

github.com/mionisation

Experience **Software Engineer (Contract), Direct Line Group** - London, UK – Oct 2020–Jun 2021

- **Work on data pipelines, processing financial information and customer data:** Ingestion and processing of legacy database deltas into AWS Glue with pySpark jobs
- **Stripe integration:** Bulk ingestion of customer credit card data into Stripe API
- **Microservices:** REST endpoints to ingest and process insurance data: API Gateway + Lambda
- **Automated batch jobs:** Conversion of financial data between JSON, XML and “BACS” fixed-width data. Workstreams included Premiums, billings, direct debits, claims, incidents, financial reports
- **Utilised:** Python, pyspark, BDD/TDD, bash, AWS Services: Lambda, Glue, S3, Athena, Dynamo, CloudFormation, CodePipeline, API Gateway

Data Engineer (Scala/Python), ClearScore - London, UK – Feb 2020–Oct 2020

- **Credit Report Events:** Porting data processing logic from AWS Lambda into microservices (Scala)
- **Zendesk metadata ingestion:** End-to-end implementation, testing and deployment of an app that consumes metadata from a Zendesk API and ingests it into a data lake
- **Enhancing “Nested Tracking” gateway:** Enhancing nested data processed in a microservice
- **End-to-end Integration Tests:** Ensuring data sent to API is ingested correctly, BDD with Pytest
- **Airflow workflows:** Adding new spark jobs using airflow operators to extract credit offer data
- **Utilised:** Scala, Python, SQL, Terraform, Spark, Jenkins, AWS: S3, Lambda, Kinesis, Redshift

Software Engineer (Data), Hive - London, UK – Jan 2019–Jan 2020

- **CRM Project:** Delivering real-time data pipelines of IoT devices and their usage for CRM / Marketing related purposes; leveraging Kafka Streams and Scala
- **Data Lake:** Ingestion of IoT hub data with Kinesis. ETL and Curation according to business requirements using Scala + Spark/Glue from IoT Hubs into S3 Data Lake. Instantiating needed infrastructure using Terraform modules.
- Automation of batch feeds of usage data on hivehome.com and mobile apps from Adobe Omniture using S3-event triggered Lambdas written in Python, sending monitoring data to Cloudfront.
- Writing CI/CD pipelines for Jenkins, scaling/monitoring K8s instances with kubectl.
- **Utilised:** Scala, Python, SQL, Kafka Streams, Spark, SBT, Jenkins, Kubernetes, Terraform, AWS: S3, Lambda, Kinesis, Athena, Redshift

Data Engineer in Supply Chain, Intern; Amazon, Luxembourg; May 2017–Nov 2017

- **Transfer Tool:** Designed algorithm to pick transfers maximising inventory selection. It generated +10 million inventory transfers during the launch of new warehouses in the UK + EU. Python + Redshift
- **VIP Dashboard:** Tracked performance of inventory optimisation model with PDF export. R Shiny
- **Root cause dashboard:** Visual analytics tool to investigate root causes for badly placed inventory
- Automation of internal site usage with Java/Selenium, SQL code generators for common queries
- Reporting / SQL queries for quantitative analysis of supply chain systems performance; close collaboration with operations research scientists, business analysts and BI teams
- **Utilised:** Python (Plotly, Dash, Pandas), Redshift, Oracle, Bash, R (+Shiny), HTML/CSS, Java + AWS SDK, PostgreSQL, Selenium, Excel

Education Technical University of Vienna; Vienna, Austria – Diploma study (MSc) in Business Informatics, 2019
Technical University of Vienna; Vienna, Austria – BSc in Computer Science, 2016

- **Master Thesis:** Designing and Evaluating a Recommender System for Board Games
- Published Research Paper at RecSys Conference: <https://dl.acm.org/doi/10.1145/3341105.3375780>
- *Coursework:* Adv. Software Engineering, Software Testing, Business Intelligence, Econometrics,...