# This sets plot images to a nice size

options(repr.plot.width = 6, repr.plot.height = 6)

# Loading packages

library(dplyr)

library(tidyr)

library(ggplot2)

# Loading data

life\_expectancy <- read\_excel("~/UNdata.xlsx")

# Taking a look at the first few rows

head(life\_expectancy)

# Subsetting and reshaping the life expectancy data

subdata <- life\_expectancy %>%

filter(Year=='2000-2005') %>%

select(Country, Subgroup, Value) %>%

spread(Subgroup, Value)

# Taking a look at the first few rows

head(subdata)

# Plotting male and female life expectancy

ggplot(subdata,aes(x=Male, y=Female)) +

geom\_point()

Chart, scatter chart

Description automatically generated

# Adding an abline and changing the scale of axes of the previous plots

ggplot(subdata, aes(x=Male, y=Female)) +

geom\_point() +

geom\_abline(intercept = 0, slope=1, linetype=2) +

scale\_x\_continuous(limits=c(35,85)) +

scale\_y\_continuous(limits=c(35,85))

Chart, scatter chart

Description automatically generated

# Adding labels to previous plot

ggplot(subdata, aes(x=Male, y=Female))+

geom\_point(colour="white", fill="chartreuse3", shape=21, alpha=.55, size=5)+

geom\_abline(intercept = 0, slope = 1, linetype=2)+

scale\_x\_continuous(limits=c(35,85))+

scale\_y\_continuous(limits=c(35,85))+

labs(title="Life Expectancy at Birth by Country",

subtitle="Years. Period: 2000-2005. Average.",

caption="Source: United Nations Statistics Division",

x="Males",

y="Females")

# Subseting data to obtain countries of interest

top\_male <- subdata %>% arrange(Male-Female) %>% head(3)

top\_female <- subdata %>% arrange(Female-Male) %>% head(3)

# Adding text to the previous plot to label countries of interest

ggplot(subdata, aes(x=Male, y=Female, label=Country))+

geom\_point(colour="white", fill="chartreuse3", shape=21, alpha=.55, size=5)+

geom\_abline(intercept = 0, slope = 1, linetype=2)+

scale\_x\_continuous(limits=c(35,85))+

scale\_y\_continuous(limits=c(35,85))+

labs(title="Life Expectancy at Birth by Country",

subtitle="Years. Period: 2000-2005. Average.",

caption="Source: United Nations Statistics Division",

x="Males",

y="Females")+

geom\_text(data=top\_male, size=3) +

geom\_text(data=top\_female, size=3) +

theme\_bw()

Chart, scatter chart

Description automatically generated

# Taking a look at the first few rows

head(subdata2)

# A tibble: 6 x 10

Country Source Unit `Value Footnote~ Female\_1985\_1990 Female\_2000\_2005

<chr> <chr> <chr> <dbl> <dbl> <dbl>

1 Afghanistan UNPD\_Worl~ Years NA 41 42

2 Albania UNPD\_Worl~ Years NA 75 79

3 Algeria UNPD\_Worl~ Years NA 67 72

4 Angola UNPD\_Worl~ Years NA 42 43

5 Argentina UNPD\_Worl~ Years NA 75 78

6 Armenia UNPD\_Worl~ Years NA 71 75

# ... with 4 more variables: Male\_1985\_1990 <dbl>, Male\_2000\_2005 <dbl>,

# diff\_Female <dbl>, diff\_Male <dbl>

# Doing a nice first version of the plot with abline, scaling axis and adding labels

ggplot(subdata2, aes(x=diff\_Male, y=diff\_Female, label=Country))+

geom\_point(colour="white", fill="chartreuse3", shape=21, alpha=.55, size=5)+

geom\_abline(intercept = 0, slope = 1, linetype=2)+

scale\_x\_continuous(limits=c(-25,25))+

scale\_y\_continuous(limits=c(-25,25))+

labs(title="Life Expectancy at Birth by Country in Years",

subtitle="Difference between 1985-1990 and 2000-2005. Average.",

caption="Source: United Nations Statistics Division",

x="Males",

y="Females")+

theme\_bw()

Chart, scatter chart

Description automatically generated

# Subseting data to obtain countries of interest

top <- subdata2 %>% arrange(diff\_Male+diff\_Female) %>% head(3)

bottom <- subdata2 %>% arrange(-(diff\_Male+diff\_Female)) %>% head(3)

# Adding text to the previous plot to label countries of interest

ggplot(subdata2, aes(x=diff\_Male, y=diff\_Female, label=Country), guide=FALSE)+

geom\_point(colour="white", fill="chartreuse3", shape=21, alpha=.55, size=5)+

geom\_abline(intercept = 0, slope = 1, linetype=2)+

scale\_x\_continuous(limits=c(-25,25))+

scale\_y\_continuous(limits=c(-25,25))+

geom\_hline(yintercept=0, linetype=2)+

geom\_vline(xintercept=0, linetype=2)+

labs(title="Life Expectancy at Birth by Country",

subtitle="Years. Difference between 1985-1990 and 2000-2005. Average.",

caption="Source: United Nations Statistics Division",

x="Males",

y="Females")+

geom\_text(data=top, size=3)+

geom\_text(data=bottom, size=3)+

theme\_bw()

Chart, scatter chart

Description automatically generated