I. COURSE INFORMATION

LIS 640-01: Data Analytics and Publication
Thursday, 6:30pm – 9:20pm r.611
Office hours: Mon: 10:00-11:30am; Th: 6:50-6:30pm (other days/times available by appointment)

Instructor Information:
Irene Lopatovska
Email: ilopatov@pratt.edu
GA: Katrina Rink krink@pratt.edu
Phone: (will be given in class)
Listserv: Pratt’s LMS

II. COURSE DESCRIPTION

Bulletin Description:
This course introduces quantitative and qualitative data analysis methods used in basic and applied information science research and guides students through the development of the core components of a social science publication. The course focuses on the topics related to research ethics; content analysis and statistics; data representation, modeling and visualization; publication and presentation of findings; research/grant proposal writing. The data collection techniques are also discussed.

Additional Description:
Data analytics course develops skills to uncover data patterns for making informed decisions, shaping actions and beliefs. The course introduces statistical and linguistic data analysis techniques that are foundational for all information-centric fields, and prepares students to ask the right questions, make inferences and present the results. A core course for students in the Data Analytics and Visualization (DAV) program, it also supports interests in digital humanities, data mining, GIS, and general research. No previous knowledge of statistics is required.

Course Goal(s):
The course aims to introduce students to the foundations of research methodology, data analysis and publication process appropriate for the information science disciplines; enable students to perform research, analyze data and make evidence-based decisions in their professional careers and well as their more advanced degrees; enhance professional writing and presentation skills.

Student Learning Objectives:
By the end of the course, student will be able to:
- Demonstrate understanding of how research is conducted and reported in order to conduct their own data analysis, with the intent of disseminating findings
- Demonstrate understanding of descriptive and inferential statistics
- Demonstrate understanding content analysis techniques
- Demonstrate ability to choose appropriate data analysis techniques to answer research questions, determine the degree of confidence in the answer, and the generalizability of the answer
- Demonstrate ability to use existing datasets to address research questions
- Demonstrate ability to prepare data for analysis
- Demonstrate understanding of the data modeling and visualization techniques for representing results from data analysis
- Demonstrate ability to prepare the research/grant proposals, research publications and presentations
<table>
<thead>
<tr>
<th>Week #</th>
<th>Date</th>
<th>Discussion Topics</th>
<th>Readings</th>
<th>Assignments</th>
</tr>
</thead>
</table>
The role of IRB [http://www.youtube.com/watch?v=8fNlyCB3foc](http://www.youtube.com/watch?v=8fNlyCB3foc)  
Research question formulation: [http://www.youtube.com/watch?v=DNOR4-9xS_c ~6min] | Introduction to term project  
Start working on the term project (A6) |
| 3      | Feb 2  | Research designs  | Babbie – The Practice of Social Research (11th edition)  
Ch. 8 Experiments (pp. 221-225)  
Ch. 9 Survey Research (pp. 244-251)  
Ch. 10 Qualitative Field Research (pp. 286-289)  
Babbie – The Practice of Social Research (11th edition in library)  
Ch. 11 Unobtrusive Research (pp. 330-338)  
Sample of dataset sources [you may identify additional sources of data that can address your research question]  
• U.S. Census data [http://www.census.gov/data.html](http://www.census.gov/data.html)  
• American Community Survey data [http://www.census.gov/acs/www/](http://www.census.gov/acs/www/) | Practice designing a study around the problem (in class); have your RQs handy! |
<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 9</td>
<td>WORKSHOP</td>
<td>Open Refine and MS Excel</td>
</tr>
<tr>
<td>Feb 16</td>
<td>Log file analysis Lab</td>
<td></td>
</tr>
<tr>
<td>Feb 23</td>
<td>Qualitative data analysis: content analysis and grounded theory Lab</td>
<td></td>
</tr>
</tbody>
</table>

### Data Sources

- Federal Government data ([data.gov](http://data.gov))
- Pew Research Internet Project ([http://www.pewinternet.org](http://www.pewinternet.org))
- NYC Open Data ([http://nycopendata.socrata.com](http://nycopendata.socrata.com))
- The Inter-university Consortium for Political and Social Research, Institute for Social Research, University of Michigan ([http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp](http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp))

Optional:


### Data Sources

<table>
<thead>
<tr>
<th>Date</th>
<th>Day</th>
<th>Session/Activity</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Mar 16</td>
<td></td>
<td>Spring Break</td>
<td></td>
</tr>
<tr>
<td>10 Mar 23</td>
<td></td>
<td>Basic concepts of inferential statistics -- Quantitative data analysis: comparing samples --</td>
<td>Vaughan – <em>Statistical Methods for the Information Professional</em> Ch. 5 What is statistically significant (pp. 45-66) Ch. 9 Are two samples significantly different? (pp.111-124) Ch.10 Are three or more samples significantly different Optional: Babbie - <em>The Practice of Social Research</em> (11th edition in library) Ch. 16 Statistical Analyses (pp. 450-455) Wagner – <em>Using IBM SPSS Statistics for Research Methods and Social Science Statistics</em> Ch. 2 Transforming Variables Ch. 3 Selecting and Sampling Cases Ch. 4 Organization and Presentation of Information Ch. 5 Charts and Graphs Ch. 11 Editing Output Ch. 12 Advanced Applications Qualitative data analysis (A3) due Quantitative data clean-up (A4) discussed in class</td>
</tr>
<tr>
<td>11 Mar 30</td>
<td></td>
<td>Quantitative data analysis: Inferential statistics Lab</td>
<td>Vaughan – <em>Statistical Methods for the Information Professional</em> Ch. 8 Examining relationships for internal and ratio data – correlation and regression (pp. 93-110) Optional: Babbie - The Practice of Social Research (11th edition in library) Ch. 16 Statistical Analyses (pp. 459-470, 476-478)</td>
</tr>
<tr>
<td>Date</td>
<td>Topic</td>
<td>Resource</td>
<td>Notes</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Apr 6 | Quantitative data analysis: tests for nominal/ordinal data lab      | Wagner – *Using IBM SPSS Statistics for Research Methods and Social Science Statistics*  
Ch. 7 Correlation and Regression Analysis | Status update on the term paper (A6)                                                             |
| Apr 13| Results presentation results and discussion lab                     | Vaughan – *Statistical Methods for the Information Professional*  
Ch. 7 Examining relationships for nominal and ordinal data (pp. 75-92)  
Optional:  
Wagner – *Using IBM SPSS Statistics for Research Methods and Social Science Statistics*  
Ch. 6 Cross-Tabulation and Measures of Association for Nominal and Ordinal Variables | Quantitative data summary due (A5)                                                              |
McCandless D. (2010) TED Talk "The beauty of data visualization" Retrieved from:  
Underwood, J. (2013). Data Visualization best practices. Available at:  
| Apr 27| Publishing data and findings preparing publication links to future work | Vaughan – *Statistical Methods for the Information Professional*  
Ch. 13 Getting sophisticated (pp. 163-182)  
IV. COURSE REQUIREMENTS

Textbooks, Readings, and Materials:


Optional:


Additional Readings: A list of additional readings can be found on LMS course site.

Project(s), paper(s), assignment(s):

**Research Question (Assignment 1):**
Develop a problem statement and a research question/hypothesis. Use personal observations, readings from your other courses, and popular science literature for inspiration. Look for inconsistencies and gaps in existing knowledge and prior research; analyze personal experiences and ask how they can be improved. Explain inspiration and rationale for your question(s): why is it interesting/important, what would be the implications/significance/benefits of the study that would try to address this question or test your hypothesis? Think what data you would need to answer the question(s): are these data already available or need to be collected? What would be expected findings of your study?

The assignment will be discussed in class. The problem statement and a research question(s)/hypothesis should be emailed to instructor prior to the lecture and brought to class for discussion on a due date.

**Data Sources (Assignment 2):**
Identify three (3) sources of data for addressing your research question/problem. The data sources can be all primary, all secondary or a combination of both types. You will need to specify the data type and justify why you choose these particular sources as opposed to others (e.g., availability, convenience, quality). If you choose to include primary source(s), you will need to describe the method(s) you would be using to collect primary data and explain how these methods will help you to obtain necessary evidence to answer the question(s).

When you work on this assignment, think what evidence you need to address your question: raw data (where do you get it from?), somebody else’s analysis and conclusion (lit review). Imagine writing a grant: if you ask for resources/funding to obtain necessary evidence, you need to justify why currently available datasets do not provide the answers you are looking for.

The assignment will be discussed in class. The list of data sources with brief explanations should be emailed to instructor prior to the lecture.

**Qualitative data summary (Assignment 3):**
You will be given a set of qualitative data. Analyze the data using content analysis techniques and suggested software. Submit two files for this assignment: 1) coded data in Excel, AtlasTI or other format, and 2) a publishable quality summary of qualitative findings (e.g., tables/graphs and corresponding narrative). The assignment will be discussed in class. Both assignment 3 deliverables should be emailed to instructor prior to the lecture.

**Quantitative data clean-up (Assignment 4):**
You will receive a set of quantitative data. Clean up and prepare the data for statistical analysis. You do not need to submit this assignment, however, be prepared to show and discuss your dataset in class. The details of the assignment will be discussed in class.

**Quantitative data summary (Assignment 5):**
You will receive a set of quantitative data for lab exercises. Based on your lab work, prepare a 2-4 page write-up describing quantitative data using descriptive statistics (e.g., mean, median, mode, range, N/%, etc.) and identified relationships between data variables using a statistical test of your choosing (t-test, ANOVA, correlation, regression, etc.). In addition to the summarization of the data trends in a narrative, consider including tables or other visualizations to represent and summarize your data.

Submit this assignment to instructor via email and prepare to discuss in class. The details of the assignment will be discussed in class.

**Term Paper (Assignment 6):**
Prepare and submit a 3-5 page draft of a research proposal. The proposal should contain the following sections:
- Abstract
- Introduction (~0.5-1p justification for the proposed research)
- Literature review (~1p summary of main themes in previous work/literature)
- Methods (~1p description and justification for the data collection (if applicable) and analysis methods you would plan to use)
- Conclusion (~1p description of potential findings and their significance)
- References (all items in References should be cited in text, otherwise they cannot be included).

Samples of datasets, copies of data collection instrument, large tables, screen shots and other supporting materials can be included in Appendix. The recommended paper length (without References and Appendixes) is 3000-5000 words. Students can integrate previous deliverables into the term paper. Use APA style, 6th ed. (or 5ed.) for formatting your paper. The paper can be co-authored with other students (in which case you need to submit a narrative explaining each author’s contribution to the paper).

ALTERNATIVELY, you can work on individual research projects with your own data throughout the semester and submit a complete research paper for assignment 6. Intermediate deliverables will have to be discussed with instructor.

Submit this assignment to instructor via email and prepare to present it in class. The details of the assignment will be discussed in class.
SILS E-Portfolio
Starting fall 2012, all students entering the MSLIS degree program are required to complete an e-portfolio that must be approved by their advisor before they will be permitted to graduate. The e-portfolio provides students with an opportunity to showcase their best work from the courses they have taken at SILS, and an opportunity to demonstrate they have met the learning objectives of a Master of Information and Library Science. **Work completed for this course may qualify for the multiple areas of e-Portfolio.** Please talk to an instructor if you would like to include course’ projects into your e-Portfolio.

Detailed information on the learning outcomes, requirements and how to create your e-portfolio is available from: https://www.pratt.edu/academics/information/sils-eportfolio/

**Class discussions:** You are expected to read all the assigned readings and actively participate in all the class discussions. Remember, a large percent of your grade is your class participation!

**NOTE REGARDING WRITTEN ASSIGNMENTS:** all written assignments should be single-spaced. The number of pages is a recommendation, not a strict requirement (e.g., your paper can be longer or shorter than suggested in the assignment description.)

**FACULTY RESEARCH:** please note that your class participation, including your assignments, can be used for research. If you want to opt out, notify the instructor

**Assessment and Grading:**
**Student conduct and grading**
All assignments must be completed to receive a passing grade in the course.

Assignments must be submitted on time, so that they can be graded consistently and discussed in the class session following the due date. Put the date of submission on the cover page as well as your name, the course number, and the title of the assignment.

**ALL ASSIGNMENTS MUST BE SUBMITTED ELECTRONICALLY VIA EMAIL TO THE INSTRUCTOR [UNLESS OTHERWISE SPECIFIED]. THE SUBJECT LINE OF YOUR EMAIL SHOULD CONTAIN THE COURSE NUMBER AND ASSIGNMENT’S TITLE.**

Except for documented medical and family emergencies, assignments submitted late will receive a lower grade for the following reasons:

(a) students who take more time to prepare their assignments have an unfair advantage over their classmates; (b) students who submit their assignments late often benefit from the review in class of their colleagues’ errors.

Late papers will receive a grade but no comments.

**Participation**
We will be discussing readings and assignments in class. You should be prepared to discuss and/or ask and answer questions based on the readings or exercises. **The higher and more substantive participation the higher the grade!**

**Missing classes**
When you have to miss a class, please, notify the instructor and submit a 1 page summary of an article/chapter that you have read for the missed class.

*Students with 3 absences (for any reason, including documented medical reasons) cannot expect to receive an A in the course.

*Students with 4 absences or more will be asked to drop the course.
The Pratt Institute attendance policy can be found at: https://www.pratt.edu/policies-portal/11415/2016-12-05-attendance-policy-final.pdf

**Email communication**
Please, include class number (LIS 696) in all your class related correspondence. Try to be courteous and professional and ALWAYS include your name in the body of the message.

**ALL WRITTEN ASSIGNMENTS MUST BE SUBMITTED ELECTRONICALLY VIA EMAIL**

**Format**
It is important to follow consistent format for all your assignments. The suggested format is APA [the course bibliography is an example of the APA style]:


All written assignments must have the following information in the top left corner of the front page:
- student’s name
- course number
- assignment or exercise number as listed in the Syllabus
- title [when appropriate]
- date

If references to other works (articles, books...) are included in the summary, they *must* be properly cited in the summary, e.g. Tenopir (2004).

**Grading**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4.0</td>
</tr>
<tr>
<td>A-</td>
<td>3.7</td>
</tr>
<tr>
<td>B+</td>
<td>3.3</td>
</tr>
<tr>
<td>B</td>
<td>3.0</td>
</tr>
<tr>
<td>B-</td>
<td>2.7</td>
</tr>
<tr>
<td>C+</td>
<td>2.3</td>
</tr>
<tr>
<td>C</td>
<td>2.0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
</tr>
</tbody>
</table>

For details on Pratt Institute grading system please refer to the *Graduate Bulletin*.

Grades for students’ work will be determined as follows:

| Assignment 1 | 5% |
| Assignment 2 | 5% |
| Assignment 3 | 20% |
| Assignment 4 | 5% |
| Assignment 5 | 20% |
| Assignment 6 (20%paper;5%presentation) | 25% |
| Class participation | 20% |

Total: 100%

**III. POLICIES**

Institute-wide policies listed in the “Community Standards” section of the bulletin:
Students must adhere to all Pratt Institute policies listed in the Pratt Student Handbook under “Student Affairs,” which include policies on attendance, academic integrity, plagiarism, and computer and network use. Full descriptions of these policies are available in the Online Student Handbook at: http://www.pratt.edu/uploads/Online_Student_Handbook.pdf
Policy on students with disabilities:

Anyone requiring special accommodations for disabilities must obtain clearance from the Office of Disability Services at the beginning of the semester. Contact Mai McDonald, Director of Disability Services at: mcdonald@pratt.edu or by phone at (718) 636-3711.

Complete documentation on the institute’s policies and services can be found at: https://www.pratt.edu/student-life/student-affairs/office-of-the-vice-president-for-student-affairs/student-policies/