

CRPLAN 5890: Applied Data Analysis for Transportation and Planning Applications, Spring 2021

Version 2, 28 October 2020

Instructor name: Andre L. Carrel
Email: carrel.20@osu.edu
Office hours: TBD

Meeting time: Tue/Thu 2:20 pm - 3:40 pm
Meeting location: Online

COURSE INFORMATION

All information in this syllabus is tentative and may be changed in the future. Changes to evaluation criteria will only be made after consultation with the class.

DESCRIPTION: This course features a broad overview of a variety of data sets, software tools, and statistical methods for transportation and urban planning applications. Emphasis is placed on the possible uses and characteristics of such data, as well as potential pitfalls, in order to build familiarity with domain-specific data. You will develop skills to process, store, visualize, and analyze such data. You are not required to have prior coding experience, but should have taken an introductory statistics course at the undergraduate level.

FORMAT: This course meets twice a week. The instructor is also available during office hours and by appointment.

REQUIRED MATERIALS

- Personal laptop
- R, available for free at www.r-project.org
- RStudio Desktop version, available for free at www.rstudio.com/products/rstudio/download/
- SQLite Studio, available for free at www.sqlitestudio.pl

ASSIGNMENTS: There will be regular assignments. Announcements will be made in class. You are expected to write a report on each assignment describing how they addressed the assignment questions and presenting their results. Reports are expected to be written clearly, formatted properly, and to use full sentences, with a mix of text, applicable code, and figures. Figures should be clearly labeled. From the report, a reader should be able to understand what was done, why it was done, and what the results were.

FINAL PROJECT: To help you practice your data analysis skills, you will be asked to choose a data analysis project, involving a publicly available transportation data set or a data set that you have access to. The project may be completed in groups of two, and you will present it at the end of the semester. You should choose the project based on a set of questions that you would like to answer rather than the type of data analysis technique that you would like to apply. The final project report must apply at least 3 distinct techniques.

Project Milestone #1: Meet with the instructor. Provide a one-paragraph description of the data set and big picture outline of questions you will ask.

Project Milestone #2: Give a short (3-4 slide) presentation covering the overarching aim/objective, the questions you are asking, the methods you will be applying, and any preliminary results.

Project Milestone #3: Give a final presentation describing the project you completed, including the objective, dataset (or datasets), data transformation performed, methods applied, findings, and lessons learned from a technical point of view.

CRITERIA FOR EVALUATION OR GRADING

Assignments	40%
Final project	40%
Participation	20%

COURSE POLICIES

ATTENDANCE: Students are expected to attend all scheduled class meeting times as outlined in the course syllabus. There are five situations which constitute an “excused absence”. They are:

- **Personal illness:** Students who are too ill or injured to participate in class may be absent from class for a limited time. Students must notify the instructor of their absence as soon as they are aware of it. Upon request, students must provide written documentation from a physician stating that the student cannot participate in class.
- **Death of a member of the student’s immediate family:** Students who have missed class due to a death in the family must provide documentation of the death (death certificate, obituary, etc.).
- **Military of government duty:** Please notify the instructor prior to service.
- **University or Knowlton School sanctioned events:** Students who will be participating in University or Knowlton School sanctioned events must provide the instructor with a copy of the event announcements or schedules.
- **Major religious holiday:** Students who will be observing a religious holiday must provide date/event written notification to the instructor within the first two weeks of the semester.

DEADLINES: Students who miss deadlines due to valid and documented extenuating circumstances may submit the required work at a date agreed upon with the instructor. Students should contact the instructor to arrange a discussion within one week of the missed classes or work. Unexcused late submissions will not be accepted, incomplete projects will be evaluated in relation to their degree of completion, and a student is present only if s/he displays sufficient preparation for the course to the instructor. Students should be aware that grade records are not required to be kept longer than two terms beyond the course offering. Any issues about grades should be reported to the instructor as soon as possible, and no later than the next active term.

COMMUNICATION: Students must check their university email daily.

COURSE OUTLINE

The tentative outline for SP21 is as follows:

Lecture	Date	Content
1	01/12/21	Overview, begin R programming
2	01/14/21	R programming
3	01/19/21	R programming
4	01/21/21	Visualization
5	01/26/21	Transformation
6	01/28/21	Transformation
7	02/02/21	Transformation
8	02/04/21	Hypothesis testing
	02/09/21	<i>No class – Spring break</i>
9	02/11/21	Hypothesis testing. Project milestone #1.
10	02/16/21	Hypothesis testing
11	02/18/21	Clustering
12	02/23/21	Clustering, classification, and prediction
13	02/25/21	Classification and prediction
14	03/02/21	Classification and prediction
15	03/04/21	Classification and evaluation
16	03/09/21	Classification and evaluation
17	03/11/21	Database introduction
18	03/16/21	Data modeling. Project milestone #2.
19	03/18/21	Data modeling
20	03/23/21	Keys and normalization
21	03/25/21	Joins and indexing
22	03/30/21	Regression
23	04/01/21	Regression
24	04/06/21	Regression
25	04/08/21	Time series
26	04/13/21	Text analysis
27	04/15/21	PCA and FA
28	04/20/21	Final presentations -- Project milestone #3