RACHAEL COLE
Big Data Engineer

ABOUT ME

I am fascinated with how Big Data is becoming increasingly used and necessary within businesses. I experienced it first hand when working at Walt Disney World where Big Data is becoming more and more involved in guest experience and all guest data is stored within a wristband which replaces any form of paper at the resort.
I was so amazed by the technology that it sparked my interest in pursuing a career in Big Data even more.

During my time at university, I worked as a maths tutor and I also worked in various retail environments including the Wimbledon Tennis Championships, through which I developed strong communication and team working skills.

Outside of work, I love cheerleading, having done it competitively at university. In March 2016, my team became BCA National University Champions with a routine that I had helped to choreograph.

CONTACT US

Kubrick Group
T: 020 3866 4620
E: consultants@kubrickgroup.com
W: kubrickgroup.com

OVERVIEW

I recently graduated from the University of York with a Maths and Physics degree. My BSc project titled 'Investigating Stellar Nucleosynthesis with an Astrophysical Reaction Network', consisted of using NUGRID PPN simulation codes to collect data on the abundances of isotopes in a particular stellar environment.
Using Python to visualise the data, I found patterns in the data to make conclusions about these isotopes and the reaction network involved. I loved analysing the data I collected and finding patterns and believe my degree and Kubrick give me the perfect toolset to succeed in a career in Big Data.

WORK

BIG DATA ENGINEER – KUBRICK GROUP, LONDON
At Kubrick Group I gained experience and qualified in: Professional and Communication Skills, IOC (Investment Operations Certificate) Level 1, Advanced Excel, Web Application development (UI/UX, JavaScript, CSS and HTML), Agile and SQL.
I then specialised in the Big Data Engineering course leading to accreditation as a Cloudera CCA Developer.

October 2016 - present
I worked on Big Data projects which exercised my skills in data ingestion and processing using Hadoop EcoSystem tools such as Sqoop, Flume and OpenSource Python Libraries including Pandas and Twython. We cleansed and munged data after exploratory data analysis through Python libraries including Plotly and Pandas then moving onto Spark to process, transform, analyse and stage the data into HDFS.

Projects included:
MetOffice Weather Data: involving Python API and Twitter processing into HDFS followed by further processing in the Hadoop Ecosystem.
Police Crime Data Analysis involving SSIS processing, SQL Analysis, and presentation through custom web front end.

VARIOUS RETAIL ROLES, USA & LONDON
August 2014 – August 2016

Most notably over the Summer 2016, I lived and worked at the Walt Disney World Resort in Florida. I was placed in the Emporium in the Magic Kingdom park where I was working with people from all over the world and was able to learn about different cultures as well as share my own. The best part was creating "magical moments" for the guests, it was amazing to do this and be a part of a family's Walt Disney World experience.

EDUCATION

UNIVERSITY OF YORK, BSC (HONS) MATHEMATICS & PHYSICS, 2:1
2013 – 2016

Modules covered included Quantum Physics, Statistical Mechanics and Special & General Relativity.

TECHNOLOGY SKILLS:
LaTeX

KEY SKILLS

Python Libraries (Anaconda):
PyMongo, Pandas, Numpy, scikit learn, Plotly and Twython
Cloudera Hadoop version 5.4:
HDFS, Yarn, Sqoop, Flume, Avro, Parquet, Impala, Hive and Hue

NoSQL:	**SQL:**
MongoDB. API with PyMongo	SSIS, Advanced T-SQL and Stored Procedures

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:
Linux CentOS 6, SSH, Putty, Agile, JSON and BSON, Python Pickle and Shelve