# Carleton College <br> One North College Street <br> Northfield, Minnesota 55057 

March 28, 2023

Dear Provost Mattson and members of the FCPC,

The Department of Mathematics and Statistics requests to increase the number of tenure-track lines in statistics from four to five with the addition of a new tenure-track statistics faculty member who can support our strong statistics curriculum and increase opportunities for students to study data science at Carleton. Currently, many of our peer institutions have committed to formally supporting data science programs, and a tenure line would allow us to support a thoughtful curriculum that would allow us to stay competitive with our peer institutions. If approved, we would conduct the search during the fall of 2023 with September of 2024 as the start date.

## 1. Growth of Statistics and and Data Science Nationally

Statistics education, especially at the introductory level, has been evolving at a fast pace since the early 2000s. The first Guidelines for Assessment and Instruction in Statistics Education (GAISE) report in 2005 and its update in $2015^{1}$ influenced both what and how we teach. These guidelines, which align nicely with many of Carleton's goals for a liberal arts education, emphasize the importance of statistical thinking, conceptual understanding instead of rote calculations, being critical consumers of data, the analysis of real data using computing technology, and data ethics. This evolution of statistics pedagogy coincided with an increase in student interest in statistics with the number of bachelor's degrees awarded in statistics and biostatistics increasing $474 \%$ since $2010^{2}$.

Even more recent has been the rapid growth in Data Science programs. While the initial growth was in master's degree or certificate programs, we are now seeing similar growth in undergraduate degree programs both at universities and liberal arts colleges. The field of data science is so new that it is still evolving as a discipline, but the 2016 Park City Mathematics Institute report provides a good starting point for undergraduate program development ${ }^{3}$. Their curriculum guidelines emphasize statistical, computational, and mathematical thinking with courses drawn from the fields of statistics, computer science, and mathematics.

[^0]We identified seventeen peer institutions that offer majors in data science (sometimes under a different name), twenty-four schools that offer minors in data science, and four schools with related programs (such as certificates) in data science. To create this list, we investigated data science curricula at institutions in our core peer group, the ACM, and the top fifty liberal arts colleges in the US ${ }^{4}$.

## Majors in data science

Claremont-McKenna, Beloit, Coe, Lake Forest, Luther, Knox, Macalester, Mount Holyoke, Wellesley, the US Naval Academy, and the US Airforce Academy all offer majors called "data science." Similar majors were called "data analytics" at Denison, Dickinson, and the University of Richmond, "applied statistics and data science" at the US Military Academy, and "statistical and data sciences" at Smith. It should also be noted that Beloit offers a major in "data analytics" in addition to their major in "data science." There are also two schools that offer rather different majors from the previous list, but are worth mentioning. Bowdoin offers a major in "digital and computational studies," which has a more interdisciplinary emphasis (i.e., a reduced emphasis in statistics and computer science), and Bucknell blends a business and data science degree in their "business analytics" major.

## Minors in data science

Bryn Mawr, Claremont-McKenna, Colby, Davidson, DePauw, Lafayette, Lake Forest, Luther, Macalester, Monmouth, Mount Holyoke, Pitzer, Washington and Lee, and Whitman all have minors called "data science." Similar minors were called "data analytics" at Furman and Union, "data analysis" at Wesleyan, "statistics and data science" at Lawrence and St. Olaf, "statistical and data sciences" at the College of Wooster, and "computer and data sciences" at Ripon. In addition, Bates and Bowdoin offer minors in "digital and computational studies" and Connecticut College offers a minor in "data, information, and society". We have also heard that Middlebury has approved a "data science" minor, but it is not yet in their catalog.

## Related programs

Four schools offer programs related to data science that are not formal degrees. Mount Holyoke has a "nexus" in data science, which required reduced coursework compared to a minor but the capstone is a substantive internship in the field. Wesleyan has an applied data science certificate, the College of Wooster has a "pathway" in data exploration and communication, and the University of Richmond has a data science concentration for computer science majors, a data analytics concentration for business majors, and a general undergraduate certificate in data analytics.

[^1]
## 2. Demand for Statistics and and Data Science at Carleton

The appeal of statistical and data science skills is broad and we are struggling to meet the demand. The effects of this demand are felt not only by the students attempting to register for a statistics or data science course, but also by the statistics faculty. Seeing sections close early in the registration process followed by large waitlists can be frustrating to the faculty teaching these courses, especially when observed term after term. Managing waitlists and fielding emails from eager waitlisted students each term is also time consuming.

For at least a decade, demand for Stat 120 (Introduction to Statistics) has been high, resulting in large waitlists and large class sizes. Of the last 54 sections of this course (since 17-18), 89\% of the sections have had 30 or more students enrolled and 2 sections have had 40 or more students enrolled. In the majority of years, we have had enough staff to offer 8 sections a year but in the last two academic years we were able to offer 10 and 11 sections, respectively, with the addition of a VAP position during Laura Chihara's phased retirement. Even with this increase in sections, we still saw demand that outpaced the number of seats available, with waitlists that could accommodate 1-2 more sections each term. Table 1 shows the waitlist sizes just after registration ends between 18-19 and 21-22.

|  | Table 1: Waitlist for Stat 120 |  |  |  |
| :--- | :---: | :---: | :---: | :---: |
| Year | $18-19$ | $19-20$ | $20-21$ | $21-22$ |
| Number of sections | 8 | 9 | 8 | 10 |
| Number waitlisted after <br> registration | 147 | 191 | 197 | 217 |
| Deficit in sections assuming <br> a class size of 32 | 4.59 | 5.97 | 6.16 | 6.78 |

This high demand combined with priority registration by class year means that, historically, very few first-year students are able to take a statistics course. Figure 1 shows enrollment in Stat 120 by graduating class. The effects of increasing the number of Stat 120 sections the last two years is clear, with a substantial bump in first year students for the graduating classes of 2025 and 2026. The overall total enrollment for these two cohorts of students will increase over the next 2-3 years too, which means that we are bound to surpass the previous maximum enrollment seen in the class of 2022.

Figure 1: enrollment by grad year


One implication in the historical lack of access for first year students in Stat 120 is that it results in a reduced pool of potential statistics majors: those who find their first college-level statistics course compelling their sophomore or junior year may have already declared or started the requirements for a different major. It also means that students are using their statistical and data science skills for a smaller number of Carleton classes if they take their first statistics course later in their Carleton career. Finally, it creates an inequity among students: those who had access to AP statistics curriculum in high school have more pathways into our curriculum. Students with more advanced preparation in statistics their first year could take either Stat 120 or Stat 230 (Applied Regression) as their first statistics course. Students who are further along in the calculus sequence might opt to take Math 240 (Probability) and then take Stat 250 (Introduction to Statistical Inference) as their first statistics course. One of the College's DEI goals is to "[e]xamine disparities in curriculum structure that hinder student progression through major/minor requirements (3.1.8)" and our lack of sufficient staffing in statistics to meet the demand of the large and diverse body of students wanting to take our courses has created a barrier to entry for those with less advanced preparation.

The interdisciplinary nature of statistics and data science means that our curriculum serves not only our majors, but majors and departments across campus, even beyond STEM fields. Majors from other departments are interested in courses beyond Stat 120, and we are seeing demand from non-statistics majors in both our intermediate level courses (including Stat 220 Intro to Data Science, Stat 230 Applied Regression, and Math 240 Probability) and our electives
courses (including sampling, advanced modeling, time series analysis, and Bayesian statistics). Since 2017-18, the average class size was 25.4 for Applied Regression (18 sections), 26.2 for Probability ( 19 sections) and 27 for Intro to Data Science ( 12 sections). The majority of students taking these classes were not statistics majors. Table 2 demonstrates this with data from the graduating classes of 2023 and 2024. Roughly $60 \%$ of juniors and seniors who took Stat 120 were from majors that are not STEM majors. Just over half of these students are social science majors that require Stat 120 like Political Science ( 81 majors), Economics ( 57 majors), and SOAN ( 37 majors). The rest come from a variety of majors from across the humanities, arts and languages like History ( 17 majors), CAMS (12 majors), and Philosophy (11 majors), and English ( 10 majors). Of juniors and seniors who took Intro to Data Science, about $25 \%$ were non-STEM majors and of those two took Applied Regression, about 28\% were non-STEM majors.

|  | Table 2: Number of juniors and seniors who took Stat 120, <br> 220 and/or 230 |  |  |
| :--- | :---: | :---: | :---: |
|  | Stat 120 (Intro to <br> Statistics) | Stat 220 (Intro to <br> Data Science) | Stat 230 (Applied <br> Regression Analysis) |
| Major area | 34 | 41 | 38 |
| Math or Stat | 185 | 39 | 38 |
| STEM (not math or stat) | 320 | 27 | 30 |
| Not STEM |  |  |  |

The number of statistics majors has grown from 3 math/stat majors graduating in 2012 to 21 statistics majors graduating in 2023. Carleton ranks 75th nationally across all colleges and universities in the total number of statistics (or statistics track) bachelors degrees graduated between 2003-2020 and is one of the top liberal arts colleges in statistics degrees granted ${ }^{6}$. With this many majors, it is important to offer a consistent rotation of electives that allow for some flexibility in scheduling. We currently are only able to offer four unique elective courses per year, which is relatively small, and we have been able to accomplish this the last few years with the addition of more visitors teaching lower-level course; however, this option is not sustainable and limits the possibility of offering new statistics or data sciences courses (at any level) in the future. See Appendix A for our three-year staffing plan with only four tenure-track/tenured statisticians.

We also face pressure in statistics comps, which is either a group comps experience run by a faculty advisor or an independent reading comps. As is usually done on the math side, we try to advise a comps group two out of every three years, but our majors overwhelmingly prefer group comps. This results in senior statistics faculty consistently advising groups of 5-6 students and students being assigned to complete an independent comps project when they would have

[^2]preferred a group comps experience. These results are at odds with what we observed for math comps, which see group sizes of 3-4 students (except for Rob Thompson's applied projects) and fewer students (proportionally) being assigned their last comps preference (independent comps).

## 3. The need for a new tenure-track line in Statistics

With our current staffing of four full-time statisticians, assuming faculty take "normal" sabbatical leaves (e.g., 15 courses taught every 10 terms), we can expect each faculty member to teach, on average, 4.5 courses per year for an average of 18 courses covered by all full-time faculty in a year. With five full-time faculty, the yearly average would be 22.5. In actuality, the number of courses covered by full-time faculty will be somewhere "around" these averages depending on course releases and sabbatical replacement approvals. Table 3 shows that our four faculty members taught 19 courses in 17-18, but in 19-20 only 12 courses were covered by full-time faculty. Of the remaining 10 courses in 19-20, 5 were taught by a visiting statistician and the other 5 were taught by a combination of 3 visiting or adjunct faculty members (Christine Kohnen, Owen Biesel, Josh Davis). The 7 courses not taught by full-time faculty in 20-21 were taught by a combination of 5 visiting or adjunct faculty (Owen Biesel, Steve Scheirer, Sam Ihlenfeldt, Josh Davis, AI Garver).

|  | Table 3: Number of courses by year |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Year | $17-18$ | $18-19$ | $19-20$ | $20-21$ | $21-22$ | $22-23$ |
| No. of STAT <br> courses + Math 240 | 20 | 20 | 22 | 23 | 25 | 26 |
| No. taught by <br> visitors/adjuncts | 1 | 1 | 10 | 7 | 8 | 5 |
| Percentage taught <br> by visitors/adjuncts | $5.0 \%$ | $5.0 \%$ | $45.5 \%$ | $30.4 \%$ | $32.0 \%$ | $19.2 \%$ |

At current staffing levels (four statisticians), our increase in statistics courses since 18-19 has been supported mainly by asking visitors and adjuncts to teach more of our courses, primarily at the introductory level. While this can work out well, entrusting the majority ${ }^{7}$ of our introductory classes in any given year to visitors, some of whom do not have PhDs in Statistics, has its risks both to our department and client departments that require (or suggest) this course for their major. The costs, too, to full-time faculty in supporting this many visitors in one year should not be overlooked.

There is interest among the statistics faculty to strengthen Carleton's offerings in data science; however, this would reduce our offerings in statistics, which would adversely impact statistics majors and many students throughout the college. Statistics is a core component of data

[^3]science education and a robust offering of courses in statistics is needed to support the development of a data science program. We believe that we can start to expand our data science curriculum at Carleton with an additional tenure line, without adversely impacting our statistics curriculum. This would help our curriculum remain competitive with our peer group. It should be noted that expanding to five statisticians is not unreasonable; among several other examples from our peer institutions, Macalester has seven full-time statisticians who contribute to their statistics and data science programs, Smith has five tenure-track statistics lines and four joint appointments for their statistical and data sciences program, and Williams has seven tenure-track statistics lines to support their statistics program.

Finally, the challenges of hiring visitors in statistics should also not be ignored. PhD statisticians and data scientists have many career options besides academia, often with better compensation. In 2015, the American Statistical Association survey of business, industry, and government statisticians showed a median annual salary of about $\$ 100,000$ for statisticians with $0-5$ years experience across all levels of statistics education ${ }^{8}$. The pool of applicants that we often see for visiting positions, especially one year positions, can be extremely varied and in many years very weak. During our last two searches for full-time visitors in statistics (conducted in 20-21 and 22-23), we had 11 applicants for each search. Other departments or majors, including mathematics, might draw upon past graduates who are ABD at the University of Minnesota or recent PhD graduates who are looking to gain experience teaching at a liberal arts school. The newness of our major means this pool of past majors, while growing, is still limited.

## 4. Supporting the liberal arts at Carleton

The fields of statistics and data science equip students with the tools to think critically and communicate effectively with data from across disciplines; a philosophy that is at the heart of Carleton's liberal arts education and supports our strategic planning theme of Advancing the Liberal Arts. An additional faculty member will increase our capacity to attract and support students who wish to major in statistics. Just as important, it will also allow us to expand our support for students in other majors, including those outside STEM fields, and provide them with the statistical tools to engage in data-driven problem solving and research. These student-centered benefits would also expand to faculty: an additional faculty member would increase our capacity to collaborate with and support faculty across campus who are working on data-centric projects.

For example, we have been collaborating with Lin Winton, director of the Quantitative Resource Center, to help define and organize support for students analyzing data and using the statistical software R in courses across departments. For the last three years, Lin has hired 1-2 statistics majors a year to serve as a "comps stats consultant" to help majors from across campus who were working on comps projects (or other research) that involved data analysis. These majors are hired on the advice of the statistics faculty and at times, statistics faculty have supported

[^4]some comps consultants with their work, which can involve understanding statistical methodology that is not part of our statistics curriculum. The work of these consultants has been well-received. In addition to student support, there is also interest among faculty from other departments who would like to learn R and collaborate on data-centered projects. An additional faculty member would enhance our capacity to provide this type of support.

We also recently proposed a minor in Statistics and Data Science which we believe would be of interest to a wide variety of students, including those with little prior experience in statistics, computer science or mathematics. We've intentionally designed the minor to have a low barrier to entry; requiring 3 statistics courses and one computer science course, and notably not requiring calculus. The remaining three courses are selected from a variety of courses from across the College that ask students to think critically with and about data. While we believe that we can support this minor at our current staffing levels, adding an additional statistics faculty member would only enhance our ability to support the minor. At current staffing levels, a data science major is not possible.

## 5. Details of Proposed Hire

Our proposal is to conduct our search during the fall of 2023 with a start date of September 2024. While we would prefer candidates with PhDs in statistics or biostatistics, we would also consider applicants with PhD in related fields ${ }^{9}$ who have a strong background in statistics and data science, and whose work has a significant focus on computation and data. We would like the faculty filling this new position to have enough depth in statistics to support our current statistics curriculum but also help us expand our data science offerings and increase our capacity to collaborate with faculty in other disciplines across campus.

We will advertise our position broadly, including including to groups like the American Statistical Association's Committee on Minorities in Statistics, the National Alliance of Mathematicians (a national organization dedicated to supporting underrepresented minorities in the mathematical sciences), Women in Machine Learning and Data Science, and local and national R-Ladies groups. Over the summer of 2023, we will do outreach by connecting with graduate students and new PhDs from underrepresented groups at the Joint Statistical Meetings and the US Conference on Teaching Statistics, and via outreach to PhD programs across the US. We understand that this effort is needed to ensure a diverse pool of candidates based on the demographics of statistics PhDs.

Under our proposed timeline, both Andy Poppick and Adam Loy will be newly tenured and will join Katie St. Clair as senior faculty in statistics. While Claire Kelling is still a relatively young junior faculty, we believe that we can effectively mentor a new statistics faculty member. We are also cognizant of the fact that effective faculty mentoring is a College DEI goal and our department has a strong commitment to mentoring with a DEI lens. In our department, each

[^5]new tenure-track faculty member is assigned a tenured department member from their field (math or statistics) as their mentor at the start of the academic year. Mentoring within our department usually takes the form of conversations about teaching and research, class visits with follow-up discussions, and connecting mentees with teaching resources and others in the statistics education community. Our department also has regularly scheduled pedagogy seminars that all department members can attend, where we share new ideas and best practices with each other about topics of common interest. We have a weekly group called the "Circle of Least Confusion" that meets to work together and learn about each other's research. The department meets weekly as a whole group for official business and has a culture of valuing junior faculty contributions in department decision making. We are a large department with many service needs within the department, but are careful not to overburden junior faculty with service requirements and are developing new ways to account for hidden work that junior faculty, and particularly our BIPOC and female junior faculty, do to support students in the department. Our department is relatively young, and there is also plenty of informal mentoring and support available for a new hire from our active and diverse cohort of junior math and stat faculty.

Thank you for considering this request. Please let me know if you have any questions.

Sincerely,


Katie St. Clair
Professor of Statistics
Chair, Department of Mathematics and Statistics

## Appendix A: Three-year staffing plan

## Tenure-track position approved

The first three-year plans assume a tenure-track hire in statistics and all statistical faculty teach 5 courses per year. There would need to be some modifications in this plan due to sabbaticals (e.g. Claire Kelling will likely be on leave during her fourth year, which could be year 3 in this scenario).

With additional staffing, we would be able to continue offering 10 Stat 120 (Intro Stats) sections per year, 3 Stat 220 (Data Science) and 3 additional electives (one 200-level and two 300-level) per year. We would also add a new 200-level Statistical Learning course that would be accessible to students with minimal statistics background. This course would be of interest to students interested in both statistics and data science.

For reference: AL is Adam Loy, AP is Andy Poppick, CK is Claire Kelling and KSC is Katie St Clair.

|  | Year 1 - assumes 5 faculty @ 5 courses each |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics |  | $\begin{aligned} & \text { TT (2), } \\ & \text { KSC } \end{aligned}$ | 3 | AP, CK (2) | 4 | $\begin{aligned} & \text { TT, AP, AL, } \\ & \text { KSC } \end{aligned}$ | 10 |
| Stat 220 Data Science | 1 | CK | 1 | TT | 1 | TT | 3 |
| Stat 230 Applied Regression | 1 | CK | 1 | AL | 1 | AP | 3 |
| Math 240 Probability | 2 | AL (2) | 1 | KSC |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | AP |  | KSC | 2 |
| Stat 260 Sampling Techniques |  |  | 1 | KSC |  |  | 1 |
| Stat 285 Stat Consulting | 0.34 | AL | 0.33 | AL | 0.33 | AL | 1 |
| Stat 310 Spatial Statistics |  |  |  |  |  | CK | 1 |
| Stat 320 Time Series | 1 | AP |  |  |  |  | 1 |
| Stat 330 Advanced Modeling |  |  |  |  |  |  | 0 |
| Stat 340 Bayesian Inference |  |  |  |  |  |  | 0 |
| Stat 2xx Statistical Learning* |  |  |  |  |  |  | 0 |
|  | 8.34 |  | 8.33 |  | 8.33 |  | 25 |


|  | Year 2 - assumes 5 faculty @ 5 courses each |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics |  | CK, KSC <br> (2) | 3 | AP, CK, TT | 4 | TT, AP, AL <br> (2) | 10 |
| Stat 220 Data Science | 1 | AL | 1 | TT | 1 | KSC | 3 |
| Stat 230 Applied Regression | 1 | CK | 1 | KSC | 1 | TT | 3 |
| Math 240 Probability | 2 | AP (2) | 1 | CK |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | AL | 1 | CK | 2 |
| Stat 260 Sampling Techniques |  |  |  |  |  |  | 0 |
| Stat 285 Stat Consulting | 0.34 | AL | 0.33 | AL | 0.33 | AL | 1 |
| Stat 310 Spatial Statistics |  |  |  |  |  |  | 0 |
| Stat 320 Time Series |  |  |  |  |  |  | 0 |
| Stat 330 Advanced Modeling |  |  |  |  | 1 | AP | 1 |
| Stat 340 Bayesian Inference |  |  | 1 | KSC |  |  | 1 |
| Stat 2xx Statistical Learning* | 1 | TT |  |  |  |  | 1 |
|  | 8.34 |  | 8.33 |  | 8.33 |  | 25 |


|  | Year 3 - assumes 5 faculty @ 5 courses each |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics | 3 | AL, CK (2) |  | AP, AL, TT | 4 | $\begin{aligned} & \text { AP, TT, } \\ & \text { KSC (2) } \end{aligned}$ | 10 |
| Stat 220 Data Science | 1 | AL | 1 | CK | 1 | CK | 3 |
| Stat 230 Applied Regression | 1 | AP | 1 | KSC | 1 | AP | 3 |
| Math 240 Probability | 2 | TT (2) | 1 | AL |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | TT | 1 | CK | 2 |
| Stat 260 Sampling Techniques | 1 | KSC |  |  |  |  | 1 |
| Stat 285 Stat Consulting | 0.34 | KSC | 0.33 | KSC | 0.33 | KSC | 1 |
| Stat 310 Spatial Statistics |  |  |  |  |  |  | 0 |
| Stat 320 Time Series |  |  | 1 | AP |  |  | 1 |
| Stat 330 Advanced Modeling |  |  |  |  | 1 | AL | 1 |
| Stat 340 Bayesian Inference |  |  |  |  |  |  | 0 |
| Stat 2xx Statistical Learning* |  |  |  |  |  |  | 0 |
|  | 8.34 |  | 8.33 |  | 8.33 |  | 25 |

## Tenure-track position not approved

The second three-year plans assume no tenure-track hire in statistics and four statistics faculty teaching 5 courses per year. It also assumes that we do not add a full-time visitor to boost our course count to (roughly) 25, as it has been in the last two years with the addition of VAP Deepak Bastola. We do pull in one Math faculty, possibly Josh Davis, to teach one section of Math 240 (Probability) each year (which would reduce the number of staffed Math courses by one). Again, there would need to be some modifications in this plan due to sabbaticals.

Without additional staffing, we would reduce the number of Stat 120 (Intro Stats) sections to 8 per year, Stat 220 (Data Science) to 2 per year and we would only be able to staff 2 additional electives per year. We would not be able to develop and offer a new Statistical Learning course without elimination of one or more of our current electives. We also could not adequately staff the 5 electives that we currently offer every 1-2 years, and some of these (TBD) would need to be offered on a more sporadic basis.

|  | Year 1 - assumes 4 faculty @ 5 courses each + 1 part-time (PT) faculty (1 course) |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics | 2 | AL, KSC | 3 | AP, CK (2) | 3 | AP, AL, KSC | 8 |
| Stat 220 Data Science | 1 | CK |  |  | 1 | CK | 2 |
| Stat 230 Applied Regression | 1 | CK | 1 | AL | 1 | AP | 3 |
| Math 240 Probability | 2 | AL, PT | 1 | KSC |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | AP | 1 | KSC | 2 |
| Stat 260 Sampling Techniques |  |  | 1 | KSC |  |  | 1 |
| Stat 285 Stat Consulting | 0.34 | AL | 0.33 | AL | 0.33 | AL | 1 |
| Stat 310 Spatial Statistics |  |  |  |  |  |  | 0 |
| Stat 320 Time Series | 1 | AP |  |  |  |  | 1 |
| Stat 330 Advanced Modeling |  |  |  |  |  |  | 0 |
| Stat 340 Bayesian Inference |  |  |  |  |  |  | 0 |
| Stat $2 \times x$ Statistical Learning* |  |  |  |  |  |  | 0 |
|  | 7.34 |  | 7.33 |  | 6.33 |  | 21 |


|  | Year 2 - assumes 4 faculty @ 5 courses each + 1 part-time (PT) faculty (1 course) |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics | 2 | CK, AL | 3 | CK, KSC <br> (2) | 3 | AP (2), AL | 8 |
| Stat 220 Data Science | 1 | KSC |  |  | 1 | KSC | 2 |
| Stat 230 Applied Regression | 1 | CK | 1 | AP | 1 | KSC | 3 |
| Math 240 Probability | 2 | AP (2) | 1 | PT |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | CK | 1 | AL | 2 |
| Stat 260 Sampling Techniques |  |  |  |  |  |  | 0 |
| Stat 285 Stat Consulting | 0.34 | AL | 0.33 | AL | 0.33 | AL | 1 |
| Stat 310 Spatial Statistics |  |  |  |  | 1 | CK | 1 |
| Stat 320 Time Series |  |  |  |  |  |  | 0 |
| Stat 330 Advanced Modeling |  |  | 1 | AL |  |  | 1 |
| Stat 340 Bayesian Inference |  |  |  |  |  |  | 0 |
| Stat $2 \times x$ StatistieatLearning* |  |  |  |  |  |  | 0 |
|  | 6.34 |  | 7.33 |  | 7.33 |  | 21 |


|  | Year 3 - assumes 4 faculty @ 5 courses each + 1 part-time (PT) faculty (1 course) |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Course | fall |  | winter |  | spring |  | total |
| Stat 120 Intro to Statistics | 2 | AL, CK | 3 | AP, AL, CK | 3 | AP, KSC (2) | 8 |
| Stat 220 Data Science | 1 | AL |  |  | 1 | CK | 2 |
| Stat 230 Applied Regression | 1 | AP | 1 | KSC | 1 | AP | 3 |
| Math 240 Probability | 2 | CK, PT | 1 | AL |  |  | 3 |
| Stat 250 Intro to Stats Inference |  |  | 1 | CK | 1 | AL | 2 |
| Stat 260 Sampling Techniques | 1 | KSC |  |  |  |  | 1 |
| Stat 285 Stat Consulting | 0.34 | KSC | 0.33 | KSC | 0.33 | KSC | 1 |
| Stat 310 Spatial Statistics |  |  |  |  |  |  | 0 |
| Stat 320 Time Series |  |  | 1 | AP |  |  | 1 |
| Stat 330 Advanced Modeling |  |  |  |  |  |  | 0 |
| Stat 340 Bayesian Inference |  |  |  |  |  |  | 0 |
| Stat $2 \times$ Statistical Learning* |  |  |  |  |  |  | 0 |
|  | 7.34 |  | 7.33 |  | 6.33 |  | 21 |

## Appendix B: Historic enrollment and capacity data

Note that Stat 285 Statistical Consulting is a repeatable, 2 credit class that is waitlist only. The instructor will determine the class size based on the number of available consulting projects and student interest.

| 2020-21 Academic Year |  |  |  | Number of students registered |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| term | class | Title | Seats | section_1 | section_2 | section_3 | section_4 |
| 20FA | STAT 120 | Introduction to Statistics | 32 | 31 | 32 |  |  |
| 20FA | STAT 230 | Applied Regression Analysis | 26 | 24 |  |  |  |
| 20FA | STAT 260 | Intro to Sampling Techniques | 28 | 30 |  |  |  |
| 20FA | STAT 285 | Statistical Consulting | 0 | 12 |  |  |  |
| 21WI | STAT 120 | Introduction to Statistics | 32 | 30 | 31 | 32 |  |
| 21WI | STAT 220 | Introduction to Data Science | 30 | 27 |  |  |  |
| 21WI | STAT 230 | Applied Regression Analysis | 25 | 20 |  |  |  |
| 21WI | STAT 250 | Intro to Statistical Inference | 28 | 28 |  |  |  |
| 21 WI | STAT 285 | Statistical Consulting | 0 | 7 |  |  |  |
| 21WI | STAT 330 | Advanced Statistical Modeling | 20 | 15 |  |  |  |
| 21SP | STAT 120 | Introduction to Statistics | 32 | 33 | 32 | 31 |  |
| 21SP | STAT 220 | Introduction to Data Science | 30 | 38 |  |  |  |
| 21SP | STAT 230 | Applied Regression Analysis | 25 | 23 |  |  |  |
| 21SP | STAT 250 | Intro to Statistical Inference | 30 | 16 |  |  |  |
| 21SP | STAT 285 | Statistical Consulting | 0 | 14 |  |  |  |
| 21SP | STAT 320 | Time Series Analysis | 20 | 18 |  |  |  |


| 2021-22 Academic Year |  |  |  | Number of students registered |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| term | class | Title | Seats | section_1 | section_2 | section_3 | section_4 |
| 21FA | STAT 120 | Introduction to Statistics | 32 | 31 | 29 | 24 |  |
| 21FA | STAT 220 | Introduction to Data Science | 30 | 25 |  |  |  |
| 21FA | STAT 230 | Applied Regression Analysis | 25 | 29 |  |  |  |
| 21FA | STAT 285 | Statistical Consulting | 0 | 9 |  |  |  |
| 21FA | STAT 340 | Bayesian Statistics | 20 | 25 |  |  |  |
| 22WI | STAT 120 | Introduction to Statistics | 32 | 31 | 30 | 31 |  |
| 22WI | STAT 220 | Introduction to Data Science | 30 | 13 |  |  |  |
| 22WI | STAT 230 | Applied Regression Analysis | 25 | 24 |  |  |  |
| 22WI | STAT 250 | Intro to Statistical Inference | 28 | 30 |  |  |  |
| 22WI | STAT 260 | Intro to Sampling Techniques | 28 | 21 |  |  |  |
| 22WI | STAT 285 | Statistical Consulting | 0 | 6 |  |  |  |
| 22SP | STAT 120 | Introduction to Statistics | 32 | 34 | 32 | 30 | 33 |
| 22SP | STAT 220 | Introduction to Data Science | 30 | 25 |  |  |  |
| 22SP | STAT 230 | Applied Regression Analysis | 25 | 23 |  |  |  |
| 22SP | STAT 250 | Intro to Statistical Inference | 28 | 21 |  |  |  |
| 22SP | STAT 285 | Statistical Consulting | 0 | 11 |  |  |  |
| 22SP | STAT 320 | Time Series Analysis | 20 | 22 |  |  |  |


| 2022-23 Academic Year |  |  |  | Number of students registered |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| term | class | Title | Seats | section_1 | section_2 | section_3 | section_4 |
| 22FA | STAT 120 | Introduction to Statistics | 32 | 31 | 32 | 30 |  |
| 22FA | STAT 220 | Introduction to Data Science | 30 | 16 |  |  |  |
| 22FA | STAT 230 | Applied Regression Analysis | 28 | 25 |  |  |  |
| 22FA | STAT 285 | Statistical Consulting | 0 | 12 |  |  |  |
| 22FA | STAT 340 | Bayesian Statistics | 20 | 18 |  |  |  |
| 23WI | STAT 120 | Introduction to Statistics | 32 | 31 | 32 | 31 | 31 |
| 23WI | STAT 220 | Introduction to Data Science | 30 | 27 |  |  |  |
| 23WI | STAT 230 | Applied Regression Analysis | 28 | 20 |  |  |  |
| 23WI | STAT 250 | Intro to Statistical Inference | 30 | 27 |  |  |  |
| 23WI | STAT 285 | Statistical Consulting | 0 | 10 |  |  |  |
| 23WI | STAT 330 | Advanced Statistical Modeling | 20 | 13 |  |  |  |
| 23SP | STAT 120 | Introduction to Statistics | 32 | 33 | 32 | 32 | 33 |
| 23SP | STAT 220 | Introduction to Data Science | 30 | 30 |  |  |  |
| 23SP | STAT 230 | Applied Regression Analysis | 25 | 28 |  |  |  |
| 23SP | STAT 250 | Intro to Statistical Inference | 30 | 25 |  |  |  |
| 23SP | STAT 285 | Statistical Consulting | 0 | 9 |  |  |  |
| 23SP | STAT 310 | Spatial Statistics | 20 | 17 |  |  |  |


[^0]:    ${ }^{1}$ https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege Full.pdf
    ${ }^{2}$ https://magazine.amstat.org/blog/2021/10/01/undergrad-stats-degrees-up/
    ${ }^{3}$ https://www.amstat.org/asa/files/pdfs/EDU-DataScienceGuidelines.pdf

[^1]:    ${ }^{4}$ The top 50 liberal arts colleges were based on the 2022 National Liberal Arts Rankings from US News and World Report. This allowed for a systematic investigation of their curricula where we visited the website of each school, looked over the lists of majors, minors, and areas of study, and read descriptions of the curricula.

[^2]:    ${ }^{5}$ There is some fuzziness in these numbers since double majors outside of math and stat majors could be counted twice in the STEM or Not STEM categories. Double majors with a math or stat major would only be counted in the Math or Stat row.
    ${ }^{6}$ https://ww2.amstat.org/misc/StatsBachelors2003-MostRecent.pdf

[^3]:    ${ }^{7}$ In 19-20, 7 of 9 Stat 120 sections were taught by visitors and in 20-21, 5 of 8 Stat 120 sections were taught by visitors. Only one visitor, Owen Biesel, taught Stat 120 in both years.

[^4]:    ${ }^{8}$ https://www.amstat.org/asa/files/pdfs/YCR-SPAIGsalarysurvey15.pdf

[^5]:    ${ }^{9}$ There are very few PhD programs in Data Science and 2020 was the first year that the government started tracking degrees (of any type) awarded in the field of Data Science.

