CSC 105 Digital Humanities - Spring 2023 Syllabus

Marcus Birkenkrahe

January 7, 2023

1 General Course Information

- Meeting Times: Tue/Thu from 2.30 - 3.45 pm
- Meeting place: Derby science building computer lab 209
- Professor: Marcus Birkenkrahe
- Office: Derby Science Building 210
- Phone: (870) 307-7254 (Office) / (501 422-4725 (Private)
- Office Hours: Tue/Thu 4-4.30 pm, Mon/Wed/Fri 4.15-4.45 pm
- Textbooks: Text mining in Practice with R (Kwartler, 2017), Text mining with R - A Tidy Approach (Silge/Robinson, 2017)
- Lesson plan: follows "Text mining with R" track in DataCamp

2 Standard and course policies

- **Standard Lyon College Policies** are incorporated into this syllabus and can be found at: lyon.edu/standard-course-policies.
- The **Assignments and Honor Code** and the **Attendance Policy** are incorporated into this syllabus also and can be found at: tinyurl.com/LyonPolicy.

3 Objectives

Understanding and practicing the basis of current AI technologies used in Natural Language Processing applications like DALL-E, ChatGPT, Alexa or Siri. Applying text mining techniques to sentiment analysis, word association, and pattern matching.

4 Student learning outcomes

Students who complete CSC 105, "Digital Humanities - working with text data", will be able to:

- Wrangle, visualize and model text data analysis problems
- Manipulate string data with the `stringr` R package
- Mine text data with the Bag-of-Words technique
- Carry out simple sentiment analysis of text data
- Know how to effectively present assignment results
- Complete an independent research project and present results
- Master data science infrastructure for editing and graphing

5 Course requirements

None. Some basic proficiency in R and the "Tidyverse" packages is useful but not required (use DataCamp lessons or Matloff's "fasteR" online tutorial on GitHub).

Some knowledge of, and experience with computers is useful but not critical. Curiosity and discipline are essential. You will gain data literacy skills by taking this course. The course will prepare you for further studies in computer and data science, or in other disciplines that use modern computing, i.e. every discipline, from accounting to zoology).

6 Grading system

You should be able to see your current grade at any time using the Schoology gradebook for the course.

REQUIREMENT	UNITS	PPU	TOTAL	% of TOTAL
Final exam	0	0	0	0
Home assignments	10	15	150	30.
Class assignments	10	10	100	20.
Final project	1	150	150	30.
Multiple-choice tests	10	10	100	20.
TOTAL			500	100.

7 Grading table

This table is used to convert completion rates into letter grades. for the midterm results, letter grades still carry signs, while for the term results, only straight letters are given (by rounding up).

%	**MIDTERM GRADE**	**FINAL GRADE**
100-98	A+	
97-96	A	A (PASSED -
95-90	A-	VERY GOOD)
89-86	B+	
85-80	B	B (PASSED -
79-76	B-	GOOD)
75-70	C+	
69-66	C	C (PASSED -
65-60	C-	SATISFACTORY)
59-56	D+	
55-50	D	D (PASSED)
49-0	F	F (FAILED)

8 Schedule and session content

For important dates, see the 2022-2023 Academic Calendar at: catalog.lyon.edu/202223-academic-calendar

For this course, we use some lessons from the DataCamp track "Text mining with R".

WEEK	DATE	ASSIGNMENT	TESTS
1	Jan 10,12	Wrangling text	
2	Jan 17,19		Test 1
3	Jan 24,26	Visualizing text	Test 2
4	Jan 31, Feb 2		Test 3
5	Feb 7,9	Sentiment analysis	
6	Feb 14,16		Test 4
7	Feb 21,23	Topic modeling	Test 5
8	Mar 2		Test 6
9	Mar 7,9	Intro to Bag-of-Words	
10	Mar 14,16		Test 7
11	Mar 28,30	Word clouds & other visuals	Test 8
12	Apr 4,6		Test 9
13	Apr 11,13	Clustering & tokenization	
14	Apr 18,20		Test 10
15	Apr 25,27	HR analytics case study	
16	May 2		