

Youssef DIR

Graduated with a Master's degree in **Economics and Data science**(ranked 6th nationally) from the University of Tours, I am actively seeking a position as a data analyst-data scientist available Beginning sep 2025.

✉ | youssefdir37@gmail.com | ☎ 0666936752 | 📄 [Dirtoussef](#) | 🌐 Youssef Dir | 🌐 <https://youssefdir.netlify.app/> |

Education

2022-2024	University of Tours ,M2 Economics, Data science : Big Data,Data Mining,Maching learning ,Web scraping, Exploratory Data Analysis (Python , Rstudio),Microeconometrics,APIS,Pricing.SQL Data Management.	Tours, France
2018-2022	University of Tours , Bachelor's Degree in Applied Economics: Econometrics,Statistics,Probability,R,Python, Microeconomics ,Excel (VBA) ,HTML - CSS,Macroeconomics,Mathematics.	Tours, France

Experiences

AVANCI MVGROUP	Tours, France
Data scientist - Internship (6 month)	May 2024 - Oct 2024
Mission 1 : Creation and Management of a Simulated Database - SQL	
<ul style="list-style-type: none">• Generation of realistic data using the Faker library (Python) to simulate a dataset tailored for business analyses.• Implementation of data on SQL Server,managing relationships and structuring tables for efficient access.• Data cleaning: handling missing values,detecting inconsistencies, and preparing data for exploratory analysis.	
Mission 2 : Development of Dynamic Dashboards and Data Visualization with Streamlit - Python	
<ul style="list-style-type: none">• Integration of visualization tools such as Plotly and Matplotlib to present insights (segmentation, business performance).• Creation of advanced visualizations: interactive map to analyze regional audience, charts for customer segmentation,and performance tracking.• Implementation of a dedicated tab for data quality visualization.	
Mission 3 : Predictive Analysis and Advanced Modeling	
<ul style="list-style-type: none">• Statistical analysis and data cleaning (handling missing values, detecting anomalies).• Selection of key features (top features) by evaluating their relative importance in predictions.• Statistical analysis and data cleaning (handling missing values, detecting anomalies).• Redundancy detection with PCA between variables to optimize the selection of relevant features.• Implementation of a sampling strategy to balance the dataset in case of imbalanced classes.• Training of Random Forest and Boosting (XGBoost) models to predict target variables related to customer behaviors and characteristics.• Evaluation of model performance using key metrics such as accuracy and AUC ROC.	
PROBTB Finance	Paris, France
Data analyst Finance Internship (3 month)	June 2022 - Sep 2023
<ul style="list-style-type: none">• Development of a VBA script to automate the reading of PDF files, extract key information,and organize it into structured tables.• Development of a VBA-based tool to track SCPI (Real Estate Investment Company) values, including purchase values, Realization values,subscription prices, etc.	

Project

Academic Projects	Tours, France
Data Scraping and Car Price Prediction	2023 / 2024
<ul style="list-style-type: none">• Data extraction through web scraping on AutoScout (used car sales) using BeautifulSoup.• Cleaning,Data cleaning, processing,and analysis with Pandas and NumPy.• Modeling and prediction of prices using Machine Learning algorithms (Boosting,Random Forest,K-Means,SVM).• Documentation of the project in RMarkdown for a structured and clear presentation.	
RShiny Application for Text Analysis and Interactive Visualization	
<ul style="list-style-type: none">• Data cleaning,processing,and analysis.• Sentiment analysis of users by country.• Statistics and interactive visualization table of results.	

Skills

🌐	Languages	French (Bilingual), English (B2),Arabic(Bilingual)
📂	Programming	Python,Rstudio,Shell ,Dataiku
⚙️	Data science	Boosting (Xgboost),Random Forest,NLP,Linear Regression,Clustering (Sklearn – Tidymodels)
📊	Data Vis	Streamlit -Rshiny-Power BI
🗄️	Data Bases	SQL Server,My SQL,PostgreSQL,SAS
🔧	Devops	Gitlab,Github
☁️	Cloud Data	AWS,Snowflake,Scala
📄	Reportings	Quarto,Rmarkdown,Miktex