

# Lista de Exercícios

2025-11-16

## Questão 1

Um entusiasta gastronômico está em uma jornada para explorar restaurantes listados no guia Michelin. O viajante começou sua aventura no Restaurante Piano 35 em Turin, Italy. Após saborear as delícias do restaurante, ele segue para a cidade mais próxima que abriga um restaurante reconhecido no guia Michelin.

Para calcular a distância entre os restaurantes, utilizaremos a equação de Haversine, uma fórmula eficaz para medir distâncias entre pontos de latitude e longitude na superfície esférica da Terra.

A equação de Haversine para calcular a distância ( $d$ ) entre dois pontos na superfície de uma esfera (como a Terra) a partir das coordenadas de latitude, em radianos,  $(\varphi_1, \varphi_1)$  e longitude  $(\lambda_1, \lambda_2)$  é dada por

$$d = 2r \arcsin \left( \sqrt{\sin^2 \left( \frac{\varphi_2 - \varphi_1}{2} \right) + \cos \varphi_1 \cos \varphi_2 \sin^2 \left( \frac{\lambda_2 - \lambda_1}{2} \right)} \right)$$

em que  $r$  é o raio da Terra (aproximadamente 6,371 km). Neste contexto, responda as questões a seguir.

1. Dica: Implemente a função `haversine_distance` para calcular a distância entre dois pontos na superfície da Terra. Verifique a necessidade de converter as coordenadas de latitude e longitude para radianos.

```
library(tidyverse)
library(magrittr)

haversine_distance <- function(lat1, long1, lat2, long2){

  lat1_rad <- lat1 * pi / 180
  lat2_rad <- lat2 * pi / 180

  dist_lat <- (lat2 - lat1) * pi / 180
  dist_lon <- (long2 - long1) * pi / 180

  d <- 2 * asin(sqrt((sin(dist_lat / 2))^2 +
    cos(lat1_rad) * cos(lat2_rad) *
    (sin(dist_lon / 2))^2)) * 6371

  return(d)
}
```

2. Utilize o banco de dados `micelin.csv` para obter as informações necessárias.

```
micelin <- read_csv("micelin.csv") %>% ##carrega os dados
  separate(Award, c("star", "color"), sep = ",") %>% ##separa a variável em duas
  separate(Location, c("city", "country"), sep = ", ")
```

```
## Rows: 7047 Columns: 13
## -- Column specification -----
## Delimiter: ","
## chr (10): Name, Address, Location, Price, Cuisine, Url, WebsiteUrl, Award, F...
## dbl (3): Longitude, Latitude, PhoneNumber
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

## Warning: Expected 2 pieces. Missing pieces filled with 'NA' in 6695 rows [2, 3, 4, 5, 6,
## 7, 8, 9, 12, 15, 17, 18, 19, 20, 21, 22, 23, 24, 27, 28, ...].

## Warning: Expected 2 pieces. Missing pieces filled with 'NA' in 309 rows [40, 41, 42, 51,
## 53, 55, 269, 270, 271, 272, 273, 274, 290, 291, 292, 298, 299, 300, 301, 302,
## ...].
```

```
dados_rest_inicial <- michelin %>%
  filter(Name == "Piano 35") ## restaurante inicial

michelin %<>%
  mutate(dist_rest_inicial = haversine_distance( ## aplica a função
    lat1 = dados_rest_inicial$Latitude,
    long1 = dados_rest_inicial$Longitude,
    lat2 = michelin$Latitude,
    long2 = michelin$Longitude))
```

- Quantos quilômetros o entusiasta precisa viajar para visitar o próximo restaurante com 1 Star ?

```
michelin %>%
  filter(star == "1 Star" & dist_rest_inicial > 0) %>%
  filter(dist_rest_inicial == min(dist_rest_inicial)) %>%
  pull(dist_rest_inicial)
```

```
## [1] 1.053176
```

- Existe, no total, quantos restaurantes com 1, 2 ou 3 estrelas Michelin a uma distância de 100 km do restaurante inicial?

```
vetor <- c("1 Star", "2 Stars", "3 Stars")

michelin %>%
  filter(star %in% vetor & dist_rest_inicial <= 100) %>%
  nrow()
```

```
## [1] 54
```

- O estusiasta decidiu celebrar seu aniversário em um restaurante com pelo menos uma estrela Michelin. Ele pode gastar até 2 dinheiros locais em um próximo restaurante que esteja a uma distância máxima de 3000 km, fora da cidade inicial. Nesse contexto, ele deveria escolher um restaurante dentre quantos?

```

michelin %>%
  filter(dist_rest_inicial <= 3000, star %in% vetor, str_length(Price) <= 2, city != dados_rest_inicial)
  nrow()

```

```
## [1] 36
```

- Caso o entusiasta queira saborear algo da culinária Italian Contemporary qual a distância mínima (em km) que ele precisará se deslocar?

```

michelin %>%
  separate(Cuisine, c("cuisine_1", "cuisine_2"), sep = ", ") %>%
  filter((str_detect(cuisine_1, "^Italian Contemporary$") | str_detect(cuisine_2, "^Italian Contemporary$"))
  summarise(dist_minima = min(dist_rest_inicial))

```

```
## Warning: Expected 2 pieces. Missing pieces filled with 'NA' in 4839 rows [1, 2, 3, 7, 8,
## 13, 14, 15, 18, 23, 24, 25, 27, 28, 31, 32, 40, 41, 42, 43, ...].
```

```

## # A tibble: 1 x 1
##   dist_minima
##   <dbl>
## 1       1.43

```

## Questão 2

Para as perguntas a seguir, utilize o banco de dados proveniente do TidyTuesday de 2021, semana 48, que contém informações sobre a série de TV Dr. Who. Para carregar o banco, utilize a função abaixo.

```
tuesdata <- tidyTuesdayR::tt_load(2021, week = 48)
```

Note que o resultado é uma lista contendo quatro objetos, cada um contendo um banco de dados. Os objetos são: `writers`, `directors`, `episodes` e `imdb`.

```

library(tidyverse)
library(magrittr)

tuesdata <- tidyTuesdayR::tt_load(2021, week = 48)

```

```

## ---- Compiling #TidyTuesday Information for 2021-11-23 ----
## --- There are 4 files available ---
##
##
## -- Downloading files -----
##
## 1 of 4: "writers.csv"
## 2 of 4: "directors.csv"
## 3 of 4: "episodes.csv"
## 4 of 4: "imdb.csv"

```

Caso necessário, junte os bancos utilizando apenas informações comuns à todos os bancos.

```
dr_who <- left_join( tuesdata$writers, tuesdata$directors, by = "story_number") %>%
  left_join(tuesdata$episodes, by = "story_number") %>%
  left_join(tuesdata$imdb, by = c("season_number" = "season", "episode_number" = "ep_num")) ##junta tod

glimpse(dr_who)
```

```
## Rows: 185
## Columns: 18
## $ story_number      <chr> "157", "158", "159", "160a", "160b", "161", "162", "16~
## $ writer            <chr> "Russell T Davies", "Russell T Davies", "Mark Gatiss",~
## $ director          <chr> "Keith Boak", "Euros Lyn", "Euros Lyn", "Keith Boak", ~
## $ era               <chr> "revived", "revived", "revived", "revived", "revived",~
## $ season_number     <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, ~
## $ serial_title      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ episode_number    <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, NA, 1, 2, 3~
## $ episode_title     <chr> "Rose", "The End of the World", "The Unquiet Dead", "A~
## $ type              <chr> "episode", "episode", "episode", "episode", "episode",~
## $ first_aired       <date> 2005-03-26, 2005-04-02, 2005-04-09, 2005-04-16, 2005--
## $ production_code   <chr> "1.1", "1.2", "1.3", "1.4", "1.5", "1.6", "1.7", "1.8"~
## $ uk_viewers        <dbl> 10.81, 7.97, 8.86, 7.63, 7.98, 8.63, 8.01, 8.06, 7.11,~
## $ rating.x          <dbl> 76, 76, 80, 82, 81, 84, 81, 83, 84, 85, 82, 86, 89, 84~
## $ duration          <dbl> 45, 44, 44, 45, 42, 45, 44, 45, 45, 45, 45, 45, 45, 60~
## $ air_date          <chr> "17 Mar. 2006", "17 Mar. 2006", "24 Mar. 2006", "31 Ma~
## $ rating.y          <dbl> 7.5, 7.5, 7.5, 7.0, 7.0, 8.7, 7.1, 8.4, 9.2, 9.2, 7.1,~
## $ rating_n          <dbl> 8301, 7279, 6804, 6543, 6337, 7211, 6255, 6889, 8383, ~
## $ desc              <chr> "When ordinary shop-worker Rose Tyler meets a mysterio~
```

Responda as seguintes questões, considerando duas casas decimais.

- A dupla de diretor e escritor formada por Joe Ahearne e Russell T Davies é responsável por quantos episódios?

```
dr_who %>%
  filter(writer == "Russell T Davies" & director == "Joe Ahearne") %>%
  drop_na(episode_number, season_number) %>% ##filtra por episodios que foram considerados
  nrow()
```

```
## [1] 3
```

- Somente em 2010 o escritor Steven Moffat escreveu mais de dois episódios. Quantos episódios ele escreveu neste ano?

```
dr_who %<>%
  mutate(first_aired = ymd(first_aired),
         ano = year(first_aired)) ## extrai o ano

dr_who %>%
  filter(writer == "Steven Moffat", ano == 2010) %>%
  drop_na(season_number, episode_number) %>%
  nrow()
```

```
## [1] 6
```

- O escritor Stephen Thompson trabalhou na série por algum tempo. Quantos anos ele trabalhou na série?

```
dr_who %>%
  filter(writer == "Stephen Thompson") %>%
  summarise(qntd_anos = max(ano) - min(ano)) ## calcula diff de tempo
```

```
## # A tibble: 1 x 1
##   qntd_anos
##   <dbl>
## 1       3
```

- Durante o tempo que o escritor Stephen Thompson escreveu roteiros para Dr. Who, ele foi responsável por quantos roteiros?

```
dr_who %>%
  filter(writer == "Stephen Thompson") %>%
  drop_na(season_number, episode_number) %>%
  summarise(n = n())
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1     3
```

- O diretor Ben Wheatley dirigiu episódios com duração média de quantos minutos?

```
dr_who %>%
  filter(director == "Ben Wheatley") %>%
  summarise(media_minutos = mean(duration, na.rm = T)) ##calcula a média
```

```
## # A tibble: 1 x 1
##   media_minutos
##   <dbl>
## 1       55.3
```