

Pierre Munhoz

Data Engineer - GCP, dbt & Cloud Data Warehousing Expert

6 années d'expérience en data

✉ pierre.munhoz@gmail.com | 📍 Paris, FR
🌐 LinkedIn | 📄 GitHub | 📁 Portfolio

PRÉSENTATION

Je suis un Data Engineer freelance avec plus de six ans d'expérience, spécialisé dans les plateformes cloud, les pipelines de données et les architectures évolutives. J'ai une expertise dans la conception de systèmes de données efficaces et sur mesure, tout en collaborant avec les équipes pour obtenir des résultats business significatifs.

EXPÉRIENCES PROFESSIONNELLES

Betclie

Paris, FR

Senior Data Engineer

Fév 2025 - Présent

- **Conception et déploiement de pipelines data** liés aux problématiques de Rétention des utilisateurs.
- Mise en place de **processus d'ingestion, de modélisation et de transformation de données** au sein d'un Data Lake basé sur Snowflake.
- **Automatisation et industrialisation du déploiement des pipelines data** à l'aide d'Airflow et de pratiques CI/CD.
- Rédaction technique et fonctionnelle de la **documentation** associée aux projets réalisés.

</> Stack technique : AWS, Airflow, dbt, Snowflake, Terraform, GitHub

Qantev

Paris, FR

Senior Data Engineer

Nov 2024 - Fév 2025

- **Développement et déploiement d'un système backend** pour appliquer les règles d'exclusion spécifiques aux assureurs, identifiant automatiquement les traitements non couverts et réduisant ainsi les efforts de validation manuelle des demandes.
- **Conception et mis en place des pipelines ETL** avec Dagster pour extraire, transformer et intégrer les données clients dans la base de données, améliorant ainsi la fiabilité des données et leur accessibilité pour les équipes d'analyse.
- **Optimisation d'un modèle de détection de fraude**, réduisant significativement les faux positifs et augmentant la précision de la détection, renforçant ainsi la confiance dans le système et minimisant les risques financiers pour les assureurs.
- **Refactorisation de la code base**, modernisant les composants critiques du traitement des données et améliorant la maintenabilité, l'évolutivité et la performance d'exécution.

</> Stack technique : Azure, Dagster, Django, Python, Terraform, GitHub

Lucca

Paris, FR

Lead Data Engineer

Avr 2024 - Nov 2024

- Création d'un **data warehouse** "from scratch" sur **Google Cloud Platform** (BigQuery, Cloud Storage, Secrets Manager, Compute Engine).
- Mise en place de **dbt** comme outil de transformation pour nos analystes.
- Implémentation de l'outil **Airbyte** sur un **Google Compute Engine** pour les **ETL** existants ainsi que la configuration de certaines extractions en low-code.
- Mise en place des configurations **Terraform** comme **IaC** pour la configuration d'Airbyte ainsi que pour Google Cloud Platform (IAM, datasets, service accounts...).
- **Orchestration de workflows** (Airbyte, dbt, scripts Python) avec **Dagster/Airflow et Docker**.
- **Management** de 2 Analytics Engineers (formation et accompagnement technique).

</> Stack technique : GCP, BigQuery, Python, Airbyte, Dagster, Docker, PySpark, Terraform, Cloud functions, dbt, GitHub Actions

Contentsquare

Paris, FR

Senior Data Engineer

Jan 2021 - Avr 2024

- **Gestion du projet de migration du data warehouse** pour migrer plus 300 tables en 1 mois et demi réduisant ainsi le temps de maintenance par 10 (Redshift, Snowflake dbt, Python).
- Rédaction d'un **article** pour expliquer comment construire un data warehouse from scratch.
- **Mentoring et formation** de juniors au métier de data analyst et d'analytics engineer (plus de 10 personnes formées).
- Développement d'**ETL** Python performants en surmontant des défis liés aux limites de taux d'API (réduction de 40 % le temps de récupération des données).
- Réalisation d'un **web scraper** Google/LinkedIn pour mettre à jour la base de données des utilisateurs internes, avec 80 000 titres de poste mis à jour et 20 000 utilisateurs désactivés (ML, Natural Language Processing).
- Owner du reporting des KPIs sales et du suivi la **performance des ventes des produits Contentsquare pour le reporting du comité exécutif** (automatisation d'un travail manuel qui prenait 1 semaine par quarter).
- Création de tests de qualité des données qui ont permis de réduire le temps de troubleshoot et d'identifier des problèmes parmi des millions de lignes.
- Création d'un algorithme de benchmark de données e-commerce qui fût ajouté aux **produits vendus** (ML).
- Plus de **30 analyses réalisées pour divers clients du Luxe & Retail** : Van Cleef & Arpels, Breitling, Castorama, Dior, Electrolux, Galeries Lafayette, Intermarché, Ma French Bank, Moncler, Orpi, Pôle Emploi, Prada, Sandro, Sephora...

</> *Stack technique* : AWS, Lambda, Redshift, Snowflake, dbt, Python, Dagster, GitHub Actions, ClickHouse

fifty-five

Paris, FR

Data Analyst

Jan 2020 - Jul 2020

- **Extraction et transformation de données** sur BigQuery pour la mise en place de suivi de KPIs et réalisation de dashboards sur Data Studio pour le groupe Shiseido (Shiseido, bareMinerals, Laura Mercier).
- **Mise en place d'un dashboard SAP** pour les marques du groupe Richemont (Cartier, Piaget, Van Cleef & Arpels. . .).

</> *Stack technique* : GCP, SAP, Python

Le Petit Paumé

Lyon, FR

Data Engineer

Jan 2019 - Déc 2019

- **Chef de projet de la migration du site internet** par un prestataire (80,000€ de budget).
- **Migration de l'ancien CRM vers Airtable** pour faciliter l'usage par des personnes non-techniques (Python).
- **Extraction et transformation de données Google Maps API** pour mettre à jour l'annuaire des établissements lyonnais (restaurant, bar, cinéma etc. . .) en Python.
- Mise à jour des données de latitude / longitude des établissements avec Open Cage API.
- **Création d'un algorithme de cartographie** afin de déterminer le quartier d'un établissement depuis son adresse.
- Récupération des données des stations de bus, métro et tram pour trouver l'arrêt le plus proche d'un établissement.
- Mise en place d'une **pipeline pour extraire des données du data warehouse** pour un de nos prestataires en vue du préparation du guide papier (200,000 guides distribués).

</> *Stack technique* : GCP, Python, JavaScript, API

Shanti Travel

Indonesia

Data Analyst

Jun 2018 - Déc 2018

- **Extraction et transformation de données CRM & Google Analytics** pour mesurer la performance des programmes de voyage proposés par l'entreprise.
- Récupération des données par API d'un partenaire pour mesurer la qualité des leads générés.

</> *Stack technique* : Python, Data Studio, JavaScript

EDUCATION

emlyon business school | Lyon, FR

Août 2017 - Déc 2020



Master in Management / Programme Grande École


- Mémoire : *Comment les algorithmes de recommandation influencent-ils l'industrie musicale*
- Cours principaux : Python (Data Science), Machine Learning, Intelligence Artificielle

COMPÉTENCES & INTÉRÊTS

 **Languages:** Français (natif), Anglais (courant), Espagnol (intermédiaire)

 **Core Technical Skills:**


- Python
- SQL
- dbt
- Airflow / Dagster
- Docker
-  Cloud Platforms:
 - AWS
 - GCP
 - Azure
-  Data Warehouses:
 - Snowflake
 - BigQuery
 - Redshift
- CI/CD
- Machine Learning / NLP

 **Additional Skills:**

- Linux Systems
- Infrastructure as Code
- Tableau / PowerBI
- Technical Leadership


♥ **Intérêts:** : cuisine, course à pied, YouTube, randonnée

PROJETS

Geography Quiz | [Website](#) 

2024 - Présent

- Création d'un **site internet** afin de deviner à quel pays est associé le drapeau affiché
- Récupération et transformation des drapeaux, pays et capitales afin de préparer d'autres modes de jeux de façon plus scalable

caviar-airways | [Website](#) 

2020

- Création d'un **site internet** intégrant une carte du monde (Mapbox API) afin de présenter des points cliquables redirigeant vers des playlists Spotify avec une thématique dédiée
- Réalisation d'une **interface administrateur** afin de rajouter, modifier ou supprimer des points sans avoir à modifier le code du site internet
- Récupération des images des playlists automatiquement grâce à l'utilisation de l'API de Spotify

RÉFÉRENCES

- Mannaïg Bénéat (Qonto): Product Analyst Manager
- Myriam Ayache (Contentsquare): Lead Data Analyst
- Corentin Flacher (Contentsquare): Senior Analytics Engineer
- Sébastien Pavot (Contentsquare): Senior Analytics Engineer
- Manuel Olmedo Ortega (Qantev): Lead Data Scientist