What is Data Science?

Data science is the study of the mathematical and computational methods for extracting meaning from data. It involves the collection, processing, organization, quantitative analysis, visualization and modeling of data.
Why Data Science?

It empowers you to solve important, challenging problems that are otherwise impractical or impossible to solve.

Data science jobs are among the fastest growing in the country, and data scientists often land interesting, challenging and lucrative jobs directly out of college.
Examples of “Important, Challenging” Problems

Protein Folding

Self-Driving Vehicles

Accelerated Climate Modeling

Audio/Video Understanding

Natural Language Understanding

To do a "farduddle" means to jump up and down really fast. An example of a sentence that uses the word farduddle is:

One day when I was playing tag with my little sister, she got really excited and she started doing these crazy farduddles.

A "yalubalu" is a type of vegetable that looks like a big pumpkin. An example of a sentence that uses the word yalubalu is:

I was on a trip to Africa and I tried this yalubalu vegetable that was grown in a garden there. It was delicious.
Data Science Curriculum
Data Science BS Curriculum

- Eight Computer Science courses
- Seven Math and Stats courses
- Two core Data Science courses
- Data Science senior project series
- Electives from
  - Computer Science
  - Math and Stats
  - M/CS
- A science sequence
DS BS: Computer Science Courses

- CSCI 141 - Computer Programming I
- CSCI 145 - Computer Programming and Linear Data Structures
- CSCI 241 - Data Structures
- CSCI 301 - Formal Languages and Functional Programming
- CSCI 305 - Analysis of Algorithms and Data Structures I
- CSCI 330 - Database Systems
- CSCI 345 - Object Oriented Design
- CSCI 405 - Analysis of Algorithms and Data Structures II
DS BS: Mathematics Courses

- MATH 124 - Calculus and Analytic Geometry I *(or equivalent)*
- MATH 125 - Calculus and Analytic Geometry II *(or equivalent)*
- MATH 204 - Elementary Linear Algebra
- MATH 224 - Multivariable Calculus and Geometry I
- MATH 304 - Linear Algebra
- MATH 341 - Probability and Statistical Inference
- MATH 342 - Statistical Methods I
DS BS: Data Science Courses

- Core classes *(details in a few slides)*:
  - DATA 311 Fundamentals of Data Science
  - DATA 371 Machine Learning

- DATA 491/492/493 Senior Project
  - Must center around data science
DS BS: Electives

Computer Science Electives
CSCI 400 - Directed Independent Study
CSCI 402 - Artificial Intelligence
CSCI 404 - Natural Language Processing
CSCI 424 - Social Network Analysis
CSCI 436 - Technology for Social Good
CSCI 471 - Advanced Machine Learning
CSCI 474 - Bioinformatics
CSCI 476 - Computer Vision
CSCI 477 - Data Mining
CSCI 479 - Spoken Language Processing
CSCI 481 - Deep Learning
CSCI 496 - Undergraduate Research
CSCI 497* - Temporary Courses

Math and Stats Electives
MATH 443 - Linear Statistical Models
MATH 444 - Categorical Data Analysis
MATH 445 - Computational Statistics
MATH 447 - Multivariate Statistics
MATH 456 - Applied Time Series Analysis
MATH 457 - Bayesian Statistics
MATH 458 - Stochastic Processes
MATH 473 - Numerical Linear Algebra

M/CS Electives
M/CS 335 - Linear Optimization
M/CS 375 - Numerical Computation
M/CS 435 - Nonlinear Optimization
M/CS 475 - Numerical Analysis C
DS BS: Science Sequence

- BIOL 204, 205, 206, or
- CHEM 161, 162, 163, or
- GEOL 211, 212 and one in {303, 308, 309, 314, 315, 340}, or
- PHYS 161, 162, 163
DATA 311: Fundamentals of Data Science

Format: Three weekly lectures + one weekly lab (4 credits)

Course Description: Introduction to the fundamentals of data science, focusing on techniques for collecting, processing, visualizing and organizing data. Applied machine learning concepts will also be covered, including fundamentals of machine learning experimentation and the use of libraries to perform clustering, classification and regression. Includes lab.

Prereqs: CSCI 141; MATH 112 or equivalent

Capacity: 35-40 seats
DATA 371: Machine Learning

**Format:** Four weekly lectures (4 credits)

**Course Description:** Key machine learning theory, algorithms and experimentation techniques. Generalization, bias and variance. Classification, regression, clustering and probabilistic modeling. Linear models and neural networks. Discrete and continuous optimization algorithms. Ethical issues in machine learning.

**Prereqs:** CSCI 241; MATH 204; MATH 224; and MATH 341. DS major status.

**Capacity:** 35 seats
Becoming a Data Science Major
If you do not already have a declared premajor:

1. Login to Web4u
2. Click “Students”
3. Click “Student Records”
4. Click “Academic Interest Update Form”
5. Indicate Data Science

This will ensure you receive all relevant information about DS.
Step 1: Take Pre-major Classes

- You cannot apply until the quarter you will complete your pre-major courses:
  - CSCI 141 (offered every quarter)
  - CSCI 145 (offered every quarter)
  - CSCI 241 (offered every quarter)
  - CSCI 301 (offered every quarter)
  - DATA 311 (offered Fall 2021 and Spring 2022)
Step 2: Apply

- Students should apply to the major by the 5th week of the quarter in which they will complete the pre-major courses.

- We plan to admit majors twice yearly:
  - First round of applications will be taken in Fall 2021.
  - Second round of applications will be taken in Spring 2022.

- During admission process, students will meet with Math advisor; once admitted, they will be assigned an advisor from the Computer Science department.

- Additional details (e.g., specific e-forms) are still being ironed out.
How to get into Fall 2021 DATA 311

- DATA 311 in Fall 2021 will have an override restriction

- Who is eligible for an override during Phase 1? Those who have...
  - completed or have transfer equivalent of CSCI 141
  - not been admitted to any major – including the Computer Science major

- To request an override: complete the Fall 2021 DATA 311 Override Request Form
  - Requests received by May 21, 2021 will receive priority processing.
How to get into Spring 2022 DATA 311

- In Spr 2022, Phase 1 registration for DATA 311 likely limited to pre-majors

- We are still establishing procedures for admissions to the DS Pre-Major
  - Will be advertised on department website
  - Will be distributed to those indicating DS Interest
Data Science Advising
Where to find advising information

- **Website:**
  - For the most up-to-date and detailed information
  - Includes curriculum flowchart and example plans of study
- **Program Advisor:** Prof. Brian Hutchinson
  - Advising email: datascience.advising@wwu.edu
  - Later: a dedicated staff advisor
- **Catalog:** [https://catalog.wwu.edu/preview_program.php?catoid=18&poid=8909](https://catalog.wwu.edu/preview_program.php?catoid=18&poid=8909)
  - For the official program requirements
Reminder to Declare DS Interest

If you do not already have a declared premajor:

1. Login to Web4u
2. Click “Students”
3. Click “Student Records”
4. Click “Academic Interest Update Form”
5. Indicate Data Science

This will ensure you receive all relevant information about DS.
Data Science BS Q&A
DS BS vs CS BS

- **Cuts**
  - CSCI 247 Computer Systems I
  - CSCI 347 Computer Systems II
  - CSCI 367 Computer Networks I
  - CSCI 447 Operating Systems
  - CISS 346 Secure Software Development

- **Adds**
  - MATH 224 Multivariable Calculus and Geometry I
  - MATH 304 Linear Algebra
  - MATH 342 Statistical Methods I
  - DATA 311 Fundamentals of Data Science
  - DATA 371 Machine Learning