Paper 4705-2020

Show Me the Money! Preparing Economics Students for Data Science Careers

Steven C. Myers, Ph.D., University of Akron

ABSTRACT

Economists make great data scientists. Economics offers as a discipline many valuable skills such as problem-solving ability and storytelling. When economic theory and deep business acumen combine with applied econometric analytics plus an obsession to understand the data generating process and methods of dealing with dirty data, you have a lot of research savvy. Add a layer of deep SAS® programming and analytics, and you have the beginnings of a great data scientist.

We partnered in 2015 with the SAS Global Academic Program to offer a joint certificate in economic data analytics. Every graduate and undergraduate since have received the certificate. This paper will discuss the methodologies taught and pedagogies used in our program. Students learn programming in a team- and problem-based learning environment working on real data and real problems.

A revision of the undergraduate program in 2003 made a significant commitment to bring the success of MA students to undergraduates with much success. Starting in fall 2019, we accept students into an even more powerful curriculum leading to a BBA in Business Data Analytics, requiring analytic courses in economics and business.

The paper will discuss the state of the saturated field of analytics in higher education and how the University of Akron is adapting to remain competitive and to produce highly demanded students. The talk will highlight student success stories and provide a guide to others on how to prepare students for data science careers.

INTRODUCTION

I am pleased to write this paper in response to the invitation from the Global Forum Content Advisory Team to tell of the journey taken by us in the economics department of the University of Akron and how a partnership with SAS has dramatically enhanced our students' learning and employability. I will try to highlight a roadmap that others may learn from as they revise their curriculum to meet the needs of today's market. Economics offers as a discipline many valuable skills such as problem-solving ability and storytelling. When economic theory and deep business acumen combine with applied econometric analytics, and each graduate has an obsession with understanding the data generating process and methods of dealing with dirty data, they exhibit a lot of research savvy. Add a layer of deep SAS® programming and analytics, and you have the beginnings of a great data scientist.

CURIOSITY AND THE ECONOMIST'S OBSESSION WITH "WHY?"

Economists are curious and obsessed with the question: "Why?" I have always been obsessed in that way, and it led me to study economics. It also often got me into trouble; not everyone likes to be asked "why," especially those in authority or holding power. Nevertheless, it's a wise leader who understands the power of that question.

For most economists, it is not enough to predict or forecast the future; economists seek to know why it is occurring. Especially in non-time series problems, economists use economic theory and common sense to solve problems *in the absence of data* and then use data evidence and analysis to confirm or refute their conjectures.[1] This statistical and analytical side is strong with economists. Still, they are certainly capable of the predictive analytics that is attended to by machine learning and other derivatives from the data mining world.

Economists may not always seek causality well, and from time-to-time may fall into correlation-alone-association studies. Still, we instinctively know one thing: Data alone cannot solve problems, for they miss the reality behind the data. As economists explore the truth behind, but not in the data, they deep-dive into the business acumen that is so necessary to not only understand and solve the problem at hand. They also learn to communicate better, indeed, to translate between the data-only data-science types and the interested parties in the C-suite. Table 1 tries to summarize this effect.

	Using present and past data evidence:	And the future? outcomes
Correlations alone are **"Oblivious of Causation."** *Typically this requires no economic theory.*	We can provide description only. We can show "what has happened."	We can extrapolate our data into the future. These predictions can be quantitatively scored, but carry no hint of "why" the future will be as it is.
Causation is Not possible to determine without information outside of the data. *The outside-of-the-data knowledge can be economic theory, or a deep level of business acumen, or both.*	We can offer an explanation of behavior in the past. *This is the heart of "Problem-solving" and "Storytelling."*	Falsifiable predictions can occur. We can be reasonably confident that we know what will cause the desired outcomes.

Table 1: The challenge of correlation and causation

AUTHOR'S BACKGROUND – HOW DID I GET HERE?

I tell this journey with some personal history, but the relevancy for programs today starts below with the 2001-2003 revision if you care to skip through this journey.

As a young man, I was always curious about how things worked and why things were as they were. At West Virginia University, I found too many courses that would not address the "why?" question with the result that I often did not enjoy them. It was in my 4th semester in principles of microeconomics that I found for the first time in my undergraduate career a discipline that centered on both answering the causal questions and relying on both reason

[1] While I learned about causality while at OSU, but with little more than the concept, most date causality in economics as a "thing" starting with Angrist and Pischke (2010). This documents the hard turn in research to better empirics and quasi-experimental design. This is the start of the credibility revolution in economics.

and evidence. It was aptly taught by an economics professor, Parvis Jenab, who always had an answer backed up by rigorous thought and evidence. He is why I majored in economics. My buddy, Mick, and I would ask him "why" questions in class just to see if he had an answer, and he always did.

My first work-study job had me working at the Regional Research Institute, punching computer cards, and taking them to access the mainframe at the computer center. I did this so much that I became a volunteer operator on the computer platform, able to read in my cards and tear off my output. As I became an economics MA student, I had the great fortune of being the research assistant of Visiting Professor W. F. Gosling of the University of East Anglia, who taught me the FORTRAN language and to program a full working regression program from scratch. With each successful step, he would prod me to add something else. I learned regression from the Fortran side, having no econometrics course at that point. By the end of the summer, Dr. Gosling had me successfully program a Fortran subroutine (kind of like a SAS Macro) that he needed in his research and that the director of the institute sent off for publication. I became hooked on the computing side of economics from that point on.

With my BS and MA in economics from WVU, I joined the Center for Human Resource Research (CHRR) at Ohio State University, working on the most massive longitudinal data collection and analysis project, the National Longitudinal Surveys, and started my Ph.D. studies. Not only did I learn about data and especially survey data, I got my first appreciation for always uncovering the data generating process (DGP). Adding to my informal education was that CHRR had a satellite main-frame system and a bunch of programmers from whom I could learn.

My exposure to interdisciplinary empirical work at CHRR helped me to be a better applied econometrician. The practical business-type experience of working on a $3-million research project also proved to be invaluable in my econometric formation.

As an "applied" econometrician, the economic profession would not regard me as a 'real' econometrician. Still, fellow students came to me to get practical applied advice on how to complete their dissertations because the real econometricians on their dissertation committees were not well versed in the ways of applied empiricism and often minimized its value. While at OSU, I had developed two fascinations, one for large, robust datasets, and another for applied, not theoretical, econometrics.

In 1979, with nine econometrics courses under my belt, I joined the faculty of the University of Akron. My offer letter tasked me with building the econometric sequence for students in the MA economics program. I believed I knew what MA students needed to be employable, and that was a strong motivator for me. My students were required to learn deep core programming to be able to acquire and work with real data. They used their programming ability to transform and clean the data, to do analysis, and to write reports that non-econometricians could understand. The Akron MA economics graduates each produced a job market paper of original research they could share with prospective employers. That was the beginning, and it was pretty ugly in the first couple of years, but vastly improved over time.

SEEKING THE RIGHT SOFTWARE – 1979 TO EARLY 1990S

While I still knew and used FORTRAN, I had learned to love PL/1 at OSU and used it primarily for my dissertation. CHHR at OSU had many specialty programs to generate tables and do analysis, and we made use of SPSS for some runs of convenience. When I came to Akron, I discovered that the PL/1 version at UA was incompatible with the release at OSU. This incompatibility was widespread in those early days. For some time, I was using publisher supplied software and other programs (such as B34S, RATS, and LIMDEP), but thanks to a programmer at the computer center, I discovered SAS as a PL/1 substitute. He introduced it to me because he said that the programming language of SAS was nearly

identical to PL/1 (it was) and that the problems I was working on could be programmed more easily in SAS (they could). More and more, I incorporated SAS into my classes. This process stretched out over the early to mid-1980s.

SAS became very helpful to my students and often led them to SAS programming jobs as much or more often than to economic analyst jobs. SAS from the 1980s had become an essential part of econometric education at UA. I designed our program to be more data analytic focused and to involve more programming than covered at many of the other MA programs. Having a high empirical edge made for a tremendous comparative advantage for our program and led to high levels of employability.

MY DETOUR FROM FACULTY TO ADMINISTRATION AS CIO

From the mid-1990s through 2001, I spent time on the administrative side of the University, becoming the first CIO at the University of Akron and also holding an associate VP of distance learning position. I took my love of computing and thoroughly enjoyed my time in instructional technology and tolerated my duty to administrative computing.

GRADUATE EDUCATION AND SAS

To summarize, the graduate program was very successful for my first graduates in 1980 and stayed strong through its suspension in 2017.[2] For me, the hallmarks of that education, from an empirical point of view, included two required courses that I taught. The first, named initially, Statistics for Econometrics, was a theorem-based mathematical statistics course that started with sets and probability and ended up with a no-problems regression course. Students deeply explored statistical inference and model specification while ignoring the real thorny issues that a more in-depth study of econometrics requires. Those thorny issues get covered in the second course, Applied Econometrics. In the first course, students had many tasks, including learning SAS and in the early weeks using SAS to program a Monte-Carlo simulation to create draws from a probability distribution. Learning SAS and how to work with data was very much a part of that course, as was team-based and problem-based learning. Classical inference and model specification completed the bulk of the course time. SAS/STAT, including PROC TTEST and PROC REG, were the primary procedures used. To guide students through the depths of problems, Peter Kennedy's "Guide to Econometrics" was at their side in each of its editions. With the fifth edition in 2003 (and in the sixth, and sadly last in 2008), Kennedy included his insightful chapter on Applied Econometrics.[3]

---

[2] The suspension of our MA program came together as a perfect storm. We had moved to the college of business from the college of arts and sciences, TAs were eliminated because the CBA was a professional program and students should fund themselves, international students on which we counted were down campus if not country-wide, our faculty was active in changing the program from one based on assistantships to one based on paying customers, a total redesign was underway, and at the same time a university-wide program review was seeking out under performing programs. During our transition, a 100 percent placement rate did not top the falling numbers and our program was suspended. The good news is the current Provost and Dean have asked for it to be brought back because of the obvious value.

[3] See Myers (2020) for a view of Kennedy as an ethicist in the use of data. I personally used Kennedy as a guiding light for students separating the brightnesss of applied econom stricts from the dimness of a more theoretical approach. With his passing in 2010 that light is sadly fading.

The second course picks up where the first course left off, now considering all of the ways that problem articulation, data cleaning, and model specification could and do go wrong. Again SAS was a significant part of this course as we went deeper into SAS/EST, including PROCs AUTOREG and ARIMA, and in later years since 2000, the fabulous PROC QLIM. When SAS routines could not be applied, we explored SAS/IML as a way to get answers we desired.

The most important part of my students' education was the requirement to write a "job-market paper." This paper represented an original piece of research where they had to articulate the problem on which they were writing and formulate their hypotheses. They were responsible for acquiring, cleaning, and transforming the data and conducting appropriate analysis on it. Students were encouraged to write in a clear employer-friendly manner. The job-market papers were often used to show students' capabilities and helped many get responsible analytical jobs.

UNDERGRADUATE REVISION IN 2001-2003

Our programs at the graduate and undergraduate levels were strong, but the undergraduate program had become directionless by the early 2000s. Those who taught the optional undergraduate econometrics used SAS, but most of us thought that employers would seek only graduate talent for real analysis. Many students did not take econometrics since it was optional and therefore avoided both the knowledge and employment opportunities that came from its study. Economic forecasting was also optional and less selected then econometrics. We spent time trying to encourage students to seek a graduate degree rather than preparing them for empirical careers with their bachelor's degree. By the 2000s, more and more jobs for data analysts were appearing, courses in data could no longer be optional, at least for the segment of our BA students that wanted to compete for jobs.

We employed consultants to review and suggest a future direction for our programs. From the success of the graduate program came quickly accepted recommendations for (1) a required econometric and strongly encouraged forecasting course, (2) a need to integrate SAS programming and other computer skills more deeply into the undergraduate curriculum, as well as (3) the need for students to pursue independent research on real data in a capstone class.

This 2003 revision then had three significant component changes.

1. Each undergraduate major was required to take a brand new class called Computer Skills for Economic Analysis, which featured over the years many different 'skills,' but always included SAS programming. Students did a SAS tutorial training followed by real experiences to read in a variety of formats of data (especially Excel) and to concatenate and merge datasets. They learned to transform the data and clean it. PROCs featured were CONTENTS, PRINT, UNIVARIATE, MEANS, FREQ, and in later years, TABULATE and SGPLOTS.
2. Every undergraduate student major in economics was now required to take econometrics where SAS handled all statistical inference and economic analysis. Procedures featured PROC TTEST and PROC REG. The course, economic forecasting, typically taught with R, became a strongly encouraged elective and regularly offered during the summer.
3. Every student major had to write an individual and independent research project using SAS. The first of these appeared in 2006. Presentation of the senior projects in front of faculty and peers was the culmination of their efforts. These senior projects gained some notoriety, one of the first earning first place in a competition and presented and published in the Atlantic Journal of Economics, Wilson (2007). The most recent published senior project is Angeletti (2017). Five students have had their research published, many other students have presented at outside conferences

and the spring showcase held at the Cleveland Federal Reserve. See the appendix for some student comments on this process.

This revision to the curriculum was very successful as a program review in 2016 showed that over half of all majors ended up placed in occupations where the title included the words "analyst" or "analytics." Why only 50%, well economics is a great major for the law, and other students do go on to additional graduate work. The next major revision did not occur until recently when we moved to the College of Business Administration and formulated a specialty degree to sit alongside the BA Economics. The Bachelors of Business Administration (BBA) in Business Data Analytics took form in 2016, and we accepted the first students in fall 2019.

WHY HIRE ECONOMISTS FOR YOUR DATA SCIENCE TEAM

While jobs do advertise for the economics major, it does not have the vocational sound of accounting or marketing, and employers may not understand the training that many economists have. So many job ads for which our students are qualified, leave out the economics major. The education that economic majors receive suggests they employers should look at economics more often. We do encourage our students to apply for jobs not listing economics, but it is worth raising awareness and thereby hope that more job ads do include economists over time.

Our graduates are nearly all employed and doing some form of data analytics (50% of our undergraduates and 100% of our graduates[4]). All of our students have experience working in teams, programming, and working with large datasets. They are used to unstructured problem solving, data cleaning, and economic and econometric model specification. They have the soft skills as well as the hard analytic skills that employers want. They have an interest in understanding the data generating process (DGP) and making sense of why a relationship exists. They know that a perfect statistical fit is of little help when that fit goes against and rejects the economic model and hypotheses generated from that modeling. A strong statistical fit that predicts an upward sloping demand curve comes to mind as an excellent example of a misspecified model.[5] Let's explore the skills that a typical economics graduate will possess.

ECONOMISTS ARE PROBLEM SOLVERS

Economists are curious and, when presented with a problem, will seek to understand the question and articulate that problem in a way that an econometric process can answer. In the classroom, unstructured problem solving combined with little guidance from the professor exposes students to real-world issues and helps them hone their problem-solving skills. Students grapple with the problem and convert it into one or more hypotheses that are subject to test. Data are not necessary to solve a problem, and indeed analysis without data (theoretical problem solving) means students of economics can use economic theory to conjecture answers without data. Economists are highly skilled to answer critical what-if scenarios. One former student and current data scientist has his team 'solve' all problems theoretically, before allowing them to look at the data, precisely to assist in choosing a more optimal feature list for his machine learning models.

ECONOMISTS KNOW CAUSALITY

Actors in the economy behave in their self-interest under constraints. Economists study these actions and develop economic theory and common sense to reason underlying

[4] These percentages do not count those who have gone on the higher education or have chosen to enter institutions like the military or the peace corp.

[5] For those who forgot their principles of economics, demand curves never slope upward.

behaviors. This effects how they view the analysis of data on economic behavior. Why a consumer makes a purchase or not, why an investor decides to buy a property or invest in an instrument, why a worker leaves a job, or why one who is unemployed refuses a job offer are all the purview of the economist. More and more sophisticated techniques have developed in the last 20 years that bring causal calculus (Pearl and Mackenzie (2018)) and quasi-experimental design (Angrist and Pischke (2010)) to the solution to causal problems.

ECONOMISTS MAKE GOOD STORYTELLERS.

A good storyteller must deeply understand the business or policy case; they must have a high degree of business acumen, and often must deepen that acumen. They recognize that no problem is complete without this background. They also use this to translate data analysis into a language understood and spoken by the stakeholder. Economists make good translators between the data science geeks on one side and the C-suite on the other.[6]

ECONOMISTS HAVE EXPERIENCE WORKING IN DIRTY DATA

Data is never clean as experimental data or results of a quality control investigation. Economists have to learn to deal with data that is observational and subject to all sorts of dirtying influences. Harvard economist, Zvi Griliches, once remarked that if it weren't for dirty data, economists wouldn't have jobs. This "dirty" data is to say that lots of influences are in the data and that one must pay strict attention to the data generation process. Knowing data well and understanding its generation allows an economist to model a solution that goes beyond the forecast of a trend or the display of a correlation. They have the opportunity to predict behavior.

Data may be subject to all sorts of biases, such as selection bias and confirmation bias. Transactional data will need manipulation before use. Each type of data: Cross-section, time series, panel data, and longitudinal data will have issues with which the economist must deal. Missing observations may be prevalent, and some variables may missing but exert an omitted variable bias on your model.

ECONOMISTS UNDERSTAND THE DATA GENERATING PROCESS

How do the data generate? Is it survey data that comes by way of a standard interviewing process? Are the observations random or not-random? Understanding the DGP goes beyond standard descriptions. Data will lie to you; it does not possess truth in and of itself. For economists, especially, the data is not repeatable. There is no yesterday again; we have to deal with the historical data as experienced. We have to get it right every time.

To prepare to discover the DGP, economists need to protect themselves by becoming data skeptics (O'Neil, (2013)) and they defend themselves by assuring they are playing with an ethical set of rules such as in Kennedy (2008) and Myers (2020).

ECONOMISTS HAVE A STRONG REGRESSION TOOLKIT

Economists know well the full stack of regression models with multiple variables, some quantitative and some qualitative, in linear and nonlinear forms. They understand the use of binary and categorical variables on both sides of the equal sign and other types of limited variables. Models are specified to analyze the return on investment, calculate treatment

[6] One of the reasons I have required Haider (2016) in my first SAS course is summarized here: "This book is as much about storytelling as it is about analytics. I believe the data scientist … uses findings to tell the most convincing and compelling story. Most books on analytics are tool or method focused. ... Seldom a book attempts to teach the reader the art and craft of turning data into insightful narratives. (p. xxiii)." In my opinion, it is too heavily STATA and insufficiently SAS based, but I like its focus and have reviewed it here https://econdatascience.com/tag/stata/.

effects, discover structural breaks, and more. Knowledge of least squares, maximum likelihood, and other estimation techniques allow for analysis of cross-section, time series, panel, and longitudinal data. The same toolkit that makes economists strong in the evaluation space makes them strong in the forecasting space. The field of applied econometrics is broad and deep. Those who criticize economists for focusing on regression (often meaning an over-application of OLS) fail to see the expanse of applications that any comprehensive econometrics text covers.

ECONOMISTS KNOW MODELING

Predictive Modeling

The data analytic world has made the name predictive analytics well known. Often this is the result of data mining and extended into machine learning methods, predictions result, and targets scored. But this is not what an economist typically means by the word prediction.

In the data analytic world, a predictively accurate analysis would be considered by an economist as the result of a descriptive exercise. In such descriptive analysis, the researcher mines the data, describe the data, and models the patterns in the data. The data analyst so trained would seek the best fit and accurate representation of a next or future data point. When this analysis is on time-series data, an economist refers to this as a forecast. Many economists do participate and excel in forecasting, and more and more are being attracted to the Machine Language and AI world.

Statistical modeling

A statistical model is not an economic model. Statistical significance is a necessary but not a sufficient condition for economic significance. Statistics doesn't go far enough. Judea Pearl picks up the discussion here: "For this reason, some statisticians to this day find it extremely hard to understand why some knowledge lies outside the province of statistics and why data alone cannot make up for lack of scientific knowledge. (Pearl and Mackenzie (2018), p. 19)" Economists are statisticians that apply outside knowledge of theory and practice to model the world as it is.

Economic Modeling

The economist uses a statistical model, but in context, addressing causality and the DGP (data generating process). Combining this model with the economic background is known as the econometric model.

The economist creates an econometric model of the actors using data where the DGP is known. So if retail point-of-sale transactions are analyzed, the economist seeks to model the demand of the consumer for the product and must also take into account the supply of products from the retail establishment. The econometric model would apply the economic theory of demand and supply and would formulate hypotheses that would be expected to be seen in the data. Necessarily the economist is concerned not just with how many items are sold, but why were they bought, and how is the price determined? The economist is models the casual relationships. This verified theory is the story. A story tells why consumers come into a retail establishment and why they purchase items while simultaneously considering the other influences on the decision such as the income of the consumers, other prices of complementary and substitutive goods, and more.

WHY HIRE APPLIED ECONOMETRICIANS

What skills then do economists need to be data analysts?

They need to be trained in econometrics; you want to hire an econometrician.

These econometricians must be economists first, statisticians second, and finally computer programmers. What makes econometricians so valuable in data analysis is that they are highly competent in all three areas: economics, statistics, and computer programming. The best economists never get that order backward.

Theoretical econometricians and classically trained econometricians will have strong skills in statistical inference. Statistical inference is the use of statistics to estimate and infer results and requires a classical training in how to correctly estimate and interpret results (such as correlations and regressions) and how to perform hypothesis testing (such as will sales next period statistically and significantly exceed last period). All economists who have taken econometrics know this classical inference of estimation and hypothesis testing.

Applied econometrics, which gets into the hot employable data science and analytic space, goes well beyond the classical training in inference briefly discussed above. As shown in Table 2, Kennedy (2008) describes the three pillars of applied econometrics as (1) problem articulation, (2) data cleaning, and (3) model specification. It is these three pillars that a data science employer should seek to find among his/her data analysts. Well trained applied econometricians should offer all three.

Classical econometricians	Applied econometrician also know (**Kennedy's three pillars of applied** econometrics)
Estimation techniques	Problem articulation
Hypothesis testing techniques	Data cleaning
	Model specification

Table 2: Classical versus Applied Econometrics

ARTICULATING THE PROBLEM

Applied econometricians can take an unstructured problem and make sense out of it. An unstructured problem is likely a vital business question but not articulated in a way that can be analyzed. As economist W. Lee Hansen has observed, the highest level of learning is the ability to ask the right question. A key asset of an applied econometrician is the derivation of questions to answer the unstructured problem. Economists use economic reasoning and common sense, and a story will emerge as to how and why the economic agents behave. Testable hypotheses on and about that economic model that will determine that the economic story is "accepted" or rejected. The economist's obsession for understanding the underlying explanation for problems leads then to formulate causal sequences.

DATA CLEANING

Kennedy refers to the second of these pillars as Data Cleaning. An unsatisfying term as this second step is much more than mere cleaning of a dataset, but the wording is instructive in that economists regard all observational data as dirty. And dirty data will need to be cleaned. Cleaning data can often be the most time-consuming part of the analysis process and is an essential part.

But more than just what might be considered cleaning, the economist must discover the essential data generating process (DGP). Is the data non-random, censored, truncated, self-selected? Are relevant variables available, or must proxies or instruments be discovered? Are adequate instruments available? Among the variables of the model, are they exogenous or endogenous, and if the latter, can the data support the system-wide economic model?

MODEL SPECIFICATION

Model specification is the third pillar of applied econometrics. The applied econometrician builds a model using data variables that match the theoretical model as much as possible. Model tests of specification and robustness help to hone in on the most appropriate model. Given the dirty nature of the data and the attempts to clean it, the model specification gives rise to many other issues (some of which are a direct result of the DGP and the quality of the variables). Other issues come directly from the nature of the problem as articulated, from the economic model itself. If there is known endogeneity of right-hand side variables, then this affects the model specification and hence the estimation and testing strategy. Variables that are poorly measured or missing altogether create another level of challenge.

SUMMARY AND THE ROLE OF DATA TRANSLATION

Applied econometricians understand economic theory and reason, not just statistical significance and goodness of fit. They know that the truth is not in the data alone and can interpret results in plain language to communicate insights to business leaders and those in the C-suite. The essential soft skill of storytelling, plus their understanding of variation, knowledge of economics, and the drive to learn the business, should make them highly sought after. The role of being a data translator who can speak to both the geeks of the data team and the leaders in the c-suite may indeed be the essential skill and a primary reason to hire economists. One of our undergraduates started at a company as a data analyst. Within three years, that data analyst was the Director of Analytics precisely because he could turn insight from his data analysis into actionable intelligence and communicate it directly to the CEO in words the CEO could understand.

The difference between staying in the cubical and being invited into the C-suite is as simple as being able to translate data and tell the story.

2014: DISCOVERY OF THE SAS GLOBAL ACADEMIC PROGRAMS

I am an economist who uses SAS extensively and teaches econometrics giving students a lot of experience with data, large and small, and especially dirty. It was Dr. Ken Sanford, a SAS employee with the excellent job title of 'SAS Econometrics Evangelist,' who first confirmed my knowledge that economists make good data scientists. In Sanford (2014), he posted five reasons for that.

Economists

1. understand objective functions
2. have a very strong linear regression toolkit
3. own observational data and causality
4. have experience in articulating the problem and the solution
5. work with Big Data

Ken and I spoke several times and expanded our conversation to include our department chair and faculty. He also joined our Economics Advisory Board for a couple of years. Also to be fair, Ken has moderated his list in the last six years, but still maintains that economists make great data scientists.

THE SAS JOINT CERTIFICATE PROGRAM

"The SAS Joint Certificate program is designed to assist universities in preparing students to work in a data-rich business environment. The joint certificate documents students' coursework using SAS software to solve real-world business problems, giving students a competitive advantage in the job market. It's a great way to give special recognition to students who have shown excellence in using and applying SAS technologies."[7]

Ken Sanford referred me to Jerry Oglesby in the SAS education department and the SAS Joint Certificate program. We soon after that applied and partnered with SAS to award our economics majors with certificates signed by both Michael Nelson, Chair, Department of Economics, and Jerry Oglesby, Senior Director, SAS Global Academic Programs. We were allowed to choose any name for the certificate program, and SAS approved ours as the joint certificate in Economic Data Analytics. Spring 2015 saw the awarding of our first certificates and all graduates and undergraduates since have earned the award. This partnership has become an essential and critical part of our curriculum.

Recently, I asked Jerry Oglesby, who is now retired from SAS, about the history of the program. He said, "During a visit with Dr. Morgan Wang at the University of Central Florida in 2001, he suggested that I should develop a way to recognize and promote programs in analytics. I have always considered Morgan to be the person most critical in the creation of the SAS Joint Certifiable Program. I think when Dr. Goutam Chakraborty at Oklahoma State University joined the SAS Joint Certificate Program in 2004, that was the time that the program gained its momentum."[8] By my count of programs listed in SAS (2020), there are 84 undergraduate programs and 126 graduate programs partnered with SAS for purposes of awarding certificates in the Joint Certificate Program. Some of those programs offer multiple certificates for different program tracks.

Ken Sanford helped us recognize that our programs at both undergraduate and graduate levels were clearly in the data science space. Our curriculum embraces all four pillars of data science: Data Acquisition, Data Management, Data Analysis, and Data Reporting and Visualization.

To be able to award a joint certificate, a faculty member from a college or university contacts SAS and shares the syllabi of courses covering a minimum of 12 credit hours where SAS is required in all courses where analytical technology is appropriate. The student must also conduct a significant project using real data, and the student presents the final project results formally to their peers and faculty (see SAS (2014)).

Requirements for our undergraduate certificate include 12 credits (1) the computer skills class where SAS programming is taught, (2) a course in applied econometrics using SAS, (3) a course in intermediate economic theory, and (4) a three-credit course where the

[7] Quoted from https://www.sas.com/en_us/learn/academic-programs/resources/joint-certificate-program.html

[8] Conversation, March 2, 2020.

student completed a senior project of independent economic research using SAS.[9] The requirements for our graduate students are similar.

EXAMPLE OF A MAJOR TASK IN THE FIRST SAS PROGRAMMING COURSE

At what level do the students learn SAS?

As a partial answer, I offer one example of the intensive data work that students in their first course, Computer Skills of Economic Analysis, have done.[10] After an introductory tutorial in SAS programming showing examples of data handling, teams of students start with the research problem for the semester: Does Greater Economic Freedom lead to Greater Human Progress across all the countries of the world? I do limit and box them in a bit since the data are not small, and this is their first large scale project. I limit them to two years, five years apart, and not the entire time series of progress. The task then is to extract the Heritage Data on the Index of Economic Freedom[11] for years 2013 and 2018 and to merge them by country names. This merge, their first real merge, does not go well since the country names change between 2013 and 2018, and students must discover how to reconcile and get a clean merge. The data include the index, its 12 components, and some overall macro data, including per capita GDP. With a clean dataset, the students try to duplicate a chart showing that per capita GDP (as a financial measure of human progress) seems to rise dramatically with economic freedom scores across all countries of the world. Figure 1 shows results from the student teams as they make their first presentation on the merged and clean dataset.

[9] See our SAS Certificate webpage at https://www.uakron.edu/economics/academics/undergraduate/economics-department-sas-certificates

[10] This course is renamed Applied Econometrics I. The difference is that with the renaming, certain prerequisites are now required such as business statistics.

[11] See https://www.heritage.org/index/.

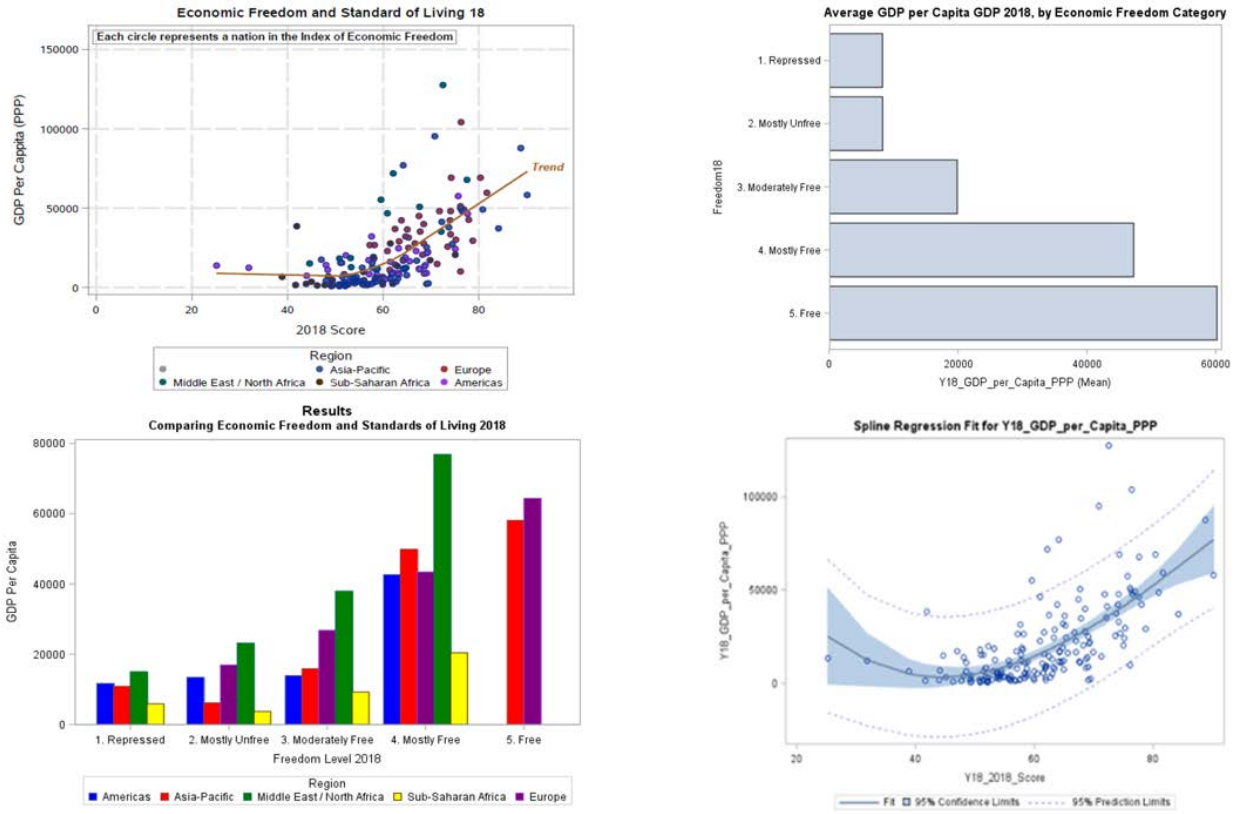

Figure 1: First student successful data merges: Index of Economic Freedom and Per Capita GDP for 2018.

The assignment was to produce a single graph and supporting statistics. I used the results to discuss good visualization practices, and I include examples with as many useful practices that I can for them to strive to meet. Figure 2 is one such example where we work through the code options to produce that graph.[12]

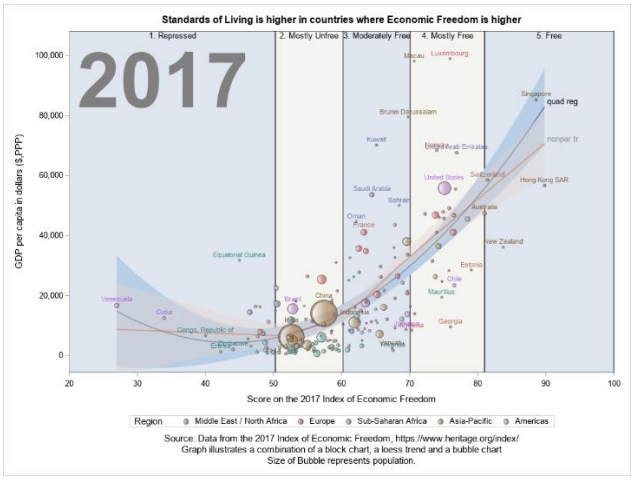

Figure 2: A SGPLOT example with code

[12] A greater explanation, including code, is available at https://econdatascience.com/economic-freedom-solve-problems-tell-stories/.

Once the student teams have the economic freedom dataset ready, the next task is for them to select five measures of human progress from the World Development Indicators dataset of the World Bank.[13] The WDI has 1600 indicators of human progress for 217 countries for the last 50 years, including child mortality, life expectancy, education levels, vaccinations, environmental levels, and so much more. The merging and cleaning of the index of economic freedom data and the WDI data proves to be a hard task for these students since the structure of the datasets are very different.

In the end, they all succeed and then proceed to analyze the overall question. I am a believer in team-based learning and problem-based learning. A larger than life task (the problem) awaits the student teams. The student teams must manage to articulate the question, clean and merge the data, and model the process, examining in every case whether the greater is economic freedom of a country, the greater is the level of human progress no matter how measured. This project takes students 6-8 weeks to accomplish. They are learning the syntax for PROC SGPLOT at the same time they are creating research data and learning to analyze it. It is real-time learning on real data.

THE STATE OF THE FIELD IN ANALYTICS HIGHER EDUCATION

Analytic programs are growing as demands for analytic talent expands. Early in 2020, Burtch Works surveyed its members and found that 67% of data science and analytic teams are planning to hire in the first half of 2020 and that 58% are hiring to increase their headcount.[14]

At the MA level, there was one analytics program in 2007 with 287 such programs today with names like Data Analytics, Business Analytics, and Data Science.[15] I cannot find a tracking of undergraduate programs, but I would expect its growth has to be similar to the graduate degrees, meaning substantial growth over time. In our department and many other economics departments, we were doing analytics and data science for years before that.

ANALYTICS CAN MEAN ANYTHING

And that is undoubtedly part of the problem of assessing the state of analytic education. Analytics means something different to everyone who uses the term, certainly from discipline to discipline and from business application to business application. Analytics to the economist means something different than analytics as understood in the IS/IT community. The communities of scholars in statistics, engineering, marketing, finance, and accounting communities also have their understandings as well.

In the context of our degree, we address the analytics of data science as applied economists have done for many years. The field is evolving, and our department is preparing our students for the future.

[13] See https://wdi.worldbank.org

[14] https://www.burtchworks.com/2020/02/26/2020-data-science-analytics-hiring-survey-results/ Accessed March 3, 2020.

[15] This according to https://analytics.ncsu.edu/?page_id=4184 at NC State.

HOW AKRON INTENDS TO REMAIN COMPETITIVE

Specifically, many things will make our degree more valuable. A list of some of those follows. The faculty already adopted many of these; others are discussed and deferred, and some are my hope.

WE ARE POSITIONED FOR SUCCESS

The Department of Economics had been in the College of Arts and Sciences since the founding of that college. Still, the faculty increasingly desired to move to the College of Business Administration. Our department joined the College of Business Administration beginning in fall 2016. The dean at the time directed us to concentrate on analytics and build contacts from the business analytic community. We believe building working relationships between the department and the local business analytics community essential. We exist in a strong place attractive to the community we serve.

WE HAVE HIRED AMAZING TALENT

While some of us are retiring from the faculty in spring 2020, we have made two great hires of new faculty, both with excellent applied econometric skills. The faculty were selected based on what specifically they brought to our latest degree program in Business Data Analytics. Our competitive edge will be that program while we continue to offer the BA in Economics. The most considerable difference is the intensity of analytic courses in the BBA.

WE WILL LEAD THROUGH BOTH UNDERGRADUATE AND GRADUATE DEGREES

We plan to bring back the master's program with the planning starting in the fall of 2020. If we pick up the last planned revision, it would be a master's in Applied Economics and Data Science. It would be like the BBA, an interdisciplinary degree rooted in economic analysis.

WE WILL CONTINUE TO OFFER A SAS JOINT CERTIFICATE IN EACH OF OUR PROGRAMS.

The SAS program and later the joint certificate have added to the desirability of our students by employers. We commit to the principle that we put the employability of our graduates in data science careers first and commit to making sure that every course offering helps achieve that end. We call this being career-ready and connected in our college.

WE WILL OFFER A TECHNICAL SUPERIOR PROGRAM

In all of our classes above the principles level, students will have at least one hands-on data experience to help students see how the topic of that class applies to the data-rich world of business and government service. We will offer a mix of skills. Our students will be well rounded in their use of programs and data access. While our economics offerings focus primarily on SAS, they also include Excel, R, and Tableau. Through the interdisciplinary offerings, students will gain other program experiences, such as with SQL, JMP, and other programs geared specifically to the chosen offerings.

This (computer) technical superiority will be matched with a serious emphasis on the ability of our students to understand and explain the economy in theoretical and practical terms.

OUR STUDENTS WILL CONDUCT REAL RESEARCH ON REAL DATA

We will continue to insist that students learn to conduct independent research using real data to assess and showcase their skills. We intend to teach and lead students in academics and research that focus on causality and participates in the credibility revolution. They will be able to model and explain both trends and patterns observed in data. They will understand prediction and scoring of explanatory models and will learn to forecast data

events. The product of the student's research will demonstrate their ability to employers. Having students involved in research will make them more employable.

WE WILL CONTINUE TO ADVOCATE FOR ECONOMISTS IN DATA SCIENCE

We intend to be at the forefront of advocating for economists in data science positions by building on our past success and making sure our curriculum stays sharp.

OUR STUDENTS WILL BE STORYTELLERS AND DATA TRANSLATORS

Our students will be well rounded in economic theory and common sense and will excel at both problem solving and story-telling.

We will focus efforts on making sure the students learn to communicate between the data and business sides of the organization for which they work. The teaching of such data translation skills will be paramount to whether they achieve positions of prominence in data science and business organizations. Students will understand that in their careers, they will need to be leaders and teach others.

BBA IN BUSINESS DATA ANALYSIS

In this section, I show the highlights of our most recent economics degree.

The Bachelor of Business Administration in Business Data Analytics degree began in fall 2019. We believe this degree will be successful because of what it offers, and because we build on our past students' career successes. A few of those successes appear in the appendix below.

The new degree will supplement and not replace our undergraduate degrees in economics as the economics course requirements of the BA in economics and the BBA are identical. Table 3 shows the outline of the curriculum for this new degree. Items marked with asterisks are the same as in the existing BA Economics program.

Curriculum components	BBA Business Data Analytics
Math and statistics prerequisites	Statistics*, Algebra*, Calculus for Business Application
Business-specific requirements	Business core courses
Analytical requirement	Database management
Required economics quantitative core*	Applied econometrics I (former computer skills for economic analysis) Applied Econometrics II Economic time series and forecasting
Required economics courses*	Intermediate Microeconomics, Intermediate Macroeconomics, and economic electives
Focused interdisciplinary analytics courses	Three courses of focused study in either (1) financial analytics, (2) marketing analytics, or (3) management analytics.

Table 3: BBA in Business Data Analytics

CONCLUSION

This paper is a combination of a walk down memory lane and a chronicle of the quality of the data analytic program and students that have come out of the University of Akron over my 41-year career. It also outlines my hope for the future of our terrific programs. The most striking changes occurred less than twenty years ago and were again enhanced in just the last few years. It is a story of modifying the program as our student's needs change. However, a constant throughout the years has been the strong place of SAS in our curriculum. The addition of the partnership with the Global Academic Programs strengthens that cause. I know the department is well-positioned for a new beginning, which will build on the success of the past and will keep a keen eye on the needs of the data science landscape.

REFERENCES

Angeletti, Richard D. III. 2016. "The Payoff of a Healthy Lifestyle while in College," Undergraduate Economic Review: Vol. 13: Iss. 1, Article 11. Available at: https://digitalcommons.iwu.edu/uer/vol13/iss1/11

Angrist, Joshua D, and Jörn-Steffen Pischke. 2010. "The Credibility Revolution In Empirical Economics: How Better Research Design Is Taking The Con Out Of Econometrics." *Journal of Economic Perspectives* 24 (2): 3-30. American Economic Association. doi:10.1257/jep.24.2.3.

Haider, M., 2015. Getting Started with Data Science: Making Sense of Data with Analytics. IBM Press.

Judea Pearl and Dana Mackenzie. 2018. The Book of Why: The New Science of Cause and Effect. Basic Books.

Kennedy, Peter. 2008. Guide to Econometrics, 6th Edition. Blackwell.

Myers, Steven C. 2020. Ethics Rules in Applied Econometrics and Data Science. In Bill Franks (ed.) 97 Things About Ethics Everyone In Data Science Should Know, O'Reilly Media. (forthcoming).

O'Neil, Cathy. 2013. On Being a Data Skeptic. O'Reilly. Access for free at https://www.oreilly.com/data/free/being-a-data-skeptic.csp.

Sanford, Kenneth. 2014. Why corporate economists are hot again and a great source for analytical talent, Subconscious Musings (blog), August 27, accessed at http://blogs.sas.com/content/subconsciousmusings/2014/08/27/why-corporate-economists-are-hot-again-and-a-great-source-for-analytical-talent/ on December 2, 2014.

SAS® Joint Certificate Program, webpage. https://www.sas.com/en_us/learn/academic-programs/resources/joint-certificate-program.html

SAS Institute, 2014. SAS® Joint Certificate Brochure: Distinguished your program and your students. Accessed at https://www.sas.com/content/dam/SAS/en_us/doc/other1/sas-joint-certificate-program-107021.pdf.

SAS Institute, 2020. SAS® Global Academic Program, Institutions that offer the SAS Joint Certificate Program. Accessed at https://www.sas.com/content/dam/SAS/documents/technical/education/en/joint-certificate-program.pdf.

Wilson, Jeffrey. 2007. Peer Effects and Cigarette Use Among College Students. Atlantic Economic Journal 35, 233–247. https://doi.org/10.1007/s11293-007-9064-z

ACKNOWLEDGMENTS

My first thanks go to Josh Horstman, a member of the content advisory team, who invited me to tell the successful Akron story. Thank you to Ken Stanford, who helped us navigate the data science field and who is such a staunch advocate for economists in data science. I am grateful for the mentorship and friendship of both Josh and Kirk Paul Laffler over these past years. Thanks also to Rochelle Fisher, our academic program manager at SAS, and Lynn Letukas, Director of the Global Academic Programs for SAS, for their support and especially for their ideas for this paper. Both have made this last year, quite a fun ride. I must thank my department chair, Michael nelson, and indeed all the faculty who have accepted my SAS advocacy over the decades. Finally, my thanks go to my wife, and best friend, Kimberly Myers, who accepts me writing when other opportunities go begging, such is her support for my present and our future.

APPENDIX: EXAMPLES OF STUDENT SUCCESS

Sam Myers, BS Economics 2016
Director of Analytics, Area Wide Protective, Akron Oh

> "Working in analytics, I have found my time with the Economics department at The University of Akron to be paramount to my success. The ascension in my career has not only been due to my ability to read data but my ability to interpret findings to c-level coworkers in a concise and direct way, as taught through the professors at UA."
>
> "Working with big data is all about being an analytical storyteller, and the economics department at UA prepares its students to not only be able to think critically with the data, but to tell its story in a way that all audiences can understand."

Teodora Lang, MA Economics, 2014
Head of Data Science & Analytics at Saint-Gobain UK & Ireland
Coventry, United Kingdom

> "Working in a data science environment is one of the 'hottest' career tracks today. Critical thinking, the ability to interrogate data, and to ask the right questions is essential in the field. Some of it is undoubtedly talent - but the rest needs to be developed through rigorous training. The Economics MA program at The University of Akron is specifically designed to provide the essential analytical skillset, but what sets it apart in the market is its unique approach to practical applications. Whether that's through project work, SAS / R training, or studying the basics of statistics and machine learning, the MA program proved to be a catalyst in my analytics career."

Erjon Gjoci, MA 2010
Economist, Consultant, Lecturer
New York City

> "At the University of Akron, I was able to identify real-world issues using empirical analysis. Engaged in diverse group conversations as a leader and as a team member. Presented research findings to classmates, guests, and faculty. And finally, UA enhanced my prioritization skills and gave me the tools to deliver efficient and effective on-time projects."

STUDENT SUCCESS COMMENTS BASED ON THEIR SENIOR RESEARCH EXPERIENCE

"The Senior Project was the biggest and most intensive school project I have ever undertaken."

"This process is much more time-intensive than I ever could have imagined. Through much labor and many hours sitting in the lab looking at SAS and manipulating my programming, I eventually was able to produce valid output. There are no words to describe what that felt like!"

"Without a project such as this, there is no evidence that a student has the ability to find, develop and synthesize a complete document using the student's own original ideas along with the implementation of a complete data set. With this project, I am now confident that I can handle a larger project."

"Working on this project was, by far, the most rewarding and challenging experience of my college career. The difficulty of the paper and the tears of frustration cried over it, however, made the final product all the more worthwhile."

"Perhaps the hardest part was staying on track because as you went along –you hit so many different obstacles, such as missing data, results that didn't seem meaningful, etc.— however, when you finish your paper, you have shown that you can complete a difficult project. You can find data, write a paper, and do econometrics on another level."

"Working with SAS was frustrating. But now I am out in the job market, I've found at least a third of the positions specifically mention wanting SAS".

"This assignment was, at times, fun and, at times, frustrating. Looking back, this was an amazing experience that will not only prepare me for graduate work but also in my future work experiences. It taught me the research process and how to manage a large scale project over a long period of time."

"Other projects and papers seem like 5th-grade art projects compared to this. Nearly 15 weeks of work, stress, and lost sleep went into making this paper the best it could be. Every student of the University of Akron should have the opportunity to create new knowledge and use their skills to their fullest potential."

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Steven C. Myers
Associate Professor and 2020 SAS Distinguished Educator
Department of Economics
College of Business Administration
The University of Akron
myers@uakron.edu
https://econdatascience.com
https://www.linkedin.com/in/stevencmyers/