# Michael. C. J. Kao

## *Curriculum Vitae*


github.com/mkao006
New Zealand/Taiwan
+(886) 917 262 817
mkao006@gmail.com
mkao006


## Personal Summary

Highly skilled data scientist with hands-on knowledge of the latest statistic and machine learning techniques 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.

I love diving, in water and in data!

## Software Skills

| | |
|---|---|
| Languages | R, Python and shell |
| Packages | mlr, shiny, pandas, numpy, scipy, sciki-learn, scrapy, 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

## Work Experience

November 2016 – July 2017
Deepblu Inc, Taiwan
*Senior Data Scientist*

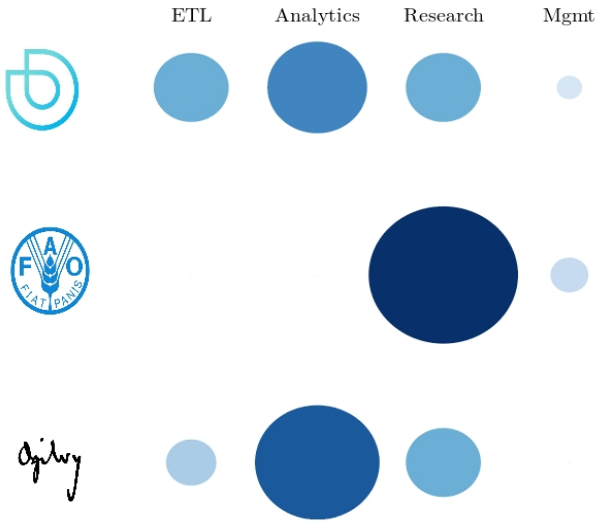
Head of newly created data team, lead the charge on building the data culture.

October 2011 – October 2016
FAO of the United Nations, Italy
*Lead Statistician*

R ambassador and project technical lead.

June 2010 – November 2011
Ogilvy & Mather, New Zealand
*Data Analyst*

Analytical consultant delivering solutions to business problems.




## Work Philosphy

"Simplicity Is The Ultimate Sophistication"

$$\min_{\beta \in R^d} \{\|y - X\beta\|_2 + \lambda\|\beta\|_1\}$$

## 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 **Data Pipeline Automation**
*Employed Airflow to automate and streamline the process of building the data lake.*

2017 **Identify Anomalies**
*Implemented the isolation forest to detect anomalies in dive logs such as dive computer failure to* ***improve the reliability of our product and the safety of diving****.*

2017 **Path to Exponential Growth**
*Anlaysed the structure and connectivity of the market and devised* ***acquisition strategy for highest potential growth*** *and* ***highlighted the areas of churn.***

2016 **The Reading Machine (Github)**
*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****.*

2014 **Food Balance Sheet (Github)**
*An update to the latest methodology for the Food Balance Sheet (FBS). The work provides a basis for* ***monitoring the suppy of demand of food*** *and ultimately the* ***estimation of the number of undernourishment around the globe.***

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

2011 **What Teacher Shortage? (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 **Marketing Optimisation Analysis**
*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 **Finding High Achievers**
*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.*

## Interests

Hobbies Scuba diving, basketball, boxing and travelling


*Places I have visited*