



The above architecture is comprised of services that may be appropriate for the customers' requirements in the below topics

- Data Ingestion
- Data preparation
- Machine Learning/Analytics Components (Training and Inferencing)
- Data visualisation

### Data Ingestion

The architecture outlines the use of a combination of AWS Data streaming ([Amazon Kinesis](#) and [Amazon MSK](#)) and IoT ([AWS IoT Core](#)) services to address the need for realtime processing of data generated by multiple sensors and cameras located in the customers on premises environment. These services also use a combination of object storage ([Amazon S3](#)) and specialised databases ([Amazon TimeStream](#) and [Amazon DynamoDB](#)) to store the various types of streamed data.

### Data preparation ML/ Analytics

Once the raw data is captured, [AWS Step Functions](#) can be used to orchestrate the series of steps need to prepare data for machine learning training. The set of steps included could leverage [Amazon ECS](#) for data processing tasks that utilize third party open source software tools, [Amazon Sagemaker](#) for training and inferencing machine learning models, [Amazon EMR](#) for performing large scale analytics/ post processing data using open source big data frameworks and [AWS Lambda](#) to invoke the ML models that have been developed. The output of this will be stored on Amazon S3.

### Data Visualisation

In order to visualise data that has been processed or run queries on that data, the architecture highlights the use of [Amazon Athena](#) (backed by [AWS Glue](#) data catalog) to query this processed data and [Amazon Quicksight](#) as a business intelligence tool.