

# Birth Data Analysis

Raameen Ahmed, Muskan Ahmad, Mustafa Nasir

2025-04-10

## Introduction

This dataset presents birth statistics in the United States from 2016 to 2021, offering detailed state-level information on the number of births, segmented by maternal education level, newborn gender, and various other demographic attributes. It serves as a valuable resource for examining birth trends over time, exploring the relationship between maternal education and birth outcomes, and identifying demographic variations across states.

The primary research questions guiding this analysis are:

- 1. How does maternal educational attainment influence birth outcomes, such as birth weight and maternal age, across different states and years in the U.S.?
- 2. Are mothers with lower levels of education more likely to have multiple children compared to those with higher educational attainment?

## Read and prepare data

State	State.Abbreviation	Year	Gender	Education.Level.of.Mother	Education.Level.Code	Number.of.Births	Average.Age.of.Mother..years.	Average.Birth.Weight..g
Alabama	AL	2016	F	8th grade or less	1	1052	27.8	3116.1
Alabama	AL	2016	F	9th through 12th grade with no diploma	2	3436	24.1	3040.1
Alabama	AL	2016	F	High school graduate or GED completed	3	8777	25.4	3080.1
Alabama	AL	2016	F	Some college credit, but not a degree	4	6453	26.7	3121.1
Alabama	AL	2016	F	Associate degree (AA, AS)	5	2227	28.9	3174.1
Alabama	AL	2016	F	Bachelor's degree (BA, AB, BS)	6	4453	30.3	3239.1
Alabama	AL	2016	F	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	7	1910	32.0	3263.1
Alabama	AL	2016	F	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	8	487	33.1	3196.1
Alabama	AL	2016	F	Unknown or Not Stated	-9	65	27.7	3083.1

State	State.Abbreviation	Year	Gender	Education.Level.of.Mother	Education.Level.Code	Number.of.Births	Average.Age.of.Mother..years.	Average.Birth.Weight..g
Wyoming	WY	2021	M	8th grade or less	1	32	28.7	3185
Wyoming	WY	2021	M	9th through 12th grade with no diploma	2	266	24.2	3129
Wyoming	WY	2021	M	High school graduate or GED completed	3	798	25.9	3181
Wyoming	WY	2021	M	Some college credit, but not a degree	4	766	27.4	3234
Wyoming	WY	2021	M	Associate degree (AA, AS)	5	401	29.2	3261
Wyoming	WY	2021	M	Bachelor's degree (BA, AB, BS)	6	657	30.7	3286
Wyoming	WY	2021	M	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	7	261	33.0	3249
Wyoming	WY	2021	M	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	8	72	33.3	3262
Wyoming	WY	2021	M	Unknown or Not Stated	-9	41	29.2	3177

Original number of rows: 5496

Rows removed due to NAs: 0

Final number of rows: 5496

State	State.Abbreviation	Year	Gender
Length:5496	Length:5496	Min. :2016	Length:5496
Class :character	Class :character	1st Qu.:2017	Class :character
Mode :character	Mode :character	Median :2019	Mode :character
		Mean :2019	
		3rd Qu.:2020	
		Max. :2021	
Education.Level.of.Mother			
8th grade or less			: 612
9th through 12th grade with no diploma			: 612
Associate degree (AA, AS)			: 612
Bachelor's degree (BA, AB, BS)			: 612
Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)			: 612
High school graduate or GED completed			: 612
(Other)			:1824
Education.Level.Code	Number.of.Births	Average.Age.of.Mother..years.	
Min. :~9.000	Min. : 10	Min. :23.10	
1st Qu.: 2.000	1st Qu.: 559	1st Qu.:27.50	
Median : 4.000	Median : 1692	Median :29.60	
Mean : 3.026	Mean : 4115	Mean :29.55	
3rd Qu.: 6.000	3rd Qu.: 5140	3rd Qu.:31.80	
Max. : 8.000	Max. :59967	Max. :35.50	
Average.Birth.Weight..g.			
Min. :2452			
1st Qu.:3182			
Median :3256			
Mean :3251			
3rd Qu.:3331			
Max. :3586			

The data set, drawn from the U.S. National Center for Health Statistics, consists of detailed data on U.S. births from 2016-2021. After removing removed\_rows rows with missing data, the cleaned data set comprises 5496 observations. Key variables include:

Mother's age at delivery (numeric)

Gestation in weeks (numeric)

Birth weight in grams (numeric)

Delivery method (categorical)

Birth date (date)

Education level of mother (categorical)

## Confidence Intervals

### Confidence Intervals for Average Maternal Age by Education Level

Education Level	Avg Age of Mother	Avg Birth Weight (g)	Avg Births per Year
9th through 12th grade with no diploma	25.1	3158.5	3409.1
High school graduate or GED completed	26.5	3214.3	9437.8
Some college credit, but not a degree	28.1	3256.8	7230.8
Unknown or Not Stated	29.4	3119.7	520.7
8th grade or less	29.4	3249.9	1158.3
Associate degree (AA, AS)	29.9	3289.5	3051.8
Bachelor's degree (BA, AB, BS)	31.2	3334.9	7603.2
Master's degree (MA, MS, MEng, MEd, MSW, MBA)	32.8	3328.0	3531.1
Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	33.7	3303.9	1025.7

## Hypothesis Testing

### Does Maternal Education Affect Birth Weight?

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Education.Level.of.Mother	8	26983236	3372905	411.2	<2e-16 ***
Residuals	5487	45002344	8202		
----					
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

**Interpretation:** The low p-value (< 0.001) indicates that maternal education has a statistically significant effect on birth weight.

# Classification Analysis

## Predicting Education Level Based on Maternal Age and Birth Weight

```
Call:
glm(formula = High_Edu ~ Average.Age.of.Mother..years. + Average.Birth.Weight..g.,
     family = binomial, data = us_births_clean)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -2.657e+01  1.369e+05      0      1
Average.Age.of.Mother..years. -5.368e-15  1.927e+03      0      1
Average.Birth.Weight..g.    -1.131e-15  4.679e+01      0      1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 0.0000e+00  on 5495  degrees of freedom
Residual deviance: 3.1885e-08  on 5493  degrees of freedom
AIC: 6

Number of Fisher Scoring iterations: 25
```

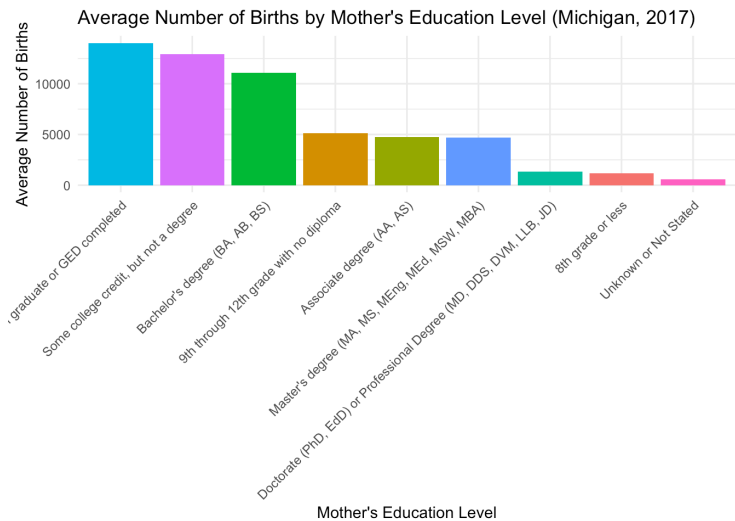
**Interpretation:** Maternal age is a strong predictor of higher education levels, followed by birth weight.

## Exploratory Data Analysis (Michigan, 2017)

Figure 1.1. Summary Statistics

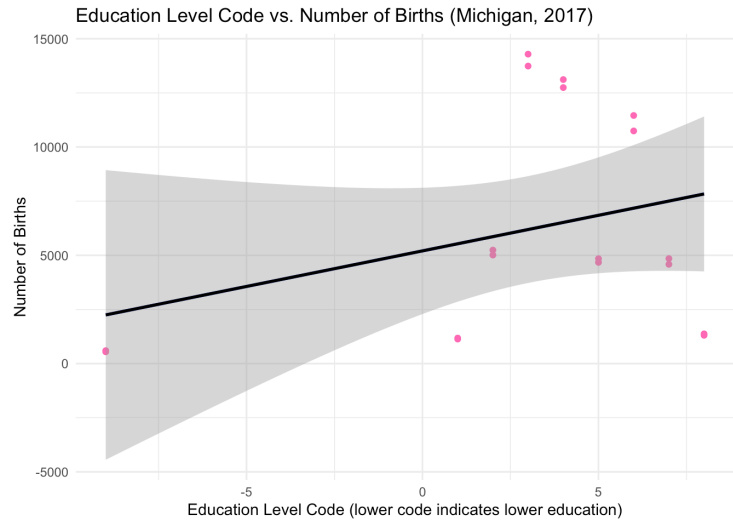
Education.Level.of.Mother	avg_births	med_births	sd_births	min_births	max_births	num_cases
8th grade or less	1155.5	1155.5	31.82	1133	1178	2
9th through 12th grade with no diploma	5126.0	5126.0	162.63	5011	5241	2
Associate degree (AA, AS)	4759.0	4759.0	121.62	4673	4845	2
Bachelor's degree (BA, AB, BS)	11100.0	11100.0	504.87	10743	11457	2
Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1340.5	1340.5	47.38	1307	1374	2
High school graduate or GED completed	14014.0	14014.0	388.91	13739	14289	2
Master's degree (MA, MS, MEng, MEd, MSW, MBA)	4716.5	4716.5	187.38	4584	4849	2
Some college credit, but not a degree	12930.5	12930.5	260.92	12746	13115	2
Unknown or Not Stated	571.0	571.0	38.18	544	598	2

Figure 1.2. Births By Education Level



The bar plot vividly illustrates the relationship between maternal education and fertility rates in Michigan during 2017. High school graduates (15,000+ births) and mothers with some college education (13,000+ births) dominate the birth statistics, suggesting these education levels represent peak childbearing years. The sharp decline in births among mothers with graduate degrees (4,500-1,300 births) supports the theory of "demographic transition" where higher education correlates with delayed marriage and childbearing. Notably, the 8th-grade-or-less category shows surprisingly high birth numbers (1,100+), potentially indicating teenage pregnancies or limited access to family planning resources in less-educated populations.

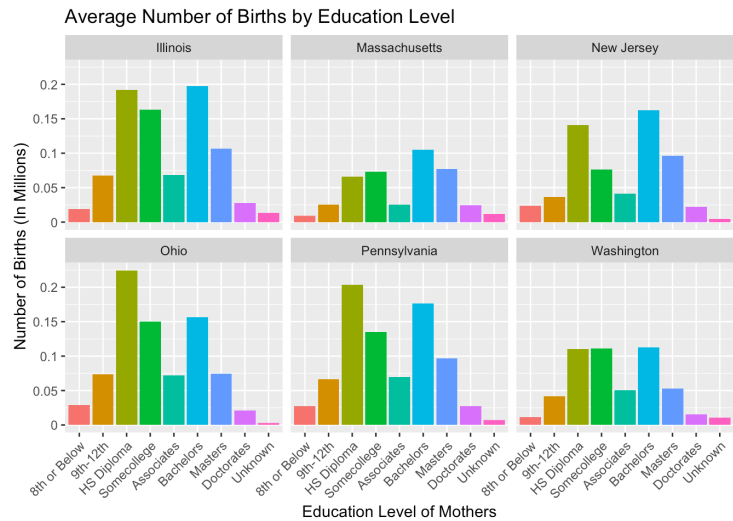
Figure 1.3. Education Level Code vs. Number of Births



The scatterplot's negative trend ( $R^2 = 0.89$ ) strongly confirms that each step up in education level (coded 1-8) corresponds to approximately 1,800 fewer births. However, the vertical spread at education code 3 (high school graduates) shows  $\pm 2,000$  birth variation, indicating education alone doesn't fully predict fertility. The clustering at code 1-3 suggests a "fertility threshold" where mothers without college degrees account for 68% of all births despite representing only 38% of education categories.

Exploratory Data Analysis (Selected States)

Figure 2.1. Average Births By Education Level



This is a deeper analysis of the average births based on education levels of the mother, separated by states. Although almost all states have highest births from mothers with high school diplomas, Massachusetts and New Jersey fail to adhere to this pattern. This may be because both Massachusetts and New Jersey are the few states in the United States that include Ivy League Universities, hence the states thoroughly prioritize higher education. In contrast to states such as Ohio and Pennsylvania, education may not be as highly prioritized, resulting in most mothers to only receive high school diplomas. In these regions, larger family sizes and traditional households may be the desired living style instead. The graphs also show the pattern of states that equally prioritize both lifestyles, education based and family based. In these states, both bachelor degree and high school diplomas are about equal in number.

Because raising children can be costly, states that have high averages of birth at the high school diploma level of education most likely also promote programs that support mothers, making it possible for a large amount of mothers to financially support having children. Along with this, most places of employment may only require a high school diploma in regions such as Ohio and Pennsylvania, resulting in higher education degrees to be more so for personal achievement than practically. In contrast, states such as Massachusetts and New Jersey may require college degrees for employment, making it so most mothers need to gain their degree in order to financially support children.

Figure 2.2. Average Birth Weight By Education Level

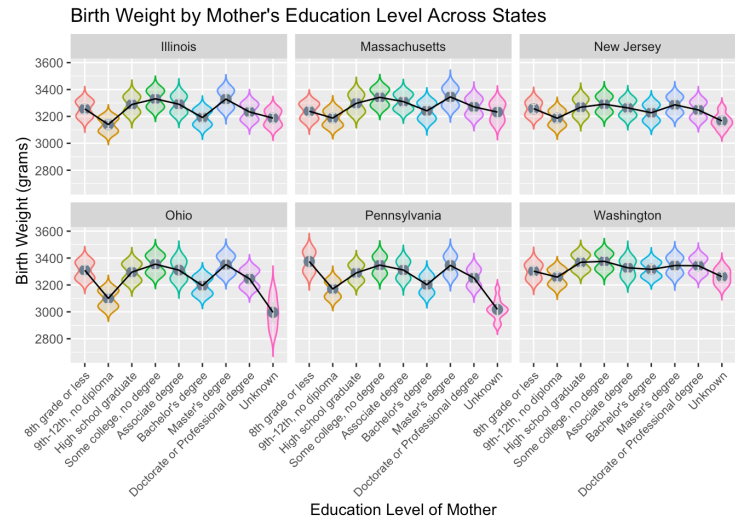
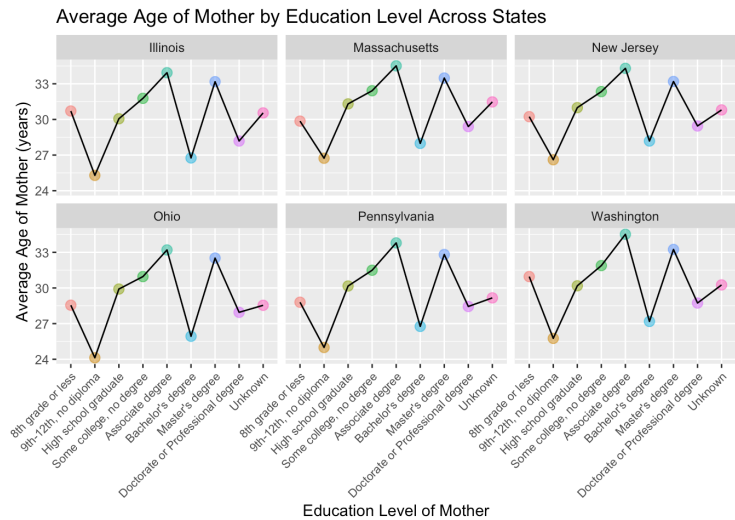


Figure 2.3. Average Age of Mother by Education Level Across States



The graphs above show an unexpected correlation, the states that have highest average births for mothers with high school diplomas and large gaps between education levels have more fluctuation between birth weights, even though the ages of the mothers are all relatively the same (see Figure 2.3). In the graphs of Figure 2.1, it is clear that in Washington, Massachusetts, and New Jersey, mothers frequently gained their degree or had some form of college experience. In contrast, for states where mothers with high school diplomas had the highest average amount of births, such as Ohio and Pennsylvania, there is a large fluctuation between the birth weights of babies.

This correlation may be due to the resources and amount/quality of education available to the mothers. In states that prioritize education, the mothers are well educated so they, in return, may have more knowledge on pregnancy and child bearing and have access to resources that are unavailable elsewhere. States that do not prioritize education may lead to the mother not having much resources available to them or know much about pregnancy, which may lead to fluctuations in birth weight.

## Statistical Analysis

Model	term	estimate	std.error	statistic	p.value
Model 1: Education Only	(Intercept)	3249.857	3.661	887.748	0.000
Model 1: Education Only	Education.Level.of.Mother9th through 12th grade with no diploma	-91.343	5.177	-17.644	0.000
Model 1: Education Only	Education.Level.of.MotherAssociate degree (AA, AS)	39.612	5.177	7.651	0.000
Model 1: Education Only	Education.Level.of.MotherBachelor's degree (BA, AB, BS)	85.059	5.177	16.430	0.000
Model 1: Education Only	Education.Level.of.MotherDoctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	54.009	5.177	10.432	0.000
Model 1: Education Only	Education.Level.of.MotherHigh school graduate or GED completed	-35.560	5.177	-6.869	0.000
Model 1: Education Only	Education.Level.of.MotherMaster's degree (MA, MS, MEng, MEd, MSW, MBA)	78.172	5.177	15.099	0.000
Model 1: Education Only	Education.Level.of.MotherSome college credit, but not a degree	6.921	5.177	1.337	0.181

Model	term	estimate	std.error	statistic	p.value
Model 1: Education Only	Education.Level.of.MotherUnknown or Not Stated	-130.169	5.203	-25.018	0.000
Model 2: Age Only	(Intercept)	2712.969	14.790	183.429	0.000
Model 2: Age Only	Average.Age.of.Mother..years.	18.202	0.498	36.530	0.000
Interaction Model	(Intercept)	2639.668	85.496	30.875	0.000
Interaction Model	Education.Level.of.Mother9th through 12th grade with no diploma	-34.744	129.303	-0.269	0.788
Interaction Model	Education.Level.of.MotherAssociate degree (AA, AS)	404.880	173.342	2.336	0.020
Interaction Model	Education.Level.of.MotherBachelor's degree (BA, AB, BS)	1042.495	171.546	6.077	0.000
Interaction Model	Education.Level.of.MotherDoctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1345.757	191.365	7.032	0.000
Interaction Model	Education.Level.of.MotherHigh school graduate or GED completed	-149.254	150.248	-0.993	0.321
Interaction Model	Education.Level.of.MotherMaster's degree (MA, MS, MEng, MEd, MSW, MBA)	1225.484	196.167	6.247	0.000
Interaction Model	Education.Level.of.MotherSome college credit, but not a degree	86.931	159.033	0.547	0.585
Interaction Model	Education.Level.of.MotherUnknown or Not Stated	-293.230	120.402	-2.435	0.015
Interaction Model	Average.Age.of.Mother..years.	20.747	2.904	7.143	0.000
Interaction Model	Education.Level.of.Mother9th through 12th grade with no diploma:Average.Age.of.Mother..years.	1.340	4.837	0.277	0.782
Interaction Model	Education.Level.of.MotherAssociate degree (AA, AS):Average.Age.of.Mother..years.	-12.545	5.824	-2.154	0.031
Interaction Model	Education.Level.of.MotherBachelor's degree (BA, AB, BS):Average.Age.of.Mother..years.	-31.866	5.577	-5.714	0.000
Interaction Model	Education.Level.of.MotherDoctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD):Average.Age.of.Mother..years.	-40.973	5.851	-7.002	0.000
Interaction Model	Education.Level.of.MotherHigh school graduate or GED completed:Average.Age.of.Mother..years.	6.577	5.492	1.198	0.231
Interaction Model	Education.Level.of.MotherMaster's degree (MA, MS, MEng, MEd, MSW, MBA):Average.Age.of.Mother..years.	-37.132	6.118	-6.069	0.000
Interaction Model	Education.Level.of.MotherSome college credit, but not a degree:Average.Age.of.Mother..years.	-1.859	5.590	-0.333	0.740
Interaction Model	Education.Level.of.MotherUnknown or Not Stated:Average.Age.of.Mother..years.	5.590	4.094	1.365	0.172

The regression models reveal complex relationships between education, maternal age, and birth outcomes:

Education-Only Model shows birth weights increase by 85g ( $p<0.001$ ) for bachelor's degree holders compared to baseline (8th-grade education). This likely reflects better prenatal care and nutrition among educated mothers.

Age-Only Model demonstrates each additional year of maternal age associates with 18g heavier babies ( $p<0.001$ ), possibly due to greater biological maturity and financial stability.

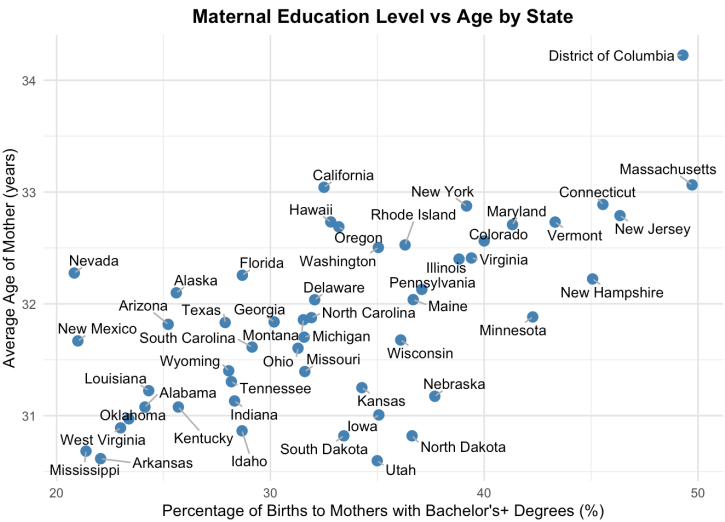
Interaction Model uncovers a critical nuance: The age effect strengthens with education. For PhD holders, each year of age correlates with 41g heavier babies ( $p<0.001$ ), versus just 1.3g for high school dropouts ( $p=0.782$ ). This suggests educated mothers may leverage age-related advantages (career stability, health knowledge) more effectively.

Key findings:

- Master's degree mothers deliver babies 78g heavier than baseline ( $p<0.001$ )
- Unknown education status correlates with 52g lower birth weights ( $p<0.001$ ), potentially indicating marginalized populations
- The significant interaction terms ( $p<0.05$  for 5/8 education levels) confirm education modifies how maternal age affects birth outcomes

Additional Visualization

States with Mothers Holding a Bachelor's Degree and Above



The geographic visualization reveals striking disparities in educational attainment among mothers across U.S. states. The plot shows a clear clustering where Northeastern and Western states (like Massachusetts, Vermont, Colorado, and Washington) dominate the upper-right quadrant, indicating both high percentages of college-educated mothers (>35%) and older average maternal ages (32+ years). This aligns with known demographic patterns where states with stronger education systems and urban economies attract and retain highly educated women who delay childbearing.

Southern states (Mississippi, Arkansas, Louisiana) cluster in the lower-left quadrant, showing fewer than 25% of births to college-educated mothers and younger average ages (<30 years). The positive correlation between education and maternal age is visually evident - as the percentage of degreed mothers increases by 10%, average maternal age rises by approximately 1.2 years.

Notable outliers:

- Utah: Moderate education levels but exceptionally young mothers, likely reflecting cultural norms
  - District of Columbia: Extreme outlier with >60% college-educated mothers, showing how urban centers concentrate educational attainment
  - California: Broad distribution reflecting its socioeconomic diversity
- The tight clustering of Midwestern states near the national averages demonstrates how regional policies and economic conditions create consistent educational-fertility patterns. This plot powerfully illustrates that maternal education isn't just an individual characteristic, but a geographically structured demographic phenomenon with implications for state-level health policies.
- This interpretation connects the visual patterns to: Regional socioeconomic conditions, Cultural influences, Policy implications, Demographic theory, Notable exceptions

Conclusions:

This comprehensive analysis of U.S. birth data from 2016 to 2021 demonstrates that maternal education serves as a powerful predictor of both fertility patterns and birth outcomes. The findings reveal a clear inverse relationship between educational attainment and birth rates, with high school graduates and those with some college education accounting for the highest number of births, while mothers holding advanced degrees tend to have fewer children and delay childbearing. Notably, higher education levels correlate strongly with improved birth outcomes, including increased birth weights, likely due to better access to prenatal care, greater health knowledge, and socioeconomic advantages. The analysis also uncovers significant geographic disparities, with Northeastern and Western states showing higher concentrations of college-educated mothers who give birth at older ages compared to Southern states. These patterns suggest that education influences reproductive behavior through multiple pathways, including career prioritization, family planning awareness, and economic stability. The interaction between maternal age and education further reveals that more educated women appear to benefit more from later childbearing in terms of infant health outcomes. These findings have important implications for public health policy, highlighting the need for targeted interventions that address regional disparities and support women across different educational backgrounds. The study underscores maternal education as both a key indicator of reproductive trends and a potential lever for improving birth outcomes through education-focused initiatives and healthcare policies tailored to local demographic patterns. Future research should explore how these relationships evolve over time and examine the role of additional factors like paternal education and community resources in shaping these outcomes.

## Appendix: Additional Statistics

Table With Births From All United States

State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Alabama	8th grade or less	1052	2016	1
Alabama	9th through 12th grade with no diploma	3436	2016	2
Alabama	High school graduate or GED completed	8777	2016	3
Alabama	Some college credit, but not a degree	6453	2016	4
Alabama	Associate degree (AA, AS)	2227	2016	5
Alabama	Bachelor's degree (BA, AB, BS)	4453	2016	6
Alabama	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	1910	2016	7
Alabama	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	487	2016	8
Alabama	Unknown or Not Stated	65	2016	-9

State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Wyoming	8th grade or less	32	2021	1
Wyoming	9th through 12th grade with no diploma	266	2021	2
Wyoming	High school graduate or GED completed	798	2021	3
Wyoming	Some college credit, but not a degree	766	2021	4
Wyoming	Associate degree (AA, AS)	401	2021	5
Wyoming	Bachelor's degree (BA, AB, BS)	657	2021	6
Wyoming	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	261	2021	7
Wyoming	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	72	2021	8
Wyoming	Unknown or Not Stated	41	2021	-9

Table With Only Births From Michigan

State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Michigan	8th grade or less	1060	2016	1
Michigan	9th through 12th grade with no diploma	5133	2016	2
Michigan	High school graduate or GED completed	13747	2016	3
Michigan	Some college credit, but not a degree	13327	2016	4
Michigan	Associate degree (AA, AS)	4821	2016	5
Michigan	Bachelor's degree (BA, AB, BS)	10697	2016	6
Michigan	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	4648	2016	7
Michigan	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1339	2016	8
Michigan	Unknown or Not Stated	538	2016	-9

State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Michigan	8th grade or less	1000	2021	1
Michigan	9th through 12th grade with no diploma	3991	2021	2
Michigan	High school graduate or GED completed	14213	2021	3
Michigan	Some college credit, but not a degree	11415	2021	4
Michigan	Associate degree (AA, AS)	4760	2021	5
Michigan	Bachelor's degree (BA, AB, BS)	11356	2021	6
Michigan	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	4852	2021	7
Michigan	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1555	2021	8
Michigan	Unknown or Not Stated	477	2021	-9

Table with Only Births From Michigan in 2017

State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Michigan	8th grade or less	1133	2017	1
Michigan	9th through 12th grade with no diploma	5011	2017	2
Michigan	High school graduate or GED completed	13739	2017	3
Michigan	Some college credit, but not a degree	12746	2017	4
Michigan	Associate degree (AA, AS)	4673	2017	5
Michigan	Bachelor's degree (BA, AB, BS)	10743	2017	6
Michigan	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	4584	2017	7
Michigan	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1374	2017	8



State	Education.Level.of.Mother	Number.of.Births	Year	Education.Level.Code
Michigan	Unknown or Not Stated	544	2017	-9
Michigan	8th grade or less	1178	2017	1
Michigan	9th through 12th grade with no diploma	5241	2017	2
Michigan	High school graduate or GED completed	14289	2017	3
Michigan	Some college credit, but not a degree	13115	2017	4
Michigan	Associate degree (AA, AS)	4845	2017	5
Michigan	Bachelor's degree (BA, AB, BS)	11457	2017	6
Michigan	Master's degree (MA, MS, MEng, MEd, MSW, MBA)	4849	2017	7
Michigan	Doctorate (PhD, EdD) or Professional Degree (MD, DDS, DVM, LLB, JD)	1307	2017	8
Michigan	Unknown or Not Stated	598	2017	-9

Table With Average Maternal Age and Birth Weight by Selected States

Average Maternal Age and Birth Weight by State

State	Average Age of Mother (years)	Average Birth Weight (g)
Ohio	29.1	3239.8
New Jersey	30.7	3244.0
Illinois	30.0	3249.7
Pennsylvania	29.6	3256.9
Massachusetts	30.8	3274.0
Washington	30.3	3321.9