

ARIF MOHAMMAD

Data Analyst · BI Developer

arif15031995@gmail.com | +33 7 66 84 57 94 | Dijon, France | arifmohammad.netlify.app
linkedin.com/in/arif-mohd57776100 | github.com/web2arif

PROFESSIONAL SUMMARY

Data Analyst and BI Developer with an MSc in Data Science (Burgundy School of Business, France) and a hands-on track record building end-to-end data products — from web scraping and NLP pipelines to deployed Power BI dashboards and production mobile apps. Proficient in Python, SQL, Power BI, R, and React. Comfortable owning a project from raw data to a finished, shareable output.

Authorised to work in France (APS — no sponsorship required).

LANGUAGES

English (C2) · French (B1) · Hindi / Telugu (C2)

TECHNICAL SKILLS

Data & Analytics: Python (Pandas, NLTK, HuggingFace Transformers, Scikit-learn, TensorFlow), SQL, R (ggplot2, Shiny), DAX, Power Query, Power BI, Tableau, Matplotlib, ggplot2, Web Scraping, ETL Pipelines, NLP, Sentiment Analysis

Engineering & Tools: React.js, React Native / Expo, Node.js / Express, Google Cloud (STT, TTS, Translate), ElevenLabs API, Azure, Databricks, Railway, Netlify, GitHub

PORTFOLIO PROJECTS

Friends TV Show — End-to-End NLP & Analytics Pipeline

Python · VADER · HuggingFace DistilRoBERTa · Power BI · DAX

- Built a complete data pipeline from scratch: scraped and parsed HTML episode transcripts, cleaned 54,008 dialogue lines across 10 seasons into a structured 11-column dataset in Python (Season, Episode, Character, Location, WordCount, IsMainCharacter, and more).
- Applied dual NLP enrichment — VADER sentiment scoring (compound, pos, neg, neu) plus HuggingFace DistilRoBERTa emotion classification across 7 classes (joy, anger, sadness, fear, surprise, disgust, neutral) — producing 6 aggregated CSV outputs covering sentiment, emotion, character interactions, vocabulary richness, and monologues.
- Loaded all 6 tables into Power BI with a custom DAX measures table spanning 7 groups, a many-to-many relationship schema, and a 6-page dark-themed dashboard with a character-coded colour palette — fully self-directed, no templates.

Olympics Historical Dashboard

Power BI · DAX · Power Query

- Built an interactive Power BI dashboard over 250,000+ athlete records spanning 120+ years of Olympic Games, with country-level drill-downs, medal trend analysis, and DAX-calculated KPIs.
- Applied Power Query transformations to normalise multiple datasets with inconsistent historical schema into a clean, unified model.

Renewable Energy Transition Dashboard

Power BI · Trend Analysis

- Analysed 500,000+ energy records to visualise Europe's shift to renewables — tracking capacity growth, investment patterns, and country-level comparisons using Power BI slicers and trend charts.

Baseball Analytics Dashboard

R · R Shiny · ggplot2

- Built an interactive R Shiny web app visualising 50,000+ game records with advanced ggplot2 charts and dynamic filters — demonstrating the full R analytics stack from data wrangling to deployed dashboard.

Vinandi — Real-Time Multilingual Translation App

React Native · Expo · Node.js · Google Cloud · ElevenLabs · Railway

- Designed and deployed a mobile real-time speech translation app: Google Cloud Speech-to-Text (LINEAR16, 16 kHz) → Google Translate → Google Cloud TTS, with ElevenLabs as TTS fallback for Farsi where Google Cloud has no native support.
- Backend in Node.js / Express deployed on Railway; frontend in React Native / Expo confirmed working on iOS — covers full audio recording pipeline, base64 encoding, API orchestration, and TTS playback end-to-end.

Animated Portfolio Website

React.js · GSAP · SVG Animation · Netlify

- Built a fully animated personal portfolio using React.js, custom GSAP transitions, SVG animations, and a modular component architecture — deployed on Netlify.

ML Modelling & SQL Data Engineering

Python · R · TensorFlow · Scikit-learn · SQL

- Trained and evaluated classification and regression models (Random Forest, gradient boosting, deep learning) across academic and personal projects using Python and R.
- Built scalable SQL ETL pipelines handling 10M+ records with automated data quality and governance workflows.

WORK EXPERIENCE

Business Analyst (MSc Internship) · Magnus Infraprojects

Oct 2022 – Jun 2023 | Hyderabad, India

Internship submitted as part of MSc programme at Burgundy School of Business.

- Identified process inefficiencies through structured data analysis, contributing to a 15% improvement in operational efficiency.
- Produced data-backed research and market analysis reports to support strategic decisions at senior level.

Data Operations Analyst (Sr. Seller Support Associate) · Amazon

Oct 2018 – Jun 2021 | Hyderabad, India

- Audited seller account data across listings, inventory, pricing, FBA reimbursement claims, and financial reconciliation (payments, chargebacks, fees) — using Excel, pivot tables, and VLOOKUP analysis to surface errors and irregularities.
- Top Q1 performer for 3 consecutive years — average case resolution under 13 minutes across 40+ daily tickets.
- Developed strong process discipline and data-system familiarity working with Amazon's internal tools at scale.

EDUCATION

Programme Grande École — Grade de Master · Burgundy School of Business

Sep 2021 – May 2024 | Dijon, France

Dual-degree programme:

- **MSc in Management** (completed 2023)
- **MSc in Data Science & Organisational Behaviour** (completed 2024)

Core modules: Machine Learning, NLP, Data Engineering, Statistics, Organisational Behaviour.

Bachelor of Engineering — Electronics & Communication · BIT Mesra

Jan 2013 – May 2017 | Ranchi, India

IEEE publication: Selection of Dominant Voice Features for Accurate Detection of Parkinson's Disease — ICBSII 2017.

CERTIFICATIONS

- Celonis Process Mining Fundamentals
- Alteryx Designer Core Certification
- Microsoft Azure Fundamentals (AZ-900)
- Bloomberg Market Concepts