Ian Huston

Data Science Leader bringing data thinking to software development teams, and agile methodology to data scientists.

ianhuston.net, github.com/ihuston

Project Highlights

Top 10 Car Maker Developed a new data driven supply chain ordering process for a global top 10 automobile manufacturer. This system is now in production across European markets and realising savings of millions of Euro per year. I led the 4 strong data science group through the 6 month project, which included guiding client data scientists in Java & Python, Test Driven Development, Pair Programming, and coordination as part of a User Centred product team. *Python, Hadoop, Java, Cloud Foundry*

Global High Street Bank Designed a new personalised messaging experience for retail customers. The tailored communications app for retail users incorporated insights and predictions from individuals' live transaction data. Driving the iOS application is a pipeline of data processing and transforming micro-services, including overdraft entry prediction, anomaly detection and simple thresholding. I led the data science efforts of our combined team of data engineers, software engineers, and product managers for 6 months. *Spring Cloud Data Flow (Java), Python, Cloud Foundry, AWS*

Gaming group Built data quality pipelines and entity resolution models to combine data sources from multiple online gaming & gambling brands into a single view of the group's customers. *PostgreSQL UDFs, R*

Leading Global Reinsurer Introduced new algorithmic techniques to reduce simulation time in a global reinsurer's catastrophe modelling from weeks to minutes. Built a parallel computation system which enabled growth of simulation size from small national regions to continent sized data sets. *R*

Premium Car Brand Analysed live car telemetry data to create a premium service appointment booking system for the UK dealer franchise of a leading European car manufacturer. *Python, Jupyter notebooks, Java, Cloud Foundry*

Further projects include data visibility in Oil & Gas sector, IoT asset management, and airport crowd visualisation.

Facilitation

I have delivered workshops on many data science and product related topics including:

- Data discovery and exploration workshops for product teams & data scientists,
- Idea generation and prioritisation workshops for CxOs & LOB leaders,
- Agile development & Lean product workshops for product teams and as part of university courses,
- Conference tutorial track on using Cloud Foundry with the Python PyData stack for practicing data scientists.

Employment

2017 - 2018 Associate Director, Pivotal Dublin

I lead a team of 15 consultants and am responsible for:

- representing Pivotal in customer presales interactions and workshops,
- working directly with Pivotal account teams in Europe and the US,
- helping our account teams design and book appropriate client projects,
- ensuring project success with delivery teams,
- acting as the liaison for client stakeholders, and escalation path,
- hiring, retention and management of our team of engineers, product managers, product designers and data scientists,
- coordinating with regional leadership on sales strategy and focus.

2016 - 2017 Principal Data Scientist, Pivotal Dublin

Lead consultant on teams using data science with Python (Scikit-learn, Matplotlib, Jupyter, Pandas, NumPy, …), Apache Spark, Apache Hawq, and PostgreSQL/Greenplum. Deploying cloud based data science micro-services using Spring Cloud Data Flow and on Cloud Foundry, AWS, and GCP.

2013 - 2016 Senior Data Scientist, Pivotal London

Consulting data scientist on multi-disciplinary teams working on data science specific and broader app development engagements using Python, R, Java, Hadoop, Apache Pig, Tableau, Greenplum, and more.

2010 - 2012 Postdoctoral Researcher, Queen Mary, University of London

Research focus on building numerical computational systems for early universe cosmology. Additional responsibilities included tutoring undergraduate mathematics courses and supervising Masters students.

Education

2006-2010 **PhD, Theoretical Cosmology**, Queen Mary, University of London

Large scale numerical simulations of physical conditions in the early universe during the inflationary phase. Created and maintained an open source Python package for numerical second order perturbation theory on large scale compute clusters which is still used in the field.

2005-2006 **MSc (with distinction), Theoretical Physics**, Imperial College London

2001-2005 **BSc (Hons), Mathematical Physics**, University College Dublin

Technical Experience

Programming Languages	**Python:** Over 10 years experience with the scientific Python stack. **Java:** Extensive use of Java in client projects over 5 years including Spring. Working knowledge of **R**, **Fortran**, **MATLAB**, **Kotlin**, **C**
Other Tech	Concourse (CI/CD), Tableau, Greenplum (MPP Database), Hadoop, Cloud Foundry