main sop draft

Statement of Purpose - PhD in Data Science (Mathematics Department)

During my undergraduate studies in mathematics, I often wondered how these mathematical concepts could be applied in the real world. Sitting in my linear algebra, probability ,and statistics classes, I would ask myself: how could all of this be helpful to solve practical problems beyond textbooks? In my second last year in university, with this curiosity, I looked for and explored various fields at the intersection of mathematics and computation, and got to know about data science. A discipline that combines mathematical reasoning, coding ,and data to solve real-world problems. This realization marked a turning point in my academic journey and showed me a path on how my mathematical interests could lead to applied problem-solving.

While a few programming languages, including C++ and MATLAB, were being taught in my course work during my degree. I started learning Python on my own and found it as my strong interest. Now, to learn more and hear from the other side on how leaders in data science are applying it in the real world, I looked for and attended various seminars and conferences in the city focused on machine learning and data science. This motivated me to learn more, and I found a great connection. This early exploration helped me see data science as not different from mathematics but as a natural extension of it.

During my bachelor's degree in mathematics, I established a solid foundation in Algebra, Calculus, Probability, and Statistics. I performed strongly, earning GPAs as high as 3.9 in several semesters. There was only one semester during my whole degree where I dropped my GPA to less than 3 as it was post-COVID and a big transition time. But I quickly got the speed back again and leveled up to end my degree with a 3.5 Cumulative GPA out of 4.0. Around this last semester, while our degree had all core maths subjects, a transformative experience was to tutor my peers to help them break down complex concepts. This act solidified my own understanding alongside my communication and analytical skills, which helped me to grow later as I moved into applied data science.

In my final semester, I was looking forward to joining any learning environment where I could get structured training to learn more about data science in depth. At the same time, via joining Data Science-related events, I got to know about a data science and AI edtech startup giving training via structured 6 months bootcamps to university students and working professionals who were from STEM backgrounds. This felt like a solid opportunity, and on reaching out to the founder to learn more about how it could be helpful for me and how it works. He recognized my strong interest and potential and offered me a merit scholarship for a six-month rigorous bootcamp. While managing my final semester and exams, I attended live nightly classes, balancing boot camp alongside my studies. The duration was a high-focus time for me but was super transformative. Linear algebra and statistics, once abstract and I used to wonder about during my classes, were now tools I applied to real datasets, reinforcing my love for both theory and practice.

Right after my graduation and while being in my second month in the bootcamp, I got done with my data analytics bootcamp covering Exploratory Data Analysis with Excel, Data Manipulation using SQL, visualization tools, and Python libraries such as Numpy, Pandas, and Matplotlib. This basic knowledge quickly helped me to secure an internship as a data analyst, where I used these learnings to help in the decision-making for stakeholders using their data. And right after my internship, I got to join the same organization as a data analyst and teaching assistant, from where I once got training to learn Data Science.

This job felt super close to me as I directly got to get into an environment where their core was Data Science. I got to learn a lot from this role. I managed a cohort of 50 students directly while I used to arrange 2-hour sessions on a weekly basis to cover what experts had done during the whole week and break down the complex problems students were facing in clear form. This cohort mainly consists of professionals transitioning to data science and undergraduate students who all come from STEM backgrounds, aged between 17 to 40. And later I have been involved as a trainer to teach SQL, Python for Data Science and visualization tools like Tableau where I am still involved as an on-demand trainer. This teaching experience helped me to make my base more stronger in the field and gave me fulfillment in helping others navigate in the same domain.

My professional experience got me numerous promotions, and I got an employee of the quarter award in recognition of my efforts, while in Q2'2025 the position for my manager's position became vacant, and I got to manage various projects independently. And during the same journey, I found the urge to give back to my community those who follow the data science path, and grow with it, and lead fellowships as well, voluntarily, to help professionals and students grow.

While being involved myself in industry to learn from peers out there, I joined local and international teams and challenged myself in various Hackathons, and competitions to test my data science and coding capabilities. I got to participate in three coding competitions. The first one was the META Hacker Cup, I could proceed only till the practice round. Right after this next month, I also got myself challenged in the Calico Competition organized by UC Berkeley and our team ranked as 128th out of 500+ teams. For the third time in the upcoming month, this competition was organized by MIT students named as an Informatics Tournament (M(IT)^2)and our team ranked 3rd position out of 300+ teams. All of these competitions strictly banned AI use which made the competition tough, but joining these competitions each month from not reaching to the next phase to scoring 3rd position globally showed the true power of consistency. The Hackathons I participated in included Hact-nation by MIT, NASA Space Apps Challenges. In the later one, our team had our team nominated globally by judges where we used the Machine Learning Model to build a product to test live Air Quality index, built in just under 48 hours.

While my previous experience in applied data science allowed me to extract actionable insights from structured datasets using statistical and machine learning methods, I realized these approaches often overlook the intrinsic shape and connectivity in complex data. Topological Data Analysis (TDA) provides a framework to capture these global structures, revealing patterns

that traditional methods may miss. By combining TDA with my mathematical background, I aim to develop methods that are both theoretically grounded and practically insightful.

I am applying to the PhD program in Data Science at the Mathematics Department of X because it provides the perfect environment to advance my understanding of mathematical foundations while engaging in cutting-edge research The Mathematical Sciences department provides a unique environment where I can explore Topological Data Analysis (TDA) alongside applications in machine learning, neuroscience, and computational biology. Working with Professor Y, whose research bridges TDA, theoretical neuroscience, and applied category theory, aligns perfectly with my goal of developing mathematically principled methods to extract structure from complex datasets. The interdisciplinary culture, active seminars, and collaborative research community at X will provide the ideal setting to advance my research, while allowing me to contribute my experience in applied data science and mentoring to the academic community.

I am particularly drawn to the opportunity to work with a diverse, collaborative community and gain deeper expertise in topological data analysis and machine learning. My goal is to contribute to meaningful research while continuing to mentor and teach, integrating my love for mathematics with practical data science applications.

Sop - draft01

-When I was studying mathematics, I often wondered how these concepts were applied in real life. After classes, I would spend time researching how mathematics connects to the world beyond textbooks, and that is when I came across data science. I was fascinated by how data and statistics could be used to solve real problems. This curiosity led me to start learning the Python programming language on my own. I enjoyed combining coding with mathematics, but I wanted to understand how coding could be applied in practical and impactful ways.

This growing interest encouraged me to attend various events and conferences related to data science. At one such event, I met the founder of Atomcamp, a leading data science and AI EdTech platform. He noticed my curiosity and engagement and offered me a merit-based scholarship for their six-month data science bootcamp. The program was rigorous, especially as I was completing my undergraduate degree at the same time, but it allowed me to connect mathematics, data, and programming through real-world projects.

Learning from the bootcamp helped me understand the core concepts of data science, starting from data cleaning to data analysis, data extraction, and manipulation using tools like SQL and Python. This eventually led me to core concepts like machine learning, NLP (Natural Language Processing), computer vision, and LLMs (Large Language Models). 2 months after the bootcamp, I got my first internship as a data analyst. Right before my internship ended, seeing my enthusiasm for data science, I got hired by the same company where I had taken the bootcamp as a trainee. Now, I work as a teaching assistant, helping train other trainees from diverse backgrounds. I assist them in learning core concepts and troubleshooting problems. This teaching experience has opened my eyes to how much I enjoy helping students.

During my time as a teaching assistant, I also had the opportunity to work on a natural language processing (NLP) project focused on detecting human emotions. In this project, I applied my Python skills to scrape large datasets from multiple sources like Facebook, X, YouTube, Reddit, and editorial sites using libraries like BeautifulSoup, Selenium, and Scrapy. I then performed sentiment analysis to evaluate textual data. This experience helped me better understand how theoretical concepts translate into applied machine learning workflows.

After completing the bootcamp, and later as a teaching assistant at atomcamp. In these roles, I enjoyed helping students refine their understanding of data science concepts and apply what they had learned in practice. While these experiences strengthened my applied skills, they also made me aware of questions I could not fully address independently, particularly related to building better models, conducting deeper analyses, and using data more rigorously to inform decisions. While my professional roles strengthened my applied skills, they also revealed methodological and theoretical questions that I could not resolve independently. Having explored industry, teaching, and research-adjacent work, I am now looking for a formal doctoral

training to rigorously investigate these challenges. mathematical foundation with advanced technical skills and to engage more deeply in research. This motivation drives my goal of studying data science at the doctoral level and contributing to meaningful, real-world decision-making through data-driven research.

Sop - draft00

When I was studying mathematics, I often wondered how these concepts are actually used in real life. After classes, I would spend time exploring how mathematics connects to the world beyond textbooks, and that's when I discovered data science. I was fascinated by how data and statistics could be used to solve real problems, and I wanted to understand how coding could make these applications practical and impactful. Motivated by this curiosity, I began learning Python on my own, combining my mathematical knowledge with programming to explore problem-solving in new ways.

To deepen my understanding, I attended various events and conferences related to data science. At one such event, I met the founder of Atomcamp, a leading data science and AI EdTech platform. He noticed my curiosity and engagement and offered me a merit-based scholarship for their six-month data science bootcamp. The program was rigorous, especially as I was completing my undergraduate degree at the same time, but it provided an incredible opportunity to connect mathematics, data, and programming through real-world projects.

During the bootcamp, I gained hands-on experience with the full data science workflow: data cleaning, data extraction, data analysis, and manipulation using SQL and Python. I gradually explored advanced topics including machine learning, NLP (Natural Language Processing), computer vision, and large language models. Two months after the bootcamp, I secured my first internship as a data analyst. Impressed by my enthusiasm, the same company later hired me and I now work as a teaching assistant, helping train other trainees from diverse backgrounds. Assisting students with core concepts and troubleshooting problems has shown me how much I enjoy teaching and mentoring skills I know will also support my growth as a researcher.

While working as a teaching assistant, I had the chance to lead an NLP project focused on detecting human emotions. I applied my Python skills to scrape large datasets from Facebook, X, YouTube, Reddit, and editorial sites using libraries like BeautifulSoup, Selenium, and Scrapy. I then performed sentiment analysis to evaluate textual data. This project not only strengthened my coding and analytical skills but also showed me how theoretical concepts translate into applied machine learning workflows.

These experiences clarified that while I had developed strong applied skills, there were deeper methodological and theoretical questions I could not fully address independently especially regarding model building, rigorous analyses, and using data to inform real-world decisions. Having explored industry, teaching, and research-adjacent work, I realized the need for formal doctoral training to develop a stronger mathematical foundation, advanced technical skills, and research expertise.

I am particularly drawn to **X's PhD in Data Science** because of its focus on combining theoretical rigor with applied research, and its collaborative research environment. Working with **Dr. Y** whose research in Topological Data Analysis, Machine Learning, and applied category theory aligns closely with my interests, presents an incredible opportunity to explore advanced methodologies and tackle complex problems. His mentorship would allow me to contribute to impactful projects that merge rigorous analytics with practical relevance.

X's PhD program provides the ideal environment to develop my expertise in data science while contributing to meaningful research. I am highly motivated to utilize my background in mathematics, coding, and applied data science to advance methodologically, develop innovative solutions and support data-driven decision-making in real-world contexts.

copy of main draft after comments

Statement of Purpose - PhD in Data Science (Mathematics Department)

When I was in my undergrad mathematics, I often wondered how these mathematical concepts could be applied in the real world. Sitting in my linear algebra and statistics classes, I would ask myself: how can they solve real problems? After my maths classes in 5th semester i came back and google all fields connected with mathemtaisc and this wa sthe tiem i a=came to know about data science, how its the cWith this curiosity led me to explore data science. I began attending tech events and conferences in Islamabad focused on machine learning and data analytics, and started learning Python on my own. Combining coding with mathematics felt natural, and I became fascinated by the possibility of using mathematical theory to extract meaningful insights from data.

Mathematics has always been my favorite subject. In high school, I loved spending hours solving problems and discovered I excelled at it, scoring 94/100 in my FSC Part II. My teachers often noted my curiosity and focus as a front-bencher, eager to absorb each concept deeply. This enthusiasm naturally led me to pursue a bachelor's degree in mathematics, where I could explore each topic Algebra, Geometry, Trigonometry, Calculus, Statistics, and Probability in greater depth.During my undergraduate studies, I performed strongly across semesters, earning GPAs as high as 3.9 in several terms, and graduated with a CGPA of 3.5. Over time, I realized I was most drawn to applied mathematics and its practical uses

One of the most transformative experiences during my undergrad was tutoring my peers. Helping classmates understand challenging courses in later semesters allowed me to question and solidify my own understanding. I realized I loved teaching and the satisfaction of breaking complex concepts into digestible ideas. This experience also strengthened my analytical and communication skills, which became invaluable as I transitioned into applied data science.

In my final semester, I wanted to explore how mathematics could directly connect to the real world. I attended various events in Islamabad focused on data science and machine learning, including a program by Atomcamp, a leading data science and AI education startup. Inspired, I approached the founder about bridging mathematics with data science. He recognized my potential and offered me a merit scholarship for a six-month rigorous bootcamp. During this time, I simultaneously self-studied Python and attended nightly classes, balancing bootcamp projects with my final semester assignments. The experience was intense but transformative. Linear algebra and statistics, once abstract, were now tools I applied to real datasets, reinforcing my love for both theory and practice.

Shortly after graduation, I secured an internship as a data analyst and was soon offered a full-time role at Atomcamp, where I also became a teaching assistant. Over four months, I led two-hour weekly sessions for 50 trainees, helping them understand complex data science concepts while managing curriculum and project assignments. I also contributed to multiple projects, applying my skills to real-world datasets. My performance earned me several

promotions and recognition, including Employee of the Quarter 2025. To continue giving back, I became a mentor on Topmate, providing 1:1 guidance to data science learners, and volunteered as a fellowship lead for Buildables, supporting early-stage professionals in skill development. Additionally, I challenged myself in competitive settings, participating in four hackathons and three coding competitions, including META, Hackt-nation by MIT, Calico by UC Berkeley, NASA Space Apps Challenges, and MIT Informatics Tournament, where my teams received global recognition and top placements.

While my previous experience in applied data science allowed me to extract actionable insights from structured datasets using statistical and machine learning methods, I realized these approaches often overlook the intrinsic shape and connectivity in complex data. Topological Data Analysis (TDA) provides a framework to capture these global structures, revealing patterns that traditional methods may miss. By combining TDA with my mathematical background, I aim to develop methods that are both theoretically grounded and practically insightful.

I am particularly drawn to the PhD program in Data Science at the X due to its strong emphasis on rigorous mathematical foundations integrated with applied research. The Mathematical Sciences department provides a unique environment where I can explore Topological Data Analysis (TDA) alongside applications in machine learning, neuroscience, and computational biology. Working with Professor Y, whose research bridges TDA, theoretical neuroscience, and applied category theory, aligns perfectly with my goal of developing mathematically principled methods to extract structure from complex datasets. The interdisciplinary culture, active seminars, and collaborative research community at X will provide the ideal setting to advance my research, while allowing me to contribute my experience in applied data science and mentoring to the academic community.

While these professional experiences have strengthened my applied skills, I realized that my core passion lies in mathematics and research. I am particularly excited about Professor Y's work in applied topology and topological data analysis at X. My strong foundation in linear algebra and applied mathematics, combined with hands-on experience in data analytics, positions me to contribute effectively to his research. I am eager to explore how mathematical structures can provide principled insights into complex datasets, and to bridge the gap between theory and real-world applications.

I am applying to the PhD program in Data Science at the Mathematics Department of X because it provides the perfect environment to advance my understanding of mathematical foundations while engaging in cutting-edge research. I am particularly drawn to the opportunity to work with a diverse, collaborative community and gain deeper expertise in topological data analysis and machine learning. My goal is to contribute to meaningful research while continuing to mentor and teach, integrating my love for mathematics with practical data science applications.

I am confident that my academic background, professional experience, and passion for applied mathematics and data science have prepared me to thrive in this program. I look forward to contributing to the research community at X, learning from Professor Y, and developing innovative solutions at the intersection of mathematics and data science.

restructured - gpt version

During my undergraduate studies in mathematics, I became increasingly interested in how abstract mathematical structures could be used to understand real-world phenomena. While courses such as linear algebra, probability, and statistics provided powerful tools, I often found myself questioning how these concepts scale to complex, high-dimensional data encountered in modern applications. This curiosity led me to explore fields at the intersection of mathematics and computation, where I discovered data science as a discipline that unifies mathematical reasoning, algorithmic thinking, and empirical insight.

Motivated to explore this intersection further, I independently learned Python alongside my coursework and began attending machine learning and data science conferences. Through hands-on experimentation, I realized that combining mathematical intuition with programming allowed me to translate theory into practice. What initially drew me to data science was not merely predictive modeling, but the opportunity to formalize patterns in data using principled mathematical tools.

I pursued a bachelor's degree in mathematics, where I developed a strong foundation in algebra, calculus, probability, and statistics, graduating with a CGPA of 3.5 and earning GPAs as high as 3.9 in multiple semesters. Over time, I found myself most drawn to applied mathematics—particularly settings where abstract theory informs practical decision-making.

Alongside my coursework, I frequently tutored peers in advanced mathematics courses, an experience that sharpened my ability to deconstruct complex ideas and communicate them clearly. This early exposure to teaching later became a defining aspect of my professional trajectory and reinforced my desire to pursue an academic research career.

In my final year, I formally transitioned into applied data science through a competitive, merit-based bootcamp at Atomcamp, where I applied linear algebra, statistics, and optimization techniques to real-world datasets. Following graduation, I joined Atomcamp as a data analyst and teaching assistant, leading weekly sessions for over 50 trainees and contributing to multiple applied projects. While this experience strengthened my applied skill set, it also revealed fundamental limitations of conventional statistical and machine learning approaches when dealing with complex, high-dimensional, and non-Euclidean data.

Through my applied work, I began to recognize that many traditional data analysis methods focus on local correlations and feature-level representations, often overlooking the global structure and shape of data. This realization led me to Topological Data Analysis (TDA), a mathematically rigorous framework that captures connectivity, geometry, and multi-scale structure in complex datasets. TDA

resonated deeply with my mathematical background, as it provides tools—such as persistent homology—that preserve structural information while remaining robust to noise. I am particularly interested in exploring how TDA can complement machine learning models by improving interpretability, stability, and representation learning.

The PhD program in Data Science at X uniquely aligns with my research goals due to its strong integration of mathematical rigor and applied data science. I am especially interested in working with Professor Y, whose research in topological data analysis, theoretical neuroscience, and applied category theory closely matches my interest in uncovering structure in complex systems. The interdisciplinary research environment within the Mathematical Sciences department provides an ideal setting to develop mathematically principled methods with real-world relevance.

My long-term goal is to conduct research at the intersection of mathematics and data science, developing theoretically grounded tools that enhance our understanding of complex data across domains such as neuroscience and biology. I am particularly motivated to pursue an academic research career that combines methodological innovation with teaching and mentorship. I am confident that my background in mathematics, applied data science experience, and growing research focus on TDA have prepared me to thrive in the PhD program at X and to contribute meaningfully to its research community.

Tab 6

During my undergraduate studies in mathematics, I often wondered how these mathematical concepts could be applied in the real world. Sitting in my linear algebra, probability ,and statistics classes, I would ask myself: how could all of this be helpful to solve practical problems beyond textbooks? During my second last year in university, with this curiosity, I looked for and explored various fields at the intersection of mathematics and computation, and got to know about data science. A discipline that combines mathematical reasoning, coding ,and data to solve real-world problems. This realization marked a turning point in my academic journey and showed me a path on how my mathematical interests could lead to applied problem-solving.

While a few programming languages, including C++ and MATLAB, were being taught in my course work during my degree. I started learning Python on my own and found it as my strong interest. Now, to learn more and hear from the other side on how leaders in data science are applying it in the real world, I looked for and attended various seminars and conferences in the city focused on machine learning and data science. This motivated me to learn more, and I found a great connection. This early exploration helped me see data science as not different from mathematics but as a natural extension of it.

During my bachelor's degree in mathematics, I established a solid foundation in Algebra, Calculus, Probability, and Statistics. I performed strongly, earning GPAs as high as 3.9 in several semesters. There was only one semester during my whole degree where I dropped my GPA to less than 3 as it was post-COVID and a big transition time. But I quickly got the speed back again and leveled up to end my degree with a 3.5 Cumulative GPA out of 4.0. Around this last semester, while our degree had all core maths subjects, a transformative experience was to tutor my peers to help them break down complex concepts. This act solidified my own understanding alongside my communication and analytical skills, which helped me to grow later as I moved into applied data science.

In my final semester, I was looking forward to joining any learning environment where I could get structured training to learn more about data science in depth. At the same time, via joining Data Science-related events, I got to know about a data science and AI edtech startup that gives training to undergraduate students or professionals who are coming from STEM backgrounds. This felt like a solid opportunity, and on reaching out to the founder to learn more about how it could be helpful for me and how it works. He recognized my strong interest and potential and offered me a merit scholarship for a six-month rigorous bootcamp. While managing my final semester and exams, I attended live nightly classes, balancing boot camp alongside my studies. The duration was a high-focus time for me but was super transformative. Linear algebra and statistics, once abstract and I used to wonder about during my classes, were now tools I applied to real datasets, reinforcing my love for both theory and practice.

Right after my graduation and while being in my second month in the bootcamp, I got done with my data analytics bootcamp covering Exploratory Data Analysis with Excel, Data Manipulation using SQL, visualization tools, and Python libraries such as Numpy, Pandas, and Matplotlib. This basic knowledge quickly helped me to secure an internship as a data analyst, where I used

these learnings to help in the decision-making for stakeholders using their data. And right after my internship, I got to join the same organization as a data analyst and teaching assistant, from where I once got training to learn Data Science.

This job felt super close to me as I directly got to get into an environment where their core was Data Science. I got to learn a lot from this role. I managed a cohort of 50 students directly while I used to arrange 2-hour sessions on a weekly basis to cover what experts had done during the whole week and break down the complex problems students were facing in clear form. This cohort mainly consists of professionals transitioning to data science and undergraduate students who all come from STEM backgrounds, aged between 17 to 40. And later I have been involved as a trainer to teach SQL, Python for Data Science and visualization tools like Tableau where I am still involved as an on-demand trainer. This teaching experience helped me to make my base more stronger in the field and gave me fulfillment in helping others navigate in the same domain.

My professional experience got me numerous promotions, and I got an employee of the quarter award in recognition of my efforts, why the position for my manager's position became vacant, and I got to manage various projects independently. And during the same journey, I found the urge to give back, to my community to follow the data science path, and grow with it, and lead fellowships as well, voluntarily, to help professionals and students grow.

While being involved myself in industry to learn from peers out there, I joined local and international teams and challenged myself in various Hackathons, and competitions to test my data science and coding capabilities. I got to participate in three coding competitions. The first one was the META Hacker Cup, I could proceed only till the practice round. Right after this next month, I got myself challenged in the Calico Competition organized by UC Berkeley and our team ranked as 128th out of 500+ teams. For the third time in the upcoming month, this competition was organized by MIT students named as an Informatics Tournament (M(IT)^2)and our team ranked 3rd position out of 300+ teams. All of these competitions strictly banned AI use which made the competition tough, but joining these competitions each month from not reaching to the next phase to scoring 3rd position globally showed the true power of consistency. The Hackathons I participated in included Hact-nation by MIT, NASA Space Apps Challenges. In the later one, our team had our team nominated globally by judges where we used the Machine Learning Model to build a product to test live Air Quality index, built in just under 48 hours.