

# Colby Ford

www.colbyford.com  
cford38@uncc.edu  
colbytylerford@gmail.com  
(828) 773-8009

## EDUCATION

### UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE

DOCTOR OF PHILOSOPHY  
Computing & Information Systems -  
Bioinformatics &  
Computational Biology  
Wayland H. Cato Doctoral Fellow

### MASTER OF SCIENCE

Data Science & Business Analytics  
First Graduate of the Program

### BACHELOR OF ARTS

Applied Mathematics  
Minor: Psychology  
Coffey Foundation Scholar  
C.L. Robbins Scholar

### CALDWELL COMMUNITY COLLEGE & TECHNICAL INSTITUTE

ASSOCIATE IN SCIENCE  
ASSOCIATE IN ARTS

## LINKS

Github:// colbyford  
LinkedIn:// colbyford  
Research Gate:// colby\_ford  
Google Scholar:// Colby T. Ford

## SKILLS

### PROGRAMMING

Object-Oriented:  
Visual Basic • Python • Java  
Functional/Statistical:  
R • Perl • SAS  
Markup/Design:  
HTML5 • CSS3 •  $\text{\LaTeX}$

### MUSIC

Instruments & Vocal:  
Piano • Organ • Percussion  
Tenor/Baritone Voice • Composition •  
Recording

## OCCUPATIONAL EXPERIENCE

### BLUEGRANITE | DATA SCIENTIST

January 2017 - Present | Charlotte, NC

- Develop machine learning models using Python, R, Microsoft ML Server, and Microsoft Azure Machine Learning.
- Utilize Microsoft ML technologies for massive parallel processing on HDInsight (Hadoop) and Databricks (Spark).
- Host client training workshops and demonstrations of Microsoft advanced analytics technologies.

### LASH GROUP | SR. DATA SCIENCE LEAD

February 2016 - February 2017 | Fort Mill, SC

- Develop machine learning experiments and data analysis workflows to aid in client analytics.
- Discover, understand, and present insights into patient drug adherence.
- Consult in company data architecture including master data management and governance for future business strategy to grow value for client accounts.

### MARINER | DATA SCIENTIST

October 2014 - February 2016 | Charlotte, NC

- Consult in the development of machine learning experiments, including parametric and non-parametric models, statistical predictions, and data mining.
- Build cloud-based solutions for data collection, processing, and storage.
  - Microsoft Azure Technologies:
    - \* Virtual Machine Management with SQL Server
    - \* Machine Learning Studio, AzureSQL Database, & Blob Storage
- Design and create interactive visualizations for both dashboarding and reporting.
  - Technologies:
    - \* Microsoft PowerBI, Tableau Desktop & Tableau Server

### NORTH CAROLINA NEW SCHOOLS | TECHNOLOGY FACILITATOR & MATHEMATICS INSTRUCTIONAL ASSISTANT

April 2014 - October 2014 | Hudson, NC

- Responsible for entire technology inventory: ordering, maintenance, management, etc.
- Design and maintenance of school website.
- Liaison between college & high school technology departments.
- Taught NCVPS mathematics courses as well as tutoring sessions.
- Hosted SAT & ACT preparatory sessions for students.

### STAPLES | STORE OPERATIONS SUPERVISOR & TECH SALES SUPERVISOR

April 2011 - June 2014 | Lenoir, NC

- Managed store planogram setup & inventory as well as employee hiring & labor scheduling.
- Supervised store P&L for EasyTech department for technology, warranty, & repair service sales.

### CALDWELL HOSPICE & PALLIATIVE CARE

### ELECTRONIC MEDICAL RECORDS DATABASE ADMIN

August 2008 - December 2011 | Hudson, NC

- Pioneered the transition from hand-written, paper medical forms to electronic data input and record keeping.
- Worked closely with HIPAA and Medicare/Medicaid guidelines for compliance.
- Served as administrator for the Suncoast database system.

## COURSEWORK

### DOCTORAL

Bioinformatics Programming - Perl  
Molecular Sequence Analysis  
Bioinformatics Programming - Python  
Energy and Interaction  
Research Rotation I  
Research Rotation II  
Drugs: Molecular & Cell Mechanics  
Computational Structural Biology  
Statistics for Bioinformatics

### MASTERS

Database Systems  
Visual Analytics  
Big Data Analytics  
Business Intelligence  
Database Design & Management  
Consumer Marketing Analytics  
Innovation Analytics  
Advanced Business Analytics  
Machine Learning  
Data Warehousing

### UNDERGRADUATE

Mathematical Dataset Modeling  
Mathematical Thought  
Applied Regression  
Financial Mathematics  
Regression Analysis  
Industrial & Organizational Psychology  
Deductive Logic  
Deterministic Modeling  
Cognitive Science - Perception  
Business Programming - Visual Basic

## RESEARCH

### PAPERS

- Ford C., Wen J., Janies D., Shi X., (2018) (*Coming Soon*) parEBEN: A Parallelized Strategy for Improving Epistasis Analysis Based on Empirical Bayesian Elastic Net Models
- Wen J., Ford, C., Janies D., Shi X., (2018) (*ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*) New strategies toward scaling up epistasis analysis on large-scale genomic datasets
- Ford, C. T. (2018). An integrated phylogeographic analysis of the Bantu migration. (Order No. 10748780, The University of North Carolina at Charlotte). *ProQuest Dissertations and Theses*, 120.
- Janies, D.A., Ford, C., Damodaran, L., Witter, Z., (2016) (*Online Journal of Public Health Informatics*) Spread of Middle East Respiratory Coronavirus: Genetic versus Epidemiological Data
- Ford, C., Xue, M., Whiteley, P.M., Wheeler, W., Janies, D.A., Shi, X. (2016) (*Society for Anthropological Sciences*) Visualizing Linguistic Disparity of Uto-Aztecan Languages and Bantu Languages
- Ford, C., Nodzak, C.M., Uppal, A., Zhao, L., (2016) (*Critical Assessment of Genome Interpretation*) Prediction of the effect of naturally occurring missense mutations on cellular n-acetyl-glucosaminidase enzymatic activity.
- Ford, C., Griggs, L., Tillinger, T., & Wood, K. (2015). QVC product airtime analysis.
- Amick, J., Ford, C., Krauss, T., & Wood, K. (2015). Toyota: preparing for the next generation Camry.
- Ford, C. (2015). CongloMed: conglomerate medical data warehousing for analysis, mining, and classification.
- Ford, C. (2015). Creating an integrated, analytical system for the modern, digital business. Masters Practicum.
- Bulusu, M., Ford, C., Ramesh, A., & Patil, P. (2014). VisUFO: organization and representation of UFO sighting data.

### OTHER WORK

- Ford, C. & Lathrop, A. (2017). (*Presentation - Analytics for Social Good, University of Cincinnati.*) Predictive modeling of vegetation density using R and a cloud data platform.
- Ford, C. (2015). Demand forecasting using machine learning to reduce working capital. *Mariner White Paper*.
- Ford, C. (2015). The allure of machine learning, now within reach in Microsoft Azure. *Mariner White Paper*.
- Ford, C., & Snyder, W. (2015). Revenue protection using machine learning for utilities management. *Mariner White Paper*.
- Ford, C., (2016). Assessment of retail out-of-stock conditions using statistical inference. *Mariner White Paper*.

### BOOKS

- Ford, C. (2010). *Caesura, late-intermediate piano technique book*. Hudson, NC: Author.  
ISBN: 978-0-557-36832-7

## TEACHING

Advanced Functions & Modeling  
Discrete Mathematics  
Precalculus  
Machine Learning  
SAT/ACT Prep

## RESEARCH

### **JANIES LAB | DISEASE TRANSMISSION RESEARCH**

August 2015 – Present | Charlotte, NC

Working with **Dr. Dan Janies** to understand the potential disparity between the phylogenetic tree of the Middle Eastern Respiratory Syndrome coronavirus and the geographic spread of the pathogen reported in confirmed cases by the World Health Organization.

### **DARPA | BANTU MIGRATION - GENOMICS AND LANGUAGE RESEARCH**

November 2015 – Present | Charlotte, NC

Working with researchers from DARPA, The American Museum of Natural History (**Drs. Ward Wheeler** and **Peter Whitley**), and The University of North Carolina at Charlotte (**Drs. Dan Janies** and **Xinghua (Mindy) Shi**) to understand the genetic, geographic, and linguistic relationship between languages in both Bantoid and Uto-Aztecan groups.

## PROFESSIONAL MEMBERSHIPS

- 2013 The Society for Industrial and Applied Mathematics
- 2014 The American Statistical Association
- 2015 UNCC Data Science Initiative Advisory Board
- 2016 Northeastern University LEVEL Board Member and Lecturer
- 2016 The Society for Anthropological Sciences

## CONFERENCE ATTENDANCES

- |           |  |                   |
|-----------|--|-------------------|
| Nov. 2014 | Microsoft Roadmap Event ( <i>Speaker</i> )                   | Charlotte, NC     |
| Sep. 2015 | Microsoft Cortana Analytics Conference                       | Seattle, WA       |
| Sep. 2016 | Advanced Pharma Analytics                                    | Newark, NJ        |
| Feb. 2017 | Analytics for Social Good - U. Cincinnati ( <i>Speaker</i> ) | Cincinnati, OH    |
| Mar. 2017 | Society for Applied Anthropology ( <i>Speaker</i> )          | Santa Fe, NM      |
| Nov. 2017 | Nissan Analytics Expo ( <i>Speaker</i> )                     | Nashville, TN     |
| Jan. 2018 | rstudio::conf 2018   | San Diego, CA     |
| Jun. 2018 | Spark+AI Summit 2018   | San Francisco, CA |

## CERTIFICATIONS

- May 2018 Databricks Certified Developer - Apache Spark 2.x for Python