

Santiago Rodriguez

[LinkedIn](#)

[Gitlab](#)

[Github](#)

Relevant Experience (+10 years total experience)

Deloitte

Remote

Data Science Manager (Global Insights & Analytics)

Oct 2022 – Present

My impact in this role was to enhance an annual people analytics pay equity initiative and build a suite of financial models that improved forecast accuracy and provided valuable contextual insights.

- People Analytics
 - Pay Equity – Refactored the code base and introduced modern software engineering practices: git for version control, using box to import libraries, RStudio project files, and renv to ensure reproducibility, YAML to record configurations, and encapsulating code in functions. Improved the statistical rigor using multilevel modeling, linear regression, beta regression, and ordinal logistic regression, as well as statistical inference techniques such as Tukey's HSD and bootstrapped confidence intervals (R)
- Finance
 - Created an automated Databricks workflow to forecast monthly sales dollars for the global CFO using statistics models, machine learning, and deep learning that improved the forecast accuracy by 10% pts over existing methods (Python, Pyspark, Nixtla)
 - Built a model to predict the probability of success of sales opportunities using Sklearn, Catboost and re-calibrated probabilities (Python, Pyspark)
- Leadership
 - Participated in the Texas A&M Buddy Program as a mentor to an aspiring analyst
 - Participated in the BOOST leadership training program

Consumers Energy

Remote

Data Scientist (Marketing)

Mar 2021 – Oct 2022

My impact in this role was to drive innovation, create an experimental design framework, and develop standards for data science, data engineering, and MLOps. Nicknamed Chief Innovation Officer by the Director of the team.

- Innovation
 - Produced two white papers on the applications of functional data analysis in energy and marketing ([2021](#), [2022](#)) and presented at the R-in-Energy group meetup in March 2022 ([YouTube](#)) ([presentation](#))
- Product Data Science
 - Designed and implemented high-performing propensity models using Python and R to enable informed decision-making and increase enrollment rates in targeted programs
 - Used multidimensional scaling, correspondence analysis, and clustering to provide product owners a deeper understanding of the data, which fostered data-driven discussions
- Experimentation
 - Spearheaded an experimental design framework for digital and non-digital marketing initiatives using frequentist hypothesis-testing methods, a Shiny app to help perform random sampling ([example](#)), which reduced the workload from weeks to minutes, and built Quarto templates for binomial response variables using categorical data analysis techniques ([example](#))

Dallas, TX

- Leadership
 - Led several efforts to create a collection of standards known as the Standards Universe
 - SQL standards to standardize queries and to ease the burden of peer reviews
 - MLOps standards to ensure models taken to production would be packaged, tracked, and registered similarly; also wrote a Python package to facilitate tracking
 - Data science standards to ensure projects are structured similarly and reproducible
 - Created a document to highlight the various capabilities of the data science team ([link](#))

Signify Health
Dallas, TX

Data Scientist (Go-To-Market)
 Jan 2020 – Feb 2021

My impact in this role was to create a COVID-19 forecast model to help senior leadership decide when to re-open corporate offices and when to send physicians back into member’s homes.

- Created a program (R) to model COVID-19 infections using the SIR model of epidemiology, cubic splines with a machine learning optimization procedure, and Facebook’s Prophet model ([app](#))
- Used hierarchical clustering to group clients into small, medium, large groups based on client population size, used bootstrapping to create a benchmark by cluster for various KPIs, and used multidimensional scaling to holistically visualize the portfolio of clients. Created a dashboard that was shared with the executive leadership team, sales, and account managers (R, SQL, Power BI)

Toyota North America
Plano, TX

Senior Data Analyst (Risk)
 Oct 2018 – Oct 2019

My impact in this role was to lead a back-end team that modernized internal financial reporting.

Southwest Airlines
Dallas, TX

Associate Data Scientist (Marketing)
 Sep 2016 – Oct 2018

My impact in this role was to create a modern account identification process and to train new analysts. Recognized as Marketing's employee of the year 2017.

Education & Certificates

- Texas A&M University: Master of Science: Statistics Dec 2021
- Florida State University: Bachelor of Science: Finance May 2014

Projects, Publications, & Presentations

- Primers on functional data analysis [2021](#), [2022](#)
- The Power of Functional Data Analysis [2023](#)
- Internet Outage Detection Using Julia [2023](#)
- Texas Parks & Wildlife Web App [2024](#) [code](#)
- Arxiv Package Analysis [2024](#)
- Functional data analysis blog [2024](#)

Technical Competencies

Analytic		Data		Visualization	Other	
Alteryx	Julia	APIs	JSON	Power BI	AWS	Linux, shell
Excel, VBA	Python	Arrow	Parquet	Shiny	Agile, Scrum	Quarto
Databricks	PySpark	CSV	SQL	Tableau	Git	VS Code
	R	Feather			Jupyter Labs	YAML