I am writing to express my interest in opportunities that align with my background in data science, statistics, and quantitative modeling. I am a statistician and data scientist with a Master of Science in Statistics and Data Science from the University of Wisconsin-Madison, where I also completed my Bachelor of Science in Mathematics and Statistics.

During my graduate studies, I explored advanced statistical methodologies such as Bayesian inference, statistical learning, and stochastic processes. These were complemented by coursework in randomized linear algorithms, financial statistics, advanced statistical computing, and data-driven dynamical systems. I developed expertise in designing scalable models for high-dimensional data, balancing theoretical development with computational efficiency. My undergraduate training, deeply rooted in proof-based mathematics, included graduate-level coursework in analysis, algebra, and topology, alongside statistical experiment design, regression analysis, and probability theory.

In Python, I am proficient in both Object-Oriented and Functional Programming paradigms, enabling the development of modular, scalable, and efficient solutions. I have extensive experience with libraries such as NumPy, Pandas, SciPy, PyTorch, and TensorFlow for implementing machine learning and deep learning models. These include supervised models like support vector machines, decision trees, and ensemble methods, as well as unsupervised techniques such as clustering, dimensionality reduction, and generative modeling. My expertise extends to advanced neural network architectures, including transformers for NLP tasks, convolutional networks for computer vision, and recurrent networks for sequence modeling.

In R, I have conducted comprehensive statistical modeling and evaluation, employing methods like linear, generalized linear, mixed-effects, and quasi-likelihood models. These have been instrumental in deriving actionable insights from structured and unstructured datasets. Additionally, I am well-versed in algorithm design, particularly in numerical optimization and symbolic computation, leveraging tools such as SageMath, Mathematica, and JAGS for both analytical and computational tasks. This technical breadth has enabled me to develop end-to-end pipelines, from raw data preprocessing to building interpretable, high-performance models.

I collaborated with research scientist at Columbia University and independent statistician from cirrus logic to develop a generative model for inferring error characteristics in binary classification systems without ground truth. I contributed to designing the model's topology, identifying redundancies, and addressing challenges related to conditional independence and heterogeneity. Utilizing symbolic computation tools, I evaluated complex multivariate derivatives, while implementing numerical optimization techniques such as BFGS and Nelder-Mead to approximate maximum likelihood estimators. I validated the model through extensive MCMC simulations using JAGS in R, executed on UW-Madison's high-performance computing infrastructure. This work was presented at the Directed Mathematics Research Program, receiving critical feedback that has guided ongoing exploration into Latent Structure Analysis with deterministic constraints as surrogates for ground truth.

In another impactful project, I partnered with Azim Premji University to support the Center for Sustainable Employment's mission of fostering inclusive job creation. Over six months, I led a coding-intensive project to scrape and process approximately 11 million rows of data from the Ministry of Skill Development and Entrepreneurship's website. I used Python's Selenium library for dynamic web scraping and leveraged Google Cloud Virtual Machines for multi-threaded, parallelized data collection. This large-scale project honed my expertise in automation, cloud computing, and handling complex datasets, delivering actionable insights into labor politics and workforce inclusion.

My professional experience includes a Quantitative Risk Internship at Nomura, where I designed and implemented a non-parametric staleness detection algorithm for time-series data. Addressing inefficiencies in the bank's existing system, I developed a dynamic threshold model that adjusted based on the weighted distribution of historical stale periods. This algorithm reduced false positives by over 90%, ensuring more reliable monitoring of diverse financial instruments, including bonds, credit swaps, and equities. This role allowed me to integrate domain-specific knowledge with statistical expertise, delivering a scalable, high-impact solution for time-series data management.

I would welcome the opportunity to discuss how my expertise and experiences align with your organization's goals. Please feel free to contact me at shrivats.sudhir@gmail.com or +1 (908)-699-0054. I look forward to the possibility of contributing to your team and driving meaningful impact.