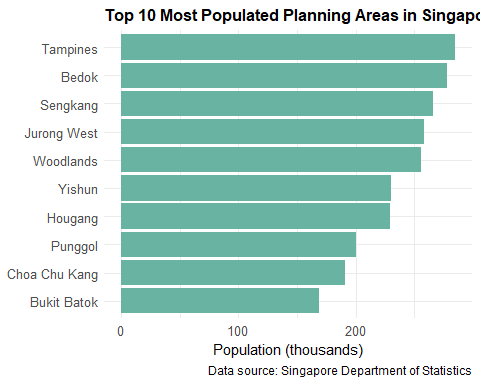
Take-Home Exercise 1: Singapore Demographic Patterns 2024

## Introduction

This project visualizes Singapore’s demographic structure and distribution in 2024, based on data from the Department of Statistics. The analysis focuses on population distribution across planning areas, age structures, and demographic patterns that reveal Singapore’s urban development and population characteristics.

## Planning Area Population Distribution

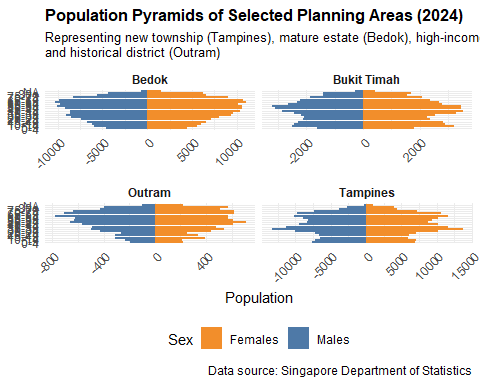
# Create planning area population bar chart  
top\_10\_pa <- pa\_summary %>%  
 top\_n(10, Total\_Pop)  
  
ggplot(top\_10\_pa, aes(x = reorder(PA, Total\_Pop), y = Total\_Pop/1000)) +  
 geom\_bar(stat = "identity", fill = "#69b3a2") +  
 coord\_flip() +  
 labs(  
 title = "Top 10 Most Populated Planning Areas in Singapore (2024)",  
 x = NULL,  
 y = "Population (thousands)",  
 caption = "Data source: Singapore Department of Statistics"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(face = "bold", size = 12),  
 axis.text = element\_text(size = 10)  
 )



Singapore’s population is heavily concentrated in a few key planning areas, with Tampines and Bedok housing the largest populations. This pattern reflects decades of urban development policy focused on creating high-density residential hubs in the east and northeast regions, connected by efficient transportation networks. These areas serve as complete townships with integrated amenities, explaining their sustained popularity despite newer developments elsewhere.

## Age Structure Comparison Across Planning Areas

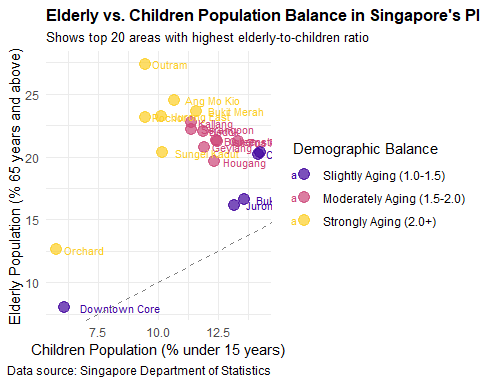
# Select representative planning areas for population pyramids  
selected\_areas <- c("Tampines", "Bedok", "Bukit Timah", "Outram")  
  
# Prepare population pyramid data  
pyramid\_data <- age\_bins %>%  
 filter(PA %in% selected\_areas) %>%  
 group\_by(PA, Age\_Group, Sex) %>%  
 summarise(Pop = sum(Pop)) %>%  
 mutate(  
 Pop = ifelse(Sex == "Males", -Pop, Pop),  
 Age\_Group = factor(Age\_Group, 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+"))  
 )  
  
# Create population pyramids  
ggplot(pyramid\_data, aes(x = Age\_Group, y = Pop, fill = Sex)) +  
 geom\_bar(stat = "identity") +  
 coord\_flip() +  
 facet\_wrap(~ PA, scales = "free\_x") +  
 scale\_fill\_manual(values = c("Males" = "#4E79A7", "Females" = "#F28E2B")) +  
 labs(  
 title = "Population Pyramids of Selected Planning Areas (2024)",  
 subtitle = "Representing new township (Tampines), mature estate (Bedok), high-income area (Bukit Timah),\nand historical district (Outram)",  
 x = NULL,  
 y = "Population",  
 caption = "Data source: Singapore Department of Statistics"  
 ) +  
 theme\_minimal() +  
 theme(  
 legend.position = "bottom",  
 axis.text.x = element\_text(angle = 45, hjust = 1),  
 strip.text = element\_text(face = "bold"),  
 plot.title = element\_text(face = "bold", size = 12),  
 plot.subtitle = element\_text(size = 10)  
 )



The population pyramids reveal distinct demographic profiles across four strategically selected planning areas. Tampines, representing newer townships, shows a balanced structure with substantial working-age population. Mature estates like Bedok display aging populations with wider tops, reflecting original residents who have aged in place. Bukit Timah, a high-income private housing area, exhibits a unique “sandglass” pattern with fewer young adults, while Outram’s pyramid demonstrates a significant elderly population characteristic of Singapore’s historical urban core.

## Demographic Balance: Elderly vs. Children Distribution

# Combine elderly and children data  
demographic\_balance <- elderly\_ratio %>%  
 select(PA, Elderly\_Ratio) %>%  
 left\_join(  
 children\_ratio %>% select(PA, Child\_Ratio),  
 by = "PA"  
 ) %>%  
 mutate(  
 Demographic\_Ratio = Elderly\_Ratio / Child\_Ratio,  
 Balance\_Category = case\_when(  
 Demographic\_Ratio >= 2.0 ~ "Strongly Aging (2.0+)",  
 Demographic\_Ratio >= 1.5 ~ "Moderately Aging (1.5-2.0)",  
 Demographic\_Ratio >= 1.0 ~ "Slightly Aging (1.0-1.5)",  
 Demographic\_Ratio >= 0.67 ~ "Balanced (0.67-1.0)",  
 TRUE ~ "Youth-dominant (<0.67)"  
 ),  
 Balance\_Category = factor(Balance\_Category,   
 levels = c("Youth-dominant (<0.67)",   
 "Balanced (0.67-1.0)",   
 "Slightly Aging (1.0-1.5)",   
 "Moderately Aging (1.5-2.0)",   
 "Strongly Aging (2.0+)"))  
 ) %>%  
 arrange(desc(Demographic\_Ratio))  
  
# Select top 20 areas for analysis  
top\_20\_balance <- demographic\_balance %>%  
 head(20)  
  
# Create scatter plot  
ggplot(top\_20\_balance, aes(x = Child\_Ratio, y = Elderly\_Ratio, color = Balance\_Category)) +  
 geom\_point(size = 4, alpha = 0.7) +  
 geom\_text(aes(label = PA), hjust = -0.2, vjust = 0.5, size = 3) +  
 geom\_abline(slope = 1, linetype = "dashed", color = "gray50") +  
 scale\_color\_viridis\_d(option = "plasma", begin = 0.1, end = 0.9) +  
 labs(  
 title = "Elderly vs. Children Population Balance in Singapore's Planning Areas (2024)",  
 subtitle = "Shows top 20 areas with highest elderly-to-children ratio",  
 x = "Children Population (% under 15 years)",  
 y = "Elderly Population (% 65 years and above)",  
 color = "Demographic Balance",  
 caption = "Data source: Singapore Department of Statistics"  
 ) +  
 theme\_minimal() +  
 theme(  
 legend.position = "right",  
 plot.title = element\_text(face = "bold", size = 12),  
 plot.subtitle = element\_text(size = 10)  
 )



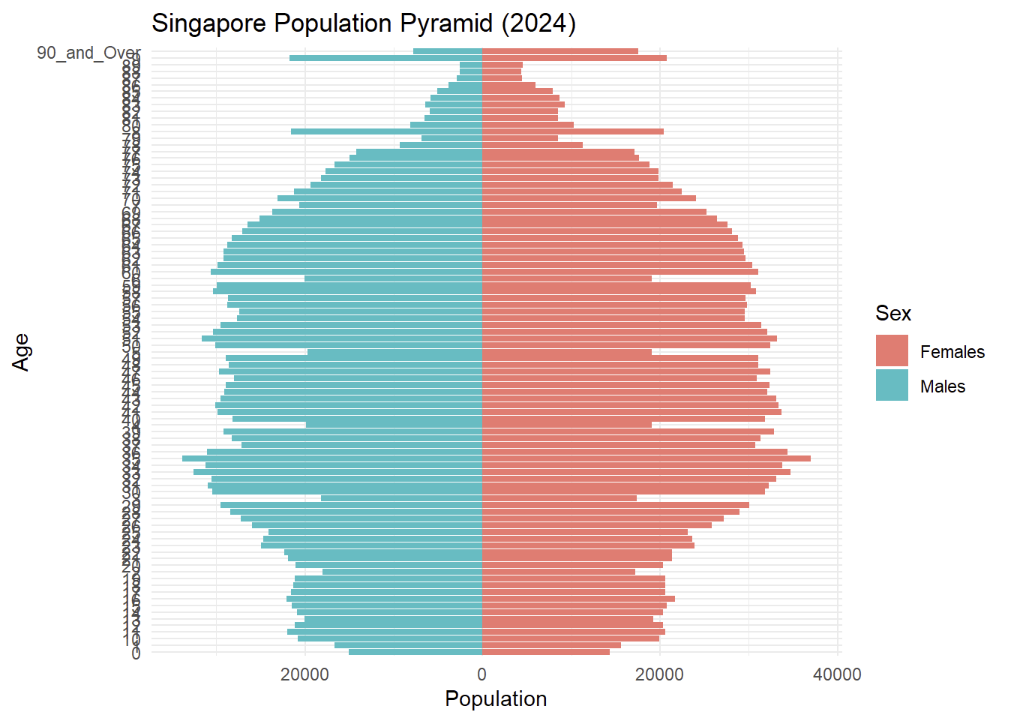
This innovative visualization reveals Singapore’s demographic transition by comparing elderly and children populations across planning areas. Areas above the diagonal line have more elderly than children, with Outram and Bukit Merah showing extreme imbalances (over 3:1 ratio). This pattern reflects Singapore’s rapid aging, particularly in first-generation public housing estates where infrastructure designed for families now serves an elderly majority. These imbalances create challenges for social services, healthcare accessibility, and school planning, requiring targeted policy responses to address these distinct neighborhood-level demographic profiles.

## Conclusion

The visualizations reveal Singapore’s complex demographic landscape in 2024, characterized by distinct spatial patterns in population distribution and age structure. Established planning areas in the east remain heavily populated, while pronounced demographic differences exist between mature and newer estates. These patterns have significant implications for urban planning priorities including: (1) targeted healthcare infrastructure development in aging neighborhoods like Outram and Bukit Merah; (2) transit connectivity improvements to address population concentration in eastern regions; and (3) housing policy adaptations to accommodate evolving household structures across different planning areas.

## Peer Review – Take-home Exercise 1 (Rajesh)

### Original visualisation



Original chart by Rajesh

### Strengths

1. **Clear narrative** – three charts move logically from overall age structure to spatial distribution and gender ratio.
2. **Chart–data fit** – population pyramid, ranked bar chart and gender-ratio plot suit their respective messages.
3. **Self-contained labels** – titles, subtitles and axes allow non-experts to grasp each visual quickly.

### Areas for improvement

1. **Colour consistency** – each chart uses a different palette, breaking visual unity.
2. **Crowded pyramid labels** – annual age ticks overlap; grouping by five-year bands would improve readability.
3. **Gender-ratio baseline** – adding a reference line at M∕F = 1 lets readers spot imbalanced areas instantly.

### Make-over (gender-ratio chart)

library(tidyverse)  
library(ggthemes)  
  
gender\_plot <- demographic\_balance %>%   
 ggplot(aes(x = reorder(PA, Demographic\_Ratio),   
 y = Demographic\_Ratio,   
 fill = Demographic\_Ratio > 1)) +   
 geom\_col(show.legend = FALSE) +   
 geom\_hline(yintercept = 1, linetype = "dashed") +   
 scale\_fill\_manual(values = c("#4E79A7", "#E15759")) +   
 labs(title = "Gender ratio by planning area (M/F)",   
 subtitle = "Bars above the dashed line indicate more males than females",   
 x = NULL,   
 y = "Male : Female") +   
 coord\_flip() +   
 theme\_economist() +  
 theme(  
 axis.text.y = element\_text(size = 7),  
 plot.margin = margin(t = 5, r = 5, b = 5, l = 20, unit = "pt"),  
 text = element\_text(size = 9)  
 )  
  
gender\_plot

