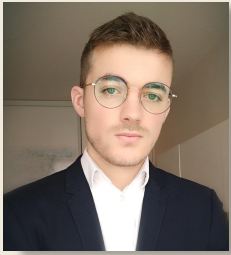


Big Data Engineer

**Julien
BONNET**



Diplômes.



Ingénieur Big Data & Machine Learning

EFREI Paris

2021



CPGE - PCSI

Lycée Marcelin Berthelot

2017

Certifications.

TOEIC

Score 850

Microsoft Azure

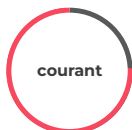
AZ-900

Databricks

Associate Developer for Apache Spark 3.0 / in scala

PT - Developer Essentials

Langues.



Anglais



Espagnol

Ingénieur en **Big Data** et **Machine Learning**, passionné par l'informatique et la data science, capable de **recupérer** et **traiter** des **volumes importants de données**, analyser à l'aide de **modèles prédictifs** puissants.

Expériences clés récentes.

Fnac Darty - Advanced Big Data Engineer

Concevoir et développer des solutions techniques d'intégration de données dans le Datalake et le Dataware RH Fnac-Darty.

Production de traitement sur Databricks tenant compte des performances et d'exploitabilité des scripts

Autonomie sur la production de flux métiers au niveau architecture et développement et TMA sur des flux en production.

Ladurée - Data Engineer

Mise en place de l'architecture Data chez Ladurée dans un contexte de carve-out très compliqué du groupe, en autonomie.

Création de différents pipelines de données avec Azure Data Factory, injectant différents flux de données (TCPOS, Adonix, D365) dans des environnements Azure SQL.

Création de reportings vitaux aux équipes métiers et à la nouvelle DSI (Reporting Finance, Comptabilité, RH, Paie etc...)

Compétences techniques.

Data Engineering

Spark 3.0, Scala, Python, Microsoft Azure, Databricks, SQL, Java, Hadoop, BeeHive, Nifi

Data Science

Python : Machine Learning / Deep Learning , C / C++

Azure Synapse Analytics

Maîtrise de Power BI et langage DAX

DevOps

Azure DevOps (boards, repos, pipelines, artifacts, test plans)

Web

HTML / CSS, Php, Angular, MongoDB

Outils

Databricks, IntelliJ, Anaconda, Power BI, Jupyter, Spyder, Virtual Studio Code, Excel, Gitlab & GitHub

Compétences fonctionnelles.

Gestion de projet

Gestion de projet agile : Azure Boards, Jira, Trello

Méthodologie

SCRUM, KANBAN

Soft Skills

Intelligence émotionnelle

Capacité d'adaptation

Secteurs d'activités

Banque / Assurances

Télécom / Energie.

Fnac Darty
Advanced Big Data Engineer**FNAC DARTY**

06/2022 - aujourd'hui

Contexte.**Objectif du projet Highway :**

Mise en place d'un data warehouse commun pour les parties RH BI de Fnac et Darty. Ce projet a de gros enjeux métiers car toute la BI RH Fnac-Darty repose sur notre data warehouse.

Responsabilité.

- **Concevoir et développer des solutions techniques d'intégration de données sur Azure (DevOps, Data Factory, SQL, Architecture Azure)**
Réalisation de pipeline de données dans Azure Data Factory afin d'alimenter un data warehouse Azure SQL. Amélioration du système de mise en production. Réalisation de POC concernant la mise en place d'une API REST depuis Azure Data Factory.
- **Conception, Développement et Mise en production de flux sortants et entrants**
Réalisation en autonomie de différents flux entrants et sortants du data warehouse via Databricks, ADF et Azure SQL. Réalisation de parsing XML nested dans un contexte métier.
- **Déploiement de résolution d'anomalies**
Maintenance du run des pipelines. Mise en place de test unitaire. Amélioration des processus de tests. Correction de scripts
- **Référent Techniques de la partie Développements**
Support et conseil auprès de la chef de projet. Force de proposition sur de nouveaux sujets. Lien entre l'équipe technique et l'équipe fonctionnelle

Livrables.

- **Pipeline de données sur ADF** contenant un ensemble de données métiers.
- **Traitement d'exécution Databricks** notebook liés aux pipelines qui produisent des traitements à grande échelle sur l'environnement RH du groupe Fnac Darty.

Environnement technique.

- **Stack technique :** Spark, Scala, Azure Data Factory, Databricks, Azure SQL serveur, Jira, Power BI.

Ladurée
Data Engineer**LADURÉE**
Paris

01/2022 - 05/2022

Contexte.**Objectif du projet ETOIL :**

Mise en place de l'architecture Data chez Ladurée afin de fournir un environnement de reporting aux opérationnels de différentes entités (RH, Paie, Comptabilité, Finance), dans un contexte de carve-out du groupe, en totale autonomie.

Responsabilité.

- **Création de pipelines de données**
Réalisation de pipeline de données dans Azure Data Factory afin d'alimenter un data warehouse Azure SQL.
- **Création d'environnements "data warehouse" cloud sur Azure**
Réalisation de plusieurs environnements Azure de test, preprod et production en tant que datawarehouse principal de Ladurée.
- **Développement, Déploiement et maintien opérationnel de reporting métier**
Analyse des besoins métiers, Production des reportings des reportings vitaux à certains opérationnels, Réalisation des évolutions des reportings.

Livrables.

- **Data Warehouses** contenant un ensemble de données métiers.
- **Reporting BI** très important aux opérationnels (Finance, Contrôle de Gestion, RH, Paie).

Environnement technique.

- **Stack technique :** Développement en DAX, SQL dans Power BI.
- **Stack Azure :** Environnement Azure Data Factory, Azure SQL, Power app

**AXA Direct Assurance
Big Data Engineer**

12/2021 - 01/2022

**Contexte.****Objectif du projet Cockpit :**

Migrer un data warehouse on-premise et des ETL dans Azure afin de faciliter aux équipes BI & Data Scientists d'AXA, leurs analyses prédictives. Ce projet est au cœur des enjeux car nous traitons toutes les données métiers que produit et ingère Direct Assurance.

Responsabilité.

- **Développement d'algorithmes d'extraction et transformation de données**
Réalisation et déploiement d'algorithmes en Spark / Scala d'extraction et transformation de données métier.
- **Développement de test unitaires**
Réalisation d'algorithmes de test unitaires des tables et transformations créés.
- **Etudes des pipelines de données dans Azure Data Factory**
Analyser, et résoudre des problèmes d'intégrations continues dans des pipelines sous Azure Data Factory.

Livrables.

- **Pull request** contenant code de transformation et test unitaires.
- **Présentation** des bugs découverts dans un ETL et résolutions associées.

Environnement technique.

- **Stack technique :** Développement en Spark / Scala dans un IDE : IntelliJ ainsi que sur Databricks, Méthodologie agile réalisée dans Azure DevOps (Boards & repository).
- **Déploiement continu** à l'aide de pipeline créés sur Azure Data Factory et gérées dans Azure DevOps ainsi qu'une surcouche on-prem SHINE.

**EDF
Data Scientist**

09/2020 - 09/2021

Contexte.**Objectif de D-Facto :**

Répondre à besoin internes de reporting autour du parc informatique d'EDF. Nous avons ainsi une équipe chargée de récupérer le besoin client, analyser sa faisabilité, récupérer de gros volumes de données nécessaires, développer des dashboards, et suivre avec le client ses évolutions possibles et leurs maintiens opérationnels.

Responsabilité.

- **Traitement de gros volumes de données**
Réalisation et injection de vue SQL dans la base postgres, création de pipeline de données avec power BI.
- **Développement d'algorithme**
Réalisation d'algorithmes d'automatisation de recettage d'une vue.
- **Réalisation de rapport**
Analyser, Créer et présenter un dashboard client et suivre son évolution et ses futures améliorations.

Livrables.

- **Rapport Power BI** sur un serveur interne, consultable depuis l'internet EDF
- **Notebook jupyter**, Fichier excel, Code et documentation etc...

Environnement technique.

- **Stack technique :** Utilisation de Power BI et jupyter (Python) comme outil de développement principal, Dbeaver pour parcourir les données principales EDF. Production de dashboard sur des environnements cloud Power BI