

statistical_significance_port.R

r3019237

2025-03-05

```
Portuguese <- read.csv("Portuguese.csv")

model_G1 <- lm(G3 ~ G1, data = Portuguese)
model_G2 <- lm(G3 ~ G2, data = Portuguese)
model_absences <- lm(G3 ~ absences, data = Portuguese)
model_studytime <- lm(G3 ~ studytime, data = Portuguese)
model_failures <- lm(G3 ~ failures, data = Portuguese)
model_famrel <- lm(G3 ~ famrel, data = Portuguese)
model_freetime <- lm(G3 ~ freetime, data = Portuguese)
model_goout <- lm(G3 ~ goout, data = Portuguese)
model_age <- lm(G3 ~ age, data = Portuguese)

model_schoolsup <- aov(G3 ~ schoolsup, data = Portuguese)
model_famsup <- aov(G3 ~ famsup, data = Portuguese)
model_paid <- aov(G3 ~ paid, data = Portuguese)
model_activities <- aov(G3 ~ activities, data = Portuguese)
model_nursery <- aov(G3 ~ nursery, data = Portuguese)
model_higher <- aov(G3 ~ higher, data = Portuguese)
model_internet <- aov(G3 ~ internet, data = Portuguese)
model_romantic <- aov(G3 ~ romantic, data = Portuguese)
model_school <- aov(G3 ~ school, data = Portuguese)
model_sex <- aov(G3 ~ sex, data = Portuguese)
model_address <- aov(G3 ~ address, data = Portuguese)
model_famsize <- aov(G3 ~ famsize, data = Portuguese)
model_Pstatus <- aov(G3 ~ Pstatus, data = Portuguese)
model_Medu <- aov(G3 ~ Medu, data = Portuguese)
model_Fedu <- aov(G3 ~ Fedu, data = Portuguese)
model_Mjob <- aov(G3 ~ Mjob, data = Portuguese)
model_Fjob <- aov(G3 ~ Fjob, data = Portuguese)
model_reason <- aov(G3 ~ reason, data = Portuguese)
model_guardian <- aov(G3 ~ guardian, data = Portuguese)
model_Dalc <- aov(G3 ~ Dalc, data = Portuguese)
model_Walc <- aov(G3 ~ Walc, data = Portuguese)
model_health <- aov(G3 ~ health, data = Portuguese)

# Extract p-values from all models and store them
all_pvalues <- c(
  summary(model_G1)$coefficients[2, 4],
  summary(model_G2)$coefficients[2, 4],
  summary(model_absences)$coefficients[2, 4],
  summary(model_studytime)$coefficients[2, 4],
```

```

summary(model_failures)$coefficients[2, 4],
summary(model_famrel)$coefficients[2, 4],
summary(model_freetime)$coefficients[2, 4],
summary(model_goout)$coefficients[2, 4],
summary(model_age)$coefficients[2, 4],
summary(model_schoolsup)[[1]]$`Pr(>F)`[1],
summary(model_famsup)[[1]]$`Pr(>F)`[1],
summary(model_paid)[[1]]$`Pr(>F)`[1],
summary(model_activities)[[1]]$`Pr(>F)`[1],
summary(model_nursery)[[1]]$`Pr(>F)`[1],
summary(model_higher)[[1]]$`Pr(>F)`[1],
summary(model_internet)[[1]]$`Pr(>F)`[1],
summary(model_romantic)[[1]]$`Pr(>F)`[1],
summary(model_school)[[1]]$`Pr(>F)`[1],
summary(model_sex)[[1]]$`Pr(>F)`[1],
summary(model_address)[[1]]$`Pr(>F)`[1],
summary(model_famsize)[[1]]$`Pr(>F)`[1],
summary(model_Pstatus)[[1]]$`Pr(>F)`[1],
summary(model_Medu)[[1]]$`Pr(>F)`[1],
summary(model_Fedu)[[1]]$`Pr(>F)`[1],
summary(model_Mjob)[[1]]$`Pr(>F)`[1],
summary(model_Fjob)[[1]]$`Pr(>F)`[1],
summary(model_reason)[[1]]$`Pr(>F)`[1],
summary(model_guardian)[[1]]$`Pr(>F)`[1],
summary(model_Dalc)[[1]]$`Pr(>F)`[1],
summary(model_Walc)[[1]]$`Pr(>F)`[1],
summary(model_health)[[1]]$`Pr(>F)`[1]
)

# Adjust p-values using the Bonferroni correction
bonferroni_pvalues <- p.adjust(all_pvalues, method = "bonferroni")

# Create a data frame with variables and their Bonferroni-corrected p-values
variables <- c("G1", "G2", "absences", "studytime", "failures", "famrel", "freetime",
              "goout", "age", "schoolsup", "famsup", "paid", "activities", "nursery",
              "higher", "internet", "romantic", "school", "sex", "address", "famsize",
              "Pstatus", "Medu", "Fedu", "Mjob", "Fjob", "reason", "guardian",
              "Dalc", "Walc", "health")

# Combine the variables and Bonferroni p-values into a data frame
result <- data.frame(Variable = variables, Bonferroni_P_value = bonferroni_pvalues)

# Print the result
print(result)

```

```

##      Variable Bonferroni_P_value
## 1         G1      5.014470e-162
## 2         G2      1.749144e-261
## 3    absences      6.167643e-01
## 4   studytime      3.381507e-09
## 5    failures      5.995875e-24
## 6     famrel      1.000000e+00
## 7   freetime      5.386200e-02
## 8     goout      7.926661e-01

```

```
## 9      age      2.049756e-01
## 10 schoolsup  1.000000e+00
## 11      famsup  1.000000e+00
## 12      paid   1.000000e+00
## 13 activities 1.000000e+00
## 14      nursery 1.000000e+00
## 15      higher  1.084894e-16
## 16      internet 3.871644e-03
## 17      romantic 6.510766e-01
## 18      school  4.855217e-12
## 19      sex     3.042739e-02
## 20      address 5.468876e-04
## 21      famsize 1.000000e+00
## 22      Pstatus 1.000000e+00
## 23      Medu    1.783191e-08
## 24      Fedu    1.586121e-06
## 25      Mjob    2.574596e-04
## 26      Fjob    3.526647e-01
## 27      reason  4.159091e-05
## 28      guardian 1.000000e+00
## 29      Dalc    4.439311e-06
## 30      Walc    1.859843e-04
## 31      health  3.642120e-01
```

```
# Check the length of the variables and p-values to ensure they match
length(variables)
```

```
## [1] 31
```

```
length(bonferroni_pvalues)
```

```
## [1] 31
```