Course: Statistics 6215.10 CRN 75711  
Semester: Fall 2022  
Class Time: Wednesday 6:10 pm-8:40 pm  
Location: ROME 459

Instructor: Prof. Reza Modarres  
Campus Address: Rome Hall 760B  
Office Phone: 202-994-9991  
E-mail: reza@gwu.edu  
Office hour: Wednesday 5-6 pm, or by appointment.  
Grader: 
Email:  
Office Hours:

Course Description:  
This is primarily a lecture course designed to introduce you to the statistical analysis of several variables, most likely dependent, following a joint normal distribution. Stat 6215 reworks much of the material in Stat 4157—4158 using matrices and vectors (chapters 1-5, parts of 6, 8 and 11). Additional topics from the literature will also be covered. The computational aspects will include the use of R and SAS/IML.

- Matrix Algebra and Random Vectors  
- Multivariate Sample Geometry  
- The Multivariate Normal Distribution  
- Inferences about a Mean Vector  
- Comparisons of Several Population Means  
- Principal component Analysis  
- Classification

LEARNING OUTCOMES:

1. Derive properties of the multivariate normal distribution.  
2. Analyze observations obtained from a multivariate normal distribution.  
3. Make inferences about the mean vector.  
4. Read, analyze and synthesize further methodology not covered in class.
Class instructions: 2 hours and half weekly
Homework and reading: average of 8 hours weekly


Grade: Your grade will be based on:
Homework 25%
Midterm 30%
Two Quizzes 15%
Final 30%

Prerequisite: Stat 4157, 4158 and Math 2184

R and SAS will be used and the computational aspects will include use of matrix algebra tools. You are expected to be familiar with R or SAS. GW labs provide access to SAS. To obtain a copy of SAS, see https://itl.gwu.edu/sas-software-distribution. To download R, see https://cran.r-project.org/

CLASS POLICIES

Attendance policy: You need to attend class and ask questions. The handouts distributed during the lecture will become available on Blackboard.

Late work: will not be accepted.

Make-up exams: Makeup examinations will only be given in exceptional circumstances (e.g. well documented medical emergency).

Homework: There will be several homework sets. Each problem counts 10 points. All graded work will be returned and discussed in class. Late submissions will not be accepted. You are expected to work individually on each problem set.

Homework is assigned to make sure you keep up with the work. You should do the homework and ask questions in class. I am available to answer questions during class, after class, during my office hours, by email, and by appointment. If you feel you do not understand a concept or are falling
behind, let me know as soon as possible and ask for help. I will also give you a reading list for each chapter. You need to read the material before coming to class and ask questions when I go over the material.

**Quizzes and Exams:** Quizzes and Exams are closed notes, but open book

**Blackboard**
I will upload all, handouts, examples, project descriptions, old exams, etc. on the Blackboard.

**SAS:** To obtain SAS for installation on their computers, students should visit https://itl.gwu.edu/sas-software-distribution

Please be inform that SAS can only be installed on a Windows OS.

**VCL:** Virtual Computer Lab (VCL) is a cloud-based service that allows you to run Windows-based software through a web browser. VCL is available only during the virtual learning period to provide special access to certain specialized academic software while students and faculty are away from campus. See https://acadtech.gwu.edu/vcl

To access SAS via VCL visit https://academiccommons.gwu.edu/virtualcomputerlab

**CCAS Cloud:** You can also access SAS (and other software) on CCAS cloud. Visit apps.ccas.gwu.edu
Login using your cloud ID. You may have to download a software called Citrix. You will see a virtual desktop and from there you can run SAS and many other apps. The Columbian College Private Cloud and CCAS Virtual Applications. Many software programs are available through the CCAS Virtual applications. https://ots.columbian.gwu.edu/columbian-college-private-cloud

**Student Blackboard Support**

GW Information Technology Support Center
Phone: 202-994-4948, option 2
Email: ithelp@gwu.edu

Support hours: Monday - Friday, 7:00am - 10:00pm*

*Blackboard support is available 24 hours by calling 202-994-4948, option 2
University policies

University policy on observance of religious holidays
In accordance with University policy, students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance. For details and policy, see “Religious Holidays” at provost.gwu.edu/policies-procedures-and-guidelines

Academic Integrity Code
Academic Integrity is an integral part of the educational process, and GW takes these matters very seriously. Violations of academic integrity occur when students fail to cite research sources properly, engage in unauthorized collaboration, falsify data, and in other ways outlined in the Code of Academic Integrity. Students accused of academic integrity violations should contact the Office of Academic Integrity to learn more about their rights and options in the process. Outcomes can range from failure of assignment to expulsion from the University, including a transcript notation. The Office of Academic Integrity maintains a permanent record of the violation. More information is available from the Office of Academic Integrity at studentconduct.gwu.edu/academic-integrity.

Virtual academic support
A full range of academic support is offered virtually in AY 21-22. See coronavirus.gwu.edu/top-faqs for updates.

Tutoring and course review sessions are offered through Academic Commons in an online format. See academiccommons.gwu.edu/tutoring Writing and research consultations are available online. See academiccommons.gwu.edu/writing-research-help

Coaching, offered through the Office of Student Success, is available in a virtual format. See studentsuccess.gwu.edu/academic-program-support
### Fall Semester 2022

- **Classes Begin**: Monday, August 29
- **Labor Day (no classes)**: Monday, September 5
- **Fall Break (no classes)**: Monday, October 24 - Tuesday, October 25
- **Thanksgiving Break (no classes)**: Wednesday, November 23 - Saturday, November 26
- **Last Day of Classes**: Monday, December 12
- **Reading/Make-Up Day**: Tuesday, December 13
- **Final Examinations**: Wednesday, December 14

### Stat 6215: Tentative Schedule

<table>
<thead>
<tr>
<th>Class</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aug 31</td>
<td>Chapter 1: Aspects of Multivariate Analysis</td>
</tr>
<tr>
<td>2</td>
<td>Sep 7</td>
<td>Chapter 2: Matrix Algebra and Random vectors</td>
</tr>
<tr>
<td>3</td>
<td>Sep 14</td>
<td>Chapter 2: Multivariate Distributions and SVD</td>
</tr>
<tr>
<td>4</td>
<td>Sep 21</td>
<td>Chapter 2: Matrix inequalities and maximization</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Quiz I</strong></td>
</tr>
<tr>
<td>5</td>
<td>Sep 28</td>
<td>Chapter 3: Sample Geometry and Random Samples</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample mean, covariance and correlation</td>
</tr>
<tr>
<td>6</td>
<td>Oct 5</td>
<td>Chapter 4: The Multivariate Normal Distribution</td>
</tr>
<tr>
<td>7</td>
<td>Oct 12</td>
<td>Chapter 4: Sampling Distribution of the mean and covariance</td>
</tr>
<tr>
<td>8</td>
<td>Oct 19</td>
<td>Chapter 4: Sampling from $N(0,\Sigma)$</td>
</tr>
<tr>
<td>9</td>
<td>Oct 26</td>
<td><strong>Midterm Exam 6:10-8:40 P.M.</strong></td>
</tr>
<tr>
<td>10</td>
<td>Nov 2</td>
<td>Chapter 5: Inference about a Mean Vector</td>
</tr>
<tr>
<td>11</td>
<td>Nov 9</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>12</td>
<td>Nov 16</td>
<td>Chapter 5, <strong>Quiz II</strong></td>
</tr>
<tr>
<td></td>
<td>Nov 30</td>
<td>Chapter 8: Principal Component Analysis</td>
</tr>
<tr>
<td>13</td>
<td>Dec 7</td>
<td>Chapter 11; Review for Final</td>
</tr>
<tr>
<td>15</td>
<td>Dec 14</td>
<td><strong>Final Exam 6:10-8:40 P.M. ROME 459</strong></td>
</tr>
</tbody>
</table>

We will also study classification and clustering techniques as time permits.
STAT 6215 OUTLINE

1. Aspects of Multivariate Analysis
   - What is multivariate analysis?
   - Differences with univariate analysis.
   - Founding fathers and their contributions
   - Data visualization of multivariate Observations: Chernoff faces, Parallel plots, Andrews plot
   - SAS IML and R matrix operations

2. Matrix Algebra and Random Vectors
   - Basic properties of vectors and vector spaces, and projections
   - Linear and quadratic form.
   - Positive definite quadratic forms matrices. Spectral decomposition.
   - Rotations and orthogonal matrices. Direction cosines.
   - Sample variance as a quadratic form, geometric interpretation
   - Partitioned vectors and matrices
   - Bivariate distributions, Box-Muller Transformation Singular value Decomposition, Cholesky Decomposition
   - Matrix inequalities and maximization

3. Sample Geometry and Random Sampling
   - The geometry of the sample mean
   - Deviation vectors
   - Mean and variance of the sample mean vector
   - Generalized sample variance and total sample variance
   - Linear combination of variables

4. The Multivariate Normal Distribution
   - The multivariate normal density and its properties
   - P-dimensional Geometry, regression lines and ellipses of constant density
   - Maximum likelihood estimation
   - Properties of the sample mean and covariance from a multivariate normal distribution, large sample properties.
   - Independence of the sample mean vector and covariance matrix
   - General covariance of the mean and variance, lack of correlation and independence
   - Conditional distributions and regression
   - Wishart distribution, Sampling distribution of the correlation coefficient
   - Assessing the normality assumption
   - QQ plot, Kolmogorov-Smirnov, Shapiro-Wilk, Cramer Von Mises, and Anderson-Darling tests, Filliben test
   - Assessing bivariate and multivariate normality
   - Measures of multivariate skewness and kurtosis
   - Box-Cox transformation and its properties
5. Inferences about a Mean Vector
   • Hotelling $T^2$ distribution and the likelihood ratio test
   • Confidence regions and simultaneous comparison of component means
   • Large sample inference about a population mean vector
   • Multivariate Quality control charts
   • Missing observation and the EM algorithm
   • Power of the Hotelling $T^2$ distribution for large $n$
   • Multivariate outlier detection techniques

6. Principal Component Analysis
   • Population principal components
   • Principal components from standardized variables
   • Covariance with special structure
   • Interpretation of principal components
   • Number of principal components
   • Large sample properties of the sample eigenvalues and eigenvectors
   • Testing for equal correlation structure
   • Interpoint distances, properties and applications
   • Graphical comparison of groups of observations