

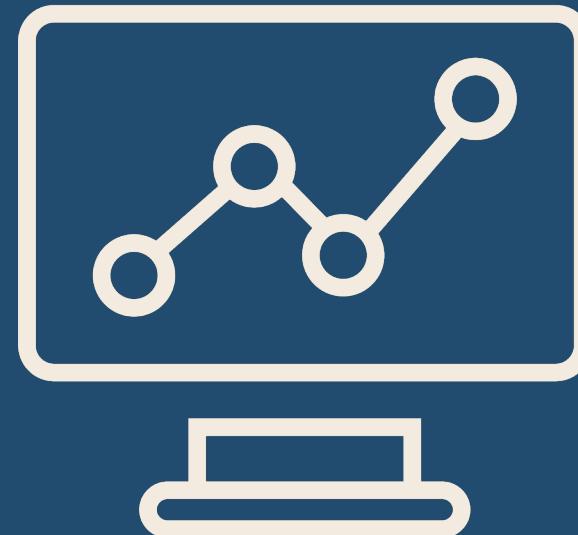
EMMANUEL OLAMIJUWON

Lecturer in Health and Data Literacy

X: eolamijuwon;

email: emmanuel@olamijuwon.com

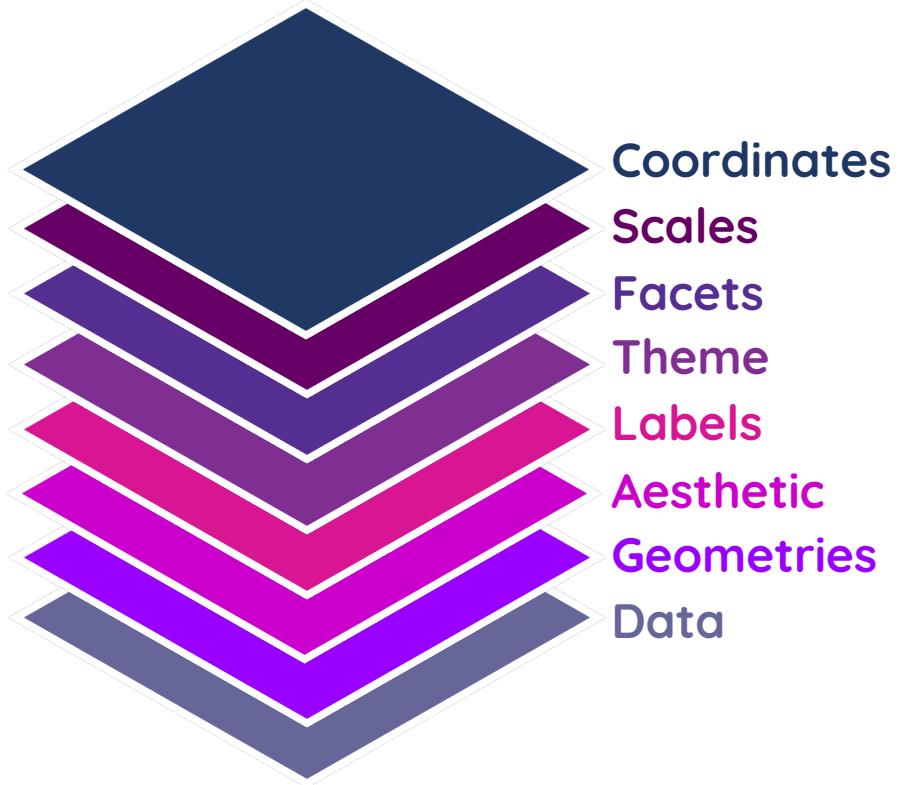
Customizing Graphics – ggplot2 tuneup



University of
St Andrews

School of Geography and
Sustainable Development

Plotting in Layers



WorldBank: <https://www.youtube.com/watch?v=LYIvh0oATeU>

Customising your graphics

Fonts: Size matters



Image source: <https://www.appletoncreative.com/>

```
library (showtext)  
library (extrafont)
```

```
font_add_google("Barlow  
Condensed", regular.wt = 200,  
family = "barlow")
```

```
font_add_google("Special  
Elite", family = "special")
```

```
font_add_google("Alegreya SC",  
family = "alegreya_sc")
```

```
showtext_auto()
```

Customising your graphics

Colors: Use them to your advantage (1/4)

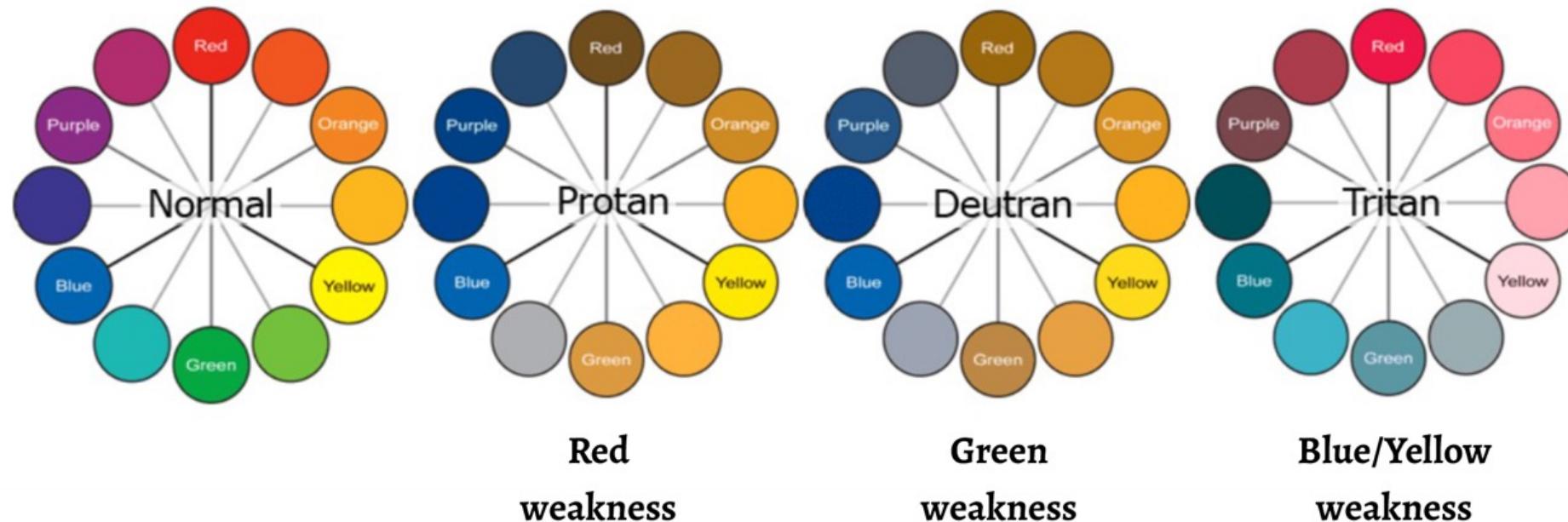


Image source: <https://theiagd.org/color-vision-deficiency-cvd/>

Customising your graphics

Colors: Use them to your advantage (2/4)

PALETTE SINGLE HUE DIVERGENT

PALETTE GENERATOR

NUMBER OF COLORS
 3 4 5 6 7 8

BACKGROUND COLOR
 LIGHT DARK



#003f5c

#444e86

#955196

#dd5182

#ff6e54

#ffa600

Image source: <https://www.learnui.design/tools/data-color-picker.html/>

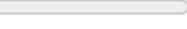
Customising your graphics

Colors: Use them to your advantage (3/4)

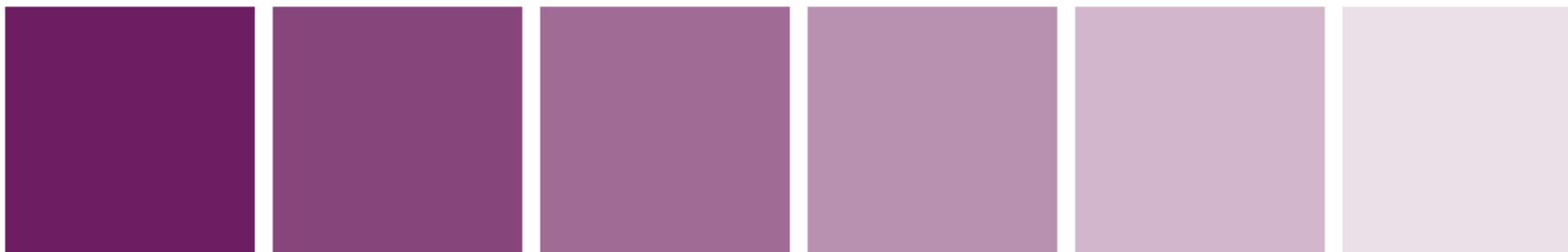
PALETTE SINGLE HUE DIVERGENT

SINGLE HUE SCALE

NUMBER OF COLORS
 3 4 5 6 7 8 9

MODIFY COLOR SCALE
BRIGHTNESS 
COLOR INTENSITY 

BACKGROUND COLOR
 LIGHT  DARK



#6d1d61

#87467b

#a06b95

#b991b0

#d2b7cc

#ebdfe8

Image source: <https://www.learnui.design/tools/data-color-picker.html/>

Customising your graphics

Colors: Use them to your advantage (4/4)

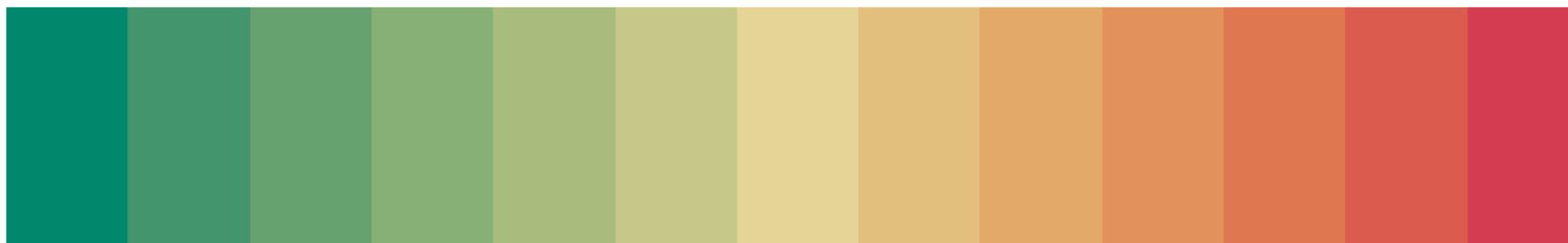
PALETTE SINGLE HUE DIVERGENT

DIVERGENT COLOR SCALE

NUMBER OF COLORS
5 7 9 11 13

MODIFY MIDPOINT COLOR
BRIGHTNESS
COLOR INTENSITY

BACKGROUND COLOR
LIGHT DARK



#488f31 #42956d #66a270 #87b075 #a7bc7d #c6c888 #e4d496 #e2bf7d #e2a968 #e19159 #df7751 #db5c4e #de425b

Mapping: Data → Aesthetics → Graphics (in R)



```
install.packages("ggplot")
library(ggplot)
```

```
Data %>% ggplot () +
geom_*(mapping = aes()) +
labs () +
theme_*( )
```

theme_bw()
theme_minimal()
theme_light()
theme_dark()
theme_gray()
theme_classic()

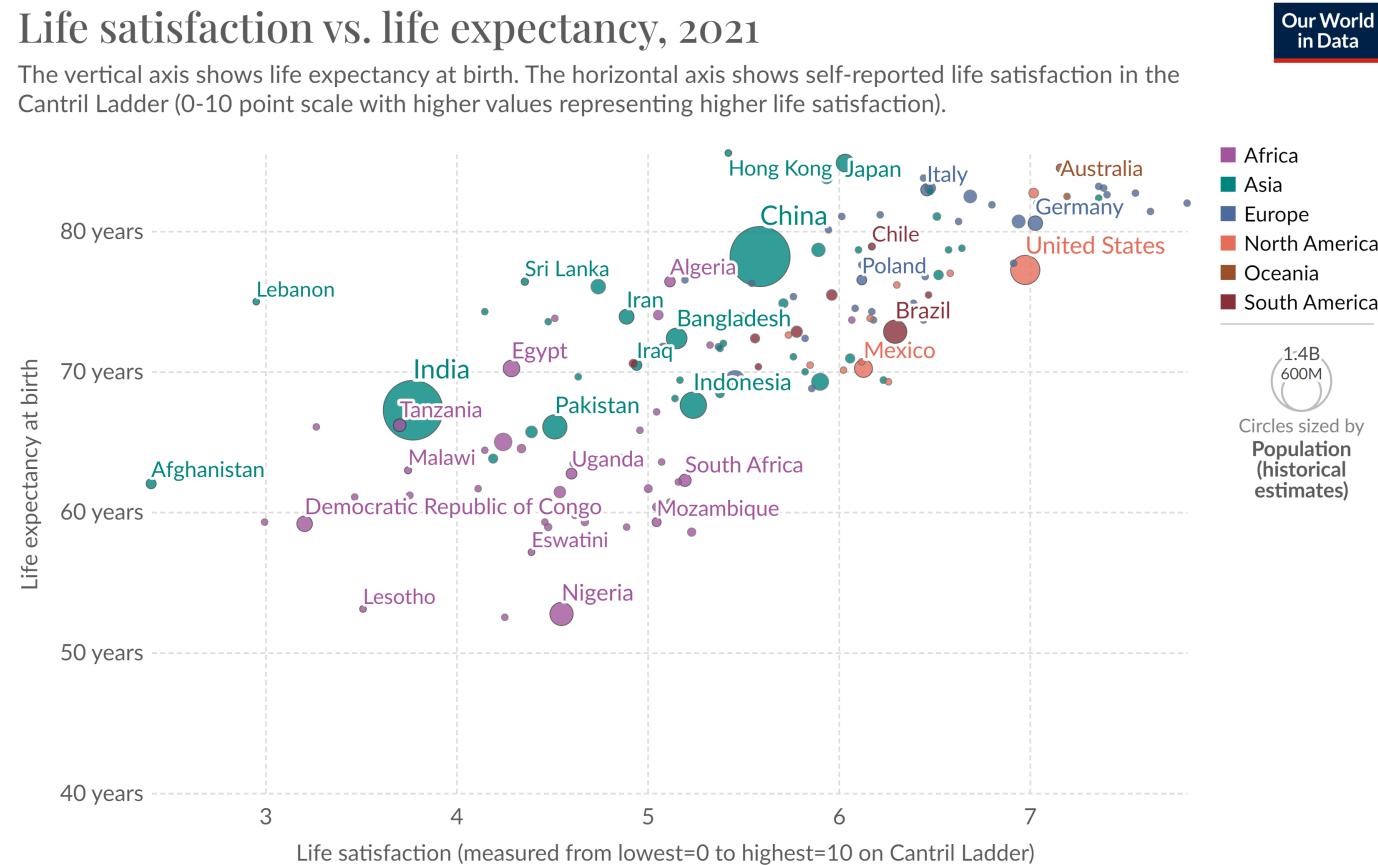
READ MORE: <https://ourworldindata.org/grapher/unadjusted-gender-wage-gap-including-unemployed>

Customising your graphics

theme(): Style It Right, Shine Bright (1/2)

Life satisfaction vs. life expectancy, 2021

The vertical axis shows life expectancy at birth. The horizontal axis shows self-reported life satisfaction in the Cantril Ladder (0-10 point scale with higher values representing higher life satisfaction).



Data source: UN, World Population Prospects (2022); World Happiness Report (2023)

OurWorldInData.org/happiness-and-life-satisfaction | CC BY

READ MORE: <https://ourworldindata.org/grapher/life-satisfaction-vs-life-expectancy?time=2021>

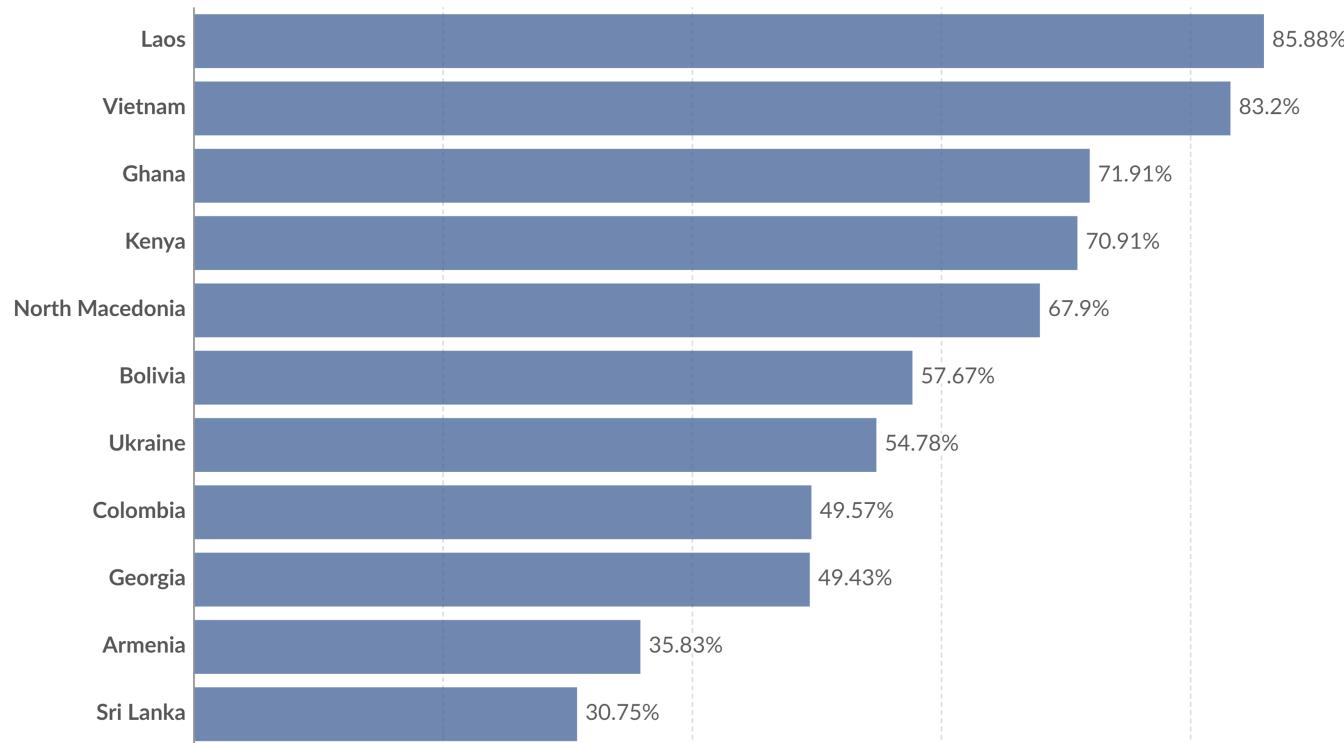
- Plot title text: `plot.title`
- Plot sub-title text: `plot.subtitle`
- Plot caption: `plot.caption`
- Axis-title: `axis.title`
- Axis-text: `axis.text`
- Axis-ticks: `axis.ticks`
- Panel grid lines: `panel.grid`
- Panel background: `panel.background`
- Legend text: `legend.text`
- Legend key size: `legend.key.size`

Customising your graphics

theme(): Style It Right, Shine Bright (2/2)

Unadjusted gender wage gap including unemployed, 2014

Shown are unadjusted pay gaps assigning zero income to individuals who are not employed. All figures correspond to women's income as a percent of men's, including wages and profits from self-employment.



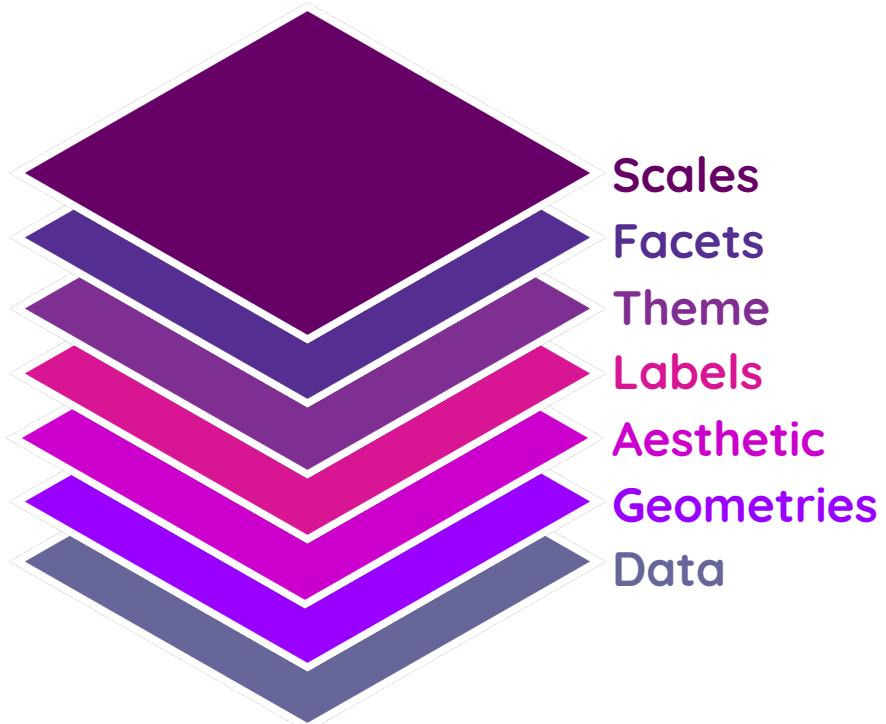
Data source: Center for Global Development (2018)

OurWorldInData.org/economic-inequality-by-gender | CC BY

READ MORE: <https://ourworldindata.org/grapher/unadjusted-gender-wage-gap-including-unemployed>

- *Plot title text: plot.title*
- *Plot sub-title text: plot.subtitle*
- *Plot caption: plot.caption*
- *Axis-title: axis.title*
- *Axis-text: axis.text*
- *Axis-ticks: axis.ticks*
- *Panel grid lines: panel.grid*
- *Panel background: panel.background*
- *Legend text: legend.text*
- *Legend key size: legend.key.size*

Plotting in Layers - Scales (2/2)



```
install.packages("ggplot")
library(ggplot)

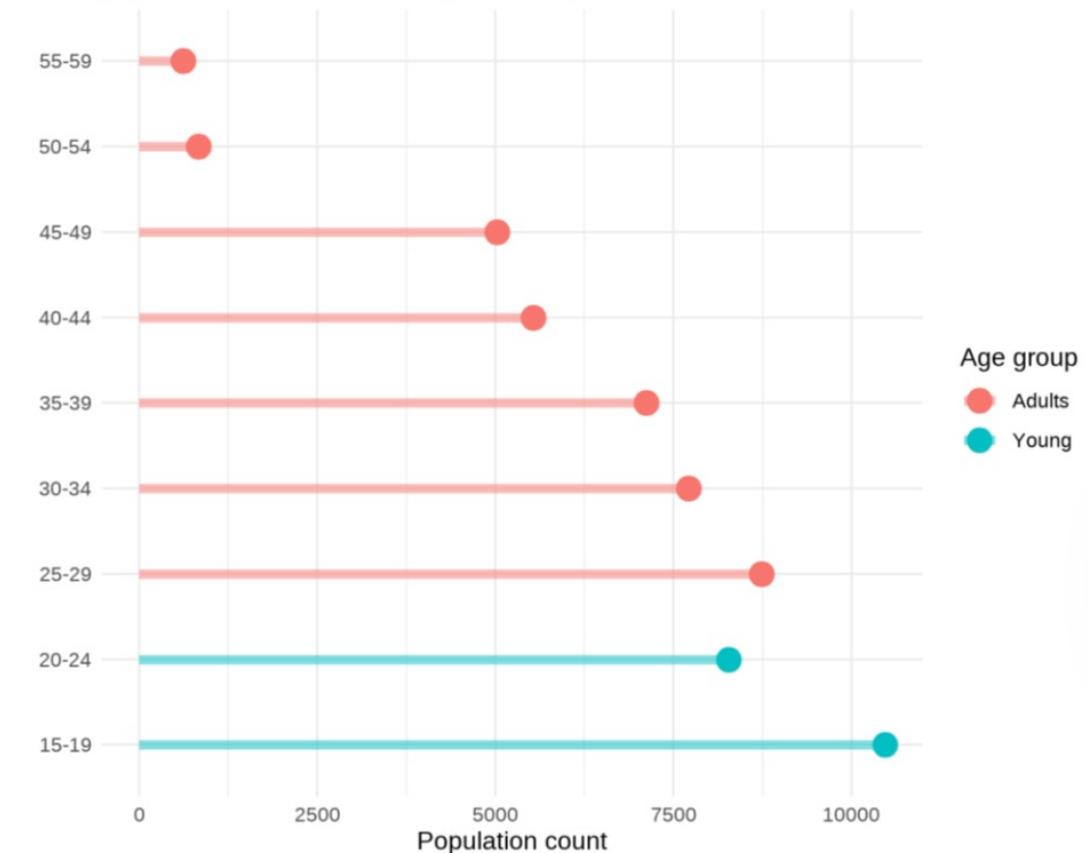
Data %>% ggplot () +
  geom_*(mapping = aes()) +
  labs () +
  theme_*( ) +
  scale_*( ) {  
  scale_x_continuous()  
  scale_fill_manual()  
  scale_color_manual()  
  scale_x_discrete() }
```

READ MORE: <https://ggplot2.tidyverse.org/reference/#scales>

Visualising distributions with lollipop plots (1/5): geom_segment()

```
new_graph <- study_dta %>%  
  ggplot () +  
  geom_point(aes(x = n, y = age_grp,  
                 color = factor(ifelse((age_grp == "15-19" |  
                  age_grp == "20-24"), "Young", "Adults"))),  
             size = 5) +  
  geom_segment(aes(y=age_grp,  
                   yend = age_grp, x=0, xend=n, color =  
                   factor(ifelse((age_grp == "15-19" | age_grp  
                   == "20-24"), "Young", "Adults"))),  
               size = 2, alpha = 0.5) +  
  labs (y = "", x = "Population count", color = "Age  
group", title = "More than a third of this  
population are young adults", subtitle =  
"Young people aged 15-24 years account for  
about a third of this population while  
older adults aged 10-59 years account for  
less than 3%.") + theme_minimal()
```

More than a third of this population are young adults
Young people aged 15-24 years account for about a third of this population while older adults aged 10-59 years account for less than 3%.



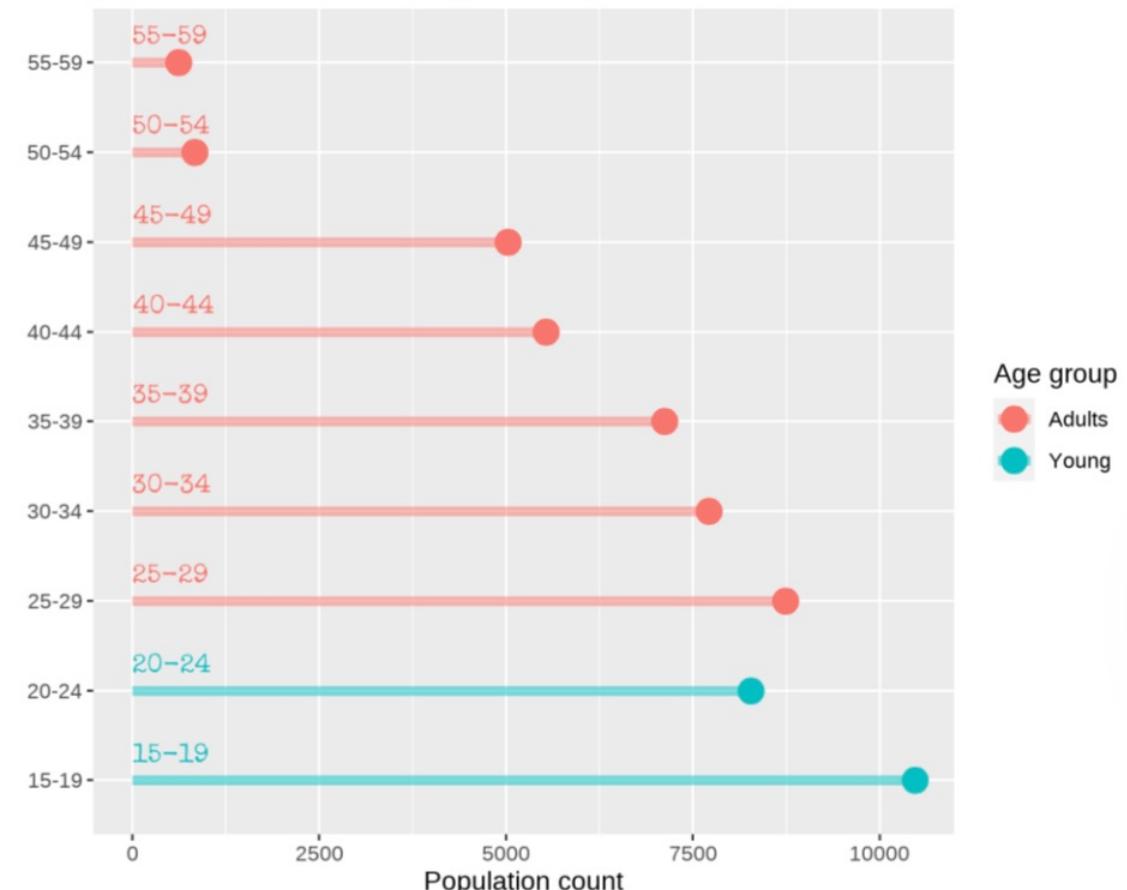
READ MORE: <https://ggplot2.tidyverse.org/>

Visualising distributions with lollipop plots (2/5): geom_segment()

```
new_graph2 <- new_graph +  
  geom_text(aes (y = age_grp, x = 0,  
    label = age_grp, color =  
      factor(ifelse((age_grp == "15-19" |  
        age_grp == "20-24"), "Young",  
        "Adults"))), vjust = -1, hjust = 0,  
    family = "special") + labs (y = "",  
    x = "Population count", color = "Age  
    group, title = "More than a third of  
    this population are young adults",  
    subtitle = "Young people aged 15-24  
    years account for about a third of  
    this population while older adults  
    aged 50-59 years account for less  
    than 3%.")
```

```
new_graph2
```

More than a third of this population are young adults
Young people aged 15-24 years account for about a third of this population while older adults aged 50-59 years account for less than 3%.

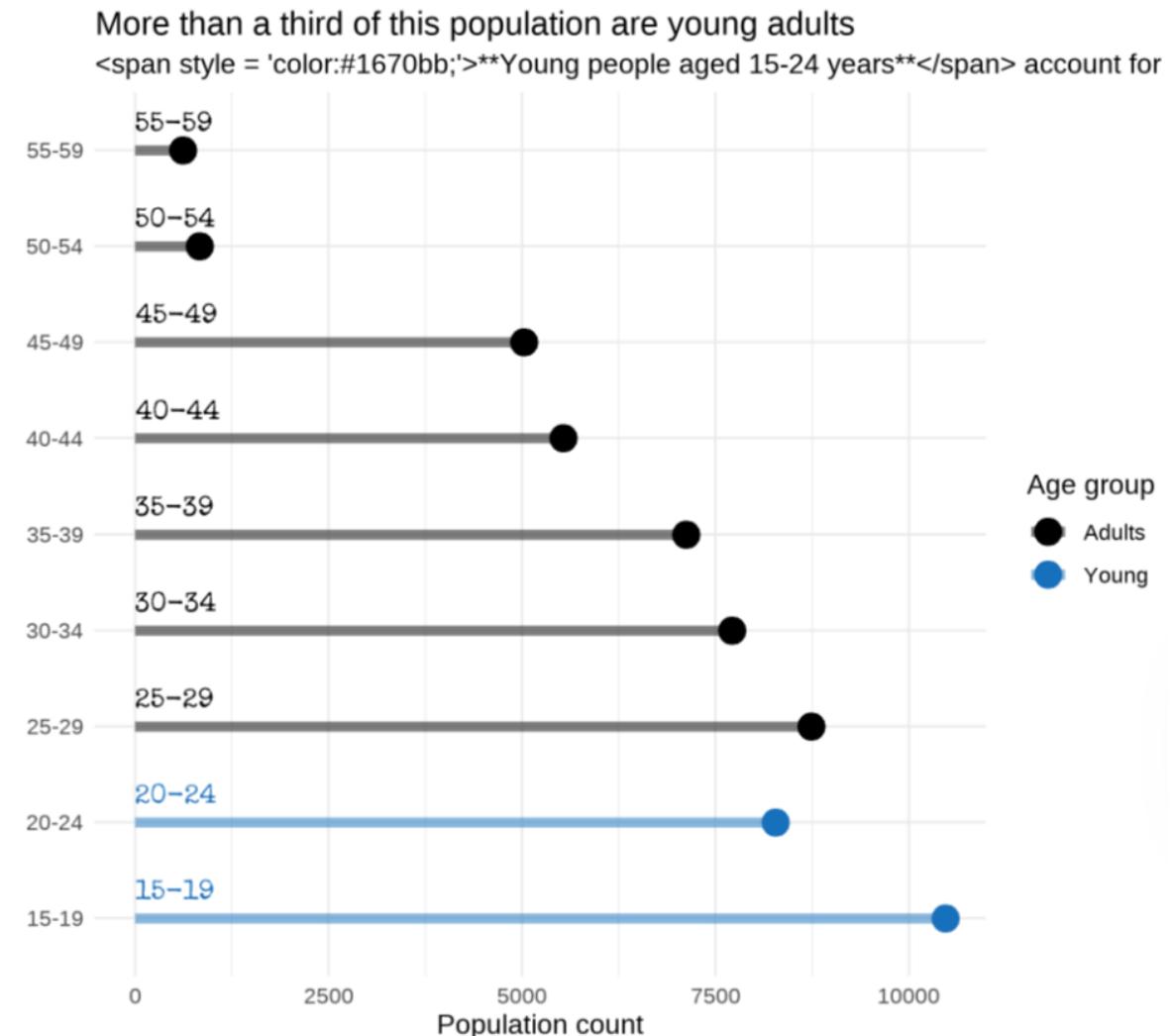


READ MORE: <https://ggplot2.tidyverse.org/>

Visualising distributions with lollipop plots (3/5): geom_segment()

```
new_graph3 <-  
  new_graph2 +  
  scale_color_manual(values =  
    c("#000000", "#1670bb")) +  
  theme_minimal()
```

```
new_graph3
```



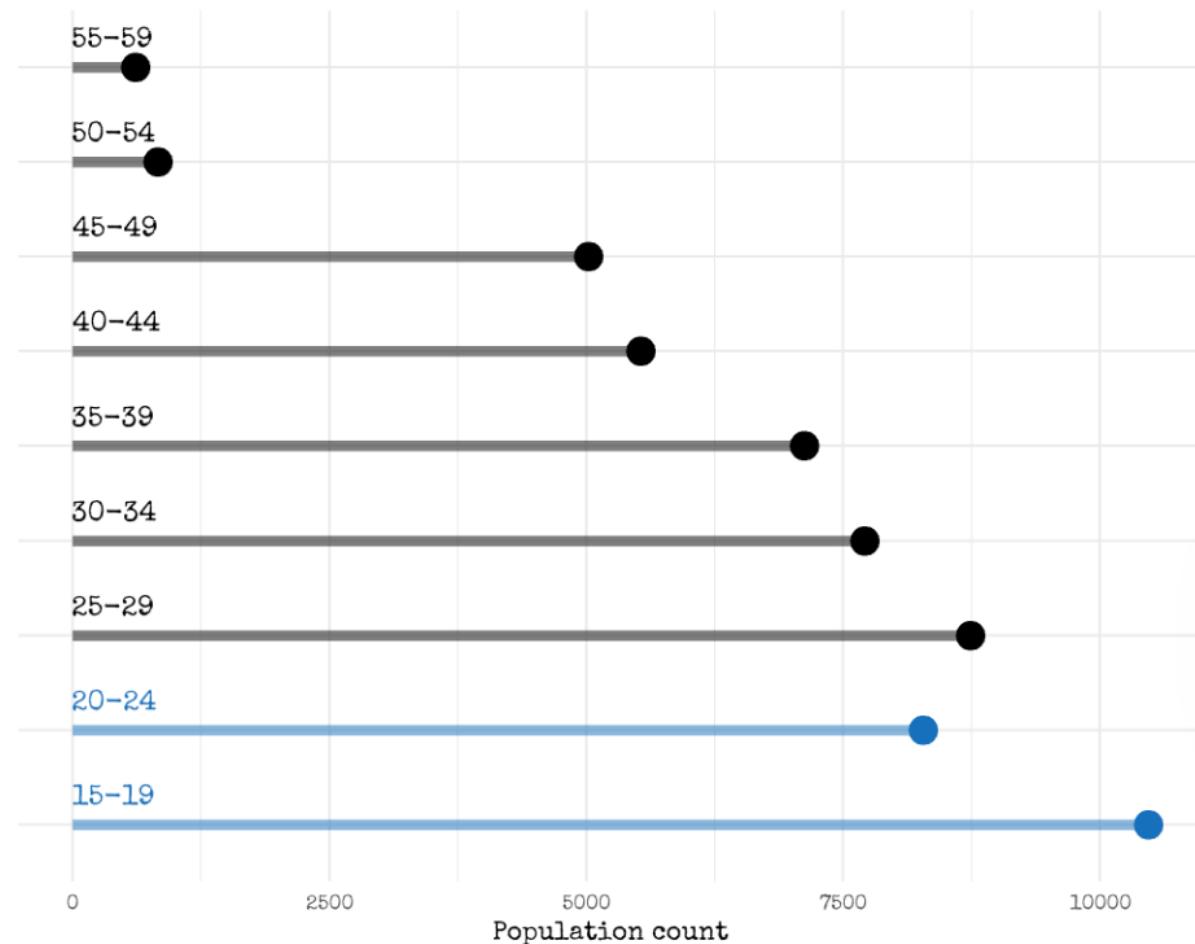
Visualising distributions with lollipop plots (4/5): geom_segment()

```
new_graph4 <-  
  new_graph3 +  
  theme (axis.text.y = element_blank(),  
         axis.ticks.y = element_blank(),  
         axis.ticks.y.left = element_blank(),  
         plot.subtitle = element_markdown  
           (hjust = 0, family = "barlow",  
            lineheight = unit(1.2, "pt"),  
            size = 14),  
         plot.title = element_text(family =  
           "alegrey_sc", size = 15),  
         axis.text = element_text(family =  
           "special", size = 8),  
         axis.title = element_text(family =  
           "special", size = 10),  
         legend.position = "none")
```

```
new_graph4
```

MORE THAN A THIRD OF THIS POPULATION ARE YOUNG ADULTS

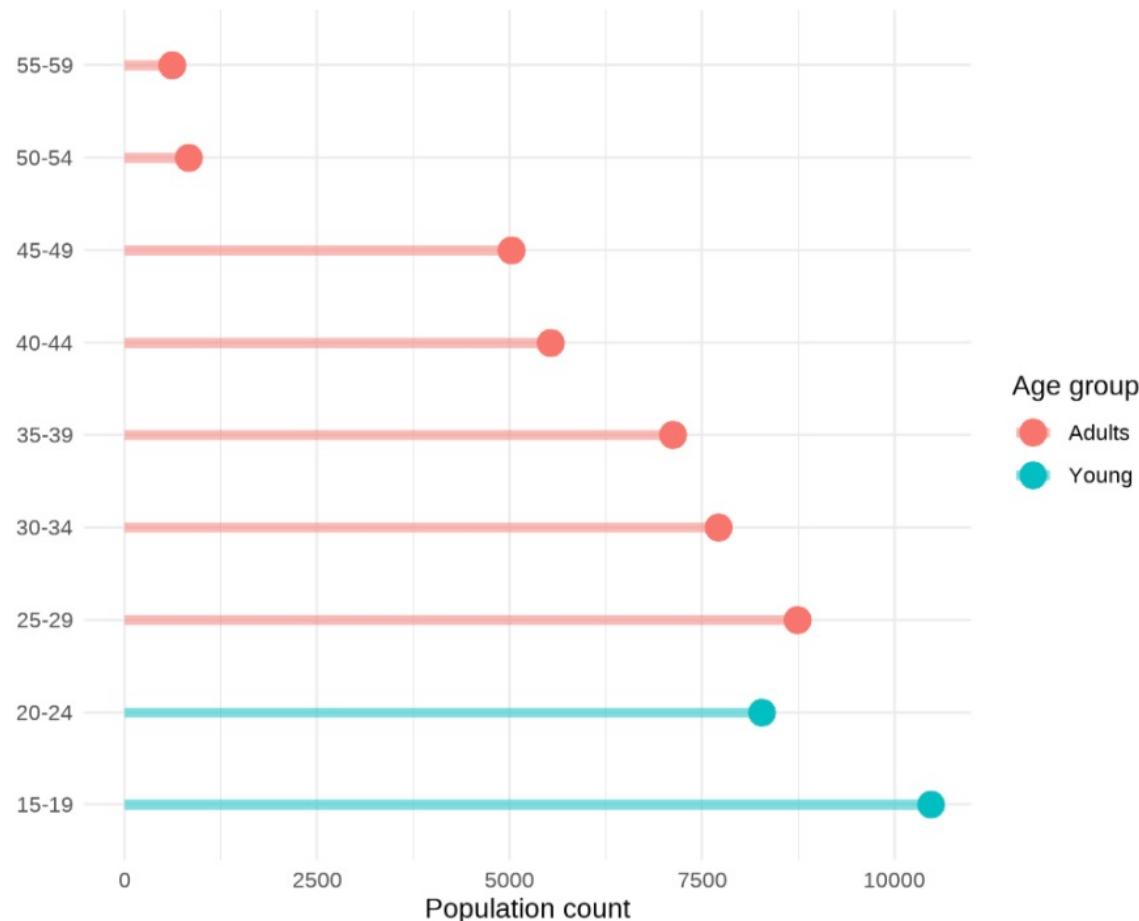
Young people aged 15-24 years account for about a third of this population while older adults aged 50-59 years account for less than 3%.



Visualising distributions with lollipop plots (4/5): geom_segment()

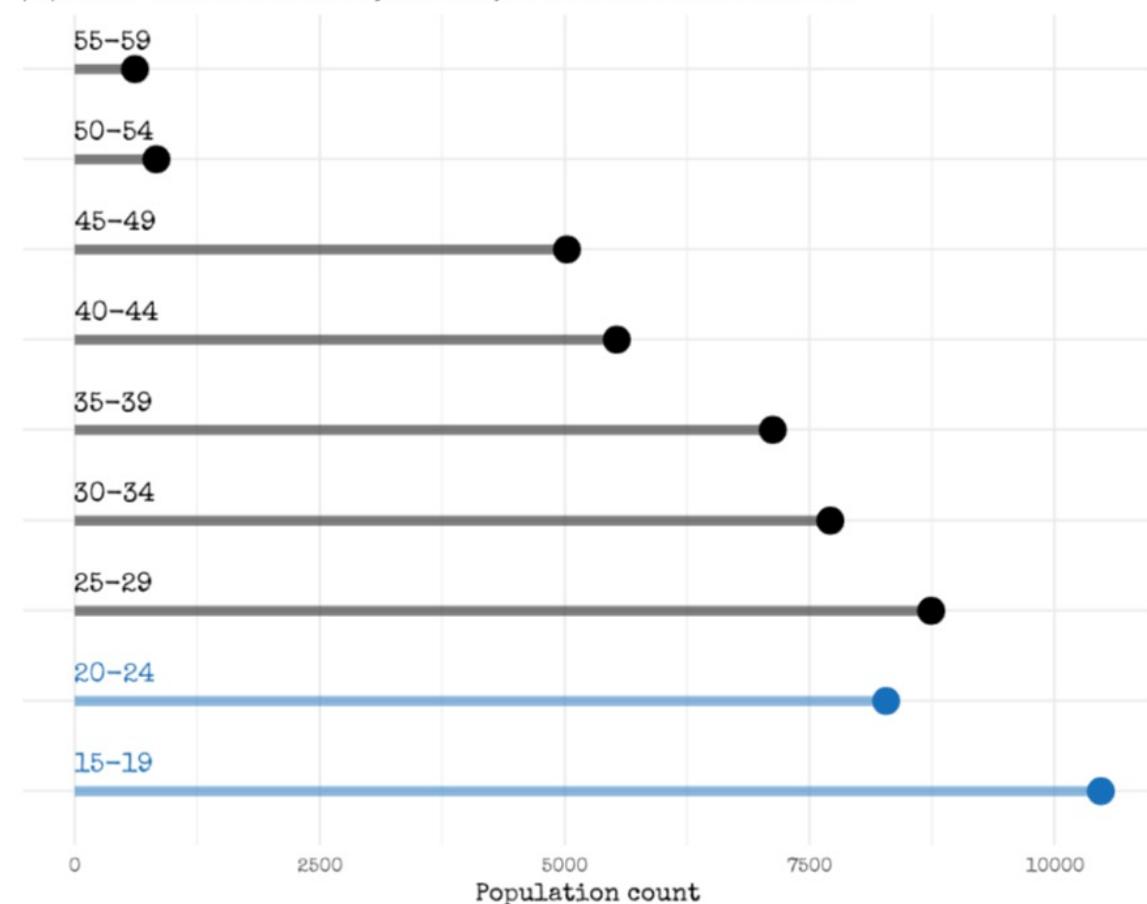
More than a third of this population are young adults

Young people aged 15-24 years account for about a third of this population while older adults aged 10-59 years account for less than 3%.



MORE THAN A THIRD OF THIS POPULATION ARE YOUNG ADULTS

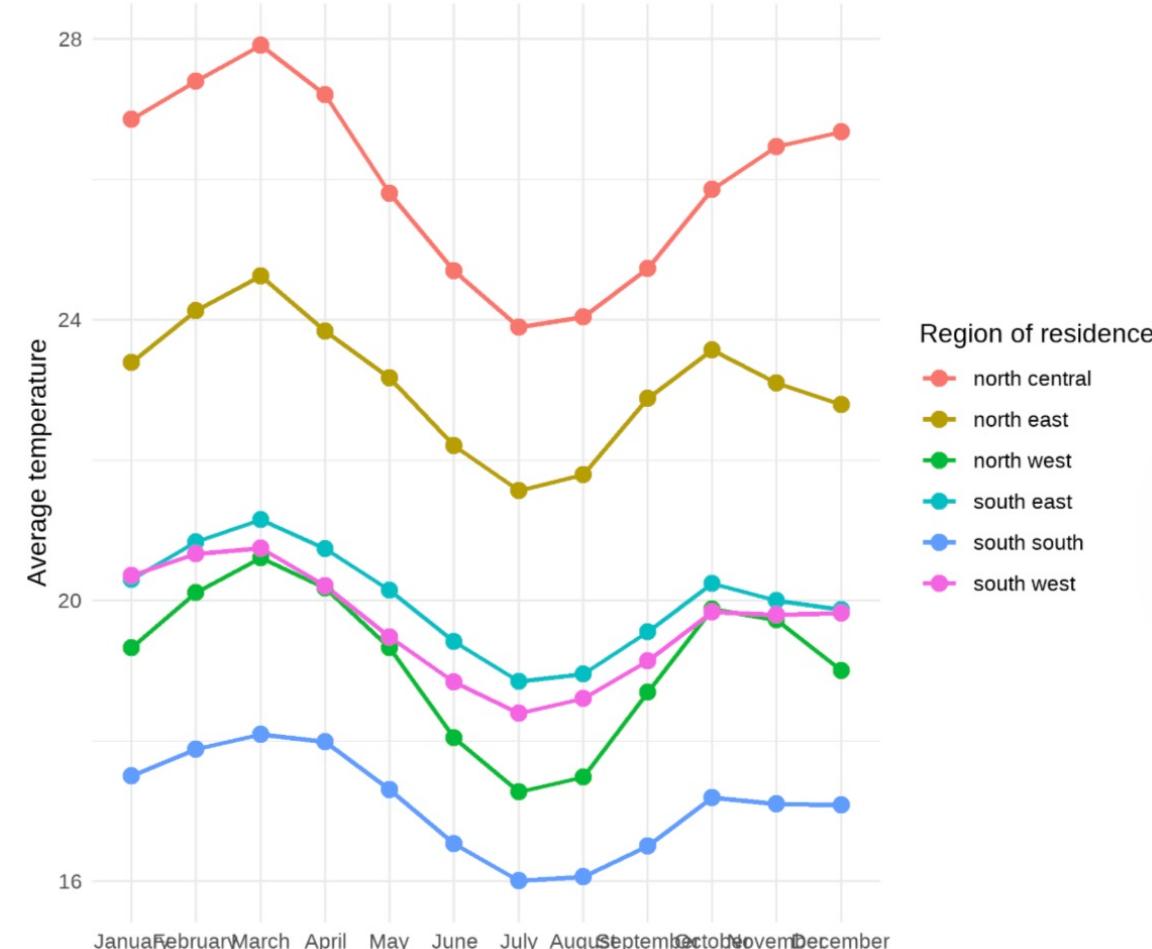
Young people aged 15-24 years account for about a third of this population while older adults aged 50-59 years account for less than 3%.



Visualising trends with lineplots (1/9): geom_line() & geom_point ()

```
study_dta %>%  
  ggplot (aes(x = month, y = temp, color =  
    region, group = region)) +  
  geom_line (size = 0.9) +  
  geom_point(size = 3) +  
  labs (y = "Average temperature",  
    x = "", color= "Region of residence",  
    title = "March is the hottest month of  
    the year in Kenya",  
    subtitle = "Average temperature varies  
    significantly across regions in Kenya.  
    \nLess hot in the South-South and hottest  
    in the North-Central regions") +  
  theme_minimal()
```

March is the hottest month of the year in Kenya
Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions

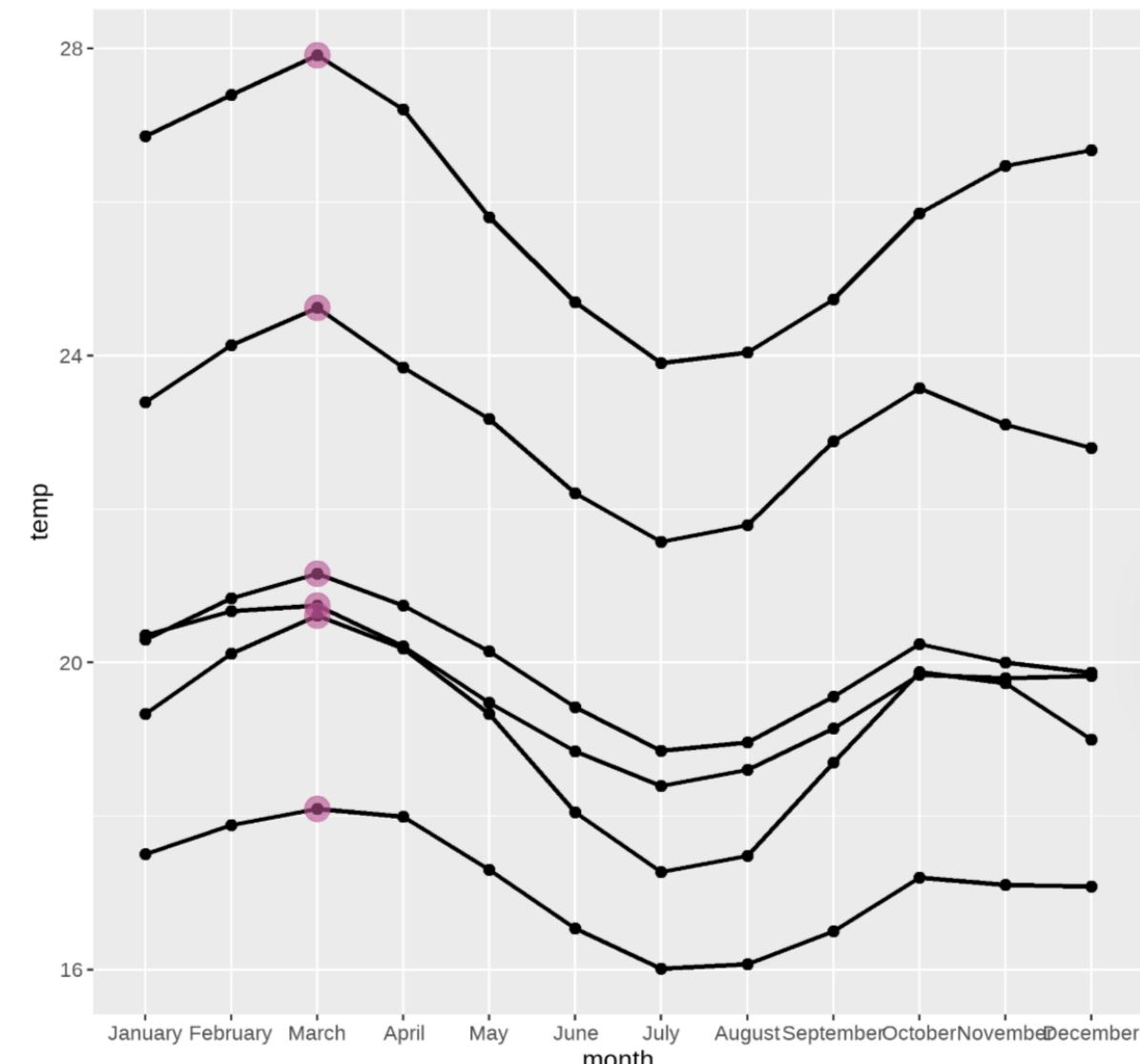


READ MORE: <https://ggplot2.tidyverse.org/>

Visualising trends with lineplots (2/9): geom_line() & geom_point ()

```
temp_graph_1 <- study_dta %>%  
  ggplot (aes(x = month, y = temp,  
             group = region)) +  
  geom_point(size = 2) +  
  geom_point(data = filter(new_data,  
                           month == "March"),  
             aes(x = month, y = temp,  
                 group = region),  
             color = "#b84e90",  
             size = 5, alpha = 0.6)
```

```
temp_graph_1
```

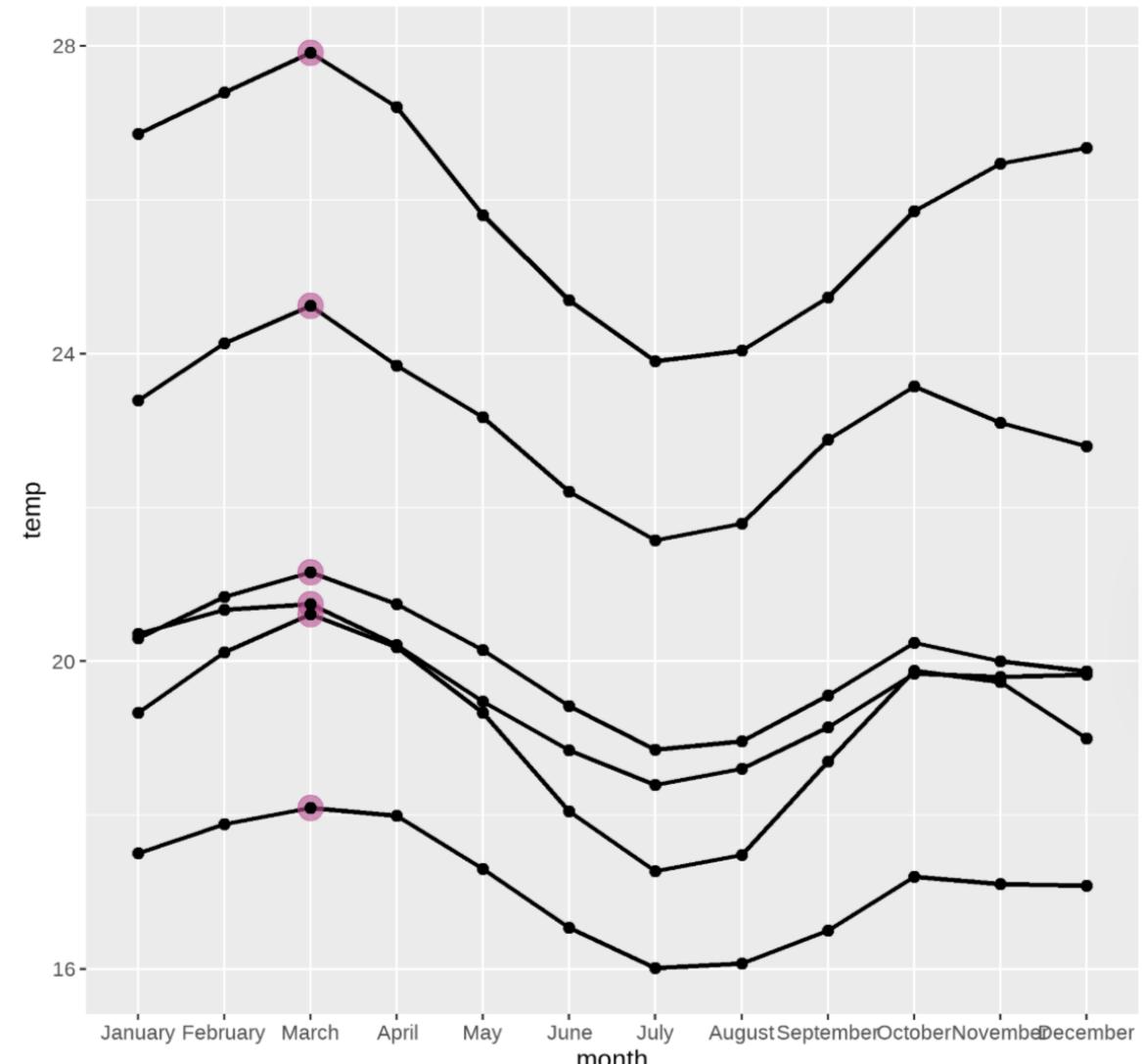


READ MORE: <https://ggplot2.tidyverse.org/>

Visualising trends with lineplots (3/9): geom_line() & geom_point ()

```
temp_graph_1 <- study_dta %>%  
  ggplot (aes(x = month, y = temp,  
              group = region)) +  
  geom_point(data = filter(new_data,  
                           month == "March"),  
             aes(x = month, y = temp,  
                  group = region),  
             color = "#b84e90",  
             size = 5, alpha = 0.6) +  
  geom_point(size = 2)
```

```
temp_graph_1
```

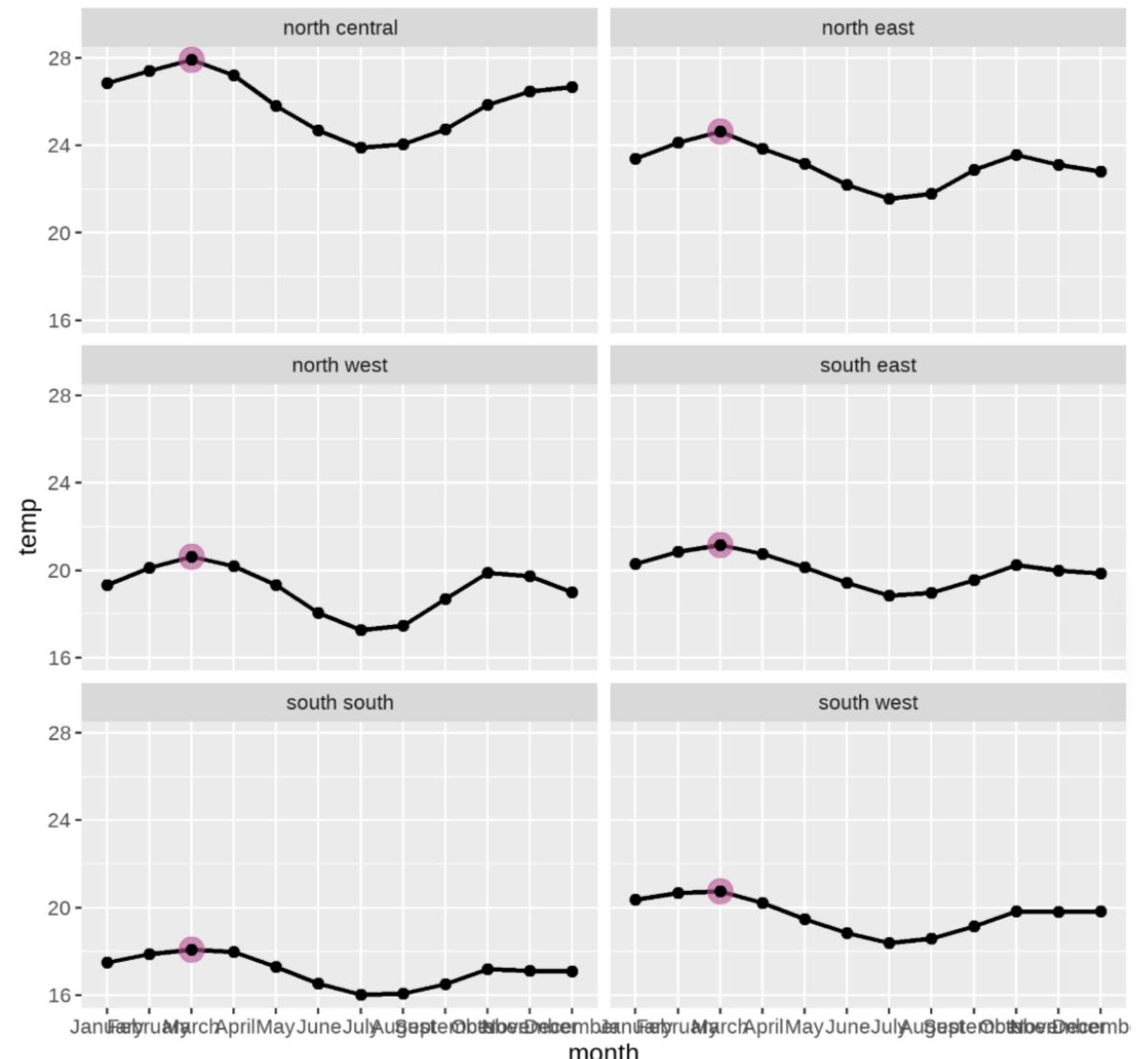


READ MORE: <https://ggplot2.tidyverse.org/>

Visualising trends with lineplots (4/9): geom_line() & geom_point()

```
temp_graph_2 <- temp_graph_1 +  
  facet_wrap(~ region, nrow = 3)
```

```
temp_graph_2
```



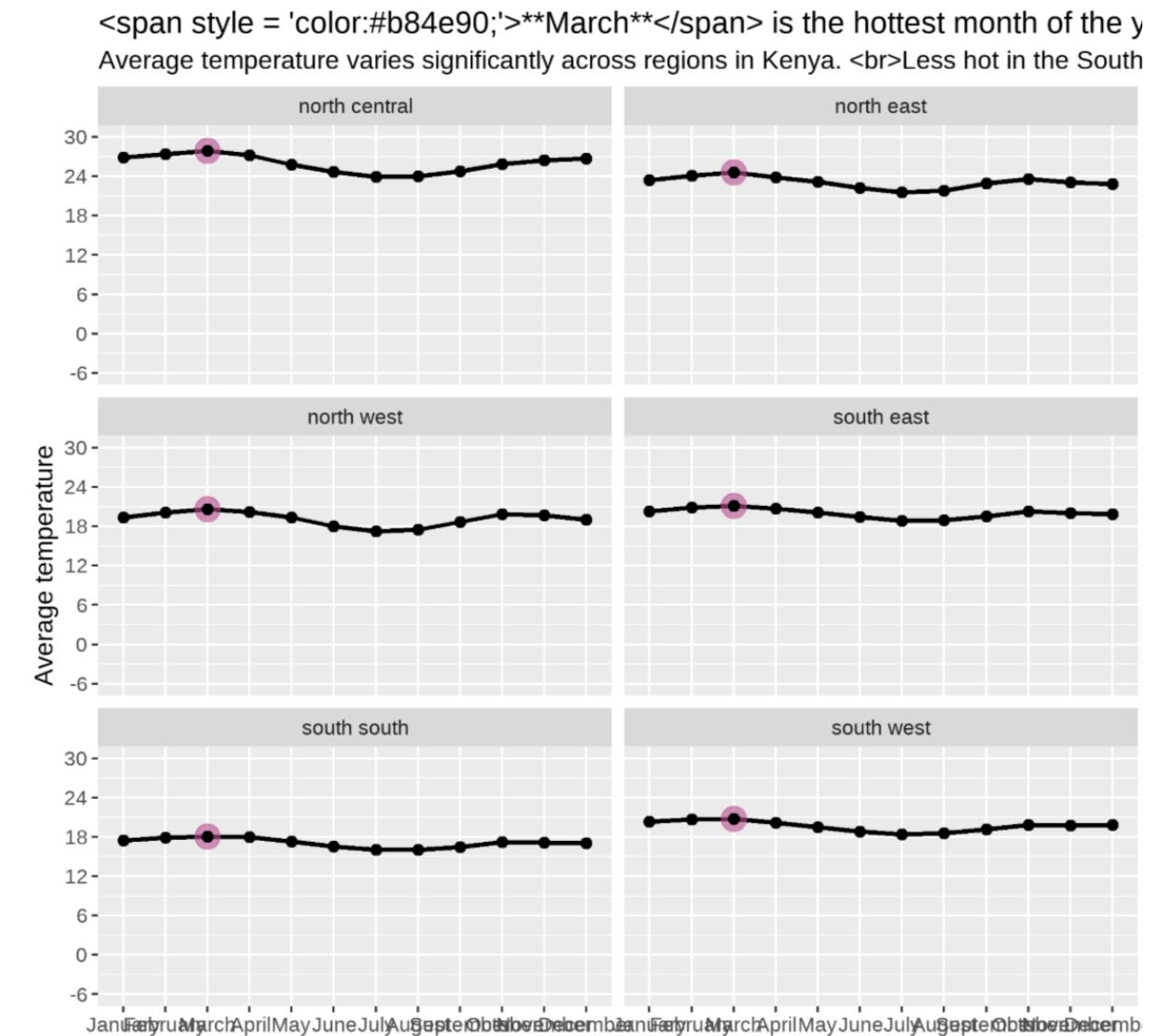
READ MORE: <https://ggplot2.tidyverse.org/>

Visualising trends with lineplots (5/9): geom_line() & geom_point ()

```
temp_graph_3 <- temp_graph_2 +  
  scale_y_continuous(breaks = c(seq(-6,  
    32, 6)), limits = c(-6, 30)) +  
  labs (y = "Average temperature",  
        x = "", color= "Region of  
        residence",  
        title = "<span style =  
          'color:#b84e90; '>**March**</span>  
          is the hottest month of the year  
          in Kenya",  
        subtitle = "Average temperature  
          varies significantly across  
          regions in Kenya. <br>Less hot in  
          the South-South and hottest in the  
          North-Central regions")
```

```
temp_graph_3
```

READ MORE: <https://ggplot2.tidyverse.org/>



Visualising trends with lineplots (6/9): geom_line() & geom_point()

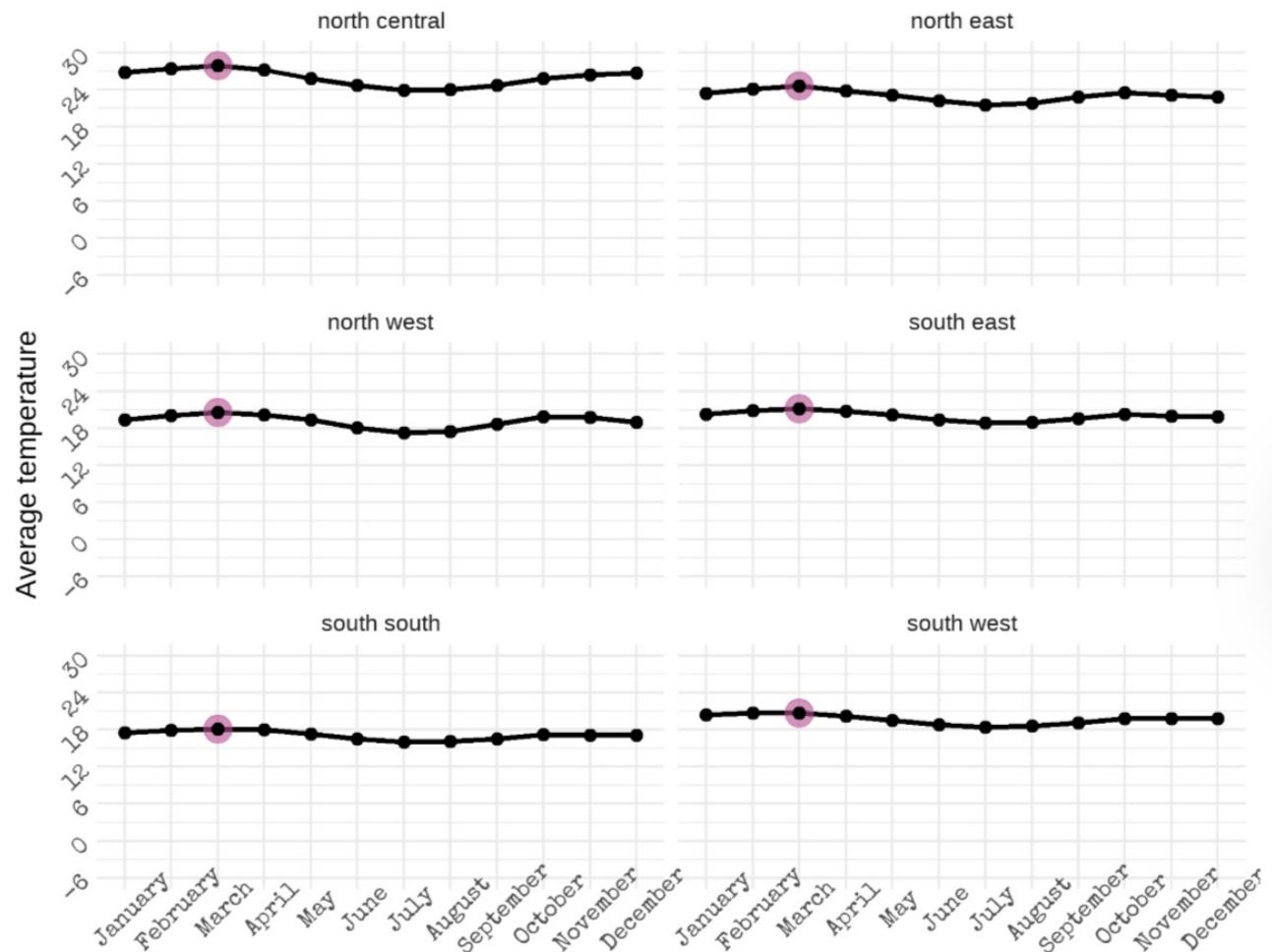
```
temp_graph_4 <- temp_graph_3 +  
  theme_minimal() +  
  theme(plot.subtitle =  
    element_markdown(hjust = 0,  
                      family = "barlow",  
                      lineheight = unit(1.2, "pt"),  
                      size = 13),  
    plot.title =  
    element_markdown(family =  
      "alegreya_sc", size = 18),  
    axis.text =  
    element_text(family =  
      "special", angle = 45,  
      size = 9))
```

```
temp_graph_4
```

READ MORE: <https://ggplot2.tidyverse.org/>

MARCH IS THE HOTTEST MONTH OF THE YEAR IN KENYA

Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions



Visualising trends with lineplots (7/9): geom_line() & geom_point()

```
temp_graph_4 <- temp_graph_3 +  
  theme_minimal() +  
  theme(plot.subtitle =  
    element_markdown(hjust = 0, family  
      "barlow",  
      lineheight = unit(1.2, "pt"), size  
      13),  
    plot.title = element_markdown(famil  
      = "alegreyas_sc", size = 18),  
    axis.text.x = element_text(family =  
      "special", angle = 45, size = 9),  
    axis.text.y = element_text(family =  
      "special", size = 8), axis.title =  
    element_text(family = "special", si  
      = 10))
```

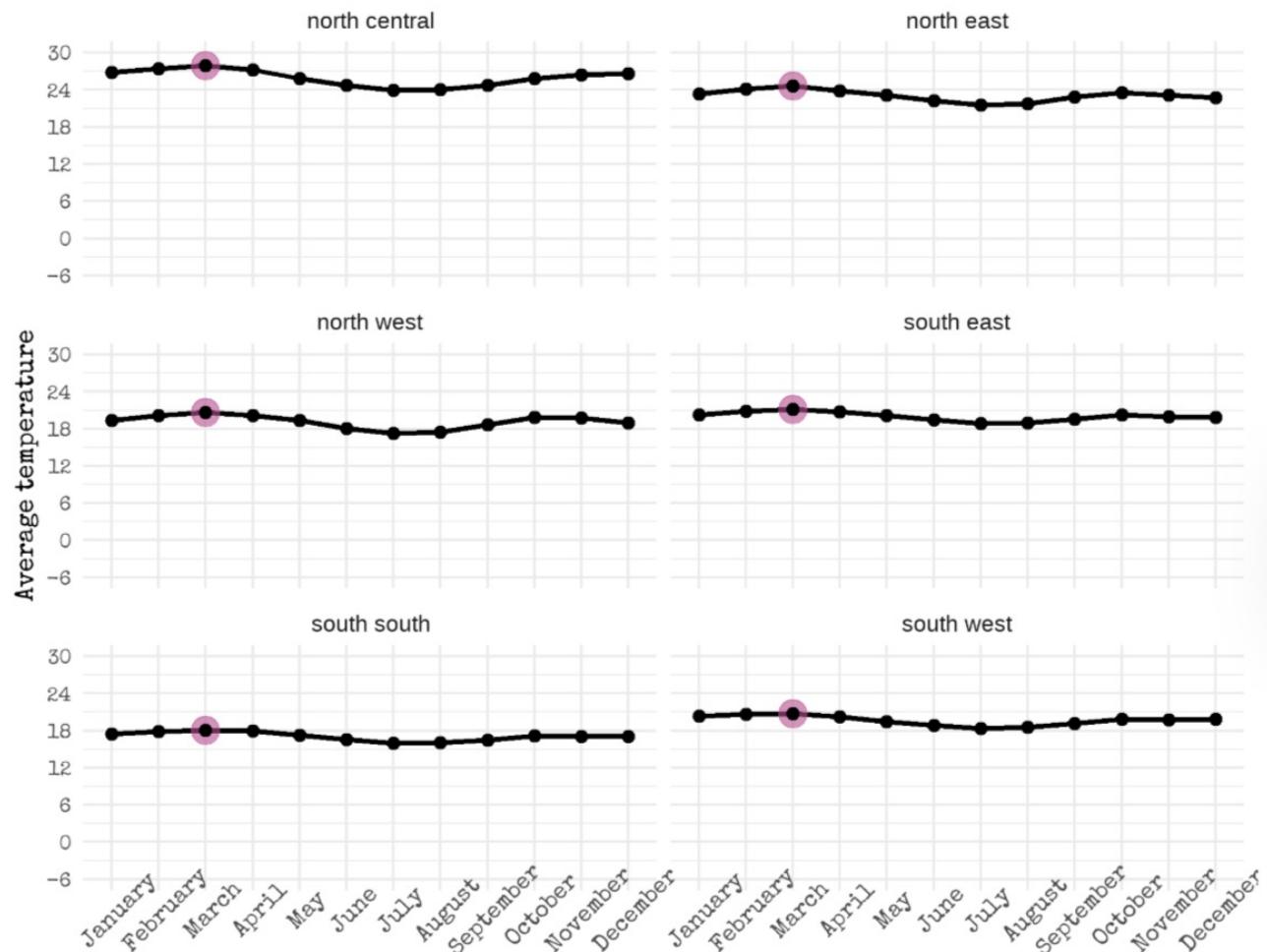


```
temp_graph_4
```

READ MORE: <https://ggplot2.tidyverse.org/>

MARCH IS THE HOTTEST MONTH OF THE YEAR IN KENYA

Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions

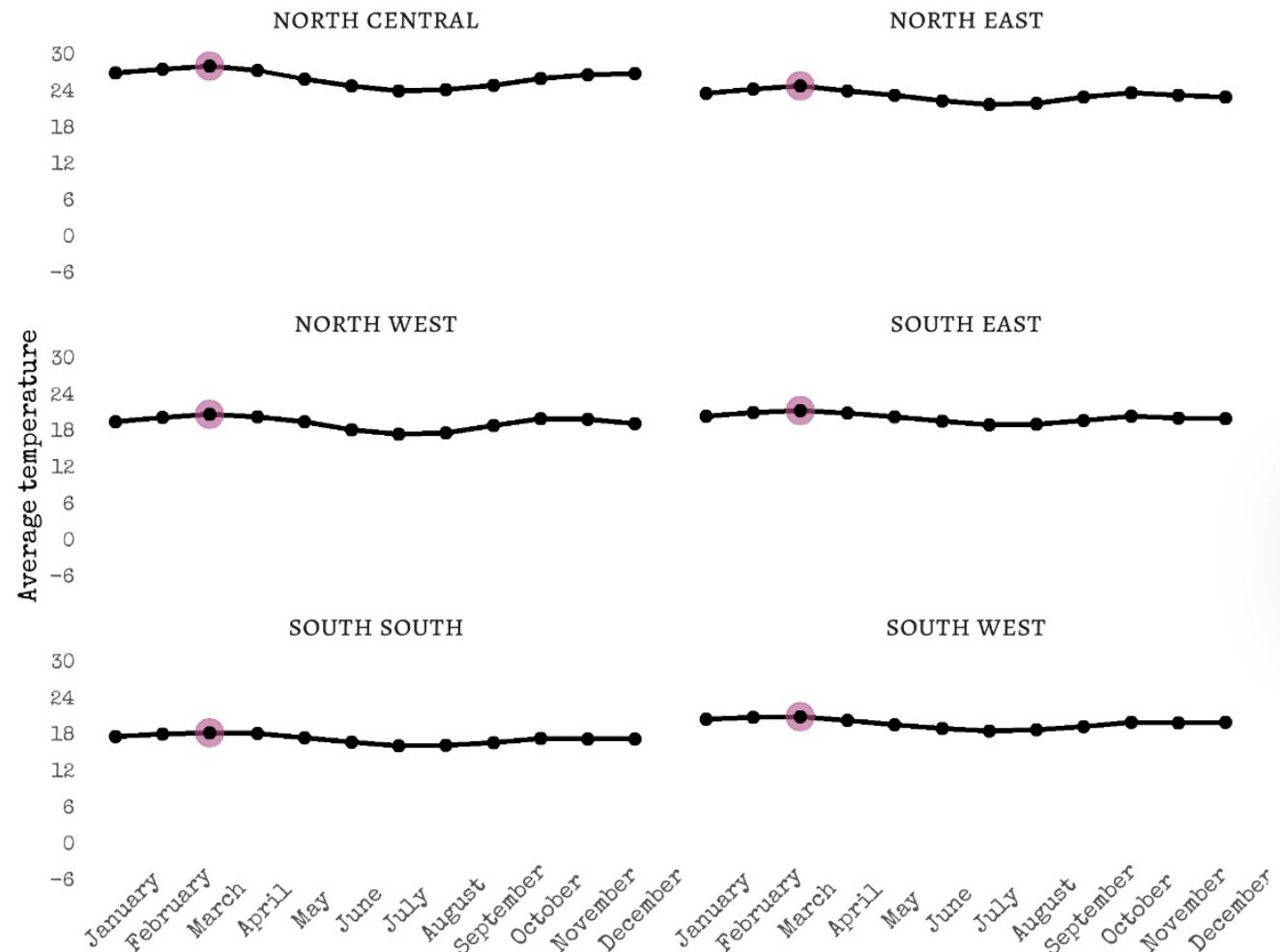


Visualising trends with lineplots (8/9): geom_line() & geom_point()

```
temp_graph_5 <- temp_graph_4 +  
  theme_minimal() +  
  theme(axis.title =  
    element_text(family =  
      "special", size = 10),  
    panel.grid = element_blank(),  
    strip.text =  
    element_text(family =  
      "alegreya_sc", size = 12))  
  
temp_graph_5
```

MARCH IS THE HOTTEST MONTH OF THE YEAR IN KENYA

Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions

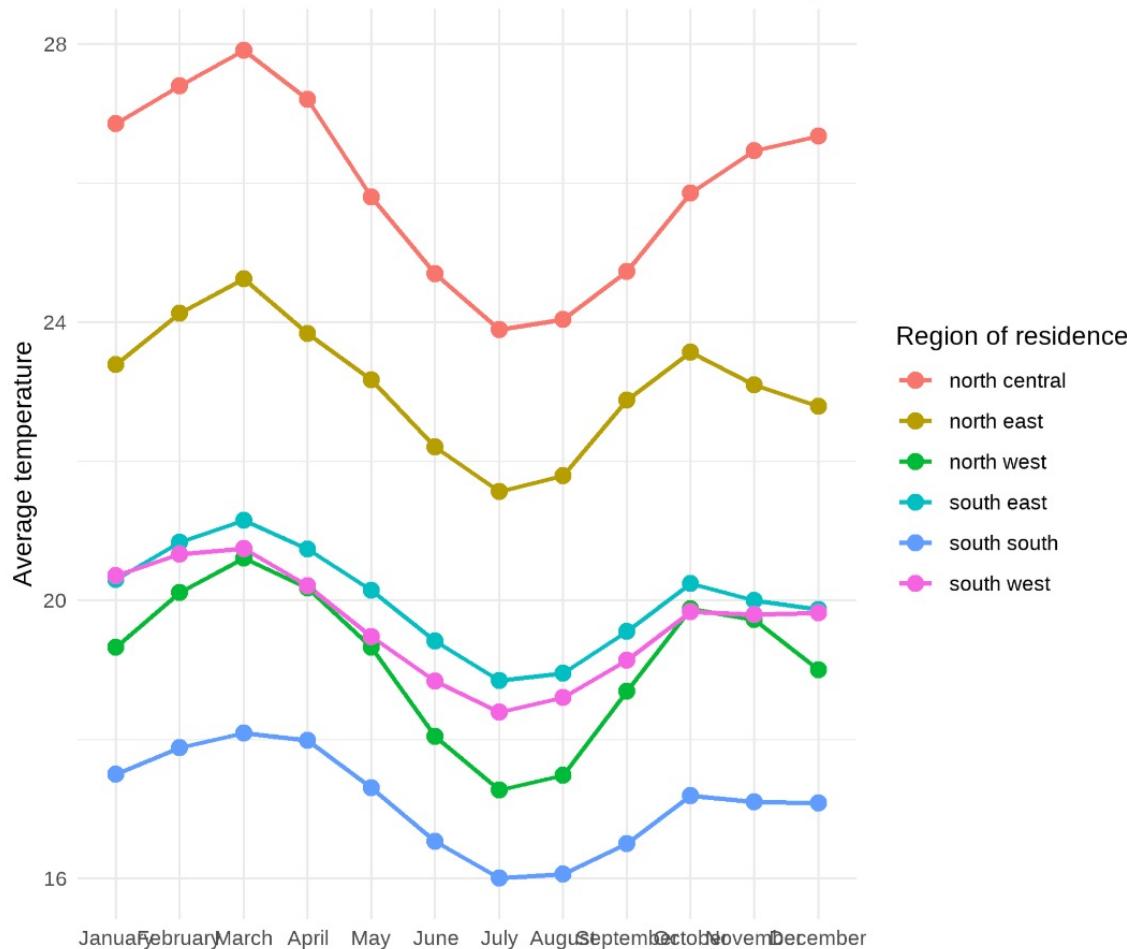


READ MORE: <https://ggplot2.tidyverse.org/>

Visualising trends with lineplots (9/9): geom_line() & geom_point()

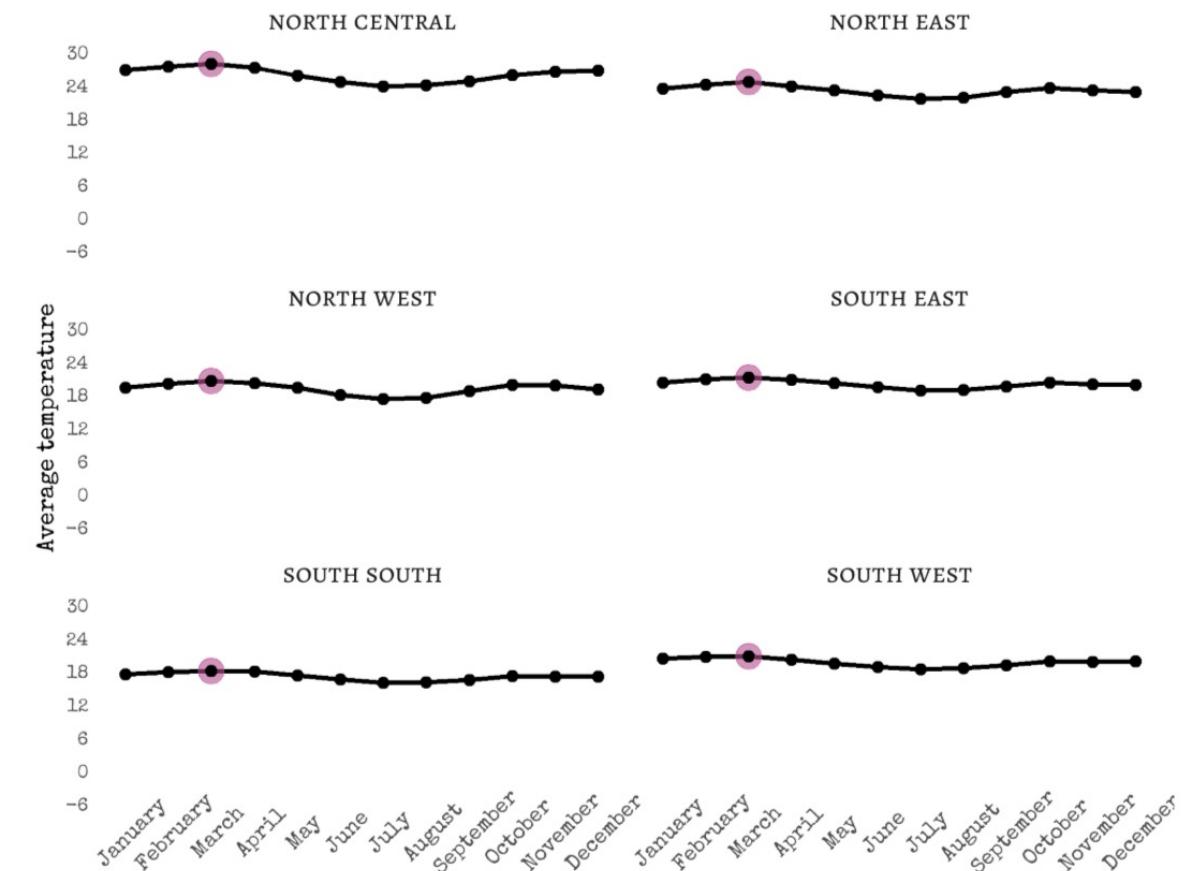
March is the hottest month of the year in Kenya

Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions



MARCH IS THE HOTTEST MONTH OF THE YEAR IN KENYA

Average temperature varies significantly across regions in Kenya.
Less hot in the South-South and hottest in the North-Central regions



READ MORE: <https://ggplot2.tidyverse.org/>

Your Turn

Open the file
`intermediate_ggplot.Rmd`
in `04-practicals`

Got questions?

Email me: `e.olamijuwon@st-andrews.ac.uk`