## ETL Project Report

**Team Members:**

Allyssa

Emily

Mrun

Nireesha

**Project Name:** Data Integration of Ecological Footprint and Human Freedom Index Datasets

**Datasets to Extract:**

1. National Footprint Accounts 2018

The Ecological Footprint of 196 Countries

Link: <https://www.kaggle.com/footprintnetwork/national-footprint-accounts-2018>

1. The Human Freedom Index

A global measurement of personal, civil, and economic freedom

Link: <https://www.kaggle.com/gsutters/the-human-freedom-index>

**ETL Steps:**

1. Extraction & Initial Data Profiling:
   1. Manual Extraction steps: downloaded .csv files from Cato Institute website (for Human Freedom Index) and Kaggle (for the National Footprint Accounts)
   2. Pandas: read csv files into Pandas DataFrames
2. Data Cleanup & Transformations:
   1. Standardize/Cleanup: updated column names and removed columns we did not plan on using.
   2. Aggregations/Joins: we merged both tables so that both had the same countries.
3. Load:
   1. Database name:
   2. Tables/Columns/Datatypes:
   3. Filter Queries (if any):