

# HW 4 607 Fall 2021- Vaccination

Mark Schmalfeld

9/26/2021

## 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>.

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. You can embed an R code chunk like this:

```
library(curl)
```

```
## Using libcurl 7.64.1 with LibreSSL/2.8.3
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.4      v dplyr   1.0.7
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   2.0.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter()      masks stats::filter()
## x dplyr::lag()         masks stats::lag()
## x readr::parse_date() masks curl::parse_date()
```

```
library(stringr)
library(ggplot2)
```

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

set up and look at df

```
# Github address for csv file
#https://raw.githubusercontent.com/schmalmr/607-Fall-2021/main/israeli_vaccination_data_analysis_startc

# target column names list for start up of project
#col_names = c("age", "not_vax_population", "fully_vax_population", "case_not_vax_100k", "case_fully_vax_100k")

urlfile<-"https://raw.githubusercontent.com/schmalmr/607-Fall-2021/main/israeli_vaccination_data_analysis_startc.csv"

# vax file holds the full set of text information
vax <- read_csv(url(urlfile), col_names= c("age", "not_vax_population", "fully_vax_population", "case_not_vax_100k", "case_fully_vax_100k"))
```

```
## Rows: 16 Columns: 6
```

```
## -- Column specification -----
## Delimiter: ","
## chr (6): age, not_vax_population, fully_vax_population, case_not_vax_100k, c...

##
## 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.
```

```
vax<-as_tibble(vax)

# Attempt to import and rename and adjust column to target data type - loss of some of the text data and
vaxnum <- read_csv(url(urlfile), col_names= c("age", "not_vax_population", "fully_vax_population", "case_not_vax_100k", "case_fully_vax_100k"))
  age = col_character(),
  not_vax_population = col_number(),
  fully_vax_population = col_number(),
  case_not_vax_100k = col_number(),
  case_fully_vax_100K = col_number(),
  efficacy_severe_disease = col_character()))

vaxnum<-as.tibble(vaxnum)
```

```
## Warning: 'as.tibble()' was deprecated in tibble 2.0.0.
## Please use 'as_tibble()' instead.
## The signature and semantics have changed, see '?as_tibble'.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_warnings()' to see where this warning was generated.
```

```
view(vaxnum)
glimpse(vaxnum)
```

```
## Rows: 16
## Columns: 6
```

```
## Warning: One or more parsing issues, see 'problems()' for details
```

```
## $ age <chr> "Age", NA, "<50", NA, ">50", NA, NA, NA, NA, "~
## $ not_vax_population <dbl> NA, NA, 1116834.0, 23.3, 186078.0, 7.9, NA, NA~
## $ fully_vax_population <dbl> NA, NA, 3501118.0, 73.0, 2133516.0, 90.4, NA, ~
## $ case_not_vax_100k <dbl> NA, 100, 43, NA, 171, NA, NA, NA, NA, NA, NA, ~
## $ case_fully_vax_100K <dbl> NA, 100, 11, NA, 290, NA, NA, NA, NA, NA, NA, ~
## $ efficacy_severe_disease <chr> "Efficacy", "vs. severe disease", NA, NA, NA, ~
```

```
vaxnum<-mutate(vaxnum,key=1:16)
#vaxnum<-rename(vaxnum, "text_description"="Age", "not_vax_population"="Population %", "fully_vax_popu
vaxnum<-mutate(vaxnum,"agenum"=age)

#Setup column for non vax and fully vax percent
vaxnum<-mutate(vaxnum, "nvp_pct"= not_vax_population)
vaxnum<-mutate(vaxnum, "fvp_pct"= fully_vax_population)

#filter to select target roles and add a lag column to include age with the percent nvp and fvp
vaxnum<-filter(vaxnum,key<=6)
vaxnum<-filter(vaxnum, key>2)
vaxnum<-mutate(vaxnum,lag(agenum,n=1))
vaxnum<-rename(vaxnum, "agenum_percent"= `lag(agenum, n = 1)`)

glimpse(vaxnum)
```

```
## Rows: 4
## Columns: 11
## $ age <chr> "<50", NA, ">50", NA
## $ not_vax_population <dbl> 1116834.0, 23.3, 186078.0, 7.9
## $ fully_vax_population <dbl> 3501118.0, 73.0, 2133516.0, 90.4
## $ case_not_vax_100k <dbl> 43, NA, 171, NA
## $ case_fully_vax_100K <dbl> 11, NA, 290, NA
## $ efficacy_severe_disease <chr> NA, NA, NA, NA
## $ key <int> 3, 4, 5, 6
## $ agenum <chr> "<50", NA, ">50", NA
## $ nvp_pct <dbl> 1116834.0, 23.3, 186078.0, 7.9
## $ fvp_pct <dbl> 3501118.0, 73.0, 2133516.0, 90.4
## $ agenum_percent <chr> NA, "<50", NA, ">50"
```

```
#Create two identical files to start filtering
vaxnum1<-vaxnum
vaxnum2<-vaxnum

#Filter to combine the specific columns with the ages
vaxnum1<-vaxnum1 %>% filter(grepl('50', age))
vaxnum2<-vaxnum2 %>% filter(grepl('50', agenum_percent))

vaxnum1<-select(vaxnum1,age, not_vax_population,fully_vax_population,case_not_vax_100k, case_fully_vax_
vaxnum2<-select (vaxnum2,nvp_pct, fvp_pct,agenum_percent)
vaxnum2<-rename(vaxnum2,"age"="agenum_percent")

#Join columns using the age >50 or <50
vaxnum_final<-full_join(vaxnum1,vaxnum2, by="age",copy_FALSE,suffix=c("age","agenum_percent"))

tibble(vaxnum_final)
```

```
## # A tibble: 2 x 7
##   age   not_vax_population fully_vax_populat~ case_not_vax_10~ case_fully_vax_1~
##   <chr>           <dbl>           <dbl>           <dbl>           <dbl>
## 1 <50             1116834             3501118             43             11
## 2 >50             186078             2133516             171            290
## # ... with 2 more variables: nvp_pct <dbl>, fvp_pct <dbl>
```

```
vaxnum_final$not_vax_population <- as.numeric(vaxnum_final$not_vax_population)

glimpse(vaxnum_final)
```

```
## Rows: 2
## Columns: 7
## $ age           <chr> "<50", ">50"
## $ not_vax_population <dbl> 1116834, 186078
## $ fully_vax_population <dbl> 3501118, 2133516
## $ case_not_vax_100k   <dbl> 43, 171
## $ case_fully_vax_100K <dbl> 11, 290
## $ nvp_pct            <dbl> 23.3, 7.9
## $ fvp_pct            <dbl> 73.0, 90.4
```

*#Q1*

*#Calculate the population. We can estimate the population eligible for the vaccine as the population of*

```
vaxnum_final<- mutate(vaxnum_final,nvp_pct1=nvp_pct/100)
vaxnum_final<- mutate(vaxnum_final,fvp_pct1=fvp_pct/100)
vaxnum_final<- mutate(vaxnum_final, population= not_vax_population/nvp_pct1+fully_vax_population/fvp_pct1)
view(vaxnum_final)
```

*# Population is approx 75% of the total population*  
(population\_vax<-sum(vaxnum\_final\$population))

```
## [1] 14304833
```

*# Total population estimate at population\_vax/(1-% not eligible for vax)*  
not\_elgible=0.25  
(total\_population=population\_vax/(1-not\_elgible))

```
## [1] 19073110
```

*# Total population is approx 19 Million based on these estimate.*

*#Q2*

*#Calcualate the efficacy vs Disease*

```
vaxnum_final<- mutate(vaxnum_final,Eff_Dis=1.0-(case_fully_vax_100K)/(case_not_vax_100k))

glimpse(vaxnum_final)
```

```
## Rows: 2
## Columns: 11
## $ age           <chr> "<50", ">50"
```

```
## $ not_vax_population    <dbl> 1116834, 186078
## $ fully_vax_population  <dbl> 3501118, 2133516
## $ case_not_vax_100k    <dbl> 43, 171
## $ case_fully_vax_100K  <dbl> 11, 290
## $ nvp_pct              <dbl> 23.3, 7.9
## $ fvp_pct              <dbl> 73.0, 90.4
## $ nvp_pctl             <dbl> 0.233, 0.079
## $ fvp_pctl             <dbl> 0.730, 0.904
## $ population           <dbl> 9589331, 4715502
## $ Eff_Dis              <dbl> 0.7441860, -0.6959064
```

```
# For those less than 50 yr of age the efficacy vs disease indicates the severe cases are higher for th
```

```
#Q3
```

```
# No - the direct comparison of just vaccination or no vaccination is not sufficient to provide the indi
```

```
““
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.