Predicting Data Breaches: Code Workflow

Load Libraries
- Mlxtend
- Scikit-learn
- SciPy
- NumPy
- Geopandas
- Matplotlib
- Pandas

Data Check
- Evaluate data integrity
- Import data csv file

Data Processing
- Encode variables into boolean values
- Outlier detection and handling
- Handle skewing in data
- Handle missing values and duplicate data

Feature Selection
- Encode features
- Select features
- Scale features
- Extract relevant features from data

Model
- Select appropriate algorithm
- Train model

Analysis
- Evaluate performance