Contact

Phone: 087 1211 867
Location: Ireland
Email:
oliverjnorton@gmail.com
Websites:
- Portfolio
- GitHub
- LinkedIn

Skills

Programming
- Python
- SQL
- R

Data Visualisation
- Power BI
- Tableau
- Excel
- matplotlib
- Plotly

Data Science Libraries
- PyTorch
- scikit-learn
- BERT
- XGBoost
- LightGBM
- Pandas
- NumPy
- Prophet
- BeautifulSoup

Data Engineering, Data Warehousing & Cloud Services
- Azure
- Docker
- Apache Kafka
- PostgreSQL
- DuckDB
- SQLite
- DBT
- Ucloud

Oliver Norton

With a master's in Social Data Science, experience as a Data Scientist, and a multitude of projects, I have the skills needed to excel in your team. I've honed my ability to communicate with non-technical stakeholders, and work as part of a team to deliver valuable business impact. For examples of my expertise, see my GitHub and website.

Experience

Feb 2023 - Jan 2024, Junior Data Scientist, *Meltwater* (Paris, France)
- Developed predictive time series models with Python Prophet to forecast demand for food ingredients in a €150,000 client project
- Performed in-depth analysis, created data visualisation dashboards and automated reporting, saving over 450 billable hours for colleagues
- Communicated clearly between cross-functional teams, and coordinated with non-technical stakeholders to provide data-driven insights
- Engineered Python scripts for ETL processes, leveraged data platform API framework, and managed Linux virtual machines (VMs)

Sep 2021 - Jun 2022, Administrative Assistant, *Copenhagen University*
- Maintained a research centre's website using HTML and a CMS, ensuring quality of content and structure
- Related with both technical staff scientists and the non-technical public, supporting scientific communication activities

Education

2021-2024, MSc in Social Data Science, *University of Copenhagen*
- Grade Average: 9.5/12 with grade 12/12 in master's thesis

2022, Academic Exchange, *Sorbonne University*
- Audited data science courses including: Probabilistic and Statistical Models and Algorithms for IT, Complexity, Databases, Networking

2017-2021, BA in Business, Economic & Social Studies, *Trinity College Dublin*
- Grade Average: 70% (First Class Honours)

2019-2020, Academic Exchange, *UNSW Sydney*

Projects

Fraud detection ML model with a large dataset using LightGBM
Dec 2024 | **Data Science** | View project on: Website, GitHub
- Developed a fraud detection model using Python package LightGBM, achieving ROC AUC of 0.932, and deployed using Kafka
- Conducted feature engineering, data wrangling, Exploratory Data Analysis (EDA), addressing class imbalance and data quality issues
- Applied gradient boosting and encoding to optimise model performance

Data pipeline and automation applied to energy data
Jul 2024 | **Data Engineering** | View project on: Website
- Designed and implemented robust data pipelines with ETL -like data process, using DBT, Docker, Python, Bash and Git
- Manipulated raw data and created models using SQL
- Implemented rigorous testing and quality checks to ensure data integrity
- Visualised data with Tableau and PowerBI dashboards to derive insights

Additional Skills

Development Tools
- Git
- JavaScript
- Confluence
- Jira
- Bash
- Linux

Other Tools
- Jupyter
- Google Data Studio

Languages
- English (Native)
- French (Fluent)
- Spanish (Inter.)
- Danish (Basic)

Projects (Continued)

Azure-based cloud architecture design for a carpooling company
Jun 2024 | **Cloud Computing** | View project on: Website
- In 72 hours, designed a data software architecture for client interested in migrating from an on-premise solution to the Cloud
- Security focused design including Identity and Access Management (IAM), encryption of data at rest/in transit and the Zero-Trust model
- Conducted a thorough risk-benefit analysis of cloud architecture, resulting in a design that enhanced data protection and compliance

Quantifying polarisation using networks and text data (Master's thesis)
Sep 2023 - Jan 2024 | **Network Science** | View on: Website

- Developed a novel multi-dimensional approach to quantify political polarisation, adhering to data governance standards
- Created a machine learning (ML) model (cross-validated Lasso regression)
- Utlised a cloud environment to access virtual machines (VM), execute code and run models

Investigating polarisation in Reddit Communities
Apr - May 2023 | **Data Science**
- Collected and cleaned over 50 GB of semi-structured Reddit data using a REST API (& Web scraping) for in-depth analysis
- Trained a multi-label sentiment classification model to categorise users

Analysing climate discourse with Natural Language Processing (NLP)
May - Jun 2022 | **Data Science** | Visualisation on: Website
- Analysed Twitter API text data, focusing on sentiment and discourse
- Developed ML classification models and used LDA topic modelling to identify key themes and trends in the conversation
- Visualised a retweet network employing community detection algorithms

Professional Development & Certifications

- **AZ-900: Microsoft Azure Fundamentals Certification**
  (In Progress)

- **Intermediate JavaScript, *Codecademy***
  (In Progress)

- **Introduction to JavaScript, *Codecademy***
  (Complete)

- **Google Analytics Individual Qualification, *Codecademy***
  (Complete)

- **Marketing Analytics Nanodegree, *Udacity***
  (Complete)

Personal Achievements

- Cycled over 4000 kilometres from Morocco to England at age 19
- Mastered French as a native English speaker
- Lived in 7 countries and travelled to over 35
- Climbed the highest mountains in 5 countries: (Morocco 3400m, Zimbabwe 2592m, UK 1345m, and Ireland 1039m)