

# Department of Statistics and Data Science

Christopher R. Genovese, Department Head  
 Rebecca Nugent, Director of Undergraduate Studies  
 Christopher Peter Makris, Programs Administrator  
 Samantha Nielsen, Academic Advisor  
 Email: statadvising@stat.cmu.edu  
 Department Office: Baker Hall 132

## Overview

Uncertainty is inescapable: randomness, measurement error, deception, and incomplete or missing information complicate all our lives. Statistics is the science and art of making predictions and decisions in the face of uncertainty. Statistical issues are central to big questions in public policy, law, medicine, industry, computing, technology, finance, and science. Indeed, the tools of Statistics apply to problems in almost every area of human activity where data are collected.

Statisticians must master diverse skills in computing, mathematics, decision making, forecasting, interpretation of complicated data, and design of meaningful comparisons. Moreover, statisticians must learn to collaborate effectively with people in other fields and, in the process, to understand the substance of these other fields. For all these reasons, Statistics students are highly sought-after in the marketplace.

Recent Statistics majors at Carnegie Mellon have taken jobs at leading companies in many fields, including the National Economic Research Association, Boeing, Morgan Stanley, Deloitte, Rosetta Marketing Group, Nielsen, Proctor and Gamble, Accenture, and Goldman Sachs. Other students have taken research positions at the National Security Agency, the U.S. Census Bureau, and the Science and Technology Policy Institute or worked for Teach for America. Many of our students have also gone on to graduate study at some of the top programs in the country including Carnegie Mellon, the Wharton School at the University of Pennsylvania, Johns Hopkins, University of Michigan, Stanford University, Harvard University, Duke University, Emory University, Yale University, Columbia University, and Georgia Tech.

## The Department and Faculty

The Department of Statistics and Data Science at Carnegie Mellon University is world-renowned for its contributions to statistical theory and practice. Research in the department runs the gamut from pure mathematics to the hottest frontiers of science. Current research projects are helping make fundamental advances in neuroscience, cosmology, public policy, finance, and genetics.

The faculty members are recognized around the world for their expertise and have garnered many prestigious awards and honors. (For example, three members of the faculty have been awarded the COPSS medal, the highest honor given by professional statistical societies.) At the same time, the faculty is firmly dedicated to undergraduate education. The entire faculty, junior and senior, teach courses at all levels. The faculty are accessible and are committed to involving undergraduates in research.

The Department augments all these strengths with a friendly, energetic working environment and exceptional computing resources. Talented graduate students join the department from around the world, and add a unique dimension to the department's intellectual life. Faculty, graduate students, and undergraduates interact regularly.

## How to Take Part

There are many ways to get involved in Statistics at Carnegie Mellon:

- The Bachelor of Science in Statistics in the Dietrich College of Humanities and Social Sciences (DC) is a broad-based, flexible program that helps you master both the theory and practice of Statistics. The program can be tailored to prepare you for later graduate study in Statistics or to complement your interests in almost any field, including Psychology, Physics, Biology, History, Business, Information Systems, and Computer Science.
- The Minor (or Additional Major) in Statistics is a useful complement to a (primary) major in another Department or College. Almost every field of inquiry must grapple with statistical problems, and the tools of statistical theory and data analysis you will develop in the Statistics minor (or Additional Major) will give you a critical edge.
- The Bachelor of Science in Economics and Statistics provides an interdisciplinary course of study aimed at students with a strong interest in the empirical analysis of economic data. Jointly administered by the Department of Statistics and Data Science and the

Undergraduate Economics Program, the major's curriculum provides students with a solid foundation in the theories and methods of both fields. (See Dietrich College Interdepartmental Majors as well later in this section)

- The Bachelor of Science in Statistics and Machine Learning is a program housed in the Department of Statistics and Data Science and is jointly administered with the Department of Machine Learning. In this major students take courses focused on skills in computing, mathematics, statistical theory, and the interpretation and display of complex data. The program is geared toward students interested in statistical computation, data science, and "big data" problems.
- The Statistics Concentration and the Operations Research and Statistics Concentration in the Mathematical Sciences Major (see Department of Mathematical Sciences) are administered by the Department of Mathematical Sciences with input from the Department of Statistics and Data Science.
- There are several ongoing exciting research projects in the Department of Statistics and Data Science, and the department enthusiastically seeks to involve undergraduates in this work. Both majors and non-majors are welcome.
- Non-majors are eligible to take most of our courses, and indeed, they are required to do so by many programs on campus. Such courses offer one way to learn more about the Department of Statistics and Data Science and the field in general.

## Curriculum

Statistics consists of two intertwined threads of inquiry: Statistical Theory and Data Analysis. The former uses probability theory to build and analyze mathematical models of data in order to devise methods for making effective predictions and decisions in the face of uncertainty. The latter involves techniques for extracting insights from complicated data, designs for accurate measurement and comparison, and methods for checking the validity of theoretical assumptions. Statistical Theory informs Data Analysis and vice versa. The Department of Statistics and Data Science curriculum follows both of these threads and helps the student develop the complementary skills required.

Below, we describe the requirements for the Major in Statistics and the different categories within our basic curriculum, followed by the requirements for the Major in Economics and Statistics, the Major in Statistics and Machine Learning, and the Minor in Statistics.

**Note:** We recommend that you use the information provided below as a general guideline, and then schedule a meeting with a Statistics Undergraduate Advisor (email: statadvising@stat.cmu.edu) to discuss the requirements in more detail, and build a program that is tailored to your strengths and interests.

## B.S. in Statistics

Academic Advisor: Samantha Nielsen  
 Faculty Advisors: Peter Freeman and Mark Schervish  
 Office: Baker Hall 132  
 Email: statadvising@stat.cmu.edu

Students in the Bachelor of Science program develop and master a wide array of skills in computing, mathematics, statistical theory, and the interpretation and display of complex data. In addition, Statistics majors gain experience in applying statistical tools to real problems in other fields and learn the nuances of interdisciplinary collaboration. The requirements for the Major in Statistics are detailed below and are organized by categories #1-#7.

## Curriculum

### 1. Mathematical Foundations (Prerequisites) 29-39 units

Mathematics is the language in which statistical models are described and analyzed, so some experience with basic calculus and linear algebra is an important component for anyone pursuing a program of study in Statistics.

Calculus\*:

Complete one of the following three sequences of mathematics courses at Carnegie Mellon, each of which provides sufficient preparation in calculus:

**Sequence 1**

21-111	Differential Calculus	10
21-112	Integral Calculus	10
and one of the following		
21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

**Sequence 2**

21-120	Differential and Integral Calculus	10
and one of the following		
21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

**Notes:**

- Other sequences are possible, and require approval from the undergraduate advisor.
- Passing the MSC 21-120 assessment test is an acceptable alternative to completing 21-120.

**Linear Algebra\*\*:**

Complete *one* of the following three courses:

21-240	Matrix Algebra with Applications	10
21-241	Matrices and Linear Transformations	10
21-242	Matrix Theory	10

\* It is recommended that students complete the calculus requirement during their freshman year.

\*\*The linear algebra requirement needs to be completed before taking 36-401 Modern Regression.

21-241 and 21-242 are intended only for students with a very strong mathematical background.

**2. Data Analysis: 36–45 units**

Data analysis is the art and science of extracting insight from data. The art lies in knowing which displays or techniques will reveal the most interesting features of a complicated data set. The science lies in understanding the various techniques and the assumptions on which they rely. Both aspects require practice to master.

The Beginning Data Analysis courses give a hands-on introduction to the art and science of data analysis. The courses cover similar topics but differ slightly in the examples they emphasize. 36-200 or 36-201 draw examples from many fields and satisfy the DC College Core Requirement in Statistical Reasoning. One of these courses is therefore recommended for students in the College. (Note: A score of 4 or 5 on the Advanced Placement (AP) Exam in Statistics may be used to waive this requirement). Other courses emphasize examples in business (36-207), engineering and architecture (36-220), and the laboratory sciences (36-247).

The Intermediate Data Analysis courses build on the principles and methods covered in the introductory course, and more fully explore specific types of data analysis methods in more depth.

The Advanced Data Analysis courses draw on students' previous experience with data analysis and understanding of statistical theory to develop advanced, more sophisticated methods. These core courses involve extensive analysis of real data with emphasis on developing the oral and writing skills needed for communicating results.

Sequence 1 (For students beginning their freshman or sophomore year)

**Beginning\***

Choose *one* of the following courses:

36-200	Reasoning with Data	9
36-201	Statistical Reasoning and Practice	9
36/70-207	Probability and Statistics for Business Applications	9
36-220	Engineering Statistics and Quality Control	9
36-247	Statistics for Lab Sciences	9

\*Or extra data analysis course in Statistics

Note: Students who enter the program with 36-225 or 36-226 should discuss options with an advisor. Any 36-300 or 36-400 level course in Data Analysis that does not satisfy any other requirement for a Statistics Major and Minor may be counted as a Statistical Elective.

**Intermediate\***

Choose *one* of the following courses

36-202	Methods for Statistics and Data Science **	9
36/70-208	Regression Analysis	9
36-309	Experimental Design for Behavioral and Social Sciences	9

\*Or extra data analysis course in Statistics

\*\* Must take prior to 36-401

**Advanced**

Choose *one* of the following courses:

36-303	Sampling, Survey and Society	9
36-315	Statistical Graphics and Visualization	9
and take the following two courses:		
36-401	Modern Regression	9
36-402	Advanced Methods for Data Analysis	9

Students can also take a second 36-46x (see section #5).

Sequence 2 (For students beginning later in their college career)

**Advanced**

Choose *two* of the following courses:

36-303	Sampling, Survey and Society	9
36-315	Statistical Graphics and Visualization	9
36-461	Special Topics: Statistical Methods in Epidemiology	9
36-462	Special Topics: Data Mining	9
36-463	Special Topics: Multilevel and Hierarchical Models	9
36-464	Special Topics: Applied Multivariate Methods	9
36-490	Undergraduate Research	9

\*\*Special Topics rotate and new ones are regularly added. See section 5 for details.

and take the following *two* courses:

36-401	Modern Regression	9
36-402	Advanced Methods for Data Analysis	9

**3. Probability Theory and Statistical Theory: 18 units**

The theory of probability gives a mathematical description of the randomness inherent in our observations. It is the language in which statistical models are stated, so an understanding of probability is essential for the study of statistical theory. Statistical theory provides a mathematical framework for making inferences about unknown quantities from data. The theory reduces statistical problems to their essential ingredients to help devise and evaluate inferential procedures. It provides a powerful and wide-ranging set of tools for dealing with uncertainty.

To satisfy the theory requirement take the following two courses:

36-225	Introduction to Probability Theory **	9
and one of the following two courses:		
36-226	Introduction to Statistical Inference	9
36-326	Mathematical Statistics (Honors)	9

\*\*It is possible to substitute 36-217 or 21-325 for 36-225. (36-225 is the standard introduction to probability, 36-217 is tailored for engineers and computer scientists, and 21-325 is a rigorous probability theory course offered by the Department of Mathematics.)

**Comments:**

(i) In order to be a Major or a Minor in good standing, a grade of at least a C is required in 36-225, 36-226 and 36-401. In particular, a grade of C or higher is required in order to be able to continue in the major.

(ii) In special cases, and in consultation with the Statistics Advisor, the theory requirement can be satisfied by taking a single *graduate level* class 36-700 Probability and Mathematical Statistics or 36-705 Intermediate Statistics which is much more mathematically rigorous. This option should be considered by strong Statistics Majors who are also majoring in Computer Science, Operations Research, or Mathematics and/or who are considering graduate study in Statistics. This option does require special permission from the advisor. Students who end up satisfying the theory requirement by taking either 36-700 or 36-705 are required to take an additional statistics elective (see category #6, Statistical Electives, below).

**4. Statistical Computing: 9 units**

36-350 Statistical Computing\* 9

\*In rare circumstances, a higher level Statistical Computing course, approved by your Statistics advisor, may be used as a substitute.

**5. Special Topics 9 units**

The Department of Statistics and Data Science offers advanced courses that focus on specific statistical applications or advanced statistical methods. These courses are numbered 36-46x (36-461, 36-462, etc.). Two of these courses will be offered every year, one per semester. Past topics included Statistical Learning, Data Mining, Statistics and the Law, Bayesian Statistics, Nonparametric Statistics, Statistical Genetics, Multilevel and Hierarchical Models, and Statistical Methods in Epidemiology. The objective of the course is to expose students to important topics in statistics and/or interesting applications which are not part of the standard undergraduate curriculum.

To satisfy the Special Topics requirement choose *one* of the **36-46x** courses (which are 9 units).

Note: All 36-46x courses require 36-401 as a prerequisite or, in rare cases, instructor permission.

**6. Statistical Elective: 9–10 units**

Students are required to take one\* elective which can be within or outside the Department of Statistics and Data Science. **Courses within Statistics** can be any 300 or 400 level course (that is not used to satisfy any other requirement for the statistics major).

The following is a partial list of **courses outside Statistics** that qualify as electives as they provide intellectual infrastructure that will advance the student's understanding of statistics and its applications. Other courses may qualify as well; consult with the Statistics Undergraduate Advisor.

10-601	Introduction to Machine Learning (Master's)	12
15-110	Principles of Computing	10
15-112	Fundamentals of Programming and Computer Science	12
15-121	Introduction to Data Structures	10
15-122	Principles of Imperative Computation	10
15-388	Practical Data Science	9
21-127	Concepts of Mathematics	10
21-260	Differential Equations	9
21-292	Operations Research I	9
21-301	Combinatorics	9
21-355	Principles of Real Analysis I	9
80-220	Philosophy of Science	9
80-221	Philosophy of Social Science	9
80-310	Formal Logic	9
85-310	Research Methods in Cognitive Psychology	9
85-320	Research Methods in Developmental Psychology	9
85-340	Research Methods in Social Psychology	9
88-223	Decision Analysis	9
88-302	Behavioral Decision Making	9

Note: Additional prerequisites are required for some of these courses. Students should carefully check the course descriptions to determine if additional prerequisites are necessary.

\* Students who end up satisfying the theory requirement using 36-700 or 36-705 are required to take *two* electives only one of which can be outside the Department of Statistics and Data Science. (In general, any waived requirement is replaced by a statistical elective.)

**7. Tracks\*:**

Self-Defined Concentration Area (with advisor's approval) 36 units

The power of Statistics, and much of the fun, is that it can be applied to answer such a wide variety of questions in so many different fields. A critical part of statistical practice is understanding the questions being asked so that appropriate methods of analysis can be used. Hence, a critical part of statistical training is to gain experience applying the abstract tools to real problems.

The Concentration Area is a set of four related courses outside of Statistics that prepares the student to deal with statistical aspects of problems that arise in another field. These courses are usually drawn from a *single* discipline of interest to the student and must be approved by the Statistics

Undergraduate Advisor. While these courses are not in Statistics, the concentration area must compliment the overall Statistics degree.

For example, students intending to pursue careers in the health or biomedical sciences could take further courses in Biology or Chemistry, or students intending to pursue graduate work in Statistics could take further courses in advanced Mathematics.

The concentration area can be fulfilled with a minor or additional major, but *not all minors and additional majors fulfill this requirement*. Please make sure to consult the Undergraduate Statistics Advisor *prior* to pursuing courses for the concentration area. Once the concentration area is approved, any changes made to the previously agreed upon coursework requires re-approval by the Undergraduate Advisor.

**Concentration Approval Process**

- Submit the below materials to the Undergraduate Statistics Advisor
  - List of possible coursework to fulfill the concentration\*
  - 150-200 word essay describing how the proposed courses complement the Statistics degree.

\* These courses can be amended later, but must be re-approved by the Statistics Undergraduate Advisor.

**Mathematical Statistics Track 46–52 units**

21-127	Concepts of Mathematics	10
21-355	Principles of Real Analysis I	9
36-410	Introduction to Probability Modeling	9

And *two* of the following:

36-700	Probability and Mathematical Statistics	12
or 36-705	Intermediate Statistics	
21-228	Discrete Mathematics	9
21-257	Models and Methods for Optimization	9
21-292	Operations Research I	9
21-301	Combinatorics	9
21-356	Principles of Real Analysis II	9

**Statistics and Neuroscience Track 45–54 units**

85-211	Cognitive Psychology	9
85-219	Biological Foundations of Behavior	9

And three electives (at least one from Methodology and Analysis and at least one from Neuroscientific Background):

**Methodology and Analysis**

36-700	Probability and Mathematical Statistics	12
or 36-705	Intermediate Statistics	
10-601	Introduction to Machine Learning (Master's)	12
18-290	Signals and Systems	12
85-314	Cognitive Neuroscience Research Methods	9
42/86-631	Neural Data Analysis	9

**Neuroscientific Background**

03-362	Cellular Neuroscience	9
03-363	Systems Neuroscience	9
15-386	Neural Computation	9
85-414	Cognitive Neuropsychology	9
85-419	Introduction to Parallel Distributed Processing	9

\* Note: The concentration/track requirement is only for students whose *primary* major is statistics and have no other additional major or minor. The requirement does not apply for students who pursue an *additional* major in statistics.

**Total Number of Units for the Major: 146-185\***

**Total Number of Units for the Degree: 360**

\* Note: This number can vary depending on the calculus sequence and on the concentration area a student takes. In addition this number includes the 36 units of the "Concentration Area" category which may not be required (see category 7 above for details).

## Recommendations

Students in the College of Humanities and Social Sciences who wish to major or minor in Statistics are advised to complete both the calculus requirement (one Mathematical Foundations calculus sequence) and the Beginning Data Analysis course 36-200 or 36-201 by the end of their Freshman year.

The linear algebra requirement is a prerequisite for the course 36-401. It is therefore essential to complete this requirement during your junior year at the latest.

## Recommendations for Prospective PhD Students

Students interested in pursuing a PhD in Statistics or Biostatistics (or related programs) after completing their undergraduate degree are strongly recommended to pursue the **Mathematical Statistics Track**.

## Additional Major in Statistics

Students who elect Statistics as a second or third major must fulfill all Statistics degree requirements except for the Concentration Area requirement. Majors in many other programs would naturally complement a Statistics Major, including Tepper's undergraduate business program, Social and Decision Sciences, Policy and Management, and Psychology.

With respect to double-counting courses, it is departmental policy that students must have at least five statistics courses that do not count for their primary major. If students do not have at least five, they typically take additional advanced electives.

Students are advised to begin planning their curriculum (with appropriate advisors) as soon as possible. This is particularly true if the other major has a complex set of requirements and prerequisites or when many of the other major's requirements overlap with the requirements for a Major in Statistics.

Many departments require Statistics courses as part of their Major or Minor programs. Students seeking transfer credit for those requirements from substitute courses (at Carnegie Mellon or elsewhere) should seek permission from their advisor in the department setting the requirement. The final authority in such decisions rests there. The Department Statistics does not provide approval or permission for substitution or waiver of another department's requirements.

If a waiver or substitution is made in the home department, it is not automatically approved in the Department of Statistics and Data Science. In many of these cases, the student will need to take additional courses to satisfy the Statistics major requirements. Students should discuss this with a Statistics advisor when deciding whether to add an additional major in Statistics.

## Research

One goal of the Statistics program is to give students experience with statistical research. There is a wide variety of ongoing research projects in the department, and students have several opportunities to get involved in a project that interests them.

Before graduation, students are encouraged to participate in a research project under faculty supervision. Students can do this through projects in specific courses (such as 36-303), through an independent study, or through a summer research position.

Qualified students are also encouraged to participate in an advanced research project through 36-490 Undergraduate Research or independent study under the supervision of a Statistics faculty advisor. Students who maintain a quality point average of 3.25 overall may also apply to participate in the Dietrich College Senior Honors Program (<http://coursecatalog.web.cmu.edu/dietrichcollegeofhumanitiesandsocialsciences/#collegeservicesandprograms>).

## Sample Programs

The following sample programs illustrate three (of many) ways to satisfy the requirements of the Statistics Major. However, keep in mind that the program is flexible enough to support *many* other possible schedules and to emphasize a wide variety of interests.

The first schedule uses calculus sequence 1.

The second schedule is an example of the case when a student enters the program through 36-225 and 36-226 (and therefore skips the beginning

data analysis course). The schedule uses calculus sequence 2, and includes two advanced electives (36-315 and 36-303), both within the Department of Statistics and Data Science. This schedule has more emphasis on statistical theory and probability.

The third schedule is an example of the Mathematical Statistics track.

In these schedules, C.A. refers to Concentration Area courses.

### Schedule 1

Freshman		Sophomore	
Fall	Spring	Fall	Spring
36-200 Reasoning with Data	36-202 Methods for Statistics and Data Science	21-256 Multivariate Analysis	36-315 Statistical Graphics and Visualization
21-111 Differential Calculus	21-112 Integral Calculus	C.A.	21-240 Matrix Algebra with Applications

Junior		Senior	
Fall	Spring	Fall	Spring
36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference	36-401 Modern Regression	36-402 Advanced Methods for Data Analysis
36-350 Statistical Computing	C.A.	C.A.	36-46x Special Topics C.A.

### Schedule 2

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-120 Differential and Integral Calculus	21-256 Multivariate Analysis	36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference
			21-240 Matrix Algebra with Applications

Junior		Senior	
Fall	Spring	Fall	Spring
36-350 Statistical Computing	36-315 Statistical Graphics and Visualization	C.A.	C.A.
36-401 Modern Regression	36-402 Advanced Methods for Data Analysis	36-46x Special Topics	36-303 Sampling, Survey and Society
C.A.	C.A.		

### Schedule 3 - Mathematics Track Only

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-120 Differential and Integral Calculus	21-256 Multivariate Analysis	36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference
	21-260 Differential Equations	21-127 Concepts of Mathematics	21-241 Matrices and Linear Transformations

Junior		Senior	
Fall	Spring	Fall	Spring
36-350 Statistical Computing	36-315 Statistical Graphics and Visualization	36-46x Special Topics	36-410 Introduction to Probability Modeling
36-401 Modern Regression	36-402 Advanced Methods for Data Analysis	21-355 Principles of Real Analysis I	36-303 Sampling, Survey and Society
21-228 Discrete Mathematics	21-341 Linear Algebra		

## B.S. in Economics and Statistics

Academic Advisor: Samantha Nielsen

Faculty Advisors: Rebecca Nugent and Edward Kennedy

Executive Director, Undergraduate Economics Program: Carol Goldberg

Associate Director, Undergraduate Economics Program: Kathleen Conway

Office: Baker Hall 132

Email: [statadvising@stat.cmu.edu](mailto:statadvising@stat.cmu.edu)

The Major in Economics and Statistics provides an interdisciplinary course of study aimed at students with a strong interest in the empirical analysis of economic data. With joint curriculum from the Department of Statistics and Data Science and the Undergraduate Economics Program, the major provides students with a solid foundation in the theories and methods of both fields. Students in this major are trained to advance the understanding of economic issues through the analysis, synthesis

and reporting of data using the advanced empirical research methods of statistics and econometrics. Graduates are well positioned for admission to competitive graduate programs, including those in statistics, economics and management, as well as for employment in positions requiring strong analytic and conceptual skills - especially those in economics, finance, education, and public policy.

All economics courses counting towards an economics degree must be completed with a grade of "C" or higher.

The requirements for the B.S. in Economics and Statistics are the following:

**I. Prerequisites 38-39 units**

**1. Mathematical Foundations 38-39 units**

**Calculus**

21-120 Differential and Integral Calculus 10

and one of the following three:

21-122 Integration and Approximation 10

21-127 Concepts of Mathematics 10

21-257 Models and Methods for Optimization 9

and one of the following:

21-256 Multivariate Analysis 9

21-259 Calculus in Three Dimensions 9

Note: Passing the MSC 21-120 assessment test is an acceptable alternative to completing 21-120.

Note: Taking both 21-111 and 21-112 is equivalent to 21-120. The Mathematical Foundations total is then 48-49 units. The Economics and Statistics major would then total 201-211 units.

**Linear Algebra**

One of the following three courses:

21-240 Matrix Algebra with Applications 10

21-241 Matrices and Linear Transformations 10

21-242 Matrix Theory 10

Note: 21-241 and 21-242 are intended only for students with a very strong mathematical background.

**II. Foundations 18-36 units**

**2. Economics Foundations 18 units**

73-102 Principles of Microeconomics 9

73-103 Principles of Macroeconomics 9

**3. Statistical Foundations 9-18 units**

Sequence 1 (For students beginning their freshman or sophomore year)

**Beginning\***

Choose one of the following courses

36-200 Reasoning with Data 9

36-201 Statistical Reasoning and Practice 9

36/70-207 Probability and Statistics for Business Applications 9

36-220 Engineering Statistics and Quality Control 9

36-247 Statistics for Lab Sciences 9

\*Or extra data analysis course in Statistics

Note: Students who enter the program with 36-225 or 36-226 should discuss options with an advisor. Any 36-300 or 36-400 level course in Data Analysis that does not satisfy any other requirement for the Economics and Statistics Major may be counted as a Statistical Elective.

**Intermediate\***

Choose one of the following courses:

36-202 Methods for Statistics and Data Science \*\* 9

36-208 Regression Analysis 9

36-309 Experimental Design for Behavioral and Social Sciences 9

\*Or extra data analysis course in Statistics

\*\*Must take prior to 36-401

Sequence 2 (For students beginning later in their college career)

**Advanced**

Choose one of the following courses:

36-303 Sampling, Survey and Society 9

36-315 Statistical Graphics and Visualization 9

36-461 Special Topics: Statistical Methods in Epidemiology 9

36-462 Special Topics: Data Mining 9

36-463 Special Topics: Multilevel and Hierarchical Models 9

36-464 Special Topics: Applied Multivariate Methods 9

36-490 Undergraduate Research 9

\*\*Special Topics rotate and new ones are regularly added.

**III. Disciplinary Core 126 units**

**1. Economics Core 45 units**

73-230 Intermediate Microeconomics 9

73-240 Intermediate Macroeconomics 9

73-270 Strategic Professional Communication for Economists 9

73-274 Econometrics I 9

73-374 Econometrics II 9

**2. Statistics Core 36 units**

36-225 Introduction to Probability Theory \*\* 9

and one of the following two courses:

36-226 Introduction to Statistical Inference \* 9

36-326 Mathematical Statistics (Honors) \* 9

and both of the following two courses:

36-401 Modern Regression \* 9

36-402 Advanced Methods for Data Analysis 9

\*In order to be a major in good standing, a grade of C or better is required in 36-225 (or equivalents), 36-226 or 36-326 and 36-401. Otherwise you will not be allowed to continue in the major.

#It is possible to substitute 36-217 or 21-325 for 36-225. (36-225 is the standard introduction to probability, 36-217 is tailored for engineers and computer scientists, and 21-325 is a rigorous Probability Theory course offered by the Department of Mathematics.)

**3. Computing 9 units**

36-350 Statistical Computing \* 9

\*In rare circumstances, a higher level Statistical Computing course, approved by your Statistics advisor, may be used as a substitute.

**4. Advanced Electives 36 units**

Students must take two advanced Economics elective courses (numbered 73-300 through 73-495, excluding 73-374 ) and two advanced Statistics elective courses (numbered 36-303, 36-315, or 36-410 through 36-495).

Students pursuing a degree in Economics and Statistics also have the option of earning a concentration area by completing a set of interconnected electives. While a concentration area is not required for this degree, this is an additional option that allows students to pursue courses that are aligned with a particular career path. The two electives that are already required for this degree could count towards your concentration area, please make sure to consult an advisor when choosing these courses.

**Total number of units for the major 191-201 units**

**Total number of units for the degree 360 units**

**Professional Development**

Students are strongly encouraged to take advantage of professional development opportunities and/or coursework. One option is 73-210 Economics Colloquium I, a fall-only course that provides information about careers in Economics, job search strategies, and research opportunities. The Department of Statistics and Data Science also offers a series of workshops

pertaining to resume preparation, graduate school applications, careers in the field, among other topics. Students should also take advantage of the Career and Professional Development Center.

## Additional Major in Economics and Statistics

Students who elect Economics and Statistics as a second or third major must fulfill all Economics and Statistics degree requirements. Majors in many other programs would naturally complement an Economics and Statistics Major, including Tepper's undergraduate business program, Social and Decision Sciences, Policy and Management, and Psychology.

With respect to double-counting courses, it is departmental policy that students must have at least six courses (three Economics and three Statistics) that do not count for their primary major. If students do not have at least six, they typically take additional advanced electives.

Students are advised to begin planning their curriculum (with appropriate advisors) as soon as possible. This is particularly true if the other major has a complex set of requirements and prerequisites or when many of the other major's requirements overlap with the requirements for a Major in Economics and Statistics.

Many departments require Statistics courses as part of their Major or Minor programs. Students seeking transfer credit for those requirements from substitute courses (at Carnegie Mellon or elsewhere) should seek permission from their advisor in the department setting the requirement. The final authority in such decisions rests there. The Department of Statistics and Data Science does not provide approval or permission for substitution or waiver of another department's requirements.

If a waiver or substitution is made in the home department, it is not automatically approved in the Department of Statistics and Data Science. In many of these cases, the student will need to take additional courses to satisfy the Economics and Statistics major requirements. Students should discuss this with a Statistics advisor when deciding whether to add an additional major in Economics and Statistics.

## Sample Program

The following sample program illustrates one way to satisfy the requirements of the Economics and Statistics Major. Keep in mind that the program is flexible and can support other possible schedules (see footnotes below the schedule).

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-120 Differential and Integral Calculus	36-202 Methods for Statistics and Data Science	21-122 Integration and Approximation **	21-240 Matrix Algebra with Applications
36-200 Reasoning with Data	21-256 Multivariate Analysis	36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference
73-102 Principles of Microeconomics	73-103 Principles of Macroeconomics	73-230 Intermediate Microeconomics	73-240 Intermediate Macroeconomics
-----*	-----	-----	73-274 Econometrics I
-----	-----	-----	-----

Junior		Senior	
Fall	Spring	Fall	Spring
36-350 Statistical Computing	36-402 Advanced Methods for Data Analysis	Statistics Elective	Economics Elective
36-401 Modern Regression	73-270 Strategic Professional Communication for Economists	Economics Elective	Statistics Elective
73-374 Econometrics II	-----	-----	-----
-----	-----	-----	-----
-----	-----	-----	-----

\*In each semester, ----- represents other courses (not related to the major) which are needed in order to complete the 360 units that the degree requires.

\*\* Students can also take 21-127 or 21-257. Students should consult with their advisor.

Prospective PhD students might add 21-127 fall of sophomore year, replace 21-240 with 21-241, add 21-260 in spring of junior year and 21-355 in fall of senior year.

## B.S. in Statistics and Machine Learning

Academic Advisor: Samantha Nielsen  
 Faculty Advisors: Ryan Tibshirani and Ann Lee  
 Office: Baker Hall 132  
 Email: statadvising@stat.cmu.edu

Students in the Statistics and Machine Learning program develop and master a wide array of skills in computing, mathematics, statistical theory, and the interpretation and display of complex data. In addition, Statistics and Machine Learning majors gain experience in applying statistical tools to real problems in other fields and learn the nuances of interdisciplinary collaboration. This program is geared towards students interested in statistical computation, data science, or "Big Data" problems. The requirements for the Major in Statistics and Machine Learning are detailed below and are organized by categories.

## Curriculum

### 1. Mathematical Foundations (Prerequisites) 49–59 units

Mathematics is the language in which statistical models are described and analyzed, so some experience with basic calculus and linear algebra is an important component for anyone pursuing a program of study in Statistics and Machine Learning.

Calculus\*:

Complete one of the following sequences of mathematics courses at Carnegie Mellon, each of which provides sufficient preparation in calculus:

#### Sequence 1

21-111	Differential Calculus	10
21-112	Integral Calculus	10

and *one* of the following:

21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

#### Sequence 2

21-120	Differential and Integral Calculus	10
--------	------------------------------------	----

and *one* of the following:

21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

#### Notes:

- Other sequences are possible, and require approval from the undergraduate advisor.
- Passing the Mathematical Sciences 21-120 assessment test is an acceptable alternative to completing 21-120

#### Integration and Approximation

21-122	Integration and Approximation	10
--------	-------------------------------	----

#### Linear Algebra\*\*:

Complete *one* of the following three courses:

21-240	Matrix Algebra with Applications	10
21-241	Matrices and Linear Transformations	10
21-242	Matrix Theory	10

\* It is recommended that students complete the calculus requirement during their freshman year.

\*\*The linear algebra requirement needs to be completed before taking 36-401 Modern Regression.

21-241 and 21-242 are intended only for students with a very strong mathematical background.

#### Mathematical Theory:

21-127	Concepts of Mathematics	10
--------	-------------------------	----

## 2. Data Analysis 45–54 units

Data analysis is the art and science of extracting insight from data. The art lies in knowing which displays or techniques will reveal the most interesting features of a complicated data set. The science lies in understanding the various techniques and the assumptions on which they rely. Both aspects require practice to master.

The Beginning Data Analysis courses give a hands-on introduction to the art and science of data analysis. The courses cover similar topics but differ slightly in the examples they emphasize. 36-200 and 36-201 draw examples from many fields and satisfy the Dietrich College Core Requirement in Statistical Reasoning. One of these courses is therefore recommended for students in the College. (Note: A score of 4 or 5 on the Advanced Placement (AP) Exam in Statistics may be used to waive this requirement). Other courses emphasize examples in business (36-207), engineering and architecture (36-220), and the laboratory sciences (36-247).

The Intermediate Data Analysis courses build on the principles and methods covered in the introductory course, and more fully explore specific types of data analysis methods in more depth.

The Advanced Data Analysis courses draw on students' previous experience with data analysis and understanding of statistical theory to develop advanced, more sophisticated methods. These core courses involve extensive analysis of real data with emphasis on developing the oral and writing skills needed for communicating results.

### Sequence 1

#### Beginning\*

Choose one of the following courses:

36-200	Reasoning with Data	9
36-201	Statistical Reasoning and Practice	9
36/70-207	Probability and Statistics for Business Applications	9
36-220	Engineering Statistics and Quality Control	9
36-247	Statistics for Lab Sciences	9

\*Or extra data analysis course in Statistics

Note: Students who enter the program with 36-225 or 36-226 should discuss options with an advisor. Any 36-300 or 36-400 level course in Data Analysis that does not satisfy any other requirement for a Statistics Major and Minor may be counted as a Statistical Elective.

#### Intermediate\*

Choose one of the following courses:

36-202	Methods for Statistics and Data Science **	9
36/70-208	Regression Analysis	9
36-309	Experimental Design for Behavioral and Social Sciences	9

\*Or extra data analysis course in Statistics

\*\*Must take prior to 36-401

#### Advanced

Choose two of the following courses:

36-303	Sampling, Survey and Society	9
36-315	Statistical Graphics and Visualization	9
36-461	Special Topics: Statistical Methods in Epidemiology	9
36-462	Special Topics: Data Mining	9
36-463	Special Topics: Multilevel and Hierarchical Models	9
36-464	Special Topics: Applied Multivariate Methods	9
36-490	Undergraduate Research	9

\*\*Special Topics rotate and new ones are regularly added.

and take the following two courses:

36-401	Modern Regression	9
36-402	Advanced Methods for Data Analysis	9

### Sequence 2

#### Advanced

Choose three of the following courses:

36-303	Sampling, Survey and Society	9
36-315	Statistical Graphics and Visualization	9

36-461	Special Topics: Statistical Methods in Epidemiology	9
36-462	Special Topics: Data Mining	9
36-463	Special Topics: Multilevel and Hierarchical Models	9
36-464	Special Topics: Applied Multivariate Methods	9
36-490	Undergraduate Research	9

\*\*Special Topics rotate and new ones are regularly added.

and take the following two courses:

36-401	Modern Regression	9
36-402	Advanced Methods for Data Analysis	9

## 3. Probability Theory and Statistical Theory 18 units

The theory of probability gives a mathematical description of the randomness inherent in our observations. It is the language in which statistical models are stated, so an understanding of probability is essential for the study of statistical theory. Statistical theory provides a mathematical framework for making inferences about unknown quantities from data. The theory reduces statistical problems to their essential ingredients to help devise and evaluate inferential procedures. It provides a powerful and wide-ranging set of tools for dealing with uncertainty.

To satisfy the theory requirement take the following two courses\*\*:

36-225	Introduction to Probability Theory	9
36-226	Introduction to Statistical Inference	9
or 36-326	Mathematical Statistics (Honors)	

\*\*It is possible to substitute 36-217 or 21-325 for 36-225. (36-225 is the standard introduction to probability, 36-217 is tailored for engineers and computer scientists, and 21-325 is a rigorous Probability Theory course offered by the Department of Mathematics.) 36-326 Mathematical Statistics (Honors) can be substituted for 36-226 Introduction to Statistical Inference and is considered an honors course.

#### Comments:

(i) In order to be a Major or a Minor in good standing, a grade of at least a C is required in 36-225, 36-226 and 36-401. In particular, a grade of C or higher is required in order to be able to continue in the major.

(ii) In special cases, and in consultation with the Statistics Advisor, the theory requirement can be satisfied by taking a single *graduate level* class 36-700 Probability and Mathematical Statistics or 36-705 Intermediate Statistics which is much more mathematically rigorous. This option should be considered by strong Statistics Majors who are also majoring in Computer Science, Operations Research, or Mathematics and/or who are considering graduate study in Statistics. This option does require special permission from the advisor. Students who end up satisfying the theory requirement by taking either 36-700 or 36-705 are required take an additional statistics elective.

## 4. Computing

## 64–67 units

Statistical modeling in practice nearly always requires computation in one way or another. Computational algorithms are sometimes treated as “black-boxes”, whose innards the statistician need not pay attention to. But this attitude is becoming less and less prevalent, and today there is much to be gained from a strong working knowledge of computational tools. Understanding the strengths and weaknesses of various methods allows the data analyst to select the right tool for the job; understanding how they can be adapted to work in new settings greatly extends the realm of problems that he/she can solve. While all Majors in Statistics are given solid grounding in computation, extensive computational training is really what sets the Major in Statistics and Machine Learning apart.

36-350	Statistical Computing *	9
15-112	Fundamentals of Programming and Computer Science	12
15-122	Principles of Imperative Computation	10
15-351	Algorithms and Advanced Data Structures	12
10-601	Introduction to Machine Learning (Master's)	12
or 10-401	Introduction to Machine Learning (Undergrad)	

\*In rare circumstances, a higher level Statistical Computing course, approved by your Statistics advisor, may be used as a substitute.

and take one of the following courses:

10-605	Machine Learning with Large Datasets	12
15-381	Artificial Intelligence: Representation and Problem Solving	9

15-386	Neural Computation	9
16-720	Computer Vision	12
16-311	Introduction to Robotics	12
11-411	Natural Language Processing	12
11-761	Language and Statistics	12

\*PhD level ML course as approved by Statistics advisor  
 \*\* Independent research with an ML faculty member

<b>Total number of units for the major</b>	<b>176–198 units</b>
<b>Total number of units for the degree</b>	<b>360 units</b>

## Recommendations

Students in the Dietrich College of Humanities and Social Sciences who wish to major or minor in Statistics are advised to complete both the calculus requirement (one Mathematical Foundations calculus sequence) and the Beginning Data Analysis course 36-200 Reasoning with Data or 36-201 Statistical Reasoning and Practice by the end of their Freshman year.

The linear algebra requirement is a prerequisite for the course 36-401. It is therefore essential to complete this requirement during your junior year at the latest!

## Recommendations for Prospective PhD Students

Students interested in pursuing a PhD in Statistics or Machine Learning (or related programs) after completing their undergraduate degree are strongly recommended to take additional Mathematics courses. They should see a faculty advisor as soon as possible. Students should consider 36-326 Mathematical Statistics (Honors) as an alternative to 36-226. Although 21-240 Matrix Algebra with Applications is recommended for Statistics majors, students interested in PhD programs should consider taking 21-241 Matrices and Linear Transformations or 21-242 Matrix Theory instead. Additional courses to consider are 21-228 Discrete Mathematics, 21-260 Differential Equations, 21-341 Linear Algebra, 21-355 Principles of Real Analysis I, and 21-356 Principles of Real Analysis II.

Additional experience in programming and computational modeling is also recommended. Students should consider taking more than one course from the list of Machine Learning electives provided under the Computing section.

## Additional Major in Statistics and Machine Learning

Students who elect Statistics and Machine Learning as a second or third major must fulfill *all* degree requirements.

With respect to double-counting courses, it is departmental policy that students must have at least six courses (three Computer Science/Machine Learning and three Statistics) that do not count for their primary major. If students do not have at least six, they typically take additional advanced electives.

Students are advised to begin planning their curriculum (with appropriate advisors) as soon as possible. This is particularly true if the other major has a complex set of requirements and prerequisites or when many of the other major's requirements overlap with the requirements for a Major in Statistics and Machine Learning.

Many departments require Statistics courses as part of their Major or Minor programs. Students seeking transfer credit for those requirements from substitute courses (at Carnegie Mellon or elsewhere) should seek permission from their advisor in the department setting the requirement. The final authority in such decisions rests there. The Department of Statistics and Data Science does not provide approval or permission for substitution or waiver of another department's requirements.

If a waiver or substitution is made in the home department, it is not automatically approved in the Department of Statistics and Data Science. In many of these cases, the student will need to take additional courses to satisfy the Statistics and Machine Learning major requirements. Students should discuss this with a Statistics advisor when deciding whether to add an additional major in Statistics and Machine Learning.

## Sample Programs

The following sample program illustrates one way to satisfy the requirements of the Statistics and Machine Learning program. Keep in mind that the program is flexible and can support other possible schedules (see footnotes below the schedule). Sample program 1 is for students who have not satisfied the basic calculus requirements. Sample program 2 is for students who have satisfied the basic calculus requirements and choose option 2 for their data analysis courses (see section #2)

### Schedule 1

Freshman		Sophomore	
Fall	Spring	Fall	Spring
36-200 Reasoning with Data	36-202 Methods for Statistics and Data Science	36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference
21-120 Differential and Integral Calculus	21-256 Multivariate Analysis	21-122 Integration and Approximation	21-241 Matrices and Linear Transformations
15-112 Fundamentals of Programming and Computer Science	15-112 Fundamentals of Programming and Computer Science	21-127 Concepts of Mathematics	15-122 Principles of Imperative Computation
-----*	-----	-----	-----
-----	-----	-----	-----

Junior		Senior	
Fall	Spring	Fall	Spring
15-351 Algorithms and Advanced Data Structures	36-402 Advanced Methods for Data Analysis	10-601 Introduction to Machine Learning (Master's)	10-605 Machine Learning with Large Datasets
36-401 Modern Regression	Stat Elective	Stat Elective	ML Elective
-----	-----	-----	-----
-----	-----	-----	-----
-----	-----	-----	-----

\*In each semester, ----- represents other courses (not related to the major) which are needed in order to complete the 360 units that the degree requires.

### Schedule 2

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-256 Multivariate Analysis	15-122 Principles of Imperative Computation	36-217 Probability Theory and Random Processes	36-226 Introduction to Statistical Inference
15-112 Fundamentals of Programming and Computer Science	21-127 Concepts of Mathematics	15-351 Algorithms and Advanced Data Structures	21-241 Matrices and Linear Transformations
-----*	-----	-----	Stat Elective
-----	-----	-----	-----
-----	-----	-----	-----

Junior		Senior	
Fall	Spring	Fall	Spring
36-350 Statistical Computing	36-402 Advanced Methods for Data Analysis	10-601 Introduction to Machine Learning (Master's)	10-605 Machine Learning with Large Datasets
36-401 Modern Regression	Stat Elective	Stat Elective	ML Elective
-----	-----	-----	-----
-----	-----	-----	-----
-----	-----	-----	-----

\*In each semester, ----- represents other courses (not related to the major) which are needed in order to complete the 360 units that the degree requires.

## The Minor in Statistics

Academic Advisor: Samantha Nielsen  
 Faculty Advisor: Peter Freeman  
 Office: Baker Hall 132M  
 Email: statadvising@stat.cmu.edu

The Minor in Statistics develops skills that complement major study in other disciplines. The program helps the student master the basics of statistical theory and advanced techniques in data analysis. This is a good choice for



deepening understanding of statistical ideas and for strengthening research skills.

In order to get a minor in Statistics a student must satisfy all of the following requirements:

**1. Mathematical Foundations (Prerequisites)**  
**29–39 units**

Calculus\*:

Complete *one* of the following two sequences of mathematics courses at Carnegie Mellon, each of which provides sufficient preparation in calculus:

**Sequence 1**

21-111	Differential Calculus	10
21-112	Integral Calculus	10

and *one* of the following:

21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

**Sequence 2**

21-120	Differential and Integral Calculus	10
--------	------------------------------------	----

and *one* of the following:

21-256	Multivariate Analysis	9
21-259	Calculus in Three Dimensions	9

Note: Other sequences are possible, and require approval from the undergraduate advisor.

Note: Passing the Mathematical Sciences 21-120 assessment test if an acceptable alternative to completing 21-120.

Linear Algebra:

Complete *one* of the following three courses:

21-240	Matrix Algebra with Applications	10
21-241	Matrices and Linear Transformations	10
21-242	Matrix Theory	10

\*It is recommended that students complete the calculus requirement during their freshman year.

\*\*The linear algebra requirement needs to be complete before taking 36-401 Modern Regression or 36-46X Special Topics.

21-241 and 21-242 are intended only for students with a very strong mathematical background.

**2. Data Analysis** **36 units**

Data analysis is the art and science of extracting insight from data. The art lies in knowing which displays or techniques will reveal the most interesting features of a complicated data set. The science lies in understanding the various techniques and the assumptions on which they rely. Both aspects require practice to master.

The Beginning Data Analysis courses give a hands-on introduction to the art and science of data analysis. The courses cover similar topics but differ slightly in the examples they emphasize. 36-200 or 36-201 draw examples from many fields and satisfy the DC College Core Requirement in Statistical Reasoning. One of these courses is therefore recommended for students in the College. (Note: A score of 4 or 5 on the Advanced Placement (AP) Exam in Statistics may be used to waive this requirement). Other courses emphasize examples in business (36-207), engineering and architecture (36-220), and the laboratory sciences (36-247).

The Intermediate Data Analysis courses build on the principles and methods covered in the introductory course, and more fully explore specific types of data analysis methods in more depth.

The Advanced Data Analysis courses draw on students' previous experience with data analysis and understanding of statistical theory to develop advanced, more sophisticated methods. These core courses involve extensive analysis of real data with emphasis on developing the oral and writing skills needed for communicating results.

Sequence 1 (For students beginning their freshman or sophomore year)

**Beginning Data Analysis\***

Choose *one* of the following courses:

36-200	Reasoning with Data	9
36-201	Statistical Reasoning and Practice	9
36/70-207	Probability and Statistics for Business Applications	9
36-220	Engineering Statistics and Quality Control	9
36-247	Statistics for Lab Sciences	9

\*Or extra data analysis course in Statistics

Note: Students who enter the program with 36-225 or 36-226 should discuss options with an advisor. Any 36-300 or 36-400 level course in Data Analysis that does not satisfy any other requirement for a Statistics Major and Minor may be counted as a Statistical Elective.

**Intermediate Data Analysis\***

Choose *one* of the following courses:

36-202	Methods for Statistics and Data Science **	9
36/70-208	Regression Analysis	9
36-309	Experimental Design for Behavioral and Social Sciences	9

\*Or extra data analysis course in Statistics

\*\*Must take prior to 36-401

**Advanced Data Analysis and Methodology**

Take the following course:

36-401	Modern Regression	9
--------	-------------------	---

and *one* of the following courses:

36-402	Advanced Methods for Data Analysis	9
36-410	Introduction to Probability Modeling	9
36-461	Special Topics: Statistical Methods in Epidemiology	9
36-462	Special Topics: Data Mining	9
36-463	Special Topics: Multilevel and Hierarchical Models	9
36-464	Special Topics: Applied Multivariate Methods	9
36-490	Undergraduate Research	9

\*\*Special Topics rotate and new ones are regularly added.

Sequence 2 (For students beginning later in their college career)

**Advanced Data Analysis and Methodology**

Take the following course:

36-401	Modern Regression	9
--------	-------------------	---

and *two* of the following courses (one of which must be 400-level):

36-303	Sampling, Survey and Society	9
36-315	Statistical Graphics and Visualization	9
36-402	Advanced Methods for Data Analysis	9
36-410	Introduction to Probability Modeling	9
36-461	Special Topics: Statistical Methods in Epidemiology	9
36-462	Special Topics: Data Mining	9
36-463	Special Topics: Multilevel and Hierarchical Models	9
36-464	Special Topics: Applied Multivariate Methods	9
36-490	Undergraduate Research	9

\*\*Special Topics rotate and new ones are regularly added.

**3. Probability Theory and Statistical Theory 18 units**

To satisfy the theory requirement take the following *two* courses:

36-225	Introduction to Probability Theory	9
36-226 or 36-326	Introduction to Statistical Inference / Mathematical Statistics (Honors)	9

\*\*It is possible to substitute 36-217 or 21-325 for 36-225. (36-225 is the standard introduction to probability, 36-217 is tailored for engineers and computer scientists, and 21-325 is a rigorous Probability Theory course offered by the Department of Mathematics.) 36-326 Mathematical Statistics (Honors) can be substituted for 36-226 Introduction to Statistical Inference and is considered an honors course.

Comments:

(i) In order to be a Major or a Minor in good standing, a grade of at least a C is required in 36-225 , 36-226 and 36-401. In particular, a grade of C or higher is required in order to be able to continue in the major.

(ii) In special cases, and in consultation with the Statistics Advisor, the theory requirement can be satisfied by taking a single *graduate level* class 36-700 Probability and Mathematical Statistics or 36-705 Intermediate Statistics which is much more mathematically rigorous. This option should be considered by strong Statistics Majors who are also majoring in Computer Science, Operations Research, or Mathematics and/or who are considering graduate study in Statistics. This option does require special permission from the advisor. Students who end up satisfying the theory requirement by taking either 36-700 or 36-705 are required take an additional statistics elective.

**Total number of units required for the minor 83 Units**

With respect to double-counting courses, it is departmental policy that students must have at least three statistics courses that do not count for their primary major. If students do not have at least three, they typically take additional advanced electives.

## Sample Programs for the Minor

The following two sample programs illustrates two (of many) ways to satisfy the requirements of the Statistics Minor. Keep in mind that the program is flexible and can support many other possible schedules.

The first schedule uses calculus sequence 1, and 36-309 to satisfy the intermediate data analysis requirement. The second schedule is an example of the case when a student enters the Minor through 36-225 and 36-226 (and therefore skips the beginning data analysis course). The schedule uses calculus sequence 2, and 36-315 as an elective (to replace the beginning data analysis course).

### Schedule 1

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-111 Differential Calculus	21-112 Integral Calculus	21-256 Multivariate Analysis	21-240 Matrix Algebra with Applications
36-201 Statistical Reasoning and Practice		36-309 Experimental Design for Behavioral and Social Sciences	
Junior		Senior	
Fall	Spring	Fall	Spring
36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference	36-401 Modern Regression	36-402 Advanced Methods for Data Analysis

### Schedule 2

Freshman		Sophomore	
Fall	Spring	Fall	Spring
21-120 Differential and Integral Calculus	21-256 Multivariate Analysis	36-225 Introduction to Probability Theory	36-226 Introduction to Statistical Inference
Junior		Senior	
Fall	Spring	Fall	Spring
21-240 Matrix Algebra with Applications	36-315 Statistical Graphics and Visualization	36-401 Modern Regression	36-462 Special Topics: Data Mining

## Substitutions and Waivers

Many departments require Statistics courses as part of their Major or Minor programs. Students seeking transfer credit for those requirements from substitute courses (at Carnegie Mellon or elsewhere) should seek permission from their advisor in the department setting the requirement. The final authority in such decisions rests there. The Department of Statistics and Data Science does not provide approval or permission for substitution or waiver of another department's requirements.

However, the Statistics Director of Undergraduate Studies will provide advice and information to the student's advisor about the viability of a proposed substitution. Students should make available as much information as possible concerning proposed substitutions. Students seeking waivers may be asked to demonstrate mastery of the material.

If a waiver or substitution is made in the home department, it is not automatically approved in the Department of Statistics and Data Science. In many of these cases, the student will need to take additional courses to satisfy the Statistics major requirements. Students should discuss this with a Statistics advisor when deciding whether to add an additional major in Statistics.

Statistics Majors and Minors seeking substitutions or waivers should speak to the Academic Advisor in Statistics.

## Faculty

DAVID CHOI, Assistant Professor of Statistics and Information Systems – Ph.D., Stanford University; Carnegie Mellon, 2004–.

ALEXANDRA CHOULDECHOVA, Assistant Professor of Statistics and Public Policy – Ph.D., Stanford University; Carnegie Mellon, 2014–.

MAX G'SELL, Assistant Professor – Ph.D., Stanford University ; Carnegie Mellon, 2014–.

CHRISTOPHER R. GENOVESE, Department Head and Professor of Statistics – Ph.D., University of California, Berkeley; Carnegie Mellon, 1994–.

JOEL B. GREENHOUSE, Professor of Statistics – Ph.D., University of Michigan; Carnegie Mellon, 1982–.

AMELIA HAVILAND, Anna Loomis McCandless Professorship of Statistics and Public Policy – Ph.D., Carnegie Mellon University; Carnegie Mellon, 2003–.

JIASHUN JIN, Professor of Statistics – Ph.D., Stanford University; Carnegie Mellon, 2007–.

BRIAN JUNKER, Associate Dean and Professor of Statistics – Ph.D., University of Illinois; Carnegie Mellon, 1990–.

ROBERT E. KASS, Professor of Statistics – Ph.D., University of Chicago; Carnegie Mellon, 1981–.

EDWARD KENNEDY, Assistant Professor – Ph.D., University of Pennsylvania; Carnegie Mellon, 2016–.

ANN LEE, Associate Professor – Ph.D., Brown University; Carnegie Mellon, 2005–.

JOHN P. LEHOCZKY, Thomas Lord Professor of Statistics – Ph.D., Stanford University; Carnegie Mellon, 1969–.

JING LEI, Assistant Professor – Ph.D., University of California, Berkeley; Carnegie Mellon, 2011–.

DANIEL NAGIN, Teresa and H. John Heinz III Professor of Public Policy – Ph.D., Carnegie Mellon University; Carnegie Mellon, 1976–.

MATEY NEYKOV, Assistant Professor – Ph.D., Harvard University; Carnegie Mellon, 2017–.

NYNKE NIEZINK, Assistant Professor – Ph.D., University of Groningen; Carnegie Mellon, 2017–.

REBECCA NUGENT, Associate Department Head, Teaching Professor – Ph.D., University of Washington; Carnegie Mellon, 2006–.

ALESSANDRO RINALDO, Associate Professor – Ph.D., Carnegie Mellon; Carnegie Mellon, 2005–.

KATHRYN ROEDER, Professor of Statistics – Ph.D., Pennsylvania State University; Carnegie Mellon, 1994–.

CHAD M. SCHAFER, Associate Professor – Ph.D., University of California, Berkeley; Carnegie Mellon, 2004–.

MARK J. SCHERVISH, Professor of Statistics – Ph.D., University of Illinois; Carnegie Mellon, 1979–.

TEDDY SEIDENFELD, Herbert A. Simon Professor of Philosophy and Statistics – Ph.D., Columbia University; Carnegie Mellon, 1985–.

COSMA SHALIZI, Associate Professor – Ph.D., University of Wisconsin, Madison; Carnegie Mellon, 2005–.

DALENE STANGL, Teaching Professor – Ph.D., Carnegie Mellon University; Carnegie Mellon, 2017–.

RYAN TIBSHIRANI, Associate Professor – Ph.D., Stanford University; Carnegie Mellon, 2011–.

VALERIE VENTURA, Associate Professor – Ph.D., University of Oxford; Carnegie Mellon, 1997–.

LARRY WASSERMAN, Professor of Statistics – Ph.D., University of Toronto; Carnegie Mellon, 1988–.

## Emeriti Faculty

GEORGE T. DUNCAN, Professor of Statistics and Public Policy – Ph.D., University of Minnesota; Carnegie Mellon, 1974–.

WILLIAM F. EDDY, John C. Warner Professor of Statistics – Ph.D, Yale University; Carnegie Mellon, 1976–.

JOSEPH B. KADANE, Leonard J. Savage Professor of Statistics and Social Sciences – Ph.D., Stanford University; Carnegie Mellon, 1969–.

## Adjunct Faculty

ANTHONY BROCKWELL, – Ph.D., Melbourne University; Carnegie Mellon, 1999–.

BERNIE DEVLIN, – Ph.D., Pennsylvania State University; Carnegie Mellon, 1994–.

APRIL GALYARDT, – Ph.D., Carnegie Mellon University; Carnegie Mellon, 2017–.

ROSS O'CONNELL, .

SAM VENTURA, – Ph.D., Carnegie Mellon University; Carnegie Mellon, 2015–.

## Special Faculty

OLGA CHILINA, Lecturer – MS, University of Toronto; Carnegie Mellon, 2006–.

PETER FREEMAN, Research Associate – Ph.D., University of Chicago; Carnegie Mellon, 2004–.

HOWARD SELTMAN, Senior Research Statistician – Ph.D., Carnegie Mellon; M.D., Medical College of Pennsylvania; Carnegie Mellon, 1999–.

ISABELLA VERDINELLI, Professor in Residence – Ph.D., Carnegie Mellon University; Carnegie Mellon, 1991–.

GORDON WEINBERG, Lecturer – M.A. Mathematics, University of Pittsburgh; Carnegie Mellon, 2004–.