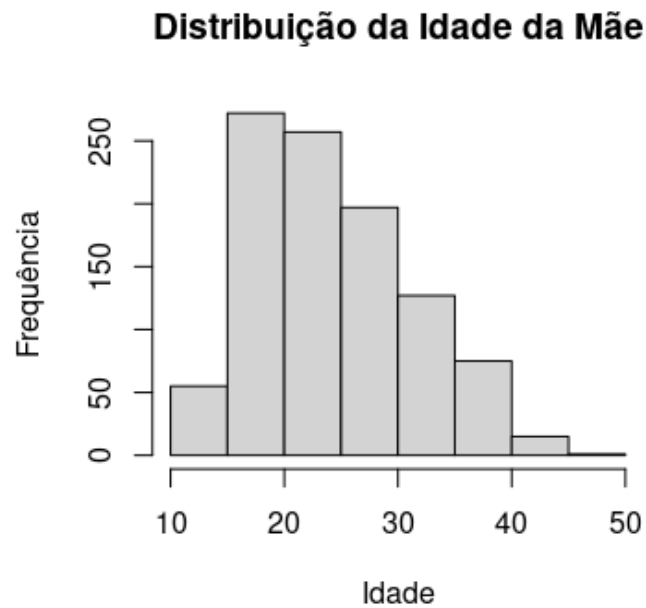


Conjunto de Dados escolhido: Nascidos Vivos 2021

Neste documento, será analisado como a idade da mãe pode estar relacionada com o peso do bebê recém nascido, e verificar se essa diferença é estatisticamente relevante.

Primeiro, faremos um histograma com a idade da mãe, e checamos sua normalidade com o teste do Shapiro:

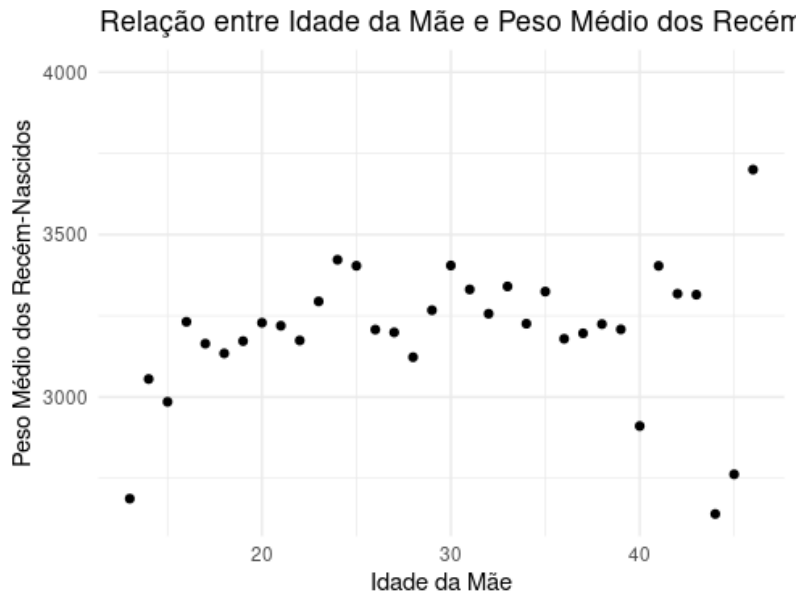


Resultado do teste de normalidade:
 $W = 0.96047$, $p\text{-value} = 8.456e-16$

Com base no valor p extremamente baixo, tem-se evidências estatísticas suficientes para rejeitar a hipótese nula de que a variável Idade Mãe segue uma distribuição normal.

A seguir, temos o cálculo da média do peso do bebe para cada idade da mãe representados em um gráfico de dispersão para mostrar a relação entre a idade das mães e o peso médio dos recém-nascidos.

Cada ponto no gráfico representa uma idade da mãe com a média correspondente do peso dos recém-nascidos.



Visualmente parece ser possível notar um aumento no peso dos recém nascidos com a idade da mãe, vamos tratar mais os dados para verificar se há uma diferença estatisticamente relevante. Utilizaremos 2 testes:

O Teste de Correlação de Pearson, para verificar se existe uma correlação linear. E o Teste de Correlação de Spearman, para verificar se existe uma correlação monotônica.

Aplicando o teste de correlação de Pearson, obtemos os seguintes valores:

Pearson's product-moment correlation

```
data: tabela_nascidos_2021$IDADEMAE and tabela_nascidos_2021$PESO
t = 2.2866, df = 996, p-value = 0.02243
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.01025485 0.13371961
sample estimates:
      cor
0.07226406
```

Analisando todos os dados individualmente:

- **Estatística de Teste (t):** O valor de t é 2.2866. Isso representa a magnitude da diferença entre a estimativa da correlação e zero, padronizada pelo erro padrão da estimativa. Valores mais altos indicam uma evidência mais forte contra a hipótese nula.
- **Graus de Liberdade (df):** O número de graus de liberdade associado ao teste. Neste caso, são 996 graus de liberdade.
- **Valor p (p-value):** O valor p é 0.02243. Este é o resultado central do teste. Ele representa a probabilidade de observar uma correlação. Um valor p pequeno (geralmente menor que 0,05) sugere que há evidência estatística para rejeitar a hipótese nula.
- **Intervalo de Confiança (95%):** O intervalo de confiança para a correlação é dado como (0.01025485, 0.13371961). Isso indica o intervalo dentro do qual a verdadeira correlação populacional é esperada para estar com 95% de confiança.
- **Estimativas da Amostra:** A estimativa da correlação amostral é 0.07226406. Isso é uma medida da força e direção da relação linear entre as variáveis. Neste caso, a correlação é positiva, indicando que à medida que a idade da mãe aumenta, o peso do recém-nascido também tende a aumentar.

Em resumo, o resultado sugere que há uma correlação estatisticamente significativa entre a idade da mãe e o peso do recém-nascido no conjunto de dados analisado. No entanto, a correlação é relativamente fraca, conforme indicado pelo tamanho do coeficiente de correlação (0.07226406).

Aplicando o teste de correlação de Spearman, obtemos os seguintes valores:

```
Spearman's rank correlation rho

data:  tabela_nascidos_2021$IDADEMAE and tabela_nascidos_20
21$PESO
S = 146752198, p-value = 0.000301
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.1141816
```

Analisando todos os dados individualmente:

- **Estatística de Teste (S):** O valor de S é 146752198. Essa é a soma dos quadrados das diferenças entre as classificações observadas e esperadas. Valores maiores indicam uma correlação de classificação mais forte.
- **Valor p (p-value):** O valor p é 0.000301. Isso representa a probabilidade de observar uma correlação de classificação tão forte ou mais forte se a verdadeira correlação de classificação na população fosse zero (hipótese nula). Um valor p pequeno (geralmente menor que 0,05) sugere que há evidência estatística para rejeitar a hipótese nula.
- **Hipótese Alternativa:** A hipótese alternativa afirma que a verdadeira correlação de classificação (ρ) não é igual a zero.
- **Estimativa da Amostra (ρ):** A estimativa da correlação de classificação na amostra é 0.1141816. Este valor ρ indica a força e direção da relação de classificação entre as variáveis. Neste caso, a correlação é positiva, indicando que há uma tendência de que, à medida que a idade da mãe aumenta, o peso do recém-nascido também aumente.

Em resumo, o resultado sugere que há uma correlação de classificação estatisticamente significativa entre a idade da mãe e o peso do recém-nascido no conjunto de dados analisado. O valor ρ é maior do que o coeficiente de correlação de Pearson calculado anteriormente, indicando uma relação de classificação um pouco mais forte entre essas variáveis. O valor p pequeno (0.000301) fornece evidência estatística para rejeitar a hipótese nula de que não há correlação de classificação.