

Bailey Zhang

Data Scientist, Statistician

Germantown, MD - Email me on Indeed: [indeed.com/r/Bailey-Zhang/a70eea8315af54ca](https://www.indeed.com/r/Bailey-Zhang/a70eea8315af54ca)

Authorized to work in the US for any employer

WORK EXPERIENCE

Data Scientist, Statistician

Synchronoss Technologies, Inc - March 2017 to June 2017

- Studied the structure of datasets in preparation for analysis to help telecom companies achieve the maximum revenue, increasing retention and reducing churn. Utilized methods such as determining the appropriate primary composite keys for each dataset by deleting unnecessary duplicate records, linking datasets together to attach additional information, and rolling the data up to get a composite summary statistic, such as a sum or mean, for a relevant group.
- Identified issues with the datasets and methods to clean them, including checking for missing values and imputing them properly, checking for outliers using a winsorising method or using the variance inflation factor, and making sure that there are no future variables in the datasets.
- Used a combination of R and SQL to analyze large datasets of up to 216 million records.
- Acquired experience in using statistical and data management methods in fast-paced working environments, where the project timeline is very short and the client demands often changed.

Quantitative Data Scientist

Booz Allen Hamilton - June 2015 to March 2017

- Created a compliance risk model to select business partnership cases in order to measure the audit risk of partnerships failing to comply to file their tax returns correctly.
- Used a combination of SAS and SQL to analyze large datasets of up to 54 million records and thousands of variables.
- Wrote multiple informative and concise Powerpoint slides in consulting decks describing the results of the analysis to be presented to the business clients, which cater towards the proper audience and use effective deck creation methods.
- Acquired experience in using statistical programs in fast-paced and high-demanding working conditions, communicating with team members about data progress and analysis results, and communicating with the clients in a professional manner to understand their requests, fulfill their needs in a timely manner, and make suitable suggestions on how to proceed.

Data Scientist

Henry Jackson Foundation - April 2012 to June 2015

- Studied the likelihood of suicide attempts among regular army service members from 2004-2009, and the correlation with factors such as demographics, mental health diagnoses, army-related variables, in particular, in relation with the Iraq War and the War on Afghanistan.
- Used a combination of SAS, SQL, and Linux to analyze large datasets up to 51 million records and hundreds of variables.
- Utilized statistical methods in linear regression, sample survey, generalized linear models, etc. such as hazard models, ROC curves, logistic regression, survey design, and spline models, among others.

- Acquired experience in using statistical methods and statistical programs in psychological and survival contexts, and experience in communicating with team members about data information, analysis methods, and results, as well discussing future research.

EDUCATION

Master of Science in Statistics

University of Virginia - Charlottesville, VA

December 2011

Bachelor of Science in Applied Mathematics

Cornell University - Ithaca, NY

December 2009

SKILLS

C+ (Less than 1 year), FUELS (Less than 1 year), SAS (4 years), Serial Attached SCSI (4 years), SQL (5 years)

ADDITIONAL INFORMATION

Computer Skills

Languages: Experienced in SAS, SQL, S-Plus, R, MATLAB, C++, Java, C, HTML, XML, LabView