

# Introduction to Data Analytics

ITP 249 - 4 Units - 2020 Spring

Section 31832R - Tuesday 3:30-6:50pm, ZHS 252

---



## Course Description

Data is now an integral part of our lives and to be successful in today's business landscape, we need to be able to leverage data to make critical business decisions. This course will teach students how to use data to make those decisions confidently.

## Objectives

After completing this course, students will be able to:

- Use Excel, SQL, NoSQL, and leading industry tools
- Collect, clean, and analyze data from multiple sources
- Pose questions, collect relevant data, analyze data, interpret data and provide insights
- Present data-driven insights using data visualization and dashboards
- Tell compelling stories with data

## Prerequisites

None

## Lectures and Lab

Tuesday 3:30PM to 6:50PM in ZHS 252

## Instructor

Calvin Nguyen <cnt.nguyen@usc.edu>

Office Hours: after class Tuesday 6:50PM to 7:50PM in ZHS 252

Contact Instructor for all Lecture-related questions.

## Teaching Assistant

Kengo Miyakoshi <miyakosh@usc.edu>

Chen Zhou <zhou763@usc.edu>

Anita Huang <[zixinhua@usc.edu](mailto:zixinhua@usc.edu)>

Office Hours: TBD

Contact Teaching Assistants for all Assignment-related questions.

## Course Notes

blackboard.usc.edu

Lecture slides and any supplemental course content will be posted to Blackboard. All announcements for the course will be posted to Blackboard.

## Textbook

No official textbook required.

### Optional Books:

Robert Stine and Dean Foster. Statistics for Business Decision Making and Analysis. Essex, UK: Pearson Education Limited, 2014. (SF) ISBN-13: 978-0134497167

Carlos Coronel, Steven Morris and Peter Rob. Database Systems: Design, Implementation, and Management. Boston, MA: Cengage Learning, 2014. (CMR) ISBN-13: 978-1111969608

Additional reference material will be provided in class as needed.

## Software

Students can use their own computers or login to the Viterbi Virtual Lab. The software will be provisioned through the Viterbi Virtual Lab. Students are introduced to several tools most used in the industry:

- Excel and Microsoft Access
- MySQL and MySQL Workbench
- NoSQL and MongoDB
- SAS

## IT Help

USC IT (ITS): <https://itservices.usc.edu/contact>

Viterbi IT: <https://viterbi.usc.edu/resources/vit/contact-us.htm>

## Grading

The weight of the graded material during the semester is listed below:

Individual Homework Assignments	30%
Team Project	10%
Exam – Midterm	30%
Exam – Final	30%

The following grading scale will be used to determine your letter grade:

A	100-95
A-	94-90
B+	89-87
B	86-83
B-	82-80
C+	79-77
C	76-73
C-	72-70
D+	69-67
D	66-63
D-	62-60
F	59 or below

## Individual Homework Assignments

This course will make use of Blackboard for assignments. All assignments will be posted to Blackboard under the “Assignments” section. Each assignment will include instructions, a due date, and a link for electronic submission. Assignments must be submitted using this link.

## Team Project

For a hands-on experience, there will be a final team project. The goal of the project is to solve real world problems using data analytics. This is a team project with each consisting of 2 to 3 members. Each team will work together to identify a problem to solve, collect the necessary data, prepare, clean and format the data, analyze the data, create visualizations, dashboards and models to analyze and understand the problem and to use insights to develop solutions.

The team project grading breakdown is listed below:

Project Statement, Methodology and Accuracy	25%
Final Report	50%
Peer Evaluation	25%

## Policies

Students are expected to attend and participate in lecture discussions, in-class exercises and team meetings. Attendance will be taken during lecture sessions electronically. Do not share the code with students that are not in the room; doing so is an academic integrity violation. If they would like to be considered for an excused absence, email the instructor and include name, week (1-15), day, date, reason, and documentation.

Students are responsible for completing individual assignments and their fair share of team assignments by stated deadlines. Assignments turned in late will have 25% of the total points deducted from the graded score for each late day.

No make-up exams (except for documented medical or family emergencies) will be offered. If they will not be able to attend an exam due to an athletic game or other valid reason, then they must coordinate with the instructor before the exam is given. They may arrange to take the exam before they leave, with an approved university personnel during the time they are gone, or within the week the exam is given. If students do not take an exam, then they will receive a 0 for the exam.

## Disability Services and Programs

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to your course instructor (or TA) as early in the semester as possible. If you need accommodations for an exam, the form needs to be given to the instructor at least two weeks before the exam.

DSP is located in STU 301 and is open from 8:30am to 5:00pm, Monday through Friday. Contact info: 213-740-0776 (Phone), 213-740-6948 (TDD only), 213-740-8216 (FAX), [ability@usc.edu](mailto:ability@usc.edu), [http://sait.usc.edu/academicsupport/centerprograms/dsp/home\\_index.html](http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html).

## Statement on Academic Conduct and Support Systems

### Academic Conduct

**Plagiarism** – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” <https://policy.usc.edu/scampus-part-b/>. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct>.

### Support Systems

**Student Counseling Services (SCS) - (213) 740-7711 – 24/7 on call**

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention. <https://engemannshc.usc.edu/counseling/>

**National Suicide Prevention Lifeline - 1-800-273-8255**

Provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week. <http://www.suicidepreventionlifeline.org>

**Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-4900 - 24/7 on call**

Free and confidential therapy services, workshops, and training for situations related to gender-based harm. <https://engemannshc.usc.edu/rsvp/>

**Sexual Assault Resource Center**

For more information about how to get help or help a survivor, rights, reporting options, and additional resources, visit the website: <http://sarc.usc.edu/>

**Office of Equity and Diversity (OED)/Title IX Compliance – (213) 740-5086**

Works with faculty, staff, visitors, applicants, and students around issues of protected class. <https://equity.usc.edu/>

**Bias Assessment Response and Support**

Incidents of bias, hate crimes and microaggressions need to be reported allowing for appropriate investigation and response. <https://studentaffairs.usc.edu/bias-assessment-response-support/>

**The Office of Disability Services and Programs**

Provides certification for students with disabilities and helps arrange relevant accommodations. <http://dsp.usc.edu>

**Student Support and Advocacy – (213) 821-4710**

Assists students and families in resolving complex issues adversely affecting their success as a student EX: personal, financial, and academic. <https://studentaffairs.usc.edu/ssa/>

**Diversity at USC**

Information on events, programs and training, the Diversity Task Force (including representatives for each school), chronology, participation, and various resources for students. <https://diversity.usc.edu/>

**USC Emergency Information**

Provides safety and other updates, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible, <http://emergency.usc.edu>

**USC Department of Public Safety – 213-740-4321 (UPC) and 323-442-1000 (HSC) for 24-hour emergency assistance or to report a crime.**

Provides overall safety to USC community. <http://dps.usc.edu>

## Course Outline

	Topics	Reading	Homework
<b>Week 1</b> <b>Jan 14</b>	<b>The Value of Data</b> <ul style="list-style-type: none"> <li>• Explanation of course objectives and tools</li> <li>• Syllabus Review</li> <li>• Discussion of the value and impact of data-driven decision making</li> <li>• Discussion of visual analytics and common presentation strategies</li> <li>• Excel Analytics</li> <li>• Brief history of databases and their role in information systems</li> <li>• Different types of databases and their organizational context</li> <li>• Survey of DBMS</li> </ul>	Check Blackboard	HW1 MS Access
<b>Week 2</b> <b>Jan 21</b>	<b>Foundations of Databases and SQL Data Modeling</b> <ul style="list-style-type: none"> <li>• Data models</li> <li>• Business rules</li> <li>• Relational and entity-relationship modeling</li> <li>• Entities, attributes, relationships</li> <li>• Keys: primary, foreign, candidate, surrogate, super</li> <li>• Minimum and maximum cardinality</li> </ul>		HW2 MS Access
<b>Week 3</b> <b>Jan 28</b>	<b>Designing Entity Relationship Diagram</b>		HW3 ERD
<b>Week 4</b> <b>Feb 4</b>	<b>Normalization</b> <ul style="list-style-type: none"> <li>• Anomalies and the need for normalization</li> <li>• Normal forms</li> <li>• First, second, third normal forms</li> <li>• Denormalization</li> <li>• Dependency Diagrams</li> </ul>		HW4 Normalization
<b>Week 5</b> <b>Feb 11</b>	<b>Introduction to SQL</b> <ul style="list-style-type: none"> <li>• Database structures</li> <li>• Introduction to SQL's SELECT statement with WHERE clauses</li> <li>• Query command tools: GROUP BY, HAVING, DISTINCT, COUNT, AND, and OR</li> <li>• Conditional operators: =, !=, &gt;, &lt;, IN, NOT IN, and BETWEEN</li> <li>• Aggregation functions: MIN, MAX, SUM, AVG, and COUNT</li> </ul>		HW5 SQL
<b>Week 6</b> <b>Feb 18</b>	<b>Combining Data in SQL</b> <ul style="list-style-type: none"> <li>• Appending similar data together</li> <li>• Combining data from different tables together</li> <li>• Commands for combining data: JOIN and UNION</li> </ul> <b>Cleaning Data and Creating Multiple Joins</b> <ul style="list-style-type: none"> <li>• Creating relationships between tables: INNER, RIGHT, FULL OUTER, EXCEPTION and CROSS JOINS</li> <li>• Optimizing queries: WHERE, LIMIT and COALESCE</li> </ul>		HW6 Joins
<b>Week 7</b> <b>Feb 25</b>	<b>Subqueries</b> <ul style="list-style-type: none"> <li>• Asking multiple questions in a single query</li> <li>• Nesting queries</li> <li>• Multi-step aggregation or filtering</li> </ul>		HW7 Subqueries
<b>Week 8</b> <b>Mar 3</b>	<b>Midterm Review</b>		
<b>Week 9</b> <b>Mar 10</b>	<b>Midterm Exam</b>		

<b>Week 10</b> <b>Mar 17</b>	<b>Big Data Analytics</b> <b>NoSQL</b> <ul style="list-style-type: none"> <li>• Drawbacks of SQL</li> <li>• Why NoSQL</li> <li>• Non-relational databases</li> </ul>		
<b>Week 11</b> <b>Mar 24</b>	<b>Fundamentals of MongoDB</b> <ul style="list-style-type: none"> <li>• Data storage</li> <li>• Data Retrieval</li> <li>• Queries</li> </ul>		HW8 MongoDB
<b>Week 12</b> <b>Mar 31</b>	<b>MongoDB Aggregation Framework</b>		HW9 Aggregation
<b>Week 13</b> <b>Apr 7</b>	<b>Basic Data Visualization</b>		HW10 Visualization
<b>Week 14</b> <b>Apr 14</b>	<b>Advanced Charts &amp; Dashboards</b>		
<b>Week 15</b> <b>Apr 21</b>	<b>Final Project</b>		
<b>Week 16</b> <b>Apr 28</b>	<b>Final Review</b> <b>Final Project Due</b>		
<b>Week 17</b> <b>May 5</b>	<b>Study Week</b>		
<b>Week 18</b> <b>May 12</b>	<b>Final Exam</b> <b>Tuesday, May 12 @2-4pm</b>		