CHAY WOODFORD
Big Data Engineer

ABOUT ME

Data analysis is at the core of any forward-thinking business and touches not only big corporations, but also currently less data driven institutions or processes in fields like education and hiring.

How Big Data can be used to shape policy addressing key issues, such as Cambridge University's work into the UK's skill gap, is a particular interest of mine. I believe social diversity and fair treatment of all is of the highest importance in today's society.

My experience as a maths teacher has shown me how to convey technical concepts, the importance of being organised and the difference that enthusiasm can make. I am determined, analytical and energetic, relish challenges and seeing the impact my solutions have.

As part of my teaching, I travelled to Dunkirk to teach maths to refugees which was a very rewarding experience.

CONTACT US

Kubrick Group

T: 020 3866 4620

E: consultants@kubrickgroup.com

W: kubrickgroup.com

kubrick
A NEW FORCE IN TECHNOLOGY CONSULTING

OVERVIEW

I graduated from Durham University in Mathematics. During my degree I concentrated mainly on statistics, where the final project looked at the Mathematics involved in modelling financial securities, with an in-depth look at CAPM's and their effectiveness.

In addition, I spent some time developing programming skills which originate from my Computing A-level. Of note the A2 year was dominated by an individual project, where I created an online exercise platform that required me to learn Javascript, PHP and SQL of my own accord.

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.

MATHS GRADUATE ASSISTANT, THE PERSE SCHOOL, CAMBRIDGE
August 2015 - August 2016

My main role was planning and teaching math lessons to some of the brightest students aged from 11 to 18. Away from classroom I kept my schedule packed, from organising Maths 'Surgery' sessions to coaching rowing along the Cam.

EDUCATION

BSC (HONS) MATHEMATICS 2:2, Durham University
2012 – 2015

ADDITIONAL TECHNOLOGY SKILLS:
VB - Java - JavaScript - HTML - CSS - PHP - SQL - Prolog - R - 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