

Análise Multivariada II

Lista I

Breno Cauã Rodrigues da Silva

1 Introdução

Os dados da presente atividade (**Lista I**) foram disponibilizados pela Profa. Dra. Marinalva Cardoso Maciel na disciplina de **Análise Multivariada II**. O conjunto de dados representa as medições dos níveis de glicose no sangue de 50 mulheres, em três semanas consecutivas. As variáveis y_1, y_2, y_3 referem-se às medições em jejum, enquanto x_1, x_2, x_3 dizem respeito às medições realizadas 1 hora após ingestão de açúcar.

O objetivo é avaliar, por meio de técnicas multivariadas, se há diferença significativa entre os níveis de glicose nos três momentos distintos e entre os dois estados fisiológicos (jejum e pós-ingestão).

2 Estruturação e Pré-processamento dos Dados

O conjunto de dados original, apresentado no ficheiro PDF “LISTA I.pdf”, passou por uma etapa inicial de estruturação e ajuste utilizando o ambiente [VS Code](#). Este procedimento foi fundamental para organizar as medições de glicose (y_1, y_2, y_3 para jejum; x_1, x_2, x_3 para pós-ingestão) num formato tabular coeso, resultando no ficheiro `df.csv` que é utilizado nas análises subsequentes neste documento.

Especificamente, os dados foram dispostos de forma que cada linha representasse uma observação para uma mulher num determinado momento e condição (jejum ou ingestão), facilitando a importação e manipulação no [R](#). A variável **condition** foi introduzida para distinguir entre as medições em jejum e após a ingestão de açúcar, e as medições semanais foram consolidadas nas colunas **S1**, **S2**, e **S3**. Este pré-processamento no [VS Code](#) garantiu que os dados estivessem prontos para as análises exploratórias e inferenciais detalhadas a seguir.

3 Análise Descritiva dos Dados

Com o conjunto de dados organizado e bem estruturado podemos começar a fazer as análises descritivas e testes de normalidade multivariada.

Começemos com as medidas de resumo de cada variável do conjunto de dados. Veja a Tabela 1 e Tabela 2 que apresentam, respectivamente, medidas descritivas dos *níveis de glicose no sangue* por semana para as mulheres em jejum e mulheres com 1 hora após a ingestão de açúcar.

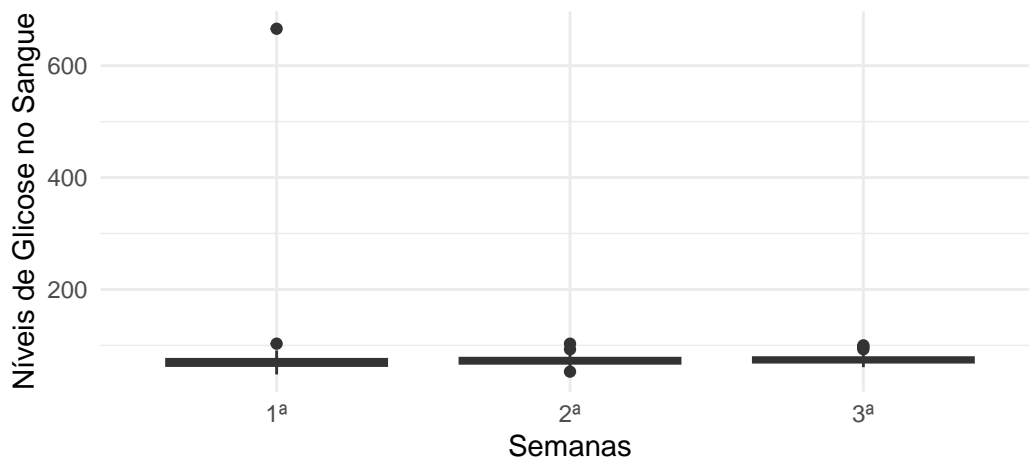
Tabela 1: Medidas de Resumo para os Níveis de Glicose no Sangue das Mulheres em Jejum.

	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Desvio Padrão
S1	48	64	70	82,39	75	666	85,678
S2	53	68	74	73,59	77	103	8,636
S3	61	70	74	75,22	78	100	8,773

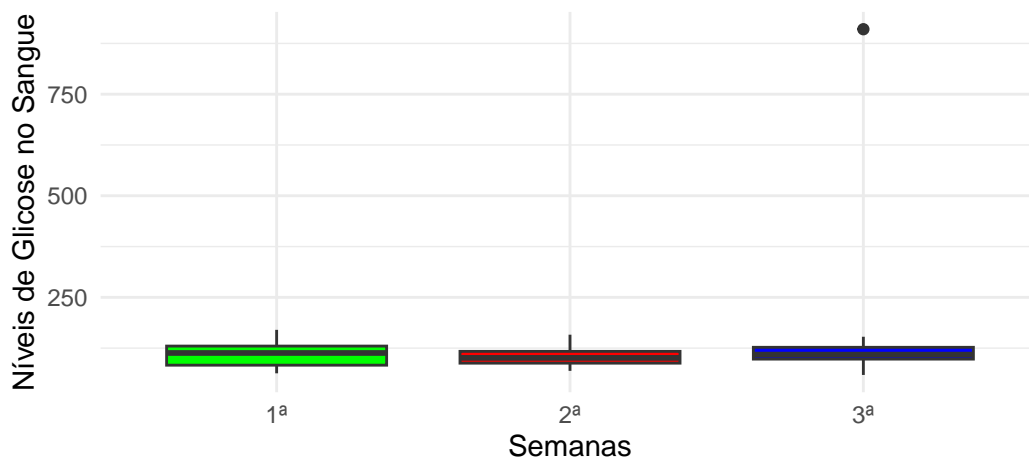
Tabela 2: Medidas de Resumo para os Níveis de Glicose no Sangue das Mulheres com 1 Hora após a Ingestão de Açúcar.

	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Desvio Padrão
S1	63	83	113	110,6	130	170	27,91
S2	69	88	101	104,6	117	158	22,58
S3	59	98	109	127,5	127	910	116,18

Uma primeira suspeita sobre os dados é sua elevada dispersão, evidenciada pelos valores do desvio padrão de algumas variáveis. Essa alta variabilidade pode ser explicada pela presença de outliers, tanto superiores quanto inferiores. Para verificar essa hipótese, foram gerados os boxplots das medições em jejum e pós-ingestão - dispostos na Figura 1 - permitindo uma visualização mais clara dos possíveis valores extremos. Além disso, valores como 666 e 910, identificados nas tabelas de medidas descritivas, requerem uma análise cuidadosa para determinar se representam erros de medição ou características legítimas dos dados.



(a) Mulheres em Jejum



(b) Mulheres em 1 Hora após a Ingestão de Açúcar

Figura 1: Boxplot's das Medições de Níveis de Glicose no Sangue de Mulhere em Jejum e 1 Hora após a Ingestão de Açúcar.

Ao analisar a Figura 1, especificamente o gráfico (a), percebe-se a presença de medidas aberrantes em

todas as semanas analisadas. O boxplot da segunda semana confirma a existência de valores extremos tanto superiores quanto inferiores, enquanto as demais semanas apresentam valores aberrantes apenas para cima.

Em contraste, a Figura 1 gráfico (b), revelou apenas um valor extremo nas medições da terceira semana, indicando menor variabilidade nos dados pós-ingestão de açúcar.

Após diálogo com o pesquisador responsável, concluiu-se que esses outliers foram erros de medição. Para evitar que tais valores comprometam a análise estatística, adotou-se a seguinte abordagem de tratamento:

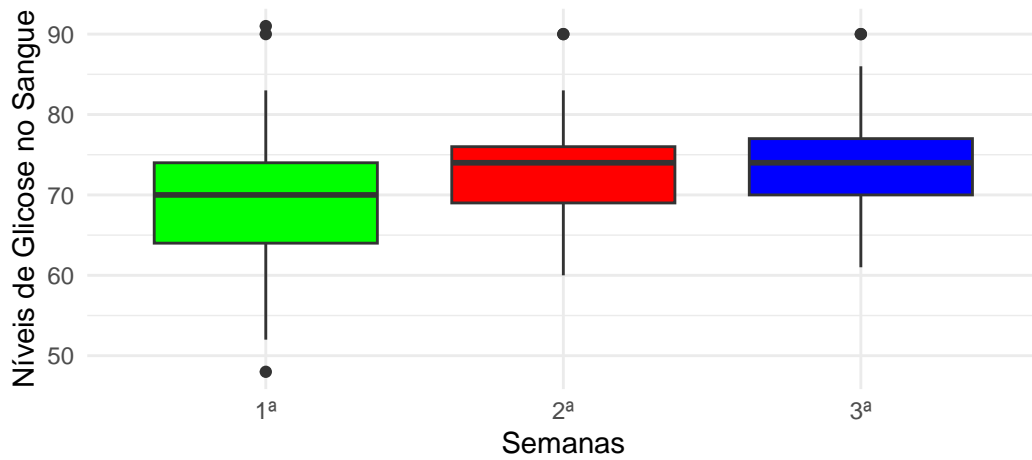
- Cópia do conjunto de dados original para preservação da informação inicial;
- Imputação dos valores aberrantes pela mediana, garantindo uma correção robusta sem distorcer a distribuição dos dados.

Tabela 3: Medidas de Resumo para Medições Ajustadas dos Níveis de Glicose no Sangue das Mulheres em Jejum.

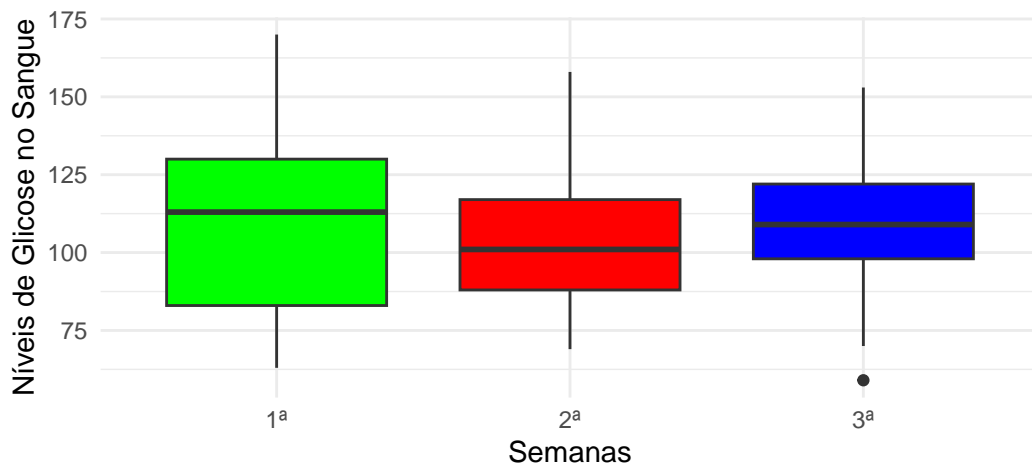
	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Desvio Padrão
S1	48	64	70	69,55	74	91	8,610
S2	60	69	74	73,04	76	90	6,291
S3	61	70	74	73,84	77	90	6,771

Tabela 4: Medidas de Resumo para Medições Ajustadas dos Níveis de Glicose no Sangue das Mulheres com 1 Hora após a Ingestão de Açúcar.

	Mínimo	1º Quartil	Mediana	Média	3º Quartil	Máximo	Desvio Padrão
S1	63	83	113	110,6	130	170	27,91
S2	69	88	101	104,6	117	158	22,58
S3	59	98	109	111,2	122	153	21,84



(a) Mulheres em Jejum



(b) Mulheres em 1 Hora após a Ingestão de Açúcar

Figura 2: Boxplot's das Medições Ajustadas de Níveis de Glicose no Sangue de Mulheres em Jejum e 1 Hora após a Ingestão de Açúcar.

Conforme pode ser visto na Figura 2 e a Tabela 3 e Tabela 4, após os ajustes realizados, a visualização dos dados tornou-se mais clara e detalhada, proporcionando uma melhor interpretação da dispersão por meio dos boxplots. No entanto, ainda há variações extremas que, possivelmente, estavam mascaradas devido às medições incorretas previamente identificadas e ajustadas. Apesar disso, esses valores aberrantes **não** parecem resultar de erros ou inconsistências nos dados, mas sim de características legítimas da variável analisada, refletindo sua própria natureza.

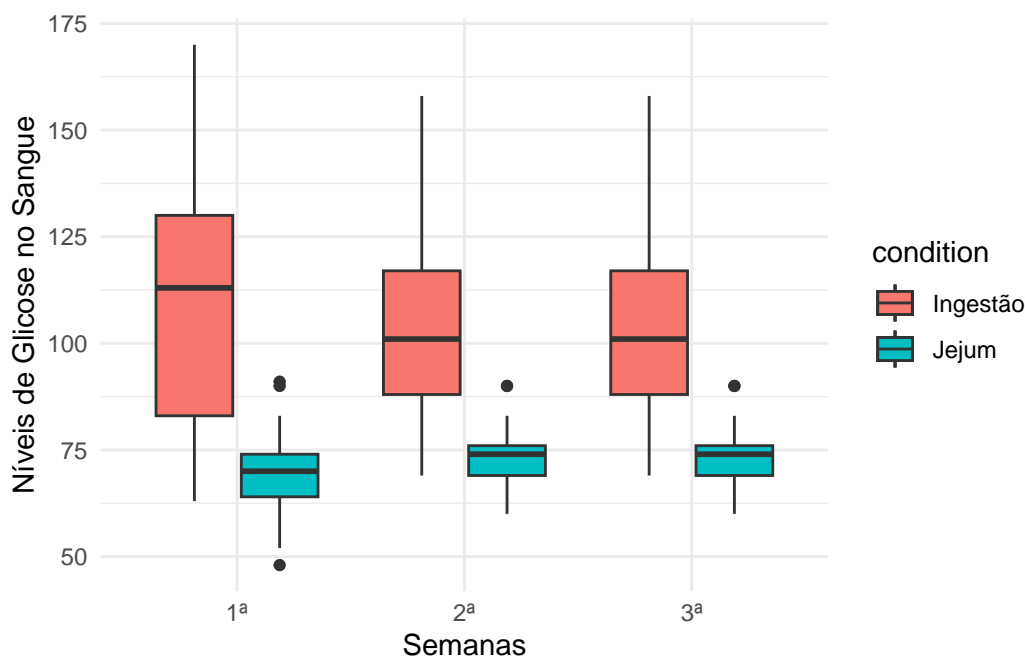


Figura 3: Boxplot's das Medições de Níveis de Glicose no Sangue de Mulheres por Condição.

Outra análise proposta é da Figura 3. Esta análise se concentra na variação dos níveis de glicose sanguínea em mulheres sob duas condições: **jejum** e **pós-ingestão**. O objetivo é avaliar as diferenças entre grupos e semanas, identificando padrões relevantes.

Os resultados mostram que as medições de glicose das mulheres em **jejum** apresentaram uma mediana inferior às das mulheres **uma hora após a ingestão de açúcar**, o que está alinhado com expectativas fisiológicas. Além disso, **não apenas a mediana difere entre os grupos, como também a variação dentro de cada grupo**.

Um ponto interessante é a **diferença na variabilidade entre as semanas**. A **primeira semana** apresenta maior dispersão dos valores, tanto no grupo em jejum quanto no grupo pós-ingestão. Isso sugere que, no início do estudo, os níveis glicêmicos podem estar mais instáveis, possivelmente devido a ajustes metabólicos ou variações individuais. Já as **segunda e terceira semanas** mostraram uma variação semelhante entre si, indicando um possível padrão de estabilização ao longo do tempo. Esse achado pode sugerir que, após um período inicial, os níveis glicêmicos tendem a se comportar de forma mais previsível.

As conclusões obtidas através da Figura 3 são ratificadas ao olharmos para a sétima coluna da Tabela 3 e Tabela 4. Lembrando que tais conclusões são visuais e com base em medidas de resumo. Caso se faça necessário essa hipótese pode ser testada por meio do *Teste de Igualdade da Matriz de Covariâncias* (**Box's M Test**).

Outro aspecto relevante da análise é a presença de **outliers exclusivamente no grupo em jejum**, mesmo após o tratamento estatístico dos valores extremos, algo que já foi comentado anteriormente. Essa característica pode estar relacionada a fatores biológicos como **respostas metabólicas individuais, variações hormonais ou hábitos alimentares prévios**. Dado esse cenário, futuras pesquisas poderiam aprofundar a investigação sobre os possíveis **fatores que influenciam esses valores aberrantes**. Algumas abordagens interessantes incluem: - **Avaliação de variáveis individuais**: Investigação do histórico alimentar e metabólico das participantes. - **Influência hormonal**: Consideração de marcadores hormonais, como insulina e cortisol, que podem impactar os níveis glicêmicos. - **Análise de fatores externos**: Estudo do impacto de elementos como estresse e qualidade do sono na variabilidade glicêmica.

Histogramas foram desenhados na Figura 4 como uma segunda forma de avaliar a distribuição dos dados.

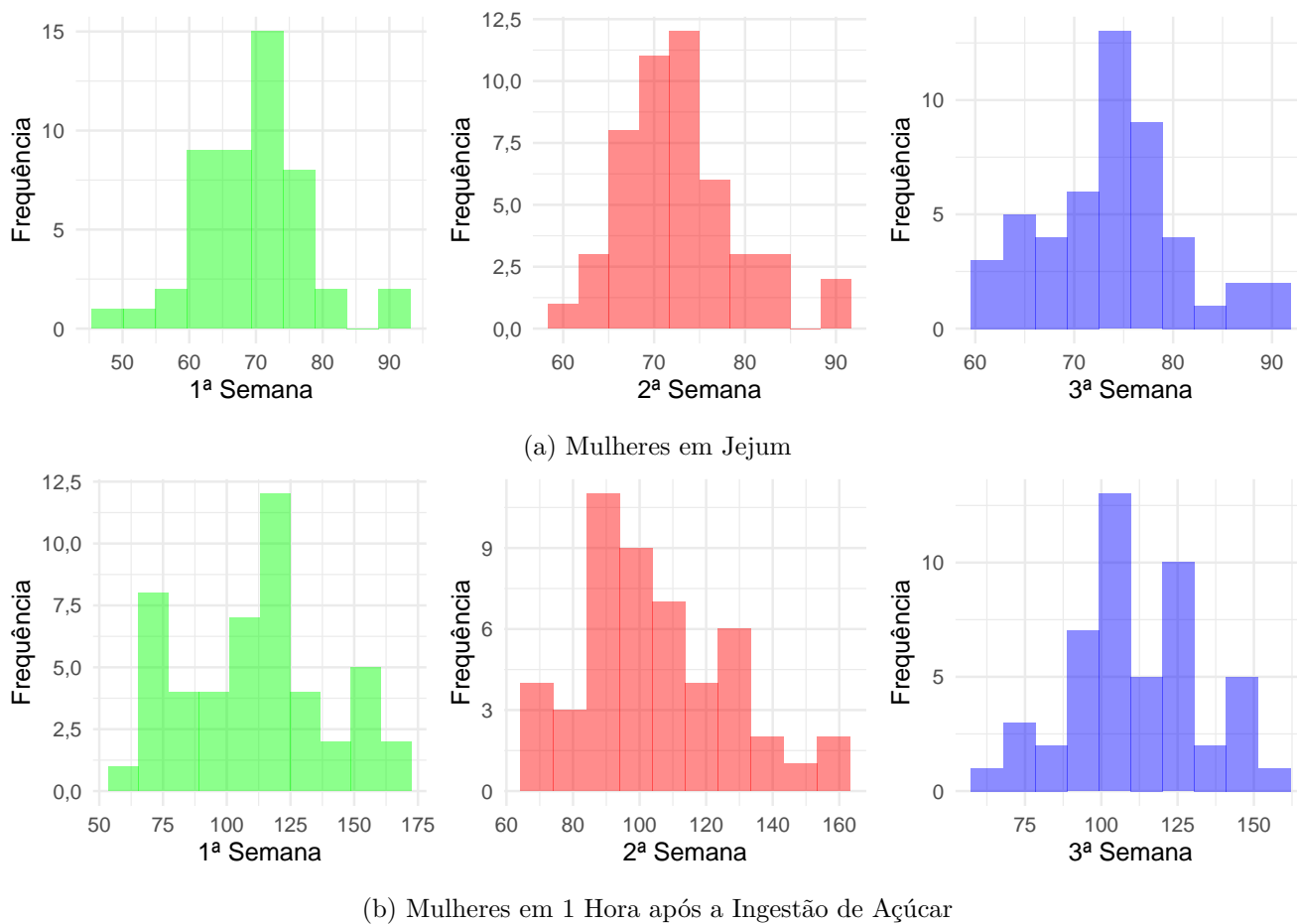


Figura 4: Histogramas dos Dados Ajustados de Mulheres em Jejum e 1 Hora após a Ingestão de Açúcar.

Os histogramas da Figura 4 indicam uma distribuição mais centrada, com valores mais equidistantes e menos extremos casando com o que a Figura 3 havia nos mostrado. Isso reforça a ideia de que os dados ajustados já não sofrem influência de valores discrepantes, garantindo uma análise mais precisa e representativa da distribuição real dos níveis de glicose.

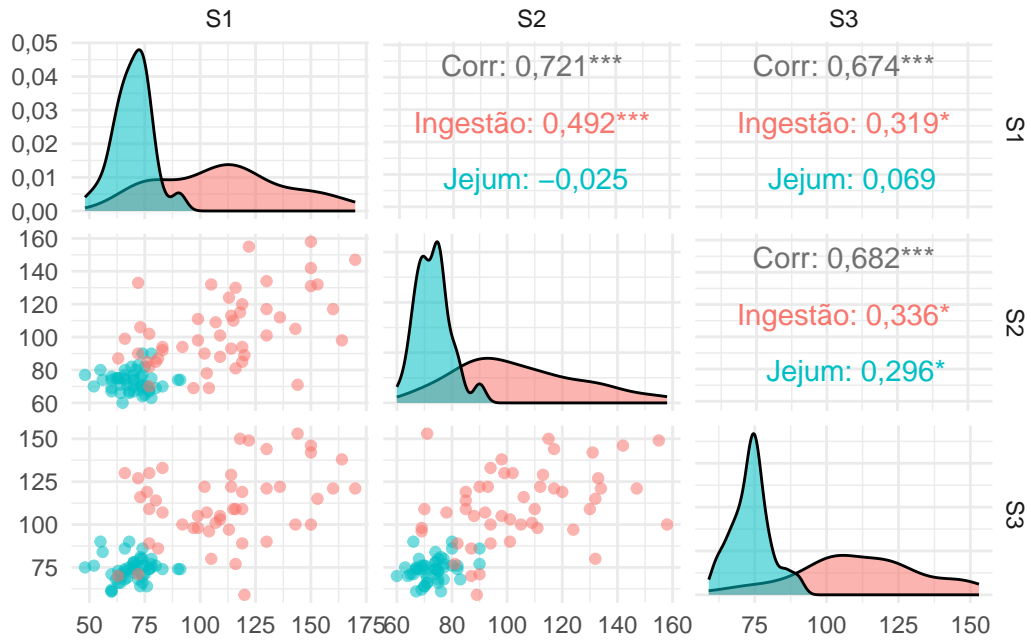


Figura 5: Análise de Correlação entre Variáveis Glicêmicas em Condições de Jejum e Ingestão.

A Figura 5 gera uma visualização rica ao correlacionar as variáveis do conjunto de dados entre si. Esse gráfico—ou, mais precisamente, uma matriz de plots—cruza as medições de cada semana, segmentadas em duas condições: jejum e pós-ingestão. A seguir, destacamos os principais pontos da análise para facilitar a interpretação.

1. Estimação de Densidade (Diagonais):

- As diagonais apresentam a densidade estimada para os dados em análise. O intuito desse gráfico é similar aos histogramas da Figura 4;
- No entanto, esse gráfico é mais adequado para avaliar a dispersão dos dados (assim julga o autor), mostrando que o grupo em jejum possui uma distribuição mais concentrada e menos variável, enquanto o grupo pós-ingestão apresenta maior dispersão;
- Essa visualização reforça e complementa as análises da Figura 3, Tabela 3 e Tabela 4.

2. Diagramas de Dispersão (Parte Inferior):

- A porção inferior da matriz apresenta gráficos de dispersão, ilustrando a relação entre as variáveis do conjunto de dados.
- A segmentação por condição evidencia diferenças claras: enquanto os dados gerais indicam uma relação linear positiva, no grupo em jejum essa tendência não sustenta, visualmente, de forma tão concreta para todas as combinações dois a dois das variáveis.
- Assim, a Figura 5 contribui para reforçar as distinções entre os grupos e destacar padrões de variação nos dados.

3. Correlação Entre Variáveis (Parte Superior)

- Como terceiro e último ponto a ser destacado, temos a parte triangular superior do gráfico. Na qual está apresentada a mostrado o *Coefficiente de Correlação Linear de Pearson* e um código que informa a significância estatística do *Teste de Linearidade*. De forma geral, esta parte do gráfico apresenta os seguintes resultados;

- Correlações gerais (todas as participantes):
 - S1 e S2: 0,721 (correlação positiva)
 - S1 e S3: 0,674 (correlação positiva)
 - S2 e S3: 0,682 (correlação positiva)
- Correlações no grupo em jejum:
 - S1 e S2: -0,025 (correlação negativa)
 - S1 e S3: 0,069 (correlação positiva)
 - S2 e S3: 0,296 (correlação positiva)
- Correlações no grupo pós-ingestão:
 - S1 e S2: 0,492 (correlação positiva)
 - S1 e S3: 0,319 (correlação positiva)
 - S2 e S3: 0,336 (correlação positiva)
- Estes coeficientes indicam que a relação entre as variáveis é mais evidente no grupo pós-ingestão, enquanto no grupo em jejum os vínculos são mais fracos ou negativos em algumas combinações. Essa diferença reforça a necessidade de considerar os grupos separadamente na análise.

Apenas para se ter noção a cerca da força da correlação entre as variáveis, foi elaborada a Figura 6.

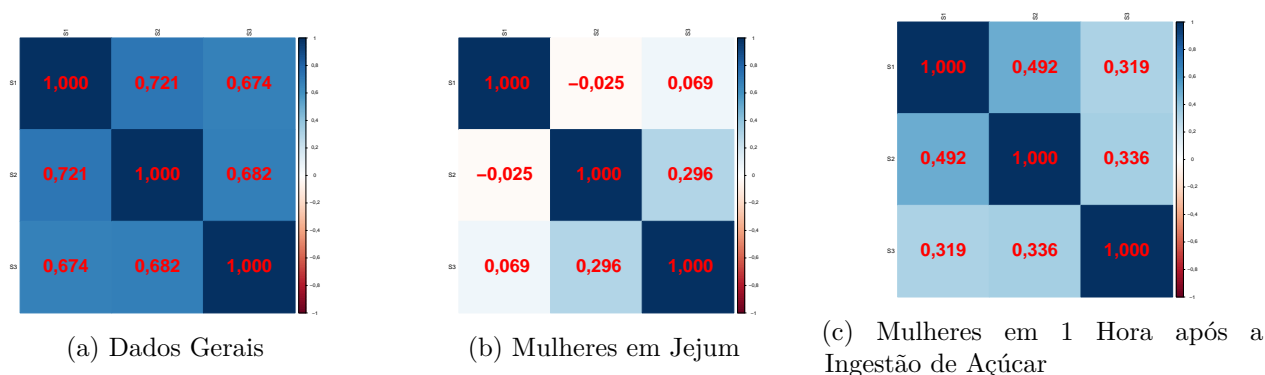


Figura 6: Mapa de Calor de Matriz de Correlação a partir do Coeficiente de Correlação Linear de Pearson.

Os valores de correlação indicam que as variáveis estão fortemente associadas no conjunto completo de dados, mas que essa relação diminui consideravelmente no grupo em jejum. Isso sugere uma mudança estrutural no comportamento glicêmico conforme o estado fisiológico das participantes.

4 Testes de Normalidade

Tabela 5: Teste de Mardia para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em Jejum.

Teste	Estatística de Teste	Valor p	Resultado
Mardia Skewness	13,1898750670407	0,213248633912465	YES
Mardia Kurtosis	0,241052123891276	0,809514718322033	YES
MVN	NA	NA	YES

Tabela 6: Teste de Mardia para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em 1 Hora após a Ingestão de Açúcar.

Teste	Estatística de Teste	Valor p	Conclusão
Mardia Skewness	7,71494336627685	0,656658258938476	YES
Mardia Kurtosis	0,0749255258944212	0,940273966909819	YES
MVN	NA	NA	YES

Tabela 7: Teste de Henze-Zirkler's para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em Jejum.

Teste	HZ	Valor p	Conclusão
Henze-Zirkler	0,842	0,1114	YES

Tabela 8: Teste de Henze-Zirkler's para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em 1 Hora após a Ingestão de Açúcar.

Teste	HZ	Valor p	Conclusão
Henze-Zirkler	0,5846	0,6724	YES

Tabela 9: Teste de Royston para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em Jejum.

Teste	H	Valor p	Conclusão
Royston	4,063	0,2551	YES

Tabela 10: Teste de Royston para Avaliação de Normalidade Multivariada das Medições de Níveis de Glicose no Sangue de Mulheres em 1 Hora após a Ingestão de Açúcar.

Teste	H	Valor p	Conclusão
Royston	4,83	0,1877	YES

A *Normalidade Multivariada* foi aceita em todos os testes. Chega-se a tal conclusão através dos valores p de cada teste. Isto é, $p - \text{valor} > \alpha$, deste modo, não se pode rejeitar H_0 .

5 Anexo: Códigos Utilizados

5.1 Estruturação e Pré-processamento dos Dados

```
import pandas as pd

j1 = pd.read_csv("C:/Users/user/Documents/PROJETOS/VS Code/MULTIVARIADA - VS/LISTAS/LISTA I/DATAS
j2 = pd.read_csv("C:/Users/user/Documents/PROJETOS/VS Code/MULTIVARIADA - VS/LISTAS/LISTA I/DATAS
```

```

# Ajustes prévios a serem feitos
d1 = [j1[:i].values.T[0] if i == 3 else j1[i-3:i].values.T[0] for i in range(3, len(j1) + 1, 3)]
d2 = [j2[:i].values.T[0] if i == 3 else j2[i-3:i].values.T[0] for i in range(3, len(j2) + 1, 3)]

# Formatação de DataFrame e Coluna Indicadora
d1 = pd.DataFrame(d1)
d1["condition"] = "Jejum"

d2 = pd.DataFrame(d2)
d2["condition"] = "Ingestão"

# Ajustes Finais
df = pd.concat([d1, d2], ignore_index=True)
df.columns = ["S1", "S2", "S3", "condition"]

# Salvar Conjunto de Dados ajustados
df.to_csv("C:/Users/user/Documents/PROJETOS/VS Code/MULTIVARIADA - VS/LISTAS/LISTA I/DATASETS/df.

# -----
# CONFIGURAÇÕES INICIAIS DO R
# -----
options(OutDec = ",", digits = 4)
library(ggplot2)
library(gridExtra)
library(dplyr)
library(MVN)

# -----
# DEFINIÇÕES DE FUNÇÕES USADAS
# -----
AIQ <- function(x) {
  AIQ <- quantile(x, 0.75) - quantile(x, 0.25)
  LI <- quantile(x, 0.25) - 1.5 * AIQ
  LS <- quantile(x, 0.75) + 1.5 * AIQ

  return(data.frame(LI = LI, LS = LS))
}

ajust.iter <- function(x) {
  # Número de iterações necessárias
  n.iter <- ncol(x)

  # Iterando de Imputação de Dados
  for (col in seq(1, n.iter)) {
    # Obtendo Limites para Ajuste
    ajj <- AIQ(x[[col]])

    # Ajustando os Dados
    x[[col]] <- ifelse((x[[col]] >= ajj$LI) & (x[[col]] <= ajj$LS), x[[col]], median(x[[col]]))
  }
}

```

```

}

# Retornar Dados Ajustados
return(x)
}

plot.histogram <- function(df) {
  p11 <- ggplot(data = df, aes(x = S1)) +
    geom_histogram(bins = 10, fill = "green", alpha = 0.45) +
    labs(x = "1ª Semana", y = "Frequência") +
    theme_minimal()

  p12 <- ggplot(data = df, aes(x = S2)) +
    geom_histogram(bins = 10, fill = "red", alpha = 0.45) +
    labs(x = "2ª Semana", y = "Frequência") +
    theme_minimal()

  p13 <- ggplot(data = df, aes(x = S3)) +
    geom_histogram(bins = 10, fill = "blue", alpha = 0.45) +
    labs(x = "3ª Semana", y = "Frequência") +
    theme_minimal()

  return(grid.arrange(p11, p12, p13, nrow = 1))
}

```

```

# -----
# LEITURA DOS DADOS
# -----

# Caminho do arquivo
path <- "C:/Users/user/Documents/PROJETOS/MIX/ANÁLISE DE MULTIVARIADA/MULTI II/df.csv"

# Leitura do arquivo
df <- readr::read_csv(file = path)

# Separação dos Dados
y <- df[df$condition=="Jejum",]
x <- df[df$condition=="Ingestão",]

```

5.2 Análise Descritiva dos Dados

```

# Medidas de Resumo
describe <- rbind(
  matrix(summary(y$S1), nrow = 1, ncol = 6),
  matrix(summary(y$S2), nrow = 1, ncol = 6),
  matrix(summary(y$S3), nrow = 1, ncol = 6)
)

```

```

stds <- sqrt(diag(var(y[, -4])))

describe <- cbind(
  describe,
  as.matrix(stds, ncol = length(stds))
)

# Formatação de Tabela
knitr::kable(
  describe,
  col.names = c("Mínimo", "1º Quartil", "Mediana", "Média", "3º Quartil", "Máximo", "Desvio Padrão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Medidas de Resumo
describe <- rbind(
  matrix(summary(x$S1), nrow = 1, ncol = 6),
  matrix(summary(x$S2), nrow = 1, ncol = 6),
  matrix(summary(x$S3), nrow = 1, ncol = 6)
)

stds <- sqrt(diag(var(x[, -4])))

describe <- cbind(
  describe,
  as.matrix(stds, ncol = length(stds))
)

# Formatação de Tabela
knitr::kable(
  describe,
  col.names = c("Mínimo", "1º Quartil", "Mediana", "Média", "3º Quartil", "Máximo", "Desvio Padrão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

ggplot(data = y) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +
  geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme_minimal()

ggplot(data = x) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +

```

```

geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
theme_minimal()

# Criando cópias dos Data Frames
y.copy <- dplyr::as_tibble(y)
x.copy <- dplyr::as_tibble(x)

# Ajustando os Dados
y.copy[, -4] <- ajust.iter(y.copy[, -4])
x.copy[, -4] <- ajust.iter(x.copy[, -4])

# Medidas de Resumo
describe <- rbind(
  matrix(summary(y.copy$S1), nrow = 1, ncol = 6),
  matrix(summary(y.copy$S2), nrow = 1, ncol = 6),
  matrix(summary(y.copy$S3), nrow = 1, ncol = 6)
)

stds <- sqrt( diag( var(y.copy[, -4]) ) )

describe <- cbind(
  describe,
  as.matrix(stds, ncol = length(stds))
)

# Formatação de Tabela
knitr::kable(
  describe,
  col.names = c("Mínimo", "1º Quartil", "Mediana", "Média", "3º Quartil", "Máximo", "Desvio Padrão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Medidas de Resumo
describe <- rbind(
  matrix(summary(x.copy$S1), nrow = 1, ncol = 6),
  matrix(summary(x.copy$S2), nrow = 1, ncol = 6),
  matrix(summary(x.copy$S3), nrow = 1, ncol = 6)
)

stds <- sqrt( diag( var(x.copy[, -4]) ) )

describe <- cbind(
  describe,
  as.matrix(stds, ncol = length(stds))
)

```

```

# Formatação de Tabela
knitr::kable(
  describe,
  col.names = c("Mínimo", "1º Quartil", "Mediana", "Média", "3º Quartil", "Máximo", "Desvio Padrão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Mulheres em Jejum
ggplot(data = y.copy) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +
  geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme_minimal()

# Mulheres em 1 Hora após a Ingestão de Açúcar
ggplot(data = x.copy) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +
  geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme_minimal()

ggplot(data = rbind(y.copy, x.copy), aes(fill = condition)) +
  geom_boxplot(aes(x = "1ª", y = S1)) +
  geom_boxplot(aes(x = "2ª", y = S2)) +
  geom_boxplot(aes(x = "3ª", y = S2)) +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme(legend.title = "Condição") +
  theme_minimal()

# Mulheres em Jejum
ggplot(data = y.copy) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +
  geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme_minimal()

# Mulheres em 1 Hora após a Ingestão de Açúcar
ggplot(data = x.copy) +
  geom_boxplot(aes(x = "1ª", y = S1), fill = "green") +
  geom_boxplot(aes(x = "2ª", y = S2), fill = "red") +
  geom_boxplot(aes(x = "3ª", y = S3), fill = "blue") +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme_minimal()

```

```

ggplot(data = rbind(y.copy, x.copy), aes(fill = condition)) +
  geom_boxplot(aes(x = "1ª", y = S1)) +
  geom_boxplot(aes(x = "2ª", y = S2)) +
  geom_boxplot(aes(x = "3ª", y = S2)) +
  labs(x = "Semanas", y = "Níveis de Glicose no Sangue") +
  theme(legend.title = "Condição") +
  theme_minimal()

# Plot das Mulheres em Jejum
plot.histogram(y.copy)

# Plot das Mulheres em 1 Hora após a Ingestão de Açúcar
plot.histogram(x.copy)

GGally::ggpairs(rbind(y.copy, x.copy),
                mapping = aes(colour = condition, alpha = 0.75),
                columns = c("S1", "S2", "S3")) +
  theme_minimal()

# Plot dos Dados Gerais
corrplot::corrplot(cor(rbind(y.copy, x.copy)[,1:3]), method = "color", addCoef.col = "red", tl.col = "black",
                  number.digits = 3, number.cex = 3)

# Plot das Mulheres em Jejum
corrplot::corrplot(cor(y.copy[,1:3]), method = "color", addCoef.col = "red", tl.col = "black",
                  number.digits = 3, number.cex = 3)

# Plot das Mulheres em 1 Hora após a Ingestão de Açúcar
corrplot::corrplot(cor(x.copy[,1:3]), method = "color", addCoef.col = "red", tl.col = "black",
                  number.digits = 3, number.cex = 3)

```

5.3 Testes de Normalidade

```

#| label: Testes de Normalidade - R

# Teste de Mardia
mardia.test.y <- mvn(y.copy[, -4], mvnTest = "mardia")
mardia.test.x <- mvn(x.copy[, -4], mvnTest = "mardia")

# Teste de Henze-Zirklers
hz.test.y <- mvn(y.copy[, -4], mvnTest = "hz")
hz.test.x <- mvn(x.copy[, -4], mvnTest = "hz")

# Teste de Royston
royston.test.y <- mvn(y.copy[, -4], mvnTest = "royston")
royston.test.x <- mvn(x.copy[, -4], mvnTest = "royston")

# Formatação de Tabela

```

```

knitr::kable(
  mardia.test.y$multivariateNormality,
  col.names = c("Teste", "Estatística de Teste", "Valor $p$", "Resultado"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Formatação de Tabela
knitr::kable(
  mardia.test.x$multivariateNormality,
  col.names = c("Teste", "Estatística de Teste", "Valor $p$", "Conclusão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Formatação de Tabela
knitr::kable(
  hz.test.y$multivariateNormality,
  col.names = c("Teste", "HZ", "Valor $p$", "Conclusão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Formatação de Tabela
knitr::kable(
  hz.test.x$multivariateNormality,
  col.names = c("Teste", "HZ", "Valor $p$", "Conclusão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Formatação de Tabela
knitr::kable(
  royston.test.y$multivariateNormality,
  col.names = c("Teste", "H", "Valor $p$", "Conclusão"),
  escape = FALSE,
  align = "c",
  booktabs = TRUE
)

# Formatação de Tabela
knitr::kable(
  royston.test.x$multivariateNormality,
  col.names = c("Teste", "H", "Valor $p$", "Conclusão"),
  escape = FALSE,

```



```
align = "c",  
booktabs = TRUE  
)
```