

# HR & Disclosures Synthetically Generated Data

This notebook generates synthetic data needed for ESG equity fund fact sheets, specifically for the team summary and compliance disclosures sections.

- For HR, synthetic data simulates what would be manually provided by the HR department (as per business advisor guidance). This includes names, titles, tenure, and team affiliations. The data is output as a CSV and prepared for GPT-based summary generation to describe the team managing each fund.
- For compliance, a reusable list of regulatory disclosure footnotes has been synthetically created. These footnotes are tagged with unique IDs and labeled as either mandatory or optional. For each fund, the compliance team will provide a list of relevant disclosure IDs, which are then pulled and added to the fund's fact sheet.

## Execution Instructions

**To run this notebook:**

1. Run the notebook sequentially from top to bottom to generate and export HR and Compliance data
2. Output files will be saved locally and visible in the file panel

## File Roadmap

HR Data: Faker-generated employee names, titles, tenure, and team

Compliance Footnotes: Reusable disclosures labeled with DisclosureID and category

**Outputs:** synthetic\_employee\_hr.csv, full\_compliance\_footnotes.csv

## Next Steps

- Generate GPT-based Team Summary using the HR CSV
- Fund-Specific Disclosure DataFrame Generation

## Future Improvement:

## **Synthetic Data Enhancements**

- Add variability in job title seniority
- Group HR data by team/product alignment