

(25-a0-03.06_academia1-tarefa1.prg)

Importar arquivo "dad600.acad3"

dad=read.csv(file.choose(), header=T, sep=";", dec=".")

head(dad); str(dad); dim(dad); Altura;

##Erro: objeto 'Idade' não encontrado

attach(dad) ## acessar as variáveis individualmente

stem(Altura)

dar <seta para cima> para recuperar os códigos digitados

Finalizar -> clicar Misc -> remover todos objetos

digitar <ctrl L> para limpar console

Importar arquivo "dad.Academia600.txt" usando `file.choose()`

`dad=read.table(file.choose(), h=T)`

abrir console do R e digitar

`dad=read.table("clipboard", h=T)` ## não digitar <enter>

abrir arquivo "dad.Academia600" e copiar as linhas e colunas.

`dad=read.table("clipboard", h=T)` ## em seguida, <enter>

`install.packages(readxl)` ## Instalar pacote readxl

`library(readxl)` ## ou `require(readxl)`

Localização do “arquivo” em D:\\24-a0-03.06_academia.excel.xls

dad=read_excel("D:\\dad.Academia600.xls")

dad=read_excel(file.choose())

apresenta nome das variáveis com as primeiras cinco linhas de dados

head(dad)

verificar classe das variáveis (p. ex, character, numérico, fator, etc.)

str(dad)

'data.frame': 600 obs. of 11 variables:

\$ Idade : int 47 35 46 33 27 37 31 30 32 38 ...

\$ E.Civil : chr "casado" "solteiro" "casado" "divorciado" ...

\$ P.Treino : chr "noite" "manha" "noite" "tarde" ...

\$ Freq.Sem : int 3 5 2 1 1 3 2 1 2 2 ...

\$ Aval.Instal : chr "A" "Ar" "A" "r" ...

\$ Aval.Aparelhos: int 5 3 5 3 2 5 4 3 4 5 ...

\$ Aval.Prof : int 4 5 4 3 2 4 4 3 5 4 ...

\$ Val.Mensal : int 142 126 147 116 111 133 118 113 119 131 ...

\$ P.Troca.Acad : chr "nao" "sim" "nao" "sim" ...

\$ Peso : int 74 68 78 64 53 65 59 54 60 68 ...

\$ Altura : num 1.71 1.63 1.7 1.61 1.57 1.64 1.6 1.59 1.6 1.65 ...

Ler as variáveis individualmente

Altura

Erro: objeto 'Altura' não encontrado

comando obrigatório para a leitura individual das variáveis de interesse

attach(dad)

Altura

Error in stem(Altura) : 'x' deve ser numérico

Vamos criar um banco de dados “individual” para realização de tarefas.

Para isso, usar número email, NE=01; 02; ... ; 17 (número e-mail)

e-mail <estatcomp22@gmail.com>

123.sergio <sergio@unesp.br>

set.seed(123); sorteio123=sample(1:600, 100,replace=T); sorteio123

[1] 415 463 179 526 195 118 299 229 244 14 374 91 348 355 26 519 426 211 590 593 555 373 143 544 490
 [26] 23 309 135 224 166 217 290 581 72 588 575 141 153 294 277 463 41 431 90 316 223 528 116 456 598
 [51] 39 159 209 374 34 516 13 69 409 308 278 89 537 291 424 286 121 110 158 64 483 477 480 67 85
 [76] 165 51 74 178 362 236 330 127 212 310 243 113 477 151 160 391 155 426 5 326 280 567 238 339 39

dad2=dad[sorteio123,]

head(dad123)

	Idade	E.Civil	P.Treino	Freq.Sem	Aval.Instal	Aval.Aparelhos	Aval.Prof	Val.Mensal	P.Troca.Acad	Peso	Altura
415	40	casado	noite	3	A	5	4	140	nao	76	1.66
463	43	casado	noite	2	Ar	5	4	135	nao	72	1.68
179	46	divorciado	tarde	5	r	3	3	142	sim	72	1.70
526	20	solteiro	manha	4	Ar	4	5	97	sim	41	1.52
195	29	viuvo	tarde	1	r	2	2	109	sim	53	1.58

118 35 viuvo tarde 1 r 2 2 124 sim 65 1.63

IMC=dad2\$Peso/(dad2\$Altura^2) ; head(IMC)

length(IMC)

[1] 100

summary(IMC)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
16.67	22.01	23.63	23.64	25.51	29.76

IMC.categ <- ifelse(IMC <= 23.63, "IMC<=Q2", "IMC>Q2")

table(IMC.categ)

```
IMC.categ1 <- ifelse(IMC <= 23.63, "IMC<=Q2", "IMC>Q2")
table(IMC.categ1)
```

```
IMC.categ1
```

```
IMC<=Q2 IMC>Q2
```

```
47    53
```

```
IMC.categ2 =matrix(0,100)
for (i in 1:100)
{
  if (IMC[i] < 22.01) IMC.categ2[i]=0
  else if (IMC [i]>= 22.01 & IMC[i]< 23.63) IMC.categ2[i]=1
  else if (IMC [i]>= 23.63 & IMC[i]< 25.51) IMC.categ2[i]=2
  else if (IMC [i]>= 25.51) IMC.categ2[i]=3
}
IMC.categ3=factor(IMC.categ2); IMC.categ3
```

```
[1] 3 3 2 0 0 2 0 2 3 1 2 2 3 0 3 0 0 3 2 2 3 3 1 1 0 0 3 0 3 2 0 2 1 1 0 3 1 3 3 1 3 0 2 0 3 1 2 1 0 2 2 0 3 2 1 3 3
```

```
[58] 2 2 2 1 2 0 0 1 3 0 0 3 3 1 1 1 2 1 1 3 2 3 2 0 3 2 3 2 1 2 1 2 1 1 3 0 0 2 0 0 1 0 2
```

Levels: 0 1 2 3

```
IMC.categ4=factor(IMC.categ3,levels=0:3,labels=c("1.Baixo","2.Normal",
"3.Elevado","4.Alto"))
```

```
table(IMC.categ4)
```

IMC.categ4

1.Baixo 2.Normal 3.Elevado 4.Alto

25 22 27 26

```
dad123Mod <- cbind(dad123, IMC, IMC.categ, IMC.categ4) ;
```

```
head(dad123Mod)
```

Idade	E.Civil	P.Treino	Freq.Sem	Aval.Instal	Aval.Aparelhos	Aval.Prof	Val.Mensal
P.Troca.Acad	Peso	Altura					
415	40	casado	noite	3	A	5	4 140 nao 76 1.66

463	43	casado	noite	2	Ar	5	4	135	nao	72	1.68
179	46	divorciado	tarde	5	r	3	3	142	sim	72	1.70
526	20	solteiro	manha	4	Ar	4	5	97	sim	41	1.52
195	29	viuvo	tarde	1	r	2	2	109	sim	53	1.58
118	35	viuvo	tarde	1	r	2	2	124	sim	65	1.63

	IMC	IMC.categ	IMC.categ4
415	27.58020	IMC>Q2	4.Alto
463	25.51020	IMC>Q2	4.Alto
179	24.91349	IMC>Q2	3.Elevado
526	17.74584	IMC<=Q2	1.Baixo
195	21.23057	IMC<=Q2	1.Baixo
118	24.46460	IMC>Q2	3.Elevado

salvar em arquivo dad123Mod.csv

write.table(dad123Mod, "D:\\dad123Mod-acad.csv", sep = ",", dec = ".")

Idade	E.Civil	P.Treino	Freq.Sem	Aval.Instal	Aval.Aparelhos	Aval.Prof	Val.Mensal	P.Troca	Acad	Peso	Altura	IMC	IMC.categ	IMC.categ4
415	40	casado	noite	3	A	5	4	140	nao	76	1.66	27.5802003193497	IMC>Q2	4.Alto
463	43	casado	noite	2	Ar	5	4	135	nao	72	1.68	25.5102040816327	IMC>Q2	4.Alto
179	46	divorciado	tarde	5	r	3	3	142	sim	72	1.724	9134948096886	IMC>Q2	3.Elevado
526	20	solteiro	manha	4	Ar	4	5	97	sim	41	1.52	17.7458448753463	IMC<=Q2	1.Baixo
195	29	viuvo	tarde	1	r	2	2	109	sim	53	1.58	21.2305720237141	IMC<=Q2	1.Baixo

Marcar a coluna em amarelo → clicar em cima da área com lado direito do mouse → clicar em excluir → clicar opção → deslocar as células à esquerda.

Idade	E.Civil	P.Treino	Freq.Sem	Aval.Instal	Aval.Aparelhos	Aval.Prof	Val.Mensal	P.Troca.Acad	Peso	Altura	IMC	IMC.categ	IMC.categ4
40	casado	noite	3	A	5	4	140	nao	76	1.66	27.5802003193497	IMC>Q2	4.Alto
43	casado	noite	2	Ar	5	4	135	nao	72	1.68	25.5102040816327	IMC>Q2	4.Alto
46	divorciado	tarde	5	r	3	3	142	sim	72	1.7	24.9134948096886	IMC>Q2	3.Elevado
20	solteiro	manha	4	Ar	4	5	97	sim	41	1.52	17.7458448753463	IMC<=Q2	1.Baixo
29	viuvo	tarde	1	r	2	2	109	sim	53	1.58	21.2305720237141	IMC<=Q2	1.Baixo
35	viuvo	tarde	1	r	2	2	124	sim	65	1.63	24.4646016033724	IMC>Q2	3.Elevado

salvar em D:\ dad123Mod2-acad.csv

Importar arquivo "dad123Mod2-acad.csv"

dad=read.csv(file.choose(), header=T, sep=";", dec=".")

head(dad); dim(dad); attach(dad) ; str(dad)

'data.frame': 100 obs. of 14 variables:

\$ Idade : int 40 43 46 20 29 35 24 30 36 38 ...

\$ E.Civil : chr "casado" "casado" "divorciado" "solteiro" ...

\$ P.Treino : chr "noite" "noite" "tarde" "manha" ...

\$ Freq.Sem : int 3 2 5 4 1 1 3 4 4 4 ...

\$ Aval.Instal : chr "A" "Ar" "r" "Ar" ...

\$ Aval.Aparelhos: int 5 5 3 4 2 2 5 4 4 4 ...

\$ Aval.Prof : int 4 4 3 5 2 2 4 5 5 5 ...

\$ Val.Mensal : int 140 135 142 97 109 124 110 109 130 131 ...

\$ P.Troca.Acad : chr "nao" "nao" "sim" "sim" ...

\$ Peso : int 76 72 72 41 53 65 46 59 69 61 ...

\$ Altura : num 1.66 1.68 1.7 1.52 1.58 1.63 1.55 1.58 1.63 1.65 ...

\$ IMC : num 27.6 25.5 24.9 17.7 21.2 ...

\$ IMC.categ : chr "IMC>Q2" "IMC>Q2" "IMC>Q2" "IMC<=Q2" ...

\$ IMC.categ4 : chr "4.Alto" "4.Alto" "3.Elevado" "1.Baixo" ...

Título: Análise Descritiva da Satisfação e Perfil dos Alunos de uma Academia de Ginástica

1. Introdução

Este relatório tem como objetivo analisar o perfil demográfico dos alunos de uma academia de ginástica, bem como seu grau de satisfação com os serviços oferecidos. A coleta de dados foi realizada por meio de questionários aplicados a uma amostra aleatória de 25 alunos. As análises estatísticas foram conduzidas utilizando o software R para fornecer insights valiosos para o dono da academia.

2. Objetivos

- Geral: Avaliar o perfil dos alunos e seu nível de satisfação com a academia.**
- Específicos:**
 - Descrever a distribuição etária, estado civil, período de treino e frequência semanal dos alunos.**
 - Avaliar a satisfação com as instalações, aparelhos e o suporte dos professores.**

- Analisar a relação entre o valor da mensalidade e a intenção de trocar de academia.

Questões e Análises

1. Qual é a distribuição do estado civil dos clientes da academia? (Unidimensional)

Objetivo: Entender o perfil demográfico dos clientes.

Código R:

Tabela de frequência

tabela_estado_civil <- table(E.Civil) ; tabela_estado_civil

E.Civil

casado	divorciado	solteiro	viuvo
50	12	30	8

Tabela 1: Distribuição de frequências absolutas dos 100 clientes atendidos na academia FCT-UNESP, no ano de 2025, segundo a variável E.Civil.

E.Civil	Pacientes	ou	f_i
casado	50		0,50
divorciado	12		0,12
Solteiro	30		0,30
viuvo	8		0,08
Total	100		100

Interpretação: A tabela apresentam a proporção de clientes casados, solteiros, divorciados e viúvos. Por exemplo, se 50% forem casados, o dono pode direcionar campanhas para casais ou famílias.

prop.table(tabela_estado_civil) ## Proporções em porcentagem

E.Civil

casado	divorciado	solteiro	viuvo
0.50	0.12	0.30	0.08

Gráfico de barras

```
barplot(tabela_estado_civil, main="Distribuição do Estado Civil",  
xlab="Estado Civil", ylab="Frequência", col=c(1:4))
```

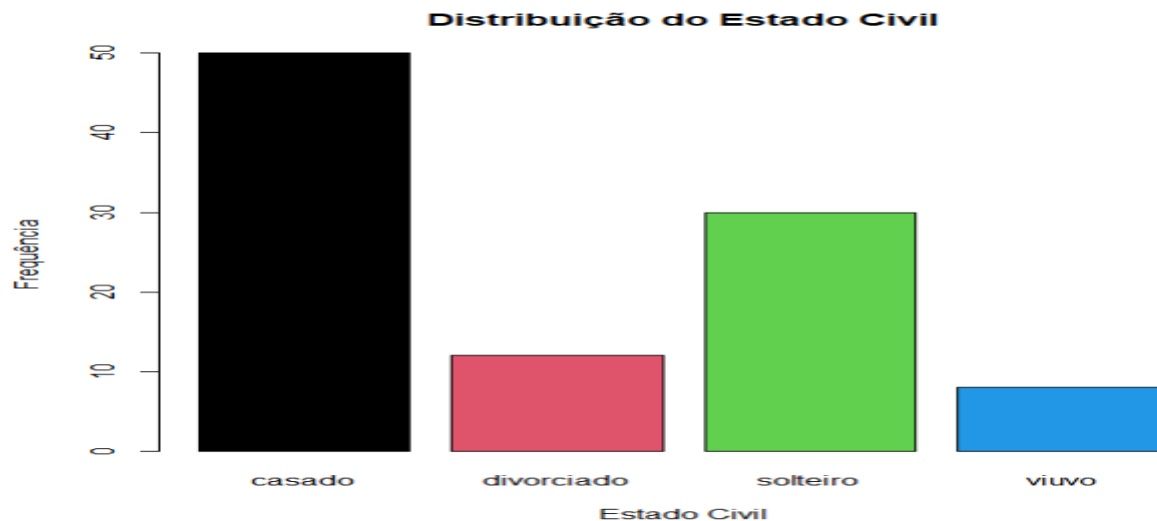


Figura 1: Distribuição de frequências absolutas dos 100 clientes atendidos na academia FCT-UNESP, no ano de 2025, segundo a variável E.Civil.

Interpretação: O gráfico apresenta a proporção de clientes casados, solteiros, divorciados e viúvos. Por exemplo, se 50% forem casados, o dono pode direcionar campanhas para casais ou famílias.

2. Qual é o período de treino mais popular entre os clientes? (Unidimensional)

Objetivo: Identificar os horários de maior demanda.

Código R:

Tabela de frequência

```
tabela_periodo <- table(P.Treino)
```

```
tabela_periodo ; prop.table(tabela_periodo)
```

P.Treino

manha noite tarde

34 50 16

P.Treino

manha noite tarde

0.34 0.50 0.16

Gráfico de pizza

```
pie(tabela_perodo, main="Distribuição do Período de Treino",  
col=rainbow(length(tabela_perodo)), labels=c("Manha (34)", "noite (50)", "tarde  
(16)"))
```

labels - comando para rotular as categorias da variável



Figura 2: Distribuição de frequências absolutas dos 100 clientes atendidos na academia FCT-UNESP, no ano de 2025, segundo a variável E.Civil.

Interpretação: Se "noite" for o período mais frequente (ex.: 50%), o dono pode aumentar a oferta de aulas ou professores nesse horário.

3. Qual é a frequência semanal dos clientes? (Unidimensional)

Objetivo: Avaliar o engajamento dos clientes.

Código R:

Tabela de frequência

```
tabela_freq <- table(Freq.Sem)
```

```
tabela_freq ; tabela_freq ; prop.table(tabela_freq)
```

Freq.Sem

1 2 3 4 5 6

9 24 32 22 10 3

Freq.Sem

1 2 3 4 5 6

0.09 0.24 0.32 0.22 0.10 0.03

Gráfico de barras

```
barplot(tabela_freq, main="Frequência Semanal", xlab="Dias por Semana",  
ylab="Frequência", col=11:16)
```

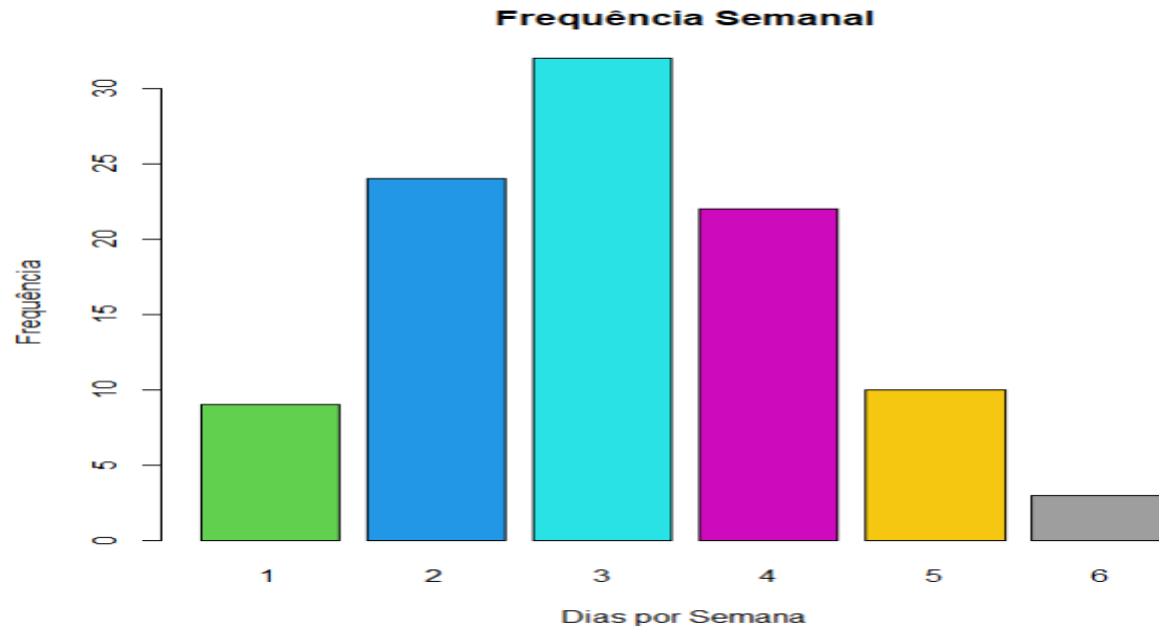


Figura ?: Título

Interpretação: Se a maioria treina 3-4 vezes por semana (ex.: 54%), isso indica bom engajamento, mas o dono pode incentivar os que treinam menos (ex.: 1-2 vezes) com promoções.

4. Como os clientes avaliam as instalações da academia? (Unidimensional)

Objetivo: Medir a satisfação com as instalações.

Código R:

Tabela de frequência

tabela_instal <- table(Aval.Instal)

tabela_instal ; prop.table(tabela_instal)

Aval.Instal

A Ar r

43 37 20

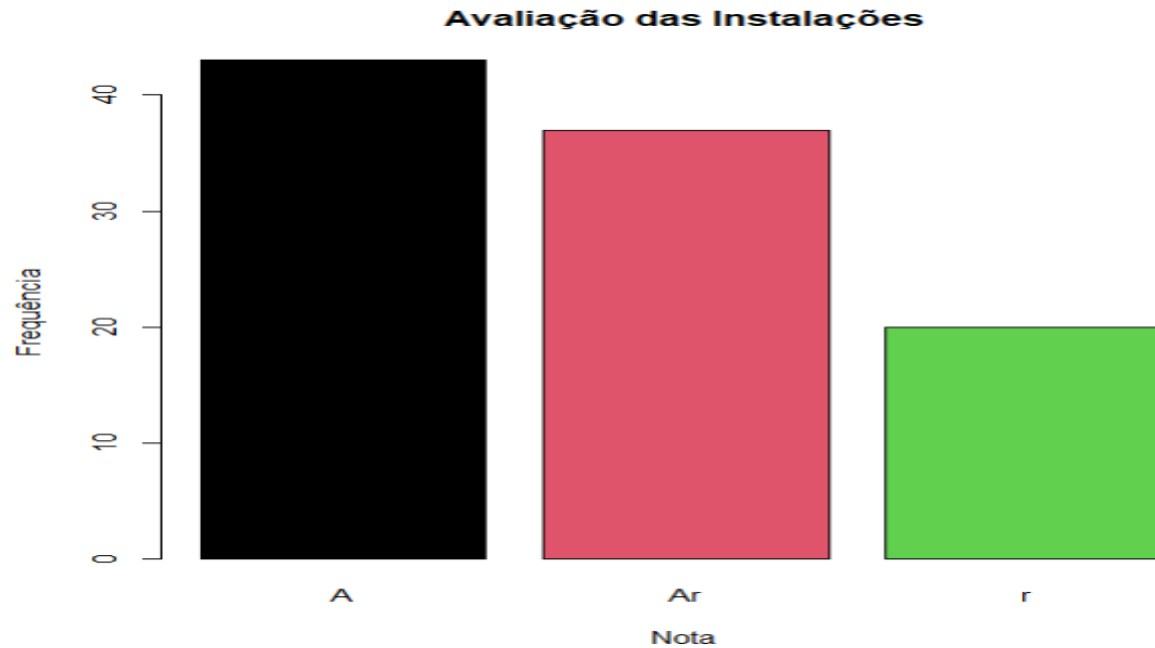
Aval.Instal

A Ar r

0.43 0.37 0.20

Gráfico de barras

```
barplot(tabela_instal, main="Avaliação das Instalações", xlab="Nota",  
ylab="Frequência", col=1:3)
```



Interpretação: Se notas A ou Ar predominam (ex.: 80%), as instalações são um ponto forte. Notas r indicam necessidade de melhorias.

5. Existe relação entre o estado civil e o período de treino? (Bidimensional)

Objetivo: Verificar se o estado civil influencia os horários preferidos.

Código R:

Tabela cruzada

tabela_cruzada1 <- table(P.Treino, E.Civil)

tabela_cruzada1 ; prop.table(tabela_cruzada1, 1) # Proporções por linha

E.Civil

P.Treino casado divorciado solteiro viuvo

manha 0 4 30 0

noite 50 0 0 0

tarde 0 8 0 8

E.Civil

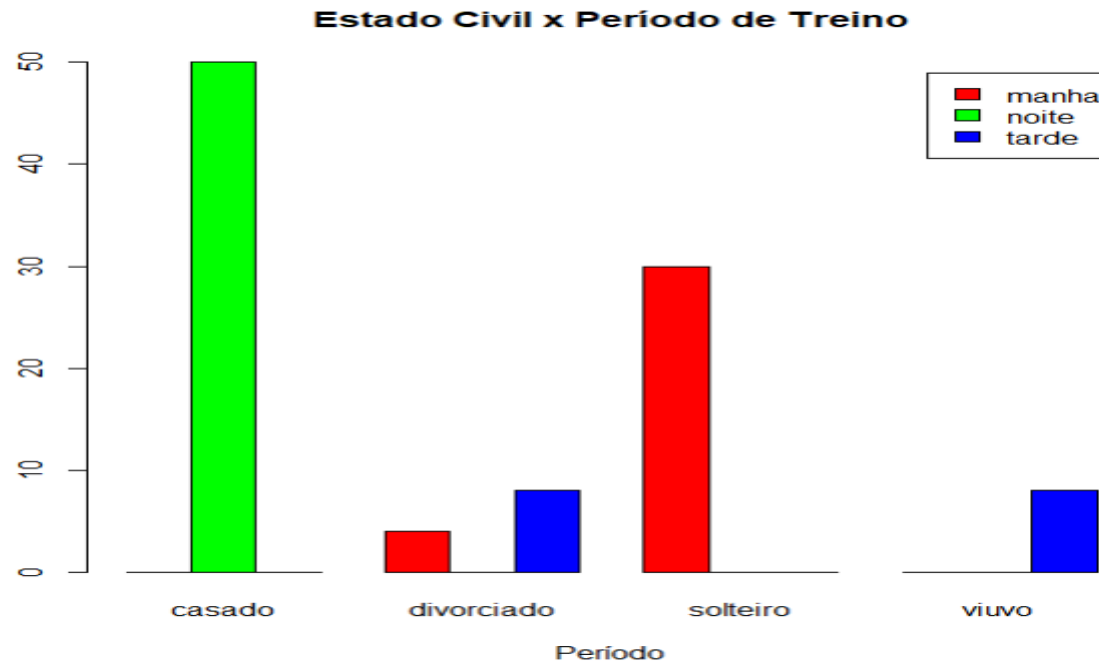
P.Treino casado divorciado solteiro viuvo

manha 0.0000000 0.1176471 0.8823529 0.0000000

noite 1.0000000 0.0000000 0.0000000 0.0000000
tarde 0.0000000 0.5000000 0.0000000 0.5000000

Gráfico de barras empilhadas

```
barplot(tabela_cruzada1, beside=T, main="Estado Civil x Período de  
Treino",  
xlab="Período",  
col=rainbow(3),  
legend=rownames(tabela_cruzada1))
```



Interpretação: Se casados preferirem "noite" (ex.: 50%) e solteiros "manhã" (ex.: 30%), o dono pode ajustar horários de aulas específicas para esses grupos.

6. A frequência semanal varia conforme o período de treino? (Bidimensional). Objetivo: Avaliar se o horário impacta o engajamento.

Código R:

Tabela cruzada

tabela_cruzada2 <- table(P.Treino ,Freq.Sem)

tabela_cruzada2 ; prop.table(tabela_cruzada2, 1)

Freq.Sem

P.Treino 1 2 3 4 5 6

manha 0 6 0 22 6 0

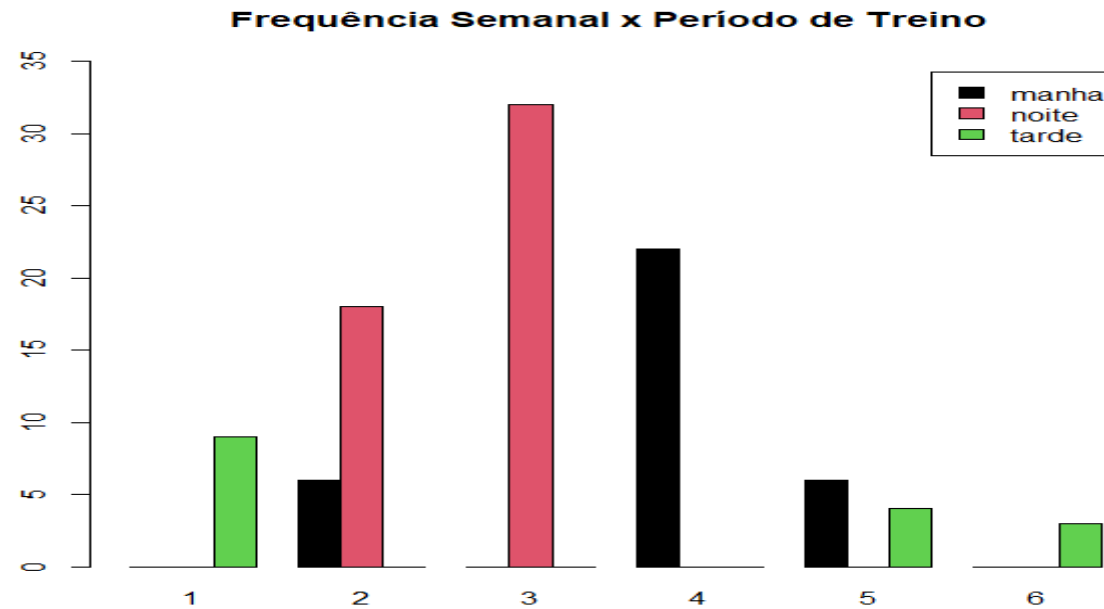
noite 0 18 32 0 0 0

tarde 9 0 0 0 4 3

Freq.Sem							
P.Treino	1	2	3	4	5	6	
manha	0.0000000	0.1764706	0.0000000	0.6470588	0.1764706	0.0000000	
noite	0.0000000	0.3600000	0.6400000	0.0000000	0.0000000	0.0000000	
tarde	0.5625000	0.0000000	0.0000000	0.0000000	0.2500000	0.1875000	

Gráfico de mosaico

```
barplot(tabela_cruzada2, beside=T, main="Frequência Semanal x Período  
de Treino ", ylim=c(0,35), col=c(1, 2, 3),  
legend=rownames(tabela_cruzada2))
```



Interpretação: Se clientes da "manha" treinam mais vezes (ex.: 4-5 dias, 28%) que os da "noite" (ex.: 2-3 dias), o dono pode investigar o que atrai maior frequência nesse período.

7. A avaliação dos professores está relacionada à propensão de trocar de academia? (Bidimensional)

Objetivo: Verificar se a satisfação com professores influencia a fidelidade.

Código R:

Tabela cruzada

```
tabela_cruzada3 <- table(P.Troca.Acad, Aval.Prof)
```

```
tabela_cruzada3 ; prop.table(tabela_cruzada3, 1)
```

fixa a tabela na linha para calcular as proporções

Aval.Prof

P.Troca.Acad 1 2 3 4 5

nao 0 0 0 55 3

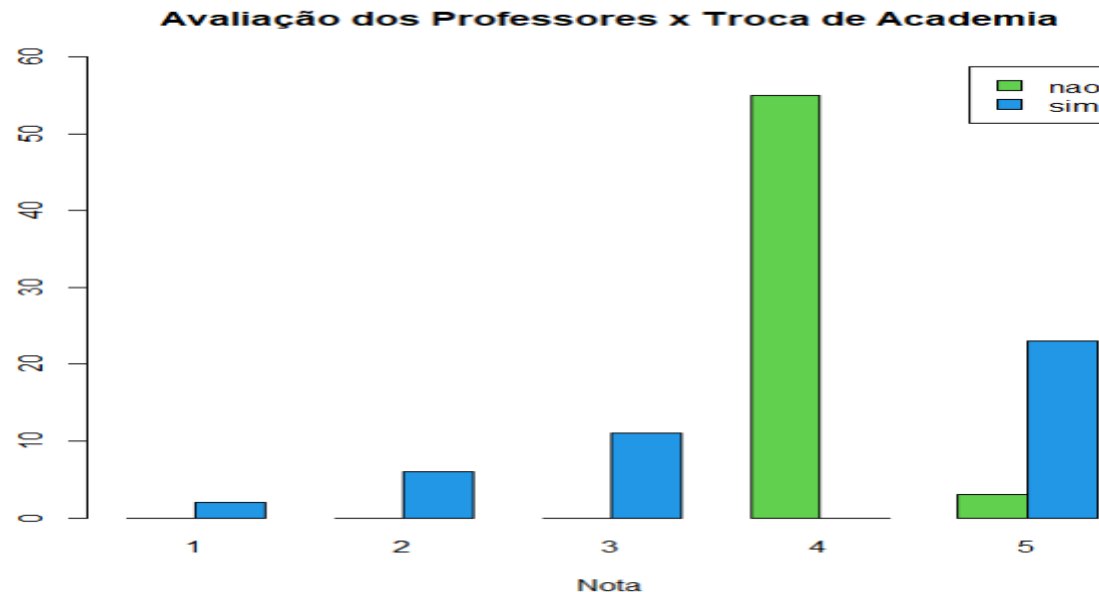
sim 2 6 11 0 23

Aval.Prof

P.Troca.Acad	1	2	3	4	5
nao	0.00000000	0.00000000	0.00000000	0.94827586	0.05172414
sim	0.04761905	0.14285714	0.26190476	0.00000000	0.54761905

Gráfico de barras empilhadas

```
barplot(tabela_cruzada3, beside=T, main="Avaliação dos Professores x Troca de Academia", xlab="Nota", col=c(3:4), legend=rownames(tabela_cruzada3), ylim=c(0,60))
```



Interpretação: Se notas baixas (1-2) têm alta propensão a trocar (ex.: 8% "sim"), o dono deve investir na capacitação dos professores.

8. O estado civil influencia a propensão a trocar de academia? (Bidimensional)

Objetivo: Identificar grupos mais propensos a deixar a academia.

Código R:

Tabela cruzada

tabela_cruzada4 <- table(P.Troca.Acad, E.Civil)

tabela_cruzada4 ; prop.table(tabela_cruzada4, 1)

E.Civil

P.Troca.Acad casado divorciado solteiro viuvo

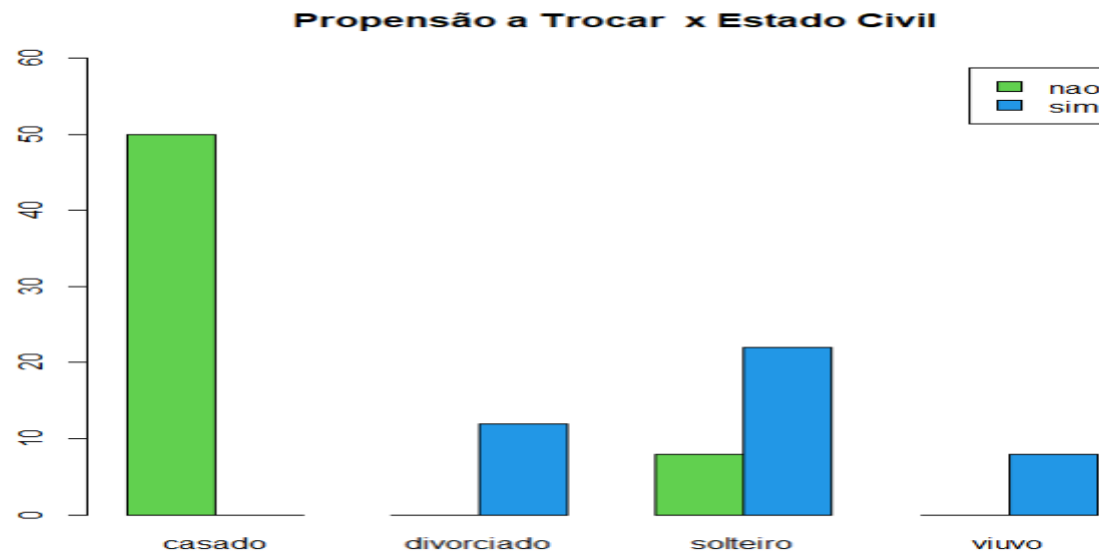
nao	50	0	8	0
sim	0	12	22	8

E.Civil

P.Troca.Acad	casado	divorciado	solteiro	viuvo
nao	0.8620690	0.0000000	0.1379310	0.0000000
sim	0.0000000	0.2857143	0.5238095	0.1904762

Gráfico de mosaico

barplot(tabela_cruzada4, beside=T, main=" Propensão a Trocar x Estado Civil", col=c(11:12), legend=rownames(tabela_cruzada4), ylim=c(0,60))



Interpretação: Se solteiros têm maior propensão a trocar (ex.: 52% "sim") que casados (ex.: 28%), o dono pode criar estratégias de retenção específicas.

9. A avaliação das instalações varia conforme o período de treino? (Bidimensional)

Objetivo: Verificar se o horário influencia a percepção das instalações.

Código R:

Tabela cruzada

tabela_cruzada5 <- table(Aval.Instal ,P.Treino)

tabela_cruzada5 ; prop.table(tabela_cruzada5, 1)

P.Treino

Aval.Instal manha noite tarde

A 0 43 0

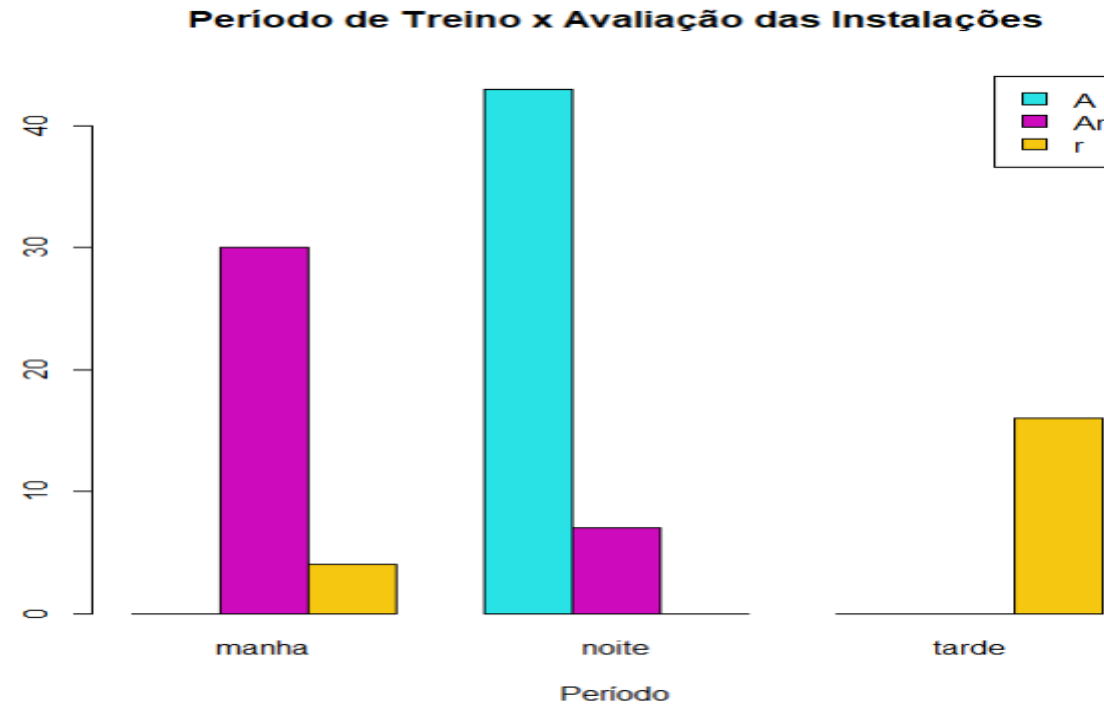
Ar	30	7	0
r	4	0	16

P.Treino

Aval.Instal	manha	noite	tarde
A	0.0000000	1.0000000	0.0000000
Ar	0.8108108	0.1891892	0.0000000
r	0.2000000	0.0000000	0.8000000

Gráfico de barras empilhadas

```
barplot(tabela_cruzada5, beside=T, main="Período de Treino x Avaliação
das Instalações", xlab="Período", col=c(5:7), legend=
c("A", "Ar", "r"), ylim=c(0,45))
```

Interpretação: Se "tarde" tem mais notas reprovadas (ex.: 80%) que "manhã" (20%), pode haver superlotação ou limpeza insuficiente à noite.

10. A frequência semanal está associada à avaliação dos aparelhos? (Bidimensional)

Objetivo: Avaliar se a qualidade dos aparelhos impacta o engajamento.

Código R:

Tabela cruzada

```
tabela_cruzada6 <- table(Aval.Aparelhos, Freq.Sem)
```

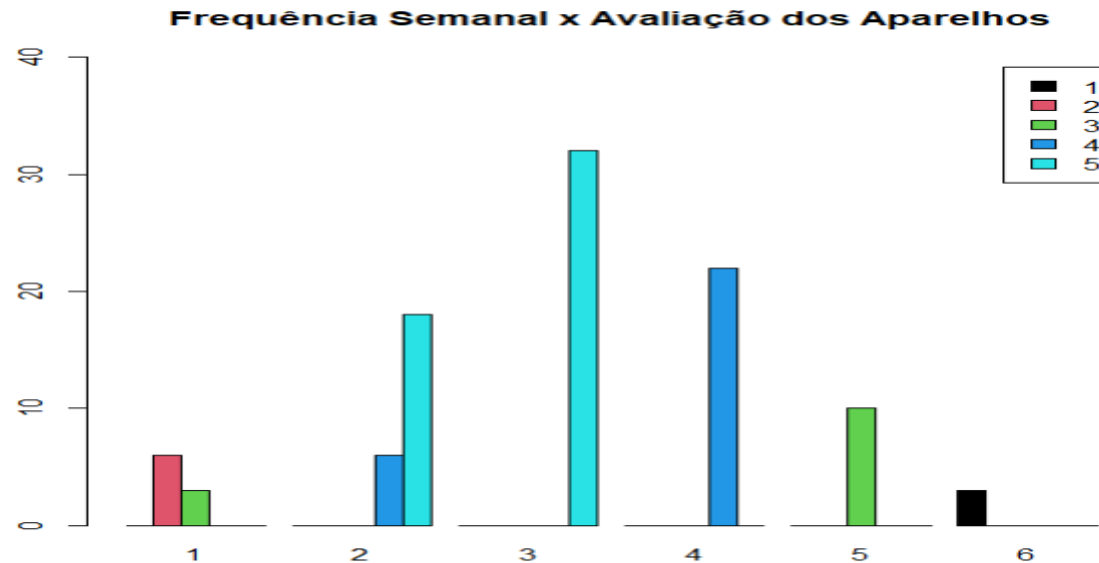
```
tabela_cruzada6 ; prop.table(tabela_cruzada6, 2) * 100
```

	Freq.Sem					
Aval.Aparelhos	1	2	3	4	5	6
1	0	0	0	0	0	3
2	6	0	0	0	0	0
3	3	0	0	0	10	0
4	0	6	0	22	0	0
5	0	18	32	0	0	0

Freq.Sem							
Aval.	Aparelhos	1	2	3	4	5	6
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	100.00000
2	66.66667	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3	33.33333	0.00000	0.00000	0.00000	100.00000	0.00000	0.00000
4	0.00000	25.00000	0.00000	100.00000	0.00000	0.00000	0.00000
5	0.00000	75.00000	100.00000	0.00000	0.00000	0.00000	0.00000

Gráfico de mosaico

```
barplot(tabela_cruzada6, beside=T, main="Frequência Semanal x
Avaliação dos Aparelhos", legend=rownames(tabela_cruzada6)
,col=c(1:5),ylim=c(0,40))
```



Considerações Finais

Essas questões e análises fornecem ao dono da academia uma visão clara sobre o perfil dos clientes, satisfação com os serviços e fatores que influenciam a fidelidade e o engajamento. Os códigos em R geram tabelas e gráficos que facilitam a visualização dos dados. Para implementar, basta carregar o arquivo CSV com os dados fornecidos (ajustando o nome do arquivo no código) e executar os comandos. Os resultados devem ser interpretados no contexto do negócio, ajustando estratégias com base nas descobertas (ex.: melhorar aparelhos, ajustar horários, focar na retenção de certos grupos).