ANLY 500: Principles of Analytics I

Harrisburg University

Late Summer

* **Instructor:** Zi(Bennie) Yang
* **email:** [zyang3@harrisburgu.edu](mailto:zyang3@harrisburgu.edu)
* **office hours by appointment**
* **classroom:** Online
* **class hours:** Tuesdays @ 8:30 PM - 10:00 PM EST

# Course Description

* The material in this course covers an overview of Data Analytics, provides a foundation in research methodology, and introduces students to the R programming language. The goal of this course is to provide an understanding and foundation of the role of Analytics in today organizational environment. The course introduces approaches to performing analytics. The student will be exposed to formal methods and functional programming techniques.
* The instructional approach will include some software demos and illustrate applications of decision support technologies to problems in finance, marketing, manufacturing, services and health care management, and information systems consulting. Students have the opportunity to work on practical exercises and projects on their own.

# Learning Objectives

The main goal of this course is to provide the students with an understanding of the data analytics process and to provide a foundation of skills to become a data analytics professional.

At the conclusion of this course, students will be able to:

* Demonstrate an understanding of the underlying methods and technologies used in business analytics;
* Analyze and applied alternate methods for designing, developing and implementing Business Analytics tools;
* Evaluate the selected alternative technology to use Business Analytics tools;
* Be prepared for taking future courses in Data analytics program and have some insights about different branches of analytics.

# Course Conduct

A few rules will help us to get the most of our investment in ANLY 500:

* Canvas is going to be our platform for almost all course activities.
* In addition, we will have at least one live session on weekly basis.
* Attending live sessions and actively participating in the discussions is included in your “Participation” grade. This is because live sessions are where the basic concept(s) of each unit are explained, assignments are discussed and your questions answered. Not participating will adversely affect your final grade.
* I anticipate that you will need 3 to 4 hours, to budget for completing assignments.
* Because this course covers a very large amount of material very quick NO graded assignments or exams will be accepted late.
* You are responsible for all the readings, even if the material is not explicitly covered in class. You should read the class materials prior to class (or live session) and be prepared to discuss and ask questions about the readings and assignments.
* You should also re-read the material after class (or live session) as not every topic will be covered during class time. Many passages in the text may need to be read several times to gain clarity. Also, taking notes on the material you are reading and reflecting on the reading and these notes will help you better understand the issues, concepts and techniques that are being presented.
* All class credit-related electronic mail must be done using Harrisburg’s electronic mail service and the student’s assigned Harrisburg University ID. By ‘credit-related’ I mean all work to be evaluated for credit. Any work submitted through a different mail system will NOT be accepted.
* All activities will be assigned individually unless mentioned in the assignment.
* Students who participate in University-sanctioned events (such as athletics) must make prior arrangements and give the instructor ample notice. Missing class (or a live session) for practice is not advised.

# Course Project Road Map - HU Core Competencies

At the conclusion of this course a student will have met the following core competencies that reflect HU’s mission:

* Critical Thinking and Problem Solving skills are demonstrated by the student’s ability to: Identify and clarify the problem, Gather information, Evaluate the evidence, Consider alternative solutions, Choose and implement the best alternative.
* Communication - The core communication skills are demonstrated by the student’s ability to: Express ideas and facts to others effectively in a variety of formats, particularly written, oral, and visual formats, Communicate effectively by making use of information resources and technology.
* Teamwork and Collaboration - The students will be working with others to increase involvement in learning and by sharing one’s own ideas and responding to others’ reactions to sharpen thinking and deepen understanding.
* Information Technology - The students will be making effective use of the .NET information resources and technology.

# Statement on Academic Integrity

According to the University’s Student Handbook: Academic integrity is the pursuit of scholarly activity free from fraud and deception, and is the educational objective of this institution. Academic dishonesty includes, but is not limited to cheating, plagiarism, fabrication of information or citations, facilitating acts of academic dishonesty by others, unauthorized possession of examinations, submitting work of another person, or work previously used without informing the instructor, or tampering with the academic work of other students. Any violation of academic integrity will be thoroughly investigated, and where warranted, punitive action will be taken. Students should be aware that standards for documentation and intellectual contribution may depend on the course content and method of teaching, and should consult the instructor for guidance in this area.

**Honor Code** - We as members of Harrisburg University community pledge not to cheat, plagiarize, steal, or lie in matters related to academic work. As a Community of Learners, we honor and uphold the HU Honor Code.

# Required Textbook

* **Title:** Discovering Statistics Using R
* **Authors:** Andy Field, Jeremy Miles, Zoe Field
* **Publisher:** Sage
* **ISBN:** 9781446200469
* **Year:** 2012

# Optional Textbooks

* **Title:** An introduction to Data Science
* **Authors:** Jefferey S. Saltz, Jefferey M. Stanton
* **Publisher:** Sage
* **ISBN:** 9781506377537
* **Year:** 2018
* **Title:** Mathematical Statistics with Applications in R
* **Authors:** Kandethody M. Ramachandran, Chris P. Tsokos
* **Publisher:** Elsevier
* **ISBN:** 9780124171138
* **Year:** 2015
* **Title:** Probability and Statistics with R - 2nd Edition
* **Authors:** Maria D. Ugarte, Ana F. Militino, Alan T. Arnholt
* **Publisher:** CRC Press
* **ISBN:** 9781466504394
* **Year:** 2016
* **Title:** An introduction to Statistics: An active learning approach - 2nd Edition
* **Authors:** Keith A. Carlson, Jennifer R. Winquist
* **Publisher:** Sage
* **ISBN:** 9781483378732
* **Year:** 2018

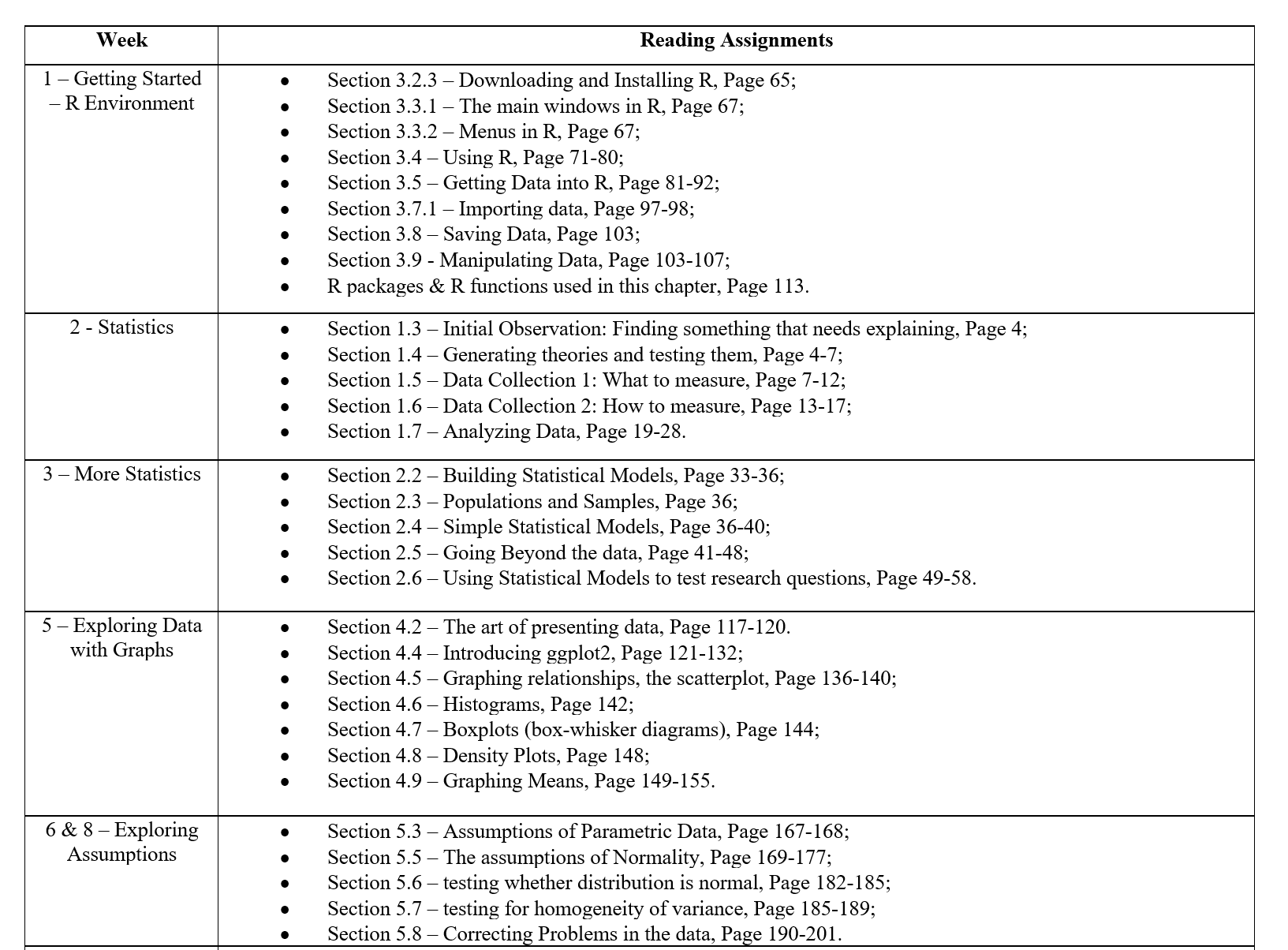
# Software Required

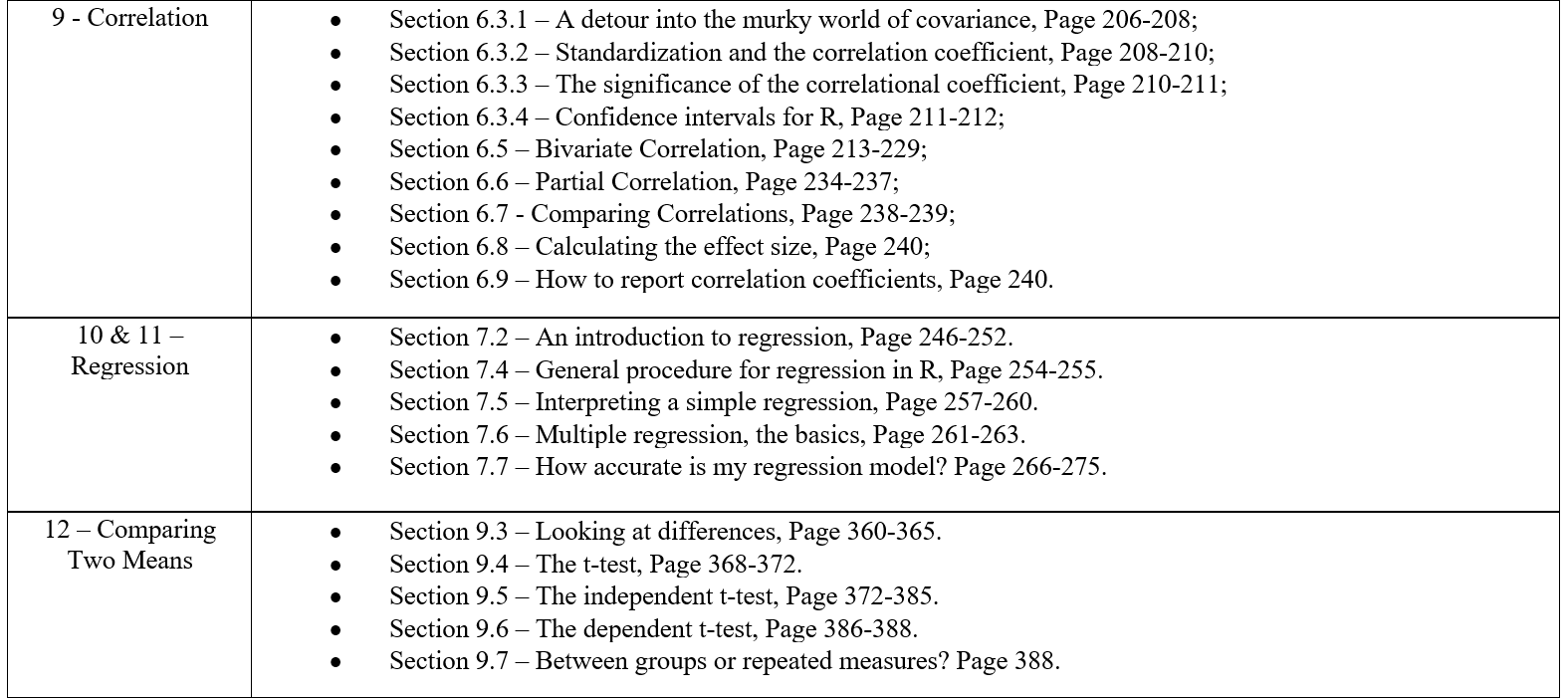
* R and RStudio available at: <https://www.r-project.org/> and <https://www.rstudio.com/>

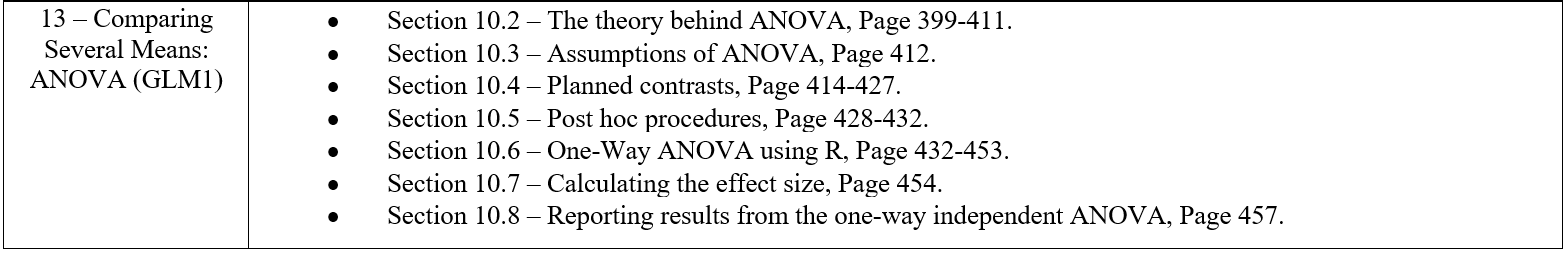
# Assignments & Grading

**Required Reading - Discovering Statistics Using R:**

Weekly reading is required to stay up to date with the course material. Not all the subject matter can be covered in the allotted time during our adobe connect sessions and executive meetings.







**Laboratories - R Practical’s:**

Lab works assignments are based on taught material and will be posted on the Canvas page.

* There are 9 assignments, which are separate labs to complete throughout the term.
* Each lab folder on Canvas contains a combination of the following items:
  + Lab Guidelines//Questions as a .rmd file template
  + Data Source
* Laboratory Topics:
  + Lab 01 - R Learning
  + Lab 03 - Introduction to Data Analytics 1,
  + Lab 04 - Introduction to Data Analytics 2,
  + Lab 05 - Graphing,
  + Lab 06 - Data Screening,
  + Lab 07 - Correlation,
  + Lab 08 - Regression,
  + Lab 09 - t tests,
  + Lab 10 - ANOVA.

**Examinations:**

* Exams are based on the taught material and will be posted on the Canvas page.
* Mid-Term Exam - *Covers Chapters 1 - 5*
* Final Exam - *Covers Chapter 6 - 10*

**Final Project:**

Please refer to the Final Project Guidelines & Report Formatting documents for more details.

* The final project consists of three deliverables:
  + 1. Project Proposal, (2) Presentation, (3) Written Report

**Other Assignments for Participation:**

To further our discussions on the course material other assignments will be posted to Canvas including the following:

* ***Forums*** - It is important that the class keeps communication ’open to discuss the various topics for the class. A forum will allow constant communication among us all during the term. Post questions, ideas, start a chat, it’s a free open space for our collective thoughts.
* ***Discussions*** - The field of data analytics is constantly improving as new information is discovered. This requires us to stay up to date on material within the analytical world. Various reading materials and questions will be posted to Moodle for the class to engage in a discussion. Discussion may be based on the required reading or another resource.
  + All forums and discussions will be announced when posted on Canvas and will be due within one week from being posted/announced.

**Final Grading:**

Your final grade will be calculated using the following:

* Course Grading Policy
  + Mid-Term Exam 20%
  + Final Exam 20%
  + Final Project 30%
  + Laboratories 20%
  + Participation 10%
* Final Grading Policy
  + 100.00 %-90.00 % = A
  + 89.99 % - 80.00 % = B
  + 79.99 % - 70.00 % = C
  + 69.99 % - 0 % = F
* Participation consists of attendance, Lab 1 submission, on time assignment submissions, and overall performance.

# Class Schedule

Here’s a best guess at the course schedule - the exams and holidays will not change, but lecture material may be moved around depending on course speed. The exercises will be posted on Canvas, along with all notes and scripts created for you to use to work from.

## Week 01

* Lecture 1 - R Learning:
  + Introduction to the course.
  + Install R and RStudio (the free versions).
  + Start coding with R.
* Assignments/Meetings:
  + Executive Session #1- In person on campus, if applicable.
  + Adobe Connect Session\_0
  + Software Installation, R.
  + Review Syllabus.
  + Review Canvas Course Page.

## Week 02

* Lecture 2- Intro to Data Analytics / Foundations {part 1}:
  + Intro to data analytics.
  + What is data analytics and the definition of predictive, descriptive and prescriptive analytics
  + Examples on different type of collecting data and analyzing them in real world.
* Assignments/Meetings:
  + Adobe Connect Session\_1
  + Lab 1 due by Week 03. please submit on Canvas.

## Week 03

* Lecture 3- Intro to Data Analytics / Foundations {part 2}
  + What is the research Process?
  + What are qualitative and quantitative data?
  + Identifying and measurement of variables.
* Assignments/Meetings:
  + Adobe Connect Session\_2
  + Lab 2 due by Week 04. please submit on Canvas.

## Week 04

* Lecture 4- Intro to Data Analytics / Foundations {part 3}
  + Know what a statistical model is and why we use them.
  + The Mean
  + Know what the ‘fit’ of a model is and why it is important.
  + The Standard Deviation
  + Distinguish models for samples and populations
* Assignments/Meetings:
  + Adobe Connect Session\_3
  + Final Project Proposal due by Week 05- please submit on Canvas.
  + Lab 3 due by Week 05. please submit on Canvas.

## Week 05

* Lecture 5- Stats & Graphing
  + Prepare data for further analysis
  + What is accuracy and error?
  + How to handle missing values?
* Assignments/Meetings:
  + Lab 4 due by Week 06 - please submit on Canvas.

## Week 06

* Lecture 6- Data screening Part I
  + Prepare data for further analysis
  + What is accuracy and error?
  + How to handle missing values?
* Assignments/Meetings:
  + Lab 5 due by Week 07 - please submit on Canvas.

## Week 07

* Midterm Review & Introduction to Final Project
  + Mid-term Review
  + Review of Final Project Remaining Deliverables
* Assignments/Meetings:
  + Midterm Exam, due by Week 08- please submit on Canvas.

## Week 08

* Lecture 7- Data screening Part II
  + How to handle outliers?
  + What are missing values and how can we handle them?
* Assignments/Meetings:
  + Lab 6 due by Week 09 - please submit on Canvas.

## Week 9

* Lecture 8- Correlation
  + A complete review on correlation and its properties
* Assignments/Meetings:
  + Lab 7 due by Week 11.

## Week 10

* Lecture 9- Linear Regression
  + What is linear regression?
  + How is it related to correlation?
  + What are predictors and response?

## Week 11

* Lecture 10- Regression mediation and moderation
  + What if the relationship between a predictor variable and outcome variable can be explained by their relationship to a third variable (the mediator)?
* Assignments/Meetings:
  + Lab 8 due by Week 12.

## Week 12

* Lecture 11- Comparing two means
  + What is t-test and how can we use it in data analytics?
  + What are independent and dependent t-tests?
* Assignments/Meetings:
  + Lab 9 due by Week 13.

## Week 13

* Lecture 12- ANOVA
  + A brief overview on ANOVA test and its applications
* Assignments/Meetings:
  + Lab 10 due by Week 14.

## Week 14

* Review session for Final Exam
* Assignments/Meetings:
  + Final Exam due
  + Final Project [Final Presentation & Written Report] due by Week 14.

## Classes End