Should every undergraduate in NYU's College of Arts and Science be required to achieve competency in "Data and Discovery"?

Ernest Davis
Dept of Computer Science
New York University
davise@cs.nyu.edu

March 4, 2024

NOTE: The opinions expressed here are entirely my own. They do not represent the opinions of the Department of Computer Science.

1 The proposed requirement and its motivation

Among the revisions proposed for the Core Curriculum is a proposal to replace the "Quantitative Reasoning" requirement by a core competency requirement, "Data and Discovery."

The following passages in Framing Document for Proposed Revisions to the CAS College Core Curriculum presented to the faculty on March 4, 2024 are the entire description and motivation for this change given in the document.

Page 4: "Data and Discovery: During their undergraduate studies, students will acquire analytical skills and computational tools that will enable them to investigate, interpret, and evaluate the many configurations of data they will encounter in their lives. By taking an expansive view of data literacy, Data and Discovery courses provide students with the knowledge, skills, and discernment they need in order to scrutinize the roles that data play in the world today. May be satisfied with a Core course in Data and Discovery or by a pre-approved departmental course."

Page 6: "The following characteristics of the new Core were designed with student success in mind: . . . a new emphasis on urgent issues such as . . . and the rise of big data and artificial intelligence (in Complexities and Data and Discovery)."

Page 15-16: "Data and Discovery

Learning Objectives
Students who successfully complete Data and Discovery should be able to:

- access, organize, and query authentic data sets in a variety of contexts;
- demonstrate skills using computational tools for data analysis;
- create, interpret, and critique visual displays of data;
- apply statistical reasoning to construct and evaluate data-based claims;

- design a mathematical model to generate predictions;
- evaluate the consequences of different forms of data collection and utilization.

Description

The impact of data on contemporary life is pervasive, complex, and growing. Data and Discovery equips students with analytical skills and computational tools to investigate, interpret, and evaluate the many configurations of data they will encounter in their lives. Students explore multiple strategies for data visualization, and analyze how different visualizations can alter the interpretation of information. By applying statistical reasoning, they learn to construct evidence-based conclusions and discern the validity of data-driven claims in the popular press and research literature. They design mathematical models to generate predictions that illuminate the role of variable selection in determining model outcomes. Examining the processes of data collection and utilization enables students to reflect on how these procedures may have both intended and unintended consequences in the social sphere. By taking an expansive view of data literacy, Data and Discovery provides students with the knowledge, skills, and discernment they need in order to scrutinize the many roles data play in the world today."

Page 23: "**Will STEM majors be required to meet the Data and Discovery Core Competency?**
The faculty committee that conducted the initial review of the Core proposed the creation of a new Data and Discovery component to replace the existing Quantitative Reasoning requirement. This proposal was later endorsed by a faculty vote within the Faculty Steering Committee for the Foundations of Scientific Inquiry. The rationale for the Data and Discovery requirement is that all CAS undergraduate students, regardless of major, should be equipped to intellectually navigate a 21st-century world in which the generation, utilization, and societal implications of data are widespread and rapidly expanding. The Core Competency in Data and Discovery applies to STEM majors because these students are also impacted by the ubiquitous presence of data in both their disciplinary training and in their lives outside the classroom. As with Scientific Inquiry, this Core Competency can be fulfilled by a Core course or by a pre-approved departmental course. We project that a large majority of CAS students will meet the Data and Discovery core competency through their major coursework."

There is also a discussion (p. 23) of the mechanism by which courses may be proposed as meeting the requirement; since this is purely bureaucratic, I omit it here. That is all the discussion of the requirement in the Framing Document.

In the presentation to the CAS Undergraduate Curriculum Committee (on which I serve), a single course was presented as satisfying the requirement: CORE-UA.111, "From Data to Discovery: Thinking with Data." Judging from the materials presented to the UCC, most of the content of this course corresponds to a traditional introductory statistics course; the list of topics includes the Chebyshev inequality, discrete and continuous distributions, the normal distribution, the central limit theorem, linear regression, logistic regression, and so on.[1] It is distinctive among introductory statistics courses in that it deals with the mathematical foundations as lightly as possible (necessarily so, since, as a core course, it has no math prerequisites); it involves very extensive use of computational tools; it includes some exposure to visualization software; and it includes some qualitative discussion of the many kinds of biases and errors that can arise in the preparation, analysis, and application of statistical inference.

The fifth bullet point in the list of "Learning Objectives," "design a mathematical model to generate predictions" is hardly achieved in the proposed course, and would be difficult to achieve in any elementary course, unless it is interpreted modestly as applying linear regression.

Overall, it would appear that the chief "core competency" envisioned in this requirement is the ability

[1]The mathematical content of the course also follows traditional syllabi in terms of basic topics that are omitted, such as the inverse power-law distribution and the difference between frequentist and Bayesian statistics.

to create or download a dataset and to run it through a statistical software package. Subsidiary competencies are the ability to use statistical graphics software to plot data and a knowledge of a few of the potential hazards encountered in doing and applying statistical analysis. If this is a correct reading of the proposal (see section 3 below) then what is being proposed is not an "expansive view" of data literacy, but a quite narrow and technical one.

2 My overall reaction to the proposal

In my opinion, there are two problems with this proposal:

- The scope of the requirement, and in particular the range of courses that would satisfy the requirement, is unclear.
- I don't think that every CAS major should be required to take a course in Data Science, however that requirement is construed.

I'll first discuss why this proposal needs to be clarified before it is even reasonable to vote on it. Then I'll discuss my view of why liberal arts college students generally should not be required to take a data science course. Finally, I'll circle back to NYU and discuss issues specific to us.

3 It is unclear how broad this requirement is intended to be, or what kinds of courses will satisfy it.

The key issue that is not clearly stated in this proposal is, what kinds of further core courses will be accepted as satisfying this requirement, and what kinds of departmental courses can be viewed as "equivalent"? The requirement description speaks of an "expansive view of data literacy"; how expansive a view is intended? There are, after all, many computational methods for analyzing data that are quite different from traditional statistics and many methods for analyzing data that are not computational or mathematical. Is the requirement limited to statistics courses in more or less the traditional sense, as I suggested earlier? Does it include courses that teach other kinds of skills for analyzing data? Does it include courses that discuss the relation of data and knowledge, but don't teach specific skills?

How critical are the subsidiary competencies of using plotting software and knowing something of the hazards of statistical analysis? For instance, MATH-UA.234, "Mathematical Statistics" is a purely mathematical treatment: it does not teach any computational tools, or discuss visualizations, or significantly discuss real-world issues, at least as far as I can tell from quickly looking over the 2021 syllabus[2] and the textbook, Would this satisfy the requirement?

Skimming through the CAS catalog, I notice the following courses that might plausibly satisfy the requirement, depending on how narrowly that is defined. I divide these into "Traditional Statistics". "Statistics+X" "Manipulating Data" and "About Data". I also omit the many lab and fieldwork courses, all or almost all of which involve collecting and analyzing data, as well as special topics classes. I've done this quickly, so no doubt there are omissions and errors. But before the faculty is asked to vote on this proposal, it would be good if the advocates of this change could review these and give their opinions on them.

Traditional Statistics:
BIOL-UA.52 Biostatistics
ECON-UA.18 Statistics

[2] https://math.nyu.edu/media/math/filer_public/17/a4/17a4e1c7-1441-4827-befc-420769eb9003/math-ua_234_spring_2021.pdf

ECON-UA.20 Analytical Statistics
MATH-UA.234 Mathematical Statistics
MATH-UA.235 Probability and Statistics
PSYCH-UA.10 Statistics for the Behavioral Sciences
PSYCH-UA.11 Advanced Psychological Statistics
SOC-UA.302 Statistics for Social Research

Statistics+X
CORE-UA 107, Quantitative Reasoning:Problems, Statistics, & Decision Making
BIOL-UA.45 Biostatistics and Human Genetics
ENVST-UA.310 Environmental Quantitative Methods
UGPH-GU.20 Biostatistics for Public Health

Manipulating Data:
BIOL-UA.03 Bioinformatics in Medicine and Biology.
BIOL-UA.120 Computing with Large Data Sets
BIOL-UA.121 Fundamentals of Bioinformatics
CSCI-UA.60 Database Design and Implementation.
CSCI-UA.381 Programming Tools for the Data Scientist.
CSCI-UA.473 Fundamentals of Machine Learning
CSCI-UA.475 Predictive Analytics
CSCI-UA.476 Processing Big Data for Analytics Applications
CSCI-UA.479 Data Management and Analysis
DS-UA.203 Machine Learning for Language Understanding
ENVST-UA.372 Geographic Information Systems for Ecology

About Data:
CAMS-UA.120 Skepticism and Proof: Research Methods in Child Mental Health.
DS-UA.111 Data Science for Everyone.
ENVST-UA.424 Environmental Qualitative Methods.
PHIL-UA.88 How Science Works
PHIL-UA.90 Philosophy of Science
PSYCH-UA.8 Data Literacy for Psychology.
SOC-UA.301 Research Methods (Sociology).

Which of these would satisfy the requirement? And how will future courses be judged?

It may well be possible to modify some of the above courses so that they meet the requirement. Without more specifics about the requirement, it is impossible to say which courses would have to be modified in what way. However, whenever a department or an instructor feels pressured to modify its own courses away from its own preferences in order to allow its students to satisfy a core requirement, that inherently incurs a cost in pedagogical quality, at least in the judgment of the department. It may be a worthwhile cost for one reason or another, or NYU-CAS may decide that, in this matter, the judgment of the Core faculty is more reliable than the judgment of the departmental faculty; but it remains a cost that has to be taken into account.

It goes without saying that there are no hard-and-fast rules, each course has to be judged on a case-by-case basis, and things will evolve over time. Still, the faculty should have a clearer idea of what we are voting for than we now have.

4 Should every student in a liberal arts program be required to take a course in data science?

Is it good for someone with a liberal arts education to know something about data science? Absolutely, no question about it. An understanding of data science is often very valuable; more on this

below.

However, that is not the question before us. Lots of things are good to know; ideally, everybody would know everything about everything. Tragically, the human condition does not permit that. The question is this: Is the kind of data literacy under discussion here now so central to a liberal arts education, that a liberal arts college should insist that every one of its graduates spend one of their 32 courses studying data science? More specifically, should a student be required to take a course in data literacy even if they are majoring in a field that do not require such a course? I will argue that the answer is no.

There are a number of arguments that might be used to support a data science requirement, for students going to college in 2024:

- Data science is an intellectually rich subject and its study enriches a liberal education.
- A knowledge of data science is important for a well-informed citizen and a beneficent member of society.
- Competency in data science is useful in careers and in life generally.
- Knowledge of data science will be useful in navigating a world in which artificial intelligence has a large presence.

All these claims have some truth to them. The question is, are these arguments strong enough to justify a data science requirement? I will examine each in turn. The proposal for the revision of the Core Curriculum presents the last three as justifications for the requirement but does not justify them at any length. It does not mention the first, which actually seems to me the key issue. For the most part, the proposal seems to view the requirement as a self-evident good. So I will do my best to flesh out these arguments before saying why I disagree with them. in the Core Revision proposal. My apologies if some of my presentations of opposing arguments come across as straw men.

Before entering on these arguments, however, let me raise two points of contrast with other hypothetical requirements that could be proposed.

As far as I know, no one is suggesting that every student at CAS be required to take an introductory course in journalism. There is a case to be made that they should. We all read or watch the news, or at least we should, and undoubtedly one can do that with a more discerning eye if one has some personal experience of how the sausage is made. The skills learned in journalism — rapidly assimilating the background of a new subject, doggedly pursuing leads, courageously asking hard questions of powerful people who do not want to answer them, poring through documents that are deliberately obfuscating to find a key nugget of truth, quickly writing up your findings for a general audience — are valuable abilities in many other contexts. And the study of journalism is especially appropriate for students at New York University (in New York), as New York is one of the world's great centers for the news media.

As far as I know, no one is suggesting that every student at CAS be required to take an introductory drama course. There is a case to be made that they should. Students in a drama class engage directly and deeply with great artistic masterpieces. They learn invaluable skills — confident public speaking, memorization, improvisation, participation in a group undertaking. It was recently reported[3] an inordinately large fraction of the movers and shakers of the world have done drama in school or summer camp. And the study of drama is especially appropriate for students at New York University, as New York is one of the world's great centers for theater.

Obviously, I could continue in this vein *mutatis mutandis*, more or less plausibly, for all the subjects taught in CAS and for many subjects not taught here. The question, then, is the case for a data science requirement *particularly* strong? Every student who takes a data science course is taking

[3] "Grown-Up Theater Kids Run the World," Madison Malone Kircher, *New York Times*, August 20, 2023.

it *instead* of some other very worthwhile course. Are we confident that we want to demand this trade-off of all the students in CAS?

4.1 Data Science as Reflective of Profound Intellectual Values

A college education is, among other things, an attempt to express and transmit intellectual ideals that are — enlightened, humanistic, profound, transcendent, cultured, civilized, noble — whatever you want to call them, they are good and they are important. Data science, like other disciplines, is fundamentally likewise an expression of these ideals, however finicky the details and however mundane or dreary the specific applications. What data science contributes to an liberal education in this sense, and how that is reflected in specific course syllabi depends on how broadly one defines "data science".

On a broad reading, data science expresses the empiricist view of knowledge. Knowledge is gained by systematically examining evidence. A claim that is strongly supported by evidence can be accepted, at least provisionally, until countervailing evidence appears. A claim that runs strongly counter to the available evidence should be rejected, again provisionally. In Bertrand Russell's formulation (*Skeptical Essays*)," "[I]t is undesirable to believe a proposition when there is no ground whatever for supposing it true." (Whether this is the only way to gain knowledge can be debated, but at least it is an important way to gain knowledge.)

However, there is a large difference between evidence and data, and there are many avenues leading from evidence to knowledge that are unrelated to those taught in any plausible data science course. Consider the following statements:

- All terrestrial matter is made of atoms.
- Archaeopteryx lived about 150 million years ago.
- The DNA molecule is a double helix.
- The galaxy GN-z211 is about 32 billion light-years from Earth (proper distance).
- The Donation of Constantine was a forgery made centuries after the death of Constantine.
- Turkish has no gendered pronouns.
- Shakespeare's *Julius Caesar* follows the account in Plutarch closely, but makes a number of alterations.
- The Danube is longer than the Rhine.

All of these claims are supported by empirical evidence, but the evidence is hardly "data" in the narrow sense, and the techniques taught in a data science class comprise at most a small part of the explanation of how the evidence supports the conclusion. The relation between the evidence and the conclusion is specific to the discipline involved.

Alternatively, one can formulate the fundamental claims of data science more narrowly, along these lines: In many circumstances, one way of collecting reliable information about a general relation is to collect multiple data points of a pre-specified form — generally speaking, the more data points, the better. There is then a large library of well-defined mathematical techniques, mostly developed over the last two centuries, that sometimes allow you to draw conclusions from the data, and to evaluate the reliability of those conclusions. The collection and analysis of the data requires care; there are many different potential flaws in how the data is collected or analyzed that can invalidate the conclusion.

I would claim that understanding the above in some depth is also part of an enlightened worldview. Statistical analysis is now one of the major sources of our collective knowledge; in some important

areas, such as pharmaceuticals, it would be crippling to exclude it. The development of statistical techniques is one of major accomplishments of applied mathematics over the last 200 years. The understanding that, on the one hand, conclusions supported by statistical analysis of many data points are generally more reliable than conclusions supported only by anecdotal evidence, and, on the other hand, that one does not have to examine all, or even a large fraction, of a population to draw conclusions about it is important in understanding the world.

A shallow understanding of this is common coin.[4] Everyone knows in a general way about opinion polls and drug testing and has a general idea of the issues involved. However, it seems to me fair to say that an enlightened worldview includes a deeper understanding of the principles and issues involved; and that someone who rejects statistical analysis *tout court* is, in that respect, an obscurantist, and someone with only a shallow understanding of these principles is, in that respect, an ignoramus.

Of course, the actual content of a data science course is not mostly concerned with grand principles, but that is no objection. The only way to gain a deep understanding and appreciation of any broad principle is to engage with its fussy and tedious details; and in statistics as in much else, the most effective way of learning is by doing.

However, there are, potentially, many different important aspects of an enlightened viewpoint, and it is not reasonable to demand that each individual, or any individual, attains them all. We are all ignoramuses in one respect or another. The question is, is the worldview associated with data science in this narrower sense especially central to an enlightened viewpoint? I certainly don't have any decisive arguments on this. But my own answer would be "No", and I expect that the answer of most of the NYU-CAS community and most educated people would be "No." Statistics as a discipline is valued but it isn't enormously valued. The figures who created the field were great scientists, but they are not, for most of us, our premiere intellectual heroes.

4.2 Data Science and the Well Informed Citizen

A more narrowly-focused version of the above argument is that an understanding of data science is vital for us as citizens and as people who contribute to the social good. For people who live in our society, and particularly for people with the kind of power and influence that we hope and expect graduates of NYU will by and large possess, it is extremely valuable to have a basic understanding of the power and limitations of data analysis. They should understand that conclusions based on statistically sound evidence is generally better than anecdotal evidence, and understand the cases where it is irrelevant. They should not be fooled by tawdry misleading graphical presentations of statistical data or slipshod statistical arguments. When they serve on a jury where a lawyer has presented statistical evidence to support their case, or if they are considering a political candidate who is using statistical arguments to argue for a policy, they should be able to evaluate the cogency of the argument. They should understand the difference between a statistically well-established connection and a large effect; the difference between the probability of A given B and the probability of B given A; the difference between correlation and causation. They should be aware of sampling bias, of confounding effects, of the file-drawer problem, and so on.

All true, certainly; but the question is, how important is data science for the citizen, among the many things that can be taught? Consider some of the major policy questions currently engaging the American citizen:

- What should be the law governing abortions?
- How far should the US commit itself to various allies who are at war or fear that they may become at war?

[4] This is new, historically speaking; it wasn't at all true in 1724 and probably much less true in 1924.

- Should ranked-choice voting be adopted?
- Should student loans be forgiven?
- What limits should be placed on police discretion?
- How can students who strongly object to the opinions of a faculty member or a speaker legitimately express their objection?

The techniques of data science has little to contribute to the discussion of any of these. NYU does, of course, teach courses that bear on most of these, particularly in the Politics and Philosophy departments, but, returning to my refrain at the start of this section, no one is proposing that these be required in the Core.

4.3 Data Science as a Useful Career- or Life-Skill

It can be argued that students will find the skills taught in this requirement useful as they go through life or pursue a career.

The argument about a career, I think, can be easily dismissed. In general if a student ends up pursuing a career that seriously requires them to carry out data analysis, then they will have to learn the techniques that are appropriate to that field. If they take a major that leads naturally to such a career, then in all likelihood the major already requires a data analysis course. A general introductory data science course is not going to prepare them for this kind of specialized requirements. I feel confident in saying that a large majority of the NYU CAS graduates that major in a subject that does not require a course in statistics will go through their entire careers without ever having to run a dataset through a statistical package.

The argument about usefulness in everyday life seems to me equally flimsy. When someone read personal advice in the newspaper or on the internet on things like diet, exercise, and so on, there are various things they can do to evaluate whether it is reliable and pertinent, but only extremely rarely will it involve their actually getting hold of the raw data and personally running statistical packages over it, or even evaluating, at a technical level, how the analysis was carried out. "I did my own research," is at this point largely a joke, and to the extent that it's not a joke, it almost never means, "I personally did a statistical analysis of the data."

4.4 Data Science and Artificial Intelligence

A final argument is that Big Data and Artificial Intelligence will soon penetrate every corner of our lives. Won't we all need data literacy in order to be able to succeed in this new world?

Actually, we won't; at least, not as "data literacy" is construed in this requirement. AI has become very prominent in the last four years, but it is anyone's guess how much impact and what kind of impact it will have over the long run. More importantly, however that plays out, knowing the kind of data science that this proposal envisions will be of essentially no value in dealing with it. If you want to know how you, as a lay person living your life, can gain the most benefit and avoid the most harm from the chatbots, image generation programs, robots, self-driving cars and so on that will exist in 2044, there is pretty much no elementary course now taught at NYU that will be of any value. Perhaps the Steinhardt course "Evolution of Technology" might help give a little historical perspective. Certainly none of the core courses under consideration as satisfying the data competency requirement are at all pertinent; and it is not at all clear that a new course that would be relevant would be considered to satisfy the requirement.

5 Back to NYU specifically

5.1 Data as Discovery vs. Quantitative Reasoning

The Data as Discovery requirement replaces the current Quantitative Reasoning requirement. I think it is safe to say that there will be few courses that satisfy Data Science that do not satisfy the Quantitative Reasoning requirement. On the other hand, there are many ways of satisfying Quantitative Reasoning that presumably will not satisfy the new requirements including: some of introductory math courses, being a science major, getting over 700 on the math subject SAT or 4 or 5 on the calculus or statistics AP tests. So this amounts to a very substantial tightening of that part of the the requirements.

5.2 The CAS Curriculum

The following CAS majors require a course in statistics:[5] Data Science, Economics, Global Public Health, Language and Mind, Neural Science, Psychology, Sociology, and the joint degrees with Tandon of Math/Civil Engineering, Math/Mechanical Engineering and Physics/Civil Engineering. Additionally Biochemistry, Biology, Computer Science, Environmental Studies, Journalism, and Math accept courses in statistics as elective credit toward the major.

That's all. I don't see why a history major, a French literature major, a math major, a chemistry major, or an anthropology major should be required to take a data science course if they feel that their time is better spent doing something else.

5.2.1 Computer Science Majors

My view on this is, naturally, very much informed by my own field of computer science.

At many schools, the computer science major requires some kind of course in data analysis; at many, it does not. At NYU, the Computer Science department in Courant/AS has no such requirement for degrees at any level; the CS departments at Tandon do require it for the undergraduate majors. Our department offers a number of courses, both introductory and advance, that deal with data in a broad sense; I have enumerated these in section 3 above. It is not clear whether any courses currently offered in the CS or Math department would be deemed to satisfy the proposed requirements or how much they would have to be modified in order to satisfy the requirement (see section 3).

My own feeling is that it is valuable for computer scientists to know something of probability theory; in fact, I regularly teach probability theory as part of "Artificial Intelligence" (CSCI-UA.0473 and CSCI-GA.2560) and of "Mathematical Techniques for Computer Science Applications" (CSCI-GA.1180), and I encourage all computer science students to take a probability course from the math department. However, it seems to me that the case for requiring CS majors to know traditional statistical techniques is much weaker. I feel quite confident in saying that most professional software engineers or computer programmers never once in their entire career run a data set through a statistical package and that many of my colleagues on the faculty of the Computer Science department have never done that.

5.3 Other Universities

I checked the math/data science general requirements in the liberal arts undergraduate program at twelve other prestigious American universities. What I found was:

[5] Again, this was done quickly and probably has errors.

- Harvard[6] has a requirement with the title "Quantitative Reasoning with Data" but which in fact include all kinds of courses with a mathematical component, such as "Introduction to Calculus", "Introduction to Computer Science" and "Sounds of Language" (a phonology course).
- U. Penn[7] has a requirement with the narrow sounding title "Quantitative Data Analysis" but which in practice is similarly broad. In spring 2024, there are 42 courses that satisfy the requirement, including "The Solar System, Exoplanets, and Life", "General Chemistry Lab", and "Introduction to Socio-Linguistics."
- M.I.T.[8] requires Calculus 1 and 2.
- Yale[9], Princeton[10], Cornell[11], U. Chicago[12] and Stanford[13] require a course in Quantitative Reasoning, somewhat like our current requirement, but often much more flexible; for instance, the Stanford requirement can be satisfied by "Earthquakes and Volcanoes", "The Mathematics of Sport", "Remote Sensing of the Oceans" and so on.
- Columbia[14], Dartmouth[15], Brown[16], and Berkeley[17] have no math- or data-oriented requirement.

Thus, none of these twelve universities requires their undergraduates has a data science requirement nearly as narrow to the one being proposed here. Now, of course, there may be other universities that do, and even if there are none, that doesn't show that we would be making a mistake in creating such a requirement. But there certainly does not seem to be a mass movement among top schools to create such a requirement, as far as I can determine.

5.4 Our Image

Finally there is the question of the image that this change presents to the world generally and to prospective applicants in particular.

Even if this requirement is interpreted expansively, it seems to me that the only possible effect on the applicant pool would be to shrink it; and if it is interpreted narrowly, as meaning, more or less, a statistics course, then it will drive away in droves such applicants as are paying attention. There are, certainly, people who really enjoy statistics as a discipline, but they are rather rare. Mostly, in the disciplines that require statistics, it is one of the majors' least favorite topics[18] and the distaste among students who plan to go into other disciplines will generally be much greater. Even if the requirement is interpreted expansively, I would think that it will give pause to some of the applicants.

In terms of NYU's image more broadly: I think it is fair to say that data science is widely viewed as primarily a tool of state bureaucracy and corporate technocracy, and though of course that view is one-sided, it certainly has an element of truth to it. The government and Big Tech have the greatest

[6] `https://oue.fas.harvard.edu/quantitative-reasoning-data`

[7] To find the courses at U. Penn that satisfy this requirement this semester: Go to `https://courses.upenn.edu/` Under Advanced Search in the left column, under "course attribute", select "College FND Quantitative Data Analysis (AUQD)". Thanks to Mitch Marcus for these directions.

[8] `https://registrar.mit.edu/registration-academics/academic-requirements/general-institute-requirements`

[9] `https://science.yalecollege.yale.edu/student-resources/academics/science-qr-courses/qr-courses/qr-courses-without-prerequisite`

[10] `https://ua.princeton.edu/policies-resources/general-education-requirements`

[11] `https://courses.cornell.edu/preview_program.php?catoid=31&poid=15425`

[12] `http://collegecatalog.uchicago.edu/thecollege/mathematicalsciencescore/`

[13] `https://ways.stanford.edu/about/ways-categories/applied-quantitative-reasoning-aqr`

[14] `https://bulletin.columbia.edu/columbia-college/core-curriculum/`

[15] `https://students.dartmouth.edu/ugar/academic-advising/faculty/resources/general-academic-requirements-graduation`

[16] `https://college.brown.edu/design-your-education/complete-your-degree`

[17] `https://guide.berkeley.edu/undergraduate/education/`

[18] I have no concrete evidence for this claim, but I am quite sure of it.

capacity for collecting and processing data, and they have the greatest opportunities for exploiting it. By announcing that data science is now a core competency required of every graduate of NYU CAS, we increase the degree to which NYU will be viewed as a ally or even a tool of the state and the corporate structure. I don't think that that's what we are, at least not primarily, and I don't think that's how we want to be viewed. I hope that, on consideration, the faculty of FAS will oppose making a change to our undergraduate core requirements that signals a shift in that direction.

Acknowledgements

I am grateful to Martin Daughtry and Trace Jordan for taking the time to discuss the requirement with me.