Proposal for New Program:

BS in Data Science: Genomics

1. “Rationale...” The proposed “Data Science: Genomics” major is designed for students interested in developing expertise in data science, with specialization in genomics. Data Science is an interdisciplinary field about methods and systems to extract knowledge or insights from large quantities of data coming in various forms. Data science employs techniques and theories drawn from many disciplines within the broad areas of mathematics, statistics, and computer and information sciences and applies them on a wide range of data-rich domains such as biomedical sciences, physical science, geoscience, social science, engineering, business, and education.

Data science is a very broad and multifaceted field and it is not realistic to expect that a B.S. program could provide students with deep expertise in all aspects of the field. The “Genomics” specialization will give students a strong background in mathematics, computational thinking, and biological data analysis, and will enable students to analyze large quantities of data to discover new knowledge and facilitate decision making. As part of the broader data science program, genomics is a critical track to include as a sub-area of focus for students interested in biology, ecology, evolution, human health and disease, and precision medicine. Over the past decade, the emergence of next-generation sequencing technologies has facilitated the rapid growth of genomic data; however, undergraduate training in big data management, big data processing, and big data analysis has not kept up with this rapid growth in large-scale genomic data generation.

Graduates of this program will have multiple career opportunities. Some of them will choose to find data science jobs in a private or a public sector and some will choose to continue with graduate studies either to deepen their overall data science expertise or learn how to better use their analytics skills in genomics.
# Appendix A
## Data Science: Genomics BS Semester Sequence Proposal

<table>
<thead>
<tr>
<th>Semester</th>
<th>Courses</th>
</tr>
</thead>
</table>
| **Freshman – Fall** (16 cr.) | Chem 1031&1033 General Chemistry I (3+1 cr.)  
CIS 1068 Program Design and Abstraction (4 cr.)  
Math 1041 Calculus I (4 cr.)  
Gen Ed English 0802 (4 cr.) |
| **Freshman – Spring** (15 cr.) | Chem 1032&1034 General Chemistry II (3+1 cr.)  
CIS 1166 Mathematical Concepts in Computing I (4 cr.)  
Math 1042 Calculus II (4 cr.)  
Gen Ed IH 0851 (3 cr.) |
| **Sophomore – Fall** (17 cr.) | Biol 1111 Introduction to Biology (4 cr.)  
Chem 2201&2203 Organic Chemistry I (3+1 cr.)  
CIS 2168 Data Structures (4 cr.)  
Gen Ed IH 0852 (3 cr.)  
Gen Ed World Society (3 cr.) |
| **Sophomore – Spring** (17 cr.) | Biol 2112 Introduction to Biology (4 cr.)  
Chem 2202&2204 Organic Chemistry II (3+1 cr.)  
Math 3031 Probability Theory I (3 cr.)  
Gen Ed US Society (3 cr.)  
Gen Ed Behavior (3 cr.) |
| **Junior – Fall** (15 cr.) | Biol 3101F Evolution (3 cr.)  
CIS 2166 Mathematical Concepts in Computing I (4 cr.)  
Math 3032 Probability Theory II (3 cr.)  
Gen Ed Arts (3-4 cr.)  
Elective (2-1 cr.) |
| **Junior – Spring** (14 cr.) | Biol 2296S Genetics (4 cr.)  
CIS 3715S Principles of Data Science (4 cr.)  
DS: Genomics Elective (3 cr.)  
Gen Ed Race (3 cr.) |
| **Senior – Fall** (15 cr.) | Biol 3111F Genomics in Medicine (3 cr.)  
DS: Genomics Elective (3 cr.)  
Elective (3 cr.)  
Elective (3 cr.)  
Elective (3 cr.) |
| **Senior – Spring** (14 cr.) | SCTC5 xxxx Advanced Data Visualization (3 cr.)  
DS: Genomics Elective (3 cr.)  
Elective (4 cr.)  
Elective (4 cr.) |

Credits in the major: 80

Credits in General Education: 25-26
Elective credits: 18-17
Total credits: 123
Appendix B

New Data Science: Genomics B.S. Degree.

Intro Science Requirements:
- Chem 1031&1033 General Chemistry I (3+1 cr.)
- Chem 1032&1034 General Chemistry II (3+1 cr.)
Subtotal 8 credits

Calculus Requirements:
- Math 1041 (4 cr.) Calculus I
- Math 1042 (4 cr.) Calculus II
Subtotal 8 credits

Math Methods in Computing Requirements:
- CIS 1166 (4 cr.) Mathematical Concepts in Computing I
- CIS 2166 (4 cr.) Mathematical Concepts in Computing II
Subtotal 8 credits

Probability and Statistics Requirements:
- Math 3031 (3 cr.) Probability Theory I
- Math 3032 (3 cr.) Probability Theory II
Subtotal 6 credits

Programming Requirements:
- CIS 1068 (4 cr.) Program Design and Abstraction
- CIS 2168 (4 cr.) Data Structures
Subtotal 8 credits

Specialty Course Requirements:
- Biol 1111 Introduction to Biology I (4 cr.)
- Biol 2112 Introduction to Biology II (4 cr.)
- Biol 3101 Evolution (3 cr.)
- Biol 3111 Genomics in Medicine (3 cr.)
- Chem 2201&2203 Organic Chemistry I (3+1 cr.)
- Chem 2202&2204 Organic Chemistry I (3+1 cr.)
- CIS 3715 (4 cr.) Principles of Data Science
Subtotal 26 credits

WI/Capstone Requirements:
- Biol 2296 Genetics (4 cr.)
- SCTC xxxx (3 cr.) Advanced Data Visualization
Subtotal 7 credits

Elective Course Requirements (9 credits required):
- Biol 2227 Principles Of Ecology (3 cr.)
Biol 3112 Foundations Of Evolutionary Genomic Medicine (3 cr.)
Biol 3114 Evolutionary Ecology (3 cr.)*
Biol 3128 Genomics and Infectious Disease Dynamics*
Biol 3201 Human Genetics (3 cr.)
Biol 3211 Human Evolution (3 cr.)
Biol 3225 Evolutionary Genetics (3 cr.)
Biol 3241 Genomics & Evolutionary Biology Of Parasites (3 cr.)
Biol 3225 Evolutionary Genetics (3 cr.)
Biol 3241 Genomics & Evolutionary Biology Of Parasites (3 cr.)
Biol 3321 Plant Community Ecology (3 cr.)*
Biol 3322 Plant Genetics (3 cr.)
Biol 3324 Molecular Biology (3 cr.)
Biol 3328 Virology (3 cr.)
Biol 3368 Biology of Cancer (3 cr.)**
Biol 3379 Biotechnology (3 cr.)**
Biol 3403 Genomics (3 cr.)
Biol 4xxx UG Computational Bio courses being developed (3 cr.)
CEE 3048 Probability, Statistics & Stochastic Methods (3 cr.)
CIS 4523/9664 Data Mining (3 cr.)

Subtotal 9 credits

Total 80 credits

*This course requires an additional prerequisite of BIOL 2227
**This course requires an additional prerequisite of BIOL 3096
CST Data Science BS programs Common Core

- Chem 1031&1033&1032&1034 General Chemistry I&II (3+1+3+1 cr.) OR Biol 1111&2112 Introduction to Biology I&II (4+4 cr.) OR Phys 1061&1062 (4+4 cr.) Elementary Classical Physics I&II (plus variants)
- CIS 1068 (4 cr.) Program Design and Abstraction
- CIS 2168 (4 cr.) Data Structures
- CIS 1166 (4 cr.) Mathematical Concepts in Computing I
- CIS 2166 (4 cr.) Mathematical Concepts in Computing II
- Math 1041 (4 cr.) Calculus I
- Math 1042 (4 cr.) Calculus II
- Math 3031 (3 cr.) Probability Theory I
- Math 3032 (3 cr.) Probability Theory II
- CIS 3715 (4 cr.) Principles of Data Science
- SCTC xxxx (3 cr.) Advanced Data Visualization

Data Science: Genomics BS Comparison to CST Data Science Common Core

Changed
- Chem 1031&1033&1032&1034 General Chemistry I&II (3+1+3+1 cr.) (plus variants) specified

Added
- Biol 1111 Introduction to Biology I (4 cr.)
- Biol 2112 Introduction to Biology II (4 cr.)
- Biol 2296 Genetics (4 cr.)
- Biol 3101 Evolution (3 cr.)
- Biol 3111 Genomics in Medicine (3 cr.)
- Chem 2201&2203 Organic Chemistry I (3+1 cr.)
- Chem 2202&2204 Organic Chemistry I (3+1 cr.)

Data Science: Modeling Physical Systems BS Comparison to CST Data Science Common Core

Changed
- Phys 1061&1062 (4+4 cr.) Elementary Classical Physics I&II (plus variants) specified

Added
- Math 3045 (4 cr.) Probability Theory I or Math 2101 (3 cr.) Linear Algebra or Math 2103 (4 cr.) Linear Algebra with Lab
- CIS 3223 (4 cr.) Data Structures and Algorithms
- Math 2043 (4 cr.) Calculus III
- Math 3043 (3-4 cr.) Numerical Analysis I
- Phys 2501 (3 cr.) Computing for Scientists
• Phys 2502 (4 cr.) Mathematical Physics
• Phys 2796 (4 cr.) Introduction to Modern Physics

Data Science: Computational Analytics BS Comparison to CST Data Science Common Core

Added
• Math 3045 (4 cr.) Probability Theory I or Math 2101 (3 cr.) Linear Algebra or Math 2103 (4 cr.) Linear Algebra with Lab
• CIS 2107 (4 cr.) Computer Systems and Low-Level Programming
• CIS 3223 (4 cr.) Data Structures and Algorithms
• CIS 4331 (4 cr.) Principles of Database Systems 4cr
• CIS 4526 (3 cr.) Foundations of Machine Learning
• CIS 4517 (4 cr.) Data-Intensive and Cloud Computing 4cr
• Math 2043 (4 cr.) Calculus III
• ENG 2696 (3 cr.) Technical Writing

Comparison of CST Data Science BS programs Common Core to Statistics and Data Science BS

Common
• CIS 1068 (4 cr.) Program Design and Abstraction
• Math 1041 (4 cr.) Calculus I
• Math 1042 (4 cr.) Calculus II

Alternates
• CIS 2168 (4 cr.) Data Structures versus CIS 1051. Introduction to Problem Solving and Programming in Python (4 cr.)
• Math 3031 (3 cr.) Probability Theory I versus Stat 2103/2903 Statistics for Business Analytics (4)
• Math 3032 (3 cr.) Probability Theory II versus Stat 2512 Intermediate Statistics- (3)

Statistics and Data Science BS only
• BA 2196/2996 Business Communications – (3)
• BA 2104 Excel for Business Applications (1)
• HRM 1101/1901 Leadership and Organizational Management- (3)
• ECON 1102/1902 Microeconomic Principles (3)
• ECON 1101/1901 Macroeconomics Principles (3)
• RMI 2101/2901. Introduction to Risk Management – (3)
• MKTG 2101/2901. Marketing Management- (3)
• Acct 2101/2901 Financial Accounting- (3)
• Stat 2501 Quantitative Foundations for Data Science –(3) new course
• Stat 2521 Data Analysis and Statistical Computing- (3)
• Stat 2523 Design of Experiments & Quality Control- (3)
• Stat 2522  Survey Design and Sampling- (3)
• Stat 3503 Intermediate Business Statistics- (3)
• Stat 3505 Introduction to SAS for Data Analytics (3) – new course
• Stat 3502 Regression and Predictive Analytics(3) – new course
• Stat 3504 Time Series and Forecasting Models(3) – new course
• Stat 3506 Nonparametric and Categorical Data Analysis(3) – new course
• Stat 4501 Capstone: Statistical methods and Data Analytics (use of SAS (with Online certification) and R (3) – new course
• 2 elective courses

CST Data Science BS programs Common Core only
• Chem 1031&1033&1032&1034 General Chemistry I&II (3+1+3+1 cr.) OR Biol 1111&2112 Introduction to Biology I&II (4+4 cr.) OR Phys 1061&1062 (4+4 cr.) Elementary Classical Physics I&II (plus variants)
• CIS 1068 (4 cr.) Program Design and Abstraction
• CIS 1166 (4 cr.) Mathematical Concepts in Computing I
• CIS 2166 (4 cr.) Mathematical Concepts in Computing II
• CIS 3715 (4 cr.) Principles of Data Science
• SCTC xxxx (3 cr.) Advanced Data Visualization
• 7-8 required courses based on particular CST Data Science BS program
• 3 elective courses based on particular CST Data Science BS program

Comparison to Biology BS

Comparison of Data Science: Genomics BS to Biology BS
Removed
• Phys 2021&2022 (4+4 cr.) General Physics I&II (plus variants)
• Biol 2227 Ecology (3 cr.)
• Biol 3096 Cell Structure and Function (4 cr.)
• 4 Biology electives

Specified
• Biol 3101 Evolution (3 cr.) as 1 elective
• Biol 3111 Genomics in Medicine (3 cr.) as 1 elective

Added
• CIS 1068 (4 cr.) Program Design and Abstraction
• CIS 2168 (4 cr.) Data Structures
• CIS 1166 (4 cr.) Mathematical Concepts in Computing I
• CIS 2166 (4 cr.) Mathematical Concepts in Computing II
• Math 3031 (3 cr.) Probability Theory I
• Math 3032 (3 cr.) Probability Theory II
• CIS 3715 (4 cr.) Principles of Data Science
• SCTC xxxx (3 cr.) Advanced Data Visualization