HOW TO CHOOSE COURSES IN THE 2018 ICPSR SUMMER PROGRAM
FIRST FOUR-WEEK SESSION

When you enroll in a Four-Week Session of the ICPSR Summer Program, you choose a set of courses that you plan to attend. These selections should be based upon your own substantive and methodological interests, your previous course work in mathematics and statistics, your current capabilities, and your research objectives. Sometimes individuals make their initial course selections based upon course titles, without fully comprehending the implications of the course contents. Sometimes course selections are based upon suggestions from a faculty member or fellow student at your home institution. These advisors and colleagues probably have the best intentions, but they may be identifying ICPSR courses based upon what they already know, what they wish you to learn, or what they want you to know when you return home (in order to help them). As you can imagine, this is not necessarily the course selection strategy that is best for you!

In this document, you will find an informal course-by-course discussion of what is involved to ensure the successful completion of any given class. Please take some time to review these comments. You will have an opportunity to discuss your course selections with a counselor at registration/check-in on the first day of the session. Rest assured that we will try to help you select the set of classes that best meets your personal and professional needs.

Computing

Introduction to Computing is a lecture course that will provide a basic overview of the three major statistical software packages used most frequently in the social sciences: SPSS, STATA, and SAS (they will be covered in that order). We recommend that you go to the first Computing lecture, if at all possible. On that day, the instructor will provide the schedule for the entire course. The instructor will also cover the basics of the ICPSR Summer Program computing environment (i.e., useful information for all participants, whether they attend the rest of the course or not). The instructors in the statistical workshops use different software packages, depending upon their own needs and interests. So, you may only want to attend the Introduction to Computing sessions that deal with the packages used in your other courses. On the other hand, you are welcome to attend as many of the computing lectures as you wish.

Along with the other statistical software packages, we offer a separate lecture course on the R statistical computing environment. This software is utilized in some of our more advanced workshops, including Maximum Likelihood Estimation I: Generalized Linear Models, Bayesian
Modeling for the Social Sciences I: Introduction and Application, and Scaling and Dimensional Analysis. These lectures will be offered in the early evening during the first two weeks of the session. The R lectures are intended for those who have had little or no prior experience with this software. But, they do cover enough material to facilitate proficiency in a software system that is often regarded as the “Lingua Franca” of modern statistics.

Mathematics
We recommend that everyone attend a math lecture. The choice will depend upon your previous math background and the statistic courses that you will be attending.

Mathematics for Social Scientists, I is the best complement to the Statistics and Data Analysis I: Introduction workshop. This lecture class is appropriate for individuals with little or no background in math (at least since undergraduate days).

Mathematics for Social Scientists, II provides brief overviews of two topics: matrix algebra and calculus. The knowledge of matrix algebra is essential for all of the statistics workshops from Regression II all the way up through Regression III, MLE, and Bayesian. The calculus is useful—and some would say essential—for the MLE and Bayesian courses.

Mathematics for Social Scientists, III covers probability distributions and calculus (integration). This information is useful (and, again, often considered essential) for any of our more advanced courses, such as MLE, Bayesian, and Advanced Multivariate Methods.

Participants with Little or No Statistics or Math Background
If you have had little or no prior training in, or experience with, statistics then there is only one place to start— at the beginning! Please do not fret about this. It is the one thing we all have in common. We all have to take the first step!

Statistics and Data Analysis I: Introduction provides a basic introduction to statistics, probability, and data analysis. Topics include data acquisition/management, classification, and summarization; basic probability; exploration of common distribution used in statistics; along with confidence intervals and hypothesis testing. If you are staying for the entire eight-week program, then you also should take Statistics and Data Analysis II: The Basics of Regression in the second session. (These two courses comprise an integrated sequence).

Along with the introductory statistics workshop you should also attend the Mathematics for Social Scientists, I lectures. Many individuals stumble in their first statistics course because they have been away from mathematics for a long time. So, it will actually help your statistics learning to also refresh (or learn anew) the various mathematical skills covered in this class. In addition, you should attend at least the first week or so of the Introduction to Computing lectures.
Regression Analysis
There are three Summer Program workshops that cover multiple regression analysis; they are designated I, II, and III. Each has a different intended target audience.

Regression Analysis I: Introduction is best suited for those who have had a basic introduction to statistics that covered topics up to the beginning of simple bivariate regression (i.e., the usual coverage of the first-semester statistics course). The course gives a straightforward presentation of how to use and interpret multiple regression (in scalar notation). It is best suited for those who have not been exposed to the topic before or may have struggled with it in a previous course. It is also an excellent course for those who want to refresh themselves on the basic logic and application of regression analysis to feel comfortable with one of the main building blocks of social science research.

Regression Analysis II: Linear Models is one of the most popular courses in the ICPSR Summer Program, and it is the workshop that is appropriate for many graduate students. This course provides solid and fairly comprehensive coverage of the general linear model. It presents multiple regression in matrix form and devotes a great deal of attention to strategies for dealing with violations of the basic regression assumptions. The presentations include both the mathematical foundations and substantive applications of multiple regression. Many Summer Program participants have probably taken a similar course at their home institution (often during the first year of graduate school). Even so, a second exposure to the subject matter is often very useful as a review. This workshop is also a “gateway” course in the sense that the material it covers is prerequisite for most of the more advanced workshops in the Summer Program.

Regression Analysis III: Advanced Methods goes beyond the standard multiple regression courses, into new and modern and alternative forms of analysis using graphical, nonlinear, and nonparametric techniques. This course is intended for those who feel comfortable with the general linear model and want to explore its extensions. It provides useful perspectives on many aspects of regression analysis that often do not receive much attention, although they are often encountered in everyday research (i.e., nonlinearity and outlier observations). The course takes a modern data-analytic approach and relies heavily on the use of graphical tools to facilitate more accurate and complete interpretations of regression models.

Other Methodological Approaches
In addition to the Summer Program’s offerings on regression analysis, we also have two courses that cover other methods and analytical strategies.

Meta-Analysis: Introduction and Application shows you how to systematically assess the results of previous research to derive stronger, more powerful conclusions. It is appropriate for anyone—regardless of discipline or field—who wants to make sense of the conflicting
findings that are often presented in existing studies, as well as for those who want to undertake their own analyses to provide more compelling results. The workshop covers the various strategies of combining evidence across different research studies; integrating multiple studies into a single statistical framework; yielding more precise estimates of effect sizes; allowing for unique treatment comparisons; explaining differences across various study results; and identifying areas for future research. The workshop assumes a basic knowledge of statistics and linear regression analysis, but it does not require any additional preparation or mathematical training. Participants who need to review the fundamentals of regression models are also encouraged to take the two-week workshop on Regression Analysis I offered during the first half of the First Session.

Network Analysis I: Introduction focuses on relationships between social entities. However, the paradigm requires a new and different set of concepts and analytic tools, beyond those provided by standard quantitative statistical methods. The key idea is that the entities under investigation can interact with each other. Network analysis provides tools for representing these interactions in ways that reveal interesting characteristics of the observations. Familiarity with matrix algebra and linear regression are helpful, but not required, for this workshop. If you need a refresher on these topics, you should also attend the first half of the Mathematics for Social Scientists II lecture (which covers matrix algebra) and the two-week workshop on Regression Analysis I (which provides an excellent foundation in linear models) during the First Four-Week Session of the Summer Program.

Beyond Regression: More Advanced Statistical Methods
All of the courses discussed below presuppose a strong background in multiple regression and a working familiarity with basic matrix notation. This is not merely a recommendation; it should be regarded as a requirement. Those participants without a rigorous background in regression analysis will face a very steep learning curve in any of the more advanced and demanding multivariate courses.

Maximum Likelihood Estimation I: Generalized Linear Models (or MLE) is the most popular advanced course in the Summer Program curriculum. It is considered by many participants (especially those in political science and economics) to be second only to regression in the importance and utility of its subject matter. Major topics covered in this course include logit and probit models, as well as other extensions of the general linear model for categorical, ordered, and limited dependent variables.

Bayesian Modeling for the Social Sciences I: Introduction and Application (or, simply, Bayesian Methods) is a relatively new, powerful, and increasingly popular methodological strategy. It is based upon likelihood methods for inference, but it also brings prior information into the estimation procedures. This workshop assumes a very thorough understanding of multiple
regression, matrix algebra, and the principles of MLE. Some calculus would also be very helpful.

*Scaling and Dimensional Analysis* covers strategies for creating geometric representations of multivariate data. These methodologies are useful for data reduction, evaluating sources of variability within data, optimizing the measurement properties of a dataset, and producing graphical depictions of data. Techniques covered in this class include summated rating (or “Likert”) scales, unfolding methods, principal components, factor analysis, and multidimensional scaling. Participants taking this course should be familiar with the multiple regression model; knowledge of matrix algebra is very useful, but not absolutely required.

*Time Series Analysis I: Introduction* covers regression analysis of data that have been collected over time. Because the units of analysis are sequential observations on the same entities, they cannot be regarded as a random sample. This violates some of the fundamental assumptions in regression analysis and therefore requires special methodological techniques. Participants in economics, business administration or public policy often find this an appropriate course selection.

*Multivariate Statistical Methods: Advanced Topics* covers statistical techniques for dealing with multiple dependent variables in a single model. Specific techniques covered in this workshop include principal components analysis, factor analysis, canonical correlation, and cluster analysis. Note that the title of this course sometimes can be a bit confusing. Many participants believe they want to learn “multivariate” techniques in order to model the effects of several independent variables on a single dependent variable. That is NOT what this course is about! Again, this workshop covers methods that are used to deal with multiple dependent variables.

*Multilevel Models: Introduction and Application* covers regression and similar models for data that are clustered within groups (e.g, students within classes, voters in different precincts, survey respondents in different nations, etc.). Such models are known by many synonyms, including hierarchical linear models, general linear mixed models, and clustered data models. The defining feature of these models is their capacity to provide quantification and prediction of random variance due to multiple sampling dimensions (across occasions, persons, or groups, or other clusters or contextual layers such as location).

**Mathematical Modeling: Game Theory, Rational Choice, and Social Choice Theory**

There are three formal theory courses offered in the first session. *Game Theory I: Introduction* covers the analysis of strategic choice and provides a broad overview of non-cooperative games. The workshop on *Rational Choice Theories of Politics and Society* investigates the ways that actions taken by multiple self-interested actors and decision-makers often lead to stable aggregate outcomes. *Social Choice Theory* focuses on the content and proofs of the classic
theorems (Arrow, Gibbard-Satterthwaite, etc.). These results speak not only to the difficulties in aggregating the preferences of individuals, but also to the difficulties in aggregating any set of diverse criteria, with the aim of examining the properties of different collective choice procedures and characterizing procedures that yield desirable outcomes. While there are no mathematical prerequisites for these courses, students should be prepared for the use of mathematical notation and concepts such as sets, functions, equations, and inequalities.

**Substantive Course: Race, Ethnicity, and Quantitative Methodology**
*Race, Ethnicity, and Quantitative Methodology* provides an overview of the major theories and empirical approaches to the study of intergroup attitudes. The workshop focuses on different methods in scaling and dimensional analyses, and their applications in the corresponding literature.

The workshop assumes a basic knowledge of statistics and familiarity with linear regression. As we focus on measurement theories, participants are strongly encouraged to enroll in *Scaling and Dimensional Analysis* or *Multivariate Statistical Methods: Advanced Topics* or both.

**The Hubert M. Blalock Lecture Series**
The Blalock Lecture series is offered throughout the four-week session. These presentations cover a wide variety of topics in advanced quantitative methods (e.g., Data Mining; Statistical Graphics, and so on), race and ethnicity (e.g., previous presentations have included Latinos and the Changing of America and The Riddle of Black Conservatism), and professional socialization (e.g., Writing Grant Proposals, Academic Publishing in Books and Journals, and Teaching Statistics). The Blalock Series is completely optional. Participants should attend the presentations, as their schedules allow, on the topics they find interesting. Note, however, that the Blalock series strives to include sessions on subjects that are of great interest to the social science research community and the presenters are excellent public speakers. For these reasons, participant reactions to the Blalock Lecture series have always been extremely positive. So, we recommend that you make them part of your own ICPSR Summer Program experience!

**How Many Courses Should You Take?**
Summer Program participants are often tempted to elect many more courses than they can safely navigate. While this might seem to be an attractive approach, it might not be the most rewarding or useful. It is physically possible to attend Summer Program classes from 9:00 a.m. until 9:00 p.m. every day. But, those who attempt such a schedule usually experience intellectual (and emotional) burn-out. ICPSR Summer Program courses are just too crammed with material, too demanding, and too time-consuming for that strategy to be successful on a regular basis.
The general guideline is that most Summer Program participants take two workshops, along with one or more additional lecture classes. With respect to the lectures, almost everyone takes one of the mathematics courses and attends at least a week or so of the computing lectures (depending upon the software requirements in their workshops). The choice of workshops, of course, depends upon your own methodological and substantive interests. Some participants take one workshop per four-week session. That is a perfectly reasonable schedule, and it would comprise what some people consider to be a “full course load.”

If you decide to participate in two workshops per session, you may want to designate one as your “primary” course and keep up with all the work (i.e., attend all of the sessions, participate in class, complete homework exercises, etc.) in it throughout the entire session. You could then audit the other “secondary” course (e.g., attend the classes but not complete some or all of the homework assignments) and still receive good exposure to the material.

A few hardy souls participate in three workshops during a single four-week session. While we understand their motivation and interpret their plans as a compliment to the quality of Summer Program courses, we still caution you about choosing to follow this route. As stated earlier, three workshops would require an enormous amount of work, even if you only audit two of them. Most participants who try to do this stop attending one or more of the workshops partway through the session. And, we often find that the net results actually are less satisfactory than would have been the case if the participant had elected a smaller number of workshops at the outset.

**Note:** Three of the workshops in the First Session are offered in a special two-week format: *Regression Analysis I* is only offered during the first two weeks of the session; *Meta-Analysis* and *Network Analysis I* are only offered during the second half of the First Four-Week Session. This gives you the opportunity to take a combination of two of these workshops during the Four-Week Session. And, you are still able to take another workshop, as well as math and computing lectures, during the entire four weeks.

We want you to select the combination of Summer Program courses that is most relevant and useful for your interests and professional objectives. If you are undecided about exactly which workshops and lectures to take, you will have opportunities to talk with counselors who can advise you about your choices. We also encourage participants to “shop around” during the first day or two of the session; if you really cannot decide between two classes, then attend each one on consecutive days, and use that to guide your choice. Our instructors expect participants to do this, so you won’t offend them! We do recommend that you decide on your course schedule as early as possible—certainly no later than the third day of the session.

We hope that you find these comments useful when you are electing or amending your Summer Program course schedule. Please do not hesitate to ask or consult or email with the
staff during this process. We thank you for your participation in the ICPSR Summer Program and sincerely hope that your experience with us will prove to be an experience that is uniquely positive from all perspectives—academic, professional, and social.