

## Internship - Statistician/Data Scientist

Hi Fidelity Genetics seeks an intern in statistics/data science for the summer of 2016.

About us: Hi Fidelity Genetics is a small start up based in Durham, NC. We develop tools to measure and predict plant traits. Such tools are essential for maintaining a sufficient supply of food as the world's population grows and its climate changes.

About the internship: Hi Fidelity Genetics is interested in understanding the relationship between plant traits, genetics, and the environment. To that end, Hi Fidelity Genetics is acquiring and generating agricultural data and developing statistical/machine-learning tools for their analysis. The intern will help in both domains, from creating data-wrangling scripts to performing statistical analyses. The specific responsibilities of the intern will depend on her/his prior knowledge, but will likely include constructing models and implementing algorithms.

It is essential that the candidate:

- Is an expert in R.
- Has a strong background in statistics or data science.
- Works in a Unix-based environment and is familiar with tools like awk, sed, and bash.
- Can data-wrangle.

It will be very helpful if the candidate:

- Knows predictive modeling, including linear models with regularization/shrinkage priors, non-parametric models, and mixed models.
- Is familiar with supervised and unsupervised classification.
- Has experience with PostgreSQL or similar.
- Knows other languages, e.g. Python.

Bonus points will be given to a candidate that

- Has experience with bioinformatics, e.g. GWAS.
- Has experience with GIS.
- Knows C/C++.
- Has experience with cloud computing.

Please send a resume, a portfolio or a link to your portfolio on Github or similar, and a *short* cover letter (in plaintext) explaining why you are a good match for this internship to [jobs@hifidelitygenetics.com](mailto:jobs@hifidelitygenetics.com). If selected we may ask for additional information, like recommendations.