Q1. What is the mean father's age?

28.90

Q2. What is the mean father's age for low birthweight babies?

24.83

Q3. Is the father's age normally distributed? Justify your answer.

No, the Shapiro Wilk test of this variable has a significance value of 0.039 therefore it is not normally distributed.

Q4. If you apply the log transformation to the father's age, what is the mean score of the transformed variable?

1.45

Q5. Is the above mean score a good representation of the real value? Justify your answer.

It is not, because when you transform the variable you chance the values.

Q6. Is the new variable (log transform of father's age) normally distributed? Justify your answer.

It is, the Shapiro Wilk test result is 0.129 which is bigger than 0.05.

Q7. Is the variable "years father was in education" normally distributed?

No, it is not.

Q8. Mentioning the null and alternative hypotheses, explain the above answer.

The null hypotheses in this case would be that "years father was in education" is normally distributed or that the sig. value in the Shapiro-Wilk test is larger than 0.05 whereas the alternative hypotheses would be that "years father was in education" is not normally distributed or that the sig. value in the Shapiro-Wilk test is smaller or equal to 0.05. Since the value of the sig. for this variable in the Shapiro- Wilk test is <0.001 the null hypotheses is rejected, therefore the alternative hypotheses is accepted.

Q9. What is the mean score for the variable "years father was in education" after you apply the Box-Cox transformation?

13.71.

Q10. Is this new variable normally distributed? Explain.

No , it isn't. The sig. value of the Shapiro-Wilk test is <0.001.

Q11. What is the mean score for this new variable (B-C transformed fathers' years in education) for mothers aged under 35?

13.55.

Q12. Which test would you use to investigate the relationship between birth weight and father's age?

- **Pearson product-moment correlation**
- **Spearman's Rank order correlation**
- **Point-Biserial correlation**
- **Phi-Coefficient**

Pearson product-moment correlation.

Does the variable needs to be a normal distribution? YES!!!

Q13. Justify the above choice in terms of the distribution of data and the nature of the test.

Because neither of the variables is dichotomous, weight is a ratio variable and father's age is an interval variable and their measures are linear, therefore I would use the Pearson product-moment correlation.

Q14. What is the direction of that relationship?

Positive.

Q15. What is the form of that relationship?

Linear

Q16. What is the degree of that relationship?

Weak, negligible correlation.

Q17. What test would you use to investigate the relationship between smoking and birth weight?

- **Pearson product-moment correlation**
- **Spearman's Rank order correlation**
- **Point-Biserial correlation**
- **Phi-Coefficient**

Point-biserial correlation.

Q18. Report on the above results including information about direction/form/degree of the relationship.

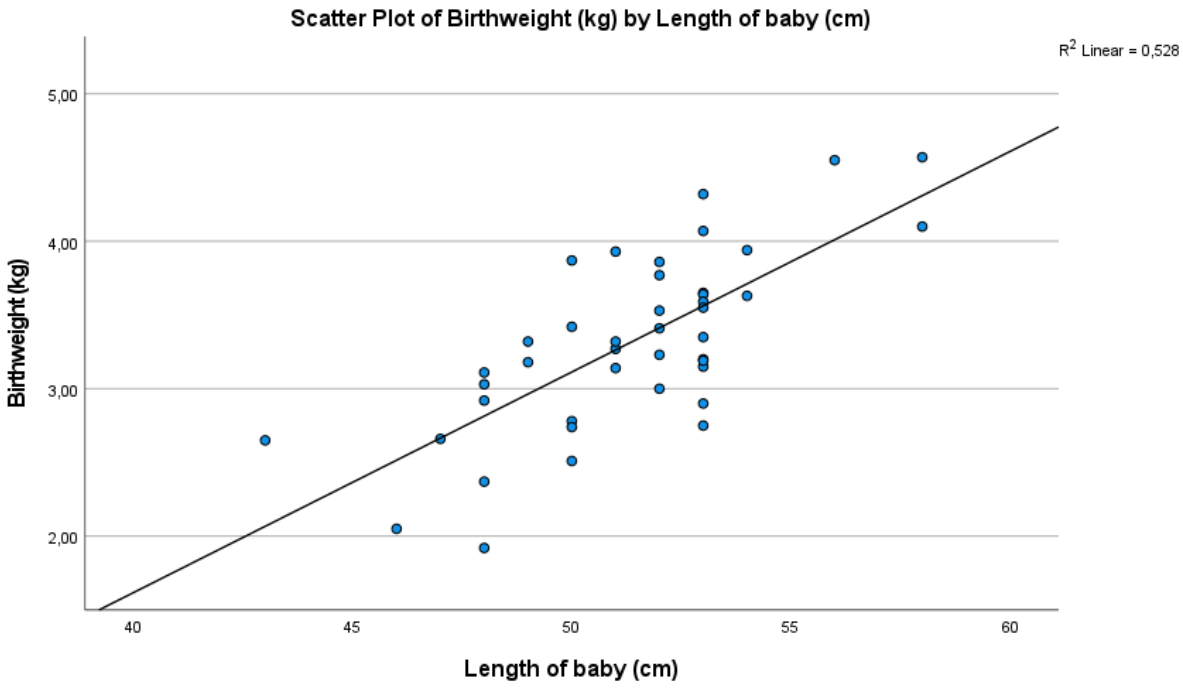
It is a low negative correlation.

Q19. If you wanted to see the effect of the length of a baby on birthweight, what would your independent variable be?

- **Length of baby**
- **Birthweight**

Length of Baby.

Q20. In statistics, when creating a scatterplot, it is a common practice to put the independent variable on the x-axis and the dependent variable on the y-axis. With this in mind, create a scatterplot for the above case and provide the regression line. For homework submitted using MS Word, insert a picture of the scatterplot.



Q21. Is the relationship between the length of baby and birthweight linear?

- **Yes**
- **No**

Yes.

Q22. Justify the above choice.

From the graph we can see that the data looks similar to a line instead of a curve.

Q23. Is there any evidence to suggest that the birth weight, length of baby, and head circumference are related?

- **Yes**
- **No**

Yes.

Q24. Justify the above choice.

Using the Pearson product-moment correlation once all the variables are not dichotomous, they are a ratio variable or an interval variable, their measures are linear and they all have normal distribution, we get the correlation coefficient of 0.563 between head circumference and length of baby, 0.685 between head circumference and birthweight and 0.727 between length of baby and birthweight, which indicates respectively a moderate positive correlation for the first two and a high positive correlation for the last one.

Q25. Describe the above relationship in your own words and provide evidence for your claims.

The results from the Pearson product-moment correlation indicate that all variables were related at some degree with the correlation coefficient of 0.563 between head circumference and length of baby, 0.685 between head circumference and birthweight and 0.727 between length of baby and birthweight, but after conducting a partial correlation test using birthweight as a control variable the correlation coefficient between head circumference and length of baby dropped to 0.131, which indicates that although we have a strong relationship between birthweight and head circumference, and birthweight and length of baby, the relationship between head circumference and length of baby is a negligible correlation or there isn't a relationship between these two variables.

Correlations

		Birthweight (kg)	Length of baby (cm)
Length of baby (cm)	Pearson Correlation	,727**	
	Sig. (2-tailed)	<,001	
	N	42	
Head circumference (cm)	Pearson Correlation	,685**	,563**
	Sig. (2-tailed)	<,001	<,001
	N	42	42

**. Correlation is significant at the 0.01 level (2-tailed).

Correlations

Control Variables			Head circumference (cm)	Length of baby (cm)
Birthweight (kg)	Head circumference (cm)	Correlation	1,000	,131
		Significance (2-tailed)	.	,414
		df	0	39
	Length of baby (cm)	Correlation	,131	1,000
		Significance (2-tailed)	,414	.
		df	39	0