

statistical_significance_math.R

r3019237

2025-03-05

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

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

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

# 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      2.790443e-88
## 2         G2      2.363973e-146
## 3    absences      1.000000e+00
## 4   studytime      1.000000e+00
## 5    failures      4.543555e-12
## 6     famrel      1.000000e+00
## 7   freetime      1.000000e+00
## 8      goout      2.550990e-01
## 9       age      3.941459e-02

```

```
## 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    8.270805e-03
## 16      internet      1.000000e+00
## 17      romantic    3.010945e-01
## 18        school      1.000000e+00
## 19          sex      1.000000e+00
## 20      address      1.000000e+00
## 21      famsize      1.000000e+00
## 22      Pstatus      1.000000e+00
## 23        Medu      4.141933e-04
## 24        Fedu      7.377915e-02
## 25        Mjob      1.610372e-01
## 26        Fjob      1.000000e+00
## 27      reason      1.000000e+00
## 28    guardian      1.000000e+00
## 29        Dalc      1.000000e+00
## 30        Walc      1.000000e+00
## 31      health      1.000000e+00
```

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

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

```
length(bonferroni_pvalues)
```

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