

1. Overview

This repository contains the datasets and analytical materials used in an empirical research project examining circular economy business models, sustainability performance, and network dynamics within the fashion industry.

The repository supports quantitative analysis conducted as part of a Master's-level dissertation in Business Analytics and Sustainability, focusing on the relationship between circular design practices, digital capability, ecosystem connectivity, and firm-level performance outcomes.

All files included here were used directly in the analytical pipeline reported in the dissertation. No additional data sources were introduced outside those documented below.

2. Data Sources

The datasets were compiled from secondary, publicly available sources, including:

- Corporate sustainability reports and disclosures
- Company websites and public product/service descriptions
- Industry reports on circular fashion and sustainable business models
- Publicly visible indicators of partnerships, certifications, and digital capabilities
- Aggregated sentiment and visibility proxies derived from online presence metrics

No proprietary, confidential, or personally identifiable data are included.

The data represent 90 circular-fashion-related firms operating across multiple regions and business model types (e.g., resale, rental, repair, subscription, and hybrid models)

3. Files Included in This Repository

3.1 circular_fashion_companies_90.csv

This file contains firm-level attributes and performance indicators for 90 circular fashion companies.

Key categories of variables include:

- Company characteristics (e.g., founding year, geographic region, business model type)
- Circular design and operational practices (e.g., modular design, recycled material usage, take-back systems)
- Digital capability indicators (e.g., technology integration, AI/ML usage, blockchain traceability)
- Credibility and transparency measures (e.g., sustainability certifications, transparency scores)
- Environmental performance metrics (e.g., carbon reduction percentage, waste diversion percentage, material circularity index)
- Financial and growth indicators (e.g., revenue levels, revenue growth, employee growth)
- Visibility and sentiment proxies (e.g., social mentions, review counts, average sentiment scores)

All variables are defined and interpreted strictly according to the accompanying codebook.

3.2 circular_fashion_company_edges.csv

This file represents the inter-organisational partnership network among the companies included in the dataset.

Each row describes a relationship between two firms, including:

- Source company ID
- Target company ID
- Type of relationship (e.g., logistics, recycling, technology, NGO, material supplier, authentication)
- Edge weight representing relationship intensity or strength

This dataset is used to construct network graphs and compute network metrics such as degree, betweenness centrality, eigenvector centrality, and community structure.

3.3 codes.pdf

This file documents the complete analytical pipeline used in the study.

It includes:

- Data loading and cleaning procedures
- Variable construction and index creation
- Exploratory data analysis (EDA)
- Regression modelling
- Network analysis
- Clustering and dimensionality reduction
- Sentiment analysis validation and comparison

The codebook embedded in this document defines all variables, transformations, and modelling choices used in the dissertation. All reported results, figures, and interpretations in the written analysis are derived exclusively from this code and its outputs.

4. Analytical Scope and Reproducibility

The repository is designed to support transparency and reproducibility:

- All variables referenced in the dissertation correspond directly to fields in the provided datasets
- No simulated or synthetic performance outcomes are used in the core empirical analysis
- Analytical outputs (figures, tables, and model results) are reproducible using the supplied data and code documentation

Users seeking to replicate or extend the analysis should follow the variable definitions and methodological constraints outlined in codes.pdf.

5. Intended Use

This repository is intended for:

- Academic review and examination
- Replication of empirical results
- Methodological transparency
- Educational and research purposes related to circular economy analytics

It is not intended for commercial benchmarking or investment decision-making.

6. Contact and Citation

If this dataset or analytical structure is reused for academic purposes, it should be cited in accordance with standard academic practice, referencing the associated dissertation and underlying literature.