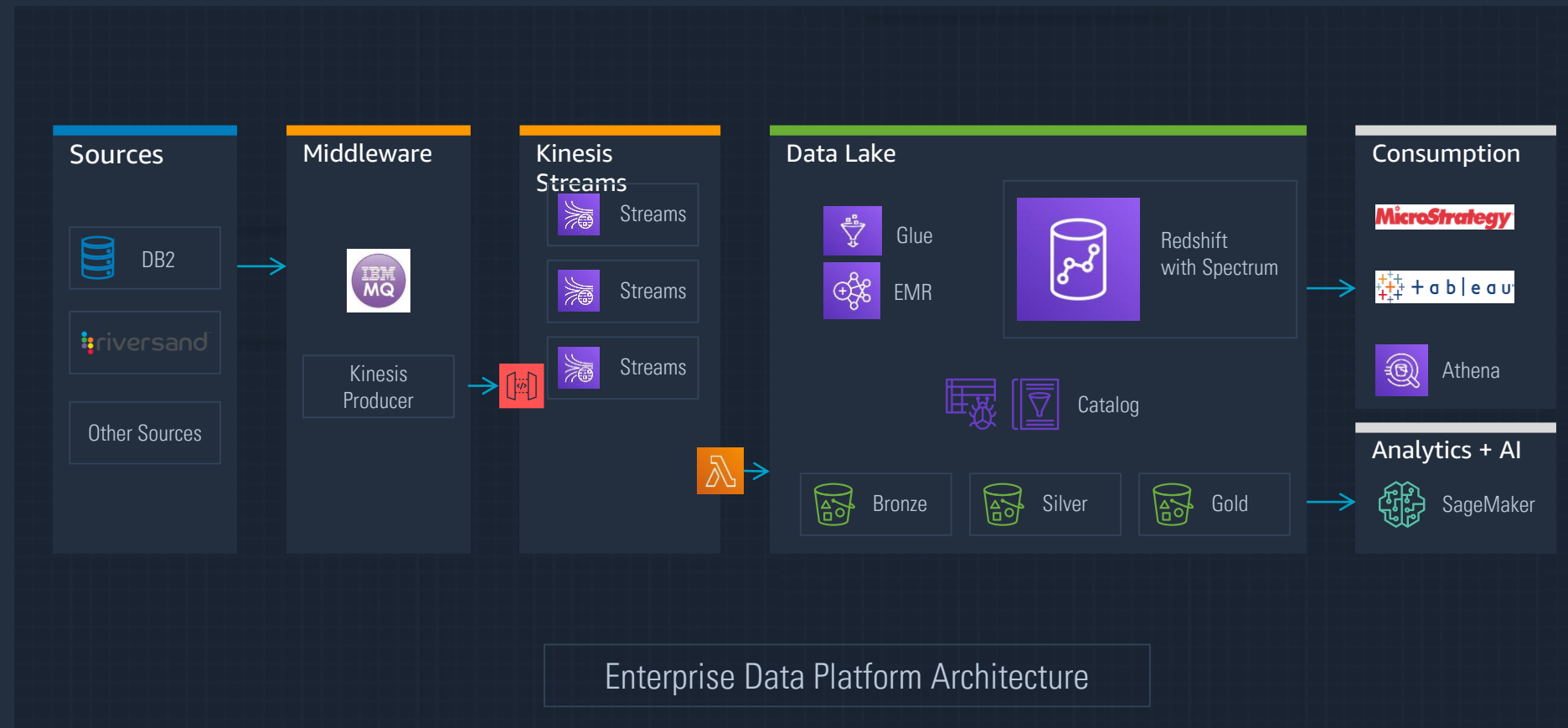


Real Time Streaming Analytics

- Implementing Industry 5.0 for Pharma Manufacturing process and Insurance Analytics

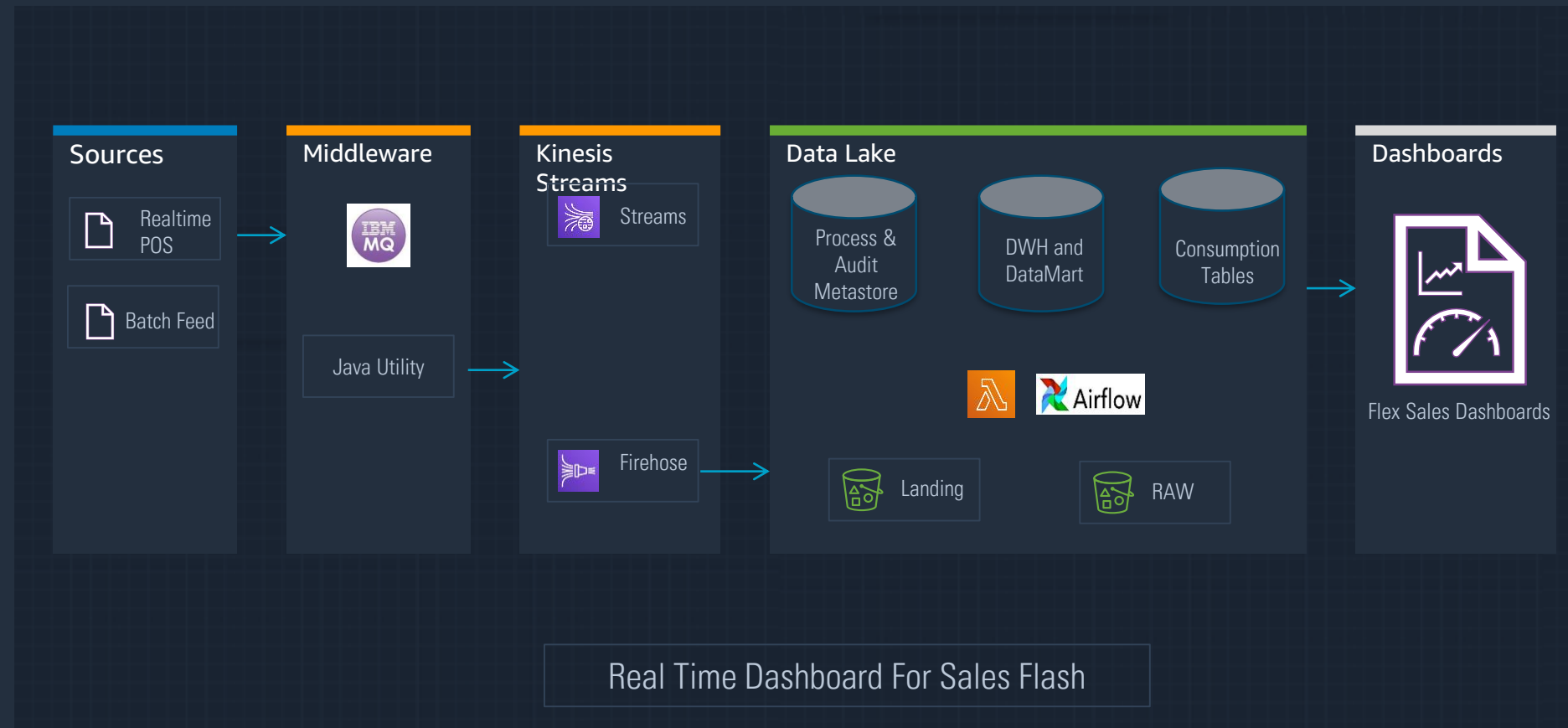
Streaming Data into Data Lake

- Customer is a US based fashion retailer with operations around the globe who undertook a data estate modernization initiative
- Data captured using Change Data Capture (CDC) process from sources on premise and MDM serviced as SaaS on another cloud platform among others
- Arrives in enterprise-wide Information Bus to cater to variety of data integration needs
- Data tapped by a bridge agent and flown real time into Kinesis Streams
- Serverless processes pick it up from streams to massage and deliver to data lake
- After further processing using Glue and EMR on data lake, the data is available for consumption using BI tools and Athena and advanced analytics using SageMaker



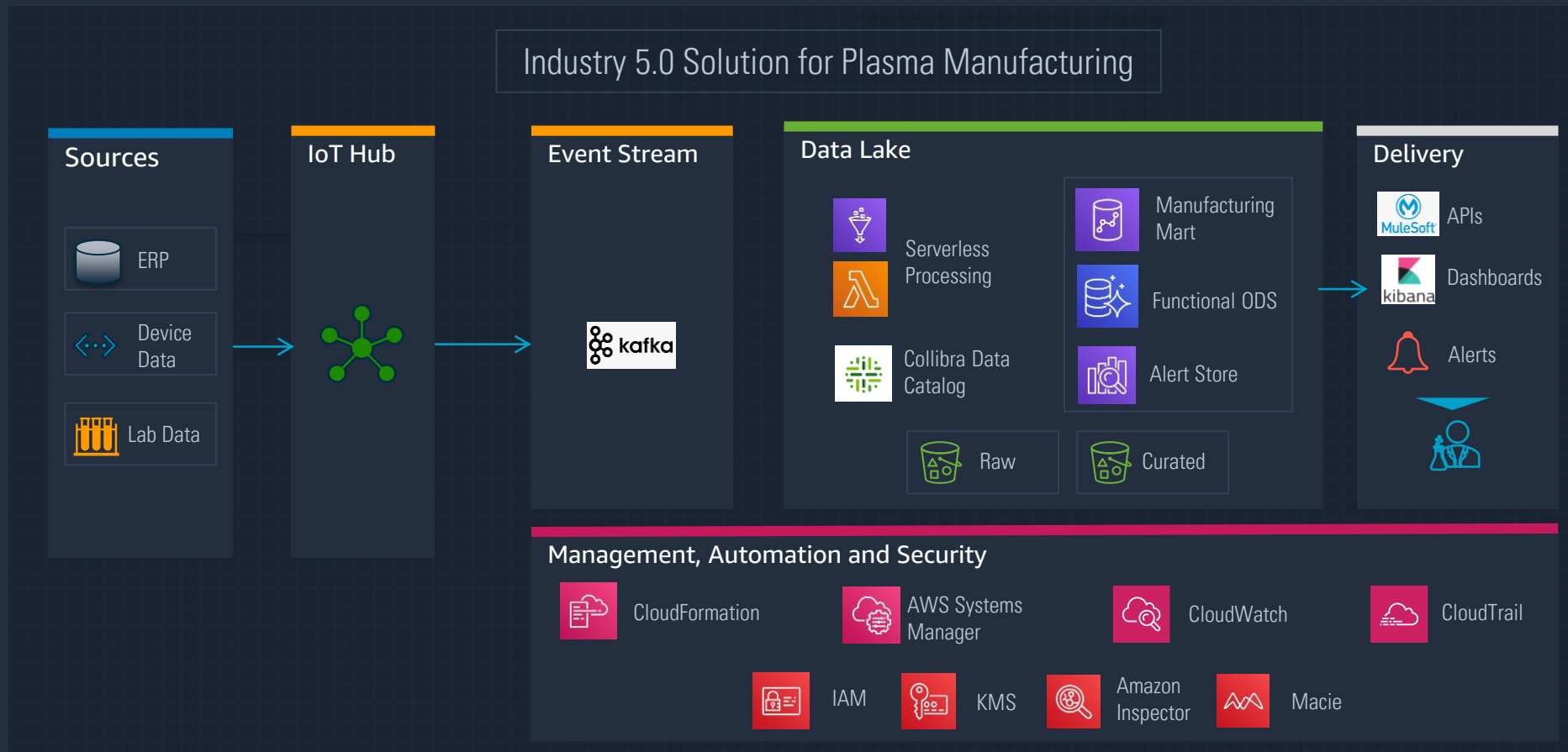
Streaming for Real Time Dashboards

- Customer is a US based premium fashion retailer who undertook a data platform modernization initiative.
- The unaudited real-time POS sales data is consumed by Amazon Kinesis and spilled over to S3 Bucket which in turn is ingested to Dashboard Layer using Lambda and Airflow every 15 minutes.
- The audited sales data as of yesterday is received every day in S3 Bucket. This is processed and loaded to DWH and Datamart tables using Lambda and Airflow once in a day. This is summarized and loaded to Dashboard layer.
- This unaudited and audited sales data is available for end user for consumption as Flex Sales Dashboard.



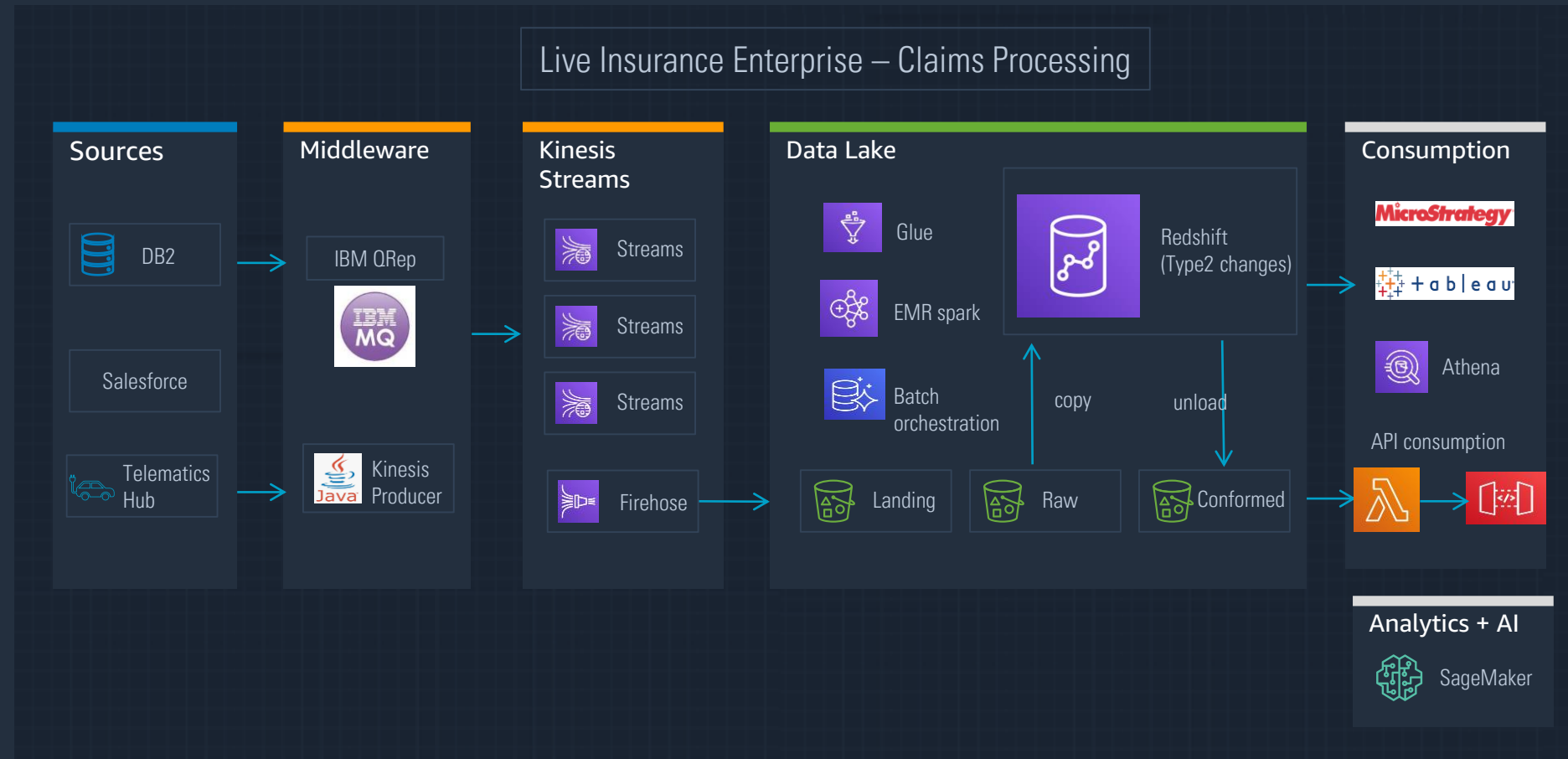
Real time analytics and alerting for Pharma manufacturing process

- During the plasma manufacturing process, the content gets moved and retained for 2-3 days for verification in the cleaning lab between the purification line and filling line.
- Maintaining pressure, temperature and other critical parameters is important for the health of the batch.
- Solution is to build an MQTT collection to bring and analyze measures in real time and produce functional marts and alert store
- These are delivered as dynamic alerts for lab personnel to act upon or as APIs for systems to trigger workflows should the thresholds for such action be met



Streaming real time for Analytics and API consumption

- Source data from Telematics hub, includes home and Auto – mobile, beacons, beaconless, home telematics
- Enterprise claim System on DB2 to produce transactional data
- IBM Queue replication process to read DB2 transaction data to message Queue
- Kinesis producer library reads data from Queue, aggregates and write it to Kinesis data stream
- Serverless processes transform and write it to API consumption for claims data analytics
- The firehose also writes data into S3 landing, which is further cleansed and transformed to raw and redshift for confirmation
- From Conformed layer, data is further processed for machine learning using data sage maker for predictive modeling



Thank you!

