Berlin Firms Comparison - Table

knitr::opts\_chunk$set(echo = TRUE, message = FALSE, warning = FALSE)  
  
# Load required libraries  
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.0 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(knitr)  
  
# Load data  
df <- readRDS("/workspaces/rct25/data/generated/Orbis\_Berlin\_Data/orbis\_panel\_berlin.rds")  
  
stopifnot(is.data.frame(df))  
  
# Clean and prepare data  
df <- df %>%  
 filter(city\_native == "Berlin") %>%  
 mutate(  
 total\_assets = toas,  
 equity\_ratio = shfd / toas,  
 log\_total\_assets = log1p(toas),  
 group = if\_else(postcode == 10437, "Postal Code 10437", "Other Berlin Firms")  
 ) %>%  
 filter(!is.na(total\_assets), !is.na(equity\_ratio))  
  
n\_total\_original <- nrow(df)  
df <- df %>% filter(total\_assets >= 1000)  
n\_removed\_assets <- n\_total\_original - nrow(df)  
  
# Summary stats  
summary\_stats <- df %>%  
 group\_by(group) %>%  
 summarise(  
 Mean\_Total\_Assets = mean(total\_assets, na.rm = TRUE),  
 SD\_Total\_Assets = sd(total\_assets, na.rm = TRUE),  
 Mean\_Equity\_Ratio = mean(equity\_ratio, na.rm = TRUE),  
 Median\_Equity\_Ratio = median(equity\_ratio, na.rm = TRUE),  
 SD\_Equity\_Ratio = sd(equity\_ratio, na.rm = TRUE),  
 n = n(),  
 .groups = "drop"  
 )  
  
get\_stat <- function(var, group\_name) summary\_stats %>% filter(group == group\_name) %>% pull({{ var }})  
  
# Extract values  
mean\_total\_assets\_10437 <- get\_stat(Mean\_Total\_Assets, "Postal Code 10437")  
mean\_total\_assets\_other <- get\_stat(Mean\_Total\_Assets, "Other Berlin Firms")  
mean\_equity\_ratio\_10437 <- get\_stat(Mean\_Equity\_Ratio, "Postal Code 10437")  
mean\_equity\_ratio\_other <- get\_stat(Mean\_Equity\_Ratio, "Other Berlin Firms")  
median\_equity\_ratio\_10437 <- get\_stat(Median\_Equity\_Ratio, "Postal Code 10437")  
median\_equity\_ratio\_other <- get\_stat(Median\_Equity\_Ratio, "Other Berlin Firms")  
sd\_equity\_ratio\_10437 <- get\_stat(SD\_Equity\_Ratio, "Postal Code 10437")  
sd\_equity\_ratio\_other <- get\_stat(SD\_Equity\_Ratio, "Other Berlin Firms")  
sd\_total\_assets\_10437 <- get\_stat(SD\_Total\_Assets, "Postal Code 10437")  
sd\_total\_assets\_other <- get\_stat(SD\_Total\_Assets, "Other Berlin Firms")  
n\_10437 <- get\_stat(n, "Postal Code 10437")  
n\_other <- get\_stat(n, "Other Berlin Firms")  
  
# Statistical tests  
t\_total\_assets <- t.test(total\_assets ~ group, data = df)  
t\_equity\_ratio <- t.test(equity\_ratio ~ group, data = df)  
t\_sd\_equity\_ratio <- t.test(df$equity\_ratio[df$group == "Postal Code 10437"], df$equity\_ratio[df$group != "Postal Code 10437"])  
t\_sd\_total\_assets <- t.test(df$total\_assets[df$group == "Postal Code 10437"], df$total\_assets[df$group != "Postal Code 10437"])  
t\_median\_equity\_ratio <- wilcox.test(equity\_ratio ~ group, data = df)

Comparison of Firms in Postal Code 10437 vs. Other Berlin Firms

| rowname | Postal Code 10437 | Other Berlin Firms | Difference | P-Value | Significance |
| --- | --- | --- | --- | --- | --- |
| Total Assets (Mean) | 2.704.204 | 18.458.902 | -15.754.698 | <0.001 | \*\*\* |
| Total Assets (SD) | 18.686.565 | 551.759.614 | -533.073.049 | <0.001 | \*\*\* |
| Equity Ratio (Mean) | -56,00% | -91,15% | 35,15% | 0,049 | \*\* |
| Equity Ratio (Median) | 36,77% | 32,80% | 3,96% | 0,044 | \*\* |
| Equity Ratio (SD) | 887,61% | 3.967,58% | -3.079,97% | 0,049 | \*\* |
| Number of Firms (N) | 2.842 | 385.328 | NA |  |  |