

**Is Your Baby on Track?**  
**An Exploration on Potentially-defining Features for Infant Birth Weight**

**Summary of the Problem**

Decades worth of scientific research has concluded that birth weight is indicative of a baby's future health condition. In fact, a research published in *Proceedings of the National Academy of Sciences* or PNAS suggested how it could affect future brain development (Goisis et al, 2017). If birth weight is, indeed, evocative of such consequences, does a big baby automatically correspond to a healthy baby? If so, which prenatal factors are relevant enough, or more importantly, controllable by parents?

The birthweight dataset presents us with 196 observations along with 18 variables describing parental demographics, apgar scores, prenatal lifestyle, prenatal care, and actual infant birth weight. The team's analysis will give light on whether or not these factors (or a combination of some) do affect birth weight. Knowing such influences could help direct attention to infants who may have special medical problems that need immediate attention. This analysis will explore variable correlations and their effect on future birth weights based on predictive models. This is done with the goal of providing parents greater assurance and an opportunity for supervised care because a "healthy" baby is more than just his or her dimpled thighs or chubby cheeks.

**Assumptions**

The group worked on the general assumption that this particular birthweight dataset was conducted among US residents only. As for the year, our initial observation on the *npvis* (number of prenatal visits) feature was that mothers have only attended prenatal visits during

the first trimester of pregnancy leading them to further look into insurance policies and its progression through the years. Research showed us that the ObamaCare Act required all insurance companies to cover full maternity packages and this was only implemented after 2014 which could situate our data timeline before that. However, this still does not guarantee that all insurance companies strictly follow the health care law.

The group also acknowledges possible falsification of information regarding cigarette smoking and alcohol intake and that these pieces of information are taken only from the mother. Our other initial assumptions are based on external scientific research such as the increased medical risk of pregnancy beyond age 32, the harmful effects of cigarettes and alcohol during pregnancy, the 7-9 normal *one-minute* and *five-minute apgar score* range, and the 2500g-4000g range for normal birth weight.

### Exploration and Modeling Process

In an attempt to determine significant variables from the birthweight dataset, proper exploratory data analysis was conducted. This involved imputing missing values, flagging outliers, and analyzing correlations. After this, a regression analysis was done in order to build a best performing model that would predict birth weight given any unknown dataset. This model was determined by exploring results using different regression model types and finding a specific mix of variables that would provide satisfactory prediction accuracy using R-squared as the main metric.

### Actionable Key Insights

#### 1. *Parental Race/Ethnicity*

Our predictive model reflected high positive coefficients for both parents' ethnicities (white, black, and other), more specifically for mothers who are of African-American descent. It tells how the race of the parent/s could be a big predictor of a

newborn's birth weight. Scientific studies have backed up the possibility of ethnicity as an influencer. According to Cyril Ferdynus, a biostatistician in Reunion, France, it is well documented in Europe and Canada that ethnicity plays a significant role in fetal growth during pregnancy and that weight references "must be developed and adapted to local populations" (Doyle, 2013). Considerably, the country is highly multicultural which implies that there is a huge immigrant population present, but despite progressive practices, there is still a significant number of African-American female immigrants who may not understand the importance of proper prenatal care (Keung, 2012). This factor alone, however, is not absolute. Other variables, even those beyond the given dataset, when combined with ethnicity, could pose a potentially greater effect on birth weight.

## *2. Father's Age*

Our results indicate a high negative coefficient for the outliers above the upper fence of the father's age. This factor is often overlooked when evaluating birth risks but there have been evidences that suggest the presence of adverse effects to the baby's overall health due to advanced paternal age. According to a study conducted by researchers at Stanford School of Medicine, babies born to fathers above the age of 35 were more prone to low birth weight and other complications. Doctor Michael Eisenberg, associate professor of Urology, further explains that every year that a man ages, he accumulates on average two new mutations in the DNA of his sperm, which in turn increases the risk during pregnancy (Armitage, 2018). Although effects associated with paternal age are less certain than those of maternal age, it is still important to consider as it may indirectly affect the mother's health as well.

### 3. *Start of Prenatal Care*

Another significant category that reflected a high negative coefficient would be the outliers for the month prenatal care began, which ranges from the 5th to 8th month of pregnancy. This suggests that as women start their prenatal care beyond the first trimester of their pregnancy, the more it negatively affects the newborn's birth weight. Prenatal care aims to prevent potential health problems for both the fetus and the mother through regular check-ups, lifestyle consultation, and medical intervention. By identifying and addressing health concerns early on in the pregnancy, we are reducing the risk of any complications in the future. The effect of prenatal care timing on birth weight, however, is not as straightforward as it seems. It may either be underestimated or overestimated depending on certain underlying factors such as socioeconomic status, cultural beliefs, quality and availability of prenatal care, delayed recognition of pregnancy, and previous experience with pregnancy, to name a few.

### Implementation Recommendations

No parent wants their child to be born unhealthy. It is vital not to create any unnecessary worry and overreaction which could lead to misguided prognosis and treatment. With the group's data exploration and modeling analysis, they have determined which controllable features have the highest potential to positively or negatively affect a newborn's birth weight. The group recommends hospitals to include the features we have deemed significant (maternal ethnicity, paternal age, and the month prenatal care was started) in their birth weight charts. Recognized birth weight charts are already implemented in hospitals in some countries, however these are only used for research, and not yet for actual clinical care, in the United States. Modifications and further feature exploration are required in order to legitimize and confirm these charts before nationwide implementation. The team's model contributes an affirmation into the significant impact of parental ethnicity, paternal age, and commencement of

prenatal care, adding value to existing prenatal practices and protocols, thereby decreasing the risk of delivering unhealthy infants.

## References

Chung, E. (2014, July 1). Child's birth weight depends more on dad's ethnicity. Retrieved from <https://www.cbc.ca/news/health/mixed-race-babies-with-asian-dads-are-born-smaller-1.2693364>

Doyle, K. (2013, August 2). Healthy newborn size varies with ethnicity: study. Retrieved from <https://www.reuters.com/article/us-healthy-newborn/healthy-newborn-size-varies-with-ethnicity-study-idUSBRE9710TC20130802>

Goisis, A., Özcan, B., & Myrskylä, M. (2017, January 3). Decline in the negative association between low birth weight and cognitive ability. Retrieved from <https://www.pnas.org/content/pnas/114/1/84.full.pdf>

Keung, N. (2012, February 15). New birth weight curves tailored to baby's ethnicity. Retrieved from [https://www.thestar.com/life/parent/2012/02/15/new\\_birth\\_weight\\_curves\\_tailored\\_to\\_babys\\_ethnicity.html](https://www.thestar.com/life/parent/2012/02/15/new_birth_weight_curves_tailored_to_babys_ethnicity.html)

Armitage, H. (2018, October 31). Older fathers associated with increased birth risks. Retrieved from <http://med.stanford.edu/news/all-news/2018/10/older-fathers-associated-with-increased-birth-risks.html>