



**Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Científica
Departamento de Estatística**

Relatório Trabalho de ME610

**Eliane Ramos de Siqueira RA:155233
Guilherme Pazian RA:160323**

Professor: Maurício Zevallos

Campinas-SP, 28 de Abril de 2017

1. Introdução

O objetivo deste trabalho é verificar esta opinião através da análise do peso da criança ao nascer

Os dados utilizados foram coletados através de entrevistas com as mães durante a gravidez e correspondem a um ano de estudo. Eles incluem 1236 nascimentos onde a criança sobreviveu pelo menos 28 dias. Além do peso das crianças, foram coletadas informações de outras variáveis que influenciam no peso das crianças.

Para este estudo estamos considerando como variável resposta, o peso da criança ao nascer.

2. Objetivo

Escreva o objetivo aqui

3. Metodologia

Escreva a metodologia aqui

4. Resultados

Escreva os resultados aqui

5. Discussão

Escreva a discussão aqui

5. Anexos

Coloque os anexos aqui

7. Referências

Escreva as referencias aqui

Limpando os dados

Excluimos as variáveis: Sexo (apenas masculino), Qtd_feto (apenas 5) e Vivo (apenas 1).

Na variável Cor_mae, temos uma observação com valor desconhecido (99)

Na variável Idade_mae temos duas observações desconhecidas (99)

Na variável Educacao_mae temos uma observação desconhecida (9)

Na variável Altura_mae temos 22 observações desconhecidas (99)

Na variável Peso_mae temos 36 observações desconhecidas (999)

Na variável Cor_pai, temos 5 observações com valor desconhecido (99)

Na variavel Idade_pai, temos 7 observações com valor desconhecido (99)
 Na variavel Educacao_pai, temos 13 observações com valor desconhecido (9)
 Na variavel Altura_pai temos 492 observações desconhecidas (99)
 Na variavel Peso_pai temos 499 observações desconhecidas (999)
 Na variavel Rendimento_anual temos 124 observações desconhecidas (98)
 Na variavel Fuma temos 10 observações desconhecidas (9)
 Na variavel Tempo_sem_fumar temos 9 observações desconhecidas (98) e 1 não perguntado(99)
 Na variavel Numero_cigarros temos 10 observações desconhecidas (98)

Tiramos as variaveis Peso_pai e Altura_pai devido a quantidade elevada de observações desconhecidas.

Variáveis disponiveis ainda:

Data_nasc
 Tempo_gestacao
 Peso
 Cor_mae
 Educacao_mae
 Peso_mae
 Peso_pai
 Altura_pai
 Estado_civil 2 desconhecidas(0)
 Fuma
 Numero_cigarros Numero_gestacoes Idade_mae Idade_pai Altura_mae Cor_pai Educacao_pai Rendimento_anual
 Tempo_sem_fumar

Convertendo unidade de medidas

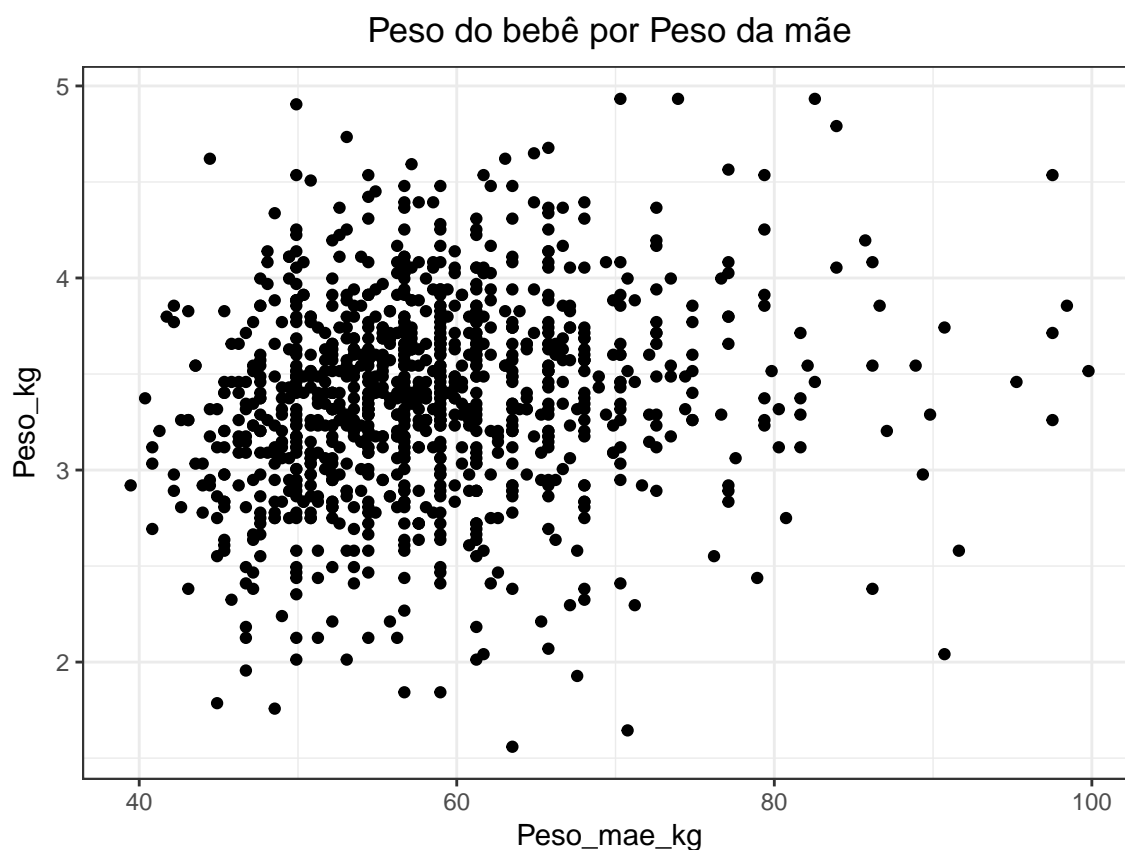
Frequencias

Varíavel Fuma	Frequência
Nunca	463
Fuma agora	401
Até a gravidez atual	79
Uma vez, não agora	91

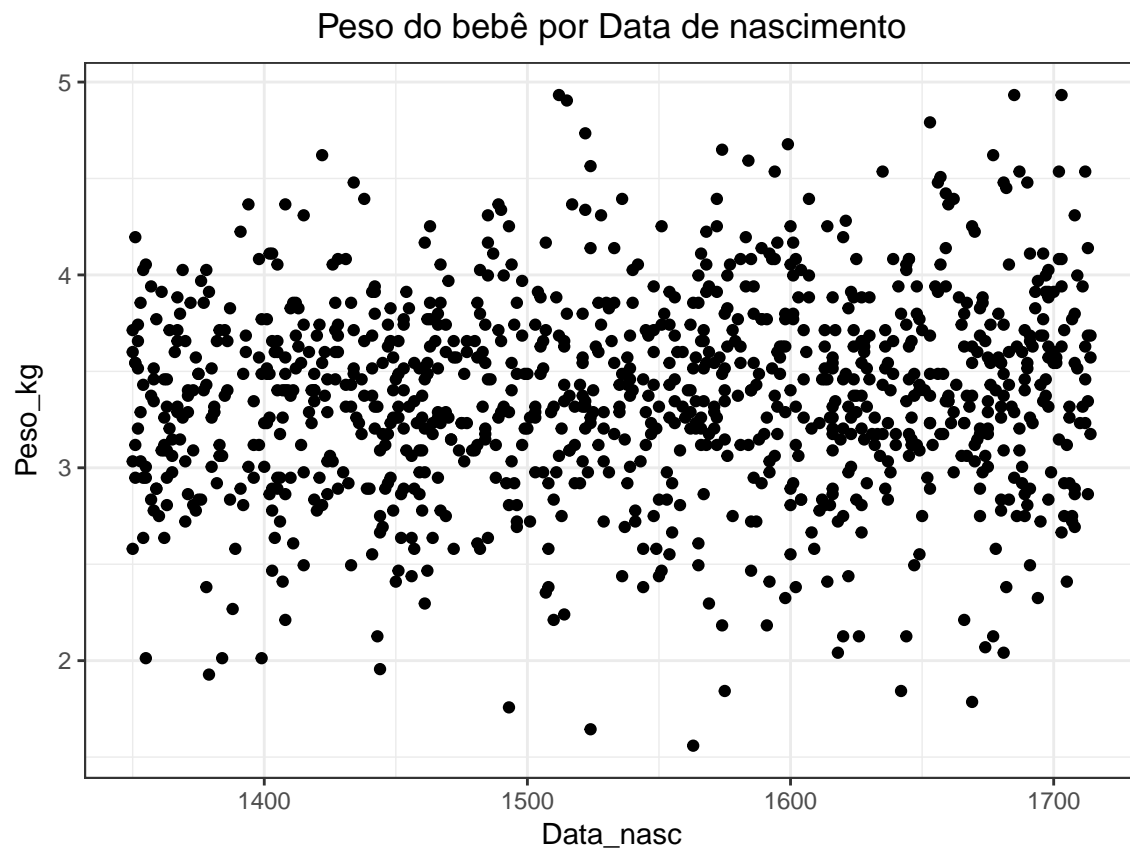
Análise Descritiva

Com base na visualização de análises descritivas (gráficos de dispersão e boxplots), observou-se as relações entre as variáveis explicativas com a variável resposta (peso do bebê ao nascer). Notou-se que existe pouca relação entre as variáveis explicativas “Altura_mae_cm”, “Idade_pai”, “Idade_mae” e “Data_nasc” em relação ao peso do bebê, ou seja, o peso do bebê parece não estar correlacionado com estas variáveis. Portanto optou-se em não incluir estas variáveis explicativas mesmo num modelo inicial. Em relação às demais variáveis explicativas, todas apresentaram indícios de uma possível correlação com o peso dos bebês, de maneira que a distribuição dos dados referentes à variável resposta muda conforme os valores observados nas variáveis explicativas. Nenhuma variável contínua/discreta aparentou ter uma relação não-linear com o peso dos bebês.

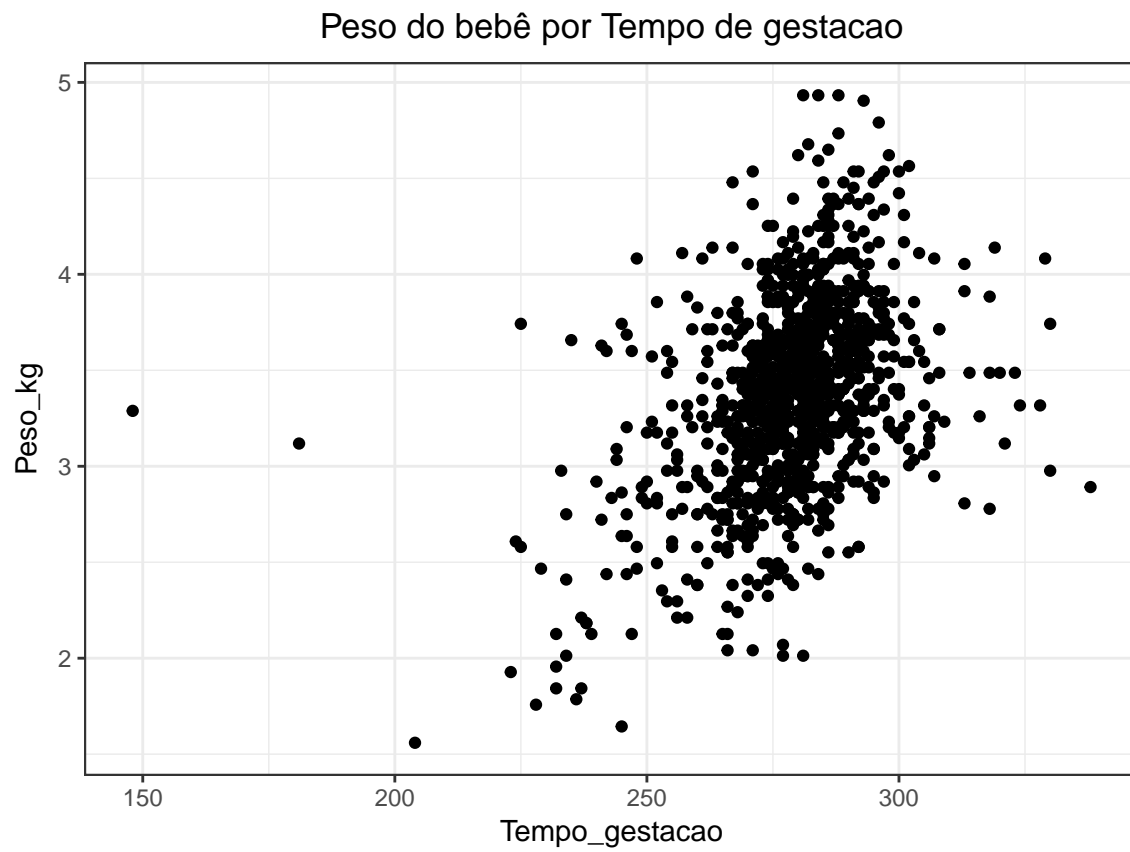
Gráfico de Dispersão



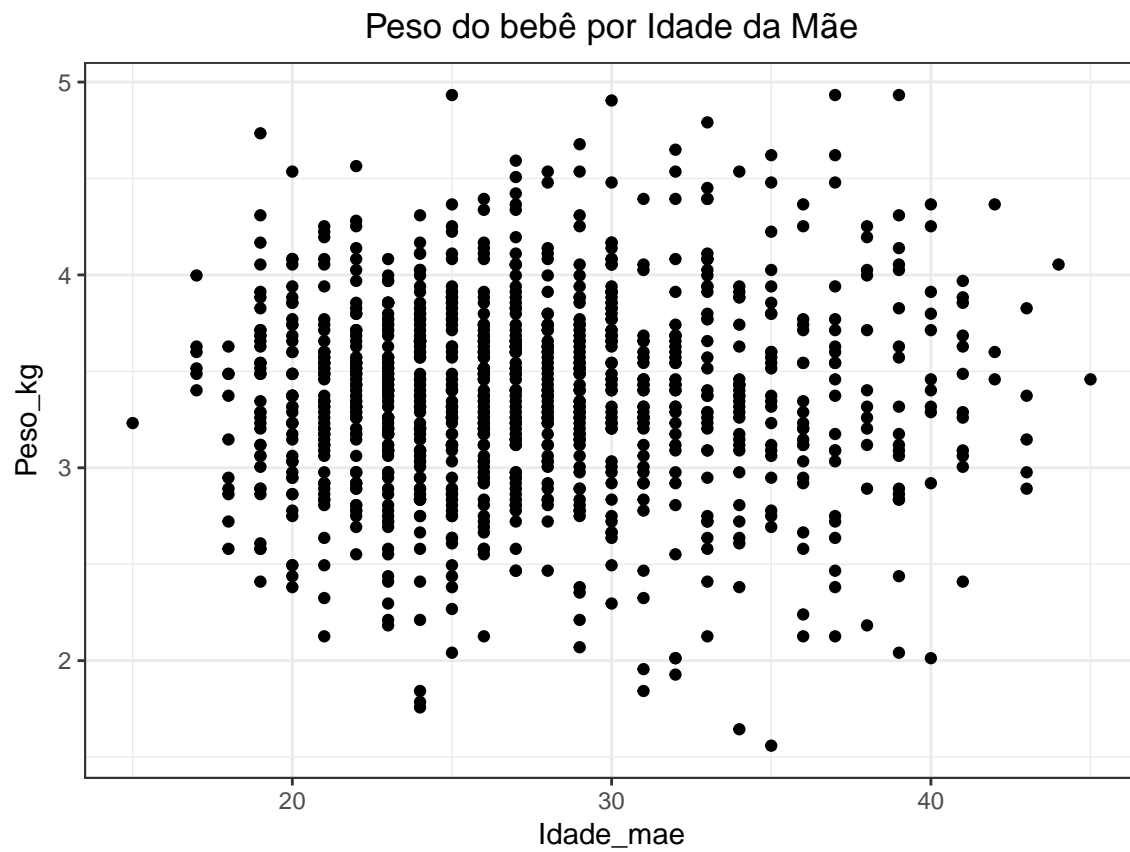
MOderado



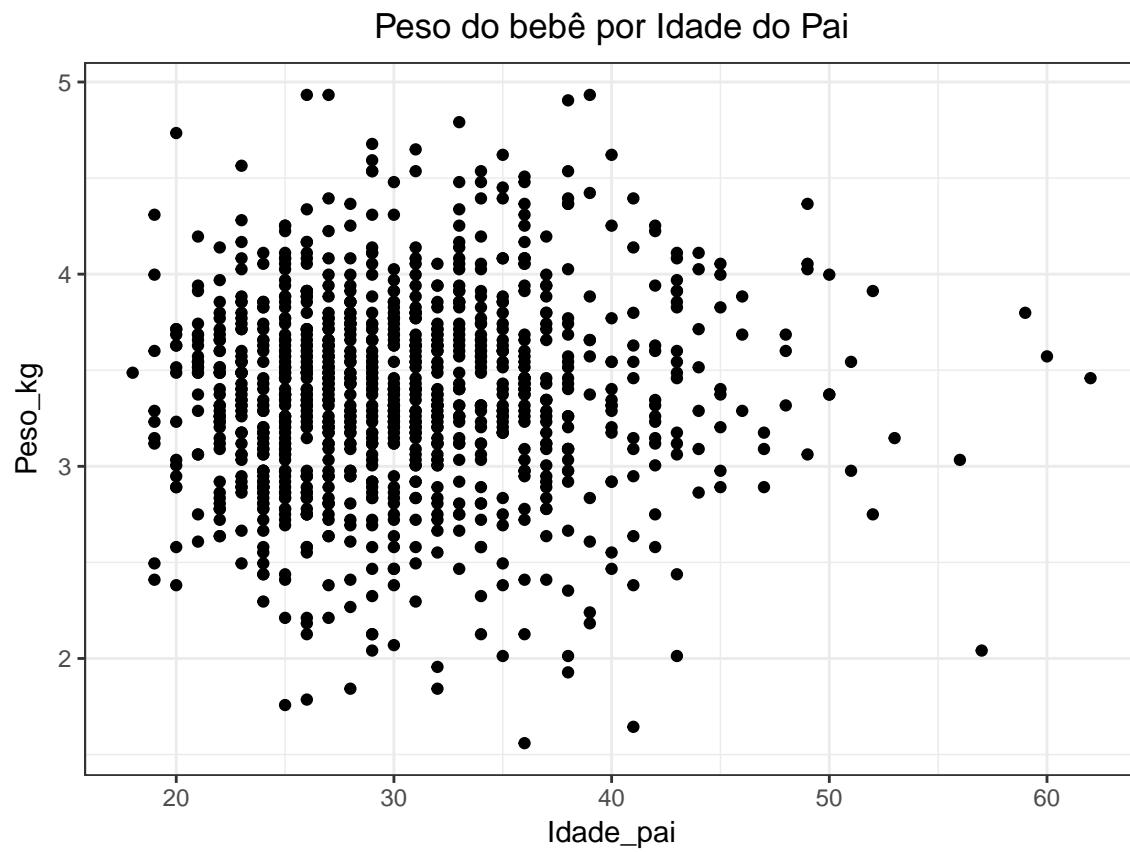
FRACA relaLinear



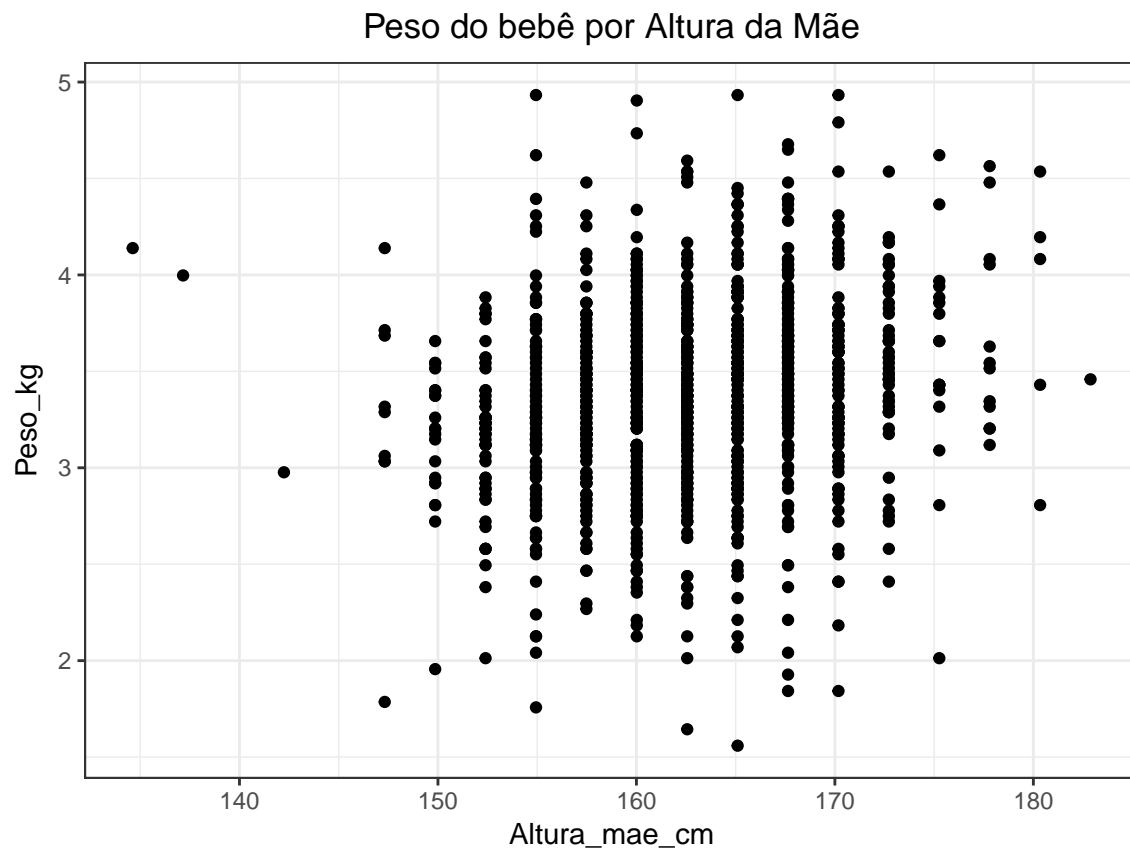
FORTE



Pouca relação linear

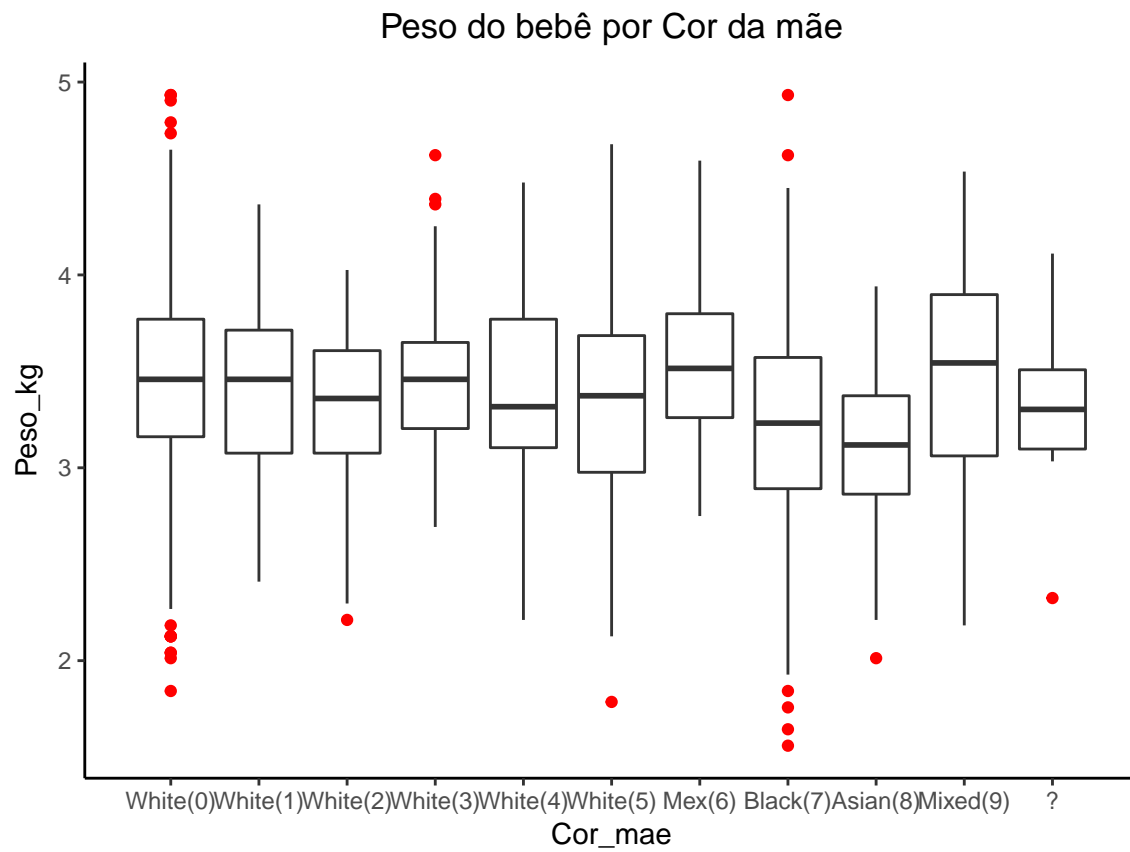


Pouca relação linear

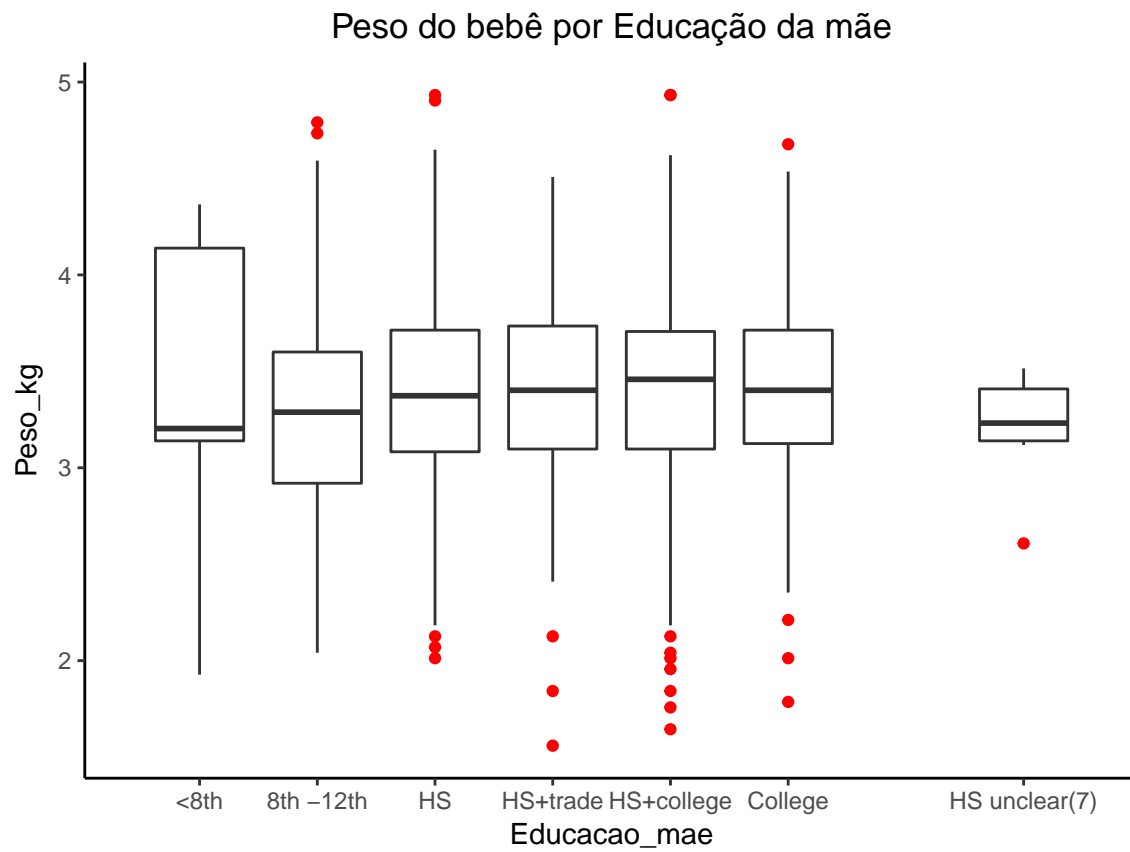


Pouca relação linear

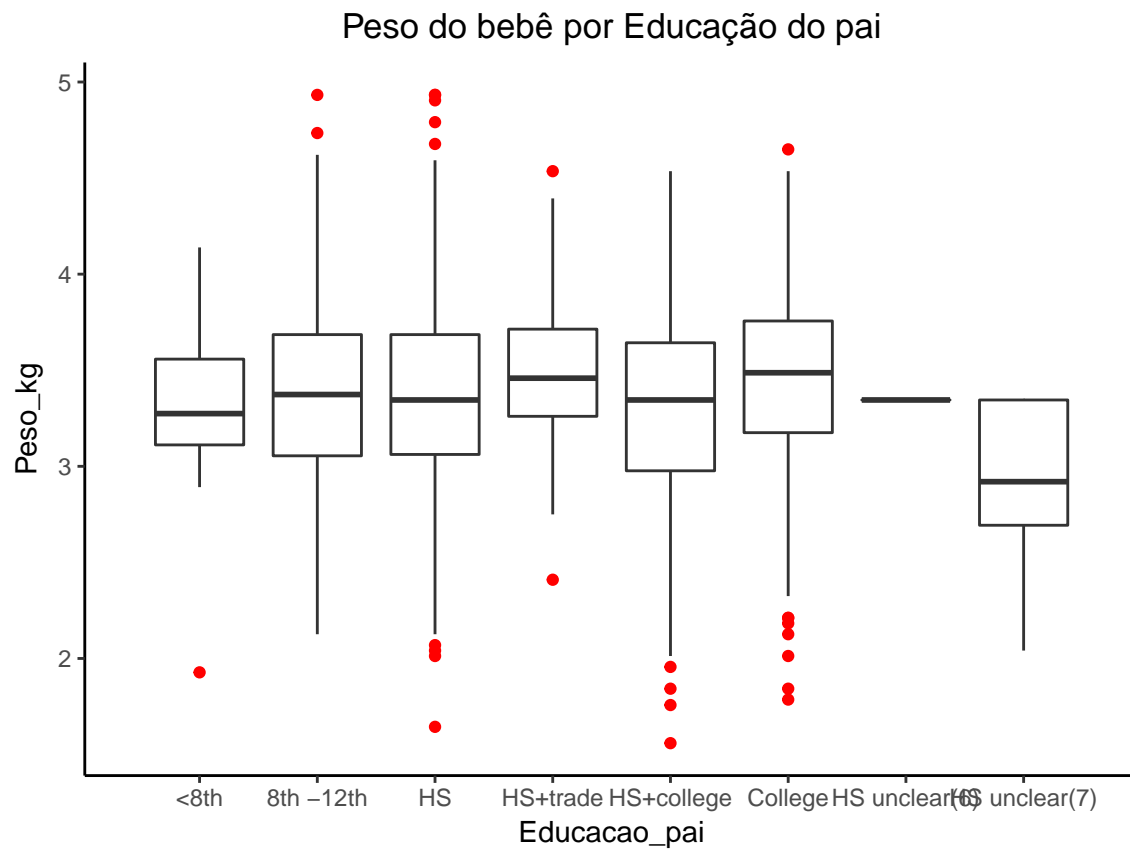
Boxplot para as variáveis categóricas e Dispersão para continua



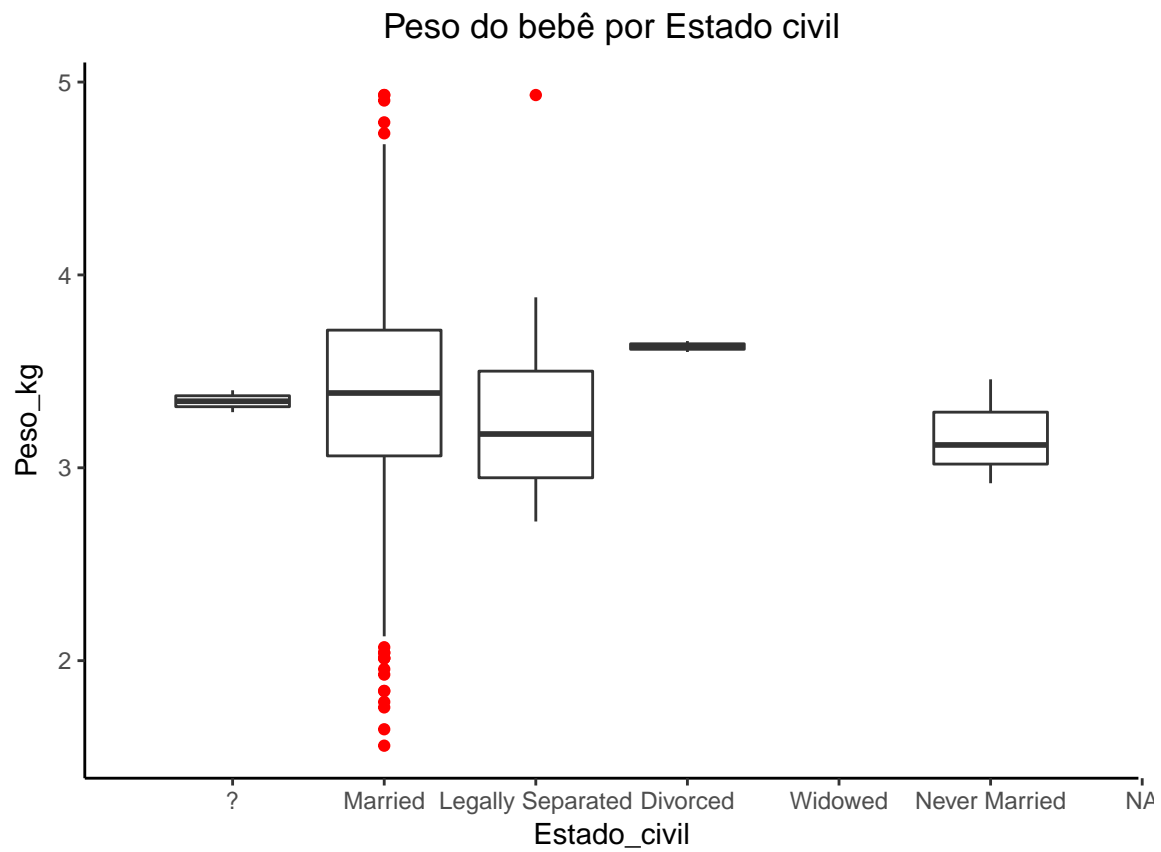
INFLUENCIA



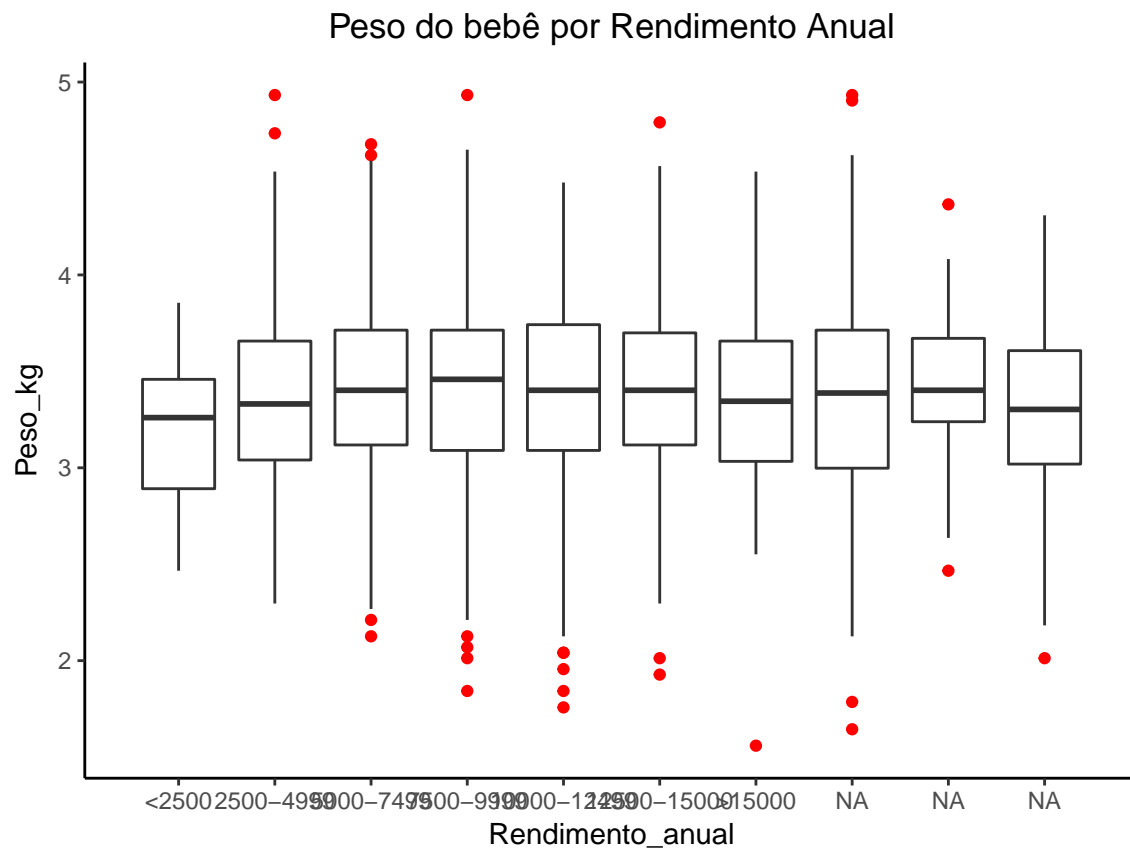
INFLUENCIA



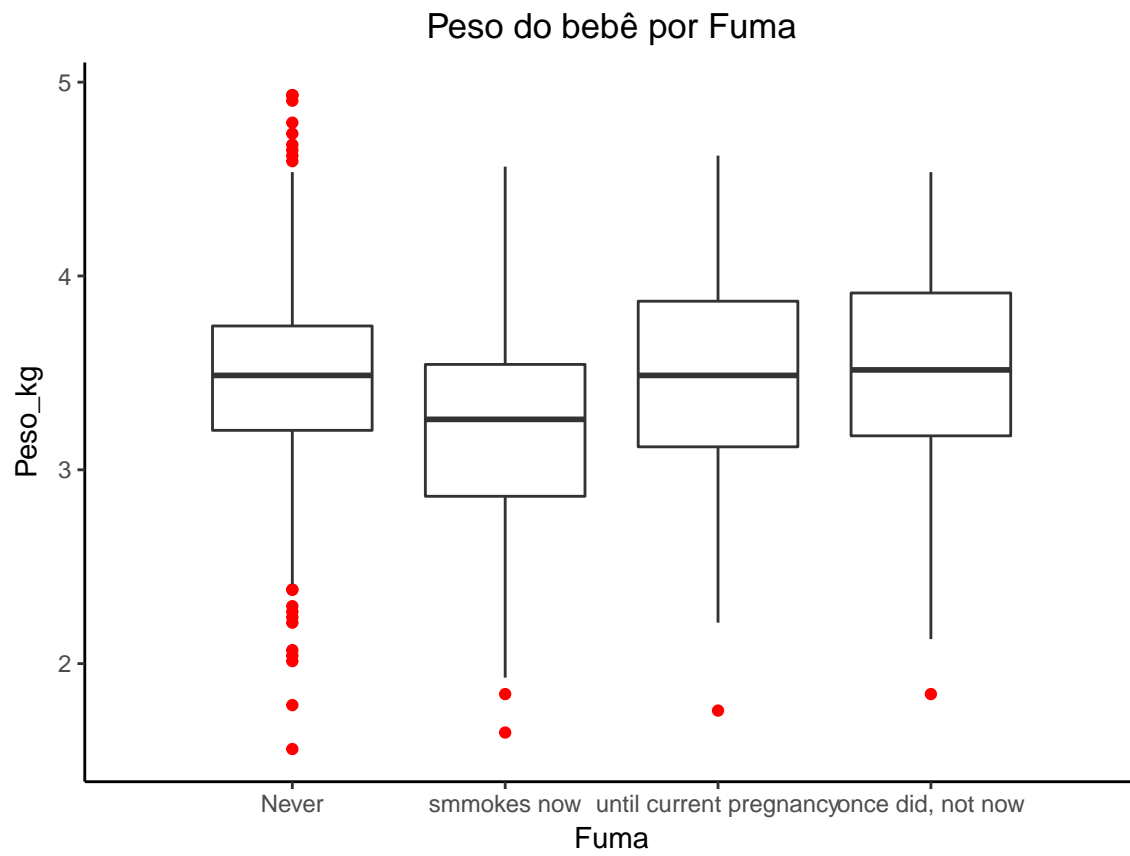
INFLUENCIA



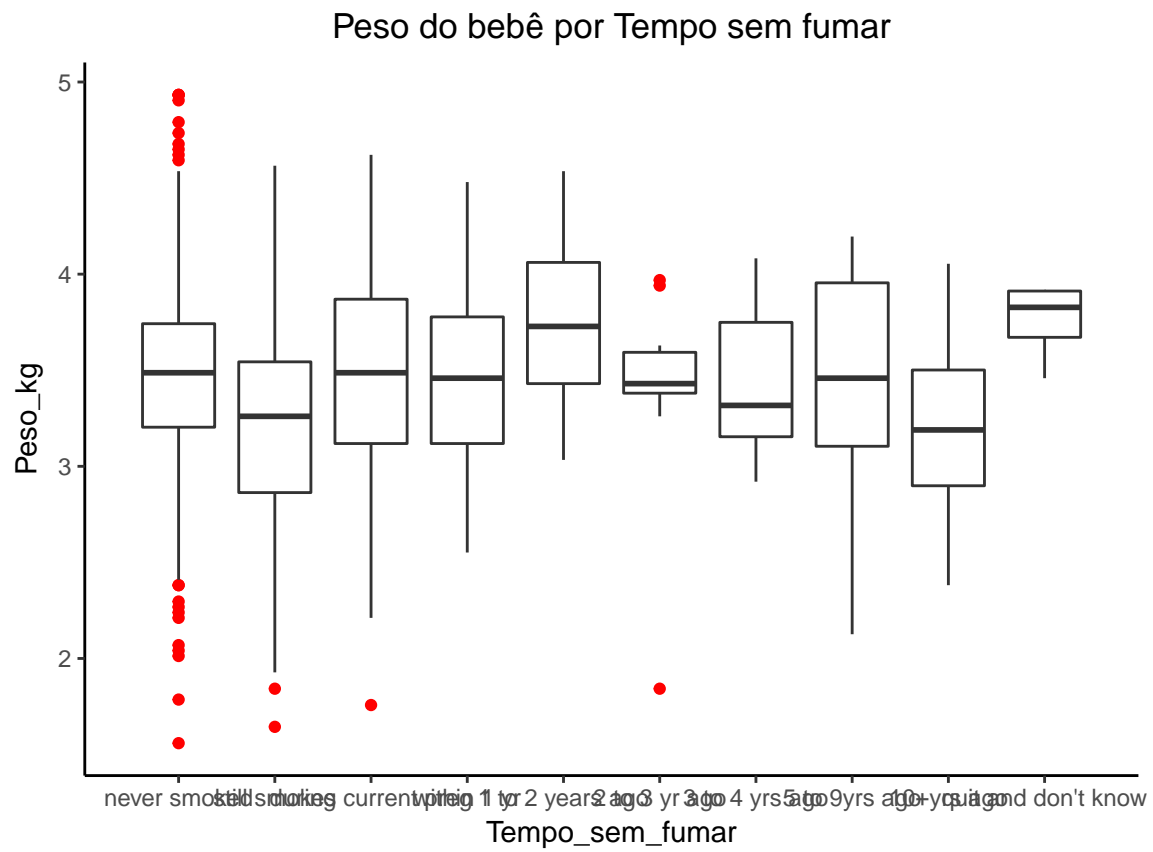
INFLUENCIA



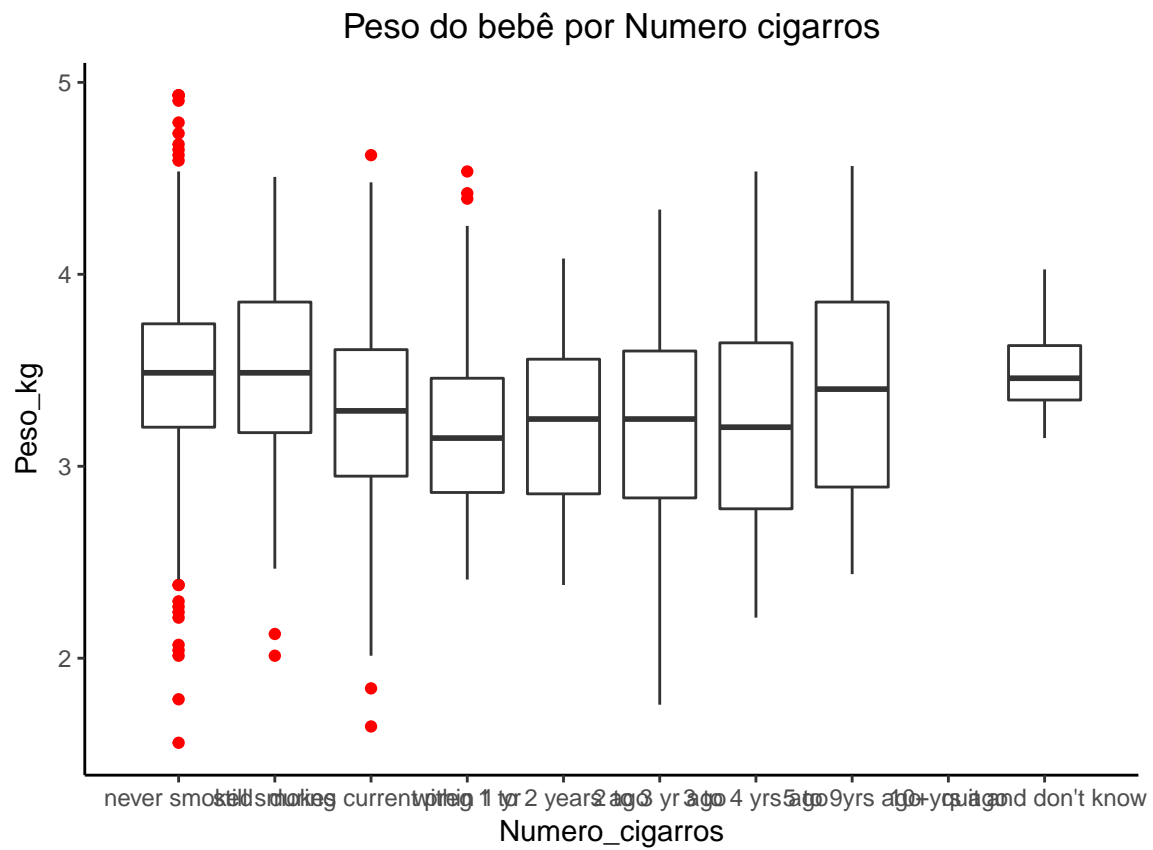
INFLUÊNCIA****



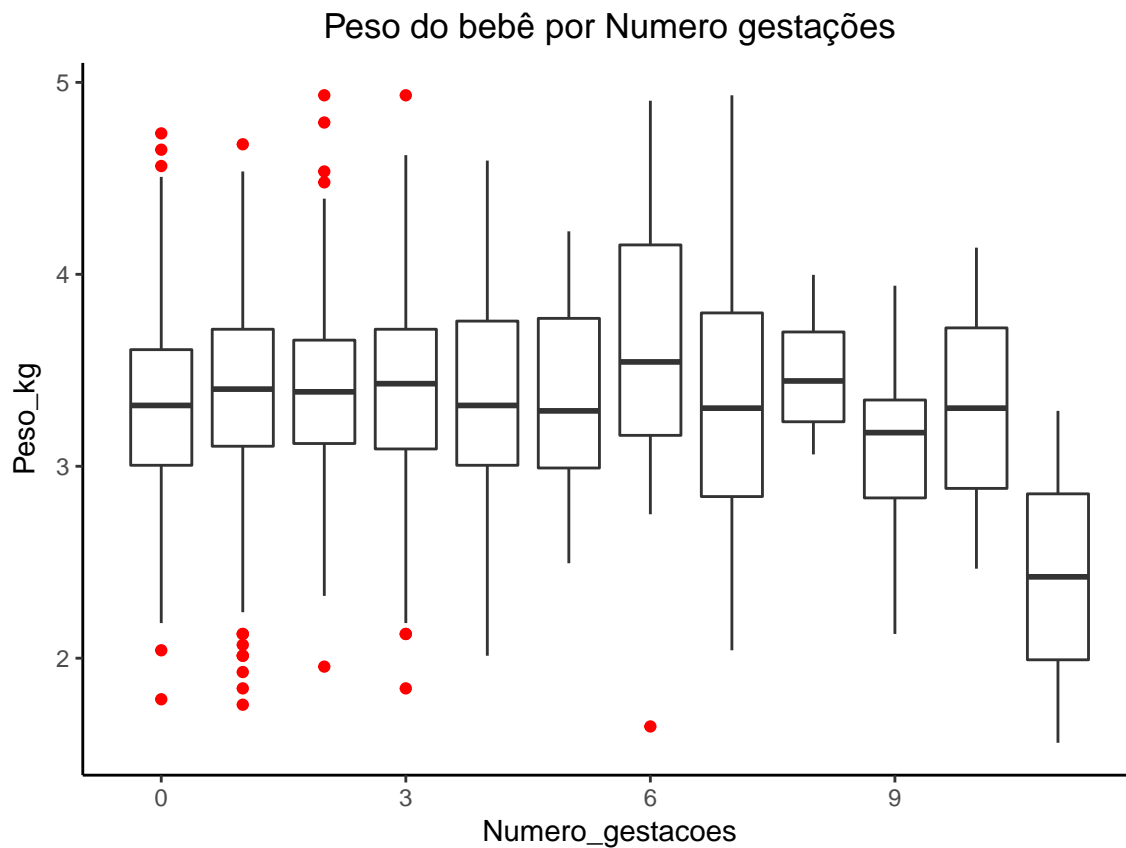
INFLUENCIA



INFLUENCIA



INFLUENCIA



INFLUENCIA

```
##
## Call:
## lm(formula = Peso_kg ~ Numero_gestacoes + Numero_cigarros + Tempo_sem_fumar +
##     Fuma + Rendimento_anual + Estado_civil + Educacao_pai + Educacao_mae +
##     Cor_mae + Peso_mae_kg + Tempo_gestacao, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37722 -0.26574 -0.01522  0.27571  1.26084
##
## Coefficients: (4 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0586076  0.4352717   0.135  0.89292
## Numero_gestacoes1 0.0580850  0.0396833   1.464  0.14360
## Numero_gestacoes2 0.1177377  0.0434704   2.708  0.00688 **
## Numero_gestacoes3 0.1383834  0.0488604   2.832  0.00472 **
```

## Numero_gestacoes4	0.1543007	0.0625778	2.466	0.01385 *
## Numero_gestacoes5	0.1338330	0.0729121	1.836	0.06673 .
## Numero_gestacoes6	0.2860307	0.0916527	3.121	0.00186 **
## Numero_gestacoes7	0.2888116	0.1398631	2.065	0.03919 *
## Numero_gestacoes8	0.3387117	0.2315378	1.463	0.14383
## Numero_gestacoes9	-0.0407316	0.1748052	-0.233	0.81580
## Numero_gestacoes10	0.3220887	0.3321328	0.970	0.33241
## Numero_gestacoes11	-0.2539511	0.3195925	-0.795	0.42704
## Numero_cigarros1	0.3250444	0.2335089	1.392	0.16424
## Numero_cigarros2	0.2677818	0.2353026	1.138	0.25539
## Numero_cigarros3	0.1309156	0.2385906	0.549	0.58334
## Numero_cigarros4	0.1099265	0.2524132	0.436	0.66330
## Numero_cigarros5	0.1148393	0.2347983	0.489	0.62488
## Numero_cigarros6	0.0889014	0.2311570	0.385	0.70062
## Numero_cigarros7	0.3106381	0.2567834	1.210	0.22668
## Numero_cigarros9	0.2717725	0.2563061	1.060	0.28925
## Tempo_sem_fumar1	-0.4316391	0.2314280	-1.865	0.06247 .
## Tempo_sem_fumar2	-0.2461489	0.2359696	-1.043	0.29714
## Tempo_sem_fumar3	-0.2826659	0.2469010	-1.145	0.25255
## Tempo_sem_fumar4	-0.0001411	0.2470891	-0.001	0.99954
## Tempo_sem_fumar5	-0.4331588	0.2715509	-1.595	0.11101
## Tempo_sem_fumar6	-0.2827085	0.2782441	-1.016	0.30986
## Tempo_sem_fumar7	-0.3982239	0.2558625	-1.556	0.11994
## Tempo_sem_fumar8	-0.5148121	0.2973397	-1.731	0.08370 .
## Tempo_sem_fumar9	NA	NA	NA	NA
## Fuma1	NA	NA	NA	NA
## Fuma2	NA	NA	NA	NA
## Fuma3	NA	NA	NA	NA
## Rendimento_anual1	0.1329547	0.0900658	1.476	0.14022
## Rendimento_anual2	0.1463628	0.0907939	1.612	0.10728
## Rendimento_anual3	0.1142922	0.0911980	1.253	0.21042
## Rendimento_anual4	0.1122543	0.0931710	1.205	0.22857
## Rendimento_anual5	0.1249592	0.0939074	1.331	0.18361
## Rendimento_anual6	0.1014644	0.1000525	1.014	0.31078
## Rendimento_anual7	0.0885252	0.0929443	0.952	0.34110
## Rendimento_anual8	0.1244386	0.1301481	0.956	0.33925
## Rendimento_anual9	-0.0619275	0.1322463	-0.468	0.63969
## Estado_civil1	-0.5011590	0.3251432	-1.541	0.12356
## Estado_civil2	-0.4368714	0.3505988	-1.246	0.21304
## Estado_civil3	-0.1878212	0.4544412	-0.413	0.67948

```

## Estado_civil5      -0.5426381  0.4199475  -1.292  0.19661
## Educacao_pai1      0.1123057  0.1011287   1.111  0.26705
## Educacao_pai2      0.1050759  0.0988281   1.063  0.28795
## Educacao_pai3      0.1587731  0.1279375   1.241  0.21490
## Educacao_pai4      0.0837987  0.1024635   0.818  0.41365
## Educacao_pai5      0.1151975  0.1044005   1.103  0.27012
## Educacao_pai6     -0.0909244  0.4550793  -0.200  0.84168
## Educacao_pai7     -0.3078015  0.2209748  -1.393  0.16396
## Educacao_mae1     -0.2024538  0.1784052  -1.135  0.25674
## Educacao_mae2     -0.1574916  0.1766910  -0.891  0.37297
## Educacao_mae3     -0.1521103  0.1865097  -0.816  0.41495
## Educacao_mae4     -0.1649667  0.1791586  -0.921  0.35739
## Educacao_mae5     -0.1558880  0.1820564  -0.856  0.39206
## Educacao_mae7     -0.2703484  0.2532178  -1.068  0.28594
## Cor_mae1          -0.0539097  0.0733380  -0.735  0.46247
## Cor_mae2          -0.1412936  0.0944784  -1.496  0.13511
## Cor_mae3           0.0142274  0.0665174   0.214  0.83068
## Cor_mae4           0.0062446  0.0685887   0.091  0.92748
## Cor_mae5          -0.0770093  0.0478548  -1.609  0.10789
## Cor_mae6           0.0759322  0.0890790   0.852  0.39420
## Cor_mae7          -0.2322562  0.0431071  -5.388  8.95e-08 ***
## Cor_mae8          -0.2169337  0.0809553  -2.680  0.00749 **
## Cor_mae9          -0.0371313  0.1060425  -0.350  0.72630
## Cor_mae10         -0.0764406  0.1455103  -0.525  0.59948
## Peso_mae_kg        0.0079415  0.0016027   4.955  8.53e-07 ***
## Tempo_gestacao     0.0121095  0.0009389  12.898  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4379 on 968 degrees of freedom
## Multiple R-squared:  0.3217, Adjusted R-squared:  0.2762
## F-statistic: 7.065 on 65 and 968 DF,  p-value: < 2.2e-16

```