DATA 620 9040 Data Management and Visualization (2172)
DATA-620

Spring 2017  Section 9040  6 Credits  01/30/2017 to 04/23/2017

Faculty Contact

Caroline Beam Carrie.Beam@faculty.umuc.edu

Course Description

Prerequisite: DATA 610. A presentation of the fundamental concepts and techniques in managing and presenting data for effective data-driven decision making. Topics in data management and design include data design approaches for performance and availability, such as data storage and indexing strategies; data warehousing, such as requirement analysis, dimensional modeling, and ETL (extract, transform, load) processing; and metadata management. Topics in data visualization include understanding data types; data dimensionalities, such as time-series and geospatial data; forms of data visualization to include heat maps and infographs; and best practices for usable, consumable, and actionable data/results presentation.

Course Introduction

This is a six-credit course.

There are 12 weeks in the semester.

The last date to withdraw is given in your online classroom.

Course Outcomes

At the end of this course, students should be able to:

1. Design, construct, and query simple databases using SQL.
2. Understand metadata, ETL, and security issues for databases and data warehouses.
3. Construct and present a wide variety of charts using Tableau to support data-driven business decisions.
4. Analyze and restructure basic numeric and text data using simple Python code.
5. Analyze and present data from a data warehouse to solve a business problem.

Course Materials

Click to access your course materials information (http://webapps.umuc.edu/grcmBook/BPage.cfm?C=DATA%20620&S=9040&Sem=2172)

Grading Information

Grading Criteria
NOTE: For all graded assignments/projects, students will be provided with written feedback and a numeric grade.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Type of Assignment</th>
<th>Frequency of Assignment</th>
<th>Points</th>
<th>Percentage of Total Grade</th>
<th>Outcomes Addressed by Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management Homework</td>
<td>Individual</td>
<td>4 @ 5 points apiece</td>
<td>30</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 @ 10 points apiece</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Management Homework Team</td>
<td>Team</td>
<td>2 @ 10 points</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Management Project Team</td>
<td>Team</td>
<td>1</td>
<td>200</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Data Management Total</td>
<td></td>
<td></td>
<td>250</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>SQL Homework</td>
<td>Individual</td>
<td>5 @ 10 points apiece</td>
<td>50</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SQL Project</td>
<td>Individual</td>
<td>1</td>
<td>200</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>SQL Total</td>
<td></td>
<td></td>
<td>250</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>Visualization Homework</td>
<td>Individual</td>
<td>5 @ 10 points apiece</td>
<td>50</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Visualization Project</td>
<td>Individual</td>
<td>1</td>
<td>200</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Visualization Total</td>
<td></td>
<td></td>
<td>250</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>Python Homework</td>
<td>Individual</td>
<td>5 @ 10 points apiece</td>
<td>50</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Python Project</td>
<td>Individual</td>
<td>1</td>
<td>200</td>
<td></td>
<td>3, 4</td>
</tr>
<tr>
<td>Python Total</td>
<td></td>
<td></td>
<td>250</td>
<td>25%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>1000</td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>
NOTE: During the weeks between assessments, there will be class sessions designed to provide key lecture, discussion, and software information and assistance necessary for successful completion of the assessments.

Course Specific Grading Policies

Each student is expected to:

1. Review and honor all UMUC and TGS policies listed in the classroom and syllabus.
2. Invest approximately 15-20 hours throughout the week.
3. Complete readings and research in advance of each week.
4. View any recorded video to remain updated on key information and course requirements.
5. Participate in the assignments posted for specific weeks.
6. Consistently demonstrate professionalism in your interactions with peers and faculty.
7. Produce high quality, original work that properly cites the words and ideas of others.
8. Uphold the integrity of your UMUC degree by not plagiarizing or cheating.
9. Accept feedback and suggestions with an open mind, taking time to reflect, and then use what you have learned to further improve your performance in future assessments.
10. Notify your professor immediately if you confront a personal emergency that puts you at risk of submitting a late assessment.
11. Honor the rule that "everything written or said within the classroom stays within the classroom."
12. Stay current on news items and postings by visiting the classroom regularly.

The average grade for UMUC graduate courses is a "B." If you complete each assessment at a level consistent with graduate work, you can expect to earn a "B." This does not mean you have to be an expert on the subject matter, but you are expected to put in as much time as is necessary to demonstrate an acceptable level of accomplishment. If you have questions about the subject matter, or if you do not understand an assessment, do not hesitate to get in touch with your faculty member. She or he is there to help you understand and apply this material. You should let your faculty member know how to make your learning experience more rewarding.

If you want to increase your chances of earning an "A," you must pay close attention to the quality of your writing. Superior writing skills certainly can help you to earn a better grade. The degree to which you provide authoritative support in your writing and presentations can also improve your chances of earning an "A." You are expected to refer to the required materials to support your analysis in an assessment. In some cases you will be expected to conduct research and use relevant articles from scholarly journals in the UMUC library's electronic databases as additional sources of information. An example of exceeding expectations is when you consistently take the initiative to seek out and reference such work to help make and strengthen your contributions to the assessments. In addition, you should look for opportunities to incorporate the materials from previous weeks in this course as well as readings and materials from previous courses into your assessments. Following these suggestions can go a long way in helping you to earn an "A."

At the other end of the grade distribution, if you do not take the time to develop an acceptable level of accomplishment of the subject material or do not complete your assessments on time, you could earn a "C." If you do not complete assessments, demonstrate little or no understanding of the subject material, and/or routinely fail to complete assessments on time, you could earn an "F."

For team assessments: Each person on a team will typically receive the same grade for an assessment. The highest possible grade on the team project will be based on what is submitted as the team end-product. However, a team member's grade may be adjusted downward for a specific assessment if the faculty member determines the quality of that person's participation to be substandard. To allow your faculty member to review team member contributions to all team assessments, each team member is required to post all contributions in the team's discussion area.

Review the posted assessment rubrics carefully. They will help you understand the faculty's expectations regarding assessment grading. If you have any questions about a grade you received, contact your faculty member. Your faculty member is here to help you understand the material, your grade, and how you may improve your performance in this course.

Due Date Extension/ Late Penalty Policy

Timely completion of all assessments is critical to student success in a graduate program. Each assessment in this program has a submission due date. Students should take these due dates seriously. It is highly recommended that each student review both the syllabus and the discussion forums thoroughly at the beginning of each course, in order to understand the nature and magnitude of
the work required—then plan in advance to allocate sufficient time to complete each assessment on or before the required due date.

For assessments a faculty member may, at his or her discretion, elect to grant an extension when a student is unable to meet a due date because of an unexpected business, health, or personal emergency beyond the student’s control. For such an extension to be considered, a student must make an extension request before the assessment due date (unless this is physically impossible). The student must support this request with a compelling rationale, indicating why such an extension is justified. If a faculty member grants an extension, he or she will set a new due date for submission of the assessment. There will be no extensions of the new due date.

For late assessment submissions that have not been approved by the faculty member, and for assessments submitted after an extension due date, the penalty will be a five-point reduction in grade for each day that the assessment is late. Faculty will not accept assessments that are twenty or more days late; such assessments will be assigned a grade of zero.

Note that there will be no extensions for the assessments due the last week of class.

**Extra Credit and “Redoing” Assessments**

This is a very fast-paced course, and assessments are due at planned times in the course schedule. The syllabus is designed to build upon each week, and the assessments are chosen carefully to develop skills and build knowledge. It is important that you do a good job on all of the assessments and that you hand them in on time. There is no "extra credit" granted to individuals in this course, and students will not be able to redo assessments.

**Original Work**

Resubmission of the coursework from previous classes (whether or not taken at UMUC), partially or in its entirety is unacceptable, unless prior approval is obtained from the instructor for the specific item. Using coursework, partially or in its entirety, from a previous course without explicit prior approval of your section professor will result in a grade of zero for the assessment.

---

**Project Descriptions**

All of the reading and writing assessments are contained in the Course Schedule section of this syllabus. Detailed guidance for each assessment will be included in the classroom. All assessments are due no later than midnight Eastern U.S. time Sunday unless otherwise indicated. All individual and team assessments must be submitted in pdf, rtf, or Word format, as the instructor requires.

**NOTE:** You should prepare and save all your assessments on your computer before you post them to the classroom. You should never type your assessment directly into the classroom. If your computer crashes or if you are disconnected from the Internet before you finish posting your work, you will lose everything you have not saved. In addition, shortly after the course is over you will no longer have access to any of the materials you posted in the classroom.

**Discussions**

In some weeks, there may be weekly online discussion groups. Students are encouraged to participate early and often, because interaction with each other and with the faculty is a key to learning.

---

**Academic Policies**

**Academic Policies and Guidelines**

**ACADEMIC INTEGRITY**

As a member of the University of Maryland University College (UMUC) academic community that honors integrity and respect for others you are expected to maintain a high level of personal integrity in your academic work at all times. Your work should be original and must not be reused in other courses.

**CLASSROOM CIVILITY**

Students are expected to work together cooperatively, and treat fellow students and faculty with respect, showing professionalism and courtesy in all interactions. Please review the Code of Civility for more guidance on interacting in UMUC classrooms: [https://www.umuc.edu/students/support/studentlife/conduct/code.cfm](https://www.umuc.edu/students/support/studentlife/conduct/code.cfm)
POLICIES AND PROCEDURES

UMUC is committed to ensuring that all individuals are treated equally according to Policy 040.30 Affirmative Action, Equal Opportunity, and Sexual Harassment. Students with disabilities who need accommodations in a course are encouraged to contact the Office of Accessibility Services (OAS) at accessibilityservices@umuc.edu, or call 800-888-UMUC (8682) or 240-684-2287.

The following academic policies and procedures apply to this course and your studies at UMUC.

150.25 Academic Dishonesty and Plagiarism – UMUC defines academic dishonesty as the failure to maintain academic integrity. All charges of academic dishonesty will be brought in accordance with this Policy.

Note: Your instructor may use Turnitin.com, an educational tool that helps identify and prevent plagiarism from Internet resources, by requiring you to submit assignments electronically. To learn more about the tool and options regarding the storage of your assignment in the Turnitin database go to: https://www.umuc.edu/library/libresources/turnitin.cfm.

151.00 Code of Student Conduct

170.40 The following policies describe the requirements for the award of each degree:

170.41 Degree Completion Requirements for the Graduate School

170.42 Degree Completion Requirements for a Bachelor's Degree

170.43 Degree Completion Requirements for an Associate's Degree

170.71 Policy on Grade of Incomplete - The grade of I is exceptional and only considered for students who have completed 60% of their coursework with a grade of B or better for graduate courses or C or better for undergraduate courses and request an I before the end of the term.

170.72 Course Withdrawal Policy - Students must follow drop and withdrawal procedures and deadlines available at https://www.umuc.edu/ under Academic Calendar.

130.80 Procedures for Review of Alleged Arbitrary and Capricious Grading – appeals may be made on final course grades as described herein.

205.06 Calculation Of Grade-Point Average (GPA) for Inclusion on Transcripts and Transcript Requests – Note: Undergraduate and Graduate Schools have different Grading Policies (i.e. The Graduate School does not award the grade of D). See Course Syllabus for Grading Policies.

GRADING

According to UMUC’s grading policy, the following marks are used:

<table>
<thead>
<tr>
<th>Undergraduate</th>
<th>Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 90-100</td>
<td>90-100</td>
</tr>
<tr>
<td>B 80-89</td>
<td>80-89</td>
</tr>
<tr>
<td>C 70-79</td>
<td>70-79*</td>
</tr>
<tr>
<td>Grade</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------</td>
</tr>
<tr>
<td>D</td>
<td>60-69</td>
</tr>
<tr>
<td>F</td>
<td>59 or below</td>
</tr>
<tr>
<td>FN</td>
<td>Grade Pending</td>
</tr>
<tr>
<td>G</td>
<td>Passing</td>
</tr>
<tr>
<td>S</td>
<td>Unsatisfactory</td>
</tr>
<tr>
<td>I</td>
<td>Incomplete</td>
</tr>
<tr>
<td>AU</td>
<td>Audit</td>
</tr>
<tr>
<td>W</td>
<td>Withdrew</td>
</tr>
</tbody>
</table>

* The grade of "B" represents the benchmark for The Graduate School. Students must maintain a Grade Point Average (GPA) of 3.0 or higher. Classes where final grade of C or F places a student on Academic Probation must be repeated. ** The Graduate School does not award the grade of D.

** COURSE EVALUATION SURVEY **

UMUC values its students' feedback. You will be asked to complete an online evaluation toward the end of the term. The primary purpose of this evaluation process is to assess the effectiveness of classroom instruction in order to provide the best learning experience possible and make continuous improvements to every class. Responses are kept confidential. Please take full advantage of this opportunity to provide your feedback.

** LIBRARY SUPPORT **

Extensive library resources and services are available online, 24 hours a day, seven days a week at [https://www.umuc.edu/library/index.cfm](https://www.umuc.edu/library/index.cfm) to support you in your studies. The UMUC Library provides research assistance in creating search strategies, selecting relevant databases, and evaluating and citing resources in a variety of formats via its Ask a Librarian service at [https://www.umuc.edu/library/libask/index.cfm](https://www.umuc.edu/library/libask/index.cfm).

** LEARNING MANAGEMENT SYSTEM SUPPORT **

To successfully navigate the online classroom new students are encouraged to view the Classroom Walkthrough under Help in the upper right menu of the LEO classroom. Those requiring technical assistance can access Help@UMUC Support directly in LEO under the Help menu. Additional technical support is available 24 hours a day, seven days a week via self-help and live chat at [https://www.umuc.edu/help](https://www.umuc.edu/help) or by phone toll-free at 888-360-UMUC (8682).

** SYLLABUS CHANGES **

All items on this syllabus are subject to change at the discretion of the Instructor and the Office of Academic Affairs.

---

** Class & Assignment Schedule **

<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6 of 13</td>
<td></td>
</tr>
<tr>
<td>Session</td>
<td>Module/Topics</td>
<td>Readings/Assessments</td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>
| 1       | Data Management (DM): data architecture, database normalization, data security | Required Reading:  
Inmon and Linstedt, Chapters 1; 2.2; 2.4; 6.1  
Oracle Corporation, "Overview of Database Security," http://docs.oracle.com/database/121/CNCPT/cmntopc.htm#CNCPT89104 (scroll down until you see "Overview of Database Security" as the title)  
MySQL Reference Manual, Section 3.2, Section 3.3.1 through 3.3.4.8  
Optional Reading:  
Weekly Assignment(s):  
0.1. SQL: Show proof of access  
1.1. DM: Normalize the opera database (5 points)  
1.2. SQL: Operations on a single table (CREATE, INSERT, SELECT); run from prompt and script. (10 points).  
1.3. DM: Reading Comprehension Quiz from Inmon and Oracle (5 points) |
<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 2 | Data Management: entity relationship diagrams, metadata management, SQL: relational databases, primary/foreign keys, operations on two tables, joins | Required Reading:  
Inmon and Linstedt, Chapters 5  
Soper, Daniel. "Database Lesson #4 of 8 - Data Modeling and the ER Model," [https://www.youtube.com/watch?v=flfaqiHqjo](https://www.youtube.com/watch?v=flfaqiHqjo) (58 minutes)  
Edina, (N.D.). "What is Metadata?" (6 minutes) [https://www.youtube.com/watch?v=96ncNPPohjo](https://www.youtube.com/watch?v=96ncNPPohjo)  
Jagdish, Deepak, and Daniel Smilkov, "The Revealing Nature of Metadata," at TEDx, [https://www.youtube.com/watch?v=i2a8pDbCabg](https://www.youtube.com/watch?v=i2a8pDbCabg)  
MySQL Reference Manual, Section 3.3.4.9 through Section 3.6.6  
Tips on SQL Join by Professor Majed Al-Ghandour (in classroom)  
Optional Readings:  
Database Analysis [http://db.grussell.org/section004.html](http://db.grussell.org/section004.html)  
Weekly Assignment(s):  
2.1. DM: create an entity-relationship diagram of a given problem (5 points)  
2.2. DM: reading comprehension quiz on Inmon and Linstedt and metadata (5 points)  
2.3. SQL: run a series of queries on a multi-table database (use OrderEntryDatabase script given in class.) (10 points).  
2.4. SQL: use JOIN to extract useful information from a two-table database (Order Entry database will be used) (10 points). |
| 3 | SQL: operations on more complex databases | Required Reading:  
Review previous weeks' SQL and data readings.  
Optional Reading:  
Weekly Assignment(s):  
<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 4       | Data Management: data quality, database design and operations, data governance, data storage and indexing | **Required Reading:**  
Inmon and Linstedt, Chapters 5.5 (Data Governance)  
Schwartzrock, Todd, “Chrysler’s Data Quality Management Case Study,” [https://www.youtube.com/watch?v=N78HpiCD0k](https://www.youtube.com/watch?v=N78HpiCD0k) (10 minutes)  
Edureka (N.D.). 'Data Quality Concepts, Data Quality Tutorial, Data Warehousing Tutorial.' [https://www.youtube.com/watch?v=HWaBdqmTqEA](https://www.youtube.com/watch?v=HWaBdqmTqEA) (25 minutes)  
Narasimhan, Girija, “Entity Integrity Constraints,” [https://www.youtube.com/watch?v=0r-Lu48ZTD0](https://www.youtube.com/watch?v=0r-Lu48ZTD0) (2 minutes)  
Fezalar Computer, "Database Basics Lesson 17 – Referential Integrity and Relationship Rules." [https://www.youtube.com/watch?v=pBFvfKqAOp0](https://www.youtube.com/watch?v=pBFvfKqAOp0) (9 minutes)  
Soper, Daniel. "Database Lesson #7 of 8 - Database Indexes," [https://www.youtube.com/watch?v=Xk3cgUdoieU](https://www.youtube.com/watch?v=Xk3cgUdoieU) (40 minutes)  
Star Schema slides from Professor Majed Al-Ghandour (in classroom)  
MySQL Reference Manual, Section 13.2.6, Section 13.2.9.1  
**Optional Readings:**  
Managing and Sharing Your Data: Best Practice for Researchers, [http://www.data-archive.ac.uk/media/2894/managingsharing.pdf](http://www.data-archive.ac.uk/media/2894/managingsharing.pdf)  
**Weekly Assignment(s):**  
4.1ABC. DM: Team Exercise on data integrity, data quality, and star vs relational schema (10 points).  
4.2. DM: assignment on data indexing (10 points)  
4.3. SQL: create script to read in a file, arrange the data, and write to an outfile (10 points) |
### Session 5: Data Management: Introduction to Data Warehousing, Requirements Analysis

#### SQL: continued practice

<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 5       | Data Management: Introduction to Data Warehousing, Requirements Analysis | **Required Reading:**

- Inmon and Linsteadt, Chapters 3, 6.3, 6.4
- Oracle Corporation, [http://docs.oracle.com/database/121/CNCPT/cmntopc.htm#CNCPT89104](http://docs.oracle.com/database/121/CNCPT/cmntopc.htm#CNCPT89104) – Overview of Data Warehousing and Business Intelligence (halfway down the page). Focus on Table 19-1.
- Overview of Extraction, Transformation, and Loading (ETL), from Oracle Corporation. [http://docs.oracle.com/database/121/CNCPT/cmntopc.htm#CNCPT1468](http://docs.oracle.com/database/121/CNCPT/cmntopc.htm#CNCPT1468)

| | | **Optional Readings:**
|---|---|---|
| | | Data Warehousing and Data Mining lecture slides, by Michael Bohlen, Free University of Bolzano, [https://files.ifi.uzh.ch/boehlen/dis/teaching/DWDM08/slides/dw1.pdf](https://files.ifi.uzh.ch/boehlen/dis/teaching/DWDM08/slides/dw1.pdf)

| | | **Weekly Assignment(s):**
|---|---|---|
| | | 5.1. (Team) DM Assignment: class activity on data warehousing (10 points)
| | | 5.2. SQL: write scripts to implement GROUP BY and ORDER BY commands (10 points).
| | | 6.X. (Team) Read through next week’s Assessment. As a team, formulate a rough work plan and post it to your study group. (0 points, but critical to next week’s 200-point assignment.)
<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 6       | Data Management: ETL, data preparation, data cleaning | Required Reading:  
http://docs.oracle.com/cd/B19306_01/server.102/b14223/ettov.htm  
Titirisca, A. ETL as a Necessity for Business Architectures,  
Weekly Assignment(s):  
| 7       | Introduction to Visualization       | Required Reading:  
Severance, Chapters 1, 2  
Weekly Assignment(s):  
7.1. Tableau graphs: bar charts, XY scatterplots, line graphs, time series analysis, trend lines, covariation on XY scatter, seasonality, outliers (10 points).  
| 8       | Visualization: multidimensional visualization | Required Reading:  
Severance, Chapters 3, 4, and 5  
Pareto Charts, from Tableau Software.  
http://www.tableau.com/learn/tutorials/on-demand/pareto-charts  
Weekly Assignment(s):  
8.1. Tableau graphs: Pareto charts, re-expression of axes, correlation, scatterplot matrices (10 points).  
8.2. Python: Write program(s) in which user enters a score and program converts to a letter grade and/or string (10 points).  
8.3 Python: Write program(s) to use while/for/continue to take a list of numbers from the user, and when given text string ‘done’ as input will compute and print the max, min, and average values (10 points). |
<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 9       | Visualization: time series analysis | Required Reading:  
Analyzing Time Series, https://www.youtube.com/watch?v=aaaljNPHSs  
Review previous weeks’ readings.  
Optional Readings:  
Weekly Assignment(s):  
9.1. Tableau graphs: create and present to the class a 5-minute PowerPoint presentation on a time series analysis (200 points). |
| 10      | Visualization: histograms, boxplots, heatmaps | Required Reading:  
Severance, Chapters 6 and 7  
Weekly Assignment(s):  
10.1. Tableau graphs: analyze multidimensional data using histograms, boxplots, and heatmaps (10 points)  
10.2. Python: extract a series of numbers from inside a formatted string and convert to floating point; do for a large input file and compute average of all such floating point numbers (10 points). |
<table>
<thead>
<tr>
<th>Session</th>
<th>Module/Topics</th>
<th>Readings/Assessments</th>
</tr>
</thead>
</table>
| 11 | Visualization: Geospatial data, infographics, text display | Required Reading:  
Severance, Chapters 8 and 9  
Optional Reading:  
| | Python: lists, dictionaries, and advanced text parsing | Weekly Assignment(s):  
11.1. Tableau: create a geospatial analysis using World Bank Indicators to present global trends in business, finance, population growth, healthcare, or transit. Create infographic. (10 points).  
11.2. Python: Romeo program (read in Romeo’s “Arise, fair sun…” speech, and print a list of each unique word); Mailbox program (read unstructured mail header data, extract sender’s email and print count of number of sent messages.) (10 points)  
11.3. Wordle. Create a “wordle” (from wordle.net) of a favorite poem or essay (10 points). |
| 12 | Visualization: text visualization | Required Reading:  
| | Python: text frequency counts and analysis | Review previous weeks’ readings.  
Weekly Assignment(s):  
12.1. Tableau/Python. Students select a text which recurs over three separate time periods (such as the President’s State of the Union address or a CEO’s letter to shareholders). Write a text parser in Python to extract and count unique words in the text over three time periods. Use Tableau to present the results of the text processing analysis (200 points). |