

Quiz 2

HTHSCI 1M03: Foundations of Data Science

Jasraj_Singh_Johal_johalj11_400434346

Thursday, March 7, 2024

Question 1

```
set.seed(34)
library(dplyr)
#>
#> Attaching package: 'dplyr'
#> The following objects are masked from 'package:stats':
#>
#>   filter, lag
#> The following objects are masked from 'package:base':
#>
#>   intersect, setdiff, setequal, union
library(magrittr)
library(tidyr)
#>
#> Attaching package: 'tidyr'
#> The following object is masked from 'package:magrittr':
#>
#>   extract
library(stringr)
library(readr)
library(gt)
```

a. Here I have also changed na values from 0 to NA

```
tib_gm <- read_csv("primary_forest_area_ha.csv", show_col_types = FALSE, na=c("0"))
tib_gm |> glimpse()
#> Rows: 231
#> Columns: 29
```

```

#> $ country <chr> "Aruba", "Afghanistan", "Angola", "Anguilla", "Albania", "Ando~
#> $ `1989` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "", NA, NA, "5.2~
#> $ `1990` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "", NA, NA, "5.2~
#> $ `1991` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1992` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1993` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1994` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1995` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1996` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1997` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1998` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `1999` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2000` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2001` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2002` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2003` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2004` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2005` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2006` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2007` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2008` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2009` <chr> NA, NA, NA, NA, "84.8k", NA, NA, NA, "1.74M", "17k", NA, NA, "~
#> $ `2010` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2011` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2012` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2013` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2014` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2015` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~
#> $ `2016` <chr> NA, NA, NA, NA, "84.8k", NA, "", NA, "1.74M", "17k", NA, NA, "~

```

b.

```

tib_gm_1 <- tib_gm |>
  pivot_longer(cols = -country,
               names_to = "yr",
               values_to = "hec")

tib_gm_1 |> glimpse()
#> Rows: 6,468
#> Columns: 3
#> $ country <chr> "Aruba", "Aruba", "Aruba", "Aruba", "Aruba", "Aruba", "Aruba",~

```

```
#> $ yr      <chr> "1989", "1990", "1991", "1992", "1993", "1994", "1995", "1996"~
#> $ hec     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
```

c.

```
tib_gm_1 <- tib_gm_1 |>
  mutate(yr = as.numeric(yr))
```

```
tib_gm_1 |> glimpse()
```

```
#> Rows: 6,468
#> Columns: 3
#> $ country <chr> "Aruba", "Aruba", "Aruba", "Aruba", "Aruba", "Aruba", "Aruba",~
#> $ yr      <dbl> 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 19~
#> $ hec     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
```

d.

```
tib_gm_1 <- tib_gm_1 |>
  filter(country == "Canada", yr == 2000)
```

```
tib_gm_1
```

```
#> # A tibble: 1 x 3
#>   country   yr hec
#>   <chr>     <dbl> <chr>
#> 1 Canada   2000 206M
```

e.

```
tib_gm_1 |> glimpse()
```

```
#> Rows: 1
#> Columns: 3
#> $ country <chr> "Canada"
#> $ yr      <dbl> 2000
#> $ hec     <chr> "206M"
```