JONATHAN GWYNNE
Big Data Engineer

ABOUT ME

I enjoy seeing the different usages of Big Data in so many different industries – for example, the game developers Valve Corporation used data analysis in the development of the game 'Left 4 Dead' and changed how players view each other. This resulted in a 40% increase in cooperative play.

I am goal driven, very ambitious, constantly challenge myself and thrive to learn and succeed.
In my A Levels I pushed myself from a D in mathematics to an A* and finally a 1st Class Honours in my mathematics degree. The journey taught me how to keep a level head under pressure and that determination prevails.

I am fascinated by Japanese culture and have taught myself Japanese during university.

CONTACT US

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

OVERVIEW

I recently graduated from Reading University in Mathematics with a 1st Class Honours. During my final year project at university I used data analysis to mathematically model submarine warfare in WW2. I collected data on the number of submarines built and destroyed, the different types of weaponry, their success rates and a number of other factors and consolidated that data using Excel. I turned that data into a set of 6 stochastic differential equations using Matlab. I discovered Big Data after I completed my degree and was drawn to it by its ability to analyse complex and fast paced situations such as monitoring formula one cars mid race.

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.

INTERNSHIP: CROWE CLARK WHITEHILL 2012 – 2014 (8 weeks)
Built a spreadsheet from scratch to consolidate data on the use of the company's conference rooms to tally expenses and determine which events were taxable and which were not.

INTERNSHIP: SAMSUNG'S ELECTRONICS' RESEARCH CENTRE 2011 (2 weeks)
Tested a new website building tool for the company by developing a webpage of Olympic Games medal winners (by country) including real-time images of selected events and a video stream of live coverage. Wrote a report on the new tool in which I recommended potential software improvements.

EDUCATION

MATHEMATICS BSC (FIRST CLASS), Reading University 2013 – 2016
Achieved a First or a 2.1 in over 90% of my modules and 80% in my final year project.

A-LEVELS, Glyn School, Ewell, Surrey 2011 – 2013
Maths: A*, Physics: B, Psychology: D

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