



ANNA CASSIDY

Big Data Engineer

ABOUT ME

My interest in data was sparked by my final year project and continued outside of academia. Something that particularly sparked my interest and fascination was the rise of boohoo.com, an e-retailer who made incredible use of information. Using the tool 'Edited' the company was able to access information regarding prices set by competitors for every product on the market, and adjust their own accordingly. Boohoo.com raised their pre-tax profits by 97% in 2016 alone, attributing this success to data driven decisions.

Something that is very important to me is balancing my work and life effectively as I have passions in both. I enjoyed my university course and the responsibilities of being part of an academic society. In particular my role in the Physics Society has allowed me to give lecture-style talks about the department to all new undergraduates, strongly developing my public speaking and communication skills.

Outside of work I am part of a team who have built an educational physics website providing articles and resources written by physicists in layman terms. My role is to promote the website and articles on social media, determining which times of day are best to be active across the many platforms in order to attract the largest volume of people.

CONTACT US

Kubrick Group

T: 020 3866 4620

E: consultants@kubrickgroup.com

W: kubrickgroup.com

OVERVIEW

I graduated from Queen Mary University of London in Physics in 2017. In my final year I focused on a project based on the analysis of the Voyager 1 data from NASA. This was used to determine the reliability of the current theoretical understanding of the heliosphere by comparing the results of the plotted NASA data with the expected trend at the edge of the sphere. I concluded that the theoretical understanding was not fully complete and that it did not match the data accurately, suggesting further study of turbulent solar wind as the currently accepted model was very simplified. This resulted in a high 2:1 grade for the report.

One of the modules I enjoyed the most was Statistical Data Analysis. It introduced me to finding the solutions to problems buried within data and touched on the variety of methods used to extract and display them. Examples are visualising and quantifying data, hypothesis testing, and multivariate analysis techniques. It was the driving force behind my dissertation focus and inspired me to look into careers which would allow me to explore further.

WORK

DATA ENGINEER - KUBRICK GROUP - LONDON

September 2017 - Current

At Kubrick Group I gained knowledge and real world experience in professional skills, a wide yet in depth array of modern technical and analytic skills. We studied the agile software development lifecycles and gained detailed knowledge of platforms and infrastructure.

I specialised in Data Engineering which involved learning and developing advanced skills in Python, Spark, Hadoop, NoSQL and SQL as well as advanced Excel - these lead me to work towards accreditation as a Cloudera Hadoop CCA Developer.

As a junior data consultant I applied my skills on real client projects including:

- Seasonality correlation to sales for a large retail client.
- Feature analysis for a large financial services client providing HR with churn rate modelling against different employee profiles.

The projects were developed in agile teams using Git and Kanban boards following agile processes in the presence of an agile coach. Delivery of the projects used virtual environments and Docker. The professional skills I have acquired and matured were invaluable in these projects.

PHYSICS SOCIETY - TREASURER

2016 - 2017

The role involved the organisation of psi star events, including a European trip and the physics ball, which helped me learn how to effectively manage my time alongside my studies. In addition to this I was part of a team organising and promoting evening lectures with academics, developing the ability to communicate well with senior members of staff.

BUSINESS CLERK, INTERNAL ENERGY UK LIMITED, PONTELAND

2011 - 2017

Acting as a point of contact for the business owner, this role involved the maintenance of cashbooks, entering of receipts, payments and the filing of documents. I took business calls and recorded call details which taught me how to adopt a professional phone demeanour, and completed Internet based market research and compilation of marketing database with client details through LinkedIn.

MATERIALS TESTING AT MARSDEN QUARRY, WHITBURN

July 2014

At Marsden Quarry, I was working in the lab collecting data from different materials by performing stress tests and conducting environment testing to ensure pollution levels were minimal. This allowed me to advance my precision in laboratory work and have practice in displaying data graphically on excel and presenting findings to supervisors by explain my methods and results.

EDUCATION

QUEEN MARY UNIVERSITY OF LONDON, PHYSICS, UPPER SECOND CLASS HONOURS BSC

September 2014 - July 2017

My desire for exposure to as many aspects of physics as possible, took me on a path covering a range of subjects from electromagnetic waves & optics, to thermal & kinetic physics, including the use of Matlab and Excel to input & process findings.

Physics of Galaxies: 80.7%
 Radiation Detectors: 70.8%
 Physical Cosmology: 70.1%
 Statistical Data Analysis: 67.0%
 Space-time & Gravity: 69%

KEY SKILLS

Python Libraries:

Pandas, Numpy, SK Learn, NLTK, PyMongo and Plotly

Cloudera Hadoop version 5.4:

HDFS, Yarn, Sqoop, Flume, Avro, Parquet, Impala, Hive and Hue

NoSQL:

MongoDB, Neo4J

SQL:

SSIS, Lavastorm, Advanced SQL

Apache Spark:

RDD, DataFrames, SparkSQL, PySpark with some exposure to and understanding of Scala

Source Control and IDE:

Git, SourceTree and Git Extensions, PyCharm

Other Skills:

Datalku, Linux CentOS 6, SSH, Putty, Agile