



DataScientest • com

## DST Airlines

**Cursus concerné : Data Engineer**

**Difficulté : 8.5/10**

### Description détaillée :

De nos jours, il est capable d'avoir des informations sur les vols dans le monde entier et de traquer en temps réel un avion. Nous pouvons observer ce [site](#) en guise d'exemple. Le but ici est de s'y approcher le plus possible en passant par des API de différentes compagnies d'aviation.

Etape	Description	Objectif	Modules / Masterclass / Templates	Conditions de validation du projet
1	Récolte des données	<p>Passer par l'API de <a href="#">Lufthansa</a> pour récupérer des données sur les vols.</p> <p>Vous pouvez tester les différentes routes de l'API de Lufthansa à l'aide de ce <a href="#">lien</a>. Vous serez amené à aller puiser différentes informations comme les codes IATA (il faudra les récupérer via Webscraping).</p> <p>Cette étape est importante, vous devez comprendre les données que vous pouvez récupérer et faire un choix des routes à utiliser.</p> <p>Il y a aussi l'API de <a href="#">Internatinal Airlines</a>, mais il se peut que vous ayez des soucis pour y accéder..</p>	<p>Utilisation de la librairie requests ou de l'outil Postman (pour tester)</p> <p>Techniques de webscraping</p>	<p>Fichier explicatif du traitement et des différentes données accessible (doc / pdf)</p> <p>Un exemple de données collectées.</p>
2	Architecture des données	<p>Il y a plusieurs options qui s'offrent à nous. Lors de la précédente étape, nous avons observé qu'il y a plusieurs "types" de données. Nous allons qualifier des données de <b>fixe</b> comme les informations sur les aéroports et de <b>variables</b>, les informations sur les vols.</p> <p>Cette diversification de données va conduire à une utilisation de différentes bases de données.</p>	<p>142 - SQL</p> <p>Architecture des données</p> <p>Elasticsearch</p> <p>143 - MongoDB</p> <p>Neo4j</p>	<p>Une base de données relationnelle</p> <p>Diagramme UML</p> <p>Un fichier de requête SQL pour montrer que c'est bien fonctionnel</p> <p>Même rendu mais exemples de requêtes Elastic/Mongo</p>

[DataScientest.com](#)

Agrément organisme de formation 11755665975

+09 80 80 79 49

2 place de Barcelone, 75016 Paris



3	Consommation de la donnée	Implémenter un DashBoard à l'aide de Dash pour visualiser les données de votre base de données relationnelle. Ici, il faudra que vous sélectionnez les valeurs qui vous semble pertinentes comme un DashBoard Power BI.	Dash	Appli Dash
		Ici, il n'y a pas d'utilisation de Machine Learning sur ces données (Vous pouvez proposer des idées à votre mentor). Ce sont des données appropriées pour un outil de reporting, à la place il faudra réaliser une API pour requêter ces différentes bases de données.  En effet, avec la base de données sur Neo4j, nous pourrions représenter notre vol d'un aéroport à un autre. Faire un conteneur Docker de chaque composant du projet(BDD,API) et faire un docker-compose fonctionnel.	Plotly  FastAPI  Docker	API FastAPI  Conteneur, Fichier yaml du docker-compose
4	Automatisation des flux	Il faut récupérer les données de l'API Lufthansa selon un rythme bien défini pour l'envoyer aux différents consommateurs de la donnée.	Airflow	Fichier python du DAG
5	Soutenance	Démonstration de leur appli et explication du raisonnement effectué lors de leur projet.	X	Soutenance Rapport