

# Christopher Hamm

*Lead Data Scientist*

Carmichael, CA 95608  
USA

✉ [butterflyology@gmail.com](mailto:butterflyology@gmail.com)  
📄 [butterflyology.github.io](https://butterflyology.github.io)

## Summary

Lead Data Scientist with a strong experimental background and 10+ years experience using reproducible research and rigorous analysis to generate unique insights and create value. Involved with Software Carpentry Foundation as Instructor for R and python.

## Experience

2018–Present **Lead Data Scientist**, *Bayer - Crop Science - Decision Science*.

Naveen Singla, supervisor.

Achievements:

- Platinum Pinnacle Award for creating and implementing company-wide digital fluency curriculum
- Supervised 30+ machine learning projects
- Develop, promote, and distribute best practices curriculum for enterprise data scientists
- Organize, coordinate, mentor, and teach data fluency workshops
- Coordinate 15 digital fluency trainers
- Develop proprietary curriculum for company specific workflows
- Created neural network forecasting models for supply chain
- Member:
  - Data Science Center of Excellence
  - Model Governance Team
  - Data Management Team

2017–2018 **Data Scientist**, *Monsanto Company*.

Shawn L. Stricklin, supervisor.

Achievements:

- Developed and implemented a program to simulate the vegetable breeding pipeline
- Implemented predictive modeling to predict genotype sample submission frequency
- Organized and coordinated Software Carpentry and Instructor Training workshops
- Created gradient boosted regression model to identify disease susceptible genotypes
- Provided *ad hoc* statistical support to 10 plant breeders

2016–2017 **Postdoctoral researcher**, *University of California, Davis*.

C. Titus Brown, supervisor.

Achievements:

- Oversaw development and deployment of Reproducible Research with R lessons for the Data Carpentry Foundation

2014–2016 **Postdoctoral scholar**, *University of Kansas*.

James R. Walters, supervisor.

Achievements:

- Developed tools to analyze large empirical and simulated molecular datasets
- Implemented and interpreted complex statistical modeling on ecological and genomic data, including differential expression analysis
- Formalized novel statistical models to describe bacterial infection frequency while correcting for relatedness

2012–2014 **Postdoctoral researcher**, *University of California, Davis*, Davis, California.

David J. Begun and Michael Turelli, supervisors.

Achievements:

- Coordinated multi-laboratory effort to assess the use of a bacterial infection to control an invasive insect
- Executed comparative genomic analysis of invasive species
- Created bioinformatics pipeline to import, trim, and map genomic data

---

## Statistical methods

Deep learning in R & Python using Keras and TensorFlow

Deep learning in python with pytorch and fastai

Machine learning in R & python

Linear and mixed-effects models in R

Forecasting with ARIMA, prophet, and LSTM

Hierarchical Bayesian modeling and inference in R and stan

---

## Computer skills

Data analysis R, Python

Reproducibility docker, knitr, RMarkdown, L<sup>A</sup>T<sub>E</sub>X, jupyter notebooks,

Version control git, GitHub, GitLab

Other BASH, SQL

---

## Education

2012 **PhD**, *Michigan State University*, Entomology.

2012 **PhD**, *Michigan State University*, Ecology, Evolutionary Biology & Behavior.

2008 **MS**, *California State University, Fresno*, Biology.  
*With Distinction*

2004 **BS**, *California State University, Fresno*, Biology.  
*Magna Cum Laude*