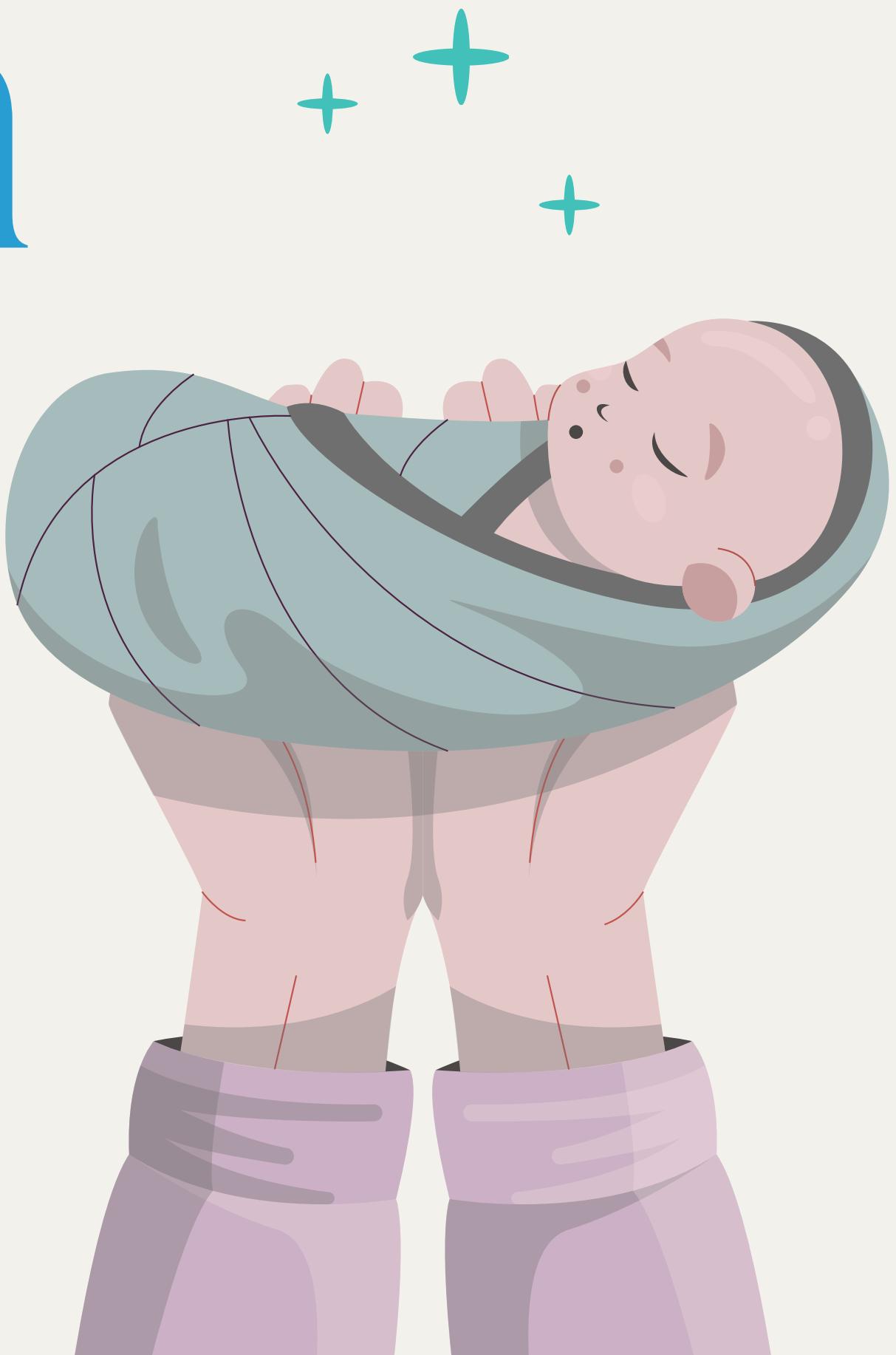


# Newborn Health

Kiran Kumar Rallabandi





# Problem

What is the impact of  
maternal health on newborn weight?

# Dataset of Interest

## 2019 Natality Public Use File



CENTERS FOR DISEASE™  
CONTROL AND PREVENTION

### The Down Side

- + Incomplete Representation of the Population
- + Limited Health Information

### The Bright Side

- + Credible Data Source (NBER)
- + Comprehensive Data Collection
- + Quantifiable Trends
- + Massive Dataset (3.75M)

# Variables

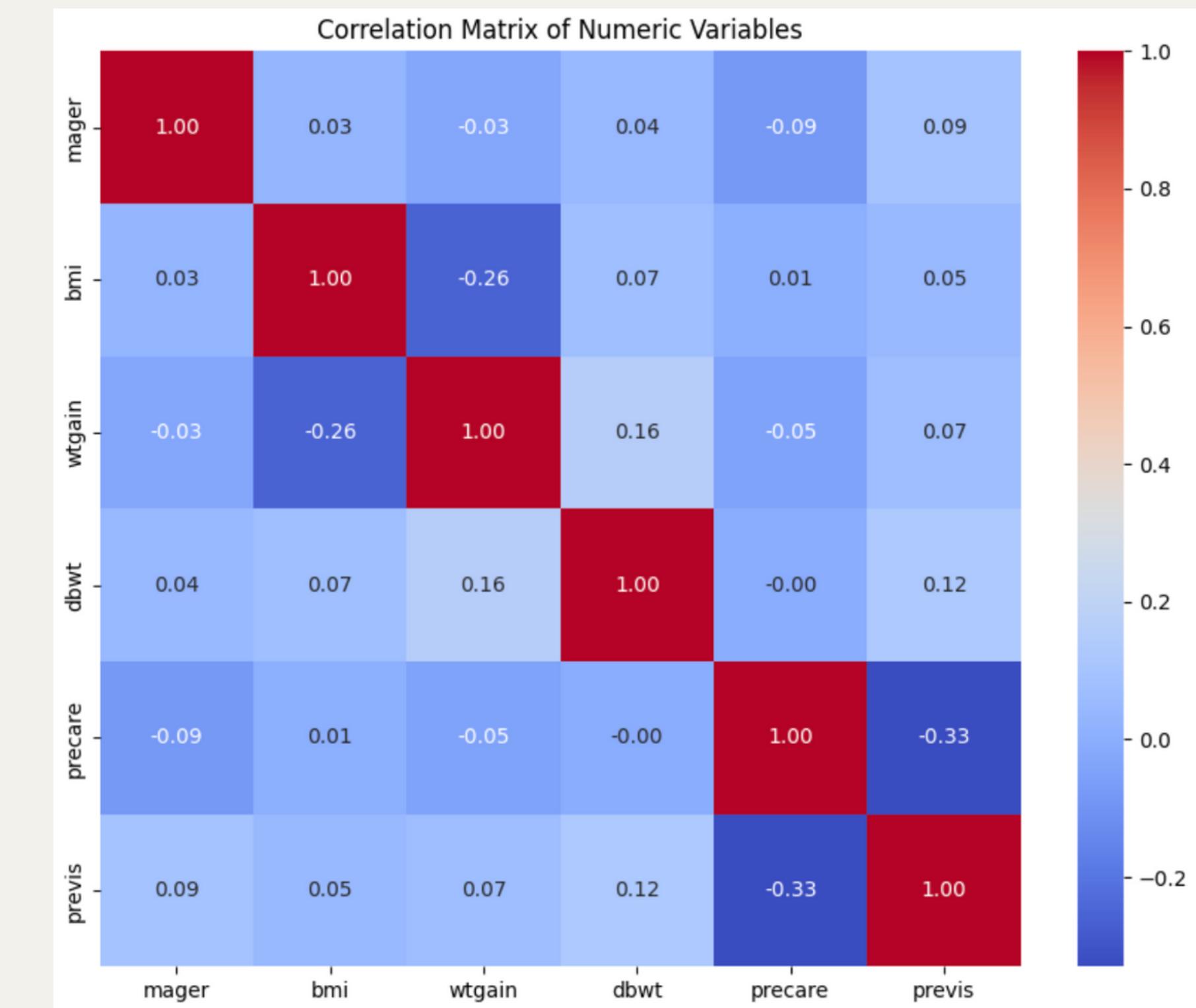
- **Mother's Years of Age**
  - MAGER
- **Cigarette Use during Pregnancy**
  - CIG\_REC (0/1)
- **Body Mass Index**
  - BMI
- **Weight Gain during Pregnancy**
  - WTGAIN (in lb)
- **Gestational Diabetes**
  - RF\_GDIAB (0/1)
- **Sex of Infant**
  - SEX (0/1)

- **Gestational Hypertension**
  - RF\_GHYPE (0/1)
- **Hypertension Eclampsia**
  - RF\_EHYPE (0/1)
- **Infertility Treatment Used**
  - RF\_INFTR (0/1)
- **Month Prenatal Care Began**
  - PRECARE
- **Number of Prenatal Care Visits**
  - PREVIS

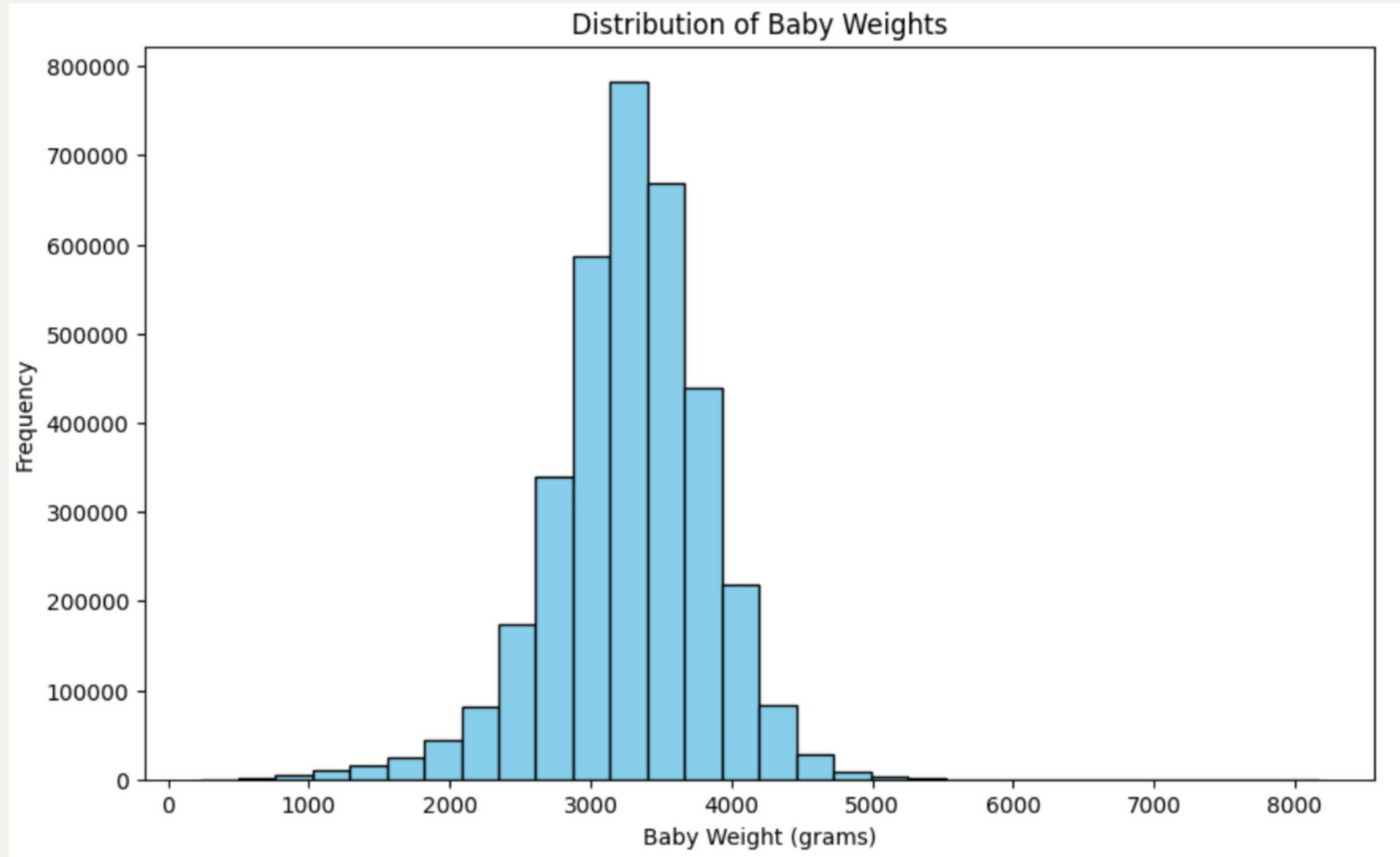
130 → 11

# Descriptive Statistics

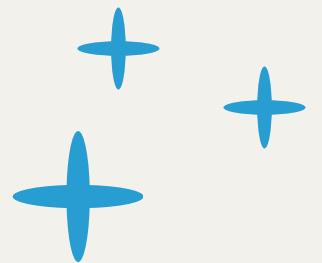
Feature	VIF
mager	15.459884
bmi	11.906884
wtgain	4.336340
previs	8.041002



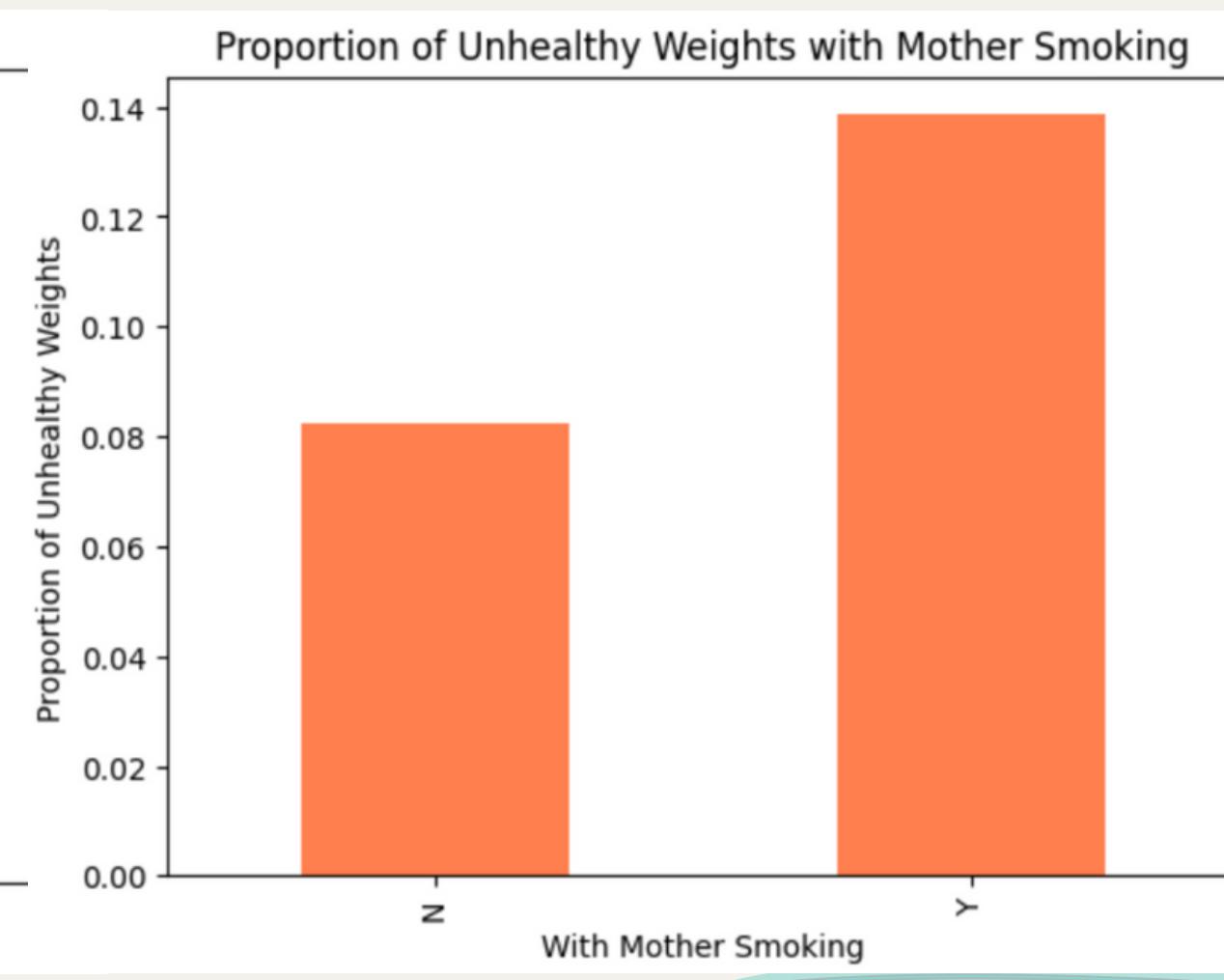
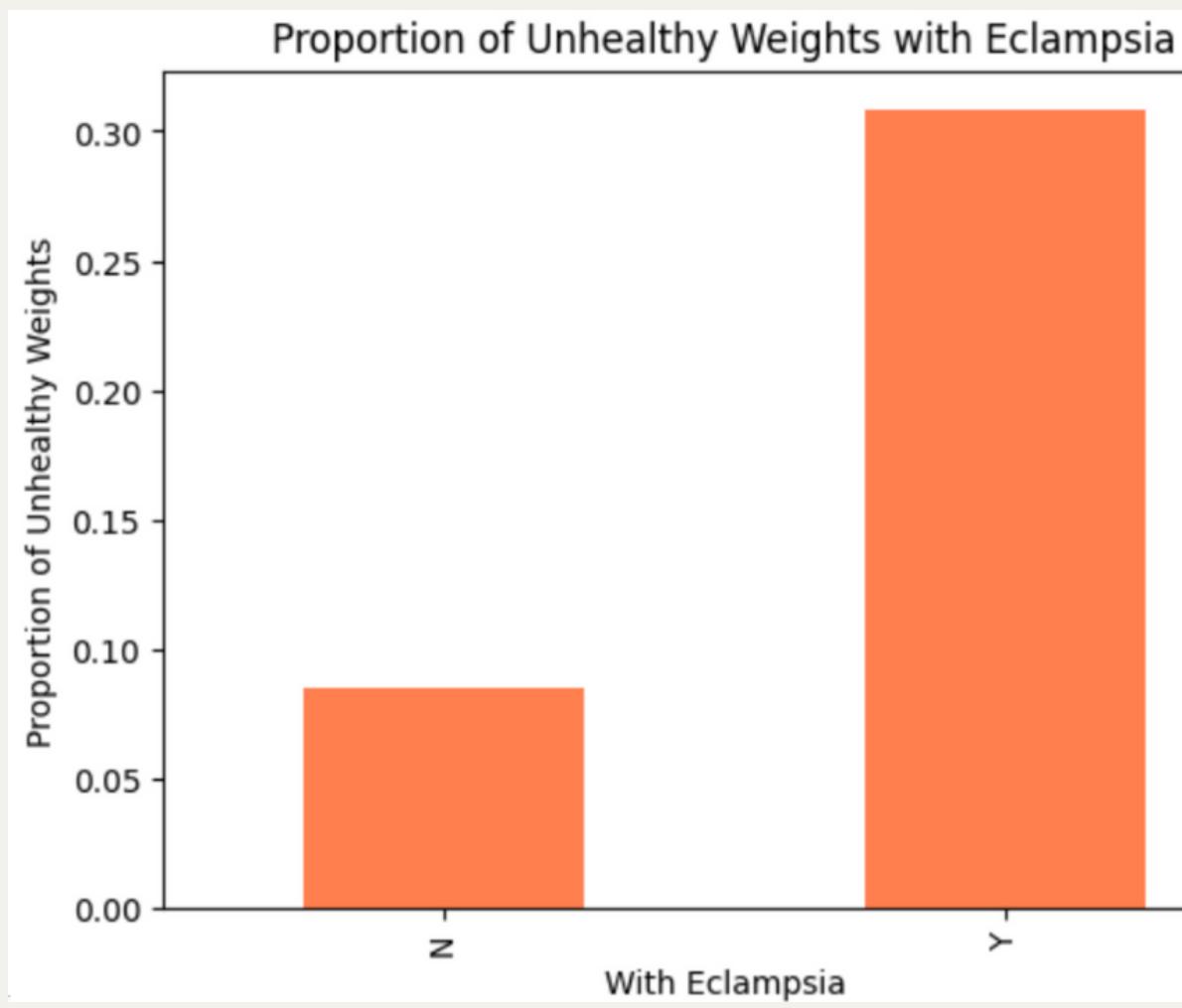
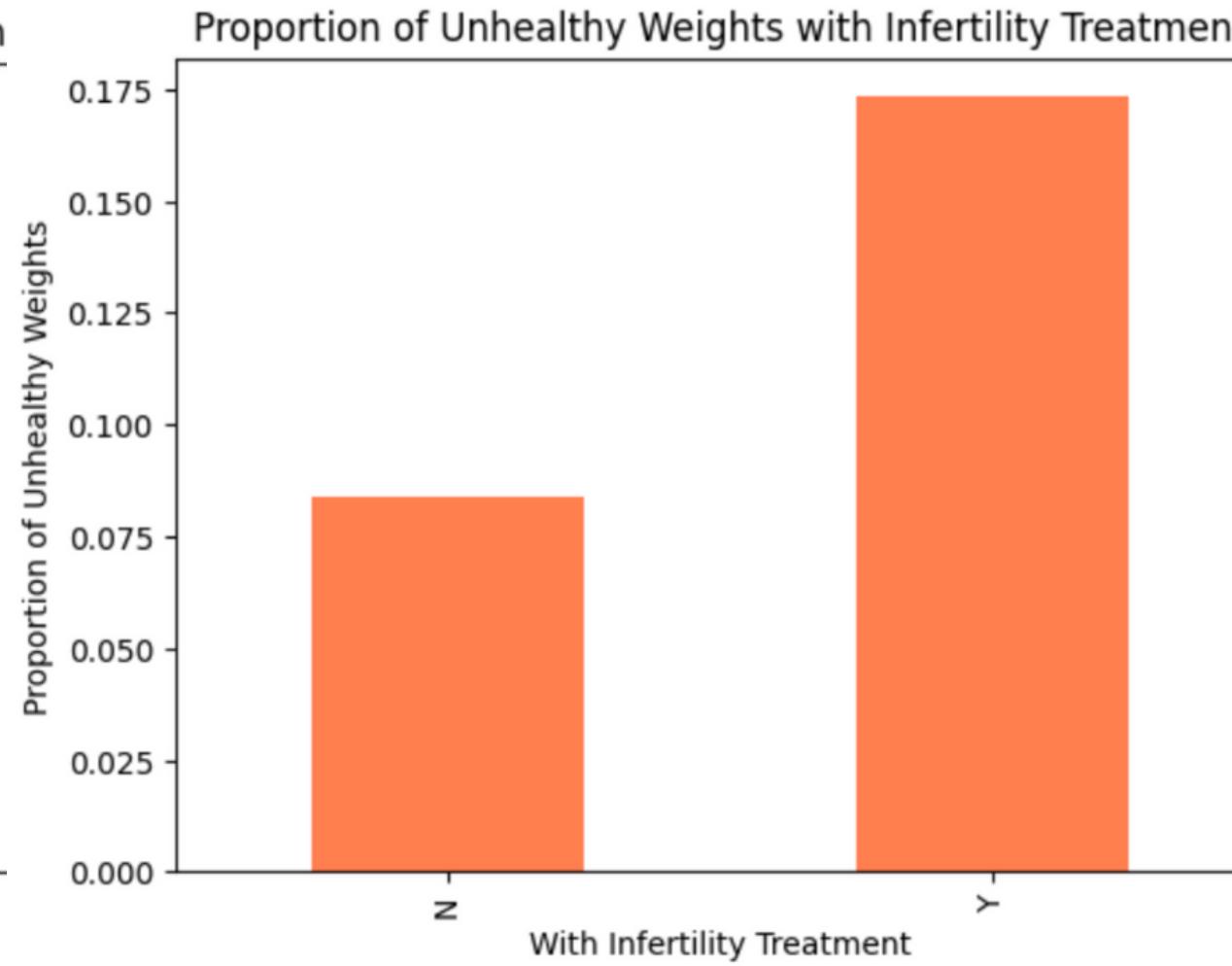
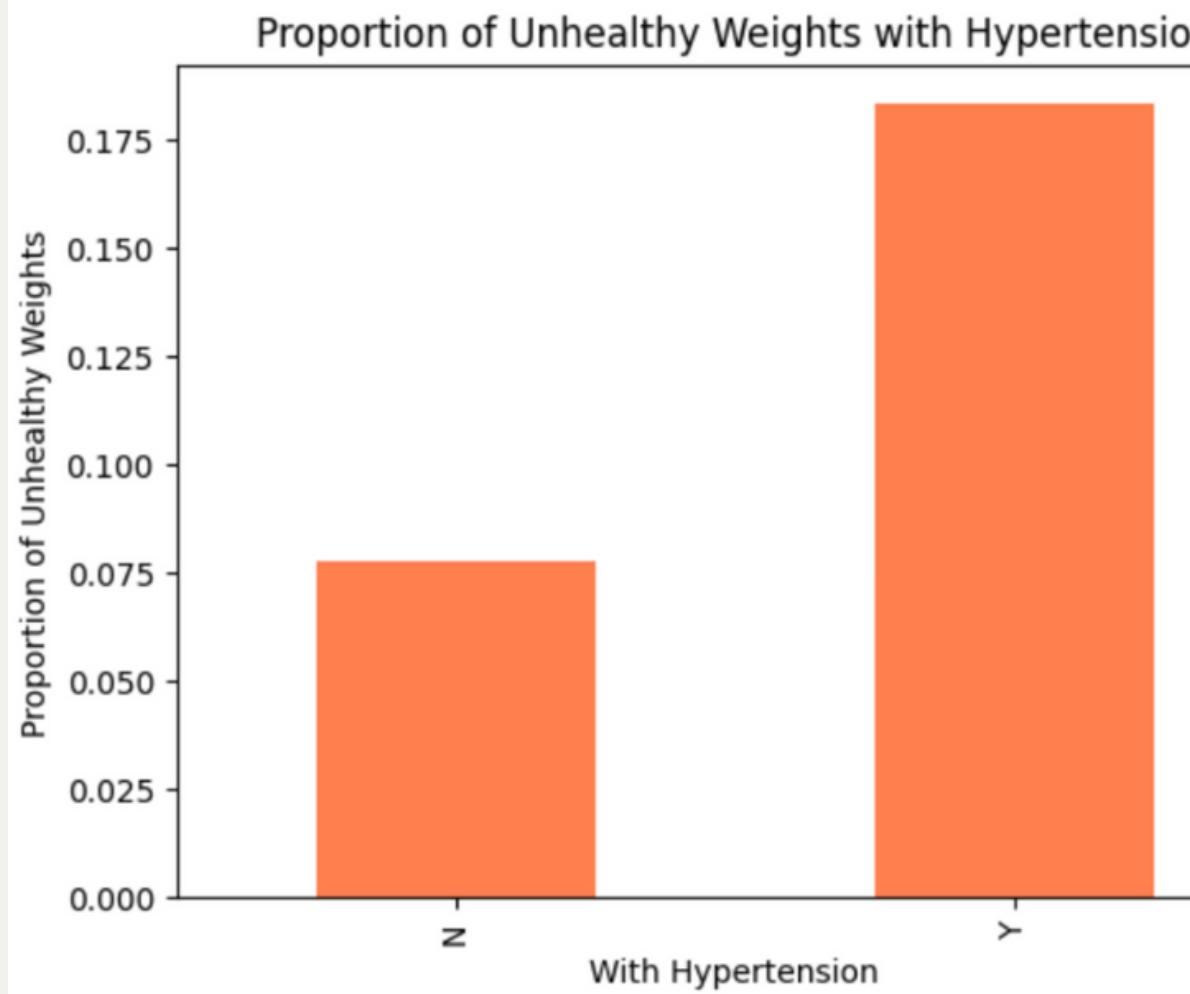
# Descriptive Statistics



# Descriptive Statistics



## Categorical Variables



# Regression Analysis

1

## OLS Regression Results

Dep. Variable:	dbwt	R-squared:	0.021			
Model:	OLS	Adj. R-squared:	0.021			
Method:	Least Squares	F-statistic:	1.544e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:23:29	Log-Likelihood:	-2.7326e+07			
No. Observations:	3536176	AIC:	5.465e+07			
Df Residuals:	3536170	BIC:	5.465e+07			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3245.8656	0.491	6604.798	0.000	3244.902	3246.829
mother_age_Old	-21.4332	0.776	-27.604	0.000	-22.955	-19.911
mother_age_Teen	-130.1887	1.417	-91.861	0.000	-132.966	-127.411
mother_age_Young	-61.2835	0.777	-78.877	0.000	-62.806	-59.761
sex_M	115.7529	0.584	198.097	0.000	114.608	116.898
cig_rec_Y	-192.9054	1.246	-154.819	0.000	-195.348	-190.463
Omnibus:	309324.120	Durbin-Watson:	1.910			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	634344.033			
Skew:	-0.581	Prob(JB):	0.00			
Kurtosis:	4.719	Cond. No.	5.82			

Different Mother's Age Categories

# Regression Analysis ②

OLS Regression Results						
Dep. Variable:	dbwt	R-squared:	0.021			
Model:	OLS	Adj. R-squared:	0.021			
Method:	Least Squares	F-statistic:	1.544e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:23:29	Log-Likelihood:	-2.7326e+07			
No. Observations:	3536176	AIC:	5.465e+07			
Df Residuals:	3536170	BIC:	5.465e+07			
Df Model:	5					
Covariance Type:	nonrobust					
coef	std err	t	P> t	[0.025	0.975]	
const	3245.8656	0.491	6604.798	0.000	3244.902	3246.829
mother_age_Old	-21.4332	0.776	-27.604	0.000	-22.955	-19.911
mother_age_Teen	-130.1887	1.417	-91.861	0.000	-132.966	-127.411
mother_age_Young	-61.2835	0.777	-78.877	0.000	-62.806	-59.761
sex_M	115.7529	0.584	198.097	0.000	114.608	116.898
cig_rec_Y	-192.9054	1.246	-154.819	0.000	-195.348	-190.463
Omnibus:	309324.120	Durbin-Watson:	1.910			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	634344.033			
Skew:	-0.581	Prob(JB):	0.00			
Kurtosis:	4.719	Cond. No.	5.82			

Sex of the Child, Cigarette Usage during Pregnancy

# Regression Analysis ③

OLS Regression Results						
Dep. Variable:	dbwt	R-squared:	0.016			
Model:	OLS	Adj. R-squared:	0.016			
Method:	Least Squares	F-statistic:	1.903e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:25:12	Log-Likelihood:	-2.7336e+07			
No. Observations:	3536176	AIC:	5.467e+07			
Df Residuals:	3536172	BIC:	5.467e+07			
Df Model:	3					
Covariance Type:	nonrobust					
coef	std err	t	P> t	[0.025	0.975]	
const	3291.2163	0.314	1.05e+04	0.000	3290.600	3291.833
rf_gdiab_Y	20.4737	1.153	17.752	0.000	18.213	22.734
rf_ghype_Y	-245.2919	1.095	-223.942	0.000	-247.439	-243.145
rf_ehype_Y	-446.4073	5.594	-79.801	0.000	-457.371	-435.443
Omnibus:	259792.210	Durbin-Watson:	1.913			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	517579.756			
Skew:	-0.509	Prob(JB):	0.00			
Kurtosis:	4.573	Cond. No.	19.2			

Gestational Diabetic, Hypertension, Eclampsia

# Regression Analysis

4

OLS Regression Results

Dep. Variable:	dbwt	R-squared:	0.025			
Model:	OLS	Adj. R-squared:	0.025			
Method:	Least Squares	F-statistic:	2.973e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:25:26	Log-Likelihood:	-2.7320e+07			
No. Observations:	3536176	AIC:	5.464e+07			
Df Residuals:	3536172	BIC:	5.464e+07			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3095.7561	0.674	4591.405	0.000	3094.435	3097.078
wtgain	5.9036	0.020	292.421	0.000	5.864	5.943
rf_gdiab_Y	73.8638	2.194	33.668	0.000	69.564	78.164
wtgain_rf_gdiab_Y	-1.6392	0.073	-22.352	0.000	-1.783	-1.496
Omnibus:	293107.078	Durbin-Watson:			1.906	
Prob(Omnibus):	0.000	Jarque-Bera (JB):			580029.940	
Skew:	-0.566	Prob(JB):			0.00	
Kurtosis:	4.630	Cond. No.			251.	

OLS Regression Results

Dep. Variable:	wtgain	R-squared:	0.006			
Model:	OLS	Adj. R-squared:	0.006			
Method:	Least Squares	F-statistic:	2.273e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:25:31	Log-Likelihood:	-1.4599e+07			
No. Observations:	3536176	AIC:	2.920e+07			
Df Residuals:	3536174	BIC:	2.920e+07			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	29.8522	0.008	3604.092	0.000	29.836	29.868
rf_gdiab_Y	-4.7343	0.031	-150.751	0.000	-4.796	-4.673
Omnibus:	241499.271	Durbin-Watson:			1.941	
Prob(Omnibus):	0.000	Jarque-Bera (JB):			364533.824	
Skew:	0.566	Prob(JB):			0.00	
Kurtosis:	4.093	Cond. No.			3.95	

Interactive Variable: Diabetic & Mothers' Weight Gain

Diabetic v.s. Non-Diabetic

# Regression Analysis ⑤

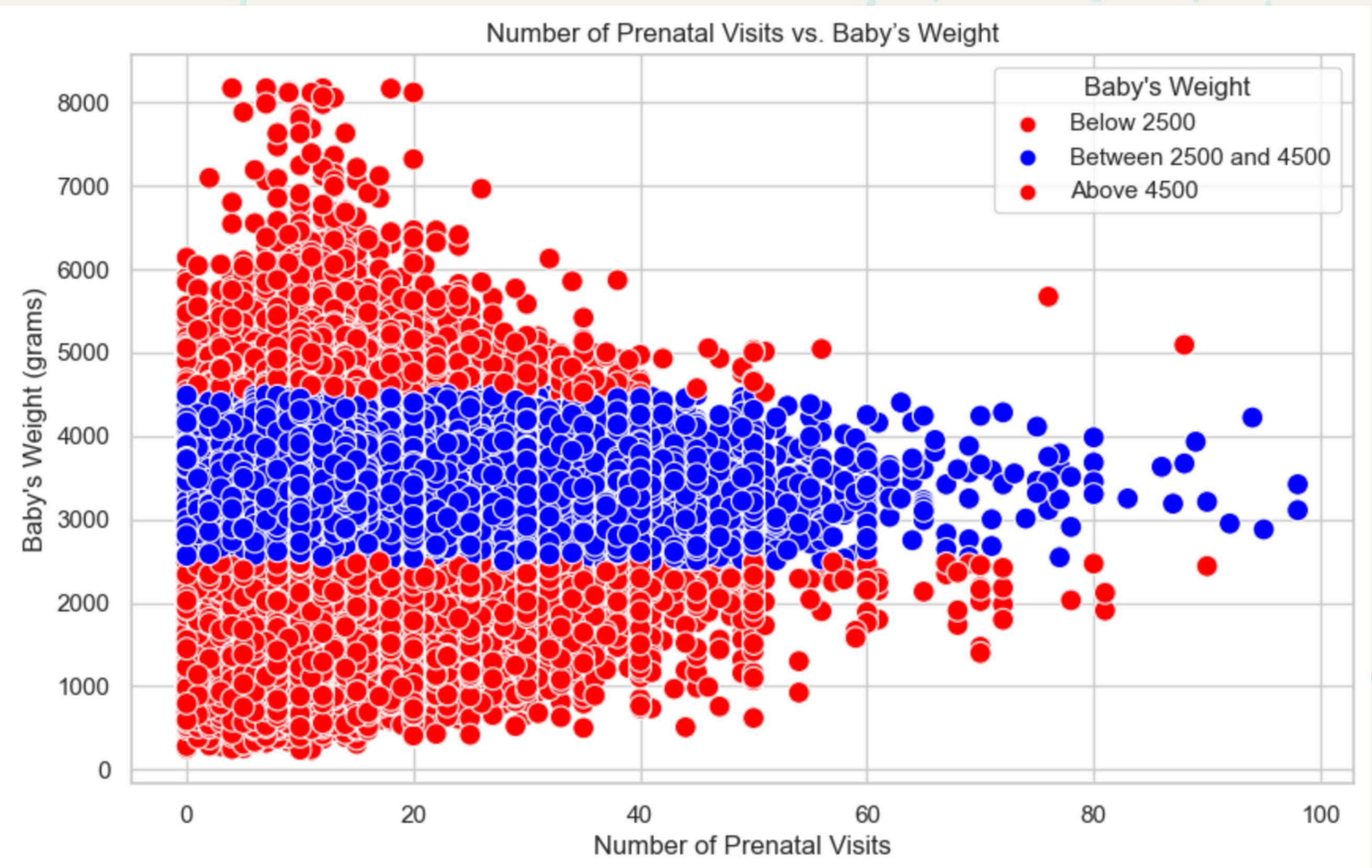


## OLS Regression Results

Dep. Variable:	dbwt	R-squared:	0.002			
Model:	OLS	Adj. R-squared:	0.002			
Method:	Least Squares	F-statistic:	5353.			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:25:35	Log-Likelihood:	-2.7361e+07			
No. Observations:	3536176	AIC:	5.472e+07			
Df Residuals:	3536174	BIC:	5.472e+07			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3275.4398	0.298	1.1e+04	0.000	3274.856	3276.024
rf_inftr_Y	-152.5960	2.086	-73.165	0.000	-156.684	-148.508
Omnibus:	281798.958	Durbin-Watson:	1.912			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	560849.984			
Skew:	-0.546	Prob(JB):	0.00			
Kurtosis:	4.617	Cond. No.	7.07			

Infertility Treatment Used

# Prenatal Care Visits



} **Healthy Baby Weight**  
**2500g-4500g**  
**(Healthline)**

# Regression Analysis

7

## OLS Regression Results

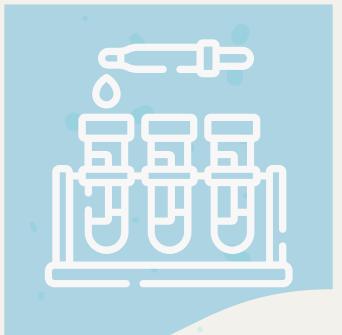
Dep. Variable:	dbwt	R-squared:	0.003			
Model:	OLS	Adj. R-squared:	0.003			
Method:	Least Squares	F-statistic:	1.173e+04			
Date:	Wed, 29 Nov 2023	Prob (F-statistic):	0.00			
Time:	19:30:37	Log-Likelihood:	-2.7358e+07			
No. Observations:	3536176	AIC:	5.472e+07			
Df Residuals:	3536174	BIC:	5.472e+07			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3225.0866	0.526	6127.048	0.000	3224.055	3226.118
previs	3.9566	0.037	108.316	0.000	3.885	4.028
Omnibus:	282041.256	Durbin-Watson:	1.911			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	558415.609			
Skew:	-0.548	Prob(JB):	0.00			
Kurtosis:	4.609	Cond. No.	25.8			

Number of Prenatal Care Visits

# Future



**Eclampsia treatment**



**More demographic-specific analysis**



**Create targeted interventions**

