

# Ye Meng

Programmer Analyst at PPD

---

## Experience

### **Programmer Analyst at PPD**

March 2016 - Present (4 months)

- Independently create, execute, maintain, and validate programs that transfer data across multiple data management systems or operating systems, combine data from a variety of sources and structures, generates and stores summary data from a variety of sources, generate reports or combine multiple databases.
- Independently create, execute, maintain, and validate programs that generate listings, tables and figures using SAS.
- Perform other programming tasks as necessary to support Clinical Data Management and Biostatistics.
- Assess and ensure the validity of all program output.
- Serve as a project lead on key deliverables within a study and oversee work of other programmers on team.
- Document all work; manage assignments to meet timelines and produce high quality deliverables.
- Function as a contributing member of a multi-disciplined team.
- Manage project teams for multiple projects. Conduct team meetings, meet project timelines, and assess resources needed, forecast and input to budgetary needs.
- Provide general infrastructure support to the Biostatistics Department. Examples include presentation / teaching at Department meetings, contributing to other general department documents and policies, or assisting with Biostatistics web page content.

### **Associate Programmer Analyst at PPD**

November 2014 - March 2016 (1 year 5 months)

### **Statistical Programmer Analyst at Keck Medicine of USC**

April 2014 - September 2014 (6 months)

- Derived and created the analysis datasets from raw dataset according to questionnaires and study design specifications through performing SAS Data Manipulation, PROC SQL, MERGE, PROC SORT
- Provided descriptive statistics using PROC MEANS, PROC UNIVARIATE; provided inferential statistics using PROC FREQ, PROC GLM
- Pulled out data from the SAS datasets, transformed data of various formats (Excel, CSV, etc) into R datasets, as well as cleaned and resolved data issues
- Loaded data, investigated missing data pattern of outcome, exposures and health questionnaire, subset data, clean/create variables, write functions in R, etc.
- Conducted model selection for large amount of genetic data using modern techniques random forests, generalized boosted models, regularized regression and Bayesian stochastic search

- Built Prediction model on Children's Health Study data, fitted model with training data, conducted cross-validation, predicted model on new cohort and evaluated model predictive ability AUC criteria in R gbm, cran, longitudinal package
- Explore GWAS data using R plug-in function in PLINK
- Revised research proposal, like giving suggestions on selecting appropriate machine learning algorithm for some model selection

---

## Certifications

### **SAS Certified Base Programmer for SAS 9**

SAS License AP010529v9

### **SAS Certified Advanced Programmer for SAS 9**

SAS License BP038126v9

### **Machine Learning**

Coursera Verified Certificates

### **Intro to Computer Science**

Udacity

### **Introduction to Data Science**

Coursera Verified Certificates

---

## Skills & Expertise

**Biostatistics**

**R**

**Machine Learning**

**Logistic Regression**

**SAS**

**Data Analysis**

**Matlab**

**SQL**

**Python**

**Java**

**Statistical Computing**

**Linear Regression**

**Epidemiology**

**Pattern Recognition**

---

## Education

### **University of Southern California**

Master's Degree, Biostatistics, 2012 - 2014

### **Xiamen University**

Bachelor's Degree, Biochemistry and Biotechnology, 2007 - 2011

### **University of California, Santa Barbara**

---

# Ye Meng

Programmer Analyst at PPD

---



[Contact Ye on LinkedIn](#)