

## OVERVIEW

I recently graduated from Swansea University in Business Management in Information Systems. For my dissertation, we designed a mobile application and I took on the role as Business Analyst, gathering and analysing the data in Excel. With a tremendous amount of data found, it was a challenge to transform the gathered data into usable information, which triggered my curiosity 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.

### INTERNSHIP: PROJECT ASSISTANT, HONG KONG TRADE DEVELOPMENT COUNCIL

#### July 2015 – September 2015

Assisted in the department of E-Commerce, using Google tools such as Google Adwords, Google Analytics, Google webmaster, and contributed to Keywords research in Search Engine Optimisation.

## EDUCATION

### BSC BUSINESS MANAGEMENT IN INFORMATION SYSTEMS, 2:1

Swansea University

2013 – 2016

I gained experience in: Project Management, Business Statistics, Business Information Systems and Search Engine Optimisation.

## 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:

MongoDB. API with PyMongo

#### SQL:

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



TIMOTHY LAU

Big Data Engineer

## ABOUT ME

Coming from a business background, the influence data analysis has on business decision is truly fascinating to me. However Big Data goes beyond playing an important role in businesses. It extends to analysing our everyday data and can help prevent illnesses at an early stage and potentially making treatment more effective and less costly. The rise in supply of data from phone apps and smart watches has enabled the healthcare sectors to draw valuable conclusions from user's geographic location, movements and sleeping patterns. Ultimately Big Data could improve human life expectancy.

During college, I regularly volunteered for the Crossroads Foundation, managing the logistics of collecting second-hand goods and distributing them to people in need internationally.

Having grown up in Hong Kong I am fluent in English, Mandarin, Cantonese. Aside from working, I enjoy playing musical instruments and have been playing the violin and piano for 15 years. I find music improves my concentration.

## CONTACT US

Kubrick Group

T: 020 3866 4620

E: [consultants@kubrickgroup.com](mailto:consultants@kubrickgroup.com)

W: [kubrickgroup.com](http://kubrickgroup.com)