

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT NAME:data analysis

PRACTICAL 12

AIM: Generating correlation matrices using `cor()` ®

PURVI ROSHAN KARKERA S086

PURVI ROSHAN KARKERA

S086

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT NAME:data analysis

RStudio

```
File Edit Code View Plots Session Build Debug Profile Tools Help
Untitled 1 AI Impact on Jobs 2030 5086 Prac 10.R 5086 Prac 11.R 5086 Prac 12.R
library(ggplot2)
library(dplyr)
library(readr)
ai_data <- read.csv("./5086 PRAC 10 TO 12/AI_Impact_on_Jobs_2030.csv")
ai_subset <- head(ai_data, 20)
correlation_data <- ai_subset[, c("Average_Salary", "Automation_Probability_2030")]
cor_matrix <- cor(correlation_data, use = "complete.obs")
print("Correlation Matrix:")
[1] "Correlation Matrix"
print(cor_matrix)
      Average_Salary Automation_Probability_2030
Average_Salary 1.0000000 0.1043887
Automation_Probability_2030 0.1043887 1.0000000
plot_scatter <- ggplot(ai_subset, aes(x = Average_Salary, y = Automation_Probability_2030)) +
  geom_point(color = "darkblue", size = 3) +
  geom_smooth(method = "lm", color = "red", se = FALSE) +
  theme_minimal() +
  labs(title = "Relationship: Average Salary vs Automation Probability",
       subtitle = "Analysis of the first 20 rows",
       x = "Average Salary",
       y = "Automation Probability (2030)")
heatmap(cor_matrix,
        main = "Correlation Matrix Heatmap",
        col = colorRampPalette(c("blue", "white", "red"))(20),
        sym = TRUE)
print(plot_scatter)
geom_smooth() using formula = 'y ~ x'
|
```

Relationship: Average Salary vs Automation Probability
Analysis of the first 20 rows

Average_Salary Automation_Probability_2030

RStudio

```
File Edit Code View Plots Session Build Debug Profile Tools Help
Untitled 1 AI Impact on Jobs 2030 5086 Prac 10.R 5086 Prac 11.R 5086 Prac 12.R
library(ggplot2)
library(dplyr)
library(readr)
ai_data <- read.csv("./5086 PRAC 10 TO 12/AI_Impact_on_Jobs_2030.csv")
ai_subset <- head(ai_data, 20)
correlation_data <- ai_subset[, c("Average_Salary", "Automation_Probability_2030")]
cor_matrix <- cor(correlation_data, use = "complete.obs")
print("Correlation Matrix:")
[1] "Correlation Matrix"
print(cor_matrix)
      Average_Salary Automation_Probability_2030
Average_Salary 1.0000000 0.1043887
Automation_Probability_2030 0.1043887 1.0000000
plot_scatter <- ggplot(ai_subset, aes(x = Average_Salary, y = Automation_Probability_2030)) +
  geom_point(color = "darkblue", size = 3) +
  geom_smooth(method = "lm", color = "red", se = FALSE) +
  theme_minimal() +
  labs(title = "Relationship: Average Salary vs Automation Probability",
       subtitle = "Analysis of the first 20 rows",
       x = "Average Salary",
       y = "Automation Probability (2030)")
heatmap(cor_matrix,
        main = "Correlation Matrix Heatmap",
        col = colorRampPalette(c("blue", "white", "red"))(20),
        sym = TRUE)
print(plot_scatter)
geom_smooth() using formula = 'y ~ x'
|
```

Correlation Matrix Heatmap

Average_Salary Automation_Probability_2030

Type here to search

24°C Mostly sunny 10:30 08-01-2026

PURVI ROSHAN KARKERA

S086

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT NAME:data analysis

PURVI ROSHAN KARKERA S086

PURVI ROSHAN KARKERA

S086