|  |  |
| --- | --- |
| **HOMEWORK -W1** | **Isabel Lecourtois  07/21/2023** |

Q1. What is the mean birth weight for babies of non-smoking mothers?

A=3.510

Q2. What is the mean birth weight for babies of smoking mothers?

A=3.134

Q3. What is the mean head circumference for babies of non-smoking mothers?

A=35.05

Q4. What is the mean gestational age at birth for babies of smoking mothers?

A= 38.95

Q5. What is the maximum head circumference for babies of non-smoking mothers?

A=39.00

Q6. What is the minimum gestational age at birth for babies of smoking mothers?

A=33.00

Q7. Based on the dataset you have, out of the two, which one would be a better bet:

* Pregnancy period in smoking mothers is shorter.
* Pregnancy period in non-smoking mothers is shorter.

Q8. Justify the above choice in a few words.

> summary(df\_nsmoker$Gestation)

Min. 1st Qu. Median Mean 3rd Qu. Max.

33.00 38.00 40.00 39.45 41.00 44.00

> summary(df\_smoker$Gestation)

Min. 1st Qu. Median Mean 3rd Qu. Max.

33.00 38.00 39.00 38.95 40.00 45.00

Pregnancy in non-smoking women is longer based on the mean, median and 3rd Qu. results, which show an increase of 1.00 compared to pregnancy in smoking women.

Q9. What is the baby birth weight range for babies of smoking mothers?

> summary(df\_smoker$Birthweight)

Min. 1st Qu. Median Mean 3rd Qu. Max.

1.920 2.743 3.185 3.134 3.545 4.570

R= 4.570-1.920 = 2.65

Q10. In your own words describe what the value of the above range for baby's birthweight tells us about smoking versus non-smoking mothers?

> summary(df\_nsmoker$Birthweight)

Min. 1st Qu. Median Mean 3rd Qu. Max.

2.650 3.140 3.385 3.510 3.933 4.550

R= 4.55-2.650= 1.90

In the case of smoking women, the range is greater than in non-smoking women, which may indicate in a very basic way that there is greater dispersion or extreme values.

Q11. Are head circumference data for babies of smoking mothers normally distributed?

p-value > 0.05: There is no significant evidence to reject the null hypothesis which means that they do follow a normal distribution.

Q12. What is the significance value for the above on the Shapiro-Wilk test?

> shapiro.test(df\_smoker$Headcirc)

Shapiro-Wilk normality test

data: df\_smoker$Headcirc

W = 0.95365, p-value = 0.3724

Q13. What is the standard score (Z-score) for head circumference of 35.05 (X=35.05) in non-smoking mothers?

> print(z\_score)

[1] 0.2632275

Q14. How are birth weight data of non-smoking mothers skewed?

> skewness(df\_nsmoker$Birthweight)

[1] 0.3333708

Q15. Are birth weight data for babies of smoking mothers normally distributed?

Yes, they are.

Q16. What is the significance value for the above on the Shapiro-Wilk test?

> shapiro.test(df\_smoker$Birthweight)

Shapiro-Wilk normality test

data: df\_smoker$Birthweight

W = 0.98242, p-value = 0.9495

Q17. Based on the dataset you have; how confident can you be in saying that a baby's birth weight will be +/- 1 standard deviation from the mean?

> shapiro.test(df$Birthweight)

Shapiro-Wilk normality test

data: df$Birthweight

W = 0.9899, p-value = 0.9685

Since the birth weight data is normally distributed, I can say that there’s approximately 69% chance that the baby’s weight will be +/- 1 standard deviation from the mean.

Q18. Based on the dataset you have, what is the probability that the birth weight for a baby of a smoking mother will be less than 4.2 kg?

> print(probability\_less\_than\_x)

[1] 0.9543497

Q19. Are data for length of baby of non-smoking mothers normally distributed?

Yes, they are.

Q20. What is the significance value for the above on the Shapiro-Wilk test?

Shapiro-Wilk normality test

data: df\_nsmoker$Length

W = 0.91225, p-value = 0.07037

Q21. What is the standard score for the length of a baby of 48.5cm for non-smoking mothers?

> print(z\_score)

[1] -0.7403141

Q22. Based on the dataset you have, what is the probability that the length of baby for non-smoking mothers will be more than 55 cm?

> print(probability\_greater\_than\_xl)

[1] 0.104352