

RECEIVED

APR 10 2015  
OFFICE OF THE  
SENATE COUNCIL  
RW**1. General Information**

1a. Submitted by the College of: BUSINESS AND ECONOMICS

Date Submitted: 11/10/2015

1b. Department/Division: B&amp;E Economics

1c. Contact Person

Name: William Hoyt

Email: whoyt@uky.edu

Phone: 257-2518

Responsible Faculty ID (if different from Contact)

Name:

Email:

Phone:

1d. Requested Effective Date: Specific Term/Year<sup>1</sup> Fall 2016

1e. Should this course be a UK Core Course? Yes

**2. Designation and Description of Proposed Course**

2a. Will this course also be offered through Distance Learning?: No

2b. Prefix and Number: ECO 624

2c. Full Title: Empirical Data Management

2d. Transcript Title: Empirical Data Management

2e. Cross-listing:

2f. Meeting Patterns

LECTURE: 33

LABORATORY: 12

2g. Grading System: Graduate School Grade Scale

2h. Number of credit hours: 3

2i. Is this course repeatable for additional credit? No

If Yes: Maximum number of credit hours:

If Yes: Will this course allow multiple registrations during the same semester?

2j. Course Description for Bulletin: This course will cover the basic ways that economic data can be obtained. The course will cover obtaining data from data from the web, from APIs, and from colleagues in various formats including raw text files, binary files and databases. It will also cover the basics of data cleaning and how to make data operational. Organized data dramatically speeds downstream data analysis tasks. The course will also cover the components of a complete data set including raw data, processing instructions, codebooks, and processed data. The course will cover the basics needed for collecting, cleaning and sharing data. Particular points of emphasis will include reading and writing datasets, data cleaning, creating, changing and labeling variables and values, automating your work, combining and reshaping files, processing observations across subgroups, .do file programming, and programming using .ado files.

2k. Prerequisites, if any: ECO 491 or its equivalent or consent of the instructor.

2l. Supplementary Teaching Component:

3. Will this course taught off campus? No

If YES, enter the off campus address:

4. Frequency of Course Offering: Fall,

Will the course be offered every year?: Yes

If No, explain:

5. Are facilities and personnel necessary for the proposed new course available?: Yes

If No, explain:

6. What enrollment (per section per semester) may reasonably be expected?: 30

7. Anticipated Student Demand

Will this course serve students primarily within the degree program?: Yes

Will it be of interest to a significant number of students outside the degree pgm?: No

If Yes, explain:

8. Check the category most applicable to this course: Traditional – Offered in Corresponding Departments at Universities Elsewhere,

If No, explain:

9. Course Relationship to Program(s).

a. Is this course part of a proposed new program?: No

If YES, name the proposed new program:

b. Will this course be a new requirement for ANY program?: Yes

If YES, list affected programs: MS, Economics

10. Information to be Placed on Syllabus.

a. Is the course 400G or 500?: No

b. The syllabus, including course description, student learning outcomes, and grading policies (and 400G-/500-level grading differentiation if applicable, from **10.a** above) are attached: Yes

## Distance Learning Form

Instructor Name:

Instructor Email:

Internet/Web-based: No

Interactive Video: No

Hybrid: No

1. How does this course provide for timely and appropriate interaction between students and faculty and among students? Does the course syllabus conform to University Senate Syllabus Guidelines, specifically the Distance Learning Considerations?

2. How do you ensure that the experience for a DL student is comparable to that of a classroom-based student's experience? Aspects to explore: textbooks, course goals, assessment of student learning outcomes, etc.

3. How is the integrity of student work ensured? Please speak to aspects such as password-protected course portals, proctors for exams at interactive video sites; academic offense policy; etc.

4. Will offering this course via DL result in at least 25% or at least 50% (based on total credit hours required for completion) of a degree program being offered via any form of DL, as defined above?

If yes, which percentage, and which program(s)?

5. How are students taking the course via DL assured of equivalent access to student services, similar to that of a student taking the class in a traditional classroom setting?

6. How do course requirements ensure that students make appropriate use of learning resources?

7. Please explain specifically how access is provided to laboratories, facilities, and equipment appropriate to the course or program.

8. How are students informed of procedures for resolving technical complaints? Does the syllabus list the entities available to offer technical help with the delivery and/or receipt of the course, such as the Information Technology Customer Service Center (<http://www.uky.edu/UKIT/>)?

9. Will the course be delivered via services available through the Distance Learning Program (DLP) and the Academic Technology Group (ATL)? NO

If no, explain how student enrolled in DL courses are able to use the technology employed, as well as how students will be provided with assistance in using said technology.

10. Does the syllabus contain all the required components? NO

11. I, the instructor of record, have read and understood all of the university-level statements regarding DL.

Instructor Name:

SIGNATURE|MKT210|Steven J Skinner|ECO 624 NEW College Review|20150327  
SIGNATURE|ZNNIKO0|Roshan Nikou|ECO 624 NEW Graduate Council Review|20150410  
SIGNATURE|JEL224|Janie S Ellis|ECO 624 NEW Senate Council Review|20150415  
SIGNATURE|WHOYT|William H Hoyt|ECO 624 NEW Approval Returned to Dept|20150430  
SIGNATURE|JEL224|Janie S Ellis|ECO 624 NEW Senate Council Review|20150504  
SIGNATURE|WHOYT|William H Hoyt|ECO 624 NEW Approval Returned to Dept|20150519  
SIGNATURE|JEL224|Janie S Ellis|ECO 624 NEW Senate Council Review|20151111  
SIGNATURE|WHOYT|William H Hoyt|ECO 624 NEW Approval Returned to Dept|20151124  
SIGNATURE|MKT210|Steven J Skinner|ECO 624 NEW College Review|20150327

## New Course Form

<https://myuk.uky.edu/sap/bc/soap/rfc?services=>

[Open in full window to print or save](#)

Generate R

## Attachments:

[Browse...](#)

Upload File

	ID	Attachment
Delete	4936	ECO 624 Syllabus (Data Management).docx

First 1 Last

(\*denotes required fields)

## 1. General Information

- a. \* Submitted by the College of:  Submission Date:
- b. \* Department/Division:
- c.
- \* Contact Person Name:  Email:  Phone:
- \* Responsible Faculty ID (if different from Contact):  Email:  Phone:
- d. \* Requested Effective Date:  Semester following approval OR  Specific Term/Year <sup>1</sup>
- e. Should this course be a UK Core Course?  Yes  No
- If YES, check the areas that apply:
- Inquiry - Arts & Creativity  Composition & Communications - II
- Inquiry - Humanities  Quantitative Foundations
- Inquiry - Nat/Math/Phys Sci  Statistical Inferential Reasoning
- Inquiry - Social Sciences  U.S. Citizenship, Community, Diversity
- Composition & Communications - I  Global Dynamics

## 2. Designation and Description of Proposed Course.

- a. \* Will this course also be offered through Distance Learning?  Yes <sup>1</sup>  No
- b. \* Prefix and Number:
- c. \* Full Title:
- d. Transcript Title (if full title is more than 40 characters):
- e. To be Cross-Listed <sup>2</sup> with (Prefix and Number):
- f. \* Courses must be described by at least one of the meeting patterns below. Include number of actual contact hours<sup>3</sup> for each meeting pattern type.
- |   |   |                                 |                                 |
|---|---|---------------------------------|---------------------------------|
| <input type="text" value="33"/> Lecture | <input type="text" value="12"/> Laboratory <sup>1</sup> | <input type="text"/> Recitation | <input type="text"/> Discussion |
| <input type="text"/> Indep. Study       | <input type="text"/> Clinical                           | <input type="text"/> Colloquium | <input type="text"/> Practicum  |
| <input type="text"/> Research           | <input type="text"/> Residency                          | <input type="text"/> Seminar    | <input type="text"/> Studio     |
| <input type="text"/> Other              | If Other, Please explain: <input type="text"/>          |                                 |                                 |
- g. \* Identify a grading system:
- Letter (A, B, C, etc.)
- Pass/Fail
- Medicine Numeric Grade (Non-medical students will receive a letter grade)
- Graduate School Grade Scale
- h. \* Number of credits:
- i. \* Is this course repeatable for additional credit?  Yes  No
- If YES: Maximum number of credit hours:
- If YES: Will this course allow multiple registrations during the same semester?  Yes  No

## J. \* Course Description for Bulletin:

This course will cover the basic ways that economic data can be obtained. The course will cover obtaining data from data from the web, from APIs, and from colleagues in various formats including raw text files, binary files and databases. It will also cover the basics of data cleaning and how to make data operational. Organized data dramatically speeds downstream data analysis tasks. The course will also cover the components of a complete data set including raw data, processing instructions, codebooks, and processed data. The course will cover the basics needed for collecting, cleaning and sharing data.

Particular points of emphasis will include reading and writing datasets, data cleaning, creating, changing and labeling variables and values, automating your work, combining and reshaping files, processing observations across subgroups, .do file programming, and programming using .ado files.

## K. Prerequisites, if any:

ECO 491 or its equivalent or consent of the instructor.

I. Supplementary teaching component, if any:  Community-Based Experience  Service Learning  Both3. \* Will this course be taught off campus?  Yes  No

If YES, enter the off campus address:

## 4. Frequency of Course Offering.

a. \* Course will be offered (check all that apply):  Fall  Spring  Summer  Winter

b. \* Will the course be offered every year?  Yes  No

If No, explain:

5. \* Are facilities and personnel necessary for the proposed new course available?  Yes  No

If No, explain:

## 6. \* What enrollment (per section per semester) may reasonably be expected? 30

## 7. Anticipated Student Demand.

a. \* Will this course serve students primarily within the degree program?  Yes  No

b. \* Will it be of interest to a significant number of students outside the degree pgm?  Yes  No

If YES, explain:

## 8. \* Check the category most applicable to this course:

Traditional – Offered in Corresponding Departments at Universities Elsewhere

Relatively New – Now Being Widely Established

Not Yet Found in Many (or Any) Other Universities

## 9. Course Relationship to Program(s).

a. \* Is this course part of a proposed new program?  Yes  No

If YES, name the proposed new program:

b. \* Will this course be a new requirement<sup>5</sup> for ANY program?  Yes  No

If YES<sup>5</sup>, list affected programs:

MS, Economics

## 10. Information to be Placed on Syllabus.

a. \* Is the course 400G or 500?  Yes  No

If YES, the *differentiation for undergraduate and graduate students must be included* in the information required in 10.b. You must include: (i) identify additional assignments by the graduate students; and/or (ii) establishment of different grading criteria in the course for graduate students. (See SR

b.  \* The syllabus, including course description, student learning outcomes, and grading policies (and 400G-/500-level grading differentiation if apply 10.a above) are attached.

<sup>5</sup> Courses are typically made effective for the semester following approval. No course will be made effective until all approvals are received.  
<sup>6</sup> The chair of the cross-listing department must sign off on the Signature Routing Log.

In general, undergraduate courses are developed on the principle that one semester hour of credit represents one hour of classroom meeting per week for a semester, exclusive of any laboratory meeting. Laboratory meeting, generally, is two hours per week for a semester for one credit hour. (from SR 52.1)

You must also submit the Distance Learning Form in order for the proposed course to be considered for DL delivery.

In order to change a program, a program change form must also be submitted.

Rev 8/09

**Economics 624: Empirical Data Management**  
Fall 2016

**Time**

**Professor:** Aaron Yelowitz  
**Office:** 303B Mathews  
**Phone:** 257-7634  
**Office Hours:** 2:30 – 3:30 MW  
**Email:** [aaron.yelowitz@uky.edu](mailto:aaron.yelowitz@uky.edu)

**Course Description:** This course will cover the basic ways that economic data can be obtained. The course will cover obtaining data from data from the web, from APIs, and from colleagues in various formats including raw text files, binary files and databases. It will also cover the basics of data cleaning and how to make data tidy. Tidy data dramatically speeds downstream data analysis tasks. The course will also cover the components of a complete data set including raw data, processing instructions, codebooks, and processed data. The course will cover the basics needed for collecting, cleaning and sharing data.

Particular points of emphasis will include reading and writing datasets, data cleaning, creating, changing and labeling variables and values, automating your work, combining and reshaping files, processing observations across subgroups, .do file programming, and programming using .ado files.

Prerequisites: ECO 491 or its equivalent or consent of the instructor.

**Course Goals:**

1. Provide students an understanding of data structures
2. Enhance skills in preparing data for econometric and other analysis
3. Develop insights into using data to answer applied economic questions

**Student Learning Outcomes:** By the end of the course student should be able to:

1. Use proper terms to describe data sets
2. Prepare data from multiple sources for analysis
3. Identify and execute data analysis approaches suitable for research questions
4. Interpret the results properly, and identify additional potential issues.

**Readings**

Required Readings

**Required Textbook:** Michael N. Mitchell, “Data Management Using Stata: A Practical Handbook” 2010, Stata Press, ISBN-13: 978-1-59718-076-4. Available at: <http://www.stata.com/bookstore/data-management-using-stata/>

Required Articles:

Edelman, Benjamin. 2012. "Using Internet Data for Economic Research." Journal of Economic Perspectives, 26(2): 189-206.



Varian, Hal R. 2014. "Big Data: New Tricks for Econometrics." *Journal of Economic Perspectives*, 28(2): 3-28.

Heffetz, Ori, and Katrina Ligett. 2014. "Privacy and Data-Based Research." *Journal of Economic Perspectives*, 28(2): 75-98.

Schwabish, Jonathan A. 2014. "An Economist's Guide to Visualizing Data." *Journal of Economic Perspectives*, 28(1): 209-34.

Gentzkow, Matthew and Jesse M. Shapiro. 2014. "Code and Data for the Social Sciences: A Practitioner's Guide." University of Chicago mimeo, <http://faculty.chicagobooth.edu/matthew.gentzkow/research/CodeAndData.pdf>.

Shapiro, Jesse M., "How to Give an Applied Micro Talk: Unauthoritative Notes," Accessed from [http://faculty.chicagobooth.edu/jesse.shapiro/research/applied\\_micro\\_slides.pdf](http://faculty.chicagobooth.edu/jesse.shapiro/research/applied_micro_slides.pdf).

#### Optional Articles:

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 28(2): 29-50.

Nickerson, David W., and Todd Rogers. 2014. "Political Campaigns and Big Data." *Journal of Economic Perspectives*, 28(2): 51-74.

Einav, Liran, and Jonathan Levin, "The Data Revolution and Economic Analysis" Accessed from <http://web.stanford.edu/~jdlevin/Papers/BigData.pdf>.

#### Evaluation:

Regular Homework Problems	Regular Through the Semester	33%
Exam 1	Midterm	33%
Exam 2	Final	33%

#### Assignments:

All assignments (including paper) are due **at the beginning of class** on the date specified when the assignment is given unless the student has an excused absence. I am a stickler about the deadline: you need to be on time for class and turn it in at the beginning. Any homework turned in after I collect it will be considered late. Assignments that are less than 3 business days late will be penalized by 10% for each day late. Homework over 3 business days late will not be accepted for credit. All homework and exams will be on a percentage basis, with 90% and above representing an A, 80% and above a B, 70% and above a C, below 70% will be given an E.

#### Grading:

Grades are assigned depending on your point accumulation throughout the term. In determining your grade, University rules regarding the marking system is followed. It is (with slight paraphrasing).

A: Represents exceptionally high achievement.

B: Represents high achievement.

- C: Represents average achievement.
- D: not used for graduate students
- E: Represents unsatisfactory performance and indicates failure of the course.

To implement these standards, the grade scale will be no more stringent than the following:

- 90 – 100% = A
- 80 – 89% = B
- 70 – 79% = C
- ≤ 69% = E

### **Excused Absences**

Students need to notify the professor of absences prior to class when possible. S.R. 5.2.4.2 defines the following as acceptable reasons for excused absences: (a) serious illness, (b) illness or death of family member, (c) University-related trips, (d) major religious holidays, and (e) other circumstances found to fit “reasonable cause for nonattendance” by the professor.

Students anticipating an absence for a major religious holiday are responsible for notifying the instructor in writing of anticipated absences due to their observance of such holidays no later than the last day in the semester to add a class. Information regarding dates of major religious holidays may be obtained through the religious liaison, Mr. Jake Karnes (859-257-2754).

Students are expected to withdraw from the class if more than 20% of the classes scheduled for the semester are missed (excused or unexcused) per university policy.

### **Verification of Absences**

Students may be asked to verify their absences in order for them to be considered excused. Senate Rule 5.2.4.2 states that faculty have the right to request “appropriate verification” when students claim an excused absence because of illness or death in the family. Appropriate notification of absences due to university-related trips is required prior to the absence.

### **Academic Integrity**

Per university policy, students shall not plagiarize, cheat, or falsify or misuse academic records. Students are expected to adhere to University policy on cheating and plagiarism in all courses. The minimum penalty for a first offense is a zero on the assignment on which the offense occurred. If the offense is considered severe or the student has other academic offenses on their record, more serious penalties, up to suspension from the university may be imposed.

Plagiarism and cheating are serious breaches of academic conduct. Each student is advised to become familiar with the various forms of academic dishonesty as explained in the Code of Student Rights and Responsibilities. Complete information can be found at the following website:

<http://www.uky.edu/Ombud>. A plea of ignorance is not acceptable as a defense against the charge of academic dishonesty. It is important that you review this information as all ideas borrowed from others need to be properly credited.

Part II of *Student Rights and Responsibilities* (available online <http://www.uky.edu/StudentAffairs/Code/part2.html>) states that all academic work, written or otherwise, submitted by students to their instructors or other academic supervisors, is expected to be

the result of their own thought, research, or self-expression. In cases where students feel unsure about the question of plagiarism involving their own work, they are obliged to consult their instructors on the matter before submission.

When students submit work purporting to be their own, but which in any way borrows ideas, organization, wording or anything else from another source without appropriate acknowledgement of the fact, the students are guilty of plagiarism. Plagiarism includes reproducing someone else's work, whether it be a published article, chapter of a book, a paper from a friend or some file, or something similar to this. Plagiarism also includes the practice of employing or allowing another person to alter or revise the work which a student submits as his/her own, whoever that other person may be.

Students may discuss assignments among themselves or with an instructor or tutor, but when the actual work is done, it must be done by the student, and the student alone. When a student's assignment involves research in outside sources of information, the student must carefully acknowledge exactly what, where and how he/she employed them. If the words of someone else are used, the student must put quotation marks around the passage in question and add an appropriate indication of its origin. Making simple changes while leaving the organization, content and phraseology intact is plagiaristic. However, nothing in these Rules shall apply to those ideas which are so generally and freely circulated as to be a part of the public domain (Section 6.3.1).

**Please note:** Any assignment you turn in may be submitted to an electronic database to check for plagiarism.

### **Accommodations due to disability**

If you have a documented disability that requires academic accommodations, please see me as soon as possible during scheduled office hours. In order to receive accommodations in this course, you must provide me with a Letter of Accommodation from the Disability Resource Center (Room 2, Alumni Gym, 257-2754, email address: jkarnes@email.uky.edu) for coordination of campus disability services available to students with disabilities.

### **Outline**

1. Getting Started: Software/fixed costs of using data (for economists) (1 week)
  - a. Software: Stata, StatTransfer, WinZIP
  - b. IRB Certification
    - i. Heffetz, Ori, and Katrina Ligett. 2014. "Privacy and Data-Based Research." *Journal of Economic Perspectives*, 28(2): 75-98.
  - c. Hard Drive Space
2. Data sources (2 weeks)
  - a. Free, publicly available data: Census, ICPSR, NBER
  - b. Restricted-access, free, public data: Versions of Census data and others
  - c. For-purchase data: I.e. Neilsen data
  - d. Web-based: Web scraping
    - i. Edelman, Benjamin. 2012. "Using Internet Data for Economic Research." *Journal of Economic Perspectives*, 26(2): 189-206.
  - e. Primary Data Collection
3. Data collection (1/2 week)
  - a. Raw files (.csv, .xlsx)

- b. Databases (mySQL)
  - c. APIs
- 4. Data formats (1/2 week)
  - a. Flat files (.csv, .txt)
  - b. XML
  - c. JSON
- 5. Making Data Tidy (3 week)
  - a. Straightforward Stata Commands
    - i. Subsetting and Sorting
    - ii. Summarizing Data
    - iii. Creating New variables
    - iv. Reshaping Data
    - v. Merging Data
    - vi. Editing Text Variables
    - vii. Regular Expressions
    - viii. Working with Dates
- 6. Distributing Data (1 week)
  - a. Gentzkow, Matthew and Jesse M. Shapiro. 2014. "Code and Data for the Social Sciences: A Practitioner's Guide." University of Chicago mimeo, <http://faculty.chicagobooth.edu/matthew.gentzkow/research/CodeAndData.pdf>.
- 7. Scripting For Data Cleaning (1 week)
- 8. Presenting Data (2 weeks)
  - a. Schwabish, Jonathan A. 2014. "An Economist's Guide to Visualizing Data." *Journal of Economic Perspectives*, 28(1): 209-34.
  - b. Shapiro, Jesse M., "How to Give an Applied Micro Talk: Unauthoritative Notes," Accessed from [http://faculty.chicagobooth.edu/jesse.shapiro/research/applied\\_micro\\_slides.pdf](http://faculty.chicagobooth.edu/jesse.shapiro/research/applied_micro_slides.pdf).
- 9. Big Data (4 weeks)
  - a. Varian, Hal R. 2014. "Big Data: New Tricks for Econometrics." *Journal of Economic Perspectives*, 28(2): 3-28.
  - b. Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 28(2): 29-50.
  - c. Nickerson, David W., and Todd Rogers. 2014. "Political Campaigns and Big Data." *Journal of Economic Perspectives*, 28(2): 51-74.
  - d. Einav, Liran, and Jonathan Levin, "The Data Revolution and Economic Analysis" Accessed from <http://web.stanford.edu/~jdlevin/Papers/BigData.pdf>.