**Journalism 502 – Data Analysis for Journalists (Fall 2015)**

**Tuesdays 3:15-5:45 pm, Ernie Pyle Hall 210**

**Instructor:** Gerry Lanosga

**Office:**  EP 103

**Office Hours:** 1-3 pm Tuesdays or by appointment

**Phone:** 812-855-0851 (office), 317-697-6083 (cell/text)

**Email:**  glanosga@indiana.edu

**Twitter:** @gerrylanosga

**Delicious:** www.delicious.com/glanosga/j502

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**About the Course**

J502 is concerned with the collection, analysis and interpretation of data in the pursuit of news. As such, this is not just a skills course. Certainly, you will get an overview of lots of digital tools to find stories in data. But more importantly, this course emphasizes how to think about, contextualize and write about the data we encounter all around us. You will generate your own quantitative analyses of data we obtain, but you will also learn to critically evaluate data analysis by others, such as scientists and interest groups. You will gain an understanding of important statistical concepts, perspective on how social science techniques can serve a journalists, and working knowledge of how to obtain data and conduct analyses using tools such as spreadsheets and databases. While this course is aimed at establishing a working journalist’s competency with data, the skills and concepts you learn will be useful in practically any branch of mass communication you choose.

**Learning Outcomes**

After taking this course students will be able to:

* Understand the scientific method and social science research.
* Employ advanced analytical techniques in the practice of journalism.
* Explain the strengths and weaknesses of various analytical techniques.
* Practice ethical standards in the conduct of research.

**Required Texts**

* Joel Best, *Damned Lies and Statistics: Untangling Numbers From the Media, Politicians, and Activists,* updated edition. (University of California Press, 2012)
* Jonathan Gray, Liliana Bounegru, and Lucy Chambers, eds., *The Data Journalism Handbook.* Free online! <http://datajournalismhandbook.org/1.0/en/>
* Brant Houston. *Computer-Assisted Reporting: A Practical Guide.* (Routledge, 2015)
* Philip Meyer. *Precision Journalism: A Reporter’s Introduction to Social Science Methods*, 4th edition. (Rowman & Littlefield Publishers, Inc., 2002)
* David Herzog. Data Literacy: A User’s Guide. (Sage Publications, 2015)
* Additional readings will be distributed through Canvas and other means.

**Assignments and grading**

 Your grade in this class will be based on the following:

* Two exams **(15% each)**
* Analysis/critique of data-driven news story **(10%)**
* Data acquisition **(10%)**
* Short news story based on data analysis **(20%)**
* Final data analysis project **(30%)**

 I use the university’s standard grading scale, calculated automatically in Canvas.

In-class exams

 Midterms will consist of short answer questions and data analysis problems.

Analysis/Critique

 You will be required to find and analyze a recent news story based largely on data. A rubric will be provided for this assignment.

News Story

You will be responsible for finding, reporting and writing a news story that should be newsworthy and of publishable quality. The story should be based on a data source that you obtain and analyze using one or more of the tools we learn about in class. There will be a detailed rubric for this assignment; for now, you should begin to think of potential topics. This is pretty wide open – sports, business, politics, features – anything except opinion pieces. Try, if you can, to pick something that has some currency. About the writing: Although this is primarily a class in reporting methods, I want you to strive to produce publishable work, and so a portion of the grade for your story will be based on writing style, grammar and usage, and overall readability.

Data Acquisition and Data Analysis

For the acquisition assignment, you will be required to identify and attempt to obtain an unpublished government database that holds potential material for a news story. For the analysis, you will be required to analyze a database (I will provide you with one) and write a detailed memo outlining an in-depth story based on your analysis.

**Course Policies**

Participation and Timeliness:

I expect you to submit assignments by the specified deadlines. Except in cases of hardship, there will be a five percent point deduction for each day an assignment is late.

Assignment Format

Unless noted otherwise, all work should be double-spaced, 12-point Times New Roman, with margins of one inch left and right, top and bottom.

Communication:

I will use Canvas to post class resources and send announcements. Please check your official IU e-mail regularly.

The best way to reach me is by e-mail. Please allow at least 24 hours for me to respond to any emails you send. If you would like to have a phone conversation, please e-mail or text your request so we can set up a time that works for both of us.

I encourage you to visit my Delicious page for this course, listed above, regularly. I use this to bookmark links of relevance to what we are doing in the course. You are also welcome to follow me on Twitter, where I frequently post items of interest to both research and professional communities (especially in my areas of interest – journalism history, media law and journalistic practice).

Academic Integrity:

I expect you to follow all University guidelines regarding academic honesty. In brief, that means plagiarism and fabrication will not be tolerated. All work you turn in must be yours and yours alone. You must accurately quote and represent all sources. Making up quotes, inventing sources, lifting material from the Web or elsewhere and presenting it as your own, and collaborating on work with another student without permission are examples of academic dishonesty in J502. You can expect that any form of academic dishonesty will result in an F either for the assignment or for the entire course, depending on the severity of the misconduct. If you are unsure about a particular practice, please see me before you turn in your assignment.

**KEY DATES**

**September 8:** Analysis/Critique due

**October 6:** Exam I held in class

**October 27** News story due

**November 10:** Exam II held in class

**November 17:** Data Acquisition Project due

**November 23-27:** Thanksgiving break; no classes

**December 15:** Final Data Analysis Project due

**Course schedule**

Note: This syllabus is a living document and subject to change depending on a number of factors. I will notify you of any changes via Canvas announcement.

**Week 1/Aug. 25 – So it begins. What is data journalism?**

* Read: Meyer preface and chapter 1 (Canvas); Houston chapter 1 (Canvas); Handbook Introduction; “Computational exploration in journalism” (Canvas); “Data Journalism in the United States” (Canvas); “Computational Journalism\_ACM” (Canvas); and skim the articles posted on Delicious for Week 1: [https://delicious.com/glanosga/j502,week1](https://delicious.com/glanosga/j502%2Cweek1)

**Week 2/Sep. 1 – Data literacy**

* Read: Meyer chapters 2 and 5; Best introduction and chapter 1; Handbook “In the Newsroom”; Paulos selections (Canvas); Herzog “Data Defined” (Canvas); “Critical Questions for Big Data” (Canvas); “Data-Drive Revelation?” (Canvas); “Understanding Data” (Canvas); Houston chapters 3-4; and skim the articles posted on Delicious for Week 2: [https://delicious.com/glanosga/j502,week2](https://delicious.com/glanosga/j502%2Cweek2)
* Hands-on: Excel refresher

**Week 3/Sep. 8 – Obtaining data both online and off (and conceptualizing a data story)**

* Read: [www.google.com/publicdata/directory](http://www.google.com/publicdata/directory)**;** Houston chapters 7, 10; Handbook “Getting Data”; Herzog “Identifying and Requesting Offline Data” (Canvas); FOI handouts (Canvas); selections from “Scraping for Journalists” (Canvas); Vallance-Jones/McKie “Organizing and Writing the CAR Story” (Canvas); browse resources at [www.google.com/publicdata/directory](http://www.google.com/publicdata/directory)
* Hands-on: Data downloads (PDFs, CSV, etc.) and basic web scraping
* **Assignment Due: Analysis/critique of data-driven story**

**Week4/Sep. 15 – Thinking like a social scientist**

* Read: Wimmer/Dominick chapters 1-4 (Canvas); “Journalism and Social Science” (Canvas); “Digital Watchdog’s First Byte” (Canvas); and the articles posted on Delicious for Week 4: [https://delicious.com/glanosga/j502,week4](https://delicious.com/glanosga/j502%2Cweek4)
* Hands-on: Advanced Excel operations

**Week 5/Sep. 22 – Math and stats (I)**

* Read: Wimmer/Dominick chapters 10-12 (Canvas); Best chapters 2-3; and the articles posted on Delicious for Week 5: [https://delicious.com/glanosga/j502,week5](https://delicious.com/glanosga/j502%2Cweek5)
* Hands-on: Math competency quiz

**Week 6/Sep. 29 – Math and stats (II)**

* Read: Meyer chapters 3-4; Best chapters 4-6 and Afterword
* Watch: “Drawing Conclusions from Data” [https://delicious.com/glanosga/j502,week6](https://delicious.com/glanosga/j502%2Cweek6)
* **Exam review**

**Week 7/Oct. 6 \*\*\*\*\*EXAM ONE\*\*\*\*\***

**Week 8/Oct. 13 – TRIPLE D (Dealing with Dirty Data)**

* Read: Houston chapter 9; Herzog “Evaluating and cleaning data” (Canvas); Ver 1.0 selections (Canvas); and skim the articles posted on Delicious for Week 8: [https://delicious.com/glanosga/j502,week8](https://delicious.com/glanosga/j502%2Cweek8)
* Hands-on: Google Refine

**Week 9/Oct. 20 – Databases (and a hint of coding)**

* Read: Meyer chapter 10; Houston chapters 5-6, 8; SQL handout (Canvas); and skim the articles posted on Delicious for Week 9: [https://delicious.com/glanosga/j502,week9](https://delicious.com/glanosga/j502%2Cweek9)
* Hands-On: Access, SQLite

**Week 10/Oct. 27 – Some words about polls and surveys**

* Read: Meyer chapters 6 and 11; Newsroom Guide to Polls (Canvas); “Strawpoll Journalism” (Canvas); and skim the articles posted on Delicious for Week 10: [https://delicious.com/glanosga/j502,week10](https://delicious.com/glanosga/j502%2Cweek10)
* Hands-on: Introduction to SPSS
* **Assignment due: News story**

**Week 11/Nov. 3 – Statistical analysis with SPSS**

* Read: Meyer chapters 7-8; Houston/SPSS (Canvas), IRE handouts (Canvas)
* **Exam Review**

**Week 12/Nov. 10 \*\*\*\*\*EXAM TWO\*\*\*\*\***

**Week 13/Nov. 17 – Analyzing data visually (I)**

* Read: Houston Appendices A and B; Herzog “Visualizing Data” (Canvas); Handouts TBD (Canvas)**;** and skim the articles posted on Delicious for Week 13: [https://delicious.com/glanosga/j502,week13](https://delicious.com/glanosga/j502%2Cweek13)
* Hands-on: Introduction to GIS
* **Assignment due: Data acquisition**

**Nov. 24 – No class; Thanksgiving break.**

**Week 14/Dec. 1 – Analyzing data visually (II)**

* Read: TBA
* Hands-on: Introduction to social network analysis

**Week 15/Dec. 8 – Flex day plus wrap-up**

**FINAL DATA ANALYSIS PROJECT DUE DECEMBER 15**