Deconstruct, Reconstruct Web Report

Deconstruct, Reconstruct Web Report

tnathu-ai

Terminology Explained

Terminology	Description
Bacterium Name	Scientific name of the bacteria.
Causes	Diseases or conditions that the bacteria typically cause.
Gram-negative?	Indicates whether the bacterium is Gram-negative (‘yes’ or ‘no’).
Individual Drugs	Resistance levels to specific antibiotics, represented as percentages.
Remarks	Additional comments about antibiotic resistance.
Resistance Score	A scale from 1-4 indicating the level of concern, with higher scores being more concerning.
N Antibiotics or Classes	Number of antibiotics or classes showing at least some resistance.
Average % Resistance	Calculated average resistance score across all considered antibiotics.

Deconstruct

Original

The original data visualisation selected for the assignment was as follows:

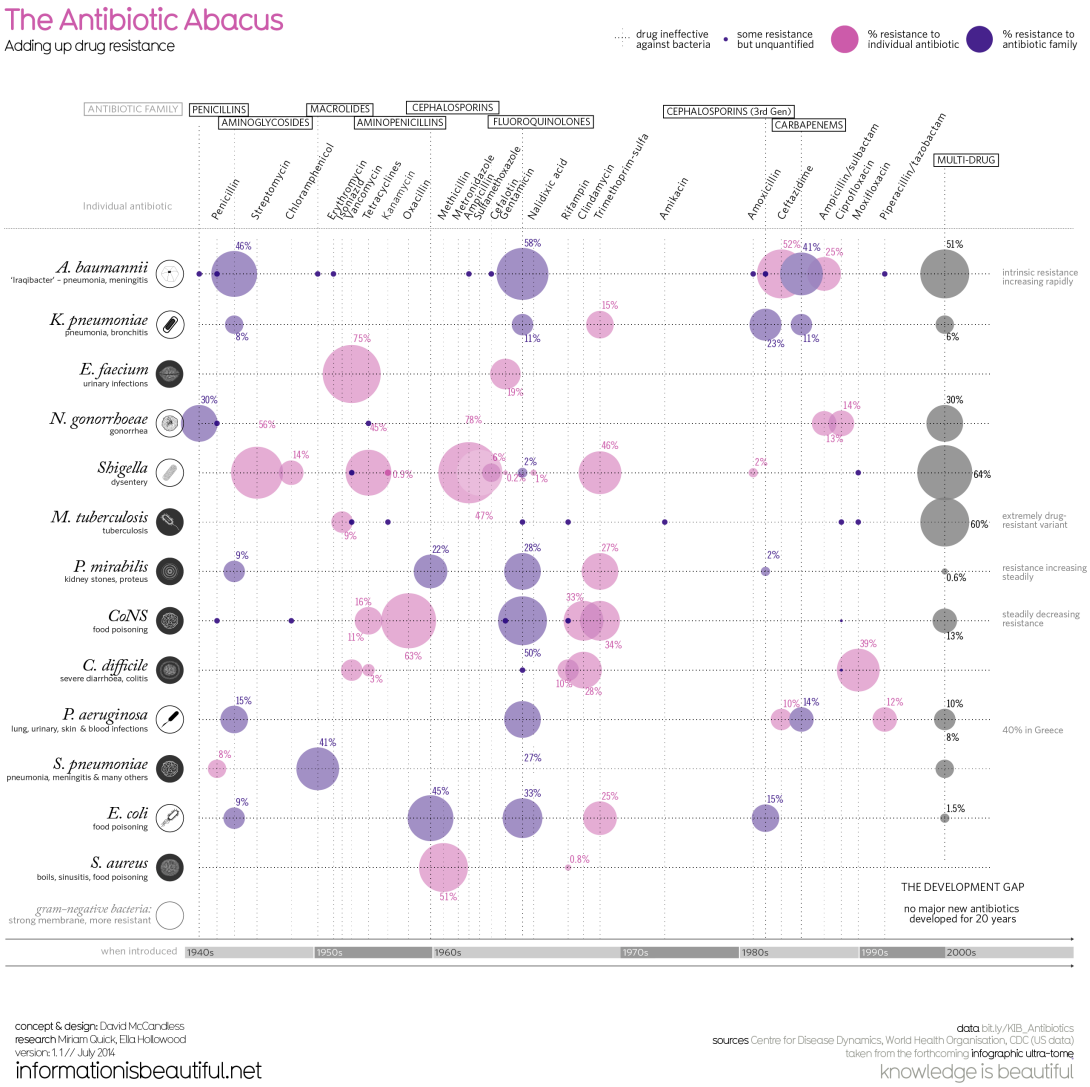

Figure 1. Original Visualization. Source: informationisbeautiful.net (https://informationisbeautiful.net/visualizations/antibiotic-resistance/) (July 2014).

Objective and Audience

The objective and audience of the original data visualisation chosen can be summarised as follows:

Objective The original data visualisation's objective is to provide a comprehensive overview of bacterial resistance to various antibiotics, highlighting the relative effectiveness of individual antibiotics against different bacteria, and emphasising the stagnation in developing new antibiotics over the years.

Audience

Given that the original visualisation is published on a visualisation-focused site, the main audience is the general public, especially those interested in understanding antibiotic resistance patterns. The visualisation does reference credible sources like the Centre for Disease Dynamics, World Health Organisation, and CDC. Such platforms typically target a wide demographic, ranging from enthusiasts who appreciate well-crafted data visualisations to individuals simply seeking to understand complex topics in a more accessible manner.

Critique

The visualisation chosen had the following three main issues:

+ Overwhelming Complexity:

Cluttered Display: The visualisation is densely populated with data, labels, and colours, potentially confusing viewers. Inconsistent circle area may mislead viewers. The visualisation must be improved to convey the relative importance due to non-uniform scales.

+ Perceptual Issues:

The pink and purple choices are problematic for those with certain colour blindness (Red-Blind/Protanopia and Monochromacy/Achromatopsia). Also, scattered small text annotations hinder quick understanding.Varying circle sizes and overlaps misrepresent data. Shading and density variations are not optimised. Different scales for bacteria and antibiotics hinder direct comparisons.

+ Audience Consideration:

Excessive detail and data points can obscure the main message for non-specialists. The distinction between individual antibiotics and antibiotic families is unclear. The line indicating “drug ineffective against bacteria” is easily misconstrued as a grid for the plot.

Reconstruct

Code

The following code was used to fix the issues identified in the original.

```
library(ggplot2)
library(reshape2)
library(RColorBrewer)
library(grid)
```

Pre-processing

Data: In July 2014, the dataset on antibiotic resistance was sourced from informationisbeautiful.net (https://docs.google.com/spreadsheets/d/1QMSQDr0vx_NuMkbSb-9wjUCBpIMZpYvxASTJh1M-NhI/edit#gid=0). We have independently verified the integrity and accuracy of this data in 2023 against original sources, including the Centre for Disease Dynamics, World Health Organisation, and CDC (US data). Subsequent data processing and formatting to generate a tailored dataframe were conducted using the R programming language.

```
# Data
antibiotic_resistance_df <- data.frame(
  Bacterium_name = c('A. baumanii', 'K. pneumoniae', 'E. faecium', 'N. gonorrhoeae',
'Shigella',
                     'M. tuberculosis', 'P. mirabilis', 'CoNS', 'C. difficile',
                     'P. aeruginosa', 'S. pneumoniae', 'E. coli', 'S. aureus'),

  Causes = c('pneumonia, meningitis', 'pneumonia, bronchitis, urinary infections', 'u
rinary infections',
             'gonorrhoea', 'dysentry', 'tuberculosis', 'kidney stones, proteus', 'foo
d poisoning',
             'severe diarrhoea, colitis', 'lung, urinary, skin, wound & blood infecti
ons',
             'pneumonia, meningitis & many other infections', 'food poisoning', 'boil
s, sinusitis, food poisoning'),

  Gram_negative = c('yes', 'yes', 'no', 'yes', 'yes', 'no', 'yes', 'no', 'no', 'yes',
'no', 'yes', 'no'),

  Penicillin = c(46, 8, NA, 30, NA, NA, 9, NA, NA, 15, 8, 9, NA),
  Streptomycin = c(NA, NA, 75, NA, 56, 9, NA, NA, NA, NA, NA, NA, NA),
  Chloramphenicol = c(NA, NA, NA, NA, 14, NA, NA, NA, 3, NA, NA, NA, NA),
  Erythromycin = c(58, 11, 19, NA, 45, NA, 22, 16, 11, 27, NA, 45, 51),
  Isoniazid = c(NA, NA, NA, NA, 0.9, NA, NA, NA, NA, NA, NA, NA, 0.8),
  Vancomycin = c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Tetracycline = c(NA, NA, NA, NA, 78, NA, NA, NA, NA, NA, NA, NA, NA),
  Kanamycin = c(NA, NA, NA, NA, 47, NA, NA, NA, NA, NA, NA, NA, NA),
  Oxacillin = c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Methicillin = c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Metronidazole = c(NA, NA, NA, NA, NA, NA, NA, NA, 10, NA, NA, NA, NA),
  Ampicillin = c(NA, NA, NA, NA, NA, NA, NA, NA, 28, NA, NA, NA, NA),
  Sulfamethoxazole = c(NA, NA, NA, NA, 6, NA, NA, NA, NA, NA, NA, 15, NA),
  Cefalotin = c(NA, NA, NA, NA, 0.2, NA, 2, NA, NA, NA, NA, NA, NA),
  Gentamicin = c(NA, NA, NA, NA, 2, NA, 1, NA, NA, 10, NA, NA, NA),
  Nalidixic_acid = c(NA, NA, NA, NA, 1, NA, NA, NA, NA, 14, NA, NA, NA),
  Rifampicin = c(NA, NA, NA, NA, 46, NA, NA, NA, NA, NA, NA, NA, NA),
  Clindamycin = c(NA, NA, NA, NA, 0.06, NA, NA, NA, NA, NA, NA, NA, NA),
  Trimethoprim_sulfa = c(NA, NA, NA, NA, 2, NA, NA, NA, NA, NA, NA, NA, NA),
  Amikacin = c(NA, NA, NA, NA, 0, NA, NA, NA, NA, NA, NA, NA, NA),
  Amoxicillin_clavulanic_acid = c(NA, NA, NA, NA, 0.06, NA, NA, NA, NA, NA, NA, NA, N
A),
  Ceftriaxone = c(NA, NA, NA, NA, 0.06, NA, NA, NA, NA, NA, NA, NA, NA),
  Ceftazidime = c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Imipenem = c(52, NA, NA, NA, 64, 60, 0.6, 13, 39, 12, 8, 1.5, NA),
  Ceftiofur = c(41, NA, NA, NA, NA, NA, NA, NA, NA, 10, NA, NA, NA),
  Ciprofloxacin = c(25, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Moxifloxacin = c(NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA),
  Piperacillin_tazobactam = c(51, 6, NA, 30, NA, NA, NA, NA, NA, 10, NA, NA, NA)
)

# Antibiotic Introduction Dates
antibiotic_intro_dates <- data.frame(
  Antibiotic = c('Penicillin', 'Streptomycin', 'Chloramphenicol', 'Erythromycin', 'Is
oniazid',
                 'Vancomycin', 'Tetracycline', 'Kanamycin', 'Oxacillin', 'Methicilli
```

```
n', 'Metronidazole',
                 'Ampicillin', 'Sulfamethoxazole', 'Cefalotin', 'Gentamicin', 'Nalidi
xic_acid',
                 'Rifampicin', 'Clindamycin', 'Trimethoprim_sulfa', 'Amikacin', 'Amox
icillin_clavulanic_acid',
                 'Ceftriaxone', 'Ceftazidime', 'Imipenem', 'Ceftiofur', 'Ciprofloxaci
n', 'Moxifloxacin',
                 'Piperacillin_tazobactam'),

  Date_introduced = c(1942, 1944, 1949, 1952, 1952, 1955, 1955, 1958, 1959, 1960, 196
0, 1961, 1961, 1964,
                      1964, 1967, 1967, 1968, 1968, 1976, 1981, 1982, 1983, 1985, 198
7, 1987, 1987, 1989)
)

# View the dataframes
head(antibiotic resistance df)
```

```
##    Bacterium_name                                 Causes Gram_negative
## 1     A. baumanii                   pneumonia, meningitis           yes
## 2   K. pneumoniae pneumonia, bronchitis, urinary infections           yes
## 3      E. faecium                   urinary infections            no
## 4  N. gonorrhoeae                             gonorrhoea           yes
## 5         Shigella                               dysentry           yes
## 6 M. tuberculosis                            tuberculosis            no
##   Penicillin Streptomycin Chloramphenicol Erythromycin Isoniazid Vancomycin
## 1         46           NA              NA           58        NA         NA
## 2          8           NA              NA           11        NA         NA
## 3         NA           75              NA           19        NA         NA
## 4         30           NA              NA           NA        NA         NA
## 5         NA           56              14           45       0.9         NA
## 6         NA            9              NA           NA        NA         NA
##   Tetracycline Kanamycin Oxacillin Methicillin Metronidazole Ampicillin
## 1           NA        NA        NA          NA            NA         NA
## 2           NA        NA        NA          NA            NA         NA
## 3           NA        NA        NA          NA            NA         NA
## 4           NA        NA        NA          NA            NA         NA
## 5           78        47        NA          NA            NA         NA
## 6           NA        NA        NA          NA            NA         NA
##   Sulfamethoxazole Cefalotin Gentamicin Nalidixic_acid Rifampicin Clindamycin
## 1               NA        NA         NA             NA         NA          NA
## 2               NA        NA         NA             NA         NA          NA
## 3               NA        NA         NA             NA         NA          NA
## 4               NA        NA         NA             NA         NA          NA
## 5                6       0.2          2              1         46        0.06
## 6               NA        NA         NA             NA         NA          NA
##   Trimethoprim_sulfa Amikacin Amoxicillin_clavulanic_acid Ceftriaxone
## 1                 NA       NA                          NA          NA
## 2                 NA       NA                          NA          NA
## 3                 NA       NA                          NA          NA
## 4                 NA       NA                          NA          NA
## 5                  2        0                        0.06        0.06
## 6                 NA       NA                          NA          NA
##   Ceftazidime Imipenem Ceftiofur Ciprofloxacin Moxifloxacin
## 1          NA       52        41            25           NA
## 2          NA       NA        NA            NA           NA
## 3          NA       NA        NA            NA           NA
## 4          NA       NA        NA            NA           NA
## 5          NA       64        NA            NA           NA
## 6          NA       60        NA            NA           NA
##   Piperacillin_tazobactam
## 1                      51
## 2                       6
## 3                      NA
## 4                      30
## 5                      NA
## 6                      NA
```

```
head(antibiotic_intro_dates)
```

```
##        Antibiotic Date_introduced
## 1      Penicillin            1942
## 2    Streptomycin            1944
## 3 Chloramphenicol            1949
## 4    Erythromycin            1952
## 5       Isoniazid            1952
## 6      Vancomycin            1955
```

Reconstruction

The following plot fixes the main issues in the original.

This heatmap provides a quick visual representation of resistance percentages for various bacteria against different antibiotics. The darker the color, the higher the resistance. The antibiotics are also ordered based on their introduction dates, which can give an idea about the evolution of resistance over time.

```
# Convert antibiotic_resistance_df to long format for ease of plotting.
melted_df <- melt(antibiotic_resistance_df, id.vars = c("Bacterium_name", "Causes",
"Gram_negative"))

# Remove duplicate antibiotic introduction dates to avoid redundancy.
antibiotic_intro_dates <- antibiotic_intro_dates[!duplicated(antibiotic_intro_dates$A
ntibiotic),]

# Order the antibiotics by their introduction date for better visualization.
antibiotic_intro_dates$Antibiotic <- factor(antibiotic_intro_dates$Antibiotic,
                                            levels = antibiotic_intro_dates$Antibiotic
[order(antibiotic_intro_dates$Date_introduced)])

# Merge the long format dataframe with antibiotic introduction dates for comprehensiv
e data.
melted_df <- merge(melted_df, antibiotic_intro_dates, by.x = "variable", by.y = "Anti
biotic", all.x = TRUE)

# Modify antibiotic names to include introduction dates and replace underscores with
spaces.
melted_df$display_name <- paste(gsub("_", " ", melted_df$variable), "(", melted_df$Da
te_introduced, ")")

# Set the order of display names based on their respective antibiotics.
melted_df$display_name <- factor(melted_df$display_name, levels = unique(melted_df$di
splay_name[order(melted_df$variable)]))

# Build the heatmap visualization.
p <- ggplot(melted_df, aes(x = display_name, y = Bacterium_name)) +
  geom_tile(aes(fill = value), color = "white", width = 0.5, height = 0.5) +  # Use t
iles for heatmap.
  scale_fill_gradient(low = "lightblue", high = "darkblue", name = "Resistance %") +
# Set color gradient for resistance.
  labs(y = "Bacterium") +  # Label y-axis.
  theme(  # Fine-tune visualization appearance.
        axis.text.x = element_text(angle = 90, hjust = 1, size = 15, margin = margin
(t = 10)),
        axis.text.y = element_text(size = 15, margin = margin(r = 10)),
        axis.title.y = element_text(face = "bold", size = 22, margin = margin(r = 2
0)),
        plot.margin = margin(150, 10, 100, 10),
        panel.grid.major = element_line(colour = "gray", size = 0.25),
        plot.background = element_rect(fill = "white", color = NA),
        panel.background = element_rect(fill = "white", color = NA),
        legend.text = element_text(size = 16),
        legend.title = element_text(size = 18, face = "bold")
  ) +
  scale_x_discrete(position = "top", name = NULL)  # Place x-axis at the top and remo
ve its default label.

# Output the plot.
print(p)

# Manually add the x-axis title.
grid::grid.text("Individual Antibiotic", x = 0.01, y = 0.77, just = c("left", "cente
```

```
r"), gp = grid::gpar(col = "black", fontsize = 22, fontface = "bold"))

# Set the main title for the visualization.
grid::grid.text("ANTIBIOTIC CHALLENGE: BACTERIAL RESISTANCE OVERVIEW", x = 0.01, y =
0.975, just = c("left", "top"), gp = grid::gpar(col = "blue", fontsize = 30, fontface
= "bold"))

# Provide citation for data sources in a darker gray.
grid::grid.text("Sources: Centre for Disease Dynamics, World Health Organisation, CDC
(US data) | Designed by Thu Tran", x = 0.01, y = 0.02, just = c("left", "bottom"), gp
= grid::gpar(col = "#2E2E2E", fontsize = 13, fontface = "italic"))

# State origin of the initial image and data collection date.
grid::grid.text("Original image: informationisbeautiful.net | Data collected by July
2014", x = 0.01, y = 0.01, just = c("left", "bottom"), gp = grid::gpar(col = "#2E2E2
E", fontsize = 13, fontface = "italic"))
```

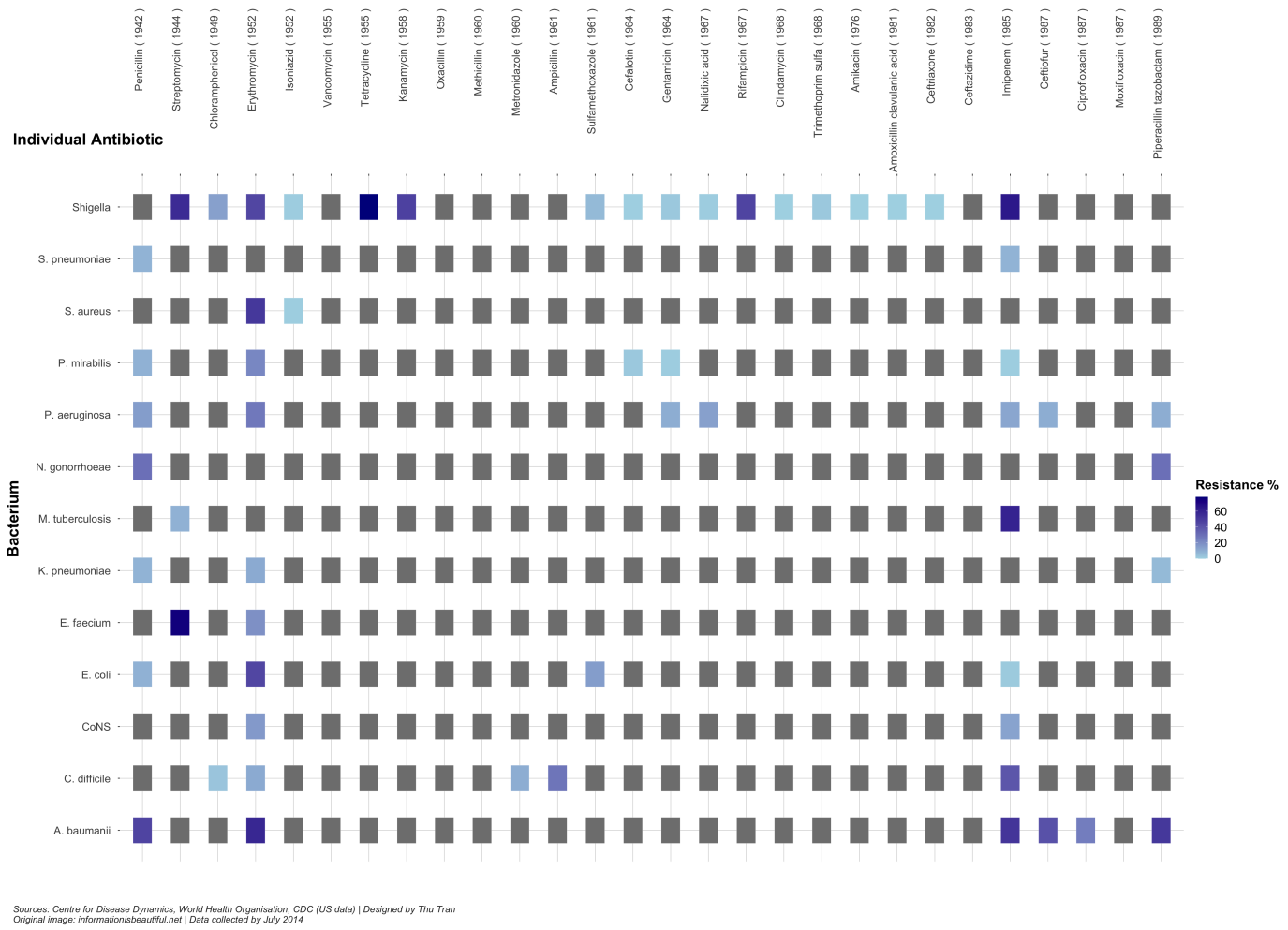

Figure 2: Antibiotic Resistance Heatmap (Sep 2023)

```
# Save the final visualization to an external image file.
ggsave(filename = "../images/antibiotic_resistance_plot.png", width = 25, height = 2
0)
```

=> The visualisation serves as a bridge between complex data and a wide audience range, from specialists to the general public, aiming to educate them about the pressing issue of antibiotic resistance. It achieves this by showcasing a heatmap of bacterial resistance percentages for each antibiotic, contextualised with their

respective introduction dates. This design choice communicates the increasing challenge of bacterial resistance to newer antibiotics.

References

- Information Is Beautiful. (n.d.). Antibiotic Resistance. https://informationisbeautiful.net/visualizations/antibiotic-resistance/ (https://informationisbeautiful.net/visualizations/antibiotic-resistance/)
- Pneumococcal Disease and Antibiotic Resistance | CDC. (2022, January 24). https://www.cdc.gov/pneumococcal/clinicians/drug-resistance.html (https://www.cdc.gov/pneumococcal/clinicians/drug-resistance.html)
- About Division of Bacterial Diseases | CDC. (2021, November 30). https://www.cdc.gov/ncird/dbd.html (https://www.cdc.gov/ncird/dbd.html)
- Immunization and Respiratory Diseases (NCIRD) Home. (2020). https://www.cdc.gov/ncird/ (https://www.cdc.gov/ncird/)
- Antibiotic - New World Encyclopedia. (n.d.). https://www.newworldencyclopedia.org/entry/Antibiotic (https://www.newworldencyclopedia.org/entry/Antibiotic)
- Antibiotics | Society | The Guardian. (n.d.). https://www.theguardian.com/society/antibiotics (https://www.theguardian.com/society/antibiotics)
- Drug-Resistant Superbugs Sweeping Across Europe - Slashdot. (2011, November 18). https://science.slashdot.org/story/11/11/18/0430245/drug-resistant-superbugs-sweeping-across-europe (https://science.slashdot.org/story/11/11/18/0430245/drug-resistant-superbugs-sweeping-across-europe)
- NHS. (2018, August 16). Medicines A-Z. https://www.nhs.uk/medicines/ (https://www.nhs.uk/medicines/)
- Wise, R., Hart, T., Cars, O., Streulens, M., Helmuth, R., Huovinen, P., & Sprenger, M. (1998). Antimicrobial resistance. BMJ, 317(7159), 609–610. https://doi.org/10.1136/bmj.317.7159.609 (https://doi.org/10.1136/bmj.317.7159.609)
- World Health Organisation. (2018, February 15). Antimicrobial resistance. https://www.who.int/en/news-room/fact-sheets/detail/antimicrobial-resistance (https://www.who.int/en/news-room/fact-sheets/detail/antimicrobial-resistance)
- Simonsen, G. S. (2018). Antimicrobial resistance surveillance in Europe and beyond. Eurosurveillance, 23(42). https://doi.org/10.2807/1560-7917.es.2018.23.42.1800560 (https://doi.org/10.2807/1560-7917.es.2018.23.42.1800560)
- Rogers, S. (2013, March 11). Antibiotics and drug resistance: how do we compare? The Guardian. https://www.theguardian.com/news/datablog/2013/mar/11/antibiotic-drug-resistance (https://www.theguardian.com/news/datablog/2013/mar/11/antibiotic-drug-resistance)
- Wickham, H. (2010). A Layered Grammar of Graphics. Journal of Computational and Graphical Statistics, 19(1), 3-28. https://doi.org/10.1198/jcgs.2009.07098 (https://doi.org/10.1198/jcgs.2009.07098)
- Riederer, Y. X., Dervieux, C., & Emily. (n.d.). Acknowledgments | R Markdown Cookbook. https://bookdown.org/yihui/rmarkdown-cookbook/acknowledgments.html (https://bookdown.org/yihui/rmarkdown-cookbook/acknowledgments.html)

- Create Elegant Data Visualisations Using the Grammar of Graphics. (2019). https://ggplot2.tidyverse.org/ (https://ggplot2.tidyverse.org/)
- Create Elegant Data Visualisations Using the Grammar of Graphics [R package ggplot2 version 3.2.1]. (2019). https://cran.r-project.org/package=ggplot2 (https://cran.r-project.org/package=ggplot2)

Acknowledgments

- Baglin, J. (2023). Data Visualisation: From Theory to Practice. In Data Visualisation and Communication (2350) [Online Textbook]. RMIT University. Retrieved from https://dark-star161610.appspot.com/secured/_book/index.html (https://dark-star161610.appspot.com/secured/_book/index.html)