

Salvatore Barbagallo

Data Analyst | Healthcare & Life Sciences Data
London, United Kingdom
+44 (0)7475200048 | barbagallo1@hotmail.it
linkedin.com/in/salvatore-barbagallo | github.com/barbavegeta

Professional Summary

Data analyst with a background in clinical and genomics laboratories (7+ years) and hands-on analytics projects in Python, R and SQL. I turn messy operational and customer data into clean, documented datasets, repeatable reporting, and stakeholder-ready Tableau dashboards. Experienced with BigQuery and Amazon Web Services, with strengths in data cleaning, SQL modelling and data quality checks. Strong focus on data cleaning, SQL modelling, and building reliable BI reporting pipelines.

Skills

Data Analytics: Python (pandas, NumPy, scikit-learn, Matplotlib, Seaborn), SQL (BigQuery), R (tidyverse), Bash, Git/GitHub

Analytics & BI: Exploratory Data Analysis (EDA), Data Cleaning, Statistical Analysis, Feature Engineering, Dashboard Design (Tableau), KPI Definition

Data Engineering: ETL/ELT concepts, Star Schema basics, Data Quality Validation, Reporting Automation, Documentation

Domain Experience: WinPath, QPulse, EPIC, iPassport, Quality Checks, Healthcare & Clinical Laboratory Data, Genomics/NGS Data, Regulated Environments (GMP, QA, Audit Readiness)

Languages: Italian (Native), English (Fluent), Portuguese (Fluent), Spanish (Intermediate)

Projects

Google Fiber BI - Market Contact Analysis (*BigQuery, SQL, Tableau*) [*Google Cloud*]

- Consolidated regional contact datasets in BigQuery, standardised fields and built repeat-contact metrics
- Designed a Tableau dashboard for cross-market monitoring, enabling drill-down into contact drivers and service issues

Bellabeat Fitbit Data Analysis - Usage & Wellness Insights (*R, SQL, Tableau*)

- Built user-level summaries from multi-source activity datasets using R and SQL
- Developed Tableau dashboards highlighting behavioural trends to support product and marketing insights

Salifort Motors - HR Analytics & Employee Turnover (*Python, scikit-learn, XGBoost*)

- Cleaned and modelled HR datasets; engineered features and compared classification models (logistic regression, random forest, XGBoost)
- Translated model findings into business-facing recommendations on employee retention drivers

TikTok Engagement Analytics - Content Performance (*Python, pandas, Plotly*)

- Analysed relationships between content attributes and engagement metrics using Python
- Delivered visual summaries and data-driven recommendations on posting strategy

AWS Three-Tier Solution Architecture

Designed high-level cloud architecture for migrating web and analytics workloads to AWS managed services using AWS services (S3, CloudFront, ECS Fargate, Aurora MySQL, ElastiCache Redis, SQS), focusing on high availability, scalability, security and cost efficiency.

Certificates

IBM: Data Engineering Professional Certificate (in progress)

Google: Data Analytics; Advanced Data Analytics; IT Automation with Python; Project Management; Business Intelligence

Google Cloud: Architecting with Google Kubernetes Engine

SAS: SAS Programming 1: Essentials; SAS Programming 2: Data Manipulation Techniques

Amazon Web Services (AWS): Cloud Practitioner Essentials; Cloud Solutions Architect

Education

Master of Science: Bioinformatics

Atlantic Technological University (Remote), September 2025 - Present (ongoing)

Master of Science: Cell and Gene Therapy

University College London, September 2021- September 2023

Bachelor: Biomedical Science

Siciliae Studium Generale - University of Catania, September 2014 - October 2017

Professional Experience

Specialist Biomedical Scientist

University College London Hospitals, Stem Cell Laboratory, September 2021 - October 2025

- Automated regulated sample and workflow tracking in Python/Excel; reduced manual reporting and reconciliation by ~10 hours/week.
- Maintained Good Manufacturing Practice (GMP) traceability for 200+ samples/week; supported audits through documentation and data validation.
- Supported 11 clinical and Advanced Therapy Medicinal Product (ATMP) trials; digitised Standard Operating Procedures (SOPs) to improve Quality Assurance (QA) readiness.

Laboratory Scientist

CooperGenomics, London, July 2019 - Sept 2021

- Processed 500+ next-generation sequencing (NGS) samples/week; automated quality-control validation steps to reduce errors and turnaround time.
- Validated and programmed SPT Labtech Mosquito and Dragonfly liquid-handling robots for scalable library preparation; trained colleagues on workflows.
- Produced ISO-compliant sequencing reports with consistent templates and documented checks.