

Banking Sector Report

1. Constructing Long Format Dataset

```
data_2021 <- read_excel("Data_LB.xlsx", sheet = "2021Q3", skip = 1) %>% mutate(date = as.Date("2021-10-01"))

## New names:
## * ' ' -> '...14'
## * ' ' -> '...15'
## * ' ' -> '...16'
## * ' ' -> '...17'
```

```
data_2022 <- read_excel("Data_LB.xlsx", sheet = "2022Q3", skip = 1) %>% mutate(date = as.Date("2022-10-01"))

## New names:
## * ' ' -> '...15'
## * ' ' -> '...16'
## * ' ' -> '...17'
```

```
data_2023 <- read_excel("Data_LB.xlsx", sheet = "2023Q3", skip = 1) %>% mutate(date = as.Date("2023-10-01"))

combined_data <- bind_rows(data_2021, data_2022, data_2023)

long_format_data <- combined_data %>%
  pivot_longer(
    cols = -c(ID, date),
    names_to = "bank.name",
    values_to = "value"
  )

head(long_format_data)
```

```
## # A tibble: 6 x 4
##   ID          date    bank.name  value
##   <chr>      <date>    <chr>    <dbl>
## 1 Debt securities 2021-10-01 SWE      1248468
## 2 Debt securities 2021-10-01 SEB       479647
## 3 Debt securities 2021-10-01 SIA       766518
## 4 Debt securities 2021-10-01 REV         88
## 5 Debt securities 2021-10-01 MED       57647
## 6 Debt securities 2021-10-01 PRB         NA
```

2. Calculate Total Banking Sector Assets

```
total_sector_assets <- long_format_data %>%
  filter(ID == "Total assets") %>%
  group_by(date) %>%
  summarise(total_assets = sum(value, na.rm = TRUE))

# Display the table
knitr::kable(total_sector_assets, caption = "Total Assets for Each Year", booktabs = TRUE) %>%
  kable_styling(bootstrap_options = c("striped", "hover"))
```

Table 1: Total Assets for Each Year

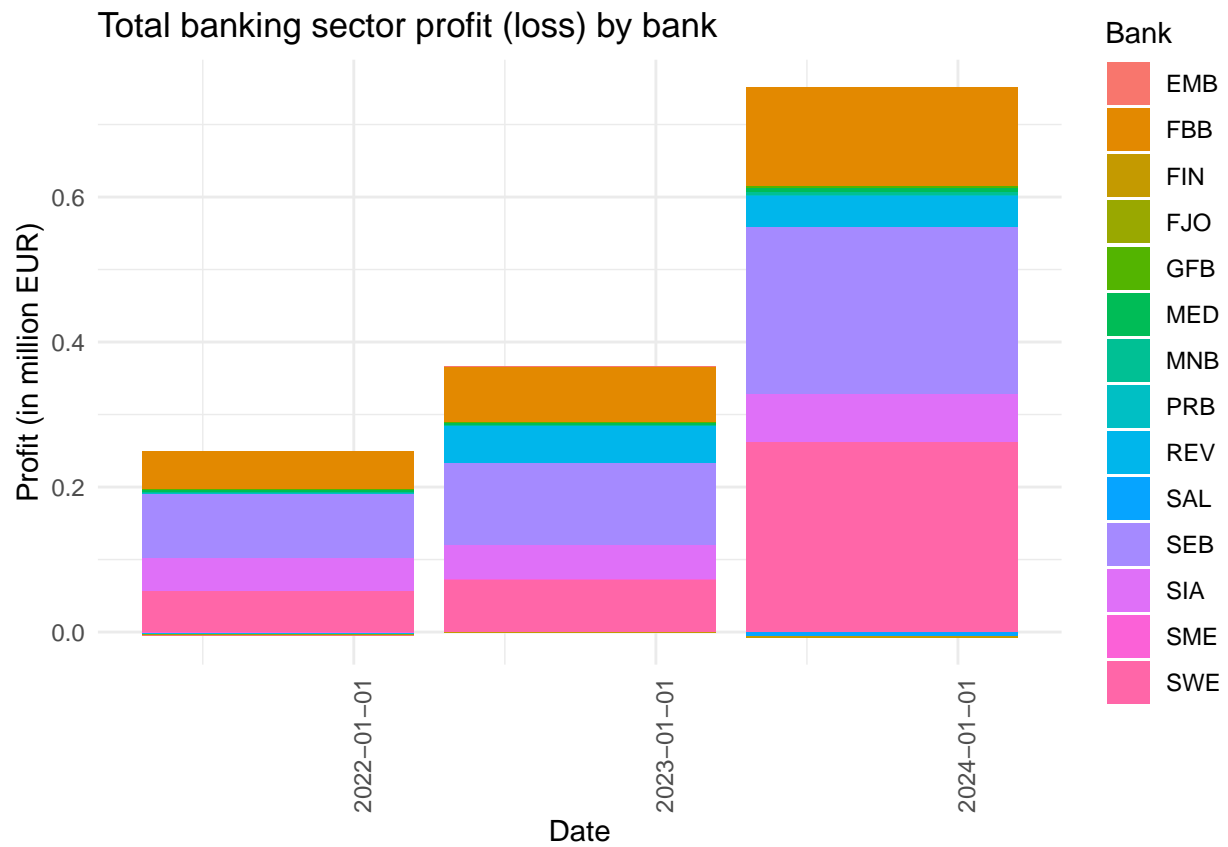
date	total_assets
2021-10-01	40546824
2022-10-01	51441948
2023-10-01	56908007

3. Stacked Column Chart of Bank Profit

```
profit_data <- long_format_data %>%
  filter(ID == "Profit (loss) of the current year" & !is.na(value))

# Create the plot
p <- ggplot(profit_data, aes(x = date, y = value/1e6, fill = bank.name)) +
  geom_col() +
  scale_y_continuous(labels = scales::comma) +
  scale_x_date(date_labels = "%Y-%m-%d", date_breaks = "1 year") +
  labs(
    x = "Date",
    y = "Profit (in million EUR)",
    title = "Total banking sector profit (loss) by bank",
    fill = "Bank"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))

print(p)
```



4. Commentary on the Results from part 2 and 3

The graph showing the banking sector's profit by each bank tells us that the profits are shared across the board, with some banks standing out more in certain years. This diversity in profit distribution could mean a variety of strategies or market responses at play.

Table 1 shows a steady climb in total assets from 2021 to 2023, indicating the banks are growing stronger financially. It's not just about the profits they make each year; it also reflects their ability to grow their resources and investments over time.

In short, the banking sector is not only making more money but also building up its financial foundation, which is a good sign for stability and future growth. It's like watching a team not only win games but also strengthen their lineup for seasons to come.

5. Subset Data for Specific Banks

```
long_format_data <- long_format_data %>%  
  mutate(value = as.numeric(gsub(",", "", value)))  
  
print(sum(is.na(long_format_data$value)))
```

```
## [1] 528
```

```
print(unique(long_format_data$bank.name))
```

```
## [1] "SWE" "SEB" "SIA" "REV" "MED" "PRB" "MNB" "GFB" "EMB"  
## [10] "SME" "FJO" "FBB" "...14" "...15" "...16" "...17" "FIN" "SAL"
```

```
specific_banks <- c("SWE", "SEB", "REV", "MED", "SIA", "MNB")
```

```
specific_banks_data <- long_format_data %>%  
  filter(bank.name %in% specific_banks, ID == "Total assets") %>%  
  mutate(value = as.numeric(value))
```

```
specific_banks_total_assets <- specific_banks_data %>%  
  group_by(date) %>%  
  summarise(total_assets_specific = sum(value, na.rm = TRUE))
```

```
total_assets_sector <- long_format_data %>%  
  filter(ID == "Total assets") %>%  
  group_by(date) %>%  
  summarise(total_assets = sum(value, na.rm = TRUE))
```

```
specific_banks_share <- specific_banks_total_assets %>%  
  left_join(total_assets_sector, by = "date") %>%  
  mutate(share_of_sector = total_assets_specific / total_assets * 100)
```

```
knitr::kable(specific_banks_share, caption = "Total Assets and Share of Specific Banks", booktabs = TRUE,
  kable_styling(bootstrap_options = c("striped", "hover")))
```

Table 2: Total Assets and Share of Specific Banks

date	total_assets_specific	total_assets	share_of_sector
2021-10-01	30827979	40546824	76.03056
2022-10-01	40855273	51441948	79.42015
2023-10-01	45475969	56908007	79.91137

6. Plot Assets by Bank and Year

```
specific_banks <- c("SWE", "SEB", "REV", "MED", "SIA", "MNB")

specific_banks_data <- long_format_data %>%
  filter(bank.name %in% specific_banks, ID == "Total assets") %>%
  mutate(value = as.numeric(value))

specific_banks_total_assets <- specific_banks_data %>%
  group_by(date) %>%
  summarise(total_assets_specific = sum(value, na.rm = TRUE))

total_assets_sector <- long_format_data %>%
  filter(ID == "Total assets") %>%
  group_by(date) %>%
  summarise(total_assets = sum(value, na.rm = TRUE))

specific_banks_share <- specific_banks_total_assets %>%
  left_join(total_assets_sector, by = "date") %>%
  mutate(share_of_sector = total_assets_specific / total_assets * 100)

assets_plot_data <- specific_banks_data %>%
  mutate(year = year(date))

assets_plot <- ggplot(assets_plot_data, aes(x = bank.name, y = value / 1e6, fill = as.factor(year))) +
  geom_col(position = position_dodge(width = 0.7)) +
  scale_fill_manual(values = c("2021" = "green", "2022" = "purple", "2023" = "orange")) +
  labs(x = "Bank", y = "Bank assets by year, mln. Eur", title = "Bank Assets by Year", fill = "Year") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))
print(assets_plot)
```

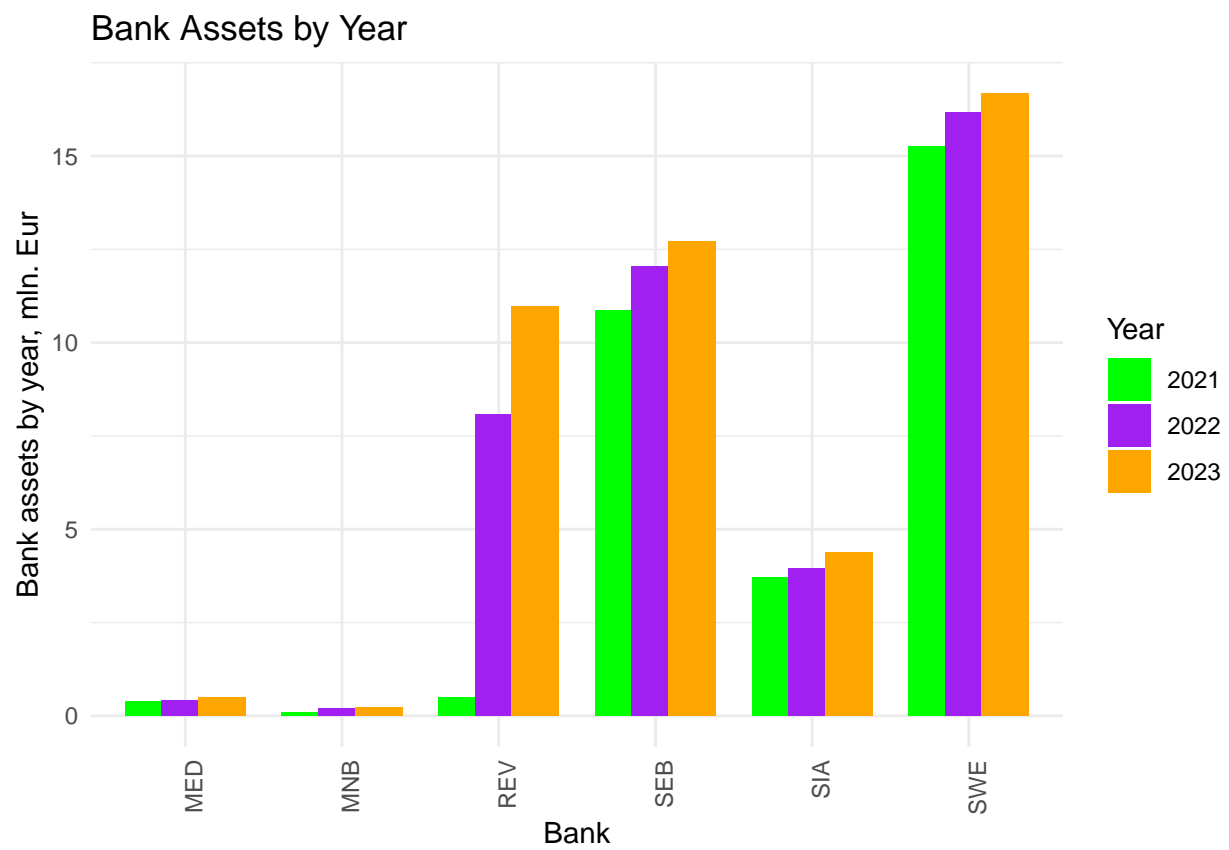


Figure 1: Assets by Bank and Year

7. Plot Profit (Loss) by Year

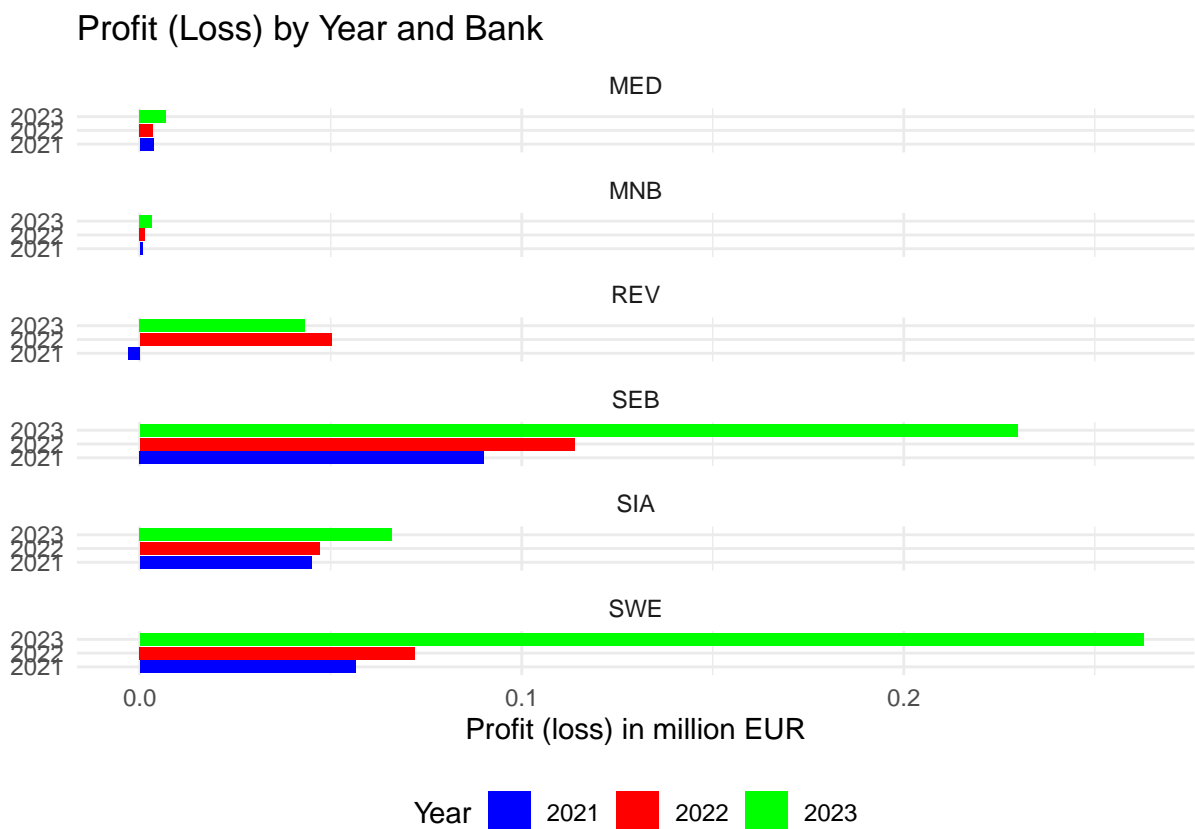
```
long_format_data <- long_format_data %>%
  mutate(value = as.numeric(gsub(",", "", value)))

specific_banks <- c("SWE", "SEB", "REV", "MED", "SIA", "MNB")
profit_data <- long_format_data %>%
  filter(bank.name %in% specific_banks, ID == "Profit (loss) of the current year") %>%
  mutate(year = year(date))

profit_plot <- ggplot(profit_data, aes(x = factor(year), y = value/1e6, fill = factor(year))) +
  geom_col(position = position_dodge()) +
  coord_flip() +
  facet_wrap(~bank.name, scales = "free_y", ncol = 1) +
  scale_fill_manual(values = c("2021" = "blue", "2022" = "red", "2023" = "green")) +
  labs(x = "", y = "Profit (loss) in million EUR", title = "Profit (Loss) by Year and Bank", fill = "Year") +
  theme_minimal() +
  theme(legend.position = "bottom")

ggsave("profit_by_year_and_bank.pdf", profit_plot, width = 11, height = 8.5)

print(profit_plot)
```



8. Ratio

```
long_format_data <- long_format_data %>%
  mutate(value = as.numeric(gsub(",", "", value)))

specific_banks <- c("SWE", "SEB", "REV", "MED", "SIA", "MNB")

specific_banks_data <- long_format_data %>%
  filter(bank.name %in% specific_banks)

specific_banks_ratios <- specific_banks_data %>%
  group_by(date, bank.name) %>%
  mutate(
    total_assets = ifelse(ID == "Total assets", value, NA)
  ) %>%
  mutate(
    Cash_to_Total_ratio = sum(ifelse(ID == "Cash balances with central banks", value, 0), na.rm = TRUE) /
    total_assets,
    Loans_to_Total_ratio = sum(ifelse(ID == "Loans and advances (including leasing)", value, 0), na.rm = TRUE) /
    total_assets
  ) %>%
  select(-value)

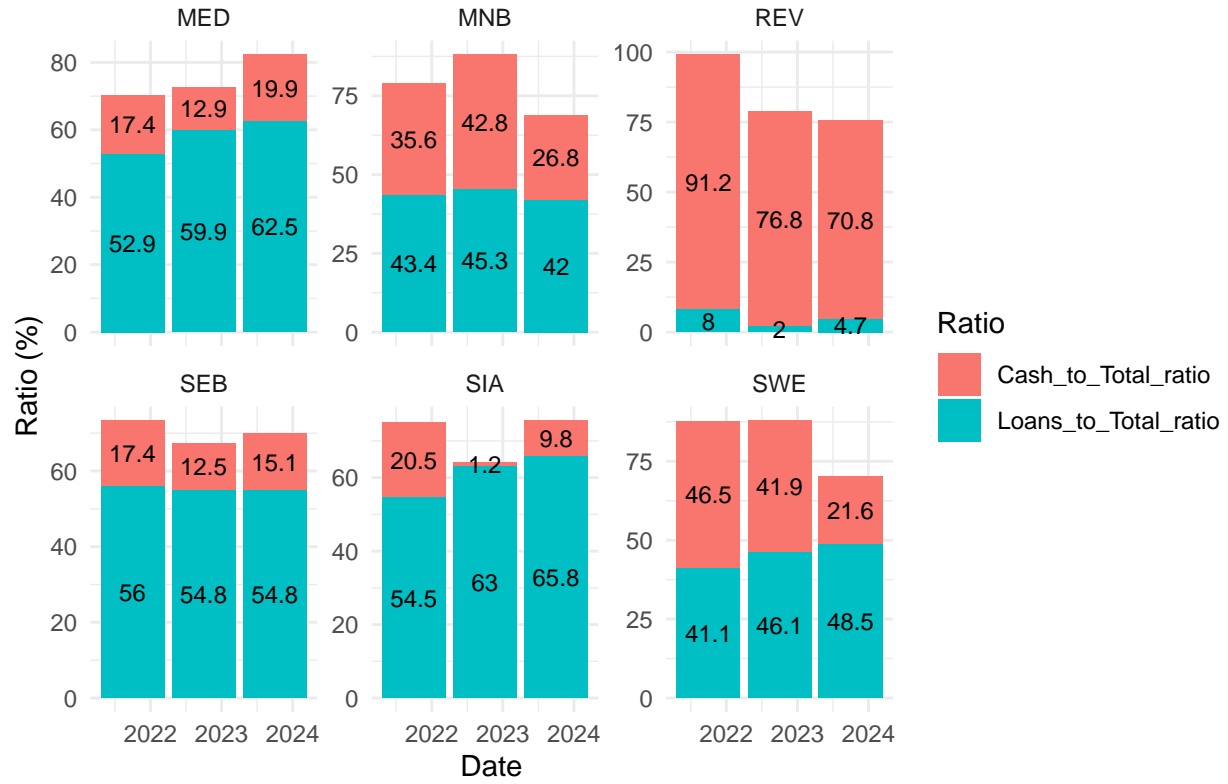
specific_banks_ratios_long <- specific_banks_ratios %>%
  pivot_longer(cols = c(Cash_to_Total_ratio, Loans_to_Total_ratio), names_to = "ratio", values_to = "value")

ratio_plot <- ggplot(specific_banks_ratios_long, aes(x = date, y = value, fill = ratio)) +
  geom_col() +
  facet_wrap(~bank.name, scales = "free_y", ncol = 3) +
  geom_text(aes(label = paste0(round(value, 1))), position = position_stack(vjust = 0.5), size = 3) +
  labs(x = "Date", y = "Ratio (%)", fill = "Ratio", title = "Cash and Loans to Total Assets Ratios by Bank") +
  theme_minimal()
print(ratio_plot)
```

```
## Warning: Removed 768 rows containing missing values or values outside the scale range
## ('geom_col()').
```

```
## Warning: Removed 768 rows containing missing values or values outside the scale range
## ('geom_text()').
```


Cash and Loans to Total Assets Ratios by Bank



9. Commentary on the Results from part 6,7 and 8

from Part 6 Plot , it is noticable that for all banks that they were consistnet over all three years assets expect for the Revolut Bank, UAB. it had a notiacble increase in its assests starting from 2022, almost an increase by 800% and it also kept increasing the folloing year.

from part 7 plot , it makes complete sense that the profit(loss) in millions matches alot the number of assets each bank has , implying a direct relation between number of assets and Profit(loss) in million for each bank bascially

from part 8 plots , we seem to have three groups , the first is those banks whose loans to total ratio is very far ahead compared to its Cash to total ratio , these banks are (UAB Medicinos bankas, B Siauliu bankas, AB SEB bankas) , the second group has a balanced percentage between Cash to total and loans to total ratios , these are (Swedbank, AB and AB "Mano bankas") and the last group is the bank that always seems to be out of formation compared to other banke ,which is Revolut Bank, UAB , its Cash to total ratio is very high compared to its Loan To total ratio , which might explain why it had low assets at the 2021 year and low profit(loss) and then it started to grow eventually .

10. Scatter plot

```
long_format_data <- long_format_data %>%
  mutate(Bank_group = ifelse(bank.name %in% specific_banks, "Selected banks", "Other banks"))

filtered_data <- long_format_data %>%
  filter(ID %in% c("Total assets", "Total equity", "Profit (loss) of the current year"))

wide_data <- filtered_data %>%
  pivot_wider(names_from = ID, values_from = value)

wide_data <- wide_data %>%
  mutate(
    Profit_to_equity_ratio = `Profit (loss) of the current year` / `Total equity` * 100,
    Profit_to_asset_ratio = `Profit (loss) of the current year` / `Total assets` * 100
  )

wide_data_2021 <- wide_data %>% filter(year(date) == 2021)
wide_data_2022 <- wide_data %>% filter(year(date) == 2022)
wide_data_2023 <- wide_data %>% filter(year(date) == 2023)

scatter_plot_2021 <- ggplot(wide_data_2021, aes(x = Profit_to_equity_ratio, y = Profit_to_asset_ratio,
  geom_point() +
  labs(x = "Profit to Equity Ratio (%)", y = "Profit to Asset Ratio (%)", color = "Bank Group", title =
  theme_minimal()

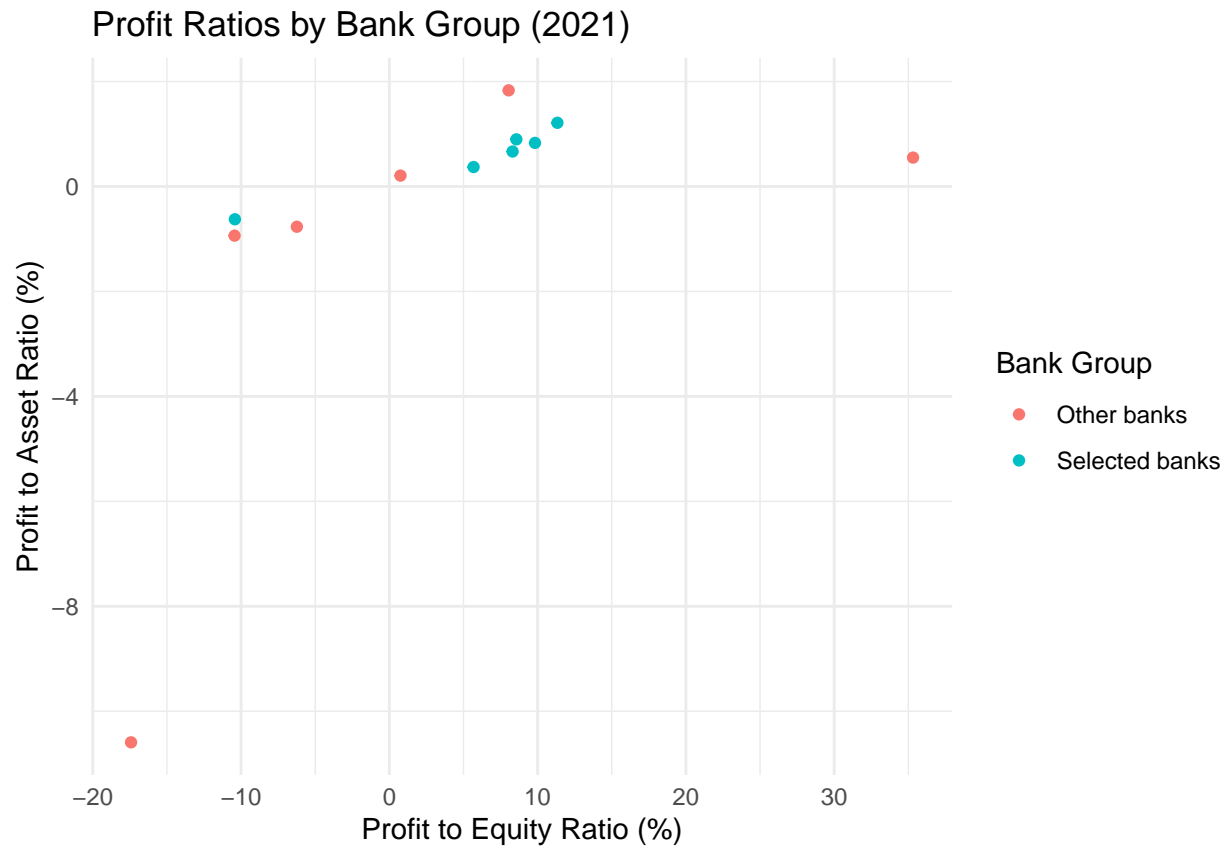
scatter_plot_2022 <- ggplot(wide_data_2022, aes(x = Profit_to_equity_ratio, y = Profit_to_asset_ratio,
  geom_point() +
  labs(x = "Profit to Equity Ratio (%)", y = "Profit to Asset Ratio (%)", color = "Bank Group", title =
  theme_minimal()

scatter_plot_2023 <- ggplot(wide_data_2023, aes(x = Profit_to_equity_ratio, y = Profit_to_asset_ratio,
  geom_point() +
  labs(x = "Profit to Equity Ratio (%)", y = "Profit to Asset Ratio (%)", color = "Bank Group", title =
```

```
theme_minimal()

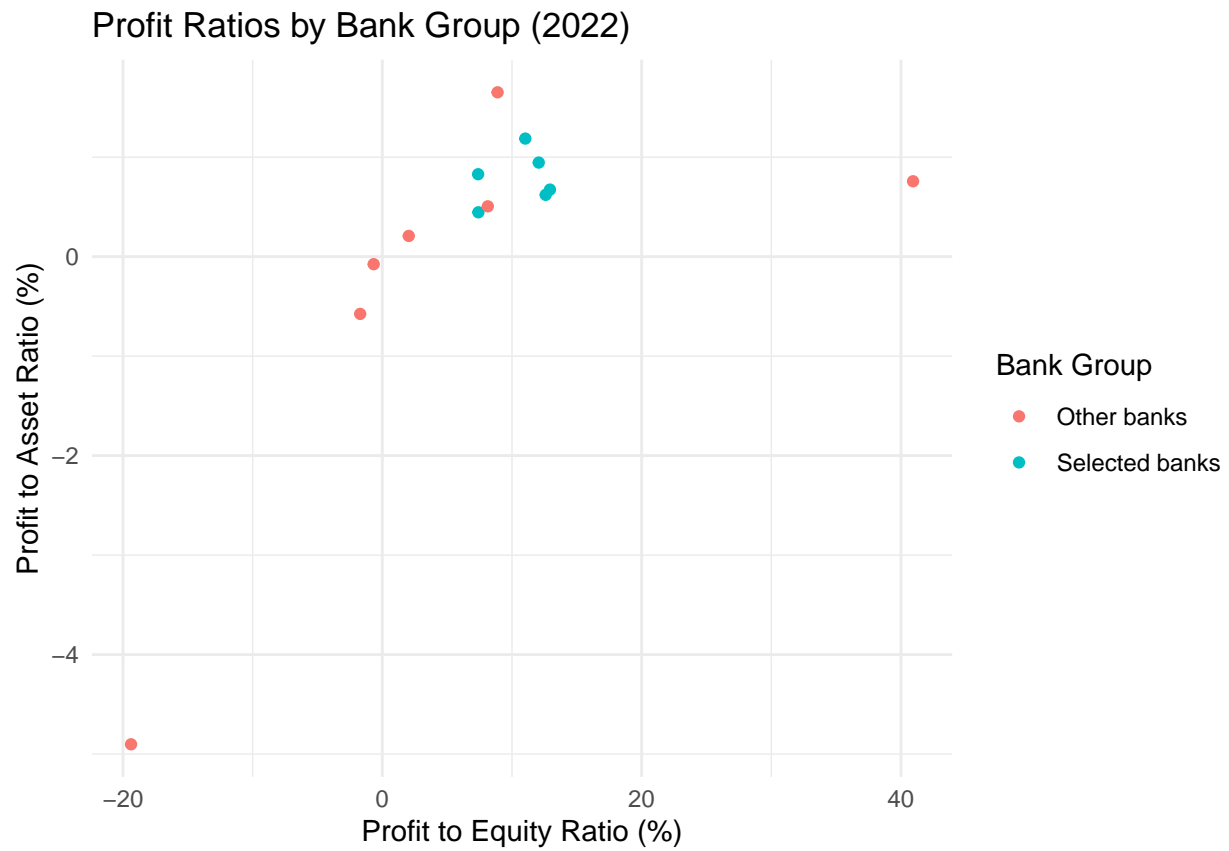
print(scatter_plot_2021)
```

```
## Warning: Removed 6 rows containing missing values or values outside the scale range
## ('geom_point()').
```



```
print(scatter_plot_2022)
```

```
## Warning: Removed 5 rows containing missing values or values outside the scale range
## ('geom_point()').
```



```
print(scatter_plot_2023)
```

```
## Warning: Removed 4 rows containing missing values or values outside the scale range  
## ('geom_point()').
```

