Problem Set 1 MACS30200

Philip Xinyu Cao

April 2018

Problem 1. *Part 1: Write a data section for your assigned data set (5 points)*

1. Describe how to access data, where it is stored, who curates it. Make sure to use the original source and curator in addition to the NBER site to which I have linked.

   (a) You can access the data by NBER website: http://nber.org/data/vital-statistics-natality-data.html, or the website of centers for disease control and prevention: https://www.cdc.gov/nchs/products/series.htm

   (b) Natality Data from National Vital Statistics System of the National Center for Health Staistics provide demographic and health data for birth for each calendar year. The micro-data is stored in the statistics offices of each State and District of Columbia.

   (c) Source: National Center for Health Statistics

2. Cite other key papers that have used this data.

   (a) Deschênes, Olivier, Michael Greenstone, and Jonathan Guryan. "Climate change and birth weight." American Economic Review 99, no. 2 (2009): 211-17.

   (b) Milligan, Kevin. "Subsidizing the stork: New evidence on tax incentives and fertility." Review of Economics and statistics 87, no. 3 (2005): 539-555.

   (c) Ruhm, Christopher J. "Are recessions good for your health?." The Quarterly journal of economics 115, no. 2 (2000): 617-650.

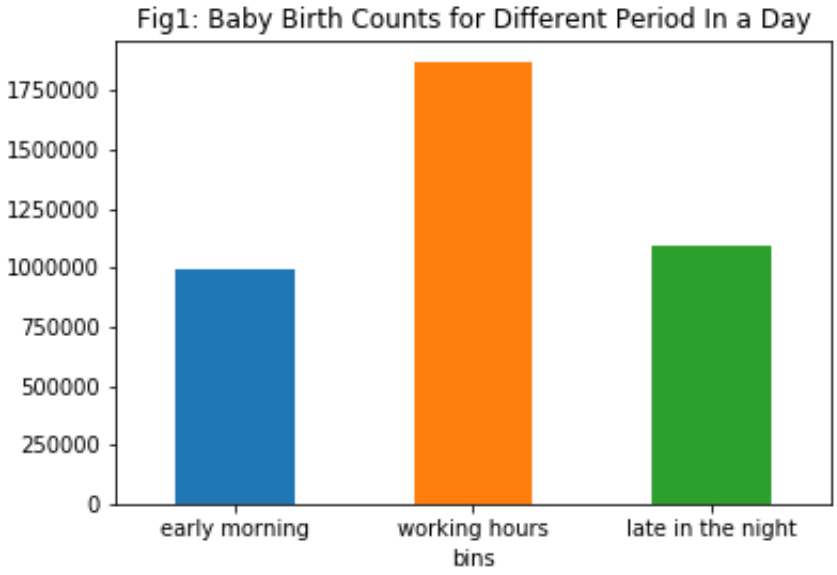
3. Describe how the data were collected.
   The vital statistics system provides the national's official vital statistics based on the collection and registration of birth and death events at the state and local level. This system contains information on all births and deaths in the united states and provides the most complete and continuous data available to public health officials at the national, state, and local levels.

4. Include a table that gives descriptive statistics for at least 8 key variables.

Table 1: Descriptive Statistics of 8 Key Variables

	dbwt	mager	meduc	fagecomb	mhtin	bmi	combgest	dobmm
count	3956112	3956112	3956112	3956112	3956112	3956112	3956112	3956112
mean	3274.6	28.7	4.4	39.6	64.4	28.8	38.7	6.6
std	634.2	5.8	1.8	22.7	4.1	13.6	3.1	3.4
min	227.0	12.0	1.0	11.0	30.0	13.0	17.0	1.0
25%	2970.0	24.0	3.0	27.0	62.0	22.1	38.0	4.0
50%	3310.0	29.0	4.0	32.0	64.0	25.5	39.0	7.0
75%	3630.0	33.0	6.0	38.0	66.0	30.8	40.0	9.0
max	9999.0	50.0	9.0	99.0	99.0	99.9	99.0	12.0

5. Show at least one conditional description of the data. This can be a table or visualization.



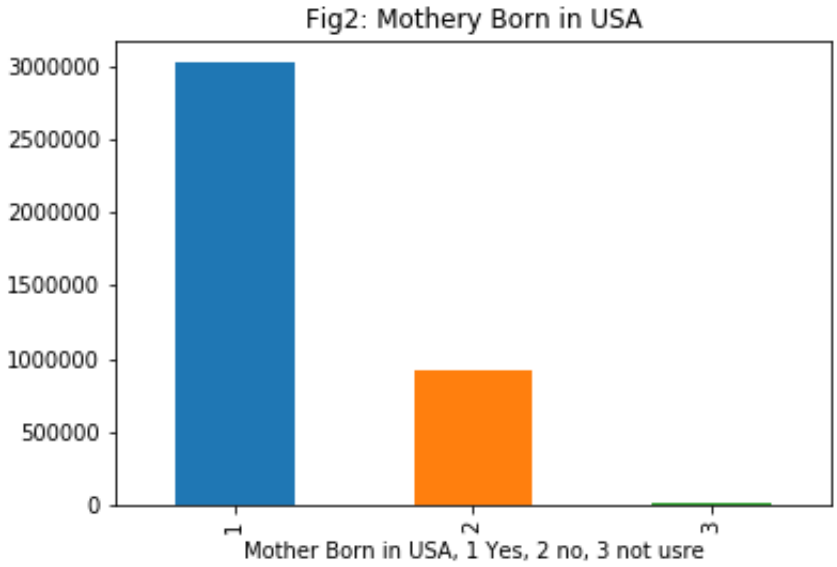


Table 2: Descriptive Statistics of 8 Key Variables of Female Infant

	dbwt	mager	meduc	fagecomb	mhtin	bmi	combgest	dobmm
count	1932663	1932663	1932663	1932663	1932663	1932663	1932663	1932663
mean	3215.6	28.7	4.4	39.6	64.4	28.8	38.7	6.6
std	616.8	5.8	1.8	22.8	4.1	13.6	3.0	3.4
min	227.0	12.0	1.0	13.0	30.0	13.0	17.0	1.0
25%	2920.0	24.0	3.0	27.0	62.0	22.1	38.0	4.0
50%	3250.0	29.0	4.0	32.0	64.0	25.5	39.0	7.0
75%	3570.0	33.0	6.0	38.0	66.0	30.8	40.0	9.0
max	9999.0	50.0	9.0	99.0	99.0	99.9	99.0	12.0

6. Note of Variable: dbwt, birth weight-details in grams; mager: mother's single years of age; meduc mother'eduction; fagecomb, father's combined age; feduc, father's eduction; mhtin mother's height in total inches; BMI, body mass index; combgest Combined Gestation – Detail in Weeks, dobmm month of born

Problem 2. *Part 2: Critique a computational research paper (5 points)*

1. The key research question is whether sentencing decisions are communicated in distinct ways?

2. The data paper use: an observational study of criminal cases in Australian lower courts of how magistrates communicate sentencing decisions in a distinct manner.

3. What theory did the paper reference in order to interpret the data?

(a) The paper trace back to Weber's theory of authority, claims to legitimacy based on 'a belief in the legality of enacted rules and right of those elevated to authority under such rules to issue commands". Goffman (1976, 1982, 1983) extend the theory to social setting constituted by professional and lay actors and driven by institutional and everyday imperatives.

(b) In court, this is a impartial adjudication as impersonal, unemotonal detachement. (Roach Anleu and Mack 2005, Mack and Roach Anleu 2010), but in lower criminal courts, the high volume of cases, time pressures, unrepresented participants and visible human emotion make this conventional performance difficult, and additional elements may be needed to accomplish legitimacy.(Mack Roach Anleu 2010, Tyler 2003, McEwan and Maiman 1986)

4. This paper is a descriptive study.

5. Just summary the descriptive statistics, basically not much computational method.

6. Give two suggestion to author as referee:

   (a) Maybe a textual analysis could useful to understand deeper for why there are distinct ways for sentencing decisions.

   (b) Maybe run a simple regression could generate more thoughts than a series of descriptive statistics, we could see how different variable cofounder the results.