Ian K. Kyle

(916) 534-8519 • iankyle1@gmail.com • ikkyle.com

Education

University of California, Davis   **Davis, CA**
B.S., Psychology   *2010–2013*
Emphasis in developmental psychology

University of California, Davis   **Davis, CA**
Minor, Statistics   *2010–2013*
Emphasis in statistical computing

American River College   **Sacramento, CA**
Community College, Psychology   *2008–2010*
Emphasis in human development

Experience

UC Davis MIND Institute   **Sacramento, CA**
Clinical Data Manager and Analyst   *2008–2015*
Maintain MySQL database of behavioral data, construct datasets for analysts, generate data reports with R and LaTeX, write scoring programs in MySQL stored procedure language, head various analysis projects.
Job duties:
- Write R code for transforming raw longitudinal data and inserting into a relational MySQL database;
- Maintain MySQL database:
  - Query database to construct datasets for analysts
  - Write data cleaning procedures in MySQL to find orphaned data, data that are out of the expected range.
  - Write data cleaning procedures in R to find outliers that may represent data entry or scoring errors
- Pioneered a method for creating matched samples for treatment data, written in R, that utilizes bagging, a method borrowed from machine learning
  - Working on a paper demonstrating that this method can be preferable to the standard method, propensity score matching, in small samples
- Analyze behavioral and psychophysiological data
  - Analyze longitudinal eye-tracking data using a variety of techniques including generalized linear mixed-effect models in R using lme4
    · I have seen this project through from the collection, processing and scoring of data, loading into a relational database, further processing in R, and analyzing using complex mixed models.
    · Currently writing a first author paper on this analysis
  - Analyze longitudinal parent-child dyadic vocal pitch data in conjunction with behavioral codes using a lag-sequential approach
    · I created the analysis dataset all in R and SQL by aggregating about 400 text files with millisecond resolution pitch data into a form that could be merged, based on timestamps, into behavioral codes with 1-5 second resolution
  - Write and maintain web applications used for dataset creation and data exploration, written in R

Baby Siblings Research Consortium   **Sacramento**
Clinical Data Manager and Statistical Programmer   *2011–2015*

Maintain relational MySQL database, process and clean data submissions, generate reports, create datasets, assist with analyses

Job duties:

○ Convert consortium data from a variety of formats (long vs. wide, excel, csv, SPSS, SAS, etc) into a format that can be inserted into the MySQL database.
  - Wrote an R package that automates much of the processing. Processing functions use meta-data on the data submission to automatically read in, save, reshape, and upload the data to the database.
○ generate datasets for use in collaborative analyses; Initially querying data with SQL and further process in R if necessary.
○ Write data reports using a R package called knitr that integrates R code with LATEX
  - The reports are written such that when data in the database changes, the report are recompile and values are updated automatically.
  - Other reports accept parameters specifying a particular consortium site and can be compiled into a report customized for that site that can be used by the PI to validate data submissions.

UC Davis Horticulture Innovation Lab   **Davis, CA**
Application programmer   *2014–2015*
Program a statistics web application to be used by agriculture researchers in developing countries.

○ The app is programmed in R and is used as a tool to analyze data and teach R programming so that researchers in developing countries can expand beyond the analysis tool to running more complex analyses.
○ The app is currently being demoed at universities in Pakistan and may expand to countries in Africa as well

Relevant Coursework

○ **Multivariate Statistics**

focused on linear algebra, multivariate normal distribution, machine learning methods. This course was heavy in R programming and was taught by Duncan Temple-Lang, one of the core authors of R

○ **Mathematical Statistics**

probability theory, Bayes theorem, continuous and discrete probability models, PDFs, limit theorems, asymptotic theory, multivariate calculus

○ **ANOVA**

methods in factorial ANOVA, post-hoc tests, using R for ANOVA

○ **Linear Regression**

methods in multivariate linear regression, model selection, linear algebra, using R for regression

○ **Statistics in psychological research**

Focused on research methods in psychology and statistical methods commonly used in behavioral research. Used R for all analyses

○ **Microbiology**

Cell structure and function, cell physiology, cell reproduction, Mendelian and molecular genetics, evolution, and ecology. Labs included PCR, electrophoresis, and bacterial transformation

○ **Evolution and Ecology**

Introduction to basic principles of ecology and evolutionary biology, focusing on the fundamental mechanisms that generate and maintain biological diversity across scales ranging from molecules and genes to global processes and patterns.

○ **Biodiversity and Phylogeny**

Introduction to organismal diversity, using the phylogenetic tree of life as an organizing theme. Lectures and laboratories covered methods of phylogenetic reconstruction, current knowledge of the tree of life, and the evolution of life's most important and interesting innovations.

○ **Physiolgical Psychology**

Hormones and hormone pathways, brain physiology

○ **Genes and Gene Expression**

Nucleic acid structure and function; gene expression and its regulation; replication; transcription and translation; transmission genetics; molecular evolution.

○ **Organic Chemistry**

Introduction to the nomenclature, structure, chemistry, and reaction mechanisms of organic compounds as well as laboratory concerned primarily with organic laboratory techniques and the chemistry of the common classes of organic compounds

○ **Introduction to Quantum Physics**

Quantum particles and astrophysics, including star birth and death, black holes, nuclear fission and fusion

Relevant Skills

○ **R code**

○ **Data analysis**

○ **Sweave/knitr**

○ **SQL**

○ **SAS code and macro language**

○ **MySQL Stored Procedures**

○ **Open Database Connectivity (ODBC) protocols**

○ **Relational Database Management**

○ **Regular Expressions**

○ **LaTeX**

○ **SPSS syntax**

○ **Visual Basic (VBA)**

○ **Knowledge of research methods**

○ **Knowledge of data analysis techniques**

References

Programming, Analysis, Data management

○ Gregory S. Young
  gnuoysg@gmail.com

Analysis, work ethic

○ Sally Ozonoff
  sally.ozonoff@ucdmc.ucdavis.edu