Irene Data analysis

Mwaura Patrick

1/27/2021

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3 v purrr 0.3.4  
## v tibble 3.0.5 v dplyr 1.0.3  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(dplyr)  
library(readxl)  
library(here)

## here() starts at E:/RESULTS/Mwaura/mwaura/WORK1/Covid19

library(skimr)  
library(kableExtra)

##   
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':  
##   
## group\_rows

library(AMR)  
library(scales)

##   
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':  
##   
## discard

## The following object is masked from 'package:readr':  
##   
## col\_factor

library(pipeR)  
library(cleaner)

## Set the working directory and load the data

setwd("E:/RESULTS/Mwaura/mwaura/WORK1/Covid19/Irene Work")  
IRENE\_DATA\_ANALYSIS\_R <- read\_csv("IRENE DATA ANALYSIS R.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## .default = col\_character(),  
## CAMEL\_ID\_NUMBER = col\_double(),  
## MDR = col\_double()  
## )  
## i Use `spec()` for the full column specifications.

## explore the data

skimr::skim(IRENE\_DATA\_ANALYSIS\_R)

Data summary

Name

IRENE\_DATA\_ANALYSIS\_R

Number of rows

302

Number of columns

33

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Column type frequency:

character

31

numeric

2

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Group variables

None

**Variable type: character**

skim\_variable

n\_missing

complete\_rate

min

max

empty

n\_unique

whitespace

ISOLATE\_ID

0

1.00

4

6

0

282

0

Ranch\_Community

0

1.00

8

13

0

5

0

Boma

267

0.12

6

16

0

10

0

Livestock\_production\_system

0

1.00

9

9

0

2

0

AGE

0

1.00

5

9

0

3

0

Gender

2

0.99

4

6

0

2

0

Sample\_Type

0

1.00

10

10

0

1

0

Transport\_media

0

1.00

10

10

0

1

0

mo

0

1.00

9

20

0

4

0

Colony\_Morphology

169

0.44

29

70

0

41

0

Gram\_staining

167

0.45

13

13

0

1

0

IMVIC\_Test

0

1.00

8

8

0

1

0

Amoxicillin\_clavulanic\_acid

166

0.45

1

1

0

2

0

Ampicillin

166

0.45

1

1

0

3

0

Ceftazidime

166

0.45

1

1

0

3

0

Cefotaxime

166

0.45

1

1

0

3

0

Ceftriaxone

166

0.45

1

1

0

3

0

Cefuroxime

166

0.45

1

2

0

4

0

Cefepime

166

0.45

1

1

0

3

0

Cefaclor

166

0.45

1

1

0

3

0

Tetracycline

166

0.45

1

1

0

3

0

Gentamicin

166

0.45

1

1

0

3

0

Spectinomycin

166

0.45

1

1

0

3

0

Chloramphenicol

166

0.45

1

1

0

1

0

Ciprofloxacin

166

0.45

1

1

0

3

0

Trimethoprim\_sulfamethoxazole

166

0.45

1

1

0

3

0

Norfloxacin

166

0.45

1

1

0

1

0

TEM

166

0.45

3

3

0

2

0

SHV

166

0.45

3

3

0

1

0

CTX-M

166

0.45

3

3

0

2

0

OXA

166

0.45

3

3

0

1

0

**Variable type: numeric**

skim\_variable

n\_missing

complete\_rate

mean

sd

p0

p25

p50

p75

p100

hist

CAMEL\_ID\_NUMBER

0

1.00

155.33

88.89

1

74.25

161.5

231.75

304

<U+2587><U+2587><U+2587><U+2587><U+2587>

MDR

228

0.25

2.23

2.06

0

1.00

1.0

4.00

7

<U+2587><U+2582><U+2583><U+2581><U+2582>

## Brief summary

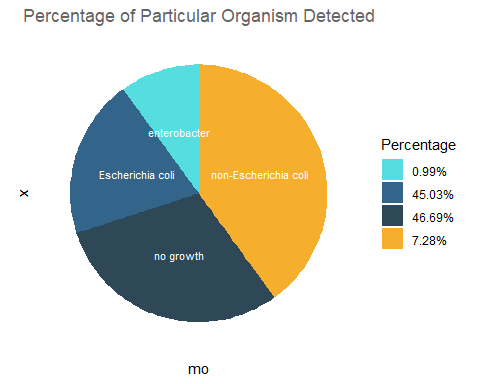
IRENE\_DATA\_ANALYSIS\_R %>% select(mo,CAMEL\_ID\_NUMBER) %>% group\_by(mo) %>% summarize(Count=n()) %>% mutate(Percentage=paste0(round(Count/sum(Count)\*100,2),"%"))

## # A tibble: 4 x 3  
## mo Count Percentage  
## \* <chr> <int> <chr>   
## 1 enterobacter 3 0.99%   
## 2 Escherichia coli 136 45.03%   
## 3 no growth 141 46.69%   
## 4 non-Escherichia coli 22 7.28%

IRENE\_DATA\_ANALYSIS\_R %>% freq(mo)

##   
##   
## \*\*Frequency table\*\*   
##   
## Class: character   
## Length: 302   
## Available: 302 (100%, NA: 0 = 0%)   
## Unique: 4   
##   
## Shortest: 9   
## Longest: 20  
##   
##   
## | |Item | Count| Percent| Cum. Count| Cum. Percent|  
## |:--|:--------------------|-----:|-------:|----------:|------------:|  
## |1 |no growth | 141| 46.69%| 141| 46.69%|  
## |2 |Escherichia coli | 136| 45.03%| 277| 91.72%|  
## |3 |non-Escherichia coli | 22| 7.28%| 299| 99.01%|  
## |4 |enterobacter | 3| 0.99%| 302| 100.00%|

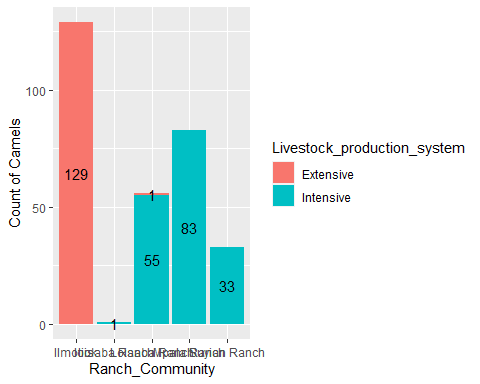
IRENE\_DATA\_ANALYSIS\_R %>% select(mo,CAMEL\_ID\_NUMBER) %>% group\_by(mo) %>% summarize(Count=n()) %>% mutate(Percentage=paste0(round(Count/sum(Count)\*100,2),"%")) %>% ggplot(aes(x="", y=mo, fill=Percentage)) + geom\_bar(stat="identity", width=1)+ coord\_polar("y", start=0) +theme\_classic()+labs( title = "Percentage of Particular Organism Detected")+scale\_fill\_manual(values=c("#55DDE0", "#33658A", "#2F4858", "#F6AE2D"))+theme(axis.line = element\_blank(),axis.text = element\_blank(),axis.ticks = element\_blank(), plot.title = element\_text(hjust = 0.5, color = "#666666"))+geom\_text(aes(y = mo, label = mo), position = position\_stack(vjust = .5),color = "white", size=3)

 ## cross tabulation by Ranch\_community vs the livestock production system

IRENE\_DATA\_ANALYSIS\_R %>% select(Ranch\_Community,Livestock\_production\_system) %>% pivot\_wider(names\_from = Livestock\_production\_system, values\_from= Livestock\_production\_system,values\_fn=list(Livestock\_production\_system=list))

## # A tibble: 5 x 3  
## Ranch\_Community Intensive Extensive   
## <chr> <list> <list>   
## 1 Mpala Ranch <chr [83]> <NULL>   
## 2 Ilmotiok <NULL> <chr [129]>  
## 3 Loisaba Ranch <chr [55]> <chr [1]>   
## 4 loisaba Ranch <chr [1]> <NULL>   
## 5 Suyian Ranch <chr [33]> <NULL>

IRENE\_DATA\_ANALYSIS\_R %>% select(Ranch\_Community,Livestock\_production\_system) %>% ggplot(aes(x=Ranch\_Community,fill=Livestock\_production\_system))+geom\_bar()+ylab("Count of Camels")+stat\_count(geom="text",aes(label=stat(count)), position = position\_stack(vjust = .5),color="black")

 ## MDRO analysis

IRENE\_DATA\_ANALYSIS\_R %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin) %>% mutate\_at(vars(Amoxicillin\_clavulanic\_acid:Norfloxacin), as.rsi)%>% mdro()%>% freq()

## Warning:   
## one unique value (covering 13.7%) could not be coerced and was considered  
## 'unknown': "non-Escherichia coli".  
## Use mo\_failures() to review it. Edit the `allow\_uncertain` argument if  
## needed (see ?as.mo).  
## You can also use your own reference data, e.g.:  
## as.mo("mycode", reference\_df = data.frame(own = "mycode", mo  
## = "B\_ESCHR\_COLI"))  
## mo\_name("mycode", reference\_df = data.frame(own = "mycode", mo  
## = "B\_ESCHR\_COLI"))

## Warning: NA introduced for isolates where the available percentage of antimicrobial  
## classes was below 50% (set with `pct\_required\_classes`)

##   
##   
## \*\*Frequency table\*\*   
##   
## Class: factor > ordered (numeric)   
## Length: 302   
## Levels: 4: Negative < Multi-drug-resistant (MDR) < Extensively drug-resistant ...   
## Available: 158 (52.32%, NA: 144 = 47.68%)   
## Unique: 2  
##   
##   
## | |Item | Count| Percent| Cum. Count| Cum. Percent|  
## |:--|:--------------------------|-----:|-------:|----------:|------------:|  
## |1 |Negative | 152| 96.20%| 152| 96.20%|  
## |2 |Multi-drug-resistant (MDR) | 6| 3.80%| 158| 100.00%|

## Resistance

IRENE\_DATA\_ANALYSIS\_R %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin) %>% mutate\_at(vars(Amoxicillin\_clavulanic\_acid:Norfloxacin), as.rsi) %>% group\_by(mo) %>% summarise(amoxiclav = resistance(Amoxicillin\_clavulanic\_acid), Ampicillin = resistance(Ampicillin), Norflaxin = resistance(Norfloxacin))

## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).

## # A tibble: 4 x 4  
## mo amoxiclav Ampicillin Norflaxin  
## \* <chr> <dbl> <dbl> <dbl>  
## 1 enterobacter NA NA NA  
## 2 Escherichia coli 0 0.0882 0  
## 3 no growth NA NA NA  
## 4 non-Escherichia coli NA NA NA

IRENE\_DATA\_ANALYSIS\_R %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin) %>% mutate\_at(vars(Amoxicillin\_clavulanic\_acid:Norfloxacin), as.rsi) %>% group\_by(mo) %>% summarise(amoxiclav = resistance(Amoxicillin\_clavulanic\_acid), Ampicillin = resistance(Ampicillin), Norflaxin = resistance(Norfloxacin)) %>% tidyr::pivot\_longer(-mo, names\_to = "antibiotic") %>% ggplot(aes(x = mo,y = value,fill = antibiotic)) +geom\_col(position = "dodge2")

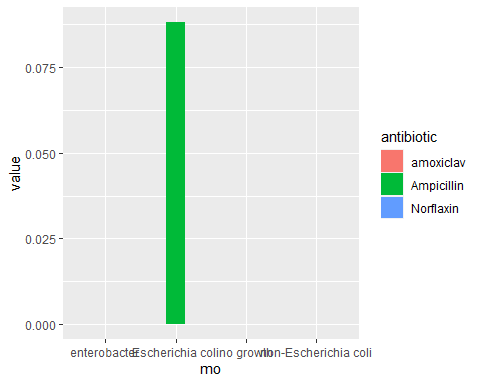
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).

## Warning: Removed 9 rows containing missing values (geom\_col).



IRENE\_DATA\_ANALYSIS\_R %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin) %>% mutate\_at(vars(Amoxicillin\_clavulanic\_acid:Norfloxacin), as.rsi) %>% group\_by(mo) %>% summarise(amoxiclav = resistance(Amoxicillin\_clavulanic\_acid), Ampicillin = resistance(Ampicillin), Norflaxin = resistance(Norfloxacin)) %>% tidyr::pivot\_longer(-mo, names\_to = "antibiotic") %>%ggplot(mapping = aes(x = mo,y = value)) +geom\_col() +labs(title = "Micro-organisms",subtitle = "Resistance to Antibiotics", x = "Organism",y = "Resistance Rates")

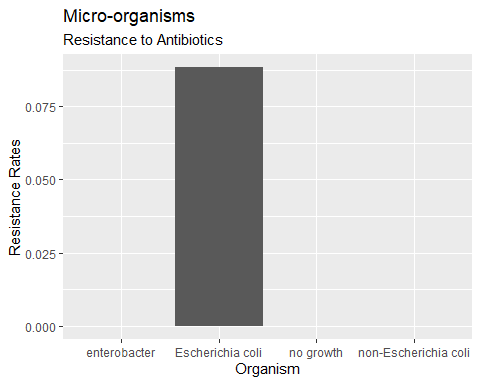
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).

## Warning: Removed 9 rows containing missing values (position\_stack).

 ## Conduct principal component analysis (PCA) for AMR

resistance\_data<-IRENE\_DATA\_ANALYSIS\_R %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin) %>% mutate\_at(vars(Amoxicillin\_clavulanic\_acid:Norfloxacin), as.rsi) %>% group\_by(mo) %>% summarise\_if(is.rsi, resistance) %>% select(mo,Amoxicillin\_clavulanic\_acid:Norfloxacin)

## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Amoxicillin\_clavulanic\_acid  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ampicillin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Ceftazidime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ceftazidime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ceftazidime (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Cefotaxime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Cefotaxime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Cefotaxime (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Ceftriaxone (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ceftriaxone (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ceftriaxone (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Cefuroxime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Cefuroxime (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Cefuroxime (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Cefepime (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Cefepime (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Cefepime (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Cefaclor (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Cefaclor (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Cefaclor (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Tetracycline (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Tetracycline (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Tetracycline (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Gentamicin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Gentamicin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Gentamicin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Spectinomycin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Spectinomycin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Spectinomycin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Chloramphenicol (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Chloramphenicol (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Chloramphenicol (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Ciprofloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ciprofloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Ciprofloxacin (`minimum` =  
## 30).

## Warning: Introducing NA: only 0 results available for Trimethoprim\_sulfamethoxazole  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Trimethoprim\_sulfamethoxazole  
## (`minimum` = 30).  
  
## Warning: Introducing NA: only 0 results available for Trimethoprim\_sulfamethoxazole  
## (`minimum` = 30).

## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).  
  
## Warning: Introducing NA: only 0 results available for Norfloxacin (`minimum` =  
## 30).

head(resistance\_data)

## # A tibble: 4 x 16  
## mo Amoxicillin\_cla~ Ampicillin Ceftazidime Cefotaxime Ceftriaxone  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 ente~ NA NA NA NA NA   
## 2 Esch~ 0 0.0882 0.0735 0.162 0.0294  
## 3 no g~ NA NA NA NA NA   
## 4 non-~ NA NA NA NA NA   
## # ... with 10 more variables: Cefuroxime <dbl>, Cefepime <dbl>, Cefaclor <dbl>,  
## # Tetracycline <dbl>, Gentamicin <dbl>, Spectinomycin <dbl>,  
## # Chloramphenicol <dbl>, Ciprofloxacin <dbl>,  
## # Trimethoprim\_sulfamethoxazole <dbl>, Norfloxacin <dbl>

resistance\_data[resistance\_data==0]<-NA  
  
view(resistance\_data)  
  
resistant\_data\_2<-data.frame(na.omit(resistance\_data))  
  
view(resistant\_data\_2)