

# vis.R

joebrew

Tue Feb 14 18:11:14 2017

```
library(readxl)
library(tidyverse)

## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr

## Conflicts with tidy packages -----

## filter(): dplyr, stats
## lag():      dplyr, stats

# Read in
df <- read_excel('stupid graph.xlsx')

# Define categories
categories <- c(rep('All TB cases', 3),
                 rep('Lab-confirmed cases', 3),
                 rep('Clinically-diagnosed cases', 3))

unique_categories <- unique(categories)

counter <- 1
for (i in 1:length(unique_categories)){
  this_category <- unique_categories[i]
  x <- read_excel('stupid graph.xlsx',
                  skip = 1)[, which(categories == this_category)]
  x$category <- this_category
  x$id <- counter:(nrow(x) + counter - 1)
  if(i == 1){
    df <- x
  } else {
    df <- bind_rows(x,
                    df)
  }
}

# Gather
df <- gather(df,
             day,
             value,
             `Day 0` : `Day 60`)
df$day <- as.numeric(gsub('Day ', '', df$day))

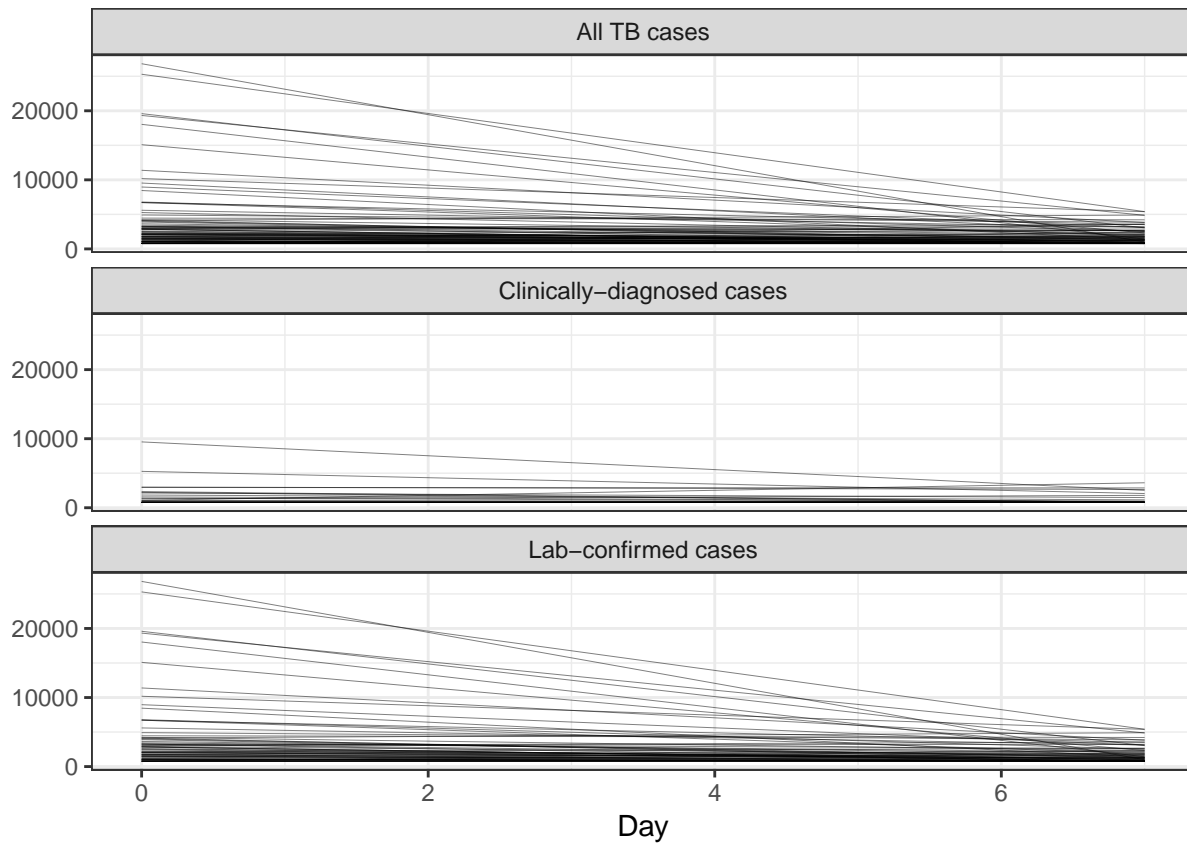
# Plot
ggplot(data = df %>%
```

```

    filter(day <= 7),
    aes(x = day,
        y = value,
        group = id)) +
    geom_line(lwd = 0.1,
              alpha = 0.5) +
    facet_wrap(~category, ncol = 1) +
    theme_bw() +
    xlab('Day') +
    ylab('')

```

## Warning: Removed 132 rows containing missing values (geom\_path).



```
ggsave('plot.pdf')
```

## Saving 6.5 x 4.5 in image

## Warning: Removed 132 rows containing missing values (geom\_path).

*# Statistical test for lab vs clinic*

```

fit <- lm(value ~ day + category,
          data = df %>%
            filter(day <= 7,
                  category != 'All TB cases'))
summary(fit)

```

##

## Call:

## lm(formula = value ~ day + category, data = df %>% filter(day <=

```

##      7, category != "All TB cases"))
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -3303.2 -1567.1  -831.2   376.1 22729.8
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)         2348.14     504.28   4.656 5.87e-06 ***
## day                -277.61      70.67  -3.928 0.000118 ***
## categoryLab-confirmed cases 1736.07     531.67   3.265 0.001287 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3516 on 199 degrees of freedom
## (634 observations deleted due to missingness)
## Multiple R-squared:  0.1159, Adjusted R-squared:  0.107
## F-statistic: 13.05 on 2 and 199 DF, p-value: 4.741e-06

```