Take Home Exercise 1

2025-05-01

# Overview

## Setting the Scene

A local online media company that publishes daily content on digital platforms is planning to release an article on demographic structures and distribution of Singapore in 2024.

## The Task

Assuming the role of the graphical editor of the media company, I am tasked to prepare at most three data visualisation for the article.

## The Data

The dataset titled “Singapore Residents by Planning Area / Subzone, Single Year of Age and Sex, June 2024”, published by the [Department of Statistics Singapore (DOS)]((https://www.singstat.gov.sg/)), provides detailed demographic data of Singapore’s resident population. It breaks down the population by age, gender, and geographic location, including planning areas and subzones defined by the Urban Redevelopment Authority (URA).

# Getting Started

### Loading Packages

The code chunk below uses the p\_load() function from the pacman package to check whether the required are installed on the computer. If the packages are already installed, p\_load() will load them into the R session. If not, it will automatically install the missing packages before loading them.

pacman::p\_load(tidyverse,patchwork,ggthemes, DT, dplyr, knitr, ggiraph, scales, ggridges)

| **Library** | **Description** |
| --- | --- |
| [**pacman**](https://cran.r-project.org/package=pacman) | A package management tool that helps load, install, and manage R packages more easily. |
| [**tidyverse**](https://www.tidyverse.org/) | A collection of R packages designed for data science, sharing an underlying design philosophy, grammar, and data structures. |
| [**patchwork**](https://patchwork.data-imaginist.com/) | Helps combine multiple ggplot2 plots into a single composite figure. |
| [**ggthemes**](https://yixuan.cos.name/ggthemes/) | Provides extra themes, scales, and geoms for customizing ggplot2 visualizations. |
| [**DT**](https://rstudio.github.io/DT/) | Provides an R interface to the JavaScript library [DataTables](https://datatables.net/) that create interactive tables on HTML pages. |
| [**dplyr**](https://dplyr.tidyverse.org) | A grammar of data manipulation, providing functions to easily filter, select, mutate, group, and summarize data in R. |
| [**knitr**](https://yihui.org/knitr/) | An R package that turns dynamic R code into elegant, reproducible reports in formats like HTML, PDF, and Word. |
| [**ggiraph**](https://davidgohel.github.io/ggiraph/) | Adds interactive features (like tooltips and hover effects) to ggplot2 visualizations using HTML widgets. |
| [**scales**](https://scales.r-lib.org/) | Provides tools to customize axis scales, labels, breaks, and color gradients in ggplot2 plots. |
| [**ggridges**](https://cran.r-project.org/package=ggridges) | Creates ridge line plots (aka joyplots) with ggplot2, useful for visualizing distributions. |

### Data Preparation

#### Import Data

data <- read\_csv("data/respopagesex2024.csv")



#### Data Overview

column\_info <- data.frame(  
 Column = names(data),  
 Class = sapply(data, class),  
 Description = c(  
 "Name of the planning area in Singapore",  
 "Name of the subzone within the planning area",  
 "Single year of age",  
 "Sex of resident: Male or Female",  
 "Number of residents in the category",  
 "Reference time period for the population data"  
 )  
)  
  
kable(column\_info, caption = "Dataset Column Overview")

Dataset Column Overview

|  | Column | Class | Description |
| --- | --- | --- | --- |
| PA | PA | character | Name of the planning area in Singapore |
| SZ | SZ | character | Name of the subzone within the planning area |
| Age | Age | character | Single year of age |
| Sex | Sex | character | Sex of resident: Male or Female |
| Pop | Pop | numeric | Number of residents in the category |
| Time | Time | numeric | Reference time period for the population data |

summary(data)

PA SZ Age Sex   
 Length:60424 Length:60424 Length:60424 Length:60424   
 Class :character Class :character Class :character Class :character   
 Mode :character Mode :character Mode :character Mode :character   
   
   
   
 Pop Time   
 Min. : 0.0 Min. :2024   
 1st Qu.: 0.0 1st Qu.:2024   
 Median : 20.0 Median :2024   
 Mean : 69.4 Mean :2024   
 3rd Qu.: 90.0 3rd Qu.:2024   
 Max. :1180.0 Max. :2024

#### Data Cleaning

##### Check Missing Values

In this section, we examine the dataset for missing values across all variables as it might impact the accuracy of our analysis.

data.frame(  
 NA\_Counts = sum(is.na(data$Pop)),  
 Pop\_Zero\_Counts = sum(data$Pop == 0, na.rm = TRUE)  
)

NA\_Counts Pop\_Zero\_Counts  
1 0 23181

Here I noticed that there are **no missing (NA)** values in the dataset. Therefore, I proceeded to check for 0 values in the numerical variable, Population. Upon inspection, I found a significant number of 0 values. To ensure data quality and meaningful analysis, I further examined whether any **Planning Areas (PA)** have only **0 population** values across **all age groups and sexes**. These areas will be removed from the dataset as they do not contribute meaningful demographic information.

data <- data %>%  
 group\_by(PA) %>%  
 filter(sum(Pop, na.rm = TRUE) > 0) %>%  
 ungroup()

##### Check for Duplicates

data[duplicated(data), ]

# A tibble: 0 × 6  
# ℹ 6 variables: PA <chr>, SZ <chr>, Age <chr>, Sex <chr>, Pop <dbl>,  
# Time <dbl>

#### Data Wrangling

##### Age Grouping

The age groups in this analysis are based on the classifications used in Singapore’s Population in Brief report by [*National Population and Talent Division*](population.gov.sg). The grouping follows 5-year intervals, ending with a final category for individuals aged 90 and above.

data <- data %>%  
 mutate(AgeGroup = case\_when(  
 Age %in% c("0", "1", "2", "3", "4") ~ "0 – 4",  
 Age %in% c("5", "6", "7", "8", "9") ~ "5 – 9",  
 Age %in% c("10", "11", "12", "13", "14") ~ "10 – 14",  
 Age %in% c("15", "16", "17", "18", "19") ~ "15 – 19",  
 Age %in% c("20", "21", "22", "23", "24") ~ "20 – 24",  
 Age %in% c("25", "26", "27", "28", "29") ~ "25 – 29",  
 Age %in% c("30", "31", "32", "33", "34") ~ "30 – 34",  
 Age %in% c("35", "36", "37", "38", "39") ~ "35 – 39",  
 Age %in% c("40", "41", "42", "43", "44") ~ "40 – 44",  
 Age %in% c("45", "46", "47", "48", "49") ~ "45 – 49",  
 Age %in% c("50", "51", "52", "53", "54") ~ "50 – 54",  
 Age %in% c("55", "56", "57", "58", "59") ~ "55 – 59",  
 Age %in% c("60", "61", "62", "63", "64") ~ "60 – 64",  
 Age %in% c("65", "66", "67", "68", "69") ~ "65 – 69",  
 Age %in% c("70", "71", "72", "73", "74") ~ "70 – 74",  
 Age %in% c("75", "76", "77", "78", "79") ~ "75 – 79",  
 Age %in% c("80", "81", "82", "83", "84") ~ "80 – 84",  
 Age %in% c("85", "86", "87", "88", "89") ~ "85 – 89",  
 Age == "90\_and\_Over" ~ "90 & Over",  
 TRUE ~ NA\_character\_  
 ))  
data

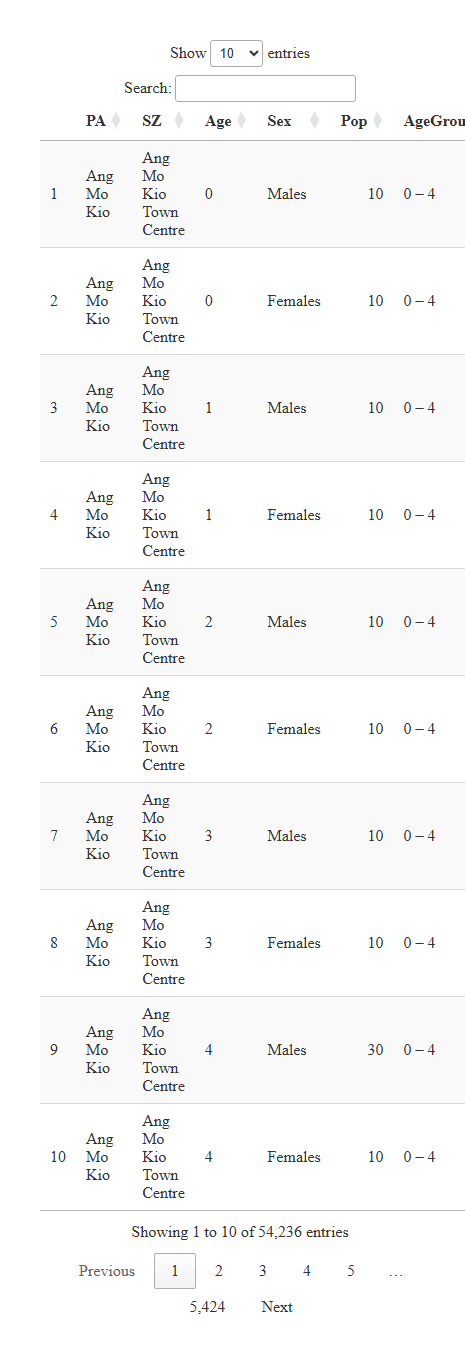
# A tibble: 54,236 × 7  
 PA SZ Age Sex Pop Time AgeGroup  
 <chr> <chr> <chr> <chr> <dbl> <dbl> <chr>   
 1 Ang Mo Kio Ang Mo Kio Town Centre 0 Males 10 2024 0 – 4   
 2 Ang Mo Kio Ang Mo Kio Town Centre 0 Females 10 2024 0 – 4   
 3 Ang Mo Kio Ang Mo Kio Town Centre 1 Males 10 2024 0 – 4   
 4 Ang Mo Kio Ang Mo Kio Town Centre 1 Females 10 2024 0 – 4   
 5 Ang Mo Kio Ang Mo Kio Town Centre 2 Males 10 2024 0 – 4   
 6 Ang Mo Kio Ang Mo Kio Town Centre 2 Females 10 2024 0 – 4   
 7 Ang Mo Kio Ang Mo Kio Town Centre 3 Males 10 2024 0 – 4   
 8 Ang Mo Kio Ang Mo Kio Town Centre 3 Females 10 2024 0 – 4   
 9 Ang Mo Kio Ang Mo Kio Town Centre 4 Males 30 2024 0 – 4   
10 Ang Mo Kio Ang Mo Kio Town Centre 4 Females 10 2024 0 – 4   
# ℹ 54,226 more rows

##### Drop Time Column

Since the Time column contains the same value for all rows, it does not contribute any analytical value. Therefore, I will remove this column to simplify the dataset.

data <- data %>% select(-Time)

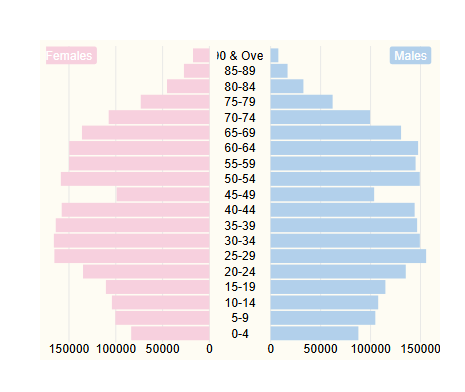
##### Final Dataset



# Data Visualization and Analysis

## Singapore Population Pyramid (2024)

pyramid\_data <- data %>%  
 group\_by(AgeGroup, Sex) %>%  
 summarise(Pop = sum(Pop), .groups = "drop")  
  
population\_pyramid\_women <- ggplot(  
 filter(pyramid\_data, Sex == "Females"),  
 aes(  
 x = Pop,  
 y = AgeGroup,  
 tooltip = paste("Population:", comma(Pop)),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(fill = "#F7d0de") +  
 scale\_x\_reverse() +  
 theme\_void() +  
 annotate(  
 geom = "label",  
 x = 150000,  
 y = "90 & Over",  
 label = "Females",  
 fill = "#F7d0de",  
 color = "white",  
 label.size = 0,  
 label.padding = unit(0.3, "lines")  
 ) +  
 theme(  
 axis.text.x = element\_text(),  
 panel.grid.major.x = element\_line(color = "grey90"),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
population\_pyramid\_men <- ggplot(  
 filter(pyramid\_data, Sex == "Males"),  
 aes(  
 x = Pop,  
 y = AgeGroup,  
 tooltip = paste("Population:", comma(Pop)),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(fill = "#b2d0eb") +  
 theme\_void() +  
 annotate(  
 geom = "label",  
 x = 140000,  
 y = "90 & Over",  
 label = "Males",  
 fill = "#b2d0eb",  
 color = "white",  
 label.size = 0,  
 label.padding = unit(0.3, "lines")  
 ) +  
 theme(  
 axis.text.x = element\_text(),  
 panel.grid.major.x = element\_line(color = "grey90"),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
age\_labels <- tibble(age = fct\_inorder(c(  
 "0-4", "5-9", "10-14", "15-19", "20-24", "25-29", "30-34",  
 "35-39", "40-44", "45-49", "50-54", "55-59", "60-64",  
 "65-69", "70-74", "75-79", "80-84", "85-89", "90 & Over"  
)))  
  
age\_labels\_plot <- ggplot(age\_labels, aes(x = 1, y = age, label = age)) +  
 geom\_text\_interactive(aes(tooltip = age, data\_id = age)) +  
 theme\_void() +  
 theme(  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = population\_pyramid\_women + age\_labels\_plot + population\_pyramid\_men +  
 plot\_layout(widths = c(7.5, 2, 7.5)) +  
 plot\_annotation(  
 theme = theme(  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA),  
 panel.border = element\_blank()  
 )  
 ),  
 options = list(  
 opts\_hover(css = "stroke:none;"),  
 opts\_hover\_inv(css = "opacity:0.2;"),  
 opts\_tooltip(  
 css = "background-color:#fef3c7;  
 color:#A94442;  
 padding:6px;  
 border-radius:6px;  
 font-weight:bold;  
 box-shadow: 0px 2px 5px rgba(0,0,0,0.2);"  
 )  
)  
)

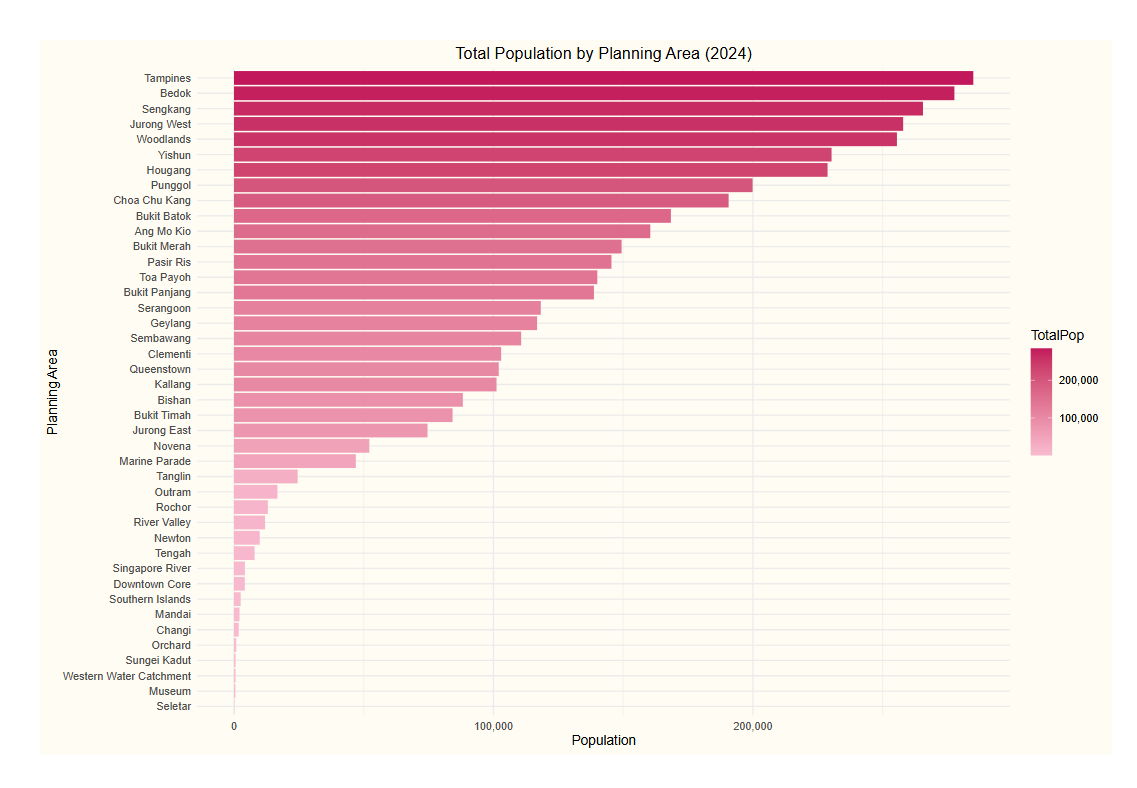


The total population for female : 2,041,480 The total population for male : 2,152,050

|  |  |
| --- | --- |
|  | **Insights**  The Singapore’s 2024 population pyramid reveals a balanced distribution between males and females, with 2,152,050 males and 2,041,480 females.  The largest population groups are concentrated in the 25–54 age range, which suggest to a strong and stable working-age demographic.  In contrast, the younger age groups make up a smaller share. |

## Singapore Total Population by Planning Area (2024)

pa\_population <- data %>%  
 group\_by(PA) %>%  
 summarise(TotalPop = sum(Pop, na.rm = TRUE)) %>%  
 arrange(desc(TotalPop))  
  
interactive\_plot <- ggplot(  
 pa\_population,  
 aes(  
 x = reorder(PA, TotalPop),  
 y = TotalPop,  
 fill = TotalPop,  
 tooltip = paste("Planning Area:", PA, "\nPopulation:", comma(TotalPop)),  
 data\_id = PA  
 )  
) +  
 geom\_col\_interactive(  
 aes(),  
 css = "cursor:pointer;",  
 hover\_css = "opacity:1;stroke:black;stroke-width:1;",  
 hover\_inv\_css = "opacity:0.3;") +  
 scale\_fill\_gradient(high = "#c2185b", low = "#f8bbd0", labels = comma) +  
 scale\_y\_continuous(labels = comma) +  
 coord\_flip() +  
 labs(  
 title = "Total Population by Planning Area (2024)",  
 x = "Planning Area",  
 y = "Population"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA),  
 plot.title = element\_text(hjust = 0.5)  
 )  
  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#fef3c7;  
 color:#A94442;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow: 0px 2px 5px rgba(0,0,0,0.2);  
 opacity:0.3;"  
 )  
 )  
)

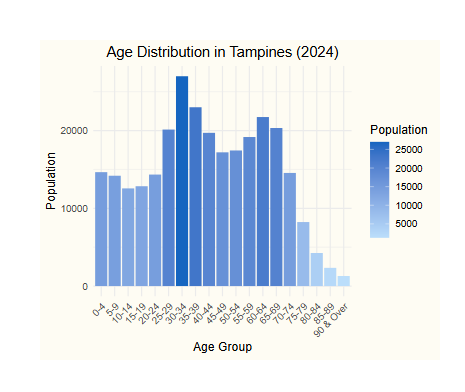


|  |  |
| --- | --- |
|  | **Insights**  Based on Singapore’s 2024 population data, **Tampines**, **Bedok**, and **Sengkang** emerge as the top three most populous planning areas. Together with **Jurong West** and **Woodlands**, these regions comprise a substantial portion of the national population. In contrast, central areas such as **Downtown Core** and **Museum** have significantly smaller populations, reflecting their primary use as commercial or institutional zones rather than residential neighborhoods. |

## Age Distribution in the Top 5 Planning Areas (2024)

### Tampines

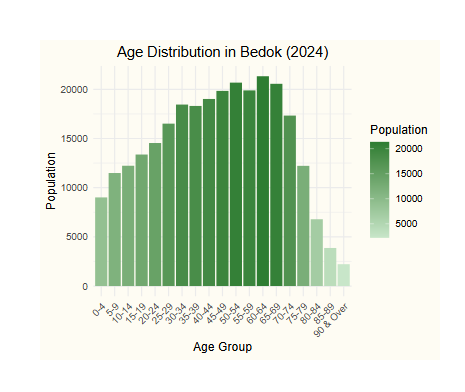
top5\_pa <- data %>%  
 group\_by(PA) %>%  
 summarise(TotalPop = sum(Pop, na.rm = TRUE)) %>%  
 slice\_max(order\_by = TotalPop, n = 5) %>%  
 pull(PA)  
  
  
top5\_data <- data %>%  
 filter(PA %in% top5\_pa) %>%  
 group\_by(PA, AgeGroup) %>%  
 summarise(Pop = sum(Pop, na.rm = TRUE), .groups = "drop")  
  
top5\_data$AgeGroup <- factor(top5\_data$AgeGroup, levels = c(  
 "0 – 4", "5 – 9", "10 – 14", "15 – 19", "20 – 24",  
 "25 – 29", "30 – 34", "35 – 39", "40 – 44", "45 – 49",  
 "50 – 54", "55 – 59", "60 – 64", "65 – 69", "70 – 74",  
 "75 – 79", "80 – 84", "85 – 89", "90 & Over"  
))  
  
interactive\_plot <- ggplot(  
 filter(top5\_data, PA == top5\_pa[1]),  
 aes(  
 x = AgeGroup,  
 y = Pop,  
 fill = Pop,  
 tooltip = paste("Age Group: ", AgeGroup, "\nPopulation: ", comma(Pop)),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(  
 aes(),  
 css = "cursor:pointer;",  
 hover\_css = "opacity:1;stroke:black;stroke-width:1;",  
 hover\_inv\_css = "opacity:0.3;"  
 ) +  
 scale\_fill\_gradient(high = "#1565c0", low = "#bbdefb") +  
 labs(  
 x = "Age Group",  
 y = "Population",  
 fill = "Population",  
 title = paste("Age Distribution in", top5\_pa[1], "(2024)")  
 ) +  
 theme\_minimal() +  
 scale\_x\_discrete(  
 labels = function(x) gsub(" – ", "-", str\_wrap(x, width = 10))  
 ) +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 plot.title = element\_text(hjust = 0.5),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#fef3c7;  
 color:#A94442;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow: 0px 2px 5px rgba(0,0,0,0.2);  
 opacity:0.3;"  
 )  
 )  
)



|  |  |
| --- | --- |
|  | **Insights**  Singapore’s 2024 demographic data highlights a significant concentration of Tampines residents in the 30-34 age range. This pattern suggests that Tampines is a hub for individuals in career-building phases of life. Despite this concentration, the population remains fairly balanced across other age groups, indicating a well-rounded and sustainable community structure. |

### Bedok

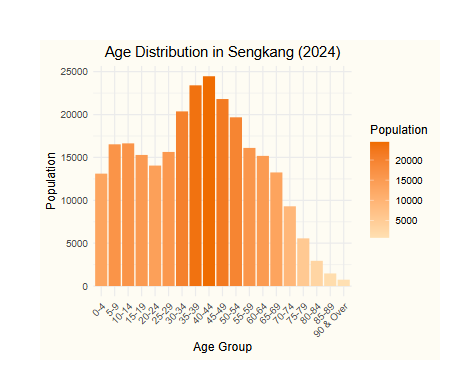
interactive\_plot <- ggplot(  
 filter(top5\_data, PA == top5\_pa[2]),  
 aes(  
 x = AgeGroup,  
 y = Pop,  
 fill = Pop,  
 tooltip = paste("Age Group: ", AgeGroup, "\nPopulation: ", comma(Pop)),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(  
 css = "cursor:pointer;",  
 hover\_css = "opacity:1;stroke:black;stroke-width:1.5;",  
 hover\_inv\_css = "opacity:0.3;"  
 ) +  
 scale\_fill\_gradient(high = "#2e7d32", low = "#c8e6c9") +  
 labs(  
 x = "Age Group",  
 y = "Population",  
 fill = "Population",  
 title = paste("Age Distribution in", top5\_pa[2], "(2024)")  
 ) +  
 theme\_minimal() +  
 scale\_x\_discrete(  
 labels = function(x) gsub(" – ", "-", str\_wrap(x, width = 10))  
 ) +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 plot.title = element\_text(hjust = 0.5),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#fef3c7;  
 color:#A94442;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow: 0px 2px 5px rgba(0,0,0,0.2);  
 opacity:0.3;"  
 )  
 )  
)



|  |  |
| --- | --- |
|  | **Insights**  Bedok exhibits a population structure heavily represented by residents aged 60-64. This mature demographic suggests Bedok is home more for aging residents. Compared to younger age groups, the higher population in older brackets may indicate that Bedok is a more settled community. |

### Sengkang

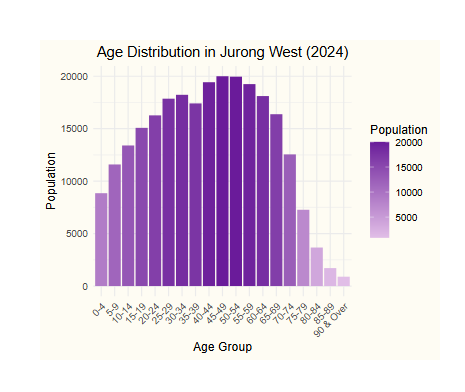
interactive\_plot <- ggplot(  
 filter(top5\_data, PA == top5\_pa[3]),  
 aes(  
 x = AgeGroup,  
 y = Pop,  
 fill = Pop,  
 tooltip = paste("Age Group:", AgeGroup, "\nPopulation:", Pop),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(  
 css = "cursor:pointer;",  
 hover\_css = "stroke:black;stroke-width:1;",  
 hover\_inv\_css = "opacity:0.3;"  
 ) +  
 scale\_fill\_gradient(high = "#ef6c00", low = "#ffe0b2") +  
 labs(  
 x = "Age Group",  
 y = "Population",  
 fill = "Population",  
 title = paste("Age Distribution in", top5\_pa[3], "(2024)")  
 ) +  
 theme\_minimal() +  
 scale\_x\_discrete(labels = function(x) gsub(" – ", "-", str\_wrap(x, width = 10))) +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 plot.title = element\_text(hjust = 0.5),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#fff3e0;  
 color:#333;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow:0 2px 5px rgba(0,0,0,0.15);"  
 )  
 )  
)



|  |  |
| --- | --- |
|  | **Insights**  Sengkang’s population is heavily concentrated in the 35–49 age group, reflecting a younger family demographic. The notable presence of children and newborns aged 0–14 further suggests that Sengkang is a preferred residential area for young families. |

### Jurong West

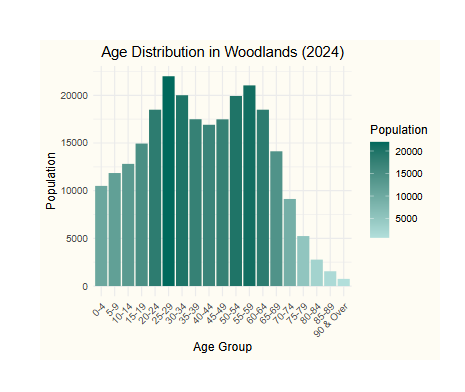
interactive\_plot <- ggplot(  
 filter(top5\_data, PA == top5\_pa[4]),  
 aes(  
 x = AgeGroup,  
 y = Pop,  
 fill = Pop,  
 tooltip = paste("Age Group:", AgeGroup, "\nPopulation:", Pop),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(  
 css = "cursor:pointer;",  
 hover\_css = "stroke:black;stroke-width:1;",  
 hover\_inv\_css = "opacity:0.3;"  
 ) +  
 scale\_fill\_gradient(high = "#6a1b9a", low = "#e1bee7") +  
 labs(  
 x = "Age Group",  
 y = "Population",  
 fill = "Population",  
 title = paste("Age Distribution in", top5\_pa[4], "(2024)")  
 ) +  
 theme\_minimal() +  
 scale\_x\_discrete(labels = function(x) gsub(" – ", "-", str\_wrap(x, width = 10))) +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 plot.title = element\_text(hjust = 0.5),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#f3e5f5;  
 color:#333;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow:0 2px 5px rgba(0,0,0,0.15);"  
 )  
 )  
)



|  |  |
| --- | --- |
|  | **Insights**  Jurong West presents a more balanced age distribution, with the largest population concentrated in the 35–59 age group.This suggests the area is home to many established families and individuals nearing retirement age. The balanced spread across other age groups also reflects a stable residential base, demonstrating Jurong West’s role as a settled community in Singapore. |

### Woodlands

interactive\_plot <- ggplot(  
 filter(top5\_data, PA == top5\_pa[5]),  
 aes(  
 x = AgeGroup,  
 y = Pop,  
 fill = Pop,  
 tooltip = paste("Age Group:", AgeGroup, "\nPopulation:", Pop),  
 data\_id = AgeGroup  
 )  
) +  
 geom\_col\_interactive(  
 css = "cursor:pointer;",  
 hover\_css = "stroke:black;stroke-width:1;",  
 hover\_inv\_css = "opacity:0.3;"  
 ) +  
 scale\_fill\_gradient(high = "#00695c", low = "#b2dfdb") +  
 labs(  
 x = "Age Group",  
 y = "Population",  
 fill = "Population",  
 title = paste("Age Distribution in", top5\_pa[5], "(2024)")  
 ) +  
 theme\_minimal() +  
 scale\_x\_discrete(labels = function(x) gsub(" – ", "-", str\_wrap(x, width = 10))) +  
 theme(  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 plot.title = element\_text(hjust = 0.5),  
 plot.background = element\_rect(fill = "#FEFCF3", color = NA)  
 )  
  
girafe(  
 ggobj = interactive\_plot,  
 options = list(  
 opts\_hover(css = "fill:#DBA39A;"),  
 opts\_hover\_inv(css = "opacity:0.3;"),  
 opts\_tooltip(  
 css = "background-color:#e0f2f1;  
 color:#004d40;  
 padding:8px;  
 border-radius:8px;  
 font-weight:bold;  
 box-shadow: 0px 2px 5px rgba(0,0,0,0.2);"  
 )  
 )  
)



|  |  |
| --- | --- |
|  | **Insights**  Woodlands displays a diverse age distribution, with notable peaks in both younger and older working-age groups. The 25–29 and 55–59 age brackets hold the highest concentrations, indicating a mix of young adults entering the workforce and seasoned professionals nearing retirement. |

# Summary

The 2024 demographic analysis of Singapore reveals several noteworthy trends. The population is nearly evenly split between males (2,152,050) and females (2,041,480), with the 25–54 age range forming the bulk of the working-age population. Among all planning areas, Tampines, Bedok, and Sengkang emerged as the most populous, suggesting high residential density in suburban towns.

Each top planning area showcases unique demographic characteristics:

* Tampines has a large concentration of young working adults (25–39).
* Bedok reflects an aging population, particularly those aged 60–64.
* Sengkang exhibits a family-oriented profile with peaks in both parents (35–49) and children (0–14).
* Jurong West shows a balanced age distribution, supporting diverse community needs.
* Woodlands features dual peaks in both early-career and near-retirement age groups.

# References

Kam, Tin Seong. (2023). Visualising Distribution. *R for Visual Analytics.* <https://r4va.netlify.app>