Os exercícios abaixo usam tabelas disponíveis nessa pasta (A-E). Todas as tabelas contém os mesmos dados, mas cada uma em um formato diferente. O desafio é conseguir fazer com que o R leia elas no formato correto. Caso uma dessas tabelas seja carregada com sucesso, ela deve possuir 6 observações, em 2 colunas. O código abaixo serve de teste (todas as linhas devem imprimir TRUE se estiver tudo certo):

dados = … # SEU CÓDIGO PRA LER A TABELA

# número de linhas e colunas: 6 e 2

nrow(dados) == 6; ncol(dados) == 2

# coluna X é numérica, coluna Y é texto/string/caracteres

class(dados$X) == “numeric”; class(dados$Y) == “character”

# média da coluna \*X\* é 30

mean(dados$X, na.rm = T) == 30

# quantidade de NAs na coluna \*Y\* é 1

sum(is.na(dados$Y)) == 1

1) Arquivo: tabelaA.csv

2) Arquivo: tabelaB.csv

3) Arquivo: tabelaC.tsv

4) Arquivo: tabelaD.xlsx

5) Arquivo: tabelaE.xlsx

Notem que no Windows, os caminhos são especificados com barra invertida (e.g. “Users\Kleber\Documentos”). O R usa barras normais (/), então cuidado na hora de copiar caminhos.

Os exercícios abaixo usam os verbos comuns de **dplyr** pra manipulação de dados.

**DATASET: mtcars (R base)**

library(tidyverse)

View(mtcars)

* Imprima na tela a coluna hp. Dica: use a função select().
* Imprima na tela todas as colunas, exceto a coluna hp.
* Crie um novo data frame mtcars2 que contém só as colunas hp, drat e wt. Dica: a função select() aceita vetores numéricos (e.g. select(5:6) seleciona da 5ª à 6ª coluna)
* Crie as seguintes colunas em mtcars: kpg (kilômetros por galão - use 1 milha = 1.6 km), kpl (kilômetros por litro - use 1 galão = 3.8 litros). Dica: use a função mutate(). Tente fazer escrevendo mutate() uma única vez.
* Crie um novo data frame mtcars3 que só tem os carros que tem mpg > 20. Dica: use a função filter().
* Crie um novo data frame mtcars4 que só tem os carros que tem mpg > 20 ou hp > 150. Dica: em R, a barra vertical (|) significa "ou" (o E comercial (&) significa "e", mas não precisa dele aqui).

**DATASET: msleep.csv (disponível na pasta)**

* Carregue os dados da tabela "msleep.csv" num data frame chamado msleep.
* Crie um data frame msleep\_primates contendo apenas os primatas.
* Faça o data frame conter apenas a coluna sleep\_total dos primatas. Escreva isso como continuação da linha de cima.
* Escreva uma linha que imprima na tela o número de primatas. Dica: use a função nrow().
* Crie uma variável sono\_medio que contenha a duração média do sono de primatas.
* Calcule a média de sleep\_total e a média de brainwt para todas as outras ordens (coluna order), excluindo primatas, e junte tudo em um data frame. Dica: use as funções group\_by() e summarise().