Manju M. Johny

Data Scientist · Statistician

manjujohny@gmail.com | mjohny.github.io | mjohny | mjohny

Summary

- Data scientist passionate about finding creative ways to solve new problems.
- 6+ years experience in research, teaching, and collaborative projects (science, engineering and business).
- Interested in using statistics and machine learning to address real world problems.

Education

Doctor of Philosophy in Statistics *Dec 2021*
Iowa State University, Ames, Iowa.

Master of Science in Statistics *Aug 2017*
Iowa State University, Ames, Iowa.

Bachelor of Arts in Chemistry and Mathematics *May 2014*
Saint Louis University, St. Louis, Missouri.

Skills

Programming R, Rstudio, JMP, Excel, Git, LaTeX (fluent); Python, Jupyter Notebook, SAS (familiar)
Data Science Time series, functional data analysis, spatial statistics, machine learning, visualization
Languages English, Malayalam (fluent); Spanish (familiar)
Other Effective public speaker with excellent teaching evaluations; creative and fast learner

Employment

Postdoctoral Researcher, Data Science *Sept 2021-Present*
NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
Developed novel statistical methods to study carbon cycle dynamics and their relation to extreme weather events.

Graduate Research Assistant *Jun-Jul 2021*
Iowa State University, Ames, Iowa.
Developed statistical methods to study plant phenology, and assessed the methods using simulation experiments.

Graduate Teaching Assistant *Aug 2014-May 2021*
Iowa State University, Ames, Iowa.
Instructed in-person and online introductory statistics courses ranging from 30-100 students.

Machine Learning Intern *Jun-Aug 2019; Aug-Dec 2018*
NASA Glenn Research Center, Cleveland, Ohio.
Developed artificial intelligence capabilities for PeTaL (Periodic Table of Life) on an interdisciplinary team.

Research Fellow *May-Aug 2014; May-Aug 2013*
US Food and Drug Administration, St. Louis, Missouri.
Performed lab experiments and statistical analysis to test for contamination of pharmaceutical materials.

Mathematics Tutor *May-Aug 2012*
Jefferson College, Hillsboro, Missouri.
Assisted undergraduate students in Algebra, Trigonometry, Calculus, Probability and Statistics.

Honors & Awards

International

- **First Place (Cleveland) and Global Nominee**, *NASA International Space Apps Challenge (Hackathon)* *2018*
  Developed a *design concept for an autonomous free-flyer* to inspect space craft damage using image classification.
- **Second Place**, *Statistical Significance Award, Joint Statistical Meetings* *2018*
  The *poster competition* highlights the contributions of statisticians to society.

- **Second Place**, *Prudsys Data Mining Cup (International Machine Learning Student Competition)* *2016*
  Big data challenge to predict returns for an online retailer (120 teams/88 universities/30 countries participated). We presented our solution at the Prudsys Personalization Summit in Berlin, Germany (leading summit for data-driven automation and artificial intelligence in retail).

Domestic

- **Teaching Excellence Award**, *Iowa State University* *2018*
  Award recognizing outstanding graduate teaching achievement.
- **Alumni Fellowship**, *Iowa State University* *2014*
- **ORISE Fellowship**, *Oak Ridge Institute for Science and Education* *2013-2014*
- **Pi Mu Epsilon Member**, *US National Mathematics Honor Society* *2014*
- **Vice President's Scholarship**, *Saint Louis University* *2010-2014*
- **Bright Flight Scholarship**, *Missouri Department of Higher Education* *2010-2014*

Media Recognition

- **SIAM News Blog**, *Article* highlighting my research on statistical approaches to monitor wildfire recovery. *2022*

Research

Postdoctoral Research *Fall 2021-Present*

NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
Advisors: Jon Hobbs, Ph.D; Vineet Yadav, Ph.D.

- Statistical data fusion of satellite measurements from different instruments (OCO-2 and TROPOMI) to study the underlying behavior of solar-induced chlorophyll fluorescence (SIF), a measure of photosynthesis, using Hierarchical Bayesian methods (manuscript under preparation).
- Utilized satellite measurements of SIF to study the effect of wildfires on vegetation and monitor post-fire vegetation recovery using functional data analysis and spatio-temporal methods (manuscript under preparation).

Ph.D Dissertation Research *Fall 2017-2021*

Iowa State University, Ames, Iowa.
Title: "Functional ANOVA-type methods with interpretable visualization for comparisons among groups of time series"
Advisor: Petruţa Caragea, Ph.D.
Committee: Emily Berg, Ph.D; Heike Hofmann, Ph.D; Kris De Brabanter,Ph.D; Daniel Nordman, Ph.D.

- Identification of significant differences between groups of time series in biological, ecological, and remote sensing applications using Analysis of Variance (ANOVA) for functional data.
- Developed this method under the presence of temporal and spatio-temporal correlation (Functional ANOVA and Spatial Functional ANOVA), and designed novel visualizations for understanding statistical significance.

Graduate Research Assistantship *Summer 2021*

Iowa State University, Ames, Iowa.
Advisor: Dan Nettleton, Ph.D.; Petruţa Caragea, Ph.D;

- Studied the seasonal dynamics of plant phenology for different areas of California using satellite measurements of solar-induced chlorophyll fluorescence.
- Funded by Plant Science Institute.

Machine Learning Internship *Fall 2018; Summer 2019*

NASA Glenn Research Center, Cleveland, Ohio.
Mentors: Vikram Shyam, Ph.D; Herbert Schilling, Ph.D.

- Developed image and pattern recognition capabilities for Periodic Table of Life (a biomimetic design tool) using deep learning. Neural networks were utilized to identify and locate organisms, and underlying patterns in images.
- Research culminated in a formal talk to NASA Glenn Research Center, and a publication.

Master's Creative Component (Thesis) Research *Spring 2016-2017*

Iowa State University, Ames, Iowa.

Title: "A Functional Anova Approach to Detecting Changes in Soil Moisture and Temperature"

Advisor: Petruţa Caragea, Ph.D

Committee: Diane Debinski, Ph.D, Yehua Li, Ph.D.

- Utilized an ANOVA for functional data to study the effects of climate change on soil parameters. The approach involved smoothing time series using non-parametric methods, and developing a parametric bootstrap procedure.

Research Fellowship *Summer 2014*

US Food and Drug Administration, Center for Drug Evaluation and Research, St. Louis, Missouri.

Mentors: Jason Rodriguez, Ph.D; Connie Gryniewicz-Ruzicka, Ph.D.

- Developed an algorithm in MATLAB to transfer laboratory methods to field instruments. Implemented rapid screening methods to identify adulteration of pharmaceutical materials on Ion Mobility Spectrometry instruments.
- Research culminated in a formal talk to DPA/CDER/FDA.

Research Fellowship *Summer 2013*

US Food and Drug Administration, Center for Drug Evaluation and Research, St. Louis, Missouri.

Mentors: Jason Rodriguez, Ph.D; Hongping Ye, Ph.D.

- Developed Raman and near Infrared spectral libraries for screening of pharmaceutical materials. Performed statistical analysis to test for ruminant contamination in heparin.
- Research culminated in a formal talk to DPA/CDER/FDA, and poster presentation.

Publications & Presentations

Publications:

- **Johny, M. M.**, Hobbs, J., Yadav, V., Nguyen, H., Johnson, M., Braverman, A., Estimation of solar-induced chlorophyll fluorescence using Bayesian Hierarchical Regression (In Preparation).
- **Johny, M. M.**, Hobbs, J., Johnson, M., Braverman, A., Monitoring post-fire vegetation recovery using functional ANOVA (In Preparation).
- **Johny, M. M.**, 2021. Functional ANOVA-Type Methods with Interpretable Visualization for Comparisons among Groups of Time Series (Publication No. 28721344). [Doctoral dissertation thesis, Iowa State University]. ProQuest.
- Vaziri, G. J., **Johny, M. M.**, Caragea, P. C., Adelman, J. S., 2019. Social context affects thermoregulation but not activity level during avian immune response, *Behavioral Ecology*, 30 (2), 383–392, https://doi.org/10.1093/beheco/ary177.
- Shyam, V., Friend, L., Whiteaker, B., Bense, N., Dowdall, J., Boktor, B., **Johny, M. M.**, Reyes, I., Naser, A., Sakhamuri, N., Kravets, V., Calvin, A., Gabus, K., Goodman, D., Schilling, H., Robinson, C., Reid, R. O., Unsworth, C., 2019. PeTaL (Periodic Table of Life) and Physiomimetics, *Designs*, 3(3), 43, https://doi.org/10.3390/designs3030043.
- Rodriguez, J. D., Skaggs, S. K., **Johny, M. M.**, Srivastiva, H. K., Loethen, Y. L.,Arzhantsev, S. L., Kauffman, J. F., Buhse, L. F., 2014, Distribution of Spectral Libraries Across Different Field Deployable Raman and Near Infrared Instruments, *Am. Pharm. Review*, 17 (1), 10-17.

Conference and Meeting Presentations:

- **Johny, M. M.**, July 2023, Monitoring the effect of wildfires using solar-induced chlorophyll fluorescence, Paris, France, *19th International Workshop on Greenhouse Gas Measurements from Space*.
- **Johny, M. M.**, March 2023, The effect of California wildfires on vegetation, Luosto, Finland, *Luosto Workshop*.
- **Johny, M. M.**, Feb 2023, Bayesian hierarchical regression for modeling solar-induced chlorophyll fluorescence, Finnish Meteorological Institute, Helsinki, Finland, *JPL-FMI-Aalto-LUT workshop on data fusion and satellite UQ*.
- **Johny, M. M.**, Feb 2023, Investigating the effect of California wildfires using satellite observations of solar-induced chlorophyll fluorescence, NASA Jet Propulsion Laboratory, Pasadena, CA, USA, *Postdoc Seminar Series*.
- **Johny, M. M.**, October 2022, Hierarchical Fourier regression for modeling solar-induced chlorophyll fluorescence, NASA Jet Propulsion Laboratory, Pasadena, CA, USA, *Uncertainty Quantification for Remote Sensing Inverse Problems Virtual Breakout Meeting*.
- **Johny, M. M.**, July 2022, Investigating the effect of wildfires on solar-induced chlorophyll fluorescence using functional ANOVA methodology, NASA Ames Research Center (Virtual), *NASA Ames Earth Science Seminar*.
- **Johny, M. M.**, April 2022, Investigating Atmospheric Carbon Dioxide and Solar-Induced Chlorophyll Fluorescence (SIF) using Functional ANOVA, Atlanta, GA, USA, *SIAM Conference on Uncertainty Quantification*.

- **Johny, M. M.**, October 2021, Functional ANOVA for Comparing Spatio-temporal Satellite Data, NASA Jet Propulsion Laboratory, Pasadena, CA, USA, *Uncertainty Quantification for Remote Sensing Inverse Problems Virtual Breakout Meeting*.
- **Johny, M. M.**, October 2020, Functional ANOVA for Satellite Data, NASA Jet Propulsion Laboratory, Pasadena, CA, USA, *Uncertainty Quantification for Remote Sensing Inverse Problems Virtual Breakout Meeting*.
- **Johny, M. M.**, October 2018, Periodic Table of Life (PeTaL): Image Classification, Cleveland, Ohio, USA, *NASA GRC Midterm Presentations*.
- **Johny, M. M.**, Caragea, P. C., Debinski, D. M., Sherwood, J., July 2018, A Functional Anova Approach to Detecting Changes in Soil Moisture and Temperature , Vancouver, British Columbia, Canada, *Joint Statistical Meetings*.
- **Johny, M. M.**, Caragea, P. C., Debinski, D. M., Sherwood, J., May 2018, A Functional Anova Approach to Detecting Changes in Soil Moisture and Temperature, Ames, Iowa, USA, *Conference on Predictive Inference and Its Applications*
- Iowa State University Team 1 (**Johny, M. M.**, Chakraborty, A., Han, Y., Li, X., Mao, X., Zhang, H.), July 2016, Data Mining Cup Solution, Berlin, Germany, *Prudsys Personalization Summit*.
- Rodriguez, J. D., Skaggs, s. K., **Johny, M. M.**, Srivastava, H. K., Loethen, Y. L., February 2015, Evaluating the Performance of Field Screening Using Portable Raman and Near Infrared Spectrometers, *IFPAC Conference*.
- **Johny, M. M.**, Skaggs, S. K., Gryniewicz-Ruzicka, C.M., Rodriguez, J. D., August 2014, Development of IMS Library for Detection of Adulterants; Standardization of Raman Spectra Across 5 Different Instruments, St. Louis, MO, USA, *FDA Summer Research Symposium*.
- **Johny, M. M.**, Ye, H., August 2013, Disaccharide Analysis to Test Ruminant Contamination of Heparin, St. Louis, Missouri, USA, *FDA Summer Research Symposium*.
- Rodriguez, J. D., Skaggs, S. K., **Johny, M. M.**, Arzhantsev, S., Loethen, Y. L., Srivastava, H. K., Kauffman, J. F., Buhse, L. F., September 2013, Developing Spectral Libraries for Domestic and Foreign Screening of Pharmaceutical Materials; White Oak, Maryland, USA, *CDER Science Day*.

Teaching

Instructor:

- STAT 330: Probability and Statistics for Computer Science (online), Iowa State University *Fall 2020 - Spring 2021*
- STAT 330: Probability and Statistics for Computer Science, Iowa State University *Spring 2019 - Spring 2020*
- STAT 101: Principles of Statistics, Iowa State University *Fall 2015 - Spring 2018*
- STAT 105XW: Introduction to Statistics for Engineers (online), Iowa State University *Summer 2017*

Lab Instructor:

- STAT 101: Principles of Statistics, Iowa State University *Fall 2014 - Spring 2015*

Grader:

- STAT 401: Statistical Methods for Research Workers (graduate level), Iowa State University *Summer 2015*
- STAT 104: Introduction to Statistics, Iowa State University *Fall 2014 - Spring 2015*

Activities & Community Outreach

- **Mentor:** NASA Club, John Marshal School of Information Technology (high school), Cleveland, Ohio *2018*
  Taught computer science concepts through video game development in an after-school program at JMIT, a high school that serves primarily minority students. Mentored students who were interested in pursuing a career in STEM.
- **Member:** American Statistical Association *2014-2018*
- **Member:** STAT-ers Club, Iowa State University, Ames, Iowa *2014-2017*
- **Member (Demonstration Captain):** Chemistry Club, Saint Louis University, St. Louis, Missouri *2012-2013*
  Performed 'fun' chemistry demonstrations at the St. Louis Science Center to promote science among children. Popular demonstrations included making 'spooky' dry ice bubbles, colorful slime, and carbon snakes.
- **Volunteer:** Lion's Club International, Saint Louis University, St. Louis, Missouri *2013*
  Performed free vision screenings around the community to identify and inform those at-risk for vision loss.
- **Volunteer:** Sunrise Senior Living, Des Peres, Missouri *2008-2013*

Graduate Courses

Theoretical Courses:

- Advanced Probability Theory and Statistical Inference (STAT 542, STAT 543, STAT 641, STAT 642, STAT 643)
  - Probability measures, L_p spaces, conditional probability, moment generating functions, convergence, central limit theorems, sufficiency, estimation, maximum likelihood, decision theory, hypothesis testing, etc
- Theory & Application of Linear Models (STAT 611)
  - Theory of least squares, best linear unbiased estimation, distribution of quadratic forms, etc.

Applied Courses:

- Advanced Statistical Methods (STAT 500, STAT 510, STAT 520, STAT 601)
  - Randomization-based inference, ANOVA, linear models, generalized linear & mixed models, estimation & inference, Monte Carlo studies, bootstrap, cross validation, latent variables, model assessment, etc
- Bayesian Statistics & Advanced Bayesian Methods (STAT 544, STAT 615)
  - Prior specification, hierarchical models, MCMC algorithms, hierarchical models, state-space models, etc.
- Nonparametric Statistics (STAT 546)
  - Smoothing methods for estimating density and regression functions, parameter selection, cross validation, etc
- Time Series Analysis (STAT 551)
  - Stationarity, temporal dependence, MA & AR structures, prediction & forecasting, etc.
- Ecological Statistics (STAT 534)
  - Estimation of abundance, survival from recapture studies, hierarchical models, etc.
- Introduction to R; Statistical Computing (STAT 579, STAT 580)
  - Programming in R, graphics, looping, function construction, introduction to C, interface of R & C

Programming:

- Introduction to R; Statistical Computing (STAT 579, STAT 580)
  - Programming in R, graphics, looping, function construction, introduction to C, interface of R & C

If you are interested in learning more about me or viewing some of my research projects, please visit my website: https://mjohny.github.io