Healthcare sector trends

Ian Mutwiri

2022-09-06

The following plots show the distribution of core healthcare workers across all counties in Kenya in 2013 and 2019. I have attached the code for reproducibility and also to demonstrate the improvements in data visualization that I have made.

Load the required packages.

```
library(tidyverse)
library(patchwork)
```

Load the health data.

```
health_13<- read_csv('health_workers_13.csv') %>%
  rename(core_per_10k_13=`Core health workforce per 10,000\rpopulation`) %>%
  mutate( County=str_to_title(County),
          County=as_factor(County))

health_19<- read_csv('health_workers_19.csv')%>%
  rename(core_per_10k_19=`Core health workforce per 10,000\rpopulation`) %>%
  mutate(County=case_when(
    County=='Tharaka-nithi'~'Tharaka Nithi',
    County=="Murang'a"~'Muranga',
    County=='Total'~'Kenya',
    TRUE~County
  ),
          County=str_to_title(County),
          County=as_factor(County))
```

Wrangle the data that it is in a form that can be plotted.

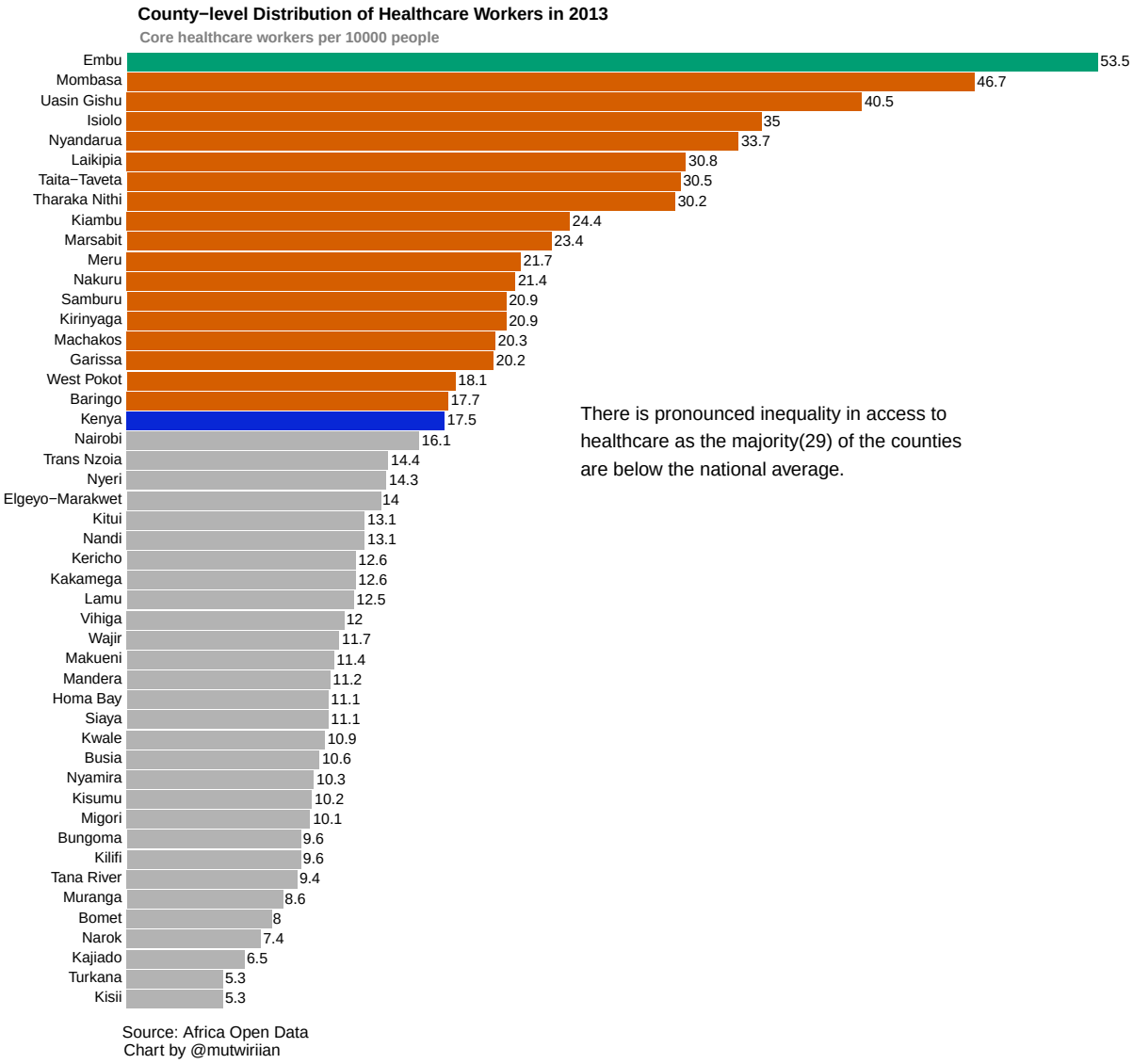
```
health <- health_13 %>%
  full_join(health_19,by='County') %>%
  pivot_longer(cols = 2:3,names_to = 'Year')
```

Set the plot annotations

```
note_13 <- "There is pronounced inequality in access to\nhealthcare as the majority(29) of the counties'
note_19 <- "Inequality still exists but overall there have\nbeen improvements in the number of core heal
```

Plot the 2013 data.

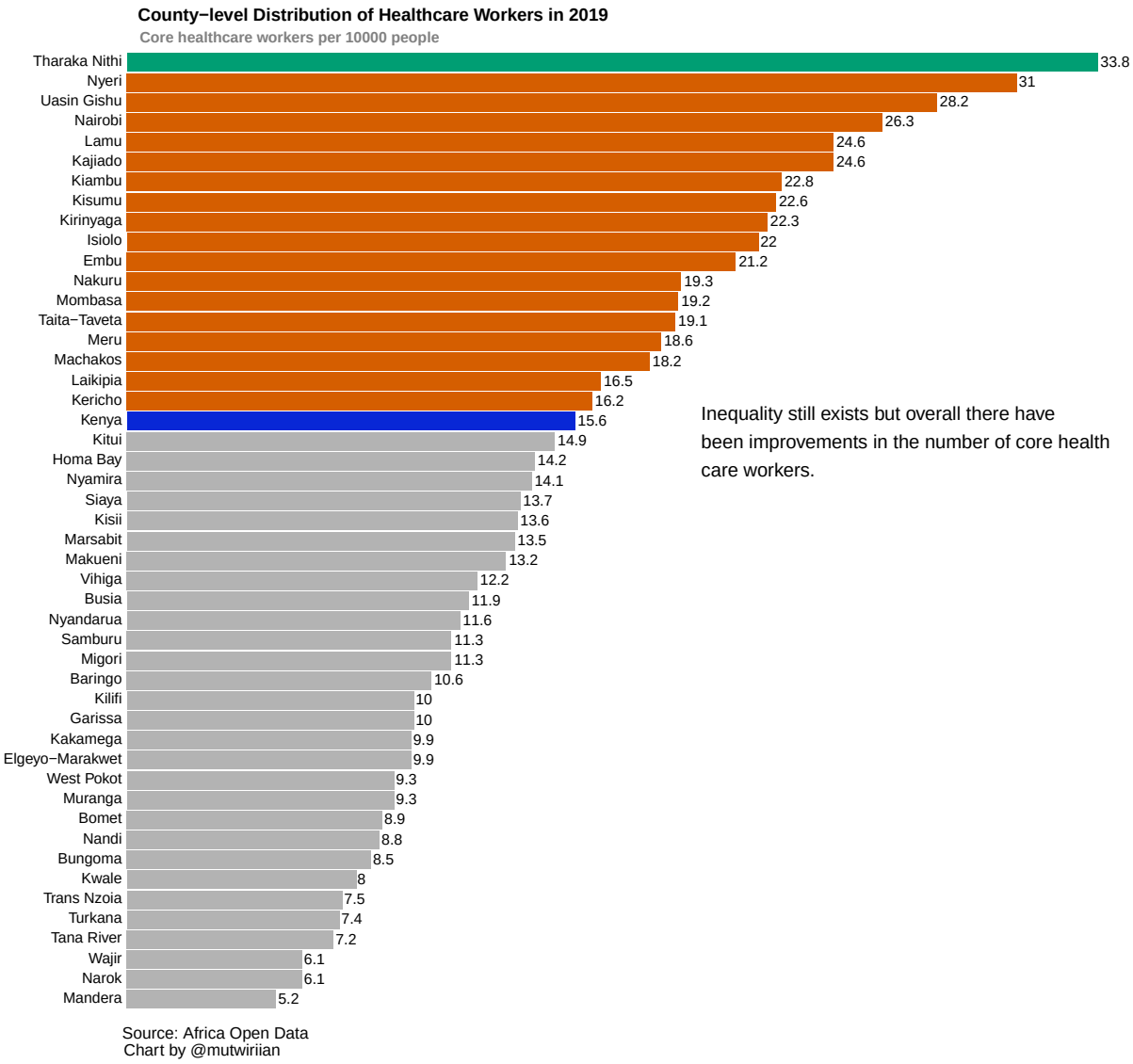
```
plot_13 <- health_13 %>%
  ggplot(aes(core_per_10k_13,fct_reorder(County,core_per_10k_13)))+
  geom_col(fill=case_when(health_13$core_per_10k_13==17.5~'#0827D6',
                          health_13$core_per_10k_13<17.5~'gray70',
                          health_13$core_per_10k_13==53.5~'#009E73',
                              TRUE~'#D55E00'))+
  geom_text(aes(label=core_per_10k_13),hjust=-.1,vjust=.4)+
  annotate('text',x=25,y='Nairobi',label=note_13,hjust=0,size=5)+
  scale_y_discrete(expand = c(0,0))+
labs(x='',y='',title = 'County-level Distribution of Healthcare Workers in 2013',
     subtitle = 'Core healthcare workers per 10000 people',
     caption ='Source: Africa Open Data\nChart by @mutwiriian')+
  theme(
    axis.ticks = element_blank(),
    panel.grid = element_blank(),
    panel.background = element_rect(fill='white'),
    axis.text.y = element_text(margin = margin(r=-35),vjust=.3),
    axis.text = element_blank(),
    plot.title = element_text(face = 'bold',hjust = .1),
    plot.subtitle = element_text(face='bold',color = 'gray50',hjust = .08),
    plot.caption = element_text(size=12,hjust = .05,vjust=4)
  )
plot_13
```



Plot the 2019 data.

```
plot_19 <- health_19 %>%
  ggplot(aes(core_per_10k_19,fct_reorder(County,core_per_10k_19)))+
  geom_col(fill=case_when(health_19$core_per_10k_19==15.6~'#0827D6',
                          health_19$core_per_10k_19<15.6~'gray70',
                          health_19$core_per_10k_19==33.8~'#009E73',
                          TRUE~'#D55E00'))+
  geom_text(aes(label=core_per_10k_19),hjust=-.1,vjust=.4)+
  annotate('text',x=20,y='Kitui',label=note_19,hjust=0,size=5)+
  scale_y_discrete(expand = c(0,0))+
  labs(x='',y='',title = 'County-level Distribution of Healthcare Workers in 2019',
       subtitle = 'Core healthcare workers per 10000 people',
       caption ='Source: Africa Open Data\nChart by @mutwiriian')+
  theme(
    axis.ticks = element_blank(),
    panel.grid = element_blank(),
```

```
    panel.background = element_rect(fill='white'),
    axis.text.y = element_text(margin = margin(r=-35),vjust=.3),
    axis.text = element_blank(),
    plot.title = element_text(face = 'bold',hjust = .1),
    plot.subtitle = element_text(face='bold',color = 'gray50',hjust = .08),
    plot.caption = element_text(size=12,hjust = .05,vjust=4)
  )
plot_19
```



Select the counties with the biggest changes.

```
top<- health %>%
  pivot_wider(names_from = Year,values_from = value) %>%
  mutate(change=core_per_10k_19-core_per_10k_13 ) %>%
  arrange(desc(change)) %>%
  select(County,change) %>%
  filter(change>10)
```

```
bot <- health %>%
  pivot_wider(names_from = Year,values_from = value) %>%
  mutate(change=core_per_10k_19-core_per_10k_13 ) %>%
  arrange(desc(change)) %>%
  select(County,change) %>%
  filter(change<=(-13))

#ad-hoc change to fit the data better on the plot grid
health<- health %>%
  mutate(Year=case_when(
    Year=="core_per_10k_13"~as.Date("2014-04-30"),
    TRUE~as.Date("2015-1-30")
  ))
```

Create the plots for the counties with the biggest changes.

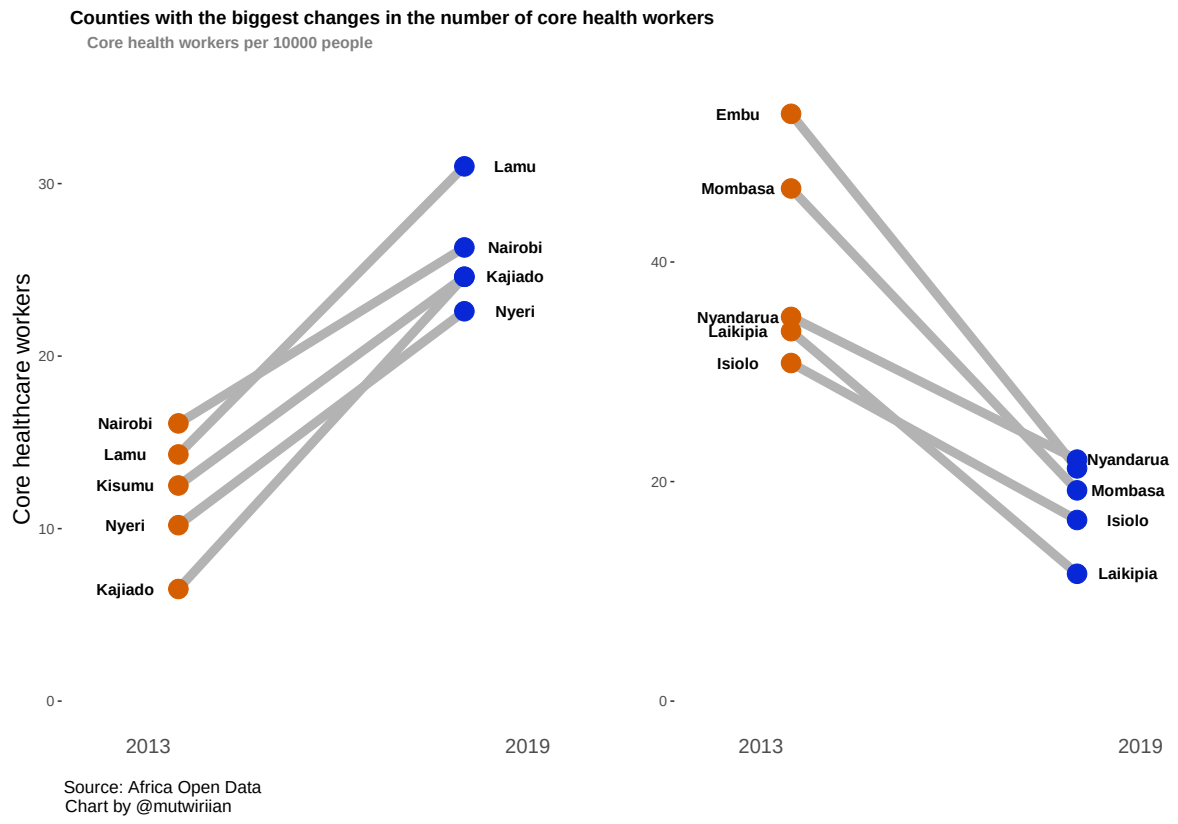
```
top_plot<- health %>%
  filter(County%in% top$County) %>%
  ggplot(aes(Year,value,group=County)) +
  geom_line(size=3,col='gray70')+
  geom_point(data = . %>% filter(lubridate::year(Year)==2014),size=6,col='#D55E00')+
  geom_point(data = . %>% filter(lubridate::year(Year)==2015),size=6,col='#0827D6')+
  geom_text(data = . %>% filter(lubridate::year(Year)<2015),
            aes(x=as.Date('2014-3-10'),y=value,label=top$County),
            fontface='bold') +
  geom_text(data = . %>% filter(lubridate::year(Year)==2015),
            aes(x=as.Date("2015-3-20"),y=value,label=top$County),
            fontface='bold',check_overlap = T)+
  scale_x_date(labels = c("",2013,rep("",3),2019,""),
    limits = c(lubridate::date('2014-2-1'),
               lubridate::date("2015-6-1"))) +
  ylim(c(0,35))+
  labs(x="",y="Core healthcare workers")+
  theme(
    axis.ticks.x = element_blank(),
    axis.text.x = element_text(size = 14),
    axis.text.y = element_text(size = 10),
    axis.title.y = element_text(size=16),
    panel.background = element_blank()
  )

bottom_plot <- health %>%
  filter(County%in% bot$County) %>%
  ggplot(aes(Year,value,group=County)) +
  geom_line(size=3,col='gray70')+
  geom_point(data = . %>% filter(lubridate::year(Year)==2014),size=6,col='#D55E00')+
  geom_point(data = . %>% filter(lubridate::year(Year)==2015),size=6,col='#0827D6')+
  geom_text(data = . %>% filter(lubridate::year(Year)<2015),
            aes(x=as.Date("2014-3-10"),y=value,label=bot$County),
            fontface='bold')+
  geom_text(data = . %>% filter(lubridate::year(Year)==2015),
            aes(x=as.Date("2015-3-20"),y=value,label=bot$County),
            fontface='bold',check_overlap = T) +
```

```
  scale_x_date(labels = c("",2013,rep("",3),2019,""),
               limits = c(lubridate::date('2014-2-1'),
                          lubridate::date("2015-6-1")))+
  ylim(c(0,55))+
  labs(x="",y="")+
  theme(
    axis.ticks.x = element_blank(),
    axis.text.x = element_text(size = 14),
    axis.text.y = element_text(size = 10),
    panel.background = element_blank()
  )
```

Align the plots side by side.

```
top_plot+bottom_plot+
  plot_annotation(
    title = "Counties with the biggest changes in the number of core health workers",
    subtitle = 'Core health workers per 10000 people',
    caption = 'Source: Africa Open Data\nChart by @mutwiriian',
    theme =   theme(
      panel.grid = element_blank(),
      panel.background = element_rect(fill='white'),
      plot.title = element_text(face = 'bold',hjust = .1),
      plot.subtitle = element_text(face='bold',color = 'gray50',
                                   hjust = .08),
      plot.caption = element_text(size=12,hjust = .05,vjust=4)
    ))
```



I believe the data tells its own story!

Thank you!