Class 4 RMarkdown Lab

Amy Winter

07 Sept 2022

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

The way you can create a file like this in RStudio is: File → New File → R Markdown and then using the default or using a template.

You can use hashtags to create headers. The more hastage the more indented the header

Below is a code chunk that will set up our packages and data.

```
library(readxl)
dtp <- read_excel("../data/DTP_Coverage_WHO.xlsx", sheet="Data")
head(dtp, n=5)
```

```
## # A tibble: 5 x 11
##   GROUP CODE  NAME   YEAR ANTIGEN ANTIG~1 COVER~2 COVER~3 TARGE~4  DOSES COVER~5
##   <chr> <chr> <chr> <dbl> <chr>   <chr>   <chr>   <chr>     <dbl>  <dbl>   <dbl>
## 1 COUN~ AFG   Afgh~  2021 DTPCV1  DTP-co~ ADMIN   Admini~ 1823296 1.70e6    93.4
## 2 COUN~ AFG   Afgh~  2021 DTPCV3  DTP-co~ ADMIN   Admini~ 1823296 1.48e6    81.2
## 3 COUN~ AFG   Afgh~  2020 DTPCV3  DTP-co~ ADMIN   Admini~ 1780564 1.52e6    85.3
## 4 COUN~ AFG   Afgh~  2020 DTPCV1  DTP-co~ ADMIN   Admini~ 1780564 1.73e6    97.2
## 5 COUN~ AFG   Afgh~  2019 DTPCV1  DTP-co~ ADMIN   Admini~ 1733907 1.64e6    94.4
## # ... with abbreviated variable names 1: ANTIGEN_DESCRIPTION,
## #   2: COVERAGE_CATEGORY, 3: COVERAGE_CATEGORY_DESCRIPTION, 4: TARGET_NUMBER,
## #   5: COVERAGE
```

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

Exercise 1

Create a chunk of code below that loads the tidyverse package (hint: use library function), but change the options so that I don't see the output of the loaded package.

Exercise 2

Create a code chunk below that creates a new data frame called `dtp1.nig.df` that only includes data from Nigeria (hint: NAME is "Nigeria") and DTP containing vaccine dose 1 (hint: ANTIGEN is "DTPCV1") and only WUENIC estimates (hint: COVERAGE_CATEGORY is "WUENIC"). Output the number of rows of data are in this new data frame, `dtp1.nig.df`?

```
dtp1.nig.df <- dtp %>%
  filter(NAME=="Nigeria" & ANTIGEN=="DTPCV1" & COVERAGE_CATEGORY=="WUENIC")
nrow(dtp1.nig.df)
```

```
## [1] 22
```

Exercise 3

Create a code chunk below that modifies the data frame such that it only induces the following variables: NAME, YEAR, ANTIGEN, COVERAGE_CATEGORY, COVERAGE. Output the first 8 rows of this data frame. Output the number of columns in this modified data frame

```
dtp1.nig.df <- dtp1.nig.df %>%
  select(NAME, YEAR, ANTIGEN, COVERAGE_CATEGORY, COVERAGE)
ncol(dtp1.nig.df)
```

```
## [1] 5
```

Exercise 4

Create a code chunk below that outputs the mean dtp dose 1 coverage from Nigeria across all years.

```
mean(dtp1.nig.df$COVERAGE)
```

```
## [1] 55.40909
```

```
dtp1.nig.df %>% select(COVERAGE) %>% summary()
```

```
##     COVERAGE
##  Min.   :42.00
##  1st Qu.:48.00
##  Median :51.50
##  Mean   :55.41
##  3rd Qu.:64.50
##  Max.   :73.00
```

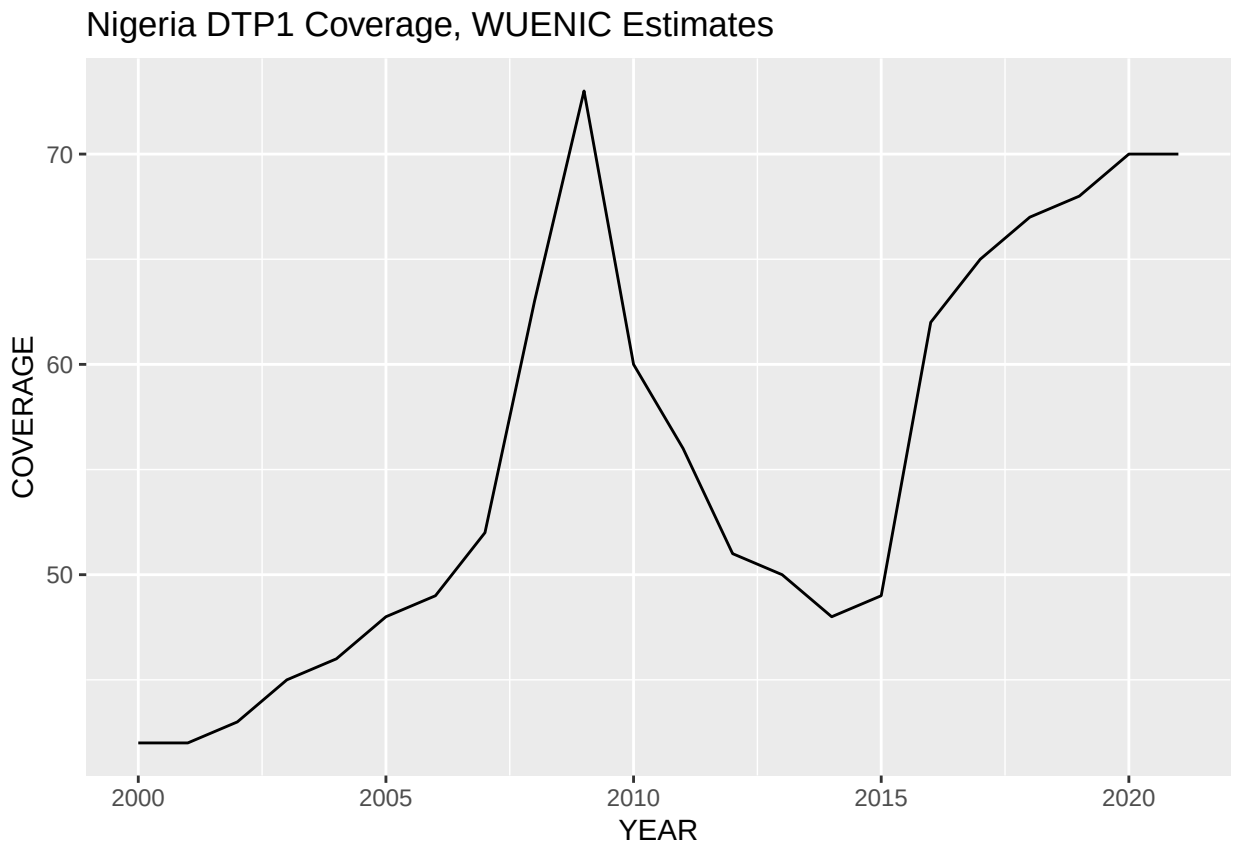
```
dtp1.nig.df %>% summarise(mean(COVERAGE))
```

```
## # A tibble: 1 x 1
##   `mean(COVERAGE)`
##              <dbl>
## 1             55.4
```

Exercise 5

Use the code chunk below to plot DTP coverage over time. Basically you just need to fill in the variable names for `x=` and `y=`. Below the code chunk interpret the plot for me.

```
ggplot(data = dtp1.nig.df) + geom_line(aes(x=YEAR, y=COVERAGE)) +
  labs(title = "Nigeria DTP1 Coverage, WUENIC Estimates")
```



Explanation of plot here. . .

Exercise 6

install.packages("tinytex") library(tinytex) install_tinytex()

Knit this file. Save the .pdf in output subdirectory of you RProject as "Day4_Lab_RMarkdown.pdf"