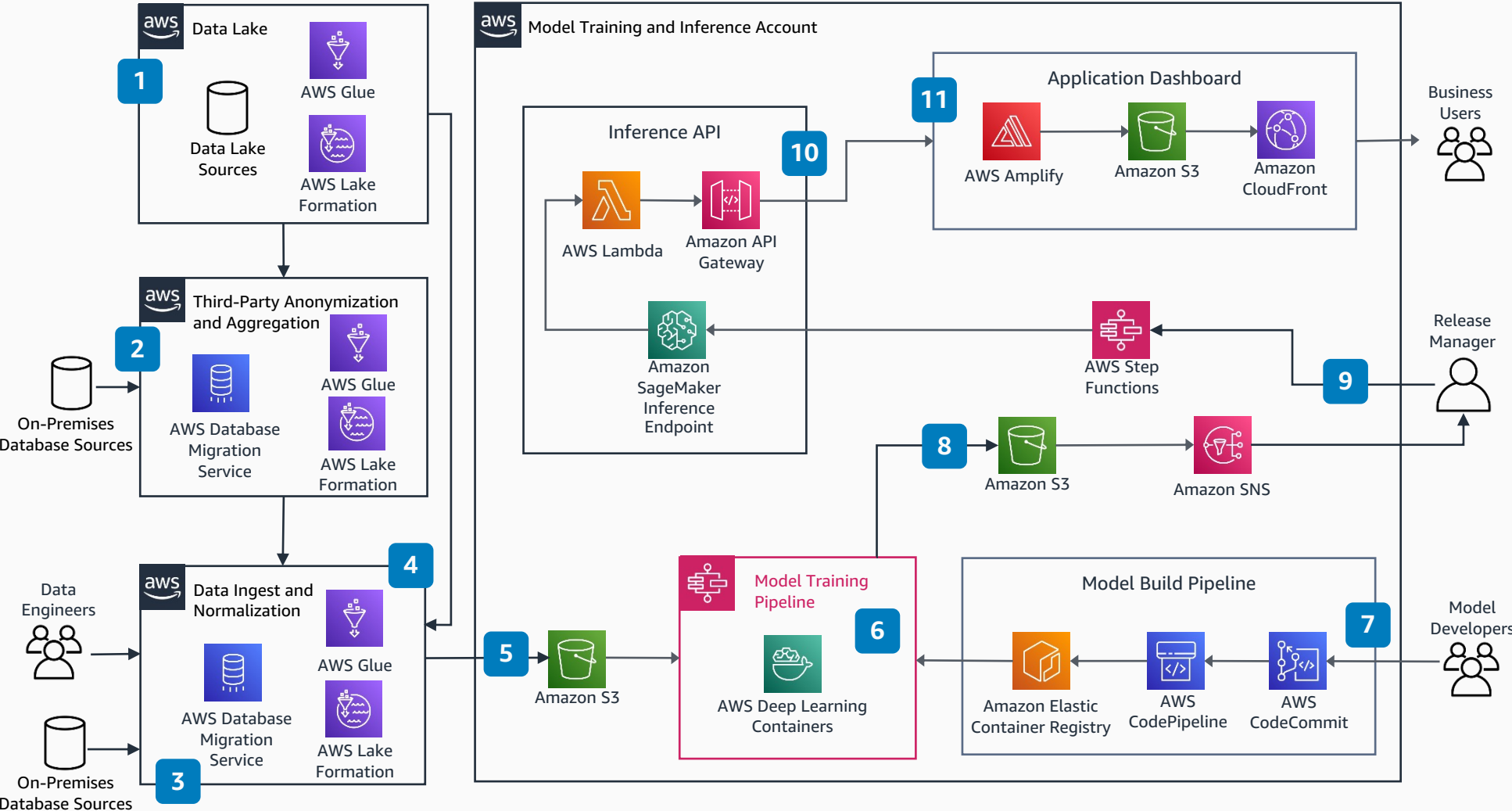


Predictive Modeling for Automotive Retail

Fine-grained Return on Investment (ROI) prediction for automotive sales incentives



- 1 A centralized data lake account can be used to accelerate development of new use cases.
- 2 Personally identifiable information is stripped. Aggregation obfuscates dealer specifics to prevent bias for or against OEM dealers.
- 3 **AWS Data Migration Service** replicates on-premises databases that aren't available through a data lake account.
- 4 Data is preprocessed by **AWS Glue PySpark Transforms**, output into master table for model training.
- 5 Training input master table pushed from data ingestion pipeline at regular intervals and stored in **Amazon Simple Storage Service (Amazon S3)**.
- 6 Input Master Table changes trigger model training pipeline—hyperparameter tuning, validation, and fit—controlled by **AWS Step Functions** utilizing **AWS Deep Learning Containers**.
- 7 Model source commits trigger container build and is stored in **Amazon Elastic Container Service**.
- 8 Versioned model outputs stored in **Amazon S3**, including training report and evaluation.
- 9 Administrator notified by **Amazon Simple Notification Service (Amazon SNS)** for review before triggering deployment with **AWS Step Functions**.
- 10 Inference requests served through **Amazon API Gateway**, **AWS Lambda**, and **Amazon SageMaker** endpoint using trained model container.
- 11 Inference dashboard built using **AWS Amplify**, static content hosted on **Amazon S3**, and served through **Amazon CloudFront**.



Reviewed for technical accuracy September 15, 2021
 © 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Reference Architecture