Danish Kings

Kristoffer Segerstrøm

2025-03-05

The task here is to load your Danish Monarchs csv into R using the `tidyverse` toolkit, calculate and explore the kings' duration of reign with pipes `%>%` in `dplyr` and plot it over time.

Load the kings

Make sure to first create an `.Rproj` workspace with a `data/` folder where you place either your own dataset or the provided `kings.csv` dataset.

1. Look at the dataset that are you loading and check what its columns are separated by? (hint: open it in plain text editor to see)
2. Create a `kings` object in R with the different functions below and inspect the different outputs.

- `read.csv()`
- `read_csv()`
- `read.csv2()`
- `read_csv2()`

```
# FILL IN THE CODE BELOW and review the outputs

kings1 <- read.csv("data/Danish_kings")
kings2 <- read_csv("data/Danish_kings")
kings3 <- read.csv2("data/Danish_kings")
kings4 <- read_csv2("data/Danish_kings")
```

Answer: 1. Which of these functions is a `tidyverse` function? Read data with it below into a `kings` object

read_csv() and read_csv2() is a part of the tidyverse-package (specifically readr). read.csv() and read.csv2() belongs to base R.

2. What is the result of running `class()` on the `kings` object created with a tidyverse function. class(kings4)

[1] "tbl_df" "tbl" "data.frame"

3. How many columns does the object have when created with these different functions?

kings1 <- read_csv2("data/Danish_kings")
kings2 <- read_csv2("data/Danish_kings")
kings3 <- read_csv2("data/Danish_kings")
kings4 <- read_csv2("data/Danish_kings")

There is 11 columns

4. Show the dataset so that we can see how R interprets each column

glimpse(Danish_kings.csv) View(Danish_kings.csv)

```
# COMPLETE THE BLANKS BELOW WITH YOUR CODE, then turn the 'eval' flag in this chunk to TRUE.

kings <- Danish_kings

class(kings)


glimpse(kings)


View(kings)
```

Calculate the duration of reign for all the kings in your table

You can calculate the duration of reign in years with `mutate` function by subtracting the equivalents of your `startReign` from `endReign` columns and writing the result to a new column called `duration`. But first you need to check a few things:

- Is your data messy? Fix it before re-importing to R
- Do your start and end of reign columns contain NAs? Choose the right strategy to deal with them: `na.omit()`, `na.rm=TRUE`, `!is.na()`

Create a new column called `duration` in the kings dataset, utilizing the `mutate()` function from tidyverse. Check with your group to brainstorm the options.

The code I used

```
kings <- kings %>%
  filter(!is.na(start_reign) & !is.na(end_reign))
kings <- kings %>%
  mutate(duration = end_reign - start_reign)
glimpse(kings)
```

Calculate the average duration of reign for all rulers

Do you remember how to calculate an average on a vector object? If not, review the last two lessons and remember that a column is basically a vector. So you need to subset your `kings` dataset to the `duration` column. If you subset it as a vector you can calculate average on it with `mean()` base-R function. If you subset it as a tibble, you can calculate average on it with `summarize()` tidyverse function. Try both ways!

- You first need to know how to select the relevant `duration` column. What are your options?
- Is your selected `duration` column a tibble or a vector? The `mean()` function can only be run on a vector. The `summarize()` function works on a tibble.
- Are you getting an error that there are characters in your column? Coerce your data to numbers with `as.numeric()`.
- Remember to handle NAs: `mean(X, na.rm=TRUE)`

The code I used

```
kings %>%
  summarise(avg_duration = mean(duration, na.rm = TRUE))
```

The average duration of reign for all rulers was 20.2 years

How many and which kings enjoyed a longer-than-average duration of reign?

You have calculated the average duration above. Use it now to `filter()` the `duration` column in `kings` dataset. Display the result and also count the resulting rows with `count()`

The code I used

```
long_reign_kings <- kings %>%
  filter(duration > average_duration)
long_reign_kings %>% count()
print(long_reign_kings)
```

24 kings enjoyed a longer-than-average duration of reign

How many days did the three longest-ruling monarchs rule?

- Sort kings by reign `duration` in the descending order. Select the three longest-ruling monarchs with the `slice()` function
- Use `mutate()` to create `Days` column where you calculate the total number of days they ruled
- BONUS: consider the transition year (with 366 days) in your calculation!

#The code I used

```
top_3_kings <- kings %>%
  arrange(desc(duration)) %>%
  slice(1:3) %>%
  mutate(days = duration * 365.25)
print(top_3_kings)
glimpse(top_3_kings)
21915.00+18993.00+15705.75
```

The answer is 56613 days

And to submit this rmarkdown, knit it into html. But first, clean up the code chunks, adjust the date, rename the author and change the `eval=FALSE` flag to `eval=TRUE` so your script actually generates an output. Well done!