**SELWYN McCRACKEN, Ph.D.**

I am a Ph.D. qualified data scientist with extensive experience at applying classical and innovative statistical techniques across the Financial Services, Government, Pharmaceutical, Healthcare, Automotive and Consumer Goods sectors.

A portfolio of examples is available on my website [http://selwyn-mccracken.com](http://selwyn-mccracken.com/)

**PROFESSIONAL EXPERIENCE**

**Ernst and Young - Fraud Investigation and Dispute Services (London)**

**Data Analytics Manager, October 2011 – September 2012.**

I co-managed a team of 6 - 10 data analysts at the heart of Ernst &Young’s largest global project, which aims to improve the public finances of a sub-Saharan African Government.

My principal role was to lead and coordinate the analytical support my team provided to the 200+ E&Y business analysts and forensic accountants that are assisting across a range of Government activities such as Budget, Revenue, Expenditure, Debt Management, IT & HR.

Examples of specific activities that I led include:

* Implementation of an interactive web-based visualisation and financial reporting portal;
* Diagnosis of system integration issues across the key Government databases;
* Design, review and validation of predictive models (e,g, identifying Ministries that will overspend their budget) ;
* Reconciliation of external debt disbursements and repayments, across the various debt monitoring systems;
* Tax payment and fine monitoring of companies granted tax breaks and exemptions;
* Quality assurance reviews and code version control processes for all requests

**Ernst and Young - Fraud Investigation and Dispute Services (London)**

**Data Analytics Executive, March 2010 – September 2011**

Completed several pieces of work on E&Y’s largest project such as:

* Tracing and visualisation of $8.5 Billion of tax revenue through more than 200 taxation accounts
* Confirmation of $5 Billion of suspected outstanding payments to Government vendors
* Account reconciliation and fuzzy transaction matching between the Government's financial system and commercial bank account statements
* Clustering of Ministries according to budget and expenditure behaviour to identify high variance outliers
* Hadoop-based large-scale parallel computation analysis proof of concept trial
* Automatic classification of financial system transactions from free-text descriptions in SQL

**Vanguard Strategy (London)**

**Senior Analyst, January 2008 – February 2010**

Provided statistical support on client engagements, research design, and presented results to clients.

Major client engagements included:

*Continental tyres (Germany, UK, Czech Republic) – Tyre market consumer dynamics:*

* Established the main drivers of tyre purchases for the client brand, brand loyalty, and the development of brand preference using logistic regression modelling.
* Estimated customer lifetime value via survival analysis (adjusting for car purchasing behaviour, preference decay and brand switching patterns).
* Forecasted market share equilibria of tyre channels

*Wyeth Pharmaceuticals (UK) – Effexor (antidepressant) end-of-life sales force reallocation*

* Saved £1.75m in salary and support costs by optimising the size of a Pharmaceutical sales force across geographic regions and product line, after taking into consideration drug lifecycles, promotional effectiveness, intensity, and affordability of sales reps.

*Coca Cola (Japan) – Georgia coffee:*

* Established that Coke Japan were incorrectly targeting canned coffee consumers that were 10-15 years older than the bulk of their customer base, thus potentially alienating the younger consumers required to maintain the longer term viability of their coffee range

**Vanguard Strategy (London)**

**Analyst-Statistician, June 2006 – December 2007:**

* Identified high-value Physician segments within the US Anti-Depressant market.
* Identified early adopters of a new Anti-Depressant drug based on past analogue uptake.
* Estimated the level of influence that a Specialist has upon their linked GPs prescribing behaviour.

**Dept. Preventive & Social Medicine, University of Otago (New Zealand)**

**Health Research Council Training Fellow, April 2001-December 2005:**

* Undertook PhD Research, epidemiology tutoring, supervisor research assistance, and Public Health training
* PhD research focused on the design and analysis of a national injury statistics database for the New Zealand construction industry, requiring:
* Regression based analysis of construction injury risk factors and injury rate fluctuations over time.
* Machine-learning classification of free-text injury descriptions.

**Dept. Preventive & Social Medicine, University of Otago (New Zealand)**

**Junior Research Fellow, 1999-2001**

* Identified risk factors for Fatal Work Related Injuries amongst Maori Workers.
* Established that fatality rate differences for Maori & Non-Maori workers is due to occupational concentration.

**TECHNICAL SKILLS**

Expert user of the **R** programming language:

* Multivariate regression modelling and prediction (linear & logistic regression).
* Tree-based segmentation and classification (e.g. random-forests)
* Loyalty/Survival/Time-to-Event modelling
* Machine Learning classification & pattern recognition (SVMs, kNN, kmeans)
* Data cleaning, missing-data interpolation and imputation
* Monte-Carlo and bootstrap analyses, kernel density estimation
* Advanced interactive graphics – e.g. multi-dimensional plots, map making and plot animation
* Time series analysis and forecasting

Seasoned **SQL** database analyst:

* PostgreSQL, MS SQL Server, MySQL, MS-Access, SQLite.

‘Big Data’ tools experience with **Hadoop**:

* Hadoop cluster deployment on Amazon’s EC2 with Apache Whirr.
* Mahout machine learning library.
* Pig.

Interactive web-based data visualisation skills:

* D3.js.
* JQuery + Javascript.

Considerable **Python** programming experience:

* Scripting, analysis automation, database ETL, data cleaning, & text processing (including XML).
* Compilation of standalone executables (py2exe)
* Automated MS PPT & Word report generation (win32com)
* Hardware interfacing (pyserial)

Capable **Linux** user and administrator (primarily the Debian and Ubuntu distros):

* Shell scripting and task automation
* Remote access and cluster computing configuration (e.g. initiating a Hadoop cluster via SSH on Amazon EC2)
* Git version control

**EDUCATION**

2007 Ph.D. – Occupational Epidemiology. University of Otago (New Zealand)

2001 MSc – Psychology, University of Otago

1999 BSc – Psychology, University of Otago

**PERSONAL DETAILS**

Nationality: New Zealand

UK Work Visa: Tier 1 General (Have just applied for Indefinite Leave to Remain)

Email: selwyn.mccracken@gmail.com Mobile: (+44) 789-686-9183

Website: [http://selwyn-mccracken.com](http://selwyn-mccracken.com/)