

Harry Baines

Data Scientist

Postgraduate with over 4.5 years experience as a Data Scientist. Proficient Python programmer with experience working on varied datasets using relational and non-relational databases. Enthusiastic about implementing solutions to complex problems and finding insights in large datasets using machine learning and statistical techniques.

WORK EXPERIENCE

Data Scientist

October 2020 – June 2022

Yordas Digital, Lancaster

- Developed and deployed time series forecasting models using machine learning to predict the probability of regulatory data updating in the next 3 months, reduced weekly data checking by 90%.
- Conducted a statistical analysis of cheese production data for a local manufacturer, provided data-driven insights and recommendations to optimize cheese yield.
- Created automation scripts to scrape data from regulatory websites, saved 20+ hours per week of manual data checks.
- Supervised 2 student internship projects: general data scraping and NER for detecting substance names in chemical literature, provided daily meetings and weekly code reviews using the scrum methodology.
- Developed a full-stack application with Django to assist regulatory experts in authoring Safety Data Sheets (SDS), reduced authoring time by over 50%.

Data Scientist Internship

June 2020 – September 2020

Yordas Digital, Lancaster

- Implemented a search engine using NLP techniques to provide smart search functionality across a database of 20,000+ complex chemical phrases, integrated into a live SDS authoring application.
- Collaborated with data scientists, software engineers and regulatory experts in an agile fashion to identify project requirements and receive iterative feedback.

Full-Stack Data Scientist (Part-Time)

June 2018 – February 2023

Cassidy and Ashton Architects, Preston

- Developed a timesheet application using the PERN stack to facilitate entry and management of time record and employee data for 60+ employees, cut down manual timesheet paperwork by 100%, and saved 20+ hours per week of manual data entry.
- Created interactive dashboards for project managers to visualize and monitor employee performance against agreed fees across 12,000+ jobs.
- Presented data visualizations and pivot tables of employee timesheet data to directors and project managers.

CONTACT

 07949 443583

 harryb0905@gmail.com

 harrybaines.net

 [linkedin.com/in/harryb0905](https://www.linkedin.com/in/harryb0905)

 github.com/harrybaines

SKILLS

Machine Learning

Statistical Modelling

Data Visualization

Natural Language Processing

Productionizing Models

Python (numpy, pandas, seaborn,

sklearn, Selenium, Flask, Celery)

R (tidyverse)

SQL (Postgres, MySQL, Redis)

Docker

Git/GitHub

Command Line

EDUCATION

Lancaster University

MSc (Hons) Data Science (Distinction)

October 2019 - September 2020

Thesis: Enhancing Phrase Retrieval
for Safety Data Sheet Authoring

Lancaster University

BSc (Hons) Computer Science (First)

October 2016 - June 2019

AWARDS

- Mathematics for Machine Learning Specialization, Coursera (2022)
- Machine Learning Scientist Track, DataCamp (2022)
- Data Science master's scholarship (2019)
- First Class degree scholarship (2019)
- Deep Learning Specialization, Coursera (2019)
- The Complete Web Developer in 2018: Zero to Mastery (2018)
- Undergraduate commendation (2017)
- A-level Computing award (2016)
- Bronze and Gold Duke of Edinburgh awards (2013, 2016)