Michael. C. J. Kao

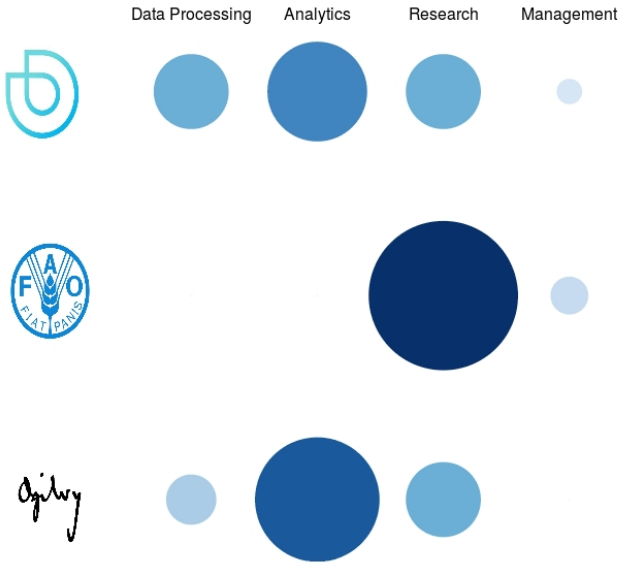
Curriculum Vitae

Personal Summary

Highly skilled data scientist with hands-on knowledge of the latest statistic and machine learning technique with a problem-solving mindset at the core.

Proficient in R and Python for analysis and data preparation. Seasoned in working with standard database and adept at scrapping unstructured data from the web.

Strong business acumen and proven records of using analytics to enhance business performance.



Work Experience

November 2016 – July 2017

Deepblu Inc, Taiwan

Senior Data Scientist

Head the creation of the data team of a scuba diving social media startup.

Laid the data infrastructure foundation, formulation of **expansion and retention strategy** and also **improved mobile app UX** with multiple custom algorithms .

October 2011 – Present

FAO of the United Nations, Italy

Lead Statistician

Lead the development of several flagship analytical projects including the latest Food Balance Sheet (FBS) for **monitoring global supply and demand of food** and "The Reading Machine" an AI program to decipher news article for **detection of food price anomalies**.

June 2010 – November 2011

Ogilvy & Mather, New Zealand

Data Analyst

Analytical consultant delivering solutions to enhance decision making and business performance. Devised novel solutions to achieve goals from **increase return on media investment** to **provide statistical evidence for policy formulation**.

Software Skills

Languages	R, Python and shell
Packages	mlr, shiny, pandas, numpy, scipy, sciki-learn, tensorflow, airflow and django
Database	*Sql, PostGIS, Mongo
Other	Linux, Docker and Git

Education

2010 – 2012 **M.Sc. in Statistic**
University of Auckland

2009 – 2010 **B.A. (1st Class Hons.) in Statistics**
University of Auckland

2005 – 2009 **B.A. & B.Sc**
University of Auckland

Personal Info

Email:	mkao006@gmail.com
LinkedIn:	https://www.linkedin.com/in/mkao
Github:	https://github.com/mkao006
Nationality:	New Zealand/Taiwan

Selected Projects

2018 **Diamond Analysis**
Scrapped diamond data from James Allen to analyse and identify the best diamond to purchase for the engagement ring.

2017 **Automated Data pipeline**
Employed Airflow to automate and streamline the process of converting raw data into analytical products.

2017 **Dive Log Validation and Classification**
Implemented the isolation forest to detect anomalies in dive logs such as dive computer failure, incorrect air mixture, and infeasible profile to ***improve the soundness and safety of diving****.*

2016 **The Reading Machine (Github)**
Automated forecasting of commodity price utilising information scrapped from web pages. Sentiment extraction and topic modeling of news article coupled with Recurrent Neural Network to forecast the commodity price. The purpose of the project is to ***identify potential food crisis and provide lead time for reaction****.*

2014 **Ensemble Imputation (Github)**
Lead research on flexible and robust imputation methodology. An ensemble learning model was formulated to impute the global production of all agricultural products. ***Over 200,000 time series were tested and imputed****.*

2013 **R package: FAOSTAT (CRAN)**
An R package providing seamless integration to the FAO Statistics database.

2011 **Project 'KARMA' (RSVP Award)**
Forecasted demand and supply of teachers and analysed the labour force to provide insights on the reality of the teaching force. The RSVP and Nexus prize was awarded for ***confirming that the teacher shortage was truly over and assisted in new policy formulation****.*

2011 **Project 'MOA'**
The project estimated the effects of various advertising channel in order to assess the respective efficiency and effectiveness. The estimations were then employed to optimise the allocation of the marketing budget for a large retail banking client. The result was a ***79% improvement in customer acquisition over the existing budget****.*

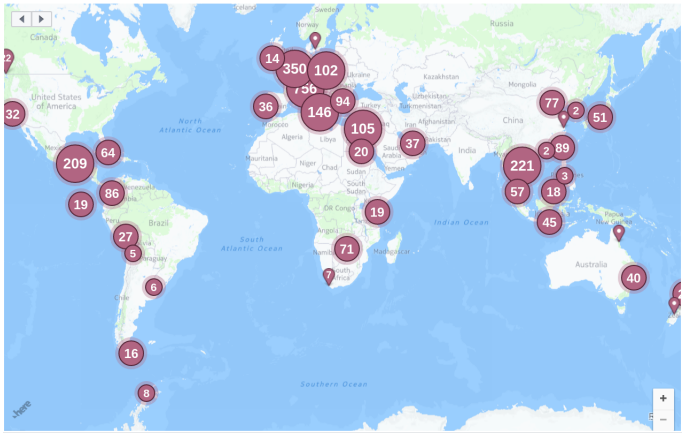
2011 **Campaign 'AND'**
The project identified segments of individuals which are high achievers from students of economically deprived background. The uses of the PRIM algorithm ***pinpoint a segment with a 70% completion rate as opposed to the pool average of 41%****. This resulted in improved utilisation of public funding.*

Language Proficiency

English	Native
Chinese	Native
Taiwanese	Native
Spanish	Conversational – EU A1

Interests

Sports Scuba diving, basketball, boxing and bouldering



Places I have visited