

BUAN 370
Descriptive Analytics & Data Management
Spring 2026

| | | | |
|--------------------------|---|-------------------------|--|
| Instructor: | Dr. Carl M. Rebman, Jr. | E-mail: | carlr@sandiego.edu |
| Office: | Olin 127 | Class Times: | 4:00-5:20 PM TTH |
| Phone: | 619-260-4135 | Classroom | MH 114 |
| Office Hours: | T: 3:25 pm - 3:55 pm (KCBE 316) W: 11:50 pm – 3:50 pm (OH 127) Th: 3:25pm – 3:55pm (KCBE 316) , or by appointment | Webpages: | http://ole.sandiego.edu http://carl.sandiego.edu/buan370 |
| Teaching Schedule | LC 106 9:30 AM – 10:35 AM, LC 106 10:45am-12:05 PM KCBE 316, 4:00-5:20 TTH | Meeting Schedule | Every 1 st /2 nd TUE of the month (faculty/area), and every 2 nd /4 th THR (senate) |

SYLLABUS

Course Description

Advances in our capability to generate and collect information coupled with decreasing disk-space prices are pushing us toward a world centered around data management. Data preparation and storage are the foundation of today's business analytics. They ensure data are properly processed for later meaningful analysis. Data preparation includes data cleansing and data transformation. The objective of data preparation is to collect the data from various sources into a single location and transform it into a form that is ready for later analysis.

Databases are at the heart of modern commercial application development for data storage. Once data is prepared and properly stored, the first step of analysis usually involves summarizing basic facts about what has happened in the past. This preliminary examination of data falls in the category of descriptive analytics (exploratory data analysis). The purpose of this course is to provide a comprehensive introduction of the data management process - from data preparation, storage, to descriptive analytics applications.

Course Outcomes

At the conclusion of this course you will be able to:

- Identify and explain analytics uses in real world business situations.
- Identify areas in which data could improve decision-making as well as viable data sources
- Interpret and communicate results to both technical and non-technical audiences.
- Access large datasets and doing the “wrangling” needed to prepare them for analysis
- Describe basic relational database concepts and write SQL queries for a variety of data definition and data manipulation scenarios
- Identify appropriate visualizations for different types of data and best practices for creating visualizations, applying appropriate analytical tools for data analysis

Prerequisites

In order to be enrolled in this course, you should have satisfactorily completed ITMG 100, ECON 216 (OR ECON 217). Some experience scripting/programming will be useful, but not required.

Class & Lab Sessions

All live class sessions will take place at the regular scheduled time for the course (see top of syllabus). These lectures will take place in person (or through Zoom in extenuating circumstances). My plan is to conduct the lectures live, but there will also be prerecord some content for supplemental viewing for outside of class between the sessions (some to be prepped before class sessions, some for afterwards). Some days we will have lab sessions where I will provide code chunks to you in advance and we will run through them together in class, so it is expected you have access to your machine (or a lab computer) during these sessions.

Workload & Expectations

Analytics is naturally a time-intensive field, which is why salaries for business analytics professionals are higher than in other fields. Therefore, you should expect to spend ***an average of 6 – 8 hours OUTSIDE*** of class per week working on and learning the material. Some weeks you may spend more than 6 – 8 hours and some weeks you may spend less. It is vitally important that you be open to “hacking” - taking advantage of open access resources on the internet (ie: AI, Stack Overflow, DataCamp, etc.) to independently supplement the course materials assigned to help you troubleshoot errors, investigate use cases, and reference open source code chunks and libraries built by the R community. If you'd like a free chat AI to assist you, I'd recommend starting an account with Uflo (<https://www.uflo.io>). Being able to experiment on your own with trial and error is an important part of studying analytics, especially when it comes to troubleshooting (nobody will be there to help you at 2am at your job - you'll need to be effective in learning how to help yourself). Before reaching out for help troubleshooting, you should always ask yourself, “did I try and resolve the issue on my own by leveraging ‘hacking’ skills before reaching out for help.” It is also critically important that you follow instructions carefully to streamline efficiency and to avoid redundancies in questions and processes.

Most easily fixable

troubleshooting issues are related to folks failing to follow instructions carefully.

An example of how you may allocate your 6 – 8 hours study time each week is as follows (everyone is unique so you may personally have a different allocation of your time; this is just an example):

- 2 hours: reading the relevant assigned readings and taking your own notes on the readings
- 1 hour: reviewing examples and notes from lecture
- 1 – 2 hours: working on assignments (worked examples, code chunks, problem sets, projects)
- 1 – 2 hours: going to office hours and/or reviewing with peers

BUAN 370
Descriptive Analytics & Data Management
Spring 2026

(groupstudy)

- 1 hour: troubleshooting and researching

Careers in Analytics

Analytics is time-intensive, but it pays off! Average “Business Analyst” salary is $\approx \$84,484$ (Indeed.com, 2024) and “Analytics Manager” salary is $\approx \$137,513$ (Salary.com, 2024). Salaries of seasoned “Data Scientists” exceed $\$156,846$ annually. (Glassdoor.com, 2024 <https://www.bls.gov/oes/current/oes152051.htm#st>).). Analytics is naturally a time-intensive field, which is why salaries for business analytics professionals are higher than in other fields. Business Analytics ranks at the highest end of the salary scale compared to any other concentration you might choose in the school of business.

Professional Certifications

The content in this course provides some of the preliminary conceptual preparation for professional analytics certifications (ie: CAP-Essentials, Pro, or Expert (<https://www.certifiedanalytics.org/>)). Completion of the DataCamp chapters will also put you in a good position to finish any remaining course chapters (not required for this class) on your own while your DC classroom license is active to add digital certifications to your portfolio - each DC course completion is accompanied by a digital certificate (pdf) that can be shared to your LinkedIn page.

Classroom Decorum & Email

To avoid distracting others in the classroom, please arrive on time and be respectful of your classmates (ie: phones on silent). I use a “call and response” style type lecture that relies on asking leading questions for students to respond to so I can gauge comprehension before moving on to the next topic. Please try and be efficient with your email volume (the more efficient you are, the faster I will be able to respond collectively). Put the course number (BUAN 314 / 370) in the subject of the email – I teach other classes and this is the fastest way for me to know who I’m responding to. Put your full name at the end of email messages as well. Your email decorum should be professional. The way to get the fastest response via email is to email me directly at carlr@sandiego.edu (DO NOT MESSAGE ME THROUGH CANVAS). I also spend about 2 hours per day responding to emails, so if you don’t get a reply immediately (I answer them in the order I get to them), please either follow-up or wait patiently. I should get back to you within no longer than three days at the most (especially if you email over the weekend).

Course Materials

[1] R Programming for Data Science

Authors: Roger D Peng

Publisher: Leanpub

Required (as a reference guide)

Available free (or you can donate) at <https://leanpub.com>

[2] ggplot2: Elegant Graphics for Data Analysis

Edition: 2nd
Authors: Hadley Wickham
Publisher: Springer
Recommended
Can be found online for *free with a quick internet search.*

[3] Database Design

Edition: 2nd
Authors: Warr and Eng
Required
Download for free from B.C. Open Textbook Collection.

[4] A First Course in Database Systems

Edition: 3rd
Authors: Ullman and Widom
ISBN: 978-0136006374
Recommended

[5] Exploratory Data Analysis with R

Authors: Roger D Peng
Publisher: Leanpub
Recommended
Available *free* (or you can donate) at <https://leanpub.com>

[6] Statistical Inference for Data Science

Authors: Brian Caffo
Publisher: Leanpub
Recommended
Available *free* (or you can donate) at <https://leanpub.com>

[7] Regression Models for Data Science in R

Authors: Brian Caffo
Publisher: Leanpub
Recommended
Available *free* (or you can donate) at <https://leanpub.com>

[8] Resource for relational database and SQL

<http://ovid.cs.depaul.edu/Classes/CSC355-S14/CSC355-links.htm>

Software

Required:

- [9] **R**, a statistical computing package that can be downloaded for free here: <https://www.r-project.org/>. You can find a useful video for installing R here: <https://youtu.be/JNy-7OjNKQ>. While this is the software that we will use, the main tool in the course will actually be item [2] below, the IDE, RStudio, that “sits on top of” R.
- [10] **RStudio**, An integrated development environment (IDE) for R, an open source statistical computing software package. You can download the most recent (desktop) version (compatible for Mac and Windows) for free at www.rstudio.com. All other packages and extensions can be loaded seamlessly from within the IDE. Here is a useful video for installing RStudio: <https://youtu.be/wvmFQfK6pyl>.
- [11] **GitHub**, a version control software that will be helpful when it comes to collaborative editing of data or documents in the cloud. Will also require PC users (but not Mac) to download Git (Mac users will need to use their terminal window). You can download Git for Windows here: <https://git-scm.com/download/win>.
- [12] **SQL** (Structured Query Language), a language used to query relational databases. We will leverage a virtual MS SQL Server as well as an open source library in R called *sqldf* that will allow us to run SQLite queries from within RStudio.

Recommended:

- [13] **SWIRL** — a supplemental package for self-teaching tutorials in R in the event you need additional practice. You can find instructions for installing the SWIRL lesson modules in code chunk #1 SWIRL INSTALLATION INSTRUCTIONS on Blackboard. Here is a useful video walking you through the installation: https://youtu.be/NI_CQCf9BE. While these lessons are optional (you can ignore the submission instructions at the end of each module), it is *highly recommended* that you complete *at least the first set of lessons (15 lessons) for the R Programming module in parallel with the first batch of DataCamp assignments*. It is also strongly recommended that you complete the following modules (multiple lessons each) outside of class in parallel with the course content (in the following order, see the schedule below): Getting and Cleaning Data (4 lessons), Exploratory Data Analysis (15 lessons), Statistical Inference (14 lessons), and Regression Analysis (13 lessons).

Guacamole (Virtual Machine) Access

If you are having trouble installing R or RStudio locally on your personal machine, you may access a virtual PC using USD's Guacamole system. To run RStudio remotely using a virtual machine (VM) using USD's Guacamole system, login to <https://guacamole.sandiego.edu/> using your AD credentials. There should be a list of forty virtual machines. When you choose one of the VMs after

logging in (that isn't being used currently - may have to check a few of them), you should see a windows Destkop screen that you can navigate as if it was your own computer (you're controlling it remotely since you've "remoted in") and RStudio is accessible from the desktop. Note: you may have issues saving your workspace / script files on the VM as well as navigating / specifying the active working directory, so all work done should be saved locally (you can email it to yourself). The virtual machines may have their memory wiped between sessions.

Professionalism (attendance, participation, originalism)

Attendance is critical for success in this class. It is in the student's best interest to attend class every day. It is understood that events may cause one to have to miss class to which advance notice is preferred. A portion of the course grade is allocated towards professionalism (includes attendance, class participation and professionalism). Failure to be in class in the event attendance is taken can result in a reduction in this grade component. This can also include being late after attendance has been taken or leaving class early. Furthermore, failure to attend more than 50% of all class lecture (and/or lab sessions) will result in failing the course. Lastly keep in mind that attendance, participation, and professionalism are three separate items, that is, you need to more than just show up; you are expected to be prepared and contribute to class activities and discussion. Also engaging in Internet surfing, messaging, games, or sending emails during class will not be tolerated and will result in a reduction in points. Plus, you will find it much easier to keep up and master the material if you attend regularly. You are responsible for all material covered in class, assigned out-of-class, or assigned during a class even if you were absent. Please contact me as soon as possible if you know in advance that you must miss a class. If you use AI to assist you in troubleshooting code (ie: ChatGPT) or use publicly accessible use cases (ie: Stackoverflow) to reference and repurpose code, this is appropriate and should be *explicitly referenced* in your report (and works cited / data sources if applicable). Otherwise, using AI to generate unoriginal content (ie: to write the report and to write your code) in lieu of using your own faculties is not acceptable and will be met with a heavy point deduction on the course assessments. If you are not sure if you are acting appropriately, consult the course instructor. Any exam missed for a legitimate, university approved and documented reason may be made up or re-weighted at the discretion of the instructor. **Canvas quizzes cannot be made up FOR ANY REASON since they are available on day 1 with deadlines fully set in the Canvas calendar. DO NOT EMAIL ME ASKING ME TO PROVIDE AN EXTENSION ON A CANVAS QUIZ.**

Quizzes

At the end of every week (starting week 2), there will be a brief (10 questions) asynchronous Canvas quiz involving a variety of question formats (including multiple choice, numeric response, and check all that apply). The quiz questions will involve a combination of conceptual questions and lab-oriented tasks related to concepts introduced that week (with a slight cumulative component due to how all of the topics are built on one another). The lab-oriented questions may involve programming and analyzing data in R / RStudio. The quizzes will be made available on Thursday and can be submitted any time prior to the Sunday 11:59pm deadlines for the associated quiz for that week. It is recommended that you complete the DataCamp exercise for

the week *prior* to taking the quiz for the week as it (the quiz) may draw on information from the DC exercise. You get as many submission attempts as you'd like and only your best attempt will be scored. You will get to automatically drop your lowest quiz score at the end of the term (a missed quiz will be your dropped quiz). *Canvas Quizzes cannot be made up for any reason since they have been available since day 1 and due dates made known on day 1 in the Canvas calendar. If you fail to submit a quiz prior to the deadline, you will not be able to see the questions and will have to ask a classmate if you would like to review the questions for study.* **DO NOT EMAIL ME ASKING TO PROVIDE LENIENCY ON A CANVAS QUIZ IF YOU MISS THE DEADLINE.**

DataCamp Programming Assignments

This course will leverage DataCamp, an online learning platform, to help introduce concepts related to both the course content (theory) and applications that will leverage a state-of-the-art software visualization tool, Tableau. You will need to enroll in the DataCamp course (you should have an email invitation in your USD inbox) in order to access and submit assignments (the first two are due before 2/11/25). There is approximately one assignment per week through the end of November. Assignments can take roughly an average of 4 hours for any given week, so you should allocate your time carefully to these activities. You will receive one "point" for completion of a chapter in Canvas following the due date of an assignment (see below). *You must complete the chapter so that the completion "bar" reads 100% for credit - without 100% completion, you will receive ZERO credit for the chapter.* Otherwise, your score on these assessments will be determined by the product of two numbers: the fraction of the modules completed and the fraction of XP (experience) points gathered in the modules - there is 58300 total possible XP for the class, but you will only be scored relative to a total of 56000 XP provided you completed all of the modules - so you can miss or "skip" roughly 2000 XP points (see the next part about "skipping"). All of the due dates for the exercises have been set and you can work on these at your leisure (so get a jump start if you'd like!). DataCamp exercises will generally involve brief tutorial /videos followed by questions and applications. Note that you can "skip" modules if you happen to get stuck, but will sacrifice the XP associated with that module (you can take hints which are less costly to help you move through the modules). Expectations are that students wouldn't need to skip more than one or two parts of any chapter. Here is the assignment summary for the course:

| Course | Due Date |
|---|-----------|
| Introduction to Statistics | 2/11/2026 |
| Introduction to R | 2/13/2026 |
| Intermediate R | 2/20/2026 |
| Introduction to Importing Data in R | 2/25/2026 |
| Introduction to Data Visualization with ggplot2 | 3/04/2026 |
| Introduction to the Tidyverse | 3/06/2026 |
| Intermediate Data Visualization with ggplot2 | 3/11/2026 |
| Introduction to SQL | 4/8/2026 |
| Intermediate SQL | 4/15/2026 |
| Joining Data in SQL | 4/22/2026 |
| Cleaning Data in R | 5/5/2026 |

Exams

There will be two midterm exams given throughout the term. The first midterm will take place in class on Tuesday, 10 March 26. The second midterm exam will take place in class on Thursday 28 April 2026. The final exam is scheduled by the registrar, but there is no formal exam - instead, the group project presentations will take place during the final exam period (as well as during the last week of classes depending on the number of groups). The exams may have cumulative elements and may require both conceptual and technical programming skills in the assessment format similar to the type of questions asked in the weekly quizzes. Sample exams and review materials will be provided one week in advance of the exams.

Course Project

The goals of the team project are (1) to integrate and apply what you have learned in the class; (2) to give you experience working in team situations; and (3) to enhance your communication and presentation skills. Throughout the course, you will complete two major analytics tasks: a descriptive task and a database task. Details for these two tasks will be made available after the appropriate material has been covered. In general, the projects will provide you the opportunity to express yourself creatively within the constraints of the task using the statistical software package to showcase your analysis and demonstrate your competency with coding and debugging your work. You should finish the course with a portfolio-worthy project to take with you to the job market (or graduate school). See the AI policy below regarding the use of artificial intelligence to assist with project tasks.

Grading:

| Grade Weights | Percent |
|--|---------|
| Assignments, Quizzes, Datacamp | 30% |
| EXAM ONE | 15% |
| EXAM TWO | 15% |
| Project | 30% |
| Professionalism (attendance, participation) | 10% |
| TOTAL PERCENT | 100% |

The course is graded on a relative curve (as is any college course). In particular, students will all be ranked from highest to lowest course score according to your final course grade calculated from the raw exam score weighting above. Letter grade assignments will depend on your percentile ranking in the class and a subjective assessment by the instructor in borderline cases (say, if there was marked improvement).

In the past, a student could typically guarantee themselves an A by ranking in the top 25% of students in the course and a B by ranking in the top 60% of students. The median score for the course curves to around a B-.

During the course of a semester a student might have concerns or challenges and wish to have a grade appeal. This is fine; however, any grade petition/appeal must be done **within two days** after the grade has been returned. All grade petitions/appeals can be done in person but MUST also be accompanied with an email petition/appeal submission. Typically, the median score for this course curves to a B+/A-. The instructor reserves the final judgement in all grade assessment matters. It is important to note that the **Canvas grade center does not automatically apply the weights above – you will need to compute your composite score manually as outlined above to apply the letter grade rubric at the end of the term.**

All grade petition/appeals must be resolved after one week. **NO GRADES or GRADE PETITION/APPEALS will be entertained after ONE WEEKS.** Once this time has elapsed the window for grade reconsideration is expired. Please note that no grade appeals will be entertained during a class session, you must either come to my office hours, or contact me via phone or email.

| Scale | Total Points |
|-------|---------------|
| A+ | >97 % |
| A | 93 % - 96% |
| A- | 90 % - 92.9% |
| B+ | 87% - 89.9% |
| B | 83% - 86.9% |
| B- | 80 % - 82.9% |
| C+ | 77% - 79.9% |
| C | 73% - 76.9% |
| C- | 70% - 72.9% |
| D+ | 67% - 69.9% |
| D | 63% - 66.9% |
| D- | 60% - 62.9% |
| F | 59% and below |

Important NOTICE: This grading scale is extremely strict. You must have the exact minimum number of points to receive the grade you desire, that is exactly 930 is an A, but 92.9 (or 929.9999) is an A-, and so

on. Therefore, note that in advance all petitions to round up to the higher point value will be denied. Also note the instructor reserves the right to modify or change any part of this syllabus at any time.

Classroom Decorum and Email

To avoid distracting others in the classroom, please arrive on time. In extreme cases the instructor reserves the right to decrease the letter grade by an entire letter for a student due to inappropriate behavior. Please restrict the use of email to the minimally necessary volume, put the course number (BUAN 370) in the subject of the email

– I teach other classes, and this is the fastest way for me to know who I'm responding to. Put your *full name* at the end of email messages as well. Your email decorum should be professional. You should make sure you address the email formally and properly and use college level grammar.

Course Evaluations

An online evaluation will be made available to you near the end of this course. Your timely and considered feedback is valuable to us and an important element of your learning experience.

Artificial Intelligence (AI)

While not prohibited in every circumstance, using AI during an exam would constitute an explicit violation of the academic integrity policy outlined above unless the instructor explicitly permits. Otherwise, students are encouraged to leverage AI in (for example) project tasks, to help debug and write code, but should cite specifically how it was used (including a log of the prompts that generated any output). AI should not be used to write any part of the project narrative and should generally not be used as a substitute for creative / critical thinking faculties. If you are using AI and are unsure of violating university policy, consult your instructor.

Academic Advising

To schedule an appointment with an advisor in Student Success, email
KSBStudentsuccess@sandiego.edu

Academic Integrity

You, your colleagues, faculty, staff, and alumni are the University of San Diego. These and many other persons have worked very hard since the founding of USD in 1949 to build a quality university. The philosophy and mission of USD <http://www.sandiego.edu/about/mission-vision-values.php> emphasizes the idea of personal and academic integrity. The following is a synopsis of the academic integrity policy. For more information click on this hyperlink <http://www.sandiego.edu/associated-students/branches/vice-president/honor-council/integrity-policy.php> or download this pdf <http://www.sandiego.edu/conduct/documents/HonorCode.pdf>

“All members of the University community share the responsibility for maintaining an environment of academic integrity since academic dishonesty is a threat to the University. Acts of academic dishonesty include: a) unauthorized assistance on an examination; b) falsification

or invention of data; c) unauthorized collaboration on an academic exercise; d) plagiarism; e) misappropriation of resource materials; f) any unauthorized access of an instructor's files or computer account; or g) any other serious violation of academic integrity as established by the instructor."

Academic Dishonesty will not be tolerated in any form. Helping each other study is anticipated. However, only original work will be accepted. There will be no sharing of materials, wearing hats, or using cell phones/PDAs during tests. All assignments unless otherwise noted are individual assignments. If an incident of academic dishonesty occurs in this course the student could receive a grade of "F" for the semester and could possibly face further disciplinary action.

Student Disability Policy

The University of San Diego complies with the American with Disabilities Act and Section 504 of the Rehabilitation Act. It is a University of San Diego Disability Services policy that when students are scheduling exams (midterm exams or final exams) in the DLDRC, they must submit an "Authorization to Administer Exam" at least one week prior to the exam date. If a student does not give DLDRC this notice, the DLDRC can deny them the right to the accommodation, as the DLDRC is not given adequate time to prepare (and set up office space during established exam scheduling times.) Please provide me (your instructor) with a Letter of Accommodation drafted by the DLDRC as soon as possible. There are no retroactive accommodations for Disabled students. As your instructor I am not obligated to provide accommodations until I receive the Letter of Accommodation drafted by the DLDRC. If a student does not give the Letter of Accommodation to me (your instructor), within adequate time to make exam arrangements, I am not obligated to fulfill any such request. If you have any questions or concerns about the process please contact, the USD Disability Services office (Serra Hall, Rm 300), phone at 619-260-4655 or via email at disabilityservices@sandiego.edu as soon as possible.

Student Athletes

You are responsible for providing me advanced written notice, (email), any time you will be unable to attend class. You are also responsible for determining alternate dates/times to make up missed work in class. Without written notice, no assignments, quizzes or tests, either due or taken in class, can be made up.

Late Policy

You will, on occasion, experience times when life intrudes and you need to turn in an assignment late. Excusable situations where I am notified ahead of time (with appropriate documentation) will be handled on a case by case basis. Otherwise, late assignments will receive a score of zero.

General Student Conduct

The University of San Diego School of Business expects its students to conduct themselves in a professional manner at all times. Its students are generally individuals who are preparing for career employment. An integral part of their career and professional development is the expectation that they will conduct themselves during the educational processes in the same manner as will be expected in an employment situation. The University of San Diego Student Code of Rights and Responsibilities is published online at <https://www.sandiego.edu/conduct/the-code/>.

Food Insecurity & Pantry

<https://www.meetatusd.com/toreros-against-hunger>

The goal of Toreros Against Hunger at the University of San Diego is to serve as occasional food relief for University of San Diego students experiencing food insecurity while actively decreasing the amount of food going to waste on campus.

Food insecurity broadly defined is “the state of being without reliable access to sufficient quantity of affordable, nutritious food.” Indicators of food insecurity include skipping meals and/or cutting the size of meals due to lack of financial resources, experiencing hunger but not eating and/or the inability to afford balanced meals.

<https://www.sandiego.edu/sustainability/initiatives/social-justice.php>

Counseling Center

<https://www.sandiego.edu/counseling-center/>

We strive to facilitate students' personal growth and enhance their academic success through accessible, culturally congruent clinical and outreach services. We work in collaboration with other Wellness and university departments and community partners.

- A counselor-on call is available to consult about after-hours urgent psychological concerns at all times. The counselor-on call can be reached by calling 619-260-4655 (24 hours a day, 7 days a week). Please contact the Department of Public Safety to access emergency services (x2222 on any campus telephone, otherwise call 619-260-2222).
- The 24-hour San Diego Access and Crisis Line (1-888-724-7240) also offers crisis intervention, information, and referrals. <https://www.sandiego.edu/conduct/the-code/rules-of-conduct.php>

Tentative Course Schedule of Select Topics (Subject to Change)

| Week | New |
|-------------|---|
| 1 | Intro to Analytics / Branches of Problem Types/ Exploratory Data Analysis /Getting Started / |
| 2 | Descriptive Statistics in R / An Example/ Boolean Logic / For & While Loops / Automating Cleaning Example |
| 3 | Simulation/ GGPLOT/ Preprocessing & Cleaning Data Fundamentals / Data Structures |
| 4 | Panel Data Example / Cross Sectional Cleaning Example w / NHIS Survey Data |
| 5 | Querying dataframes with dplyr using the Tidyverse |
| 6 | Combining dplyr and ggplot2 for descriptive analysis and data visualizations |
| 7 | Analytics Programming Exam / Stochastic Simulation / Monte Carlo |
| 8 | Intro to Database / Overview / Relational Databases / Entity-Relationship Diagram ERD |
| 9 | ERD / Database Design / Normalization |
| 10 | SQL / Connection to SQL Server / Writing SQL CRUD Queries |
| 11 | SQL Aggregates/ Joins Two Tables / Nested Queries / Views |
| 12 | SQL IIS Interactions / SQL Injections / Data Warehousing / Other OOB DB |
| 13 | SQL Skills Exam |
| 14 | Dealing with Missing Information, TIDY Data , Deletion vs. Deletion when necessary |
| 15 | Final Project Work Week |
| | FINAL EXAM - PROJECT PRESENTATION |