**Overview: Analysis-Ready FHM Dataset**

**Title:** FHM Data Filtering and Combination for Analysis
**Author:** Soren Donisvitch
**Date:** 07/06/2023

**Purpose**

The Forest Health Monitoring (FHM) dataset has been prepared to provide an analysis-ready version of tree, sapling, and seedling data. This dataset normalizes different plot structures, making comparisons and statistical analyses feasible across multiple years and locations. The processed data ensures that variations in sampling methods do not bias the analysis of tree regeneration, growth, and mortality trends.

**Data Integration and Normalization**

The FHM dataset is derived from multiple sources and database structures:

* **VMC\_fhm (Fia cluster style plots)**
* **VMC\_fhmma (Nested style plots)**

To achieve a standardized dataset, the script harmonizes data structures, applies species translations, normalizes DBH (diameter at breast height) and height measures, and calculates biomass, basal area, and density expansion factors.

**Key Processing Steps**

1. **Species Standardization:**
	* Translates species codes into Latin and common names.
	* Focuses on key species relevant to forest health monitoring.
2. **Biomass and Basal Area Calculations:**
	* Uses allometric models to compute aboveground biomass.
	* Computes basal area per hectare (BA\_MPH) and per acre (BA\_FPA).
3. **Density Normalization:**
	* Uses distinct expansion factors for trees, saplings, and seedlings per hectare and per acre.
	* Adjusts for variations in plot type (Cluster vs. Nested plots).
4. **Tree, Sapling, and Seedling Data Unification:**
	* Merges Vermont and Massachusetts datasets into unified tables.
	* Standardizes tree vigor codes pre- and post-2013.
	* Ensures consistent DBH and height measurement columns.

**Final Outputs**

The following cleaned and standardized datasets are exported as CSV files:

1. **FEMC\_FHM\_TREES.csv** – Normalized tree data with biomass, basal area, and density attributes.
2. **FEMC\_FHM\_SAPLINGS.csv** – Processed sapling data with comparable density metrics.
3. **FEMC\_FHM\_SEEDLINGS.csv** – Seedling-level data with uniform expansion factors.

**Conclusion**

This dataset provides a consistent framework for analyzing forest health trends by ensuring compatibility across different plot sampling methodologies. By standardizing expansion factors, vigor codes, and species classifications, the dataset allows for robust cross-site and temporal comparisons. The processed outputs serve as a foundation for ecological modeling, forest management assessments, and long-term trend analyses in forest regeneration and mortality.