

**Questão 1 (*Gráfico de dispersão*)***(2 pontos)*

Carregue o pacote do ggplot

```
library(ggplot2)
```

Utilizando o banco de dados Iris, Crie um gráfico de dispersão das variáveis “Sepal.Length”, e “Sepal.Width”. De acordo com as seguintes propriedades:

1. A variável “Sepal.Length” deve ser colocada no eixo x;
2. A variável “Sepal.Width” deve ser colocada no eixo y;
3. Cada ponto deve ser colorido baseado na variável “Species”
4. O título do gráfico deve ser “Sepal Length vs Sepal Width”
5. O subtítulo do “Iris Database”
6. O caption do gráfico deve ser “Fonte: Iris database (R-Studio)”
7. O nome do eixo x deve ser “Sepal Length”
8. O nome do eixo y deve ser “Sepal Width”

**Solução:**

```
library(ggplot2)

ggplot(iris)+
  geom_point(aes(x = Sepal.Length,
                 y = Sepal.Width,
                 color = Species)) +
  labs(title = "Sepal Length vs Sepal Width",
       subtitle = "Iris Database",
       caption = "Fonte: Iris database (R-Studio)",
       y="Sepal Width",
       x="Sepal Length",
       color="legenda")
```

**Questão 2 (*Gráfico de barras*)***(2 pontos)*

Utilizando o banco de dados airquality, Crie um gráfico de barras da variável “Temp”. De acordo com as seguintes propriedades:

1. A cor da borda das barras deve ser preta (colour=“black”)
2. A cor da barras deve ser azul (fill=“blue”)
3. A barra deve ter transparência de 50% (alpha=0.5)
4. O título do gráfico deve ser “New York Air Quality Measurements”
5. O subtítulo deve ser nulo (subtitle = NULL)
6. O caption do gráfico deve ser “Fonte: airquality database (R-Studio)”
7. O nome do eixo x deve ser “Temperatura [F]”
8. O nome do eixo y deve ser “Qtd”

**Solução:**

```
library(ggplot2)

ggplot(airquality) +
  geom_bar(aes(x=Temp), colour="black", fill="blue", alpha=0.5) +
  labs(title = "New York Air Quality Measurements",
       subtitle = NULL,
       caption = "Fonte: airquality database (R-Studio)",
       y="Qtd",
       x="Temperatura [F]")
```

**Questão 3 (*Histograma*)**

(2 pontos)

Utilizando o banco de dados diamonds (para isso carregue a biblioteca ggplot2), crie um histograma da variável “Carat”. De acordo com as seguintes propriedades:

1. A cor do histograma deve ser azul (fill=“blue”)
2. O histograma deve ter transparência de 50% (alpha=0.5)
3. A variável y deve mostrar a densidade (y=..density..)
4. O histograma deve ter 50 *bins*
5. O histograma deve ter facetas (*Facets*) por “cut”
6. O título do gráfico deve ser ”Carat of round cut diamonds”
7. O subtítulo deve ser “Facet by cut”

**Solução:**

```
ggplot(diamonds) +
  geom_histogram(aes(x=carat, y=..density..),
                fill="blue", alpha= 0.5, bins = 50) +
  facet_wrap(~cut) +
  labs(title = "Carat of round cut diamonds",
       subtitle = "Facet by cut")
```

**Questão 4 (*Gráfico de linha (time series)*)**

(2 pontos)

Primeiramente vamos converter o banco de dados “AirPassengers” para um data.frame, o código é mostrado abaixo.

1. O nome do banco convertido será “data”
2. Vamos nomear a coluna do banco “data” de “Passengers”
3. Vamos criar uma nova coluna chamado “date” com as datas associadas.
4. Vamos converter os dados da coluna “Passengers” para o tipo “numeric”

```
# O nome do banco convertido será "data"
data <- as.data.frame(AirPassengers)

# Vamos nomear a coluna do banco "data" de "Passengers"
colnames(data) <- "Passengers"

# Vamos criar uma nova coluna chamado "date" com as datas associadas.
data$date <- seq(from=as.Date("1949-01-01"), to=as.Date("1960-12-01"), by="month")

# Vamos converter os dados da coluna "Passengers" para o tipo "numeric"
data$Passengers <- as.numeric(data$Passengers)
```

Utilizando o banco de dados convertido, crie um gráfico de linha de acordo com as seguintes propriedades:

1. O eixo x deve conter a variável “date”
2. O eixo y deve conter a variável “Passengers”
3. Adicione uma linha, de cor vermelha, com intercept = 760, slope = 0.087
4. Adicione os seguintes elementos gráficos:
  - (a) title = “Monthly Airline Passenger Numbers 1949-1960”;
  - (b) eixo y = “Passengers”;
  - (c) eixo x = “Data”

#### Solução:

```
ggplot(data) +
  geom_line(aes(x=date, y=Passengers)) +
  geom_abline(intercept = 760, slope = 0.087, colour= "red") +
  labs(title = "Monthly Airline Passenger Numbers 1949-1960",
       y="Passengers",
       x="Data")
```

#### Questão 5 (*Gráfico de dispersão com linha*)

(2 pontos)

Utilizando o banco de dados mtcars, crie um gráfico de dispersão colocando a variável “hp” no eixo x e a variável “mpg” no eixo y. De acordo com as seguintes propriedades:

1. O título do gráfico deve ser “Scatter plot”
2. O nome do eixo x deve ser “Weight (1000 lbs)”
3. O nome do eixo y deve ser “Miles/(US) gallon”
4. Deve conter uma linha de regressão linear que aproxima os dados (para isso utilize o geoma: “geom\_smooth(method=lm)”) tracejada (linetype=“dashed”) com preenchimento em azul (fill=“blue”)
5. O tema aplicado deve retirar os *grids* (“theme(panel.background = element\_blank())”)

#### Solução:

```
library(ggplot2)

# gerando o grafico
ggplot(data = mtcars,
        mapping = aes(x=hp, y=mpg)) +
  geom_point() +
  geom_smooth(method=lm,
              linetype="dashed",
              fill="blue") +
  labs(title = "Scatter plot",
        y="Miles/(US) gallon",
        x="Weight (1000 lbs)") +
  theme(panel.background = element_blank())
```