Data Exploration

Jessica Gomez Lemus

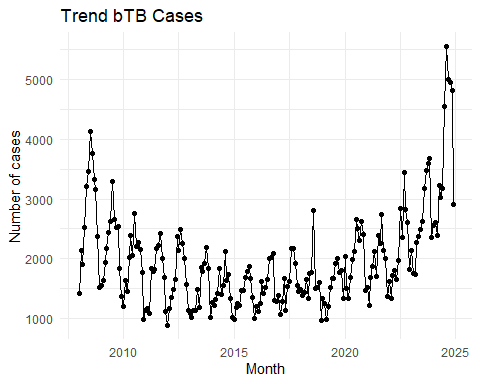
Table of contents

# 1. Load and Prepare Data

source("R\_files/01\_libraries.R")  
source("R\_files/02\_data.R")

# 2. General trend over time

avg\_per\_month <- cases\_per\_month %>%  
 group\_by(month) %>%  
 summarise(avg\_cases = mean(num\_cases), .groups = "drop")  
  
ggplot(cases\_per\_month, aes(x = year\_month, y = num\_cases)) +  
 geom\_line() +  
 geom\_point() +  
 labs(title = "Trend bTB Cases", x = "Month", y = "Number of cases") +  
 theme\_minimal()



# 3. Seasonality by Month

ggplot(cases\_per\_month, aes(x = year, y = num\_cases, group = 1)) +  
 geom\_line() +  
 geom\_point(size = 1) +  
 geom\_hline(  
 data = avg\_per\_month,  
 aes(yintercept = avg\_cases),  
 linetype = "solid",  
 color = "blue"  
 ) +  
 facet\_grid(. ~ month, scales = "free\_x", space = "free") +  
 labs(  
 title = glue::glue(  
 "Seasonality of bTB by month (2008-{max(cases\_per\_month$year)})"  
 ),  
 #add years auto  
 x = "Year",  
 y = "No. of bTB cases"  
 ) +  
 theme\_minimal() +  
 theme(  
 strip.background = element\_rect(fill = "grey90", color = NA),  
 strip.text = element\_text(size = 12, face = "bold"),  
 axis.text.x = element\_text(  
 angle = 90,  
 hjust = 1,  
 face = "bold"  
 )  
 )

