**ATIVIDADE 1**

Realizar a atividade abaixo e postar a resolução no local indicado para envio de arquivo no APRENDER.

**EXERCÍCIOS**

1. De acordo com a aula teórica as variáveis de um estudo podem ser classificadas em quantitativas ou qualitativas e suas respectivas subdivisões. Classifique as variáveis abaixo:

* Time de preferência dos alunos: QUALITATIVAS NOMINAIS
* Quantidade de salários mínimos dos funcionários de uma empresa: QUANTITATIVAS DISCRETAS
* Cidade em que nasceu: QUALITATIVAS NOMINAIS
* Número de bactérias encontrados em amostras de leite QUANTITATIVAS DISCRETAS
* Temperatura mensal da cidade de Presidente Prudente: QUANTITATIVAS CONTINUAS
* Categoria do lutador de boxe (peso-pena, peso-leve, peso-pesado, etc.). QUALITATIVAS ORDINAIS
* Nível de escolaridade: QUALITATIVAS ORDINAIS
* Estado Civil: QUALITATIVAS ORDINAIS
* Quantidade (em gramas) de proteína ingerida por paciente: QUANTITATIVAS CONTINUAS

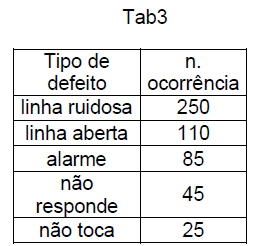
1. Utilize o conjunto “Milsa.txt”, construa a tabela de frequências das variáveis Instrução e Região.

dados$instr <- factor(dados$instr, label=c("1º Grau", "2º Grau", "Superior"), lev = 1:3, ord=T)

dados$regiao <- factor(dados$regiao, label=c("capital", "interior", "outro"), lev = c(2, 1, 3))

tabela<-matrix(c(prop.table(table(dados$instr)),prop.table(table(dados$regiao))\*100), nrow = 2, ncol = 3)

1. Numa central telefônica de uma grande empresa. havia a sensação de saturação do sistema utilizado. Para melhor representar o que ocorria foi realizado um acompanhamento com as telefonistas que teriam que responder aos problemas em que números ocorriam e lançá-los na **Lista de Verificação**. Tab3 resume os dados desta lista.



1. Classifique a variável “Tipo de defeito”. QUALITATIVAS NOMINAIS
2. Apresente um gráfico de setores para estes dados, incluindo porcentagem, legenda, cores e título.

tipo\_defeito <- c('Linha Ruidosa', 'Linha Aberta', 'Alarme', 'Não Responde', 'Não toca')

ocorrencias <- c(250, 110, 85, 45, 25)

porcentagem <- round(ocorrencias / sum(ocorrencias) \* 100, 2)

dados <- data.frame(tipo\_defeito, ocorrencias, porcentagem)

minhas\_cores <- rainbow(length(tipo\_defeito))

pie(dados$porcentagem, labels = paste(dados$tipo\_defeito, " (", round(dados$porcentagem, 1), "%)"), col = minhas\_cores, main = "Distribuição de Tipos de Defeito")

legend("topright", legend = dados$tipo\_defeito, fill = minhas\_cores, title = "Tipos de Defeito", cex = 0.8)

1. O vetor peso indica o peso médio de pintinhos com 2,4,6,8 e 10 dias de nascido respectivamente.

peso=(42,51,59,64,76).

Considere cada tempo de nascido como sendo uma categoria. Estabeleça um gráfico de barras para o peso médio dos pintinhos em cada categoria. Coloque título, legenda, nomes nos eixos, barras horizontais e utilize o seguinte esquema de cores: “blue”, “pink”, “yellow”,”green”,”red”.

peso <- c(42, 51, 59, 64, 76)

dia\_do\_nasc <- c(2, 4, 6, 8, 10)

dados <- data.frame(dia\_do\_nasc, peso)

minhas\_cores <- c("blue", "pink", "yellow", "green", "red")

barplot(dados$peso, names.arg = dados$dia\_do\_nasc, horiz = TRUE, col = minhas\_cores, xlab = "Peso Médio", ylab = "Nascimento", main = "Peso Médio dos Pintinhos por Nascimento")

legend("topright", legend = dados$dia\_do\_nasc, fill = minhas\_cores, title = "Nascimento")

1. Os dados são referentes às temperaturas diárias do mês de maio e setembro, em Fahrenheit, na cidade de Nova York em 1973.

tempm=c(67,72,74,62,56,66,65,59,61,69,74,69,66,68,58,64,66,57,68,62,59,73,61,61,57,58,57,67,81,79,76)

temps=c(91,92,93,93,87,84,80,78,75,73,81,76,77,71,71,78,67,76,68,82,64,71,81,69,63,70,77,75,76,68)

1. Converta as temperaturas do mês de maio para graus Celsius através da expressão °C = (°F − 32) / 1,8. Faça o histograma da freqüência relativa, coloque título, sombreamento de densidade 25 e cor = “dark blue”.

tempm <- c(67,72,74,62,56,66,65,59,61,69,74,69,66,68,58,64,66,57,68,62,59,73,61,61,57,58,57,67,81,79,76)

temps <- c(91,92,93,93,87,84,80,78,75,73,81,76,77,71,71,78,67,76,68,82,64,71,81,69,63,70,77,75,76,68)

temps\_celsius <- (temps - 32) / 1.8

hist(temps\_celsius, freq = FALSE, breaks = 5, col = "dark blue", main = "Histograma da Frequência Relativa das Temperaturas de Maio em Graus Celsius", xlab = "Temperatura em Graus Celsius")

lines(density(temps\_celsius), col = "dark blue", lwd = 2)

1. Calcule o nº de classes adequado, através da fórmula de Sturges, para o conjunto de temperatura do Mês de Setembro e construa um histograma.

temp\_set <- c(91,92,93,93,87,84,80,78,75,73,81,76,77,71,71,78,67,76,68)

num\_classes <- ceiling(log2(length(temp\_set) + 1))

hist(temp\_set, breaks = num\_classes, col = "dark blue", main = "Histograma das Temperaturas de Setembro", xlab = "Temperatura", ylab = "Frequência", border = "black")

1. Durante todo o mês de julho de 2018, a Sociedade Empresária Alfa realizou pesquisa diária visando medir a força da relação linear entre o número de acessos ao seu site na Internet (variável X) e o volume de vendas (em R$) de seu Produto “A” (variável Y). Sabe-se que os dados amostrais obtidos para os 7 primeiros dias de pesquisa foram:

| **Dia** | **Número de acessos ao site** | **Volume de vendas do Produto “A” (em R$)** |
| --- | --- | --- |
| **1** | 7 | 14 |
| **2** | 8 | 18 |
| **3** | 11 | 19 |
| **4** | 10 | 21 |
| **5** | 9 | 18 |
| **6** | 12 | 24 |
| **7** | 8 | 17 |

Construa um diagrama de dispersão para esse par de variáveis. Com base neste gráfico, você acredita que há relação entre o número de acessos ao site e o volume de vendas?

dia <- 1:7

acessos <- c(7, 8, 11, 10, 9, 12, 8)

volume <- c(14, 18, 19, 21, 18, 24, 17)

plot(acessos, volume, xlab = "Número de Acessos", ylab = "Volume de Vendas", main = "Diagrama de Dispersão: Acessos x Volume")

abline(lm(volume ~ acessos), col = "red")

correlacao <- cor(acessos, volume)

cat("Correlação entre número de acessos e volume de vendas:", correlacao)

PELO MOSTRADO NO GRÁFICO QUANTO MAIS ACESSO O SITE TEVE MAIS VENDAS FORAM REALIZADAS.

1. De acordo com Ministério da Educação a quantidade e alunos matriculados no ensino de 1º grau no Brasil nos de 1990 a 1996 em milhares de alunos, são: 19.720 – 20.567 – 21.473 – 21.887 – 20.598 – 22.473 – 23.564. Faça um gráfico de série temporal para apresentar esses dados.

anos <- c(1990:1996)

alunos <- c(19720, 20567, 21473, 21887, 20598, 22473, 23564)

plot(anos, alunos, type = "o", col = "blue", xlab = "Ano", ylab = "Alunos", main = "Alunos Matriculados no Ensino de 1º Grau no Brasil de 1990 a 1996")