



Data Science Degree at a Two-Year College: *Fitting it all in*

Montgomery College is a two-year college located outside of Washington, D.C., ranked ***Most Diverse TYC*** (in continental US) by Chronicle of Higher Ed, and has an enrollment of approximately 55,000 total students across three campuses. 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 200 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, 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. 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.

The Data Science Certificate

Students must complete five courses, or 16 credits, to earn the certificate: **one introductory statistics class of their choice; Introduction to Data Science (DATA 101)** which focuses primarily on data ingestion, cleaning, EDA, and ethics; **Data Visualization and Communication (DATA 110)** which includes exploring the Tidyverse, principles of visualizations, web scraping, ethics, GIS, and reproducible research; **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)** where students 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. 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. All materials are free and open sourced for the students. Data tools used include R Studio, Python, Tableau, Git/Github, and SQL.

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 (*dataMontgomery*), 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 real-world experience, and county officials receive useful, sometimes eye-opening, information.



Creating an Associate of Science Data Science Degree

We are currently working on creating a Data Science A.S. degree, and one of the greatest challenges is fitting all courses into two years with the limited 60-credits. Obviously, we must satisfy articulation agreements, including Gen Ed requirements. But beyond meeting the needs of our transfer institutions, we also are trying to provide our students with the best training and skills they can acquire from us. Those skills include mathematics, statistics, data ethics, data tools, and of course programming skills. Fitting all of those skills into course sequences that fit prerequisites is a puzzle that has necessitated negotiations with other departments such as the Mathematics department and the Computer Science department. We are navigating new territory, because Data Science does not necessarily fit with what has worked in the past.

One of the positive side effects of having to fit all the puzzle pieces together has been the great collaborations that have evolved among other departments within Montgomery College as well as discussions with colleagues at other institutions. We have greatly benefitted by directly speaking with and getting advice from folks at ADSA (Academic Data Science Alliance), UNC Charlotte, and AUCC (Atlanta University Center Consortium). Additional insight and information gathering has come from attending recent conferences such as ASA's *USCOTS21*, AUC's *W.E.B. Du Boise Symposium*, Berkeley's *National Workshop on Data Science Education*, and Databrick's *Data Science + AI Summit*.

Other challenges we perpetually face include advertising the program, identifying students who are the right fit, and connecting more students to networking and career opportunities. 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 not only will 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.