Josemari Feliciano

(310) 347-8863 jfeliciano@aya.yale.edu jfeliciano@alumni.harvard.edu

EDUCATION

Yale University, School of Public Health New Haven CT
Masters of Public Health, **Biostatistics**, Awarded: May 2020.
Key courses: Advanced Regression Analysis, Longitudinal and Multilevel Data Analysis, Applied Survival Analysis, Statistical Practice and Consulting I-II, Advanced Statistical Programming in SAS and R, Probability Theory, Theory of Statistics, Categorical Data Analysis, Spatial Statistics, Multivariate Statistics for Social Sciences.

Harvard University, Extension School Cambridge MA
Bachelor of Liberal Arts, Extension Studies, **Mathematics.** Awarded: May 2017.
Key courses: Multivariable Calculus, Linear Algebra, Numerical Analysis, Intro to Quantitative Methods, Introduction to Proofs, Website Development, Dynamic Web Applications, Political Science Research Methods.

SKILLS & INTERESTS

Statistical Skills: R, SAS, Stata, Python. **Network Skills:** Gephi, R (ggnetwork, ggnet, ggraph, igraph). **Certifications:** Intermediate R, Tidyverse in R, Supervised Machine Learning in R, Advanced Google Analytics, Python, SQL, Tableau Desktop I, Tableau Desktop II, Tableau Desktop III. **Web Development Skills:** HTML 5, CSS, SCSS, Javascript (ES6), Angular, Backbone, Bootstrap, Node, NPM Express, PHP, Laravel. **Other Skills:** Java, C++, MATLAB, LaTeX, Git, databases (SQL, NoSQL, Google Cloud, AWS Redshift), APIs (Twitter, Google Maps, YouTube, NCBI, ClinicalTrials.gov).

RECENT EXPERIENCE

Small Business Administration – Full-time (40hrs/week) Washington DC
GS-14 Data Scientist February 2023-Current

- Data scientist (GS-14) within SBA's Chief Data Officer Division with responsibilities ranging from data policy and governance to dashboard creation (e.g., PowerBI, modifying existing/underlying SQL tables).
- For the data policy aspect of the job, supports SBA's Chief Data Officer (CDO) as a data policy lead. This includes:
  - Analyzing data-related statutes, executive orders, and data guidance from key stakeholders across the federal government (e.g., Office of Management and Budget, CDO Council) to ensure SBA's regulatory compliance. This includes briefing key stakeholders at the agency (e.g., CFO, all CDO staff, OCIO) of forthcoming additional requirements for the agency (e.g., requirements from the Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence).
  - Providing comments to draft regulations, executive orders, and Government Accountability Office reports on behalf of SBA's CDO and the entire agency. For example, I reviewed OMB's draft policy guidance on Artificial Intelligence and provided a number of critical comments and suggestion to OMB as they finalize forthcoming AI policies for all federal agencies.
- For the data governance aspect of the job, supports key aspects of Evidence Act implementation within SBA and other CDO work priorities in support of the Agency's Data Governance Board. Responsibilities include:
  - Assisting the CDO in assigning work to four contractors within our team and actively engaging the Chief Information Officer team and their contractors to modernize our open data infrastructure and web analytics to monitor the public's utility and consumption of our public data assets. For example, I wrote SBA's Standard Operating Procedure (SOP) on how to publish data assets through our Open Data Platform (data.sba.gov) and provide agency-wide guidance on how to meet metadata requirements set forth by the General Services Administration (GSA) for the federal-wide data.gov platform.
  - Representing the CDO in various internal working groups (e.g., mySBA data modernization) and federal-wide working groups (e.g., Zero Trust Security Working Group).
  - Monitoring data maturity goals and progress across the agency by engaging data stakeholders across agency offices through direct engagement, survey formation and analyses.

- Creating tutorials for relevant stakeholders throughout the agency on how to engage, leverage, and analyze USASpending.gov data. This includes my creation of tutorials using R, a statistical programming language, that I singlehandedly authored with applied examples and homework using USASpending.gov data.
- For the data analysis and engineer aspect of my work, creates PowerBI dashboards with a focus on fraud analytics (e.g., monitoring fraud-related recoveries and efforts by the agency). This includes having write access to our Oracle SQL databases to include additional data, variables to existing SQL tables that feeds our fraud-related dashboards (i.e., creating or modifying procedures and pushing changes to production).

Rural Development, US Department of Agriculture – Full-time (40hrs/week) New Haven CT (100% Remote)
GS-13 Statistician (Data Scientist) March 2021- February 2023

- Data scientist (GS-13) within the Innovation Center's Chief Data Officer Division with responsibilities ranging from program evaluation (statistical and analytics support) to database management and data governance.
- For the analyst aspect of the job, performed ad hoc analyses on my agency's multi-billion investments across Rural America, fulfilled last minute data requests and analysis from the undersecretary's office and other key political appointees.
- For the statistician aspect of the job, performed statistical analysis to assess the impact of our investments using longitudinal models. Created training materials for analysts and data scientists within my division using R, Python, and SQL.
- For the data engineering and governance aspect of the job, lead a number of data modernization and transformation workstreams for Rural Development.
  - Lead the creation of SQL views and tables within our Impala databases that are used for our public-facing dashboards. This includes the creation of multiple SQL views that powers USDA's Rural Data Gateway dashboards (www.rd.usda.gov/rural-data-gateway/rural-investments).
  - Maintained my data analytic division's multi-TB MSSQL database which I lead the creation of in collaboration with the USDA IT Team (OCIO) for hosting large research-related datasets. Within my division prior to my departure, I was the only staff with write access to the MSSQL database, thereby allowing me to create more tables/data/views for our researchers.
  - Represented the Rural Development CDO in various working groups (e.g., data modernization, data requirements gathering).
- Lead a number of Customer Experience (CX) projects to improve our services to loan and grant applicants (e.g., survey formation, analysis, studying and extracting databases for customer contact information).

Yale University – Part-time (20hrs/week) New Haven CT
Teaching Fellow – Biostatistics in Public Health II January 2019-May 2020

- Taught Biostatistics labs in R under Dr. Zelterman for BIS 505b in Spring 2019, 2020.
- Taught weekly labs to perform statistical tasks using R, held regular office hours throughout the semester
- Topics include least square estimates, regression diagnostics, logistic regression, Poisson regression, survival curves, cox proportional hazards, and a number of semi-parametric and non-parametric tools.

Yale University – Part-time (20hrs/week) New Haven CT
Teaching Fellow – Regression Models August 2019- December 2019

- Taught weekly labs and a number of lectures (on GLMs) under Dr. Elizabeth Claus for BIS 621.
- Using SAS, taught sample size calculations, regression models and diagnostic tools for longitudinal and repeated measures data, and a deeper understanding of ANOVA, Poisson, Logistic, Cox regression.
- I made the Poisson Regression lecture more rigorous by teaching how to scale the Poisson model using residuals, and other count models such as negative binomial and zero-inflated models.
- Advised PhD students and MDs on statistical methods, study design for project proposals.

Merck – Full-time (40hrs/week) North Wales PA
Data Science and Business Analytics Summer Intern May 2019 – August 2019

- Leveraged statistics, machine learning on prescription-level data to generate insight on long-term medication adherence and prescription abandonment at the pharmacy.
- Technology stack includes SAS, R, SQL, Amazon Redshift. Utilized R's Tidyverse extensively.
- Generated insight via LME, GEE, poisson, quasi-poisson, NB, zero-inflated, hurdle models.

Yale School of Public Health – Part-time (10hrs/week) New Haven CT
Biostatistics Graduate Research Assistant March 2019-May 2020

- Currently a graduate research assistant for Dr. Elizabeth Claus.
- Generates insight using R, tidyverse, network packages (ggnetwork, ggnet, ggraph, igraph), personal API access (Google, Twitter APIs), Gephi to both qualify, quantify network interactions in social media among patients, providers, other key stakeholders within the brain tumor community.
- Performs analytics research to support my PI's International Low-Grade Glioma Registry to collect data on this understudied population of brain tumor patients and survivors.

OTHER RECENT EXPERIENCE

Yale School of Public Health New Haven CT
Student Body President March 2019 – May 2020

- Oversaw a $100,000 budget that the Student Association of Yale School of Public Health (SAYPH) used for student events, orgs, key initiatives.
- Organized a significant portion of student admit day events for prospective Yale students.
- Worked with other board members to organize professional development events, promote well-being.

Yale School of Public Health New Haven CT
YSPH Education Committee Voting Member September 2018 – May 2020

- Hand-picked by YSPH deans to serve as the MPH student representative, voting member in the education committee led by Dr. Melinda Pettigrew.
- Reviewed course content and requirements, and proposals for new courses, requirements, and programs.
- Made recommendations to improve curricula—both school and departmental—when deficits were found.

Warrior Scholar Project New Haven CT
Physics Content Reviewer June 2018- July 2018

- Worked with Dr. Marla Geha (Yale Astronomy, Yale Physics, WSP) and Dr. Jeremy Bradford (Yale Poorvu Center for Teaching and Learning, WSP) to review study materials (assigned readings, problem sets) developed by Dr. Bradford in revamping the Warrior-Scholar STEM curriculum.
- Vetted reading materials, kinematics problem sets to assess their appropriateness for student veterans with little to no Physics background.
- The reviewed materials were later distributed to host campuses across the US (e.g. Yale, MIT, Texas A&M, Syracuse) for WSP events.

ACADEMIC PUBLICATIONS AND OTHER WRITING

Peer Reviewed:

1. Claus EB, **Feliciano J,** Benz LS, Calvocoressi L. **Social media partnerships with patient organizations for neuro-oncology patient recruitment.** Neurooncol Pract. 2020 Mar;7(2):143-151. doi: 10.1093/nop/npz049. Epub 2019 Oct. PMID: 32626583; PMCID: PMC7318861.

2. **Josemari Feliciano**, Liz Salmi, Charlie Blotner, Adam Hayden, Edjah Nduom, Bethany Kwan, Matthew Katz, Elizabeth Claus. **A Twitter-Based Network Analysis of Brain Tumor Social Media (#BTSM)**. Neuro-Oncology, doi.org/10.1093/neuonc/noz175.330, November 2019.

3. Hauc SC, Tshering D, **Feliciano J**, Atayde AMP, Aboukhater LM, Dorjee K, Dukpa T, Rinchen P, Yoezer N, Luc CM, Adhikari RN, Lhamo K, Khoshnood K. **Evaluating the effect of village health workers on hospital admission rates and their economic impact in the Kingdom of Bhutan**. BMC Public Health. 2020 Aug 24;20(1):1277. doi: 10.1186/s12889-020-09347-4. PMID: 32838794; PMCID: PMC7445904.

4**. Feliciano JT**, Salmi L, Blotner C, Hayden A, Nduom EK, Kwan BM, Katz MS, Claus EB. **Brain Tumor Discussions on Twitter (#BTSM): Social Network Analysis**. J Med Internet Res. 2020 Oct 8;22(10):e22005. doi: 10.2196/22005. PMID: 33030435.

5. Hauc SC, Tshering D, **Feliciano J**, Atayde AMP, Aboukhater LM, Dorjee K, Dukpa T, Rinchen P, Yoezer N, Luc CM, Adhikari RN, Lhamo K, Khoshnood K. **A Cross-Sectional Survey Analyzing Community Perception and Utilization of Village Health Workers Stratified by the Urban-Rural Divide Within the Kingdom of Bhutan.** Asia Pac J Public Health. 2021 Jan;33(1):113-116. doi: 10.1177/1010539520969232. Epub 2020 Nov 11. PMID: 33174437.

6. Kwon J, Grady C**, Feliciano JT**, Fodeh SJ. **Defining facets of social distancing during the COVID-19 pandemic: Twitter analysis.** J Biomed Inform. 2020 Nov;111:103601. doi: 10.1016/j.jbi.2020.103601. Epub 2020 Oct 14. PMID: 33065264; PMCID: PMC7553881.

7. Kadakia S, Stratton C, Wu Y, **Feliciano J**, Tuakli-Wosornu YA. **The Accessibility of YouTube Fitness Videos for Individuals Who Are Disabled Before and During the COVID-19 Pandemic: Preliminary Application of a Text Analytics Approach**. JMIR Form Res. 2022 Feb 15;6(2):e34176. doi: 10.2196/34176. PMID: 35044305.

8. Burgess R, **Feliciano JT**, Lizbinski L, Ransome Y. **Trends and Characteristics of #HIVPrevention Tweets Posted Between 2014 and 2019: Retrospective Infodemiology Study**. JMIR Public Health Surveill. 2022 Aug 11;8(8):e35937. doi: 10.2196/35937. PMID: 35969453; PMCID: PMC9412898.

Pending (Working Papers and Recently Submitted) Manuscripts:

1. **Josemari Feliciano**, Vasilis Vasiliou, and Elizabeth B. Claus. **Association of County-Level Breast Cancer Incidence with 1,4-Dioxane and 16 Common Air Pollutants: An Ecological Study.**

2. Anil Rupasingha, Alex Marre, and **Josemari Feliciano**. **Place-Based Tax Incentives and Minority Employment: Evidence from the New Market Tax Credit (NMTC) Program**

External Non-Peer Reviewed Writing:

1. Racial misrepresentation in clinical trials: a call for health equity. Connecticut Mirror. Newspaper, 2020.
2. Permutation Test as an Alternative to Two-Sample t-test Using R. Medium, Blog, 2020.
3. A Pedagogical Reflection: Utilizing a Circle for Simulations and Monte Carlo Methods, Medium, Blog, 2019.
4. Javascript Exponentiation: Iteration, Recursion and an Optimized Recursion, Medium, Blog 2017.
5. Web Development: Color Selection and Counting All Possible Colors, Medium, Blog, 2017.
6. N-Queens: Describing The Problem, Medium, Blog, 2017.

PRESENTATIONS AND TALKS

1. **A Twitter-Based Network Analysis of Brain Tumor Social Media (#BTSM).** Society of Neuro-Oncology Conference, November 2019. Poster.
2. **A Twitter-Based Network Analysis of Brain Tumor Social Media (#BTSM).** Yale Day of Data Conference, December 2019. Poster. Awarded best poster presentation award.
3. **Text Analysis and Visualization in R.** Yale University Library, February 2020. Workshop.
4. **Social Media Partnerships with Patient Organizations for Neuro-Oncology Patient Recruitment.** Yale School of Nursing PhD Colloquium, March 2020. Talk.
5. **Overdispersion and Excess Zeros: Count Models from Poisson to Hurdle Models.** Medidata. July 2020. Talk.

6. **Spatial Analysis of Food Recalls in the United States (2010-2019).** Brigham and Women Computational Data Science Symposium. October 2020. Talk.
7. **Principal Component Analysis Using SAS.** Yale Biostatistics. November 2020. BIS 621 (Regression Models for Public Health) Guest Lecture.
8. **Simulation Studies Using R.** Yale Biostatistics. November 2020. BIS 679 (Advanced Statistical Programming in SAS and R) Guest Lecture.
9. **Advanced Excel Tools.** USDA Rural Development Office of Civil Rights. November 2021. Workshop.
10. **Introduction to the Tidyverse packages and the Census API.** Yale Biostatistics. November 2021. BIS 679 (Advanced Statistical Programming in SAS and R) Guest Lecture.
11. **Getting Started with Spatial Data Using R.** RLadies St Louis. April 2022. Invited talk.
12. **Principles of Data Visualization.** Yale Biostatistics. November 2022.. BIS 633 (Population and Public Health Informatics) Guest Lecture.

OTHER ACADEMIC WORK

1. **Social Distancing Twitter Analysis Dashboard.**
   Fodeh Lab, Yale School of Medicine.
   Role: Designed, created the HTML dashboard.
   Dashboard link: https://samahfodehlab.github.io

2. **HIV Prevention Dashboard.**
   Collaborative research work, Yale School of Public Health.
   Role: Designed, created the HTML dashboard.
   Dashboard link: https://jmtfeliciano.github.io/HIVPreventionDashboard

ACADEMIC CONFERENCE CONTRIBUTIONS

1. Grady C., **Feliciano J. T.**, Kwon J., Fodeh S. J. COVID-19 on Twitter: A Sentiment Analysis of Tweets from Public Health Agencies, News Media, and Normal Users. Oral presentation at: National Conference on Health Communication, Marketing, and Media; October, 2020; Atlanta, GA.
2. Kwon J., **Feliciano J. T.**, Grady C., Nowak G. The Use of Twitter to Communicate COVID-19: A Content Analysis of Tweets from CDC, WHO, and Major News Media. Oral presentation at: National Conference on Health Communication, Marketing, and Media; October, 2020; Atlanta, GA.
3. Lizbinski, L., **Feliciano, J**., Burgess, R., Ransome, Y. Discussion of HIV/AIDS Prevention Methods on Twitter: A Longitudinal Analysis from 2014-2019. 2nd Annual Computational Data Neuroscience Symposium. Harvard Medical School, Boston, MA [Virtual]. Oct 23, 2020.
4. Pender, J., Goldstein, J., Mahoney-Nair, D., Charankevich, H., **Feliciano, J.** Impacts of the Broadband Initiatives Program by Race & Ethnicity. Agricultural & Applied Economics Association (AAEA) Conference, Anaheim, CA. August 2022.
5. Anil Rupasingha, Alex Marre, and **Josemari Feliciano**. Place-Based Tax Incentives and Minority Employment: Evidence from the New Market Tax Credit (NMTC) Program. Federal Committee on Statistical Methodology Research and Policy Conference, Washington, DC. October 2022.
5. Cho, J., **Feliciano, J**., Leckrone, K., Martin, K., Rupasingha A., Wu, T. USDA Rural Development and Wellbeing of Rural America. Association Public Policy Analysis & Management (APPAM) Conference, Washington, DC. November 2022.

JOURNAL PEER REVIEW ACTIVITIES

1. **Journal of Medical Internet Research.** 2020-
2. **American Public Health Association**. 2020-
3. **Yale Journal of Biology and Medicine**. 2023-

AWARDS

Facebook Hack Reactor Fellowship ($20000); Yale Conference Travel Grants ($2000+); Outstanding Poster Award, Yale Day of Data 2019 Conference.

TEACHING PORTFOLIO (MATH/STATISTICS)

I created the following GitHub repository to include a number of files related to my teaching experience:

https://github.com/jmtfeliciano/TeachingPortfolios

This portfolio includes:

1. My teaching slides for Multiple Regression Using R.
2. My teaching slides for Poisson Regression Using SAS.
3. My teaching reviews for Yale's BIS 505b (Biostatistics II) course for Spring 2020.
4. My teaching reviews for Yale's BIS 621 (Regression Models) course for Fall 2019.
5. A homework solution guide I created for the first BIS 621 HW.
6. A practice midterm solution guide I created for Accelerated Math 97/99.