

MACS 30200

Problem Set #1 Part 1

A Data Section for U.S. Natality Data (2016)

Xi Chen

The U.S. Natality Data, which is curated by the National Vital Statistics System of the National Center for Health Statistics (NCHS), is available in the website of National Bureau of Economic Research (NBER)¹. The Natality Data, binned by year from 1968 to 2016, are based on information abstracted from birth certificates filed in vital statistics offices of each State and District of Columbia.

The Natality Data have been intensively used in the past research. Almond & Edlund (2007) used the Natality Data (1983-2001) to explore the potential role of offspring sex ratio to white mothers. Heck, et al. (1997) used the Natality Data (1969-1994) to examine delayed childbearing by material education level. Tucker, et al., (2007) used the data to explore the black-white disparity in pregnancy-related mortality. Mendoza, et al., (2013) used the Natality Data (2007-2010) to conduct pediatrics research related to early intervention programs enrollment. Schoendorf & Branum (2006) used the Natality Data to conduct perinatal and obstetric research.

This assignment only uses the Natality Data (2016) for the following analysis. The Natality Data (2016), which provide demographical and health information for the births occurring in the year of 2016, have 3956112 observations and 240 variables. Table 1 presents descriptive statistics for several demographical variables and health information, including the birth month, the birth day of week, mother's age, mother's race, mother's education, father's age, father's race, father's education, marital status, and the reported number of cigarette per day.

¹ <http://nber.org/data/vital-statistics-natality-data.html>

Table 1. Descriptive Statistics

	Birth Month	Birth Day of Week	Mother's Age	Mother's Race	Mother's Education
mean	6.562	4.062	4.343	15.239	4.368
std	3.413	1.849	1.199	11.144	1.807
min	1.000	1.000	1.000	10.000	1.000
25%	4.000	3.000	3.000	10.000	3.000
50%	7.000	4.000	4.000	10.000	4.000
75%	9.000	6.000	5.000	20.000	6.000
max	12.000	7.000	9.000	61.000	9.000

	Marital Status	Father's Age	Father's Race	Father's Education	Cigarettes Before Pregnancy
mean	1.398	5.635	2.853	4.896	0.236
std	0.489	2.358	3.05	2.326	0.808
min	1.000	1.000	1.000	1.000	0.000
25%	1.000	4.000	1.000	3.000	0.000
50%	1.000	5.000	1.000	4.000	0.000
75%	2.000	6.000	4.000	6.000	0.000
max	2.000	11.000	9.000	9.000	6.000

Figure 1. Maternal Educational level vs. Prenatal Care Begin Month

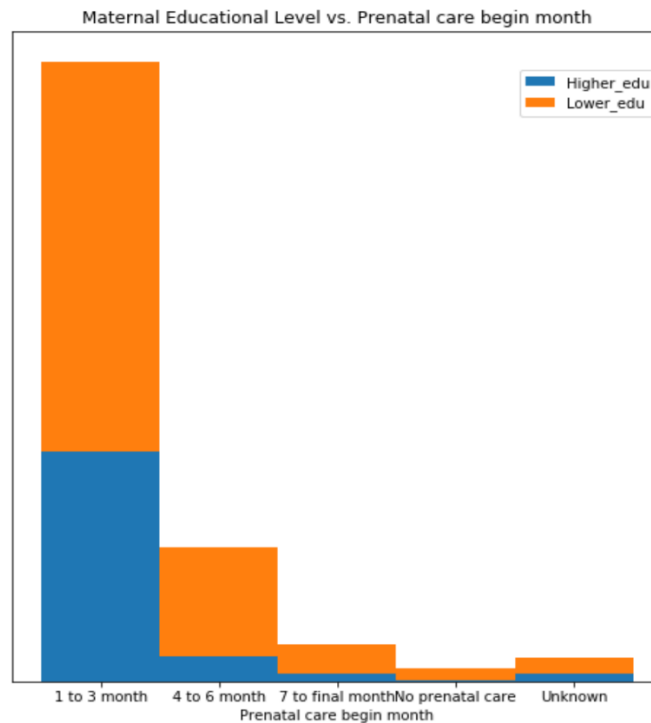


Figure 1 shows the relationship between mother's educational level and the prenatal care begin month. If the mother's education level is at least Bachelor's degree, the observation is categorized into the higher education group, otherwise it is categorized into the lower education group. Compared to the lower maternal education group, most of individuals in the higher maternal education group begin to have prenatal cares at an early date, such as 1st to 3rd month. Figure 2 shows the relationship between mother's race and the prenatal care begin month. Most of mothers in all racial groups tend to begin the prenatal care from the 1st to 3rd month.

Figure 2. Mother's Race vs. Prenatal Care Begin Month

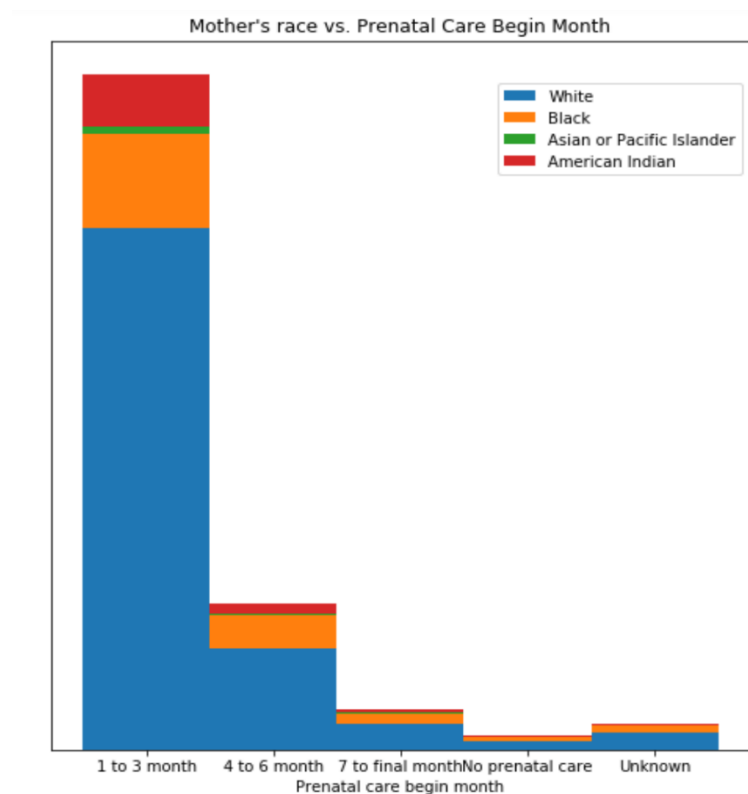


Table 2 shows the descriptive statistics by maternal education level. Similiary to Figure 1, observations are categorized into either higher maternal education group or lower maternal educational group. By comparing the two sub-tables, the average mother's age in the higher education group is greater than that in the lower education group; the higher education group also has much higher father's education level on average; the higher education group has more prenatal visits as well; but the lower educational group has higher BMI than the higher educational group.

Table 2. Descriptive Statistics by Maternal Educational Level

i. Higher Maternal Education Group

	Mother's Age	Father's Education	Number of Prenatal Visits	Body Mass Index (BMI)
mean	4.974	5.859	7.246	2.849
std	0.931	1.617	1.864	1.468
min	1.000	1.000	1.000	1.000
25%	4.000	5.000	6.000	2.000
50%	5.000	6.000	7.000	2.000
75%	6.000	7.000	8.000	3.000
max	9.000	9.000	12.000	9.000

ii. Lower Maternal Education Group

	Mother's Age	Father's Education	Number of Prenatal Visits	Body Mass Index (BMI)
mean	4.033	4.423	6.786	3.197
std	1.194	2.471	2.153	1.571
min	1.000	1.000	1.000	1.000
25%	3.000	3.000	6.000	2.000
50%	4.000	4.000	7.000	3.000
75%	5.000	6.000	8.000	4.000
max	9.000	9.000	12.000	9.000

Reference

- Almond, D., & Edlund, L. (2007). Trivers–Willard at birth and one year: evidence from US natality data 1983–2001. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1624), 2491-2496.
- Heck, K. E., Schoendorf, K. C., Ventura, S. J., & Kiely, J. L. (1997). Delayed childbearing by education level in the United States, 1969–1994. *Maternal and child health journal*, 1(2), 81-88.
- Tucker, M. J., Berg, C. J., Callaghan, W. M., & Hsia, J. (2007). The Black–White disparity in pregnancy-related mortality from 5 conditions: differences in prevalence and case-fatality rates. *American journal of public health*, 97(2), 247-251.
- Shapiro-Mendoza, C., Kotelchuck, M., Barfield, W., Davin, C. A., Diop, H., Silver, M., & Manning, S. E. (2013). Enrollment in early intervention programs among infants born late preterm, early term, and term. *Pediatrics*, 132(1), e61-e69.
- Schoendorf, K. C., & Branum, A. M. (2006). The use of United States vital statistics in perinatal and obstetric research. *American Journal of Obstetrics & Gynecology*, 194(4), 911-915.