Analysis of Population Dynamics by Age Groups and Regions

Bailey Dalton, Dhanya Dharshini Veeraraghavan, Vincent Currie

30-NOV-2024


```
#
# Set the default behaviour for displaying code in all chunks to ensure that all display their code,
# unless turned off within a specific chunk.
#
knitr::opts_chunk$set(echo = TRUE, warning = FALSE)
```

1 Introduction

This report examines the evolution of gender gaps in education, from primary to tertiary levels, through historical and modern contexts, analysing regional progress, policy impacts, and links to socioeconomic development. It highlights global trends and contrasts between developed and developing countries and explores STEM-specific disparities, including Australia's gender parity across STEM and non-STEM fields.

The authors have undertaken specific parts of the analysis: [Author 1] explores historical trends, [Author 2] investigates regional progress, and Vince Currie examines STEM-specific disparities. These individual sections come together in the conclusion to provide a comprehensive understanding of the gender gap research questions and their broader implications.

1.1 Research Questions

This section investigates critical research questions about gender gaps in education across various levels and fields. It examines global trends, regional differences, and how policies influence gender parity. The analysis emphasizes disparities in STEM and non-STEM fields, focusing on Australia's efforts to achieve equity in educational representation.

1.1.1 How does the gender gap in STEM education compare to other fields globally between 2015 and 2022? (Author: Vince Currie)

This analysis draws on data from the OECD (2024) to examine gender disparities in education, focusing on the number of enrolled students, new entries, and graduates. It assesses female representation in STEM fields relative to other areas, aiming to determine whether STEM fields experience a shortfall of females. The source data is available here (https://data-explorer.oecd.org/vis?fs%5B0%5D=Topic%2C1%7CEducation%20and%20skills%23EDU%23%7CStudents%23EDU_STU%23&pg=0&fc=Topic&bp=true&snb=31&df%5Bds%5D=dsDiss

Table 2.3 provides a detailed breakdown of various vocational areas. STEM fields are highlighted in *italics* and large font, while the broader categories of **STEM** and **OTHER** represent collections of different vocational groupings. STEM includes the highlighted fields, while OTHER covers all remaining professions.

The analysis calculates the gender gap as the difference between the number of males and females in each global category (males - females). A positive value indicates a shortfall of females in STEM (males outnumber females), while a negative value reflects more females than males, suggesting no shortfall.

1.1.2 How does gender parity in STEM education compare to non-STEM fields in Australia across different levels and fields from 2015 to 2022? (Author: Vince Currie)

This analysis narrows its focus to Australia, drawing on data from the OECD (2024) to examine gender disparities in education across different levels (e.g., Bachelor's, Master's, and Vocational) and fields of study. It evaluates gender parity by comparing the representation of females and males in STEM fields to non-STEM fields (OTHER) within the Australian context from 2015 to 2022. The aim is to identify specific areas and education levels where gender disparities are most pronounced and assess progress toward achieving parity. The source data is available here (https://data-explorer.oecd.org/vis?fs%5B0%5D=Topic%2C1%7CEducation%20and%20skills%23EDU%23%7CStudents%23EDU_STU%23&pg=0&fc=Topic&bp=true&snb=31&df%5Bds%5D=dsDiss

Table 2.3 provides a breakdown of various vocational areas in Australia. STEM fields are highlighted in italics and large font, while the OTHER category includes all non-STEM fields.

The analysis calculates the gender gap as the difference between the percentage of males and females in each category (males - females). A positive value indicates male dominance (males outnumber females), while a negative value indicates female dominance (females outnumber males). A value close to zero reflects near parity. Near parity is defined as $|Gender Gap|<= $, meaning the representation of males and females is approximately equal within this range.

2 Analysis

2.1 Data Preparation

This section describes the steps to preprocess and organise the data, including cleaning, transforming, and categorising variables. It ensures the data-set's readiness for analysis and aligns with the research objectives.

```
#
# Load the dataset.
#
fp1 <- here("Data", "population-young-working-elderly.csv")

#
# Load the Population Young Working Elderly CSV file.
#
d1 <- read_csv(fp1, show_col_types = FALSE)

#
# Set the relative file path (assuming the file is in the 'Data' directory).
# Use here() to specify the relative file path using the root of the R project
# (where the .Rproj file is located).
#
fp2 <- here("Data", "Gender_education_area2015_2022.csv")

#
# Load the Gender Education csv.
#
d2 <- read_csv(fp2, show_col_types = FALSE)

#
# Vector of the required columns (Author: Vince Currie).
#
sc2 <- c("Reference area",
         "Education level",
         "Field of education",
         "Sex",
         "OBS_VALUE")

#
# The following steps are undertaken (Author: Vince Currie):
#    Rename the column names to something more meaningful.
#    Obtain the mean percentage by gender category for country, education level and field using group by and summarise.
#    Turn the gender category into columns with the calculated means using pivot: giving "Total" and "Female" columns.
#    Calculate the "Male" %  using mutate to create the column as Male = Total - Female. The coalesce replaces NA with 0.
#
#
 rn2 <- d2[, sc2] %>%
    rename(Cntry = "Reference area",           # Country.
           Edctn_lvl = "Education level",      # Education Level.
           Fld = "Field of education",         # Field.
           Gndr_ctgry = "Sex",                 # Gender Category (Values: Female and Total)
           Obsrvtn_vl = "OBS_VALUE")  %>%      # Percentage. For example, percentage of female participation.
    group_by(Cntry, Edctn_lvl, Fld, Gndr_ctgry)  %>%
    summarise(mObsrvtn_vl = mean(Obsrvtn_vl,na.rm = TRUE),.groups = "drop") %>%
    pivot_wider(names_from = Gndr_ctgry, values_from = mObsrvtn_vl) %>%
    mutate(Male = coalesce(Total,0) - coalesce(Female,0))

#
# Define a vector of STEM-related keywords to search for to classify field as stem or not (Author: Vince Currie).
#
 stem_keywords <- c("Science", "Technology", "Engineering", "Mathematics", "ICT", "Veterinary")

#
# Add a new column 'STEM_Status' categorizing as 'STEM' or 'OTHER' (Author: Vince Currie).
#    Exclude "Total (excluding unknown)" as its not a discipline.
#    Add STEM_Status to the data frame using mutate:
#       * Collapse stem_keywords down to a string of "or" cases using paste and collapse keyword.
#       * Search the "Fld" column looking for a stem keyword in each row using "grepl" command, ignoring case.
#       *   If a stem keyword is found, assign the STEM_Status column the value "STEM",
#       *   else assign the STEM_Status column the value "OTHER" using the "case_when" command.
#       *   Make the special case of "Social sciences" "OTHER" as its not part of stem but "science" is in the name.
#    Change the STEM_Status column to categorical using mutate and,
#    Calculate the Gender Gap which is the difference between Male and Female percentages working in that field:
#       If the Gender Gap is positive, it means that Males are dominant in the field,
#       else it means that females are dominant if the field for Gender Gap < 0.
#
rn2 <- rn2 %>%
  filter(Fld != "Total (excluding unknown)") %>%
  mutate(STEM_Status = case_when(
    grepl(paste(stem_keywords, collapse = "|"), Fld, ignore.case = TRUE) &
    !grepl("social science|arts|humanities", Fld, ignore.case = TRUE) ~ "STEM",
    TRUE ~ "OTHER"
  )) %>%
  mutate(STEM_Status = factor(STEM_Status),
         Gender_Gap = round(coalesce(Male,0) - coalesce(Female,0),2))
```

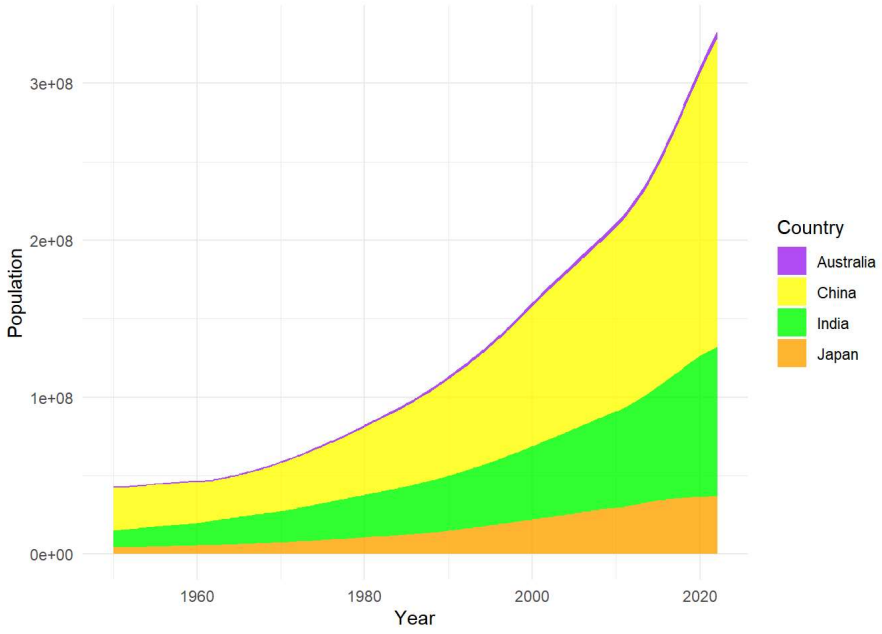
```
#
# Get the unique field areas being compared as part of Gender Gap (Author: Vince Currie).
#
unique_fields <- unique(rn2$Fld)

#
# Convert the unique fields into a data frame for display as a formatted table (Author: Vince Currie).
#
unique_fields_df <- data.frame(FieldArea = unique_fields)

#
# Rename the column from 'FieldArea' to 'Field Area' (Author: Vince Currie).
#
colnames(unique_fields_df)[colnames(unique_fields_df) == "FieldArea"] <- "Field Area"

#
# Calculate the summary statistics for gender gap by grouping by STEM_Status (Author: Vince Currie).
# Apply multiple summary functions to 'Gender_Gap' using across within summarise.
#   * ".x" represents the value of "Gender_Gap".
#   * ".names = "{.fn}_Gender_Gap" creates the columns as Lower_Gender_Gap, and so on.
#   * The results are rounded to 2 decimal places using round.
#   * The Lower and Upper are capped at -100 and 100 using pmin and pmax. There's only 100% of something.
#
summary_stats <- rn2 %>%
  group_by(STEM_Status) %>%
  summarise(
    across(
      Gender_Gap,
      list(
        Lower = ~round(pmax(min(.x, na.rm = TRUE), -100), 2),  # Ensure min is capped at -100
        Upper = ~round(pmin(max(.x, na.rm = TRUE), 100), 2),   # Ensure max is capped at 100
        Median = ~round(median(.x, na.rm = TRUE), 2),
        Mean = ~round(mean(.x, na.rm = TRUE), 2)
      ),
      .names = "{.fn}_Gender_Gap"
    )
  )

#
# Calculate Q1 (25%) and Q3 (75%) for Gender_Gap grouped by STEM_Status (Author: Vince Currie).
#
iqr_stats <- rn2 %>%
  group_by(STEM_Status) %>%
  summarise(
    Q1 = round(quantile(Gender_Gap, 0.25, na.rm = TRUE),2),
    Q3 = round(quantile(Gender_Gap, 0.75, na.rm = TRUE),2)
  )

#
# Australia only rows for visualising the Australia data (Author: Vince Currie):
#   * Filter the Australian data.
#   * Select the required columns.
australia_data <- rn2 %>%
  filter(Cntry == "Australia") %>%
  select(Cntry, Edctn_lvl, Fld, STEM_Status, Gender_Gap)

#
# Australia data table (Author: Vince Currie).
#   * Rename columns for display in table.
#   * Select the fields.
#   * Sort the data by Field and Education_Level using arrange.
#   * Display Field once per group using lag and mutate.
#
# The table is used to augment the figure by showing those
# rows where the gender gap is close to parity (|Gender Gap| <= 1).
#
australia_tab <- australia_data %>%
  rename(Field = Fld,
         Level = Edctn_lvl,
         `Gender Gap` = Gender_Gap,
         `Stem Status` = STEM_Status) %>%
  arrange(`Stem Status`, Field, Level) %>%
  mutate(
    Field = ifelse(is.na(lag(Field)) | lag(Field) != Field, Field, "")
  ) %>%
  filter(abs(`Gender Gap`) <= 1) %>%
  select( Field, `Stem Status`, Level, `Gender Gap`)
```

2.2 Visualisation

The visualisation section presents key insights through graphs and charts, illustrating gender gaps across education levels, fields, and regions. It highlights disparities in STEM and non-STEM fields, global trends, and Australia's context. These visualisations enhance understanding by showcasing patterns, variability, and progress toward achieving gender parity in education.

```
# Define countries to highlight
highlight_countries <- c("United States", "China", "India", "Australia", "Japan")

# Filter data for highlighted countries
# VC: highlighted_data <- data %>% filter(Entity %in% highlight_countries)
highlighted_data <- d1 %>% filter(Entity %in% highlight_countries)

# Create the stacked area chart
ggplot(highlighted_data, aes(x = Year, y = Population_65plus, fill = Entity)) +
  geom_area(alpha = 0.8) +  # Use geom_area for stacked chart
  labs(x = "Year",
       y = "Population",
       fill = "Country") +
  theme_minimal() +
  scale_fill_manual(values = c(
    "United States" = "red",
    "China" = "yellow",
    "India" = "green",
    "Australia" = "purple",
    "Japan" = "orange"
  ))
```

Figure 2.1: Population Aged 65+ for Selected countries (1950-2020) - (Author:Dhanya)

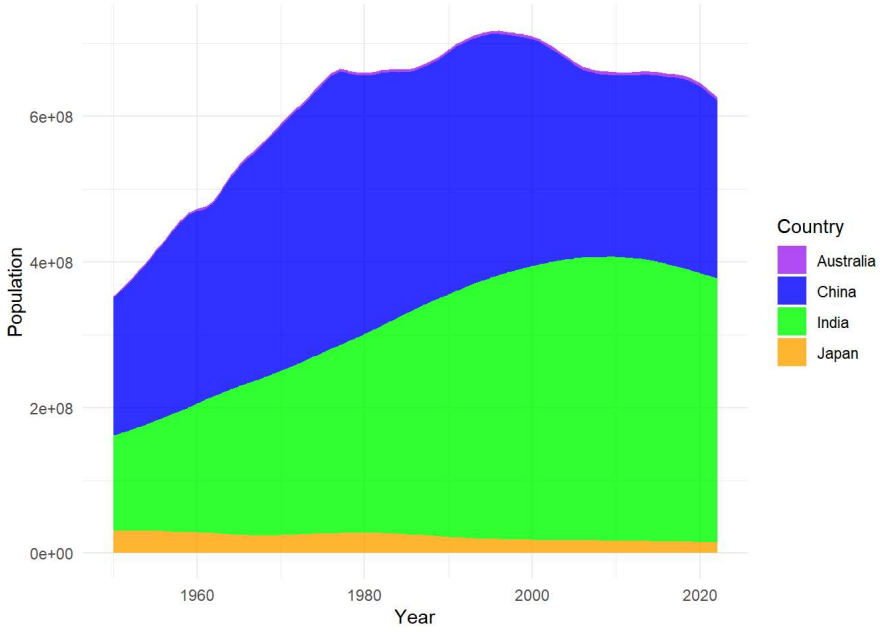



```
# Define countries to include
selected_countries <- c("United States", "China", "India", "Australia", "Japan")

# Filter data for the selected countries
# VC: filtered_data <- data %>%
filtered_data <- d1 %>%
  filter(Entity %in% selected_countries) %>%
  group_by(Entity, Year) %>%
  summarise(Population_0to14 = sum(Population_0to14, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'Entity'. You can override using the
## `.groups` argument.
```

```
# Create the stacked area chart
ggplot(filtered_data, aes(x = Year, y = Population_0to14, fill = Entity)) +
  geom_area(alpha = 0.8) +  # Stacked area chart
  labs(x = "Year",
       y = "Population",
       fill = "Country") +
  theme_minimal() +
  scale_fill_manual(values = c(
    "United States" = "red",
    "China" = "blue",
    "India" = "green",
    "Australia" = "purple",
    "Japan" = "orange"
  ))
```

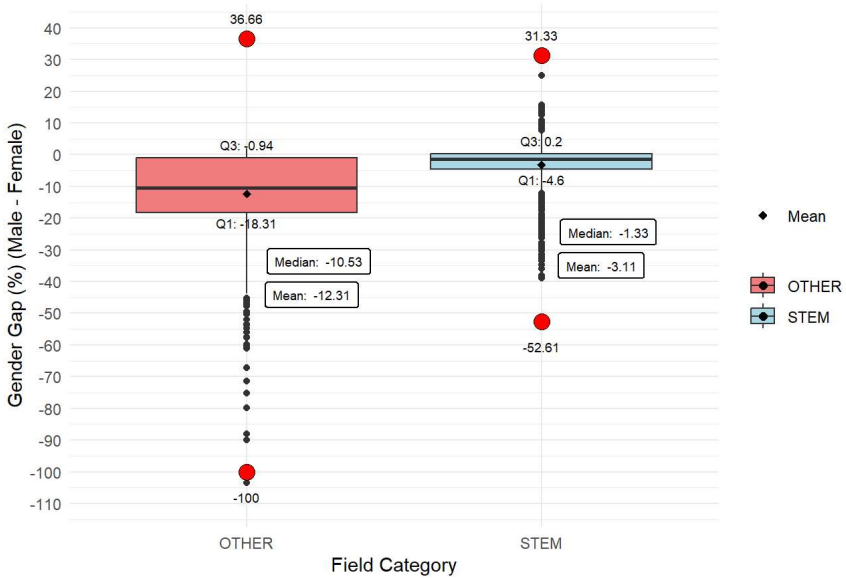
Figure 2.2: Population Aged 0-14 for Selected Countries (1950-2020) - (Author:Dhanya)



```
#
# Visualize the gender gap for STEM and non-STEM Vocations using a box plot with "geom_boxplot".
#
ggplot(rn2, aes(x = STEM_Status, y = Gender_Gap, fill = STEM_Status)) +
  geom_boxplot() +

  # Display a circle at the lower and upper ends of the whiskers.
  geom_point(
    data = summary_stats,
    aes(x = STEM_Status, y = Lower_Gender_Gap),
    shape = 21,
    fill = "red",
    size = 4
  ) +
  geom_point(
    data = summary_stats,
    aes(x = STEM_Status, y = Upper_Gender_Gap),
    shape = 21,
    fill = "red",
    size = 4
  ) +

  # Display the lower and upper values.
  geom_text(
    data = summary_stats,
    aes(x = STEM_Status, y = Lower_Gender_Gap, label = Lower_Gender_Gap),
    vjust = 3,
    size = 2.5
  ) +
  geom_text(
    data = summary_stats,
    aes(x = STEM_Status, y = Upper_Gender_Gap, label = Upper_Gender_Gap),
    vjust = -1.5,
    size = 2.5
  ) +

  # Display the middle 50%.
geom_text(
  data = iqr_stats,
  aes(x = STEM_Status, y = Q1, label = paste("Q1:", round(Q1, 2))),
  hjust = 0.5,                                # Centre horizontally above the bar.
  vjust = 1.5,                                # Adjust vertically to avoid overlap.
  color = "black", size = 2.5
) +
geom_text(
  data = iqr_stats,
  aes(x = STEM_Status, y = Q3, label = paste("Q3:", round(Q3, 2))),
  hjust = 0.5,                                # Centre horizontally above the bar.
  vjust = -0.8,                               # Adjust vertically to avoid overlap.
  color = "black", size = 2.5
) +

  # Add mean value diamond.
  geom_point(
    data = summary_stats,
    aes(x = STEM_Status, y = Mean_Gender_Gap, shape = "Mean"),
    size = 2,
    color = "black"
  ) +

  labs(x = "Field Category", y = "Gender Gap (%) (Male - Female)") +

  # A legend for the mean value.
  scale_shape_manual(name = "Statistics", values = c("Mean" = 18)) +
  theme_minimal() +

  # Add call out boxes for the mean and median.
  geom_label(data = summary_stats,
             aes(x = STEM_Status, y = Mean_Gender_Gap, label = paste("Mean: ", round(Mean_Gender_Gap, 2))),
             color = "black", size = 2.5, label.padding = unit(0.5, "lines"), label.size = 0.5,
             vjust = 4.5, hjust = -0.2, fill = "white") +  # White background for the call out label.
  geom_label(data = summary_stats,
             aes(x = STEM_Status, y = Median_Gender_Gap, label = paste("Median: ", round(Median_Gender_Gap, 2))),
             color = "black", size = 2.5, label.padding = unit(0.5, "lines"), label.size = 0.5,
             vjust = 3.4, hjust = -0.2, fill = "white") +  # White background for the call out label.

  scale_fill_manual(values = c("STEM" = "lightblue", "OTHER" = "lightcoral")) +
  theme(legend.title = element_blank()) +
```

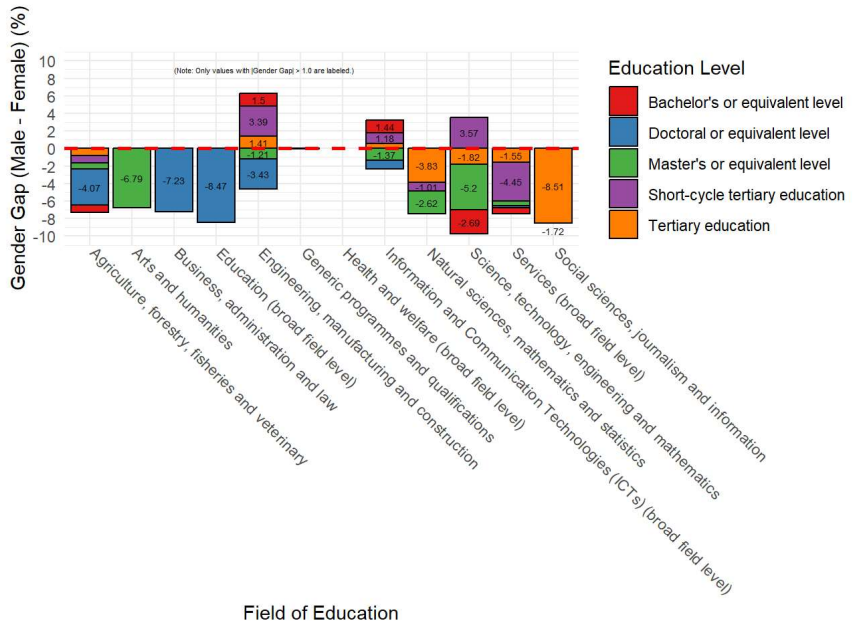
```
scale_y_continuous(breaks = seq(-110, 40, by = 10)) +  # Adjust y-axis ticks.
expand_limits(y = c(40,-110)) +                        # Ensure that the plot expands to y = (-110,40).

theme(legend.title = element_blank(),
  plot.margin = margin(t = 30, r = 10, b = 10, l = 20)
      )
```

Figure 2.3: World STEM Gender Gap(%)
(Based on 2015-2022 Data) - (Author: Vince Currie)



```
#
# Create a bar chart for Australian STEM education.
# (Useful information on colour: : https://www.r-bloggers.com/2013/09/how-to-expand-color-palette-with-ggplot-and-rcolorbrew
er/).
#
ggplot(australia_data, aes(x = Fld, y = Gender_Gap, fill = Edctn_lvl)) +
  geom_bar(stat = "identity", position = "stack", color = "black") +         # Use position = "stack" for stacking.
  scale_fill_brewer(palette = "Set1") +                                      # Distinct colours for education levels.
  geom_hline(yintercept = 0, color = "red", linetype = "dashed", size = 1) + # Add parity line.
  scale_y_continuous(
    limits = c(-10, 10),
    breaks = seq(-10, 10, by = 2)                                            # Add more y-axis ticks.
  ) +
  geom_text(
   aes(
    label = ifelse(abs(Gender_Gap) > 1.0, round(Gender_Gap, 2), "")          # Display label only if Gender_Gap > 1.0.
  ),
  position = position_stack(vjust = 0.5),                                    # Centre labels within segments.
  size = 1.8,
  color = "black"
)+
  annotate(
    "text",
    x = 3, y = 9,                                                            # Adjust the x and y positions.
    label = "(Note: Only values with |Gender Gap| > 1.0 are labeled.)",
    size = 1.5,
    hjust = 0,                                                               # Align text horizontally.
    color = "black"
  )+
  labs(
    x = "Field of Education",
    y = "Gender Gap (Male - Female) (%)",
    fill = "Education Level"
  ) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(
      angle = -45,
      hjust = 0,
      vjust = 0
    ),
    plot.margin = margin(t = 30, r = 10, b = 10, l = 20)
  )
```

Figure 2.4: Gender Gap in STEM Education
by
Field and Level in Australia (2015–2022) - (Author: Vince Currie)



2.3 Tables

The Tables section presents detailed data summaries supporting the report's analysis. It includes gender gap statistics across education levels, fields, and regions, highlighting disparities and progress. These tables provide precise numerical insights, complementing the visualizations and enabling a more profound exploration of global trends, STEM and non-STEM comparisons, and Australia's educational gender dynamics.

```
# Calculate growth in the population aged 65+
growth <- d1 %>%
# growth <- data %>%
  group_by(Entity) %>%
  summarise(Start_1950 = first(Population_65plus),
            End_2023 = last(Population_65plus)) %>%
  mutate(Growth = End_2023 - Start_1950) %>%
  arrange(desc(Growth)) %>%
  head(5)

# Display the table
kable(growth,
      booktabs = TRUE,
      align = "c",
      format = "html") %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "hover"),
    position = "center",
    full_width = FALSE
  )
```

Table 2.1: Top 5 Countries with Highest Growth in Population Aged 65+ (Author:Dhanya).

Entity	Start_1950	End_2023	Growth
China	27433876	196380099	168946223
India	10739080	95464582	84725502
Japan	4234245	36632124	32397879
Brazil	1270713	21552433	20281720
Russia	5279960	23392858	18112898

```
# Calculate growth in the population aged 0-14 for each country
growth <- d1 %>%
#growth <- data %>%
  group_by(Entity) %>%
  summarise(Start_1950 = first(Population_0to14),
            End_2023 = last(Population_0to14),
            Growth = End_2023 - Start_1950) %>%
  arrange(desc(Growth)) %>%
  head(5)

# Display the table
kable(growth,
      booktabs = TRUE,
      align = "c",
      format = "html") %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "hover"),
    position = "center",
    full_width = FALSE
  )
```

Table 2.2: Top 5 Countries with Highest Growth in Population Aged 0-14 - (Author:Dhanya).

Entity	Start_1950	End_2023	Growth
India	131348889	363307648	231958759
Nigeria	15555163	93617234	78062071

Entity	Start_1950	End_2023	Growth
Pakistan	15350690	90924190	75573500
China	189268864	244016392	54747528
Indonesia	28342278	70471815	42129537

```
#
# Create the formatted table to show the list of education fields.
# kable options:
# Here is an explanation of the bootstrap options:
#    striped: Adds alternating row colours, which improves readability, especially in long tables.
#    hover: The rows will highlight when you hover over them, making it easier to track position in the table.
#    condensed: The rows are more compact, reducing the vertical space between rows.
# Other kable options:
#    booktabs = TRUE, creates a cleaner, more professional table by replacing the default horizontal lines with well-spaced
rules.
#    align = "l", controls the alignment of the columns in the table. The value "l" stands for left alignment. Values can be
lcr for left/centre/right.
#    format = "html", the table will be rendered as an HTML table, which can be displayed in a browser.
# kableExtra options:
#    position = "centre", the table will be horizontally centre-aligned within the available space.
#    full_width = FALSE ensures the table doesn't stretch across the page, keeping it compact.
# row_spec, used to align row 0 to the centre(align = "c"), which is the column label for the table. The column name ("Field
Area") was left aligned by
# the previous kable step used to left-align (align="l") the Fld text, which was globally applied. Also, the rows 1,5, 8, 9,
10 are italic and large
# font as these represent STEM areas.
#
knitr::kable(unique_fields_df,
             booktabs = TRUE,
             align = "l",
             format = "html")  %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "hover", "condensed"),
    position = "center",
    full_width = FALSE
  ) %>%
  row_spec(row=0,align = "c") %>%
  row_spec(1, italic = TRUE, font_size = "large") %>%
  row_spec(5, italic = TRUE, font_size = "large") %>%
  row_spec(8, italic = TRUE, font_size = "large") %>%
  row_spec(9, italic = TRUE, font_size = "large") %>%
  row_spec(10,italic = TRUE, font_size = "large")
```

Table 2.3: Fields Compared for Gender Gap(%) Analysis (Based on 2015-2022 Data) - (Author: Vince Currie).

Field Area
Agriculture, forestry, fisheries and veterinary
Arts and humanities
Business, administration and law
Education (broad field level)
Engineering, manufacturing and construction
Generic programmes and qualifications
Health and welfare (broad field level)
Information and Communication Technologies (ICTs) (broad field level)
Natural sciences, mathematics and statistics
Science, technology, engineering and mathematics
Services (broad field level)
Social sciences, journalism and information

```
#
# Create the formatted table to show the summary statistics.
#
knitr::kable(summary_stats,
             booktabs = TRUE,
             align = "c",
             format = "html")  %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "hover", "condensed"),
    position = "center",
    full_width = FALSE
  )
```

Table 2.4: World STEM Gender Gap(%) Statistics (Based on 2015-2022 Data) - (Author: Vince Currie).

STEM_Status	Lower_Gender_Gap	Upper_Gender_Gap	Median_Gender_Gap	Mean_Gender_Gap
OTHER	-100.00	36.66	-10.53	-12.31
STEM	-52.61	31.33	-1.33	-3.11

```
knitr::kable(australia_tab,
             booktabs = TRUE,
             align = "l",
             format = "html")  %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "hover", "condensed"),
    position = "center",
    full_width = FALSE
  )%>%
  row_spec(row=0,align = "c")
```

Table 2.5: Gender Gap in STEM Education by Field and Level in Australia (2015–2022) that Approximates Parity - (Author: Vince Currie)

Field	Stem Status	Level	Gender Gap
Generic programmes and qualifications	OTHER	Bachelor's or equivalent level	-0.01
	OTHER	Doctoral or equivalent level	0.00
	OTHER	Master's or equivalent level	0.00
	OTHER	Short-cycle tertiary education	0.00
	OTHER	Tertiary education	-0.01
Services (broad field level)	OTHER	Bachelor's or equivalent level	-0.73
	OTHER	Doctoral or equivalent level	-0.18
	OTHER	Master's or equivalent level	-0.55
Agriculture, forestry, fisheries and veterinary	STEM	Bachelor's or equivalent level	-0.85
	STEM	Master's or equivalent level	-0.70
	STEM	Short-cycle tertiary education	-0.84
	STEM	Tertiary education	-0.83
	STEM	Doctoral or equivalent level	-0.95
	STEM	Tertiary education	0.60

2.3.1 Analysis - How does the gender gap in STEM education compare to other fields globally between 2015 and 2022 - (Author:Vince Currie)?

Figure 2.3 presents a box plot and statistical markers illustrating differences in gender gaps between STEM and OTHER fields. In STEM fields, the median gender gap is close to zero (-1.33), and the mean (-3.11) skews slightly negative, indicating modest improvements in gender balance. The relatively narrow interquartile range (IQR), with Q3 (0.2) highlighting areas of male dominance and Q1 (-4.6) reflecting progress in female representation, suggests reduced variability. However, the positive skewness of the box plot highlights persistent male dominance, particularly in disciplines like engineering and ICT.

In contrast, OTHER fields display a more pronounced negative mean (-12.31) and median (-10.53), indicating female-majority representation, particularly in health and social sciences. The broader IQR, spanning Q1 (-18.31) to Q3 (-0.94), suggests more significant variability in gender representation, with a notable skew toward female dominance.

Additionally, the extremes in Table 2.4 reveal significant male dominance in STEM (Upper:31.33%) and instances of female-majority representation (Lower:-52.61%), with narrower variability than in OTHER fields. OTHER fields show broader extremes (36.66% to -100%), reflecting greater gender diversity. Figure 2.3 depicts these extremes as red circles, emphasising the extent of gender disparities. These findings underscore the persistent underrepresentation of females in STEM and the critical need for targeted strategies to close the gender gap and foster equity in STEM occupations.

2.3.2 Analysis - How does gender parity in STEM education compare to non-STEM fields in Australia across different levels and fields from 2015 to 2022 - (Author:Vince Currie)?

Figure 2.4 illustrates the gender gap in STEM and non-STEM fields across different education levels in Australia from 2015 to 2022. The stacked bar chart reveals notable disparities, with STEM fields showing consistent male dominance across levels such as Bachelor's and Master's, particularly in disciplines like engineering and ICT. However, some fields at the vocational and PhD levels approach gender parity, with $||GenderGap|| \leq 1.0$ indicating near-equal representation. For example, in the Agriculture, Forestry, Fisheries, and Veterinary category, all levels except PhD demonstrate approximate parity. In contrast, females dominate the PhD level (Table 2.5). The table also highlights other fields and education levels where the gender gap is near parity.

Non-STEM (OTHER) fields demonstrate a contrasting trend, with a pronounced negative gender gap indicating female dominance. For instance, health and education fields exhibit significantly higher female representation across all levels, particularly in Bachelor's and Master's degrees (Figure 2.4). Variability in non-STEM fields reflects broader gender diversity, with some areas achieving parity or female dominance while others differ significantly from STEM. OTHERS differ significantly from STEM, where male dominance is more uniform and persistent, requiring targeted efforts to address gender disparities. The broader range of gender gap values in non-STEM fields underscores this diversity compared to STEM. These findings emphasize the need for targeted initiatives to encourage female participation in STEM while maintaining equity in non-STEM areas.

2.3.3 Analysis - (Author:Dhanya)

2.3.3.1 Figure 1: Population Aged 65+ for Selected Countries (1950-2020)

This chart highlights the rapid growth of elderly populations in selected countries. China shows the steepest increase due to improved healthcare and longer life expectancy (World Bank, 2023). Japan's sharp rise post-1970 reflects its aging crisis caused by low fertility and high longevity, leading to challenges in sustaining healthcare and pensions (OECD, 2023). India's slower growth suggests it remains in an earlier demographic stage, while the United States and Australia exhibit steady but less pronounced growth.

2.3.3.2 Figure 2: Population Aged 0-14 for Selected Countries (1950-2020)

India retains the largest 0-14 population due to high fertility rates (UNICEF, 2024). In contrast, China's decline post-1970s aligns with its one-child policy (Pletcher, K., 2024). Developed nations like the United States, Japan, and Australia show consistent declines, reflecting lower birth rates and the challenges of maintaining a sustainable workforce in aging populations.

2.3.3.3 Table 1: Top 5 Countries with Highest Growth in Population Aged 65+

China, India, and Japan lead in growth, driven by rising life expectancies and large populations (UN DESA, 2023). Brazil and Russia follow, reflecting similar aging trends in emerging economies as they transition demographically.

2.3.3.4 Table 2: Top 5 Countries with Highest Growth in Population Aged 0-14

India, Nigeria, and Pakistan dominate, driven by high fertility rates in developing nations. Indonesia's rise signals Southeast Asia's growing youth population, while China's moderated growth reflects earlier demographic trends during its high-fertility era.

3 Conclusion

These trends highlight aging challenges in developed nations and youth-driven opportunities in developing regions. Targeted policies in healthcare, education, and economic planning are essential to address these demographic shifts effectively.

Globally, STEM fields consistently exhibit male dominance with narrower gender variability compared to non-STEM fields. In Australia, STEM shows similar male dominance across most levels, while non-STEM fields demonstrate broader gender diversity and female-majority representation. Targeted initiatives remain critical to fostering equity in STEM both globally and within Australia.

References

OECD. (2023). *The silver economy: Policies for an aging society*. https://www.oecd.org (https://www.oecd.org)

OECD. (2024). *Distribution of enrolled students, new entrants, and graduates by field of education.*https://data-explorer.oecd.org (https://data-explorer.oecd.org)

Pletcher, K. (2024). *China's one-child policy and its effects.* https://www.britannica.com (https://www.britannica.com)

UNICEF. (2024). *State of the world's children.* https://www.unicef.org (https://www.unicef.org)

World Bank. (2023). *Population and demographic indicators.* https://www.worldbank.org (https://www.worldbank.org)