PS1

PS1

By: Bethany Bailey

Part 1: Write a Data Section

Assigned Dataset: U.S. natality data (2016)

The data used in this paper is the National Vital Statistics System's natality birth data, which was compiled through a cooperation between the Center for Disease Control and Prevention's National Center for Health Statistics (NCHS) and the States. Statistics on births from birth certificates are collect in each state and the District of Columbia's vital statistics office and compiled by the NCHS. The data can be found in multiple formats (Stata, SAS, and csv) on the National Bureau of Economic Research's site or the Center for Disease Control and Prevention's website. It is curated by the CDC, and also hosted by the NBER.

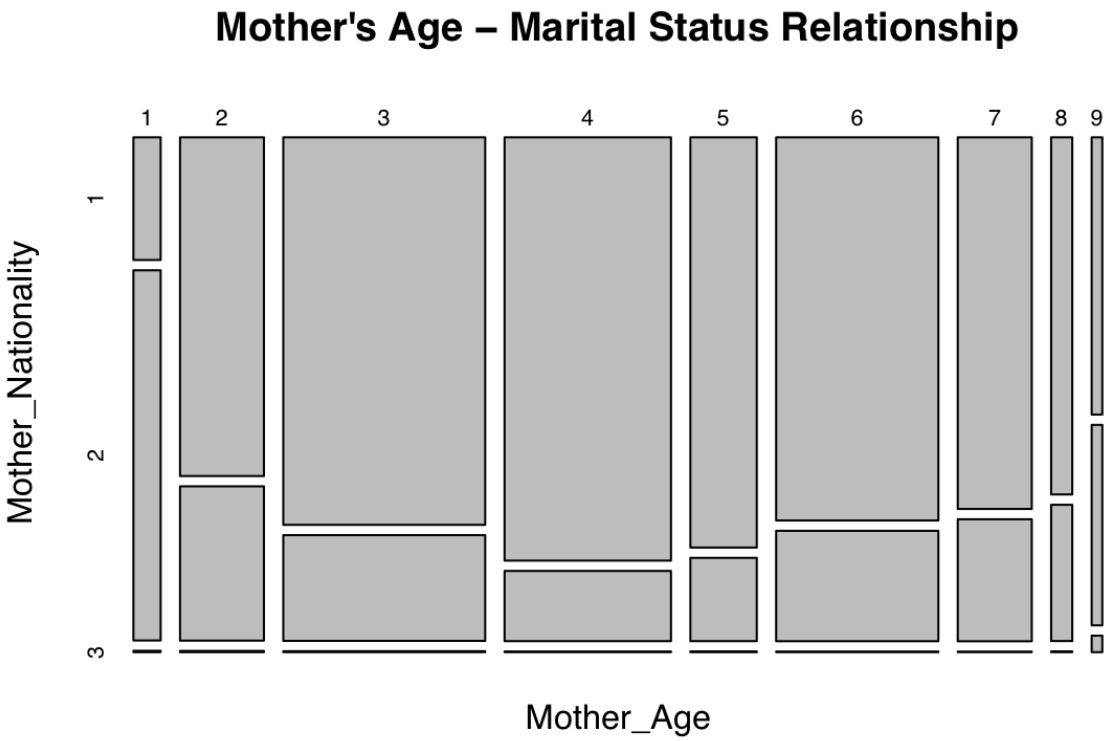
Since 1985, this data has been a 100% sample of births in all states and the District of Columbia. It includes demographic, health, and geographic data. Recently, this data has been collected using the 1989 U.S. Standard Certificate of Live Birth (from the DHHS). The CDC collects and distributes this data in an anonymized way by using this certificate and each state's Department of Vital Statistics.

Previous research using this data has looked at the effects of cigarette smoking during pregnancy (Drake, Driscoll, and Matthews 2016), teen birth rates (Kearney and Levine 2012; Hamilton, Rossen, and Branum 2016), and reproductive trends (Hamilton and Kirmeyer 2017).

Below are some summary statistics to give the reader a feel for this data. These statistics include data on age, birth place, educational attainment, race, marital status, and cigarette use.

Variables	Data Statistics
Mother's Age	
min	12
max	50
mean (sd)	28.71 ± 5.82
Father's Age	
min	1
max	11
mean (sd)	5.63 ± 2.36
Birth Place	
Hospital	3,892,038 (98.38%)
Freestanding Birth Center	21,124 (0.53%)
Home	29,621 (0.7487%)
Clinic/Doctor's Office	466 (0.01%)
Other/Unknown	3,366 (0.085083537%)
Mother's Education	
mean (sd)	4.37 ± 1.81
Father's Education	
mean (sd)	4.90 ± 2.33
Father's Race	
White	2,434,182 (61.53%)
Black	471,817 (11.93%)
AIAN	25,827 (0.65%)
Asian	223,920 (5.66%)
NHOPI	9,737 (0.25%)
More than one race	75,650 (1.91%)
Mother's Marital Status	
Married	2,383,288 (60.24%)
Unmarried	1,572,824 (39.76%)
Cigarettes Before Pregnancy	
Nonsmoker	3,566,462 (90.15%)
Smoker	96,485 (2.44%)
6-10	126,575 (3.20%)
11-20	125,529 (3.17%)
21-40	18,639 (0.47%)
41 or more	3,072 (0.08%)
Unknown or not stated	19,350 (0.49%)

As you can see from the below mosaic plot of nationality and a mother's age, young mothers (group 1, under 15 years of age) are disproportionately born outside the U.S. (group 2).



The below table shows the breakdown of the descriptive variables by nationality (mbstate_rec). Individuals born in the U.S. are 1, outside the U.S. are 2, and unknown are 3.

	mbstate_rec: 1 (N = 3027705)	mbstate_rec: 2 (N = 920102)	mbstate_rec: 3 (N = 8305)
Mother's Age			
min	12	12	13
max	50	50	50
mean (sd)	28.20 ± 5.74	30.39 ± 5.76	28.78 ± 6.14
Father's Age			
min	1	1	2
max	11	11	11
mean (sd)	5.58 ± 2.45	5.78 ± 2.02	7.49 ± 3.03
Birth Place			
Hospital	2,971,218 (98.13%)	912,718 (99.20%)	8,102 (97.56%)
Freestanding Birth Center	17,887 (0.59%)	3,220 (0.35%)	17 (0.20%)
Home	27,538 (0.9095%)	2,037 (0.2214%)	46 (0.5539%)
Clinic/Doctor's Office	417 (0.01%)	49 (0.01%)	0 (0.00%)
Other/Unknown	2,824 (0.093271967%)	504 (0.054776536%)	38 (0.457555689%)
Mother's Education			
Mother's Education	4.44 ± 1.69	4.14 ± 2.12	4.97 ± 2.59
Father's Education			
Father's Education	4.99 ± 2.27	4.58 ± 2.48	6.33 ± 2.72
Father's Race			
White	1,992,963 (65.82%)	438,774 (47.69%)	2,445 (29.44%)
Black	379,766 (12.54%)	90,786 (9.87%)	1,265 (15.23%)
AIAN	24,500 (0.81%)	1,295 (0.14%)	32 (0.39%)
Asian	45,346 (1.50%)	177,962 (19.34%)	612 (7.37%)
NHOPI	5,165 (0.17%)	4,482 (0.49%)	90 (1.08%)
More than one race	65,121 (2.15%)	10,410 (1.13%)	119 (1.43%)
Mother's Marital Status			
Married	1,752,447 (57.88%)	626,810 (68.12%)	4,031 (48.54%)
Unmarried	1,275,258 (42.12%)	293,292 (31.88%)	4,274 (51.46%)
Cigarettes Before Pregnancy			
Nonsmoker	2,653,902 (87.65%)	905,828 (98.45%)	6,732 (81.06%)
1-5	91,092 (3.01%)	5,148 (0.56%)	245 (2.95%)
6-10	123,226 (4.07%)	3,120 (0.34%)	229 (2.76%)
11-20	122,953 (4.06%)	2,332 (0.25%)	244 (2.94%)
21-40	18,316 (0.60%)	294 (0.03%)	29 (0.35%)
41 or more	2,972 (0.10%)	91 (0.01%)	9 (0.11%)
Unknown or not stated	15,244 (0.50%)	3,289 (0.36%)	817 (9.84%)

Part 2: Critique a Computational Research Paper

My Paper: Roach Anleu, S., & Mack, K. (2015). Performing authority: Communicating judicial decisions in lower criminal courts. Journal of Sociology, 51(4), 1052-1069.

Research Question

My assigned paper analyzes how magistrates perform their judicial authority through communication by looking at the delivery of face-to-face decisions in open court. The key research question that is directly stated in the paper is as follows: Are sentencing decisions communicated in distinct ways? The authors ask more specific questions about communication, such as how much legitimacy work is required in different types of sentencing decisions, operationalized by how frequently judges look at the defendant while making different decisions, how decisions are ordered, and how lawyer presence influences these behaviors.

Data

The data used in this paper was drawn from "a national court observation study of criminal cases in Australian lower courts" (1053). The data consists of 1,287 individual court matters that have information on three aspects of judicial behavior:

(1) "whether the magistrate looks at and/or speaks directly to the defendant,
(2) the magistrate's ordering of the decision and the reasons, and
(3) how these encounters are affected by the presence of a legal representative." (1053)

This data was collected by the researchers using a template to record their observations on information related to the defendant and his/her offences, the legal representation, the magistrate's interaction, and decisions and outcome information, with space for additional comment (1056). They also collected additional information, such as defendants' demographic data and offence categories, from court records (1057).

The data was covered 6 percent of the population of Australian magistrates (27 different magistrates) (1057). The magistrates were in 30 different court sessions and varied in gender, age, and level of experience. This included magistrates in 20 courts from each state and territory and from regional centers, suburbs, and capital cities. The data includes information on proceedings related to offences such as "drink driving, theft, assault and some drug offences and includes decisions on bail, adjournments, standing matters down (to be heard later in the list), setting the matter for another procedure, such as a trial, taking guilty pleas, and sentencing" (1056).

Theory

This paper drew from sociological theory on judicial interactions, including Max Weber's theory of authority and Irving Goffman's theory of performativity. The authors draw from these theories to state that judicial officers must perform their authority daily in courts in order to gain legitimacy. They also draw on more recent theory that states that judicial officers "emphasize impersonality and dispassion as central to neutrality and therefore to legitimacy and legal authority (Bandes 2009; Maroney 2011)" (1054). However, other literature states that while communicating decisions, by directly communicating with the defendant, judges are able to achieve further legitimacy (1055) (Carlen 1976; Cowan et al. 2006; McBarnet 1981; Rock 1991). These specific theories of performativity and legitimacy are the main theories discussed in this paper. Additional, more recent literature suggests nuances to these theories. These theories lead the authors to seek more evidence on how judges perform their authority/legitimacy in courts.

Type of Study

This study was a descriptive exercise that gathered a new dataset and offered an overview of the information present in the dataset, as well as a few descriptive statistics about the dataset. The authors were able to conduct a descriptive study because they collected the dataset themselves; thus, the analysis is interesting because the dataset is novel.

Computational Methods

The paper utilized simple descriptive statistics to answer the research question. Specifically, the authors found simple percentages/ratios of judicial decisions with certain characteristics (e.g. number of sentencing decisions in which the magistrate looked at the defendant). It used a chi-squared test to identify magistrate correlation. Other than that, the paper did not utilize any other computational methods. Further, it did not discuss why they chose to use their methods or provide any additional information on its computational methods.

Results

In answer to the authors research, "Are sentencing decisions communicated in distinct ways," yes, they are. The authors found that "when communicating their decisions, magistrates generally rely on an impersonal performance of the judicial role" (1056). They found that magistrates engage with defendants less when delivering sentencing decisions that when delivering other decisions. Overall, their findings suggest that engagement with the defendant and formal elements (such as logic, rationality, and impersonality) are required to perform legal authority in court. Thus, judges do perform their legal authority in court.

Suggestions for Improvement

Critique #1: The researchers did not successfully decouple their theory from their experiment. Instead, theory and data description were both included throughout the paper, and though this made the reading interesting, it made it more difficult to understand the clear operationalization of their question. As such, the specific question the researchers were asking and how they were asking it was not completely clear until the end of the paper. I would suggest that the researchers add clearer section titles in their paper (including Methods, Data, and Theory sections) to make it easier for the reader to analyze these sections without having to dig through the entire paper.

Critique #2: The Data and Methods section of the paper could have been more clearly laid out and described. The authors did not mention their methodology in the paper at all, and though it was a simple methodology, some mention of their choice would have been useful to the reader. They did not justify their method or population of choice, which opens valid critiques about the appropriateness of this method (such as the whether the lower court population is representative of the entire court population in Australia). Further, the entire dataset was not clearly laid out anywhere in the paper. They should have provided a table with the key variables they were interested in, rather than just a graph of some of their results. Additionally, an appendix (online or at the end of the paper) with the data on all 1,287 matters would have assisted with reproducibility and verifiability and allowed researchers to ask additional questions of this data.