RUTGERS
School of Business | Camden

52:623:447:90   Data Management and Analytics with R
Dr. Shankar Sundaresan
Spring 2023

Contact Information	**Office:** 438 BSB **Phone:** 856 225 6694
	e-mail: sundares@camden.rutgers.edu
Class Meetings	online
Office hours	Mondays: 5:00 pm to 6:00 pm (online)
	Thursdays: 5:00 pm to 6:00 pm (online)
	other times by appointment
Course LMS	Canvas

Course Overview:

This course will cover the fundamentals of Data Management as well as the elements of data analytics.

A good understanding of data management and analytics concepts and a working knowledge of the tools of analytics are becoming indispensable for business professionals. This course is designed to enable students to learn the underlying principles of data management and apply them using the SQL language. Moreover, data analytics concepts and functions such as data manipulation, importing, creating, modifying, filtering, summarizing, reshaping and analyzing data sets and their application will be explored using the "R" language. The course will help students think about data, its manipulation and analysis and graphical presentation in practical contexts.

Course Objectives:

Specifically the course aims to help you:

- understand the basic foundations of data management.
- understand the principles of relational databases and their role in analytics.
- create tables and other structures for storing and organizing data.
- retrieve and manipulate data in relational databases by creating SQL queries

- Formulate and answer business questions by extracting and analyzing data
- learn to filter, modify and summarize data
- learn the basics of the "R" language
- conduct data analytics operations on business data
- use visualization tools to present the data analysis results
- develop a versatile vocabulary about data management and analytics
- apply the concepts by employing managerial IT tools

- use information and data to analyze, interpret and solve business problems, specifically data analysis using SQL and R.

- demonstrate conceptual knowledge and practical application of information technology, specifically SQL and R.

Required Course Materials

Required Texts:

SQL Queries for mere mortals, 4e, by John L. Viescas, Addison-Wesley Professional, 2018.
ISBN: 9780134858333, e-text version available.

R for Everyone Advanced Analytics and Graphics, 2nd edition, by Jared P. Lander, Addison-Wesley Professional, 2017**.**
ISBN: 9780134546926; e-text version available.

Other web-based readings (articles, business clippings and web sites) will be assigned regularly.

There is a wide variety of free texts and material available on the web on these topics. In addition, **Rutgers-LinkedIn site** provides some excellent tutorials/classes both for beginning and advanced SQL and R.

Software: The "official" software for the class will be **MySQL Workbench** and **R-Studio**. **MySQL Workbench** and **R-Studio** are available in class lab computers. **Free** versions of both of these software are available for installing on personal computers. Instructions will be provided for installation and use.

Prerequisites

Please note that this course is designed as an advanced elective course in the undergraduate program and will assume that have completed

the basic business computing course, the core MIS course, and the two statistics courses. You will be working with two software programs, and are expected to install the required programs. Detailed instructions will be provided in the course materials. You will also be using these two software programs extensively to master the contents of the course. You do **not** need any prior programming experience; all the necessary skills will be taught in the course. However, it will be helpful if you have an aptitude for hands-on work, especially "algorithmic thinking" and working with code as required.

Class Materials

All class materials can be obtained via **Canvas**. Usually, the lecture notes and class discussions will be available in Canvas by the start of the week, often earlier. You are strongly encouraged to access the course via Canvas several times a week. You can imagine that the class seamlessly "runs" through from day 1 (Monday) to day 7 (Sunday).

Class Communication:

Since class attendance is not compulsory, a viable and reliable form of communication is vitally important. Note that all class communication will be via your **Rutgers e-mail** and discussion forums and other tools in Canvas. You are expected to check your Rutgers e-mail at least (equally spaced) three times every week. All class announcements can also be accessed via the 'Announcement' page in Canvas.

Email Communication:

Note that during the week, from Monday until Friday, I will try reply to all e-mails within 24 hours, unless I am traveling. Although, I check my e-mails several times a day, I may not be able to completely answer all e-mails immediately upon receiving them. Note that I may not be available on **weekends** and may not be able respond to weekend e-mails until Monday.

Course Components and Weighted Values

Assignments	**18 %**
Exam1	**16 %**
Exam2	**22 %**
Exam3	**16 %**
Exam4	**22 %**
Class Participation	**6 %**

Grading Scale

Score Range	Grade
90%–100%	A
85%–90%	B+
80%–85%	B
75%–80%	C+
70%–75%	C
60%–70%	D
Below 60%	F

If your score is in the indicated range, you will **at least** receive the grade in the corresponding Grade column. The instructor reserves the right to grade on a curve, which usually improves students' grades.

Course Requirements

Participation: Each one of you is expected to be engaged with the class by regularly logging in to Canvas site and reviewing the materials including lecture videos, turning in homework assignments, quizzes and exams in a timely fashion, responding to any posted discussions, and participate other class activities.

Assignments: There will be 10-12 assignments to be completed by each student. Many will require the use of tools such as MySQL and R, etc. The assignments for meant for improving your understanding of the course materials and application skills. Please approach me if you need help with the assignments.

Four exams will be conducted. The format and policy for each exam is expected to be as follows:

● **Exams 1, 2, and 3** will be available between **noon EST on Thursday** of the **specified week** of the exam and must be completed **by 11 p.m. EST Sunday**.

● **Exam 4** is expected to be available between **noon EST on Thursday (May 4, 2023) and 11 p.m. EST Sunday (May 7, 2023).**

● You can take each exam only once. Once started, the exam must be completed in the allowed time, in one uninterrupted session. That is, do not log out until you have completely finished the exam. Any unanswered question will automatically receive a zero grade for that question. Any unsubmitted exam will automatically be assigned a zero grade. **Note that every student might have a different exam,** because the examination questions are taken from a randomized pool of questions. Any students whose answer(s) do not pertain to their version of the exam, but to another student's version(s) will receive a zero grade for that exam.
● While taking the test, you may not refer to any other course materials, except for those specifically permitted in the individual exam instructions. Ensure that you have a reliable computer, fully charged battery, and reliable Internet connection before starting the test. Make sure that your computer meets all the requirements and that you configure the browser as suggested. **Exams may require the use of respondus / lockdown browsers.** Hence, use of video features may be required.
● Any student who missed the exam without previous approval of the instructor or a compelling reason will receive a zero grade for that exam.
● **Exam Make-Up Policy:** Because you are given a very reasonable amount of time to work on the exams, make-ups are not given. If you cannot take an exam by the scheduled deadline for a university-approved reason, you must give the professor written notice at least one week in advance so that other arrangements can be made. The professor reserves the right to request written documentation to support your absence (such as a doctor's note or military orders).
If the situation does not allow for advance notification (for example, emergency hospitalization), contact the professor as soon as possible after a missed exam.

Student Expectations

To be successful in the course, it is a critically important to complete the following activities in the learning module for each week.

- Carefully read the learning goals and complete all the assignments.
- Print out and review the online PowerPoint (PPT) lecture notes for each session. Note that any narrated PPT lectures will briefly discuss some aspects of the material.
- Carefully read the required and assigned textbook chapter(s) and write down additional notes on the printouts
- Read/review the additional required articles/websites or watch related videos. Go through the animated activities for applied understanding.
- Make sure you follow the lecture material with the associated **MySQL or R-Studio hands-on work**. **Unless you replicate the work independently, you will not be able to master the materials.**
- Complete all the homework assignments for each module. The homework problems/review problems are meant for practice, learning, and mastery of the material. If you do not complete the homeworks, you may not do well in the course. Many exam problems usually closely resemble the homework problems.
- Complete all assigned online exams because all of these count toward your final course grade.
- Participate in the assigned class discussions, if any, to get maximum credit for class participation. These discussions are designed to be a collaborative learning experience for all involved. You should log in two or three times throughout the week to participate in discussions.
- Using the **SQLR Box**, you are strongly encouraged to share any relevant class-related topics pertaining to the current business environment. Here, you may also benefit and assist other students through thoughtful interaction.
- Please use the office hours to help you understand the material better and to clarify doubts.

Canvas Conference and Direct Interaction

- Although Canvas provide good learning tools, a direct one-to-one professional interaction to resolve difficult issues is often the most effective. For example, you may have a certain technical/mathematical problem that is frustratingly difficult to solve and may not be resolved in open forums or elsewhere. There will be open online office hours (on a weekly schedule) offered, where we will use the interactive collaborative tools to

address these questions. A **Teaching Assistant** is also available for the course, who can help with different aspects of the course.

- When required, please feel free to send me an email and a contact number so that I can call you directly to resolve it. I will also be able to meet with you personally, if there is a need for it.

Etiquette Expectations from Online Students

The following protocols on the codes of behavior reflect professional business norms on manners, courtesy, and respect. (In general, you should treat others as you would like others to treat yourself. Be mindful that what is acceptable in a text or chatroom with friends may not be appropriate in a classroom or in an online conversation with an instructor.) Even though many of you are already aware of these protocols, they are explicitly stated here so that everyone is cognizant of the same protocols. These protocols should be followed by all students taking this course to help ensure that the online experiences for everyone involved are pleasant.

● If you send an email to the professor or tutor, please address the person appropriately such as "Dr. …." or "Mr./Ms. …." , not "Hey." I will address you using your first name, unless you prefer that I address you differently.
● When sending an email, you can get better attention by using the following guidelines: a) use descriptive subject lines (this will deter your message from being wrongly mistaken as spam and deleted by the receiver), b) be as brief as possible by going straight to the point, and c) if possible, limit the use of attachments.
● Never send offensive and insulting messages. If you disagree, say so and state your reasons. Social media is a very powerful tool for communication. However, it can be badly misused if it is not used correctly or professionally. For example, you may have personal and legitimate concerns with this course. However, other students, who do not have similar feelings, can be negatively influenced by your concerns, which can unnecessarily and negatively affecting their overall experience of the course. To minimize such an occurrence, directly contact the instructor/tutor first to resolve any concerns you may have to help ensure that everyone's online experience of this course is wonderful!
● Always guard against flaming when it comes to content, opinions, and the like. That is, avoid blaming or accusing others of wrongdoing. Do not start a volley of back-and-forth emails, with copies distributed to every student in the class.
● Copy the minimum number of people. That is, send email to only the people you think should receive and will benefit from it.

● Treat all emails and postings as permanent forms of written record and do not expect any your email communications to be private, unless stated otherwise. Instead, assume that all email communications are public. Do not publicize your own or others' personal information (such as email addresses, phone numbers, last names, and the like).

●Avoid using ALL CAPS, if possible, and never type messages in ALL CAPS because this is considered yelling and can be seen as a form of aggression.

Other Administrative Comments

- **Student Involvement:** You are encouraged and expected to ask questions and to interact with the instructor. If at any time during the course, you have questions regarding course-related matters, do not hesitate to contact me. If the scheduled office hours are inconvenient for you and if you want to meet in person, please contact me to arrange an alternative appointment.

- **Feedback:** I welcome your feedback on the content and style of the class. You may choose to provide your feedback in person, via e-mail, or anonymously at any time. Your suggestions will help me to make the class more useful for everybody.

- **Announcements:** Announcements will be periodically made in class about changes in schedules, assignments, exam, readings, project, policies and other class activities. Please keep abreast of the announcements and changes.

- **Late Submissions:** All work must be turned in by the due date and time, and in the manner suggested (e.g., through Canvas). **Late submissions are not accepted**.

Academic Integrity:

For the policy on Academic Integrity please see:
http://academicintegrity.rutgers.edu/academic-integrity-at-rutgers

Academic Integrity means that you must:
- •properly acknowledge and cite all use of the ideas, results, or words of others,
- •properly acknowledge all contributors to a given piece of work,
- •make sure that all work submitted as your own in a course activity is your own and not from someone else
- •obtain all data or results by ethical means and report them accurately

• treat all other students fairly with no encouragement of academic dishonesty

Adherence to these principles is necessary in order to ensure that:
   •everyone is given proper credit for his or her ideas, words, results, and other scholarly accomplishments
   •all student work is fairly evaluated and no student has an inappropriate advantage over others
   •the academic and ethical development of all students is fostered
   •the reputation of the University for integrity is maintained and enhanced.

For instance, you are responsible for preparing and entering your own work and properly referencing the work of others. Cheating, plagiarism, and other types of misconduct are not acceptable. Penalties can include expulsion from the University. You are free to discuss any part of the course materials with your classmates. However, you are not allowed to discuss (i.e., receive nor give any assistance on) any part of the exams with anyone. You may not refer to sources not permitted nor receive help from outside agencies. If any cheating is found, the most severe sanctions available will be sought.

Students are responsible for understanding the principles of academic integrity and abiding by them in all aspects of their work at the University. Students are also encouraged to help educate fellow students about academic integrity and to bring all alleged violations of academic integrity they encounter to the attention of the appropriate authorities. Violations are taken seriously and will be handled according to University policy.

Student Code of Conduct

As a student at the University you are expected adhere to the Code of Student Conduct. The University's Student Code of Conduct can be found at

http://studentconduct.rutgers.edu/university-code-of-student-conduct

Rutgers University-Camden seeks a community that is free from violence, threats, and intimidation; is respectful of the rights, opportunities, and welfare of students, faculty, staff, and guests of the University; and does not threaten the physical or mental health or safety of members of the University community, including in classroom space.

Note that the conduct code specifically addresses disruptive classroom conduct, which means "engaging in behavior that substantially or

repeatedly interrupts either the instructor's ability to teach or student learning. The classroom extends to any setting where a student is engaged in work toward academic credit or satisfaction of program-based requirements or related activities."

Violations of the Student Code of Conduct are considered serious infractions of student behavior and students who violate the code are subject to penalties relative to the level of the matter. Violations of the Student Code of Conduct should be reported to the Dean of Students office deanofstudents@camden.rutgers.edu or 856-225-6050.

If you need to interact with me in-person, we will follow applicable University's Covid guidelines.

Covid Guidelines

We will follow all Covid related guidelines that Rutgers University and the Rutgers School of Business – Camden institute. This currently includes mandatory masking in class rooms and indoor spaces. These policies are treated as part of Student Code of Conduct, and violations will be treated as violations of the Student Code of Conduct.

More details may be found at the Rutgers University/ Rutgers School of Business – Camden web sites.

Disability Services / Student Health and Well-Being Resources:

The University is committed to supporting the learning of all students and faculty will provide accommodations as indicated in a Letter of Accommodation issued by the Office of Disability Services (ODS). If you have already registered with ODS and have your letter of accommodations, please share this with me early in the course. If you have or think you have a disability (learning, sensory, physical, chronic health, mental health or attentional), please contact https://success.camden.rutgers.edu/disability-services.

Accommodations will be provided only for students with a letter of accommodation from ODS. Their services are free and confidential. Letters only provide information about the accommodation, not about the disability or diagnosis.

Our campus provides a number of student health and well-being resources. Please see https://wellnesscenter.camden.rutgers.edu for details.

Course Outline (Tentative)

(In the readings, JV refers to the "SQL Queries for mere mortals," by John Viescas; JL refers to the "R for Everyone: Advanced Analytics and Graphics," by Jared Lander)

Week 01: (Jan 17 – Jan 22)	Readings
Topic : Overview, Introduction to Data Management and Analytics	
Course Overview Introduction to Data Management Relational Databases	JV: Chapter 1
Week 02: (Jan 23 - – Jan 29)	
Topic : Relational Databases and SQL	
Design Principles for Relational Databases SQL and Query Primer	JV: Chapter 2, 3 Appendix A, B
Hands-on: MySQL Workbench	
HomeWork 1	
Week 03: (Jan 30 – Feb 05)	
Topic : Simple Queries	
Simple Queries, Filtering Results Combining Multiple conditions	JV: Chapter 4, 5, 6
Hands-on: MySQL Workbench	
HomeWork 2	
Week 04: (Feb 06 – Feb 12)	
Topic : Simple Queries, Data from Multiples Tables	
Join Operations Querying Multiple Tables	JV: Chapter 8
Hands-on: MySQL Workbench	
Homework 3	
Week 05: (Feb 13 – Feb 19)	
Topic : Exam 1	
Topic : Data from Multiples Tables - Continued	
Join Operations Querying Multiple Tables	AB: Chapter 8
Complex Joins	AB: Chapter 9
Hands-on: MySQL Workbench	
Homework 4	
Exam 1 (Due on Feb 19)	
Week 06: (Feb 20 – Feb 26)	
Topic : Aggregating Data	
Grouping data Aggregate Functions	JV: Chapter12, 13, 14

Hands-on: MySQL Workbench	
Homework 5	
Week 07: (Feb 27 – Mar 05)	
Topic : Aggregating Data Continued, Populating Databases	
Creating and Populating Databases, Updating, Deleting Data	JV: Chapter 15, 16, 17
Hands-on: MySQL Workbench	
Homework 6	
Week 08: (Mar 06 – Mar 12)	
Topic : Exam 2	
Topic : Introduction to R and Data Analytics	
R and R-Studio	JL: Chapter 1, 2, 3
Basics of R	JL: Chapter 4
Hands-on: R Studio	
Exam 2 (Due on Mar 12)	
Mar 13– Mar 19 : Spring Break … Spring Break … Spring Break	
Week 09: (Mar 20 – Mar 26)	
Topic : Data Types, Structures and Data Frames in R	
Basics of R	JL: Chapter 4
Data structures and Data Frames in R	JL: Chapter 5
Hands-on: R Studio	
Homework 7	
Week 10: (Mar 27 – Apr 02)	
Topic : Datasets and Loading Data in R	
Getting data into R	JL: Chapter 6
Loading data from SQL	
Loading data from Excel	
Hands-on: R Studio	
Homework 8	
Week 11: (Apr 03 – Apr 09)	
Topic : Datasets and Loading Data in R … Continued	
Getting data into R	JL: Chapter 6
Loading data from SQL	
Loading data from Excel	
Screen scraping and loading data from web sites	
Hands-on: R Studio	
Homework 8	
Week 12: (Apr 10 – Apr 16)	
Topic : Exam 3	

Topic : Graphics in R, Exam 3	
Basic Graphics Scatter Plots Bar graphs, pie charts, Box plots	JL: Chapter 7
Hands-on: R Studio	
Homework 9	
Exam 3 (Due on Apr 16)	
Week 13: (Apr 17 – Apr 23)	
Topic : Visualization, Advanced Graphics, Statistical Inference in R	
Basics of data visualization Advanced Graphics in R	JL: Chapter 7
Probability Distributions	JL: Chapter 14
Hands-on: R Studio	
Homework 10	
Week 14: (Apr 24 – Apr 30)	
Topic : Predictive Analytics in R	
Basic Statistical Analysis Simple Linear Regression in R	JL: Chapter 15 JL: Chapter 16
Multiple Linear Regression in R	
Hands-on: R Studio	
Homework 11	
Week 15: (May 01 – May 07)	
Topic : Reading Days and Exam 4	
May 02, 03: Reading Days **Exam 4 – (Due on May 07)**	

(Please note that the dates indicated for topics and assignments are tentative, and
are likely to change depending on the course pace.)