



# David CAPELLE

## DATA SCIENTIST / DATA ANALYST

Professionnel de l'informatique depuis 25 ans et de la Business Intelligence depuis 10 ans, je souhaite travailler en tant que data scientist / data analyst en indépendant, en partie en télétravail, en partie sur site.

### CONTACT



06 84 76 74 09



davidcapelle.professionnel@gmail.com



20 rue Cicéron 33600 Pessac - France



[Profil LinkedIn](#)



[Référentiel GitHub](#)

### SOFT SKILLS

- Rigueur, autonomie
- Esprit de synthèse, capacités d'analyse
- Capacités d'adaptation

### HARD SKILLS

- Analyse des données (EDA)
- Modélisation supervisée / non supervisée
- Feature engineering
- Architecture Big Data AWS
- Traitement chaînes ETL/ELT
- Conception / réal. univers, datawarehouses, datamarts

### OUTILS

- **Langages** : Python, Streamlit, Flask, Java
- **Base de données**: Oracle, DB2, MongoDB
- **ETL**: ODI, Talend
- **BI** : SAP BI 4, Webi, Qlik Sense
- **Big Data**: AWS (EC2, S3, Sagemaker)
- **Machine Learning**: Spark, PySpark
- **Outil recette**: HP-ALM

### CENTRES D'INTÉRÊT

- Jazz (fusion, nu jazz)
- Finance – trading
- Généalogie successorale

### EXPÉRIENCES PROFESSIONNELLES

#### Analyste BI / Data Analyst

Sopra Steria Group – du 02/2006 au 10/2020

- Avant-vente, préparation de POC, benchmarking
- Analyse des besoins, animation d'ateliers MOA
- Maintenance corrective / évolutive (chaînes ETL, univers SAP BI et restitutions Webi)
- Recette, animation comités fonctionnels.
- Assistance aux équipes de développement.

**Clients** : secteur public (Education Nationale, Défense, Ministère du développement durable).

### AUTRES EXPÉRIENCES

#### Formation Data Scientist (8 projets) - Titre RNCP niveau 7

OpenClassrooms – du 03/2021 au 02/2022

##### • Segmentation des clients d'un site e-commerce

EDA, modélisation non supervisée, réglage des hyper-paramètres, évaluation des performances du modèle afin de réaliser une segmentation des clients.

##### • Anticipation des besoins en consommation électrique

Analyse exploratoire (EDA) des données, modélisation (régression) supervisée pour prédire la consommation électrique et émissions CO2 de la ville de Seattle.

##### • Implémentation d'un modèle de scoring

Création d'un modèle de scoring (classification) avec métrique personnalisée pour accord / refus de prêts bancaires. Mise en place d'un dashboard déployé dans le cloud.



[Référentiel complet expériences professionnelles](#)



[Référentiel expériences data analyst data scientist](#)

### FORMATION

#### Titre professionnel RNCP niveau 7 (bac +5) Data Scientist

OpenClassrooms – du 03/2021 au 02/2022

**DRT Réseaux et systèmes répartis (bac +6) (1997)**

**IUP-MIAGE (1995)**

**DEUG Sciences-économiques (1993)**