



# Architecture CLOUD AWS pour l'analyse des réservations aériennes

Équipe **Smart SIMPSONS** (Jessy Durca – Jérémy Peterson – Mérick Zounon)

# Introduction et objectifs

|                        |   |
|------------------------|---|
| <b>Objectif</b>        | Concevoir et mettre en place une architecture cloud (AWS) afin de traiter et d'analyser des données de réservations aériennes pour mieux comprendre les comportements d'achat et optimiser les stratégies de vente. |
| <b>Contexte</b>        | Utilisation d'un jeu de données Kaggle comme données de la compagnie aérienne, Analyse complète du cycle de réservation   |
| <b>Contexte métier</b> | Besoin de connaître les préférences clients (types de trajets, services additionnels, moment de réservation...) pour ajuster les offres.  |
| <b>Profils</b>         | Trois profils clés :<br>Data Analyst (requêtes & analyses avancées)<br>Marketing (KPIs & campagnes ciblées)<br>CIO/Direction (Tableaux de bord stratégiques)  |

# Analyse des données disponibles

## Données comportementales

- Nombre de passagers
- Origine géographique
- Canaux de ventes
- Délai achat/départ
- Durée du séjour

Basic

## Caractéristiques du vol

- Itinéraire
- Horaire
- Durée
- Jour de départ

Pro

## Services additionnels

- Bagages supplémentaires
- Sièges préférentiels
- Repas spécifiques

Premium



# Architecture AWS (Vue d'ensemble)



Collecte et  
stockage

S3 (données brutes)  
Kinesis Data Firehose  
(flux en temps réel)



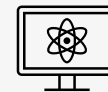
Nettoyage et  
organisation

AWS Glue (Crawler, ETL,  
Catalog)



Analyse

Amazon Athena (SQL)  
Amazon QuickSight  
(dashboards)



Machine Learning

Amazon SageMaker  
(modèles prédictifs)

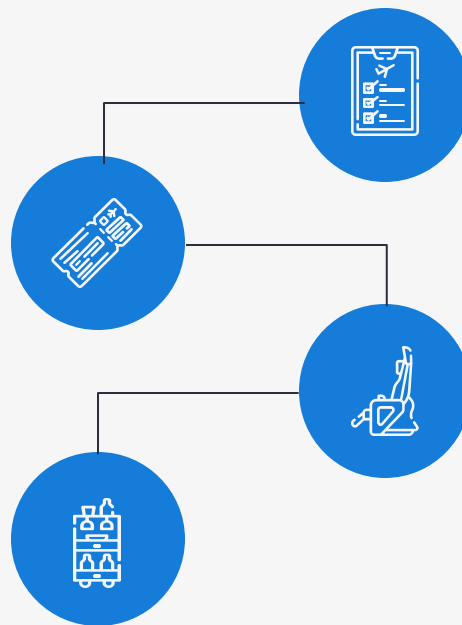
# Pipeline de données

## 2. Nettoyage et catalogue

- AWS Glue (ETL) pour uniformiser/formater les données
- Glue Catalog pour le schéma (Athena/QuickSight)

## 4. Machine learning

SageMaker : entraînement de modèles, prédiction (ex. Probabilité de booking\_complete)



## 1. Ingestion

- Données issues des clients (desktop, mobile) ou CSV Kaggle → S3 (brut)
- Kinesis Data Firehose : ingestion continue (transformations Lambda au besoin) → S3 (fichiers “propres”)

## 3. Analyse

- Amazon Athena : SQL sur les données stockées dans S3
- QuickSight : Tableaux de bord, rapports marketing

# Bénéfices pour chaque profil



## Data Analyst

Requêtes SQL via Athena,  
analyses ad hoc, exports CSV/ Excel  
Possibilité d'ETL avancé dans Glue



## Marketing

Dashboard QuickSight  
(taux de complétion,  
canaux de vente...)  
Segmentation, campagnes  
ciblées, suivi en temps réel



## CEO/ Direction

Vue d'ensemble  
stratégique  
Rapports de synthèse  
(mensuels, annuels), KPIs  
macro

# Conclusion et perspectives



## Conclusion

Architecture **scalable et unifiée**, répondant aux besoins **d'analyse et de reporting** pour la compagnie aérienne.

Différents rôles (Data Analyst, Marketing, CIO) mieux servis.

## Perspectives

Intégrer d'autres sources (API externes, logs de navigation, etc.).

Approfondir la gouvernance des données (sécurité, partitionnement, lifecycle S3). Mettre en place un pipeline MLOps plus complet (SageMaker Pipeline)

## Les avantages

Scalabilité

Automatisation

Optimisation des  
coûts

Accessibilité des  
données

Capacités  
avancées

S'adapte à de  
grands volumes  
de données

Réduction des  
interventions  
humaines

Utilisation de  
services  
serverless

QuickSight et  
Athena facilitent  
l'exploration des  
données

Intégration du  
Machine Learning  
avec SageMaker