Research Log

Student Name: Unathi Zwane
Course: Research on Corporate Transparency
Assignment Title: Empirical Assignment 1
Lecturer: Prof. Dr. Joachim Gassen

Step 1 – Identifying Firms in My Postal Code (10557 Berlin)

Date: 2025-05-14
Activity: Identifying and verifying registered firms in my local area
Objective: Find 5 firms with a registered address in postal code 10557 Berlin and confirm their official name and address via legal disclosures ("Impressum").

Process & Sources Consulted

- Used Google Search and company websites to locate firms within the 10557 Berlin postal code.
- Verified company names and addresses by consulting the "Impressum" section on each firm's official website.
- Replaced one initially selected company (Lagardère Travel Retail Deutschland GmbH, located in Wiesbaden) with a valid alternative (Spryker Systems GmbH) that meets the location criteria.

Firms Identified

Company	Address and Source
KPMG AG Wirtschaftsprüfungsgesellschaft	Heidestraße 58, 10557 Berlin `https://kpmg.com/de/de/home/misc/unternehmensangaben.html`
50Hertz Transmission GmbH	Heidestraße 2, 10557 Berlin `https://csr.50hertz.com/de/Nachhaltigkeit/Impressum`
Europace AG	Heidestraße 8, 10557 Berlin `https://europace.de/impressum/`
Total Energies Holdings Deutschland GmbH	Jean-Monnet-Straße 2, 10557 Berlin `https://w3.windmesse.de/total`

Spryker Systems GmbH	Heidestraße 9-10, 10557 Berlin `https://spryker.com/imprint/`

Reflections

The Impressum sections on official company websites proved to be a reliable source for verifying the legal names and registered addresses of firms. In contrast, some company databases and third-party listings contained outdated or incomplete location information, which underscored the necessity of consulting the official disclosures for accuracy. The experience of having to replace an initially identified firm due to an incorrect location reinforced the importance of cross-verification and attention to detail when conducting company-level research.

Step 2 – Verifying Register Information and Financial Statements

Date: 2025-05-14
Activity: Accessed `www.unternehmensregister.de` to verify register information and financial statements for selected firms.
Objective: Obtain official firm names, register numbers (HRB), and details of available financial statements.

Firm-Level Register Information

1. KPMG AG Wirtschaftsprüfungsgesellschaft

- **Register Number:** HRB 106191
- **Register Court:** Amtsgericht Berlin-Charlottenburg
- **Financial Statements:** Consolidated (2007–2023)
- **Latest Filing:** 2024-10-24

2. 50Hertz Transmission GmbH

- **Register Number:** HRB 84446
- **Register Court:** Amtsgericht Berlin-Charlottenburg
- **Financial Statements:** Single legal entity (2013–2023)
- **Latest Filing:** 2024-05-21

3. Europace AG

- **Register Number:** HRB 136078
- **Register Court:** Amtsgericht Berlin-Charlottenburg
- **Financial Statements:** Consolidated (2015)
- **Latest Filing:** 2016-09-15
- This company is wholly owned by its parent company Hypoport SE, Lübeck. In most financial reporting periods, it has been exempt through company board resolution.

4. Total Energies Holdings Deutschland GmbH

- **Official Name:** Total Deutschland GmbH
- **Register Number:** HRB 80333
- **Register Court:** Amtsgericht Berlin-Charlottenburg
- **Financial Statements:** Consolidated (2021–2023)
- **Latest Filing:** 2024-10-17

5. Spryker Systems GmbH

- **Register Number:** HRB 134310
- **Register Court:** Amtsgericht Hamburg
- **Financial Statements:** Single legal entity (2016–2019)
- **Latest Filing:** 2021-06-22

Reflections

All five firms had their register information and financial statements available on the Unternehmensregister, ensuring transparency and accessibility. The inclusion of both consolidated and single-entity financial statements offers a comprehensive perspective on each firm's financial health, allowing for nuanced analysis of their respective corporate structures and financial positions. This level of detail reinforces the value of the Unternehmensregister as a reliable resource for corporate disclosure and research.

Step 3 – Prepare Orbis Data

Date: 2025-05-19
Activity: Setting up and reviewing Orbis data

Objective: Integrate Orbis data into the research environment and verify whether financial statements identified in Step 2 are included in the Orbis dataset.

Process

- Cloned course repository: `https://github.com/joachimgassen/rct25`
- Downloaded Orbis data file from Moodle and saved it in `data/generated/`
- Loaded data in RStudio
- Filtered for selected firms using company names and postcodes
- Compared with register data

Findings

Company	In Orbis?	Years	Notes
KPMG AG	Yes	2017–2022	Orbis lacks 2007 - 2016 reports available in register
50Hertz Transmission GmbH	Yes	2013–2022	Orbis lacks 2023 report available in register
Europace AG	No	N/A	No company data is present in Orbis
Total Energies Holdings	Yes	2013–2022	Partially matches register data – register only includes data for 2021-2023, Orbis lacks 2023 report
Spryker Systems GmbH	No	N/A	No company data is present in Orbis

Reflections

Only three of the five firms were found in the Orbis dataset, with Europace AG and Spryker Systems GmbH notably absent. This highlights a key limitation of commercial databases like Orbis: while they provide a rich source of standardized financial information, they are not exhaustive and may lack coverage for certain firms. Consequently, supplementing Orbis data with official records, such as those from the Unternehmensregister, is essential to ensure a complete and accurate analysis of the corporate population under study.

Step 4 – Largest Firm in Orbis for Postal Code 10557 (Berlin)

Date: 2025-05-19
Activity: Identified and analyzed the largest firm (by total assets) located in postal code 10557 using Orbis data for the fiscal year 2021.
Objective: To determine the firm in postal code 10557 with the highest total assets in Orbis for fiscal year 2021, and compare its financial information with official register data from the Unternehmensregister.

Process

- To identify the largest firm in postal code 10557 for the 2021 fiscal year, I used the Orbis dataset and executed the following steps and code in RStudio:

```
# Filter Orbis data for 10557 postal code and year 2021
pc_10557_2021 <- subset(orbis_2021, postcode == "10557" & year == 2021)

# Find the firm with the largest total assets
largest_firm <- pc_10557_2021[which.max(pc_10557_2021$toas), ]

# View the details of the firm (Company name, Total Assets, Equity, Net Sales)
largest_firm[, c("name_native", "toas", "shfd", "turn")]
```

Results & Cross-Referencing with Register Data

- Identified **Deutsche Bahn Finance GmbH** as the firm with the largest total assets totalling €29,959,251,000
- Retrieved 2021 statements from Unternehmensregister
- Reviewed balance sheet (Bilanz), income statement (GuV), and notes

Comparison Table

Metric	Register	Orbis
Total Assets	€29,959,251,000	€29,959,251,000
Book Value of Equity	€64,132,000	€64,132,000
Net Sales	€466,928,000 (interest income)	(blank)

Reflections

The comparison revealed complete alignment between Orbis data and official filings for total assets and equity, indicating strong consistency in fundamental financial metrics. Notably, net sales were not reported in the traditional sense, as the firm operates as a financing vehicle.

Instead, interest income (Zinserträge) of €466.9 million was recorded, reflecting the nature of its business model. This consistency underscores the reliability of Orbis as a resource for basic financial benchmarking, while also highlighting the importance of understanding industry-specific reporting nuances when interpreting financial data.

Step 5 – Comparing Firms in Postal Code 10557 to All Berlin Firms

Date: 2025-05-22
Activity: Conducted a statistical comparison of firms located in postal code 10557 with the broader Berlin firm population.
Objective: To assess whether firms in my local area (10557) differ significantly from the wider Berlin population in terms of size (total assets) and capital structure (equity ratio), based on 2021 data from Orbis.

Research Design Summary

Data Source:
Used the Orbis panel dataset (`orbis_panel_berlin.rds`), which contains firm-level financial data for Berlin. The focus was on data for fiscal year 2021 to ensure temporal consistency across all firms analyzed.

Sample Selection:
Two comparison groups were defined:

- Group 1: Firms with postal code 10557 (local group)
- Group 2: All firms in the Orbis dataset (entire Berlin firm population)

Note: The analysis compares firms in postal code 10557 to the entire Berlin firm population, which includes the 10557 firms themselves. While this introduces some overlap between the groups, the large overall sample size of the Berlin population mitigates concerns about statistical dependence. The approach maintains a valid comparison for exploring whether local firms differ meaningfully from the broader population.

Financial Metrics Analyzed:

- **Total Assets:** Used as an indicator of firm size.
- **Equity Ratio:** Defined as $(Shareholders' Funds \div Total Assets) \times 100$; measures financial leverage and solvency.

Data Processing:

- Filtered the Orbis dataset to include only firms reporting in 2021.
- Calculated the equity ratio in R using: `equity / total assets`.

- Removed firms with missing data for relevant financial variables.
- Segmented data into the 10557 firms and the Berlin-wide sample.

Statistical Methods:
Welch's two-sample t-tests were employed to compare group means, as this method is robust to unequal variances and sample sizes. For each test, confidence intervals (95%) for the difference in means and degrees of freedom (df) were also reported, providing a more precise understanding of the uncertainty around the estimates. In addition to inferential tests, descriptive statistics including medians, standard deviations, and sample sizes were calculated to highlight distributional patterns and support a comprehensive interpretation of results. All p-values were assessed to determine statistical significance, and findings were contextualized with reference to confidence intervals and dispersion measures to ensure a nuanced understanding of the data.

Tools Used:
All computations and statistical tests were conducted in RStudio. A reproducible PDF report was generated using RMarkdown, including code, tables, and interpretation.

Key Findings

1. **Firm Size (Total Assets):**
    - The mean total assets for firms in postal code 10557 (€272,306,249) is higher than the Berlin-wide average (€24,510,694).
    - The difference is marginally statistically significant ($p = 0.049$), indicating a real, but borderline difference in average firm size.
    - The 95% confidence interval for the mean difference in total assets is (€1,520,454; €494,070,657), a wide range that reflects the high variability and presence of large firms.
    - The high degrees of freedom ($df = 271.46$) confirms that the Welch's t-test was used, adjusting for unequal variances and indicates a large, robust sample (272 for 10557, 28,009 for Berlin). Thus, the t-test has a high degree of precision.
    - The median total assets (€3,960,340 for 10557 vs. €1,015,124 for Berlin) reinforces the pattern: firms in 10557 tend to be larger even when considering typical values, not just means.
2. **Dispersion (Standard Deviation of Total Assets):**
    - The standard deviation of total assets is much larger in 10557 (€2.06 billion) compared to the Berlin-wide firms (€607 million).
    - This means there is more extreme variability in firm sizes within 10557 than across Berlin.
    - The presence of a few very large firms (like Deutsche Bahn Finance GmbH) in 10557 skews the distribution and inflates the standard deviation.

3. **Equity Ratio (%):**
    - The mean equity ratio for 10557 firms is -1.41%, whereas the Berlin-wide mean is a much more negative -4,632.70%.
    - While this difference is not statistically significant (p = 0.203), it suggests that the average capital structure in Berlin is skewed by extreme negative outliers (distressed firms or accounting anomalies).
    - The 95% confidence interval for the equity ratio difference is (-2495.13%; 11757.71%), indicating considerable uncertainty and variability.
    - The median equity ratios (13.32% for 10557 vs. 38.29% for Berlin) suggest that typical firms in Berlin are more equity-financed, but this difference was not tested statistically.
    - This indicates that the negative mean for Berlin is heavily influenced by a few extreme cases, whereas the typical (median) Berlin firm has a more positive equity ratio.
4. **Dispersion in Equity Ratios:**
    - The standard deviation of equity ratios is far lower for 10557 firms (202.41%) compared to Berlin firms (608,486.4%), reinforcing the observation that the Berlin population includes extreme outliers.
5. **Sample Sizes:**
    - The analysis is based on 272 firms in postal code 10557 and 28,009 firms in the Berlin population.
    - These are large enough samples to support robust conclusions.

Reflections

The comparative analysis between firms in postal code 10557 and the broader Berlin population revealed important insights into firm size and capital structure. Firms in 10557 exhibited substantially higher mean total assets (€272.3 million) compared to the Berlin average (€24.5 million), with the difference being marginally statistically significant (p = 0.049, df = 271.46). The 95% confidence interval for the mean difference (€1.5 million; €494 million) reflects a wide range, driven by the substantial variability in firm sizes. Notably, the standard deviation of total assets for 10557 firms (€2.06 billion) is much larger than for Berlin firms (€607 million), indicating greater dispersion within the 10557 group—likely influenced by the presence of a few very large firms such as Deutsche Bahn Finance GmbH. The median total assets also support this observation, with 10557 firms having a higher median (€3.96 million) compared to Berlin (€1.02 million).

For equity ratios, the mean for 10557 firms (-1.41%) is less negative than the Berlin average (-4,632.70%), although this difference is not statistically significant (p = 0.203). The wide confidence interval for the equity ratio difference (-2,495.13%; 11,757.71%) and the much

higher standard deviation for Berlin firms (608,486.40%) suggest the presence of extreme outliers or distressed firms in the Berlin sample. In contrast, 10557 firms show a more stable capital structure, reflected in a lower standard deviation (202.41%) and a more positive median equity ratio (13.32%). Overall, the findings highlight that firms in 10557 tend to be larger and exhibit greater variability in size, but more stable capital structures, compared to the broader Berlin firm population.

Final Summary

This research log documents a systematic and methodologically sound exploration of the financial characteristics of firms located in postal code 10557 Berlin, viewed within the broader context of the Berlin firm population. The process began with the identification of five firms through Impressum verification (Step 1), followed by a detailed review of legal identifiers, register data, and financial statement availability via the Unternehmensregister (Step 2). Step 3 involved integrating Orbis data, which highlighted both the value and limitations of commercial databases, underscoring the need to cross-check multiple sources for a complete analysis. In Step 4, a case study of Deutsche Bahn Finance GmbH, the largest firm in the sample by total assets, revealed consistency between Orbis data and official filings, while also drawing attention to industry-specific reporting practices, such as presenting interest income in place of traditional net sales. The final comparative analysis (Step 5) found that firms in 10557 tend to be larger and exhibit more stable capital structures than the wider Berlin population, with the analysis further illustrating how extreme outliers can distort aggregate statistics. This underscores the importance of using medians and dispersion measures alongside means and statistical tests when comparing firm populations. Overall, the research demonstrates the importance of methodological transparency, critical data interpretation, and triangulating sources to ensure robust, meaningful insights in the study of corporate transparency.