



Data Science at Montgomery College

Montgomery College is a two-year college located outside of Washington, D.C., and has an enrollment of approximately 55,000 total students taking both credit and non-credit classes. The cost for a Montgomery County resident to attend for a semester (15 credits) is about \$2700. In 2015, faculty and administrators began working to create a data science certificate program for the College. The first cohort to join the College's Data Science Certificate program in 2017 included 30 students. To date, we have over 180 students who are either currently enrolled in or have taken data classes in the past. Both full time and part time faculty have taught these students over the past few years, and those faculty have carefully adjusted the content of the courses as we evolve and gain better insight into what the students and the broader community need and expect from the classes.

Graduates of our Data Science Certificate develop not only the hard skills needed to create high quality data science products but also the soft skills critical both to effectively communicating about the data and to considering the ethical implications of those products. Our students have digital skillsets anchored in STEM, yet they possess very diverse life experiences. In fact, one of our program's greatest assets is the sheer diversity of our students' backgrounds – some students enter without any prior degree, while some students enter the program with prior undergraduate or even in some cases graduate degrees, with males and females equally represented, and age groups ranging from teens to retired professionals.

Students must complete five courses, or 16 credits, to earn the certificate: **one statistics class of their choice; Introduction to Data Science (DATA 101)** which focuses primarily on data ingestion and cleaning; **Data Visualization and Communication (DATA 110)**; **Statistical Methods in Data Science (DATA 201)** which includes advanced statistical methods, machine learning, clustering and regression; and the **Capstone Experience in Data Science (DATA 205)** to work with an industry or governmental partner. Students are able to complete this certificate in three semesters or even two semesters if they have already taken a statistics course, though some will take four semesters. In addition to the listed topics, students explore concepts such as web scraping, data ethics, geographic information science, and reproducible research. All data classes are offered during fall and spring semesters, and the 100-level courses are offered during the summer semesters, as well.

The final course, the project-based Capstone Experience, is where students display their accomplishments and the program gains some of its greatest visibility. Based on open data sets shared by Montgomery County, our first capstones have had the opportunity to present their



analyses to county officials and other stakeholders. The data sharing benefits both parties—students get excellent real-world experience, and county officials receive useful, sometimes eye-opening, information.

Data Tools

In the two 100-level classes (DATA 101 and DATA 110), students learn to program in R in the Tidyverse and they learn Python in DATA 201. In addition, during their certificate experience, they learn to use Tableau in the Data Visualization and Communication class, they are exposed to Github throughout the certificate experience, and professors also share other software within each of the classes. All resources for all data classes are free and open source materials.

Challenges

One challenge is serving students who enter with very little undergraduate-level mathematics experience. Data Science generally seems to increasingly garner broader appeal among students, so many students coming in to the program with various mathematical and statistical backgrounds are taking data classes and considering the data science certificate. As mentioned earlier, only one statistics class is the required prerequisite to take the 100-level data classes and to complete the certificate, but one statistics course may not be enough to successfully transfer to a four-year statistics or data science degree program.

Other challenges we have faced include advertising the program, identifying students who are the right fit, and connecting more students to networking and career opportunities. Although we are empowering our students with tools to work with and explore data, to perform statistical analyses and create data visualizations, and then finally to communicate their discoveries to a broader audience, we continually need to address the “what next” career transition portion for when they graduate with the certificate. We know that there is a huge need for entry level data analysts, especially in Montgomery County and the greater Washington DC community. We will not only better serve our students by connecting them with area organizations, but these organizations in turn will gain from the energy, curiosity, intelligence, talent, and strong drive that these students possess in abundance.

Accomplishments

Over the past three years, we have essentially doubled the size of our data classes. We have successfully completed a transfer agreement for students to pursue a four-year degree at a local university, and we are actively working on other transfer agreements. In addition, we are establishing partnerships with other employers and local relationships. We are working with other organizations to create opportunities for our students such as internships, networking events, and work shadowing experiences. The pandemic has slowed some of the momentum we had with this exploration, but we are still working to build these local partnerships for the students.