Table of Contents

Summary

The main aim of this study was to determine how the presence of different metals have an impact on the weight of a new born child and to see how the Maternal pre-pregnancy body mass index, Maternal weight gain during pregnancy, Gestational age at birth and Maternal age at the time of birth impact on the weight of a new born child. For this purpose, a sample of 1000 observation of pregnant ladies was used. Through this sample we concluded that child birth weight Arsenic and cadmium were the only two metals which were significant in determining birth weight. It was also determined that those two have a negative relationship with the birth weight. While the Maternal pre-pregnancy body mass index, Maternal weight gain during pregnancy and Gestational age at birth are positively correlated as well as they have a higher contribution as compare to metals.

Goals

The main aim of this report is to determine that how several factors affect the weight of a new born child. Most of the factors that affect the weight of a child are during the pregnancy period. At that time the child grows in the womb of a mother therefore factors considered in this task are several factors which affect the health of both mother and baby.

Goals of this report are:

- To determine how the presence of different metals have an impact on the weight of a new born child.
- To determine how the Maternal pre-pregnancy body mass index, Maternal weight gain during pregnancy (kg), Gestational age at birth (week) and Maternal age (years) at the time of birth impact on the weight of a new born child.

Exploratory Data Analysis

In this section of the report data will be analyzed; both independent and dependent variables used in this report will be analyzed thoroughly.

Dependent variable

In this case our dependent or the outcome variable is child birth weight. First of all, we will check the descriptive statistics of the data for which results are as follows:

```
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
1280    3050    3390    3378    3720    4850
```

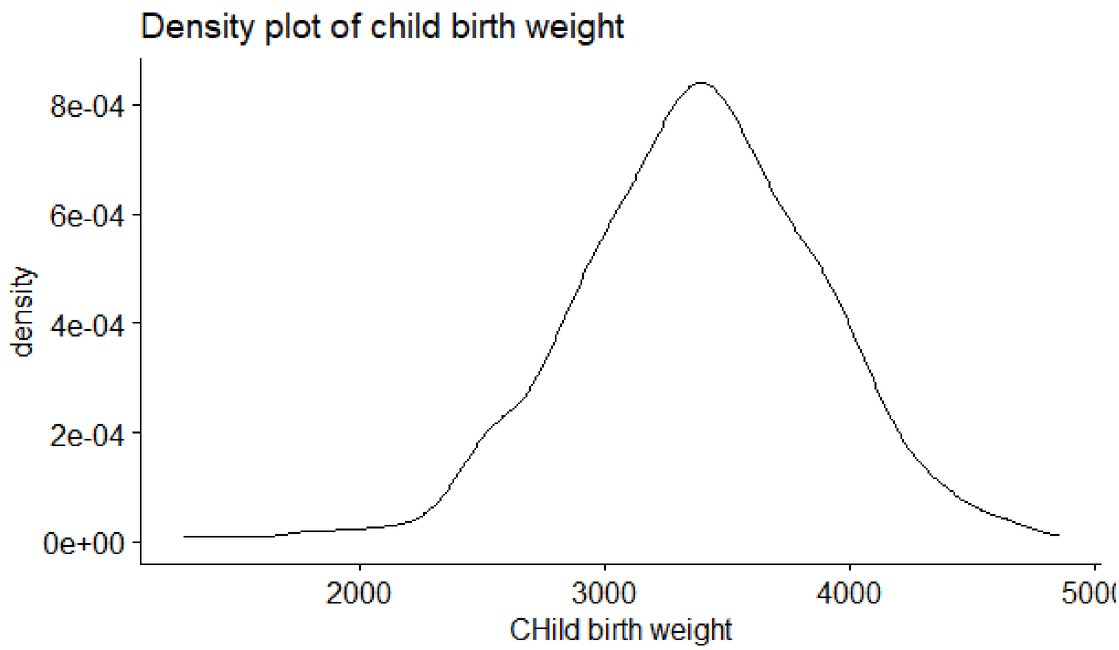
Descriptive statistics shows that the data has a minimum weight of 1280 kg and maximum of 4850 kg. The average weight determined for this sample is 3378 kg and the median weight is 3390 kg. 25% of the data falls below 3050 kg while 25% falls beyond 3720 kg.

Normality check:

It is necessary that the dependent variable should be normally distributed. To determine that whether it is normally distributed or not there are two normality checks through graphs and a normality test.

Density graph

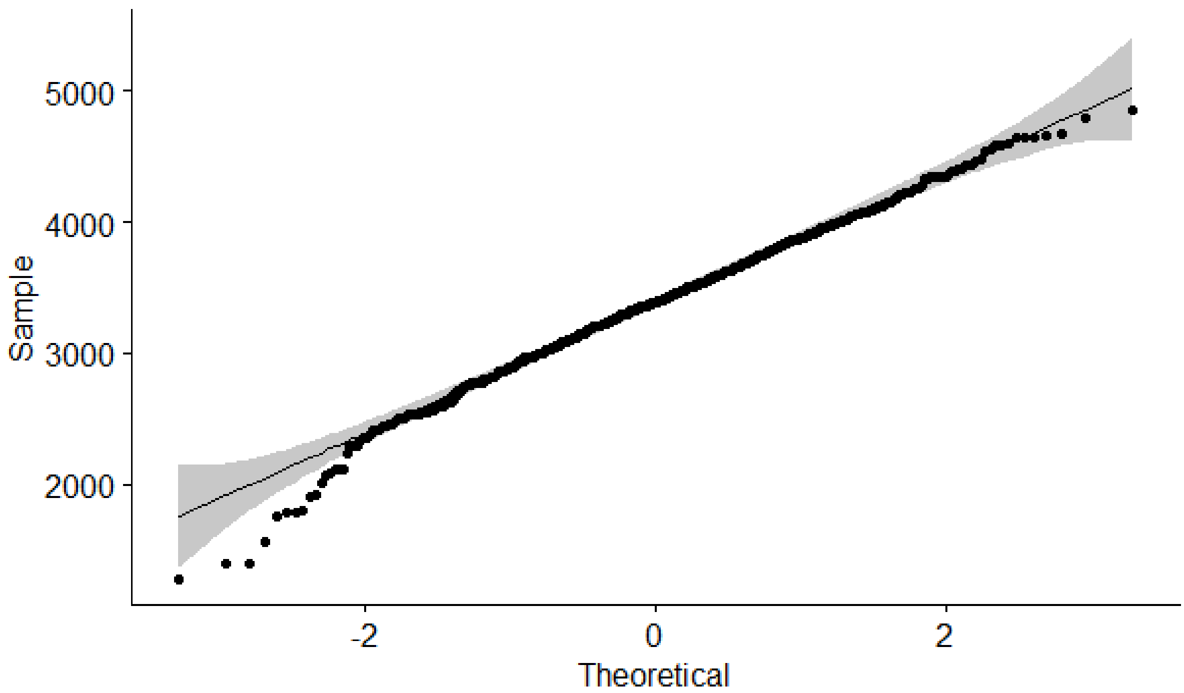
Following is the density plot of child birth weight which is used to tell us the distribution of the data.



Density plot of Child birth weight shows that the weights of the children are normally distributed, as the graph shows a bell shape. This indicates that the data falls equally on both sides of the distribution.

QQ plot

Following is the QQ plot of child birth weight which is used to determine normality of the data.



The plot shows that the data is normally distributed as the maximum data points fall on the QQ line and in its 95% range while we can see there are several outliers in the data.

Correlation matrix

Our goal is to find the correlation with the independent and the dependent variables. In this report we have several goals to attain each goal we will be using a different model.

For the goal 1 we have attained the following correlation matrix, followed by a p-value matrix for the correlations determined:

	Child birth weight	Arsenic (As) in mother	Cadmium (Cd) in mother	Cobalt (Co) in mother	Caesium (Cs) in mother	Copper (Cu) in mother	Mercury (Hg) in mother	Manganese (MN) in mother	Molybdenum (Mo) in mother	Lead (Pb) in mother
Child birth weight	1	-0.08	-0.07	-0.04	0.01	-0.01	-0.08	-0.01	-0.06	-0.04
Arsenic (As) in mother		1	0	-0.04	0.26	-0.07	0.36	-0.2	0.14	-0.13
Cadmium (Cd) in mother			1	0.02	-0.03	0.18	0.08	0.1	0.09	0.19
Cobalt (Co) in mother				1	-0.4	-0.15	-0.05	0.27	-0.01	-0.09
Caesium (Cs) in mother					1	-0.02	0.27	-0.31	0.11	0.06
Copper (Cu) in mother						1	-0.09	0.03	0.06	0.11
Mercury (Hg) in mother							1	-0.16	0.11	0.05
Manganese (MN) in mother								1	-0.05	0.11
Molybdenum (Mo) in mother									1	0
Lead (Pb) in mother										1

p-value matrix:

	Child birth weight	Arsenic (As) in mother	Cadmium (Cd) in mother	Cobalt (Co) in mother	Caesium (Cs) in mother	Copper (Cu) in mother	Mercury (Hg) in mother	Manganese (MN) in mother	Molybdenum (Mo) in mother	Lead (Pb) in mother
Child birth weight		0.0091	0.0271	0.2137	0.8597	0.8225	0.0124	0.6851	0.0622	0.1861
Arsenic (As) in mother			0.9879	0.2242	0.0000	0.0367	0.0000	0.0000	0.0000	0.0000
Cadmium (Cd) in mother				0.5044	0.3641	0.0000	0.0171	0.0021	0.0043	0.0000
Cobalt (Co) in mother					0.0000	0.0000	0.1111	0.0000	0.7497	0.0043
Caesium (Cs) in mother						0.5897	0.0000	0.0000	0.0006	0.0488
Copper (Cu) in mother							0.0027	0.3687	0.0728	0.0005
Mercury (Hg) in mother								0.0000	0.0004	0.0849
Manganese (MN) in mother									0.1272	0.0006
Molybdenum (Mo) in mother										0.9820
Lead (Pb) in mother										

Half of the matrix is clear as it will show the same results as in the upper part of the matrix. In the matrix 1 we can see that the weight has a low and a negative correlation with the covariates considered along with that there are several covariates which are not significantly contributors. Highlighted parts in matrix 1 shows that there is certain evidence of multicollinearity in the variables. To omit multicollinearity from the model several covariates were omitted from the model to get better results.

For the goal 2 we have attained the following correlation matrix, followed by a p-value matrix for the correlations determined:

	Child birth weight	Maternal pre-pregnancy body mass index (kg/m2)	Maternal weight gain during pregnancy (kg)	Gestational age at birth (week)	Maternal age (years)
Child birth weight	1	0.12	0.16	0.54	-0.02
Maternal pre-pregnancy body mass index (kg/m2)		1	-0.05	-0.01	-0.13
Maternal weight gain during pregnancy (kg)			1	0.11	-0.01
Gestational age at birth (week)				1	-0.03
Maternal age (years)					1

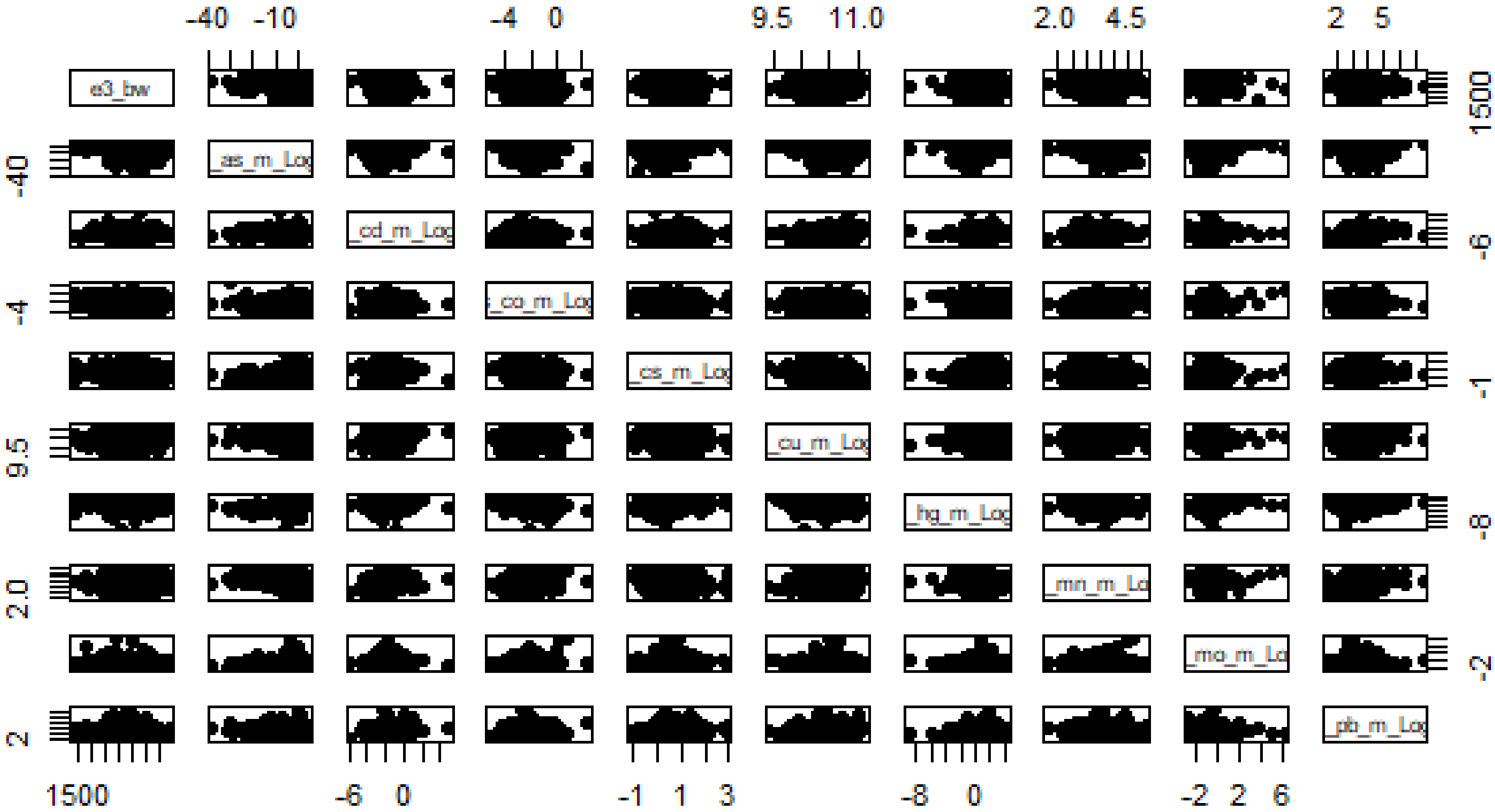
P-value table:

	Child birth weight	Maternal pre-pregnancy body mass index (kg/m2)	Maternal weight gain during pregnancy (kg)	Gestational age at birth (week)	Maternal age (years)
Child birth weight		0.0002	0.0000	0.0000	0.4679
Maternal pre-pregnancy body mass index (kg/m2)			0.1133	0.7899	0.0000
Maternal weight gain during pregnancy (kg)				0.0007	0.7473
Gestational age at birth (week)					0.3069
Maternal age (years)					

Table above shows that the Maternal pre-pregnancy body mass index and Maternal weight gain during pregnancy are positive but have a low correlation with the child birth weight. While Gestational age at birth is significantly correlated with child birth weight. These all covariates are significantly correlated with child birth weight. There is evidence of multicollinearity of Gestational age at birth with Maternal pre-pregnancy body mass index and Maternal weight gain during pregnancy but the relationship between them is negligible and these correlations are insignificant. Maternal age has a low correlation as well as it is correlated with other covariates so we omitted that variable.

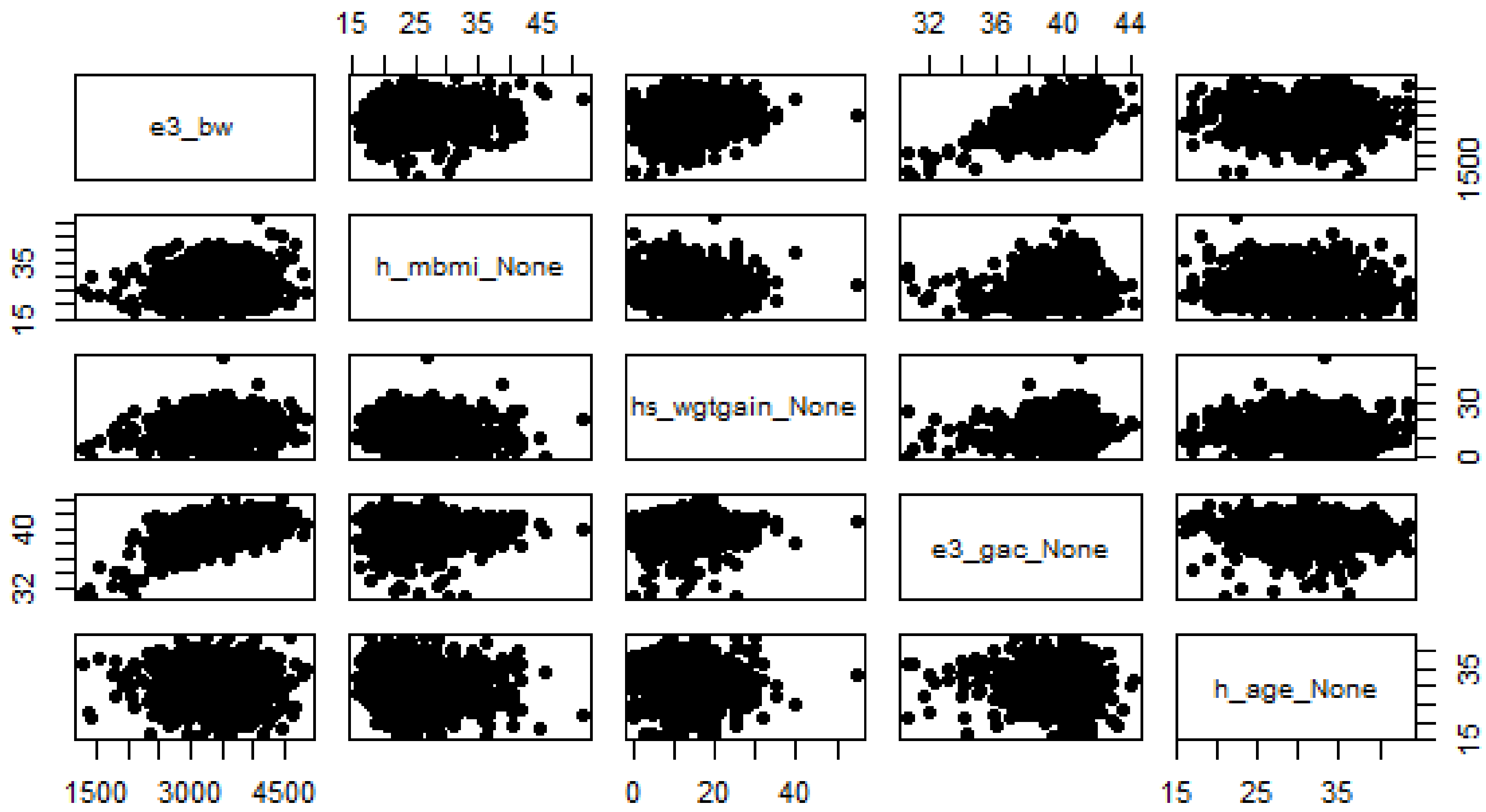
Scatter plot:

Following is the scatter plot obtained for the goal 1:



Result shows that the data is correlated but a low correlation is present. Results are similar to what we obtained from the correlation matrix.

Following is the scatter plot obtained for the goal 2:



Result shows that the data is correlated but a moderate correlation is present as we obtained from the correlation matrix.

Methods

As observed above all the assumptions of a simple linear model are fulfilled we can apply a simple linear model to the data. There is no transformation applied on the outcome variable as it is found to be normal. But there are certain independent variables which are already transformed.

Constraints that show evidence of multicollinearity in the model were omitted from the model then several models were obtained in which the model with the most significant results and high R^2 were chosen for analysis.

Results

In this section we will see the results computed for the regression analysis.

Results for model 1:

These results are computed by keeping goal 1 in mind and the results of the assumptions computed above.

Syntax:

Following is the syntax used to obtain results:

```
model=lm(pollution$e3_bw~pollution$hs_as_m_Log2+pollution$hs_cd_m_Log2)
summary(model)
```

Results:

```
Call:
lm(formula = pollution$e3_bw ~ pollution$hs_as_m_Log2 + pollution$hs_cd_m_Log2)

Residuals:
     Min       1Q   Median       3Q      Max 
-2093.05  -303.81     8.09   328.90  1429.72 

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)            3271.780     42.171  77.583  < 2e-16 ***
pollution$hs_as_m_Log2   -7.486      2.856  -2.621  0.00891 ** 
pollution$hs_cd_m_Log2  -39.539     17.801  -2.221  0.02656 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 506.8 on 997 degrees of freedom
Multiple R-squared:  0.01169,   Adjusted R-squared:  0.009712 
F-statistic: 5.899 on 2 and 997 DF,  p-value: 0.00284
```

Model

$$weight = 3271.78 - 7.486 * Arsenic - 39.539 * Cadmium$$

It is observed that the covariates have a significant contribution to the model as the p-value of the covariates are below 5%. The slope coefficient is also a significant contributor to the model. Anova table results shows that for F-statistic 5.899 the p-value is 0.00284 which is significant for all levels. This also indicates that the fitted model is a good fit.

Results indicates that even if there is no amount of metals in the body the weight of the child will be 3271.78 kg which is below the average. If there is 1-unit increase of Arsenic in mother's body, then there will a decrease of 7.486 units in the weight of the baby while keeping other variants constant. If there is 1-unit increase of Cadmium in mother's body, then there will a decrease of 39.539 units in the weight of the baby while keeping other variants constant.

Results for model 2:

Keeping the assumptions results and goal 1 in mind results for the goal are obtained below.

Syntax:

Following is the syntax used to obtain results:

model1=lm(pollution$e3_bw~pollution$h_mbmi_None+pollution$hs_wgtgain_None+pollution$e3_gac_None)

summary(model1)

Results:

```
Call:
lm(formula = pollution$e3_bw ~ pollution$h_mbmi_None + pollution$hs_wgtgain_None +
    pollution$e3_gac_None)

Residuals:
     Min       1Q   Median       3Q      Max
-1245.23  -266.74    -2.42   280.52  1429.40

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -3337.345    315.084 -10.592  < 2e-16 ***
pollution$h_mbmi_None        12.583      2.567   4.901 1.11e-06 ***
pollution$hs_wgtgain_None     9.055      2.130   4.251 2.33e-05 ***
pollution$e3_gac_None       158.382      7.803  20.298  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 419.8 on 996 degrees of freedom
Multiple R-squared:  0.3224,	Adjusted R-squared:  0.3204
F-statistic:   158 on 3 and 996 DF,  p-value: < 2.2e-16
```

Model

$weight =- 3337.345 + 12.583 * Maternal\ pre - pregnancy\ body\ mass\ index + 9.055 * Mater$

Results shows that there will be an increment of 158.382 units in weight of the child by increasing 1 unit of gestational age at birth. Keeping in mind the results above we can state the if there is 1-unit increase in Maternal pre-pregnancy body mass index, then there will an increase of 12.583 units in the weight of the baby while keeping other variants constant. We can also state the if there is 1-unit increase in Maternal weight gain during pregnancy, then there will an increase of 9.055 units in the weight of the baby while keeping other variants constant.

Results indicates that even if there is no amount of mother's body weight determined in the body then the weight of the child will be -3337.345 kg which is not feasible both theoretically and practically.

It is observed that the covariates have a significant contribution to the model as the p-value of the covariates are below 5% and 1%. The slope coefficient is also a significant contributor to the model. Anova table results shows that for F-statistic 158 the p-value is 0.00 which is significant for all levels. This also indicates that the fitted model is a good fit.

Discussion

Results from this study shows that the Arsenic and cadmium are the two metals that have a great impact in changing the weight of the child to be taken birth. The high amount of those metals will be a cause of child weight loss. It is the main concern of the scientist that during pregnancy the amount of Arsenic, lead and cadmium should be minimum in the body of the mother. Study shows that in Myanmar there is an effect of prenatal heavy metal exposure on the birth outcomes among. This examination recognized that Myanmar moms were profoundly presented to cadmium. Pre-birth maternal cadmium openness was related with an event of low birth weight (Wai, Mar, Kosaka, Umemura & Watanabe, 2021).

Pre-birth substantial metals directness adversely affects birth weight. In any case, their effect on various clinical types of fetal diminutiveness was rarely evaluated. Fetal serum levels of Cd showed a differential relationship between little embryos' clinical subclassification, which along with the expanded Cd levels in both maternal and fetal serum of the little hatchlings support the negative impact of substantial metals on birth weight. These discoveries give more freedom to check the job of weighty metals openness comparable to little embryos' subclassification (Sabra, Malmqvist, Saborit, Gratacós & Gomez Roig, 2017).

Mother weight and its gestational age are impactful of the weight of the children. Unnecessary pregnancy weight acquired seems to build birth weight and the posterity's danger for corpulence sometime down the road. In any case, this relationship might be puzzled by hereditary and other shared impacts. Maternal weight acquired during pregnancy expands birth weight autonomous of hereditary elements. Taking into account the clear connection between birth weight and grown-up weight, corpulence avoidance designated to lady during pregnancy might be justified (Ludwig & Currie, 2011).

References

Wai, K., Mar, O., Kosaka, S., Umemura, M., & Watanabe, C. (2017). Prenatal Heavy Metal Exposure and Adverse Birth Outcomes in Myanmar: A Birth-Cohort Study. Retrieved 3 August 2021.

Sabra, S., Malmqvist, E., Saborit, A., Gratacós, E., & Gomez Roig, M. (2017). Heavy metals exposure levels and their correlation with different clinical forms of fetal growth restriction. *PLOS ONE*, *12*(10), e0185645. doi: 10.1371/journal.pone.0185645

Ludwig, D., & Currie, J. (2011). The Association Between Pregnancy Weight Gain and Birthweight. *Obstetric Anesthesia Digest*, *31*(4), 219. doi: 10.1097/01.aoa.0000406669.86787.0a

Appendix:

```
install.packages("dplyr")

install.packages("devtools")

devtools::install_github("kassambara/ggpubr")

install.packages("ggpubr")

library("dplyr")

library("ggpubr")

ggdensity(pollution$e3_bw,

        main = "Density plot of child birth weight",

        xlab = "CHild birth weight")

ggqqplot(pollution$e3_bw)

summary(pollution$e3_bw)

install.packages(tidyverse)

install.packages(ggpubr)

install.packages(rstatix)

library(tidyverse)

library(ggpubr)

library(rstatix)

mydata.cor = cor(pollution, method = c("spearman"))

mydata.rcorr = rcorr(as.matrix(pollution))

mydata.rcorr

scatterplot(pollution$e3_bw, pollution$hs_as_m_Log2)

install.packages("corrplot")

library("corrplot")
```

```
mydata <- pollution[, c(1,18,19,20,21,22,23,24,25,26)]

head(mydata)

pairs(mydata2,pch=19)

rquery.cormat(mydata)

mydata.rcorr = rcorr(as.matrix(mydata))

mydata.rcorr


mydata2 <- pollution[, c(1,76,77,78,79)]

head(mydata2)

rquery.cormat(mydata2)

mydata2.rcorr = rcorr(as.matrix(mydata2))

mydata2.rcorr


lr1=lm(pollution$e3_bw~pollution$hs_as_m_Log2)

lr1

summary(lr1)

install.packages("tidyverse")

library("tidyverse")

model=lm(pollution$e3_bw~pollution$hs_as_m_Log2+pollution$hs_cd_m_Log2)

summary(model)

model1=lm(pollution$e3_bw~pollution$h_mbmi_None+pollution$hs_wgtgain_None+polluti
on$e3_gac_None)

summary(model1)
```