

### 1. Course Information:

Course : EMSE 4765.10 – Data Analysis for Scientists and Engineers  
Semester : Spring 2020 # of Credit Hours: 3.0  
Meeting Time : Tuesday from 12:45PM to 3:15PM  
Location : Tompkins Hall Computer Lab 405

### 2. Instructor and Contact Information:

Name : J. René van Dorp, Professor  
Campus Address : 800 22<sup>nd</sup> Street, Office 2800, Washington DC 20052  
Phone : 202-994-6638  
E-mail : dorpjr@gmail.com  
Office hours : Mondays 2:00PM to 4:00PM

### 3. Course Description:

Inference methods in a single dimension: estimation, confidence intervals, hypothesis testing and goodness-of-fit testing; multivariate data analysis techniques using matrices and vectors: the Hotelling T-squared test, multiple linear regression, one way ANOVA, two-way ANOVA and 2<sup>K</sup> ANOVA.

### 4. Prerequisite Requirement:

APSC 3115: Engineering Analysis III (or any other undergraduate Applied Statistics course from a physical or natural sciences program).

<http://www.seas.gwu.edu/~dorpjr/APSC3115/Intro.html>

### 5. Required Materials:

- a. No Textbook Required.
- b. Electronic Lectures notes available at:  
<http://www2.seas.gwu.edu/~dorpjr/EMSE4765/Coursefiles.html>
- c. MS Excel Software
- d. MINITAB software (available on lab computers) Six months or twelve months rental of the MINITAB Software is available for students at a discounted rate at:  
<http://www.onthehub.com/minitab/>

### 6. Recommended Materials:

- a. Recommended Text Books (two available for download on Spring Link):
  - "[A Modern Introduction to Probability and Statistics, Understanding Why and How](#)" by F.M. Dekking, C. Kraaikamp, H.P. Lopuhaä and L.E. Meester, Springer-Verlag, 2005.
  - "[Modern Mathematical Statistics with Applications](#)", by Devore, Jay L., Berk, Kenneth N., 2nd ed. 2012.
  - "[Analyzing Multivariate Data](#)" by Lattin, Carroll and Green.
- b. Recommended Software: [R](#) and [R Studio](#): Students will be introduced to [the open source R](#) statistical software environment.

## 7. Learning Outcomes

As a result of completing this course, students will be able to:

- a. Perform univariate statistical inference techniques involving confidence intervals, hypothesis test, distribution fitting and goodness-of-fit testing. Students will learn to perform these inference techniques in MS EXCEL.
- b. Perform multivariate statistical inference techniques involving estimation of the sample mean vector, the sample variance covariance matrix and use these to perform the Hotelling  $T^2$  hypothesis test on a single multivariate sample and two multivariate samples. Students will learn to perform these inference techniques in MS EXCEL.
- c. Perform regression analysis involving multiple explanatory variables using matrix algebra in MS EXCEL. Student will learn to perform and interpret regression analysis results using the software MINITAB
- d. Perform One-Way and Two-Way ANOVA using the software MINITAB.

## 8. Attendance

Regular class attendance is strongly encouraged. You will be held responsible for all the class discussions as well as the reading assignments. Here is the university policy:

<https://registrar.gwu.edu/university-policies#attendance>

## 9. Independent Learning

In a 15-week semester, including exam week, students are expected to spend a minimum of 100 minutes of out-of-class work for every 50 minutes of direct instruction, for a minimum total of 2.5 hours a week. A 3-credit course should include 2.5 hours of direct instruction and a minimum of 5 hours of independent learning or a total minimum of 7.5 hours per week. More information about GW's credit hour policy can be found at:

<https://provost.gwu.edu/policies-procedures-and-guidelines>

and click on Assignments of Credit Hour Policy (PDF), Or see the PDF pages (webpage);

[https://provost.gwu.edu/files/downloads/Resources/Assignment-of-Credit-Hours\\_Final\\_Oct-2016.pdf](https://provost.gwu.edu/files/downloads/Resources/Assignment-of-Credit-Hours_Final_Oct-2016.pdf)

## 10. Method of Instruction:

One hour and 20 minutes lecture including homework discussion, followed by a 10 minute break and a one hour lecture. Microsoft Excel and Minitab are used to perform statistical analysis during the class sessions and the homework. During class sessions the only software programs that should be open on your desktop are either Adobe Acrobat (for viewing the notes) or Microsoft Excel or MINITAB for statistical analysis. **Reading assignments will have to be completed before class. Homework will have to be completed and handed-in accordance to the outline schedule one class after it was assigned and before the class starts.** During the class sessions (except for the break of course) **a student is not to check his e-mail, the internet and should not engage in**

**instant messaging sessions. Basically, your attention should be directed towards the class material.**

### **11. Homework Grading Policy:**

Homework sets consisting of multiple homework problems will have to be completed prior to the next class for discussion **and uploaded through Blackboard**. A student may be called upon to discuss their solution for each homework problem in a homework set, so you must be prepared! The rest of the class should be involved in the discussion. Your level of effort will be graded. Not handing in a solution for a homework problem will result in **0 points**. Homework problems that are handed in on time AND demonstrate an adequate level of effort will typically be awarded **1 point**. Partial points can be awarded for homework problems in a homework set. At times a larger homework problem may be awarded a number of effort points **larger than 1**, which will be indicated. Homework sets that are handed in one day late receive a maximum of 50% of the assigned homework credit. Homework sets that are handed in more than one day late will not be awarded any credit.

### **12. Midterm Exam and Final Reports:**

Students will complete an **in-class Midterm Exam using Microsoft Excel + MINITAB** (using a lab computer or the student's laptop). Theoretical questions will be answered in an exam booklet. The MS EXCEL file, the MINITAB file and the exam booklet will be part of the grading of the midterm exam. Two Multivariate datasets will be provided to the students for analysis. Students will be required to perform multivariate data analyses using those datasets and write a final report for each dataset detailing their analysis steps, final analysis results and analysis conclusions. **Students are required to submit the electronic files associated with the final reports through blackboard** as well as **a hard copy of the final report** that will be graded. **Students are required to work on their own to perform the multivariate analysis using those multivariate data sets and write the final report on their own.**

### **13. Grading:**

10% - Class Attendance  
20% - Homework  
30% - Midterm Exam (In-Class)  
40% - Multivariate Data Analyses + Final Reports

### **14. Homework Set and Reading Assignments:**

Homework sets, Lecture notes and recommended chapters for reading will be assigned prior to class as indicated in the outline below.



**Class Schedule: Subject to change, please check the schedule regularly**

	Session	Date	Day of Week	Reading Assignments	Topics	Homework Assigned
<b>Part 1: Prob. and Stats. Review</b>	1	14-Jan	Tuesday	Ch. 1, 2, 3	Why Prob. And Stats?, Probability Calculus Review	
	1	14-Jan	Tuesday	Ch. 4, 5	Discr. and Cont. Distributions Review	Homework Set 1
	2	21-Jan	Tuesday	Ch. 7, 10	Expectation, Variance and Covariance Review	
	2	21-Jan	Tuesday	Ch. 15, 16	Exploratory data analysis: Graphical + Numerical Summaries Review	Homework Set 2
	3	28-Jan	Tuesday	Ch. 17	Basic Statistical Models Review	
	3	28-Jan	Tuesday	Ch. 23	Confidence intervals for the mean: Essentials Review	Homework Set 3
<b>Part 2: Statistical Inference</b>	4	4-Feb	Tuesday	LN S4	Estimator distributions, Confidence Intervals for mean and Variance	
	4	4-Feb	Tuesday	LN S4	Hypothesis Testing, Type I Error and Type II Errors	Homework Set 4
	5	11-Feb	Tuesday	LN S5	Goodness-of-Fit, Credibility Intervals	
	5	11-Feb	Tuesday	LN S5	Two Sample Hypothesis Testing, Joint Normal Distribution	Homework Set 5
<b>Part 3: Multivariate Estimation</b>	6	18-Feb	Tuesday	LN S6	Vectors and Matrices, Matrix Algebra, Linear Combinations,	
	6	18-Feb	Tuesday	Ch. 2	Coordinate Systems, Geometric Interpretation, Joint Normal Distribution, Multivariate Point Estimation	Homework Set 6
	7	25-Feb	Tuesday	LN S7	Single Sample Hotelling's T <sup>2</sup> Test	Homework Set 7
	7	25-Feb	Tuesday		Two Sample Multivariate Hotelling's T <sup>2</sup> Test	Practice Exam
<b>Midterm Examination</b>	8	3-Mar	Tuesday		Discuss Solution Practice Exam, Homework Set 7	
	8	3-Mar	Tuesday			
	9	10-Mar	Tuesday		<b>MIDTERM EXAM - PART 1, PART 2 and PART 3</b>	
	9	10-Mar	Tuesday		<b>MIDTERM EXAM - PART 1, PART 2 and PART 3</b>	
		17-Mar	Tuesday		<b>Spring Break</b>	
		17-Mar	Tuesday		<b>Spring Break</b>	
<b>Part 4: Regression Analysis</b>	10	24-Mar	Tuesday	LN S9, Ch. 3	Simple Linear Regression, Model Testing, Parameter Inference	Homework Set 9
	10	24-Mar	Tuesday			
	11	31-Mar	Tuesday	LN S11, Ch. 3	Multiple Regression, Residual Diagnostics	
	11	31-Mar	Tuesday			
	12	7-Apr	Tuesday	LN S12, Ch. 3	Outlier Detection, Comparing Imbedded Models	<b>REGRESSION DATA SET AVAILABLE</b>
<b>Part 5: Principle Component Analysis</b>	12	7-Apr	Tuesday		Forecasting	
	13	14-Apr	Tuesday	LN S13	One-Way Analysis of Variance (ANOVA)	Homework Set 10
	13	14-Apr	Tuesday			
	14	21-Apr	Tuesday	LN S14	Two-Way ANOVA, 2K ANOVA	
	14	21-Apr	Tuesday			<b>ANOVA DATA SET AVAILABLE</b>
	15	28-Apr	Tuesday		<b>FINAL REPORT PART 4 DUE ON APRIL 28th, 2019</b>	
	15	12-May	Tuesday		<b>FINAL REPORT PART 5 DUE ON MAY 5th, 2019</b>	

### **15. Academic integrity:**

Academic integrity is central to the learning and teaching process. Students are expected to conduct themselves in a manner that will contribute to the maintenance of academic integrity by making all reasonable efforts to prevent the occurrence of academic dishonesty. Academic dishonesty includes, but is not limited to, obtaining or giving aid on an examination, having unauthorized prior knowledge of an examination, doing work for another student, and plagiarism of all types. Ignorance is no excuse.

The number one problem that students run into with regards to academic integrity is plagiarism. It is not okay to copy, use, or otherwise exploit other people's ideas, words, or creations without giving them credit in the proper form. Sometimes this means you must use quotation marks; while other times a simple source citation will do the trick. Changing a few words in a paraphrase is not enough to turn source material into "your own words" – in fact, that's a really bad idea to even try. Changing the phrasing order of sentences is not okay and using the thesaurus to find ways to change "happy" to "glad" is also a very bad idea. It is expected that students know how to correctly quote and cite material, and also how to write well. For those students who need assistance, the GWU Writing Center is available. Please see: <https://writingcenter.gwu.edu/>

### **Academic Integrity Code:**

Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information. For the remainder of the code, see: <https://studentconduct.gwu.edu/code-academic-integrity>

### **16. What to do if the instructor does not arrive:**

If the Instructor does not arrive for the class at the designated starting time and has not notified the class of a late starting time or the cancellation of the class, the students should wait in the classroom for at least 30 minutes before departing. One member of the class should be selected to notify the EMSE Department of the Instructor's absence by calling the EMSE Department 202-994-4892 on next business day.

### **17. University Policy on Religious Holidays:**

In accordance with University Policy,

- a. 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.
- b. Faculty should extend to these students the courtesy of absence without penalty on such occasions, including permission to make up examinations.
- c. Faculty who intend to observe a religious holiday should arrange at the beginning of the semester to reschedule missed classes or to make other provisions for their course-related activities. For more details and policy and accommodations for religious holidays please see:

<https://students.gwu.edu/accommodations-religious-holidays>

## **18. Support for Students Outside the Classroom:**

### **Disability Support Services (DSS):**

Any student who may need an accommodation based on the potential impact of a disability should contact the Disability Support Services office at 202-994-8250 in the Rome Hall, Suite 102, to establish eligibility and to coordinate reasonable accommodations. For additional information please refer to:

<https://disabilitysupport.gwu.edu/>

### **Mental Health Services:**

#### **Colonial Health Services: 202-994-5300 (24Hours/7Days).**

The University's Mental Health Service offers 24/7 assistance and referral to address students' personal, social, career, and study skills problems. Services for students include: crisis and emergency mental health consultations, confidential assessment, counseling services (individual and small group), and referrals:

<https://healthcenter.gwu.edu/mental-health>

## **19. Security and Safety Policy:**

GW Campus Advisories; Students should check the GW Campus Advisories Web Site at: <https://campusadvisories.gwu.edu/> for current information related to campus conditions, closures, safety information and any other information concerning events that may disrupt normal operations. Life-Threatening Emergencies **On Campus: Call GWPD at 202-994-6111**

### **GW Alert Notifications:**

GW Campus Advisories. Students should check the GW Campus Advisories Web Site at: <http://www.campusadvisories.gwu.edu/index.cfm> for current information related to campus conditions, closures, safety information and any other information concerning events that may disrupt normal operations. All students, faculty and staff registered in the GW banner system GW will receive emergency alerts, notifications and updates sent directly to their GW email address. If individuals elect to receive these alerts on a mobile device they may logon to GWeb Information Web Site at: <https://banweb.gwu.edu/> and update their contact information to include mobile devices.