TEXAS LBJ School
The University of Texas at Austin
Lyndon B. Johnson School of Public Affairs

# Lyndon B. Johnson School of Public Affairs

Dean's Certificate in Data Science and Policy Analysis (DSPA)
DSPA Data Sheet, 2021-2022

The recent explosion of "Big Data" has led to a rapidly expanding market for graduate degree holders with data analysis and modeling skills, not only in the private sector, but also in government organizations and non-profits. Many LBJ students are interested in options that allow them to acquire these new analytical skills at the LBJ School, in a format that will provide them with a clear pathway into these new jobs. The Dean's Certificate in Data Science and Policy Analysis is an exciting opportunity that assists our graduates in finding great employment opportunities after graduation.

The DSPA Certificate has a unique, innovative structure designed to certify acquisition of a particular set of data science skills and demonstrate to prospective employers that the certificate recipient can apply those skills in defining, organizing, and executing an original data science project. It does not require completion of a particular fixed selection of courses. Selection of the three approved courses in which those skills can be acquired is flexible by design, and can utilize both qualifying sections of advanced methods classes or qualifying sections of policy-oriented classes with an empirical analysis project component. Recipients of the LBJ DSPA certificate must complete an original, individual data science project demonstrating proficiency in **all four** (A-D) required skills.

This year we are pleased to report some additional news about the DSPA certificate:

1. In its inaugural year (June 2020), five students graduated with the certificate. This year (June 2021), six students graduated with the certificate. Especially notable were a number of last-minute inquiries by graduating master's degree students during the Spring 2020 semester about eligibility of classes and projects in satisfying the DSPA certificate requirements, some of which, unfortunately, did not meet the certificate's established requirements.

2. An LBJ School team (6 LBJ students and 1 LBJ faculty) placed second in an international data analysis competition organized on the XPrize online platform: the Microsoft-ODI Open Education Data Challenge. Team LBJ opted to donate the $35,000 prize money to fund an endowed award (initially $1,000) for the best student data science and policy analysis project submitted by a graduating LBJ School student. This endowed student award fund is currently awaiting approval by the UT Board of Regents, with funds likely available for an inaugural Team LBJ Award in Data Science and Policy Analysis in the academic year following its approval.

The certificate program-- open to all LBJ School master's degree students (MPAff, MGPS, and, with the approval of their dissertation supervisor, PhD students)-- is intended to encourage LBJ students to develop new analytical data science skills as part of their training at the LBJ School. The certificate is an internal LBJ School certificate, recorded in LBJ School records, and recognized on the commencement program at graduation. This spring, in response to student inquiries, interim Dean Springer also approved its award to graduate students in other UT Austin degree programs.

**Certificate Requirements**

A total of three approved course must be completed which, collectively, demonstrate competence in the following five skills:[1]

A. Use of a modern data science software platform for empirical data analysis;
B. Demonstrated knowledge of current methods of acquiring, merging, and cleaning multiple sources of unprocessed digital data; knowledge and use of principled methods to deal with missing, miscoded, and outlier data;
C. Ability to create compelling and effective summaries and visualizations of large-scale data sets;
D. Demonstrated proficiency in application of statistical models or machine learning methods to the analysis of empirical data sets using software tools;
E. Successful completion of an original, individual policy analysis project using statistical and data science tools applied to a real-world empirical dataset, demonstrating proficiency in all the above skills.

For most students, skill requirement (E) will be satisfied by completion of a data science empirical research project as a Professional Report (PR) or master's thesis (for dual degree students). In this case an individual PR or thesis course can also count as one of the three qualifying courses required for the DSPA Certificate. The project is submitted with the application for the certificate, and the faculty member supervising the course must certify that the submitted data science project demonstrates proficiency in **all four** (A-D) required skills

In addition, LBJ policy classes, including PRPs and individual conference courses, may permit exceptional students to undertake an original empirical data analytical project on a course-related policy issue. With the faculty instructor's approval, students with a sufficient skill set can submit an original paper, or their individual contribution to a group project (in the form of a paper, chapter, report section, or other documentation, for which the student takes sole responsibility), to satisfy the DSPA certificate's empirical analysis project requirement.

Updates to the list of LBJ School classes providing organized instruction in DSPA skill areas must be submitted to the certificate coordinator prior to the first class day of the semester in which it is being taught, so your LBJ instructor is unable to have your LBJ School class added to the DSPA skill area matrix retroactively.

Graduate classes taught in other UT Austin units can be approved to count toward certificate skill areas after review and approval by the certificate coordinator/coordinating committee. Students must email a syllabus for the class, attached to this request, to the certificate coordinator/coordinating committee as soon as possible, and in all cases before the last day of classes for the semester in which the non-LBJ class is taken.

To satisfy requirement (E), data science projects must demonstrate proficiency in each of the four skill areas, (A) through (D). Introductory LBJ School classes (i.e., Analytical Methods, Intro to Empirical Methods, Applied Micro, Public Financial Management, Public Management, and Policy Development) can count toward satisfying a skill requirement if the faculty instructor has incorporated organized instruction in that skill into the class but will not count toward the 3-course certificate requirement.

[1] Students entering LBJ School degree programs with substantial data science experience can be exempted from the coursework requirement for a specific skill on a case-by-case basis, subject to approval by their graduate advisor and certificate coordinator.

**Data Science and Policy Analysis Certificate: Preliminary Qualifying 2021-22 Course List,** August 2021

| | Data Science and Policy Analysis Certificate: Preliminary Qualifying 2021-22 Course List | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | **Data Science and Policy Analysis Skill/Competency Areas:** | | | | |
| **Faculty Name** | **Academic Year (2021-22)** | **Course Number** *(e.g., PA 388R)* | **Section Name** *(e.g., Policy Analysis, Using R)* | ***A. Software Platform*** | ***B. Dirty to Clean Dataset*** | ***C. Summary & Visualization*** | ***D. Statistical/ ML Analysis*** | ***E. Individual Research Project*** | |
| | **Fall 2021** | | | | | | | | |
| Gordon Abner | Fall 2021 | PA 384C | Public Management | x | | x | x | | |
| Patrick Bixler | Fall 2021 | PA397C | Program Evaluation for Social Impact | X | X | X | X | | |
| Ji Ma | Fall 2021 | PA397G | Analytical Methods for Global Policy Studies | x | | x | x | * | |
| T Olmstead | Fall 2021 | PA 397 | IEM | X | | | X | | |
| R Wasem | Fall 2021 | PA383C | PD: Immigration Incorporation | | | X | | | |
| | | | | | | | | | |
| | **Spring 2022** | | | | | | | | |
| R Fabregas | Spring 2022 | PA 397 C | Evaluation Methods for Global Development | X | | | X | | |
| K Flamm | Spring 2022 | PA393H | Trade, Technology & Industrial Policy, Using Python | X | | X | | | |
| K Flamm | Spring 2022 | PA393K | Applied Micro for Policy Analysis, Using Python | X | | X | | | |
| K Flamm | Spring 2022 | PA397C | AEM: Intro to Data Visualization & Econometrics Using Python | X | X | X | X | * | |
| E Lentz | Spring 2022 | 397G | Analytical Methods | X | | X | | | |
| T Olmstead | Spring 2022 | PA393L | APE & PRP: Economic Eval of Health Care Programs | | | | X | | |
| R Wasem | Spring 2022 | PA383C | PD: Refugees, Asylum, & Human Security | | | X | | | |
| | | | | | | | | | |
| Any faculty member | | | | | | | | | |
| | 2021-2022 | PA 398R | Professional Report (w/qualifying data science project) | | | | | X | |
| | 2021-2022 | PA 189C | Conference Course | * | * | * | * | * | |
| | | | | | | | | | |
| *X :* | ***Organized instruction in functional skill area, applies to all students passing course.*** | | | | | | | | |
| **** :*** | ***Qualifying data science option possible, by arrangement with instructor and certificate coordinator/faculty steering committee.*** | | | | | | | | |
| Example: student can complete class requrements by choosing to submit either qualifying individual data science project, or individual empirical contribution to group project. | | | | | | | | | |
| Definition of qualifying data science project: student displays ability to define and undertake original, individual data science project including all four skill elements (A)-(D). | | | | | | | | | |
| Intro classes (PM,PFM,AMPA,IEM,AM,PD) count toward skill areas but not toward 3-course requirement. | | | | | | | | | |

Appendix 1: Academic Year 2020-21 Qualifying Course List

## Data Science and Policy Analysis Certificate: Qualifying 2020-21 Course List

| | | | | Data Science and Policy Analysis Skill/Competency Areas: | | | | |
|---|---|---|---|---|---|---|---|---|
| **Faculty Name** | **Academic Year (2020-21)** | **Course Number** *(e.g., PA 388R)* | **Section Name** *(e.g., Policy Analysis, Using R)* | ***A. Software Platform*** | ***B. Dirty to Clean Dataset*** | ***C. Summary & Visualization*** | ***D. Statistical /ML Analysis*** | ***E. Individual Research Project*** |
| Abner | Fall | PA 388L | Equity, Organizations, Data Visualization | X | X | X | | |
| Bixler | Fall | PA 397C | AEM:Program Evaluation for Nonprofit and Social Impac | X | X | X | X | |
| Pavon | Fall | PA 388K | Intro to GIS | X | X | X | X | * |
| Sankaran | Fall | PA397G | Analytical Methods for Global Policy Studies | X | X | X | X | |
| Stolp | Fall | PA 388K | Evaluation of Social Polices/Programs | X | X | X | X | * |
| Weaver; Ma | Fall/Spring | PRP | Counting What Counts: Global Development and the Politics of Data | X | X | X | | |
| Abner | Spring | PA 397C | Survey Design | X | X | X | | |
| Flamm | Spring | PA 397C | AEM: Data Vis, Statistics & Econometrics, Using Python | X | X | X | X | * |
| Flamm | Spring | PA 393K | AMP: Applied Micro, Using Python | X | | X | | |
| Flamm | Spring | PA 393H | Intl Econ:Tech,Trade and Industrial Policy, Using Python | X | | X | X | * |
| Lentz | Spring | PA 397G | Analytical Methods | X | | X | | |
| Ma | Spring | PA 397C | AEM:Data Management and Research Life Cycle | X | X | X | | * |
| Pavon | Spring | PA 388K | Intro to GIS | X | X | X | X | * |
| Rai | Spring | PA 397C | AEM: Statistical Analysis and Learning | X | X | X | X | * |
| Stolp | Spring | PA 390C | Advanced Research Methods (PhD) | X | X | X | X | * |
| Stolp | Spring | PA 397C | AEM: Econometrics for Policy Analysis | X | X | X | X | * |
| von Hippel | Spring | PA 397C | AEM: Evaluation Design | X | X | X | X | * |
| Fabregas | Spring | PA 397 C | Evaluation Methods for Global Development | X | | X | X | * |
| | | | | | | | | |
| Any faculty | | PA 398R | Professional Report (w/qualifying data science project) | | | | | X |
| | | | | | | | | |
| *X :* | ***Organized instruction in functional skill area, applies to all students passing course.*** | | | | | | | |
| **** :*** | ***Qualifying data science option possible, by arrangement with instructor and certificate coordinator/faculty steering committee.*** | | | | | | | |

Example: student can complete class requirements by choosing to submit either qualifying individual data science project, or individual empirical contribution to group project.

Definition of qualifying data science project: student displays ability to define and undertake original, individual project including all four skill elements (A)-(D).

Appendix 2: Academic Year 2019-20 Qualifying Course List

| | | | | DSPA Skill Competency Area | | | | |
|---|---|---|---|---|---|---|---|---|
| **Course Name** | **Course Number** | **Instructor** | **Semester** | **Software Platform (A)** | **Dirty to Clean Dataset (B)** | **Summary & Visualization (C)** | **Statistical/ML Analysis (D)** | **Individual Research Project (E)** |
| Survey Design | PA 397C | Abner | Fall | X | X | X | | * |
| Organization & Equity | | Abner | Spring | X | X | X | | * |
| Mixed Methods for Policy Analysis | PA 397C | Aiken | | | | X | | |
| Evaluation Methods for Global Development | PA 397C | Fabregas | | X | * | | X | * |
| Applied Micro for Policy | PA 393K | Flamm | Spring | * | | * | | |
| International Economics Using Python | PA 393H | Flamm | Spring | X | | X | X | * |
| Data Visualization, Statistics & Econometrics Using Python | PA 397C | Flamm | Spring | X | X | X | X | * |
| Linked Open Data and Computational Social Science Methods | PA 397C | Ma | | X | X | | X | * |
| Data Management and Research Life Cycle | PA 397C | Ma | | X | X | X | | * |
| Introduction to Empirical methods and Policy Analysis | PA 397 | Olmstead, T | | X | | | X | |
| Economic Evaluation of Health | PA 393L | Olmstead, T | | X | | | X | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Introduction to Geographic Information Systems | PA 388K | Pavon | | X | X | X | X | * |
| Statistical Analysis and Learning | PA 397C | Rai | | X | X | X | X | * |
| Analytical Methods for Global Policy Studies | PA 397G | Sankaran | | X | X | X | X | |
| Policymaking in a Global Age | PA 383G | Sankaran | | | | | | * |
| Statistical Reasoning | PA 388K | Stolp | | X | X | X | X | * |
| Introduction to Empirical Methods for Policy Analysis | PA 397 | Stolp | | X | X | X | X | |
| Research Design | PA 390E | Stolp | | X | | | X | * |
| Evaluation of Social Programs | PA 388K | Stolp | | X | X | X | X | |
| Urban Economics and Policy | PA 393L | Waxman | | * | * | * | * | * |
| Professional Report with Qualifying Data Science Project | PA 398R | Any faculty member | Fall and Spring | | | | | X |
| * = Qualifying data science option possible, by arrangement with instructor and certificate coordinator/faculty steering committee. | | | | | | | | |
| Example: student can complete class requirements by choosing to submit either qualifying individual data science project, or individual empirical contribution to group project. | | | | | | | | |
| Definition of qualifying data science project: student demonstrates ability to define and undertake original, individual data science project including all four skill elements (A)-(D). | | | | | | | | |

Appendix 3: Academic Year 2018-2019 Qualifying Course List

| | | | | DSPA Skill Competency Area | | | | |
|---|---|---|---|---|---|---|---|---|
| **Course Name** | **Course Number** | **Instructor** | **Semester** | **Software Platform (A)** | **Dirty to Clean Dataset (B)** | **Summary & Visualization (C)** | **Statistical/ML Analysis (D)** | **Individual Research Project (E)** |
| Mixed Methods for Policy Analysis | PA 397C | Aiken | | | | X | | |
| Evaluation Methods for Global Development | PA 397C | Fabregas | | X | | | X | |
| Applied Micro for Policy Using Python | PA 393K | Flamm | Spring | X | | X | X | |
| Data Visualization, Statistics & Econometrics Using Python | PA 397C | Flamm | Spring | X | X | X | X | |
| Analytical Methods for Global Policy Studies | PA 397G | Lentz | | X | | | | |
| Data Management and Research Life Cycle | PA 397C | Ma | | X | X | X | | * |
| Introduction to Empirical methods and Policy Analysis | PA 397 | Olmstead, T | | X | | | X | |
| Introduction to Geographic Information Systems | PA 388K | Pavon | | X | X | X | X | * |
| Statistical Analysis and Learning | PA 397C | Rai | | X | X | X | X | * |
| Analytical Methods for Global Policy Studies | PA 397G | Sankaran | | X | X | X | X | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Statistical Reasoning | PA 388K | Stolp | | X | X | X | X | * |
| Introduction to Empirical Methods for Policy Analysis | PA 397 | Stolp | | X | X | X | X | |
| Econometrics for Policy Analysis | PA 397C | Stolp | | X | X | X | X | * |
| Evaluation of Social Programs | PA 388K | Stolp | | X | X | X | X | |
| Evaluation Design | PA 397C | von Hippel | | X | X | X | X | * |
| Urban Economics and Policy | PA 393L | Waxman | | * | * | * | * | |
| * = Qualifying data science option possible, by arrangement with instructor and certificate coordinator/faculty steering committee. | | | | | | | | |
| Example: student can complete class requirements by choosing to submit either qualifying individual data science project, or individual empirical contribution to group project. | | | | | | | | |
| Definition of qualifying data science project: student demonstrates ability to define and undertake original, individual data science project including all four skill elements (A)-(D). | | | | | | | | |