

Documentation for welfare_analysis_framework.R

3/07/2025

Overview

The `welfare_analysis_framework.R` file implements a comprehensive modular framework for analyzing welfare data across different species and time periods. The framework supports multiple welfare quantification methods (NC - Neural Count and WR - Welfare Range) and welfare level methods (isoelastic and 32-82), providing flexible analysis capabilities for comparative welfare research.

Core Analysis Functions

`analyze_welfare_data()`

Purpose: Main orchestration function that processes and analyzes welfare time series data

Key Parameters:

- `data_path`: Path to input Excel file containing welfare data
- `welfare_level_method`: Choice between "isoelastic" or "3282" methods
- `output_dir`: Directory for saving results and visualizations
- `create_visualizations`: Boolean flag for generating plots

Process Flow:

1. Reads and preprocesses data from Excel
2. Ensures both NC and WR columns exist in dataset
3. Calculates net series aggregations
4. Computes correlations and elasticities between human and animal welfare
5. Generates comprehensive visualizations if requested
6. Saves all results to Excel files with method-specific suffixes

Returns: List containing processed data, net series, correlations/elasticities, and factor changes

`ensure_nc_columns()` & `ensure_wr_columns()`

Purpose: Data validation and calculation functions that ensure required welfare metrics exist

NC (Neural Count) Calculations:

- **NC_potential1**: Relative neural capacity compared to humans ($\text{forebrain_neurons} / \text{human_neurons}$)
- **NC_utility**: $\text{Population} \times \text{neural potential} \times \text{welfare level}$
- **NC_tot**: Total neural mass ($\text{population} \times \text{forebrain_neurons}$)

WR (Welfare Range) Calculations:

- **WR_utility**: $\text{Population} \times \text{welfare range potential} \times \text{welfare level}$

Data Requirements: Both functions require **aliveatanytime** (population), **Welfare_level1**, and species-specific neural/welfare data

Statistical Analysis Functions

calculate_net_series()

Purpose: Aggregates welfare metrics across all categories by year, focusing on temporal overlap periods

Methodology:

- Identifies common time window across all species categories
- Filters data to overlapping years only
- Sums welfare utilities and populations by year
- Handles missing values appropriately

Outputs: Annual totals for WR_utility, NC_utility, NC_tot, and total_population

calculate_correlations_elasticities()

Purpose: Computes detailed correlation and elasticity metrics between human and non-human categories

Metrics Calculated (for each animal category):

- **Population correlations:** **cor_pop**, **e_pop**
- **WR utility relationships:** **WR_cor_u**, **WR_cor_hpop_au**, **WR_cor_hu_apop**
- **WR elasticities:** **WR_e_u**, **WR_e_hpop_au**, **WR_e_hu_apop**
- **NC utility relationships:** Similar structure for NC method
- **NC elasticities:** Complete parallel set for neural count method

Robustness Features:

- Handles varying time ranges between species
- Validates data sufficiency before calculations

- Uses error handling for statistical operations

calculate_correlations_elasticities_extended()

Purpose: Performs aggregated group analysis using four predefined animal groupings

Groupings Analyzed:

- **tot:** All categories including humans
- **tot_non_human:** All animals excluding humans and problematic wild categories
- **w:** All wild animals (birds, mammals, fish, arthropods)
- **f:** All farmed animals

Data Source: Reads extended dataset from

`fourth_pass/dat/extended_integrated_calc_tseries.xlsx`

calculate_factor_changes()

Purpose: Computes population change factors from first to last year for each category

Calculation: `max_year_population / min_year_population` for each species category

Data Extension and Preprocessing

extend_animal_trends()

Purpose: Extrapolates animal population data to match human time coverage using intelligent trend analysis

Extension Methods:

- **Backward extension:** Uses early-period linear trends
- **Forward extension:** Chooses between stable mean or linear extrapolation based on coefficient of variation
- **Stability threshold:** $CV < 0.1$ triggers mean-based projection

Parameters:

- `target_year_range:` Desired coverage (default: 1960:2025)
- `endpoint_years:` Number of years for trend calculation (default: 5)

Quality Assurance: Prevents negative population values and maintains data continuity

create_trend_extension_plots()

Purpose: Generates quality control visualizations showing original vs. extended data

Features:

- Identifies categories that were extended
- Color-codes original (black) vs. extended (red) data points
- Saves individual plots for each extended category
- Provides visual validation of extension quality

Visualization System

Master Visualization Function

`create_utility_visualizations()`

Purpose: Orchestrates complete visualization suite creation

Visualization Categories:

1. Individual category utility trends (NC and WR methods)
2. Net utility comparisons with/without humans
3. Total neural mass series
4. Disaggregated plots with aggregate total lines

NC (Neural Count) Visualization Functions

`create_nc_utility_plots()`

Purpose: Generates progressive category exclusion plots for NC utility trends

Plot Progression:

- All categories included
- Excluding wild terrestrial arthropods
- Excluding wild arthropods + wild fish
- Sequential exclusion of humans, wild mammals, farmed fish, chickens, wild birds

Features: Direct line labels, publication-ready formatting, consistent styling

`create_nc_net_utility_comparisons()`

Purpose: Creates comparative net utility analysis plots

Comparison Types:

- With vs. without humans (all categories)
- Human vs. non-human vs. total (filtered categories)
- No wild animals scenario

Time Window: Configurable analysis period (default: 1960-2019)

`create_nc_net_tot_series()`

Purpose: Analyzes total neural mass trends across time

Metrics: Absolute neural mass aggregations with and without human contributions

WR (Welfare Range) Visualization Functions

`create_wr_utility_plots()`

Purpose: Parallel WR utility analysis with same progressive exclusion logic as NC plots

Additional Exclusions: Includes bee-specific filtering for WR method

`create_wr_net_utility_comparisons()`

Purpose: WR-specific net utility comparisons with extended filtering options

Unique Features:

- Bee exclusion scenarios
- Wild animal vs. farmed animal comparisons
- Multi-level aggregation analysis

Specialized Visualization Functions

`create_disaggregated_plots_with_totals()`

Purpose: Creates individual category plots with overlay total trend lines

Plot Types:

- **NC_tot trends:** Basic neural mass by category
- **NC_tot with totals:** Individual categories + aggregate total line
- **NC utility with totals:** Individual utilities + aggregate
- **WR utility with totals:** Individual utilities + aggregate with/without bees

Total Line Calculation: Uses configurable time window (default: 1990-2017) for total calculations

`prepare_data_for_net_series()`

Purpose: Comprehensive data preparation pipeline for net series analysis

Process Flow:

1. Extends animal trends to match human time coverage
2. Creates quality control extension plots

3. Saves extended dataset for reuse
4. Returns prepared data for net series calculations

Output Management: Creates organized directory structure and saves intermediate results

Data Management and Utilities

`universal_ggsave()`

Purpose: Advanced plot saving function with flexible output format support

Features:

- Always saves PDF (standard format)
- Conditional presentation image creation
- Configurable image formats, dimensions, and DPI
- Pattern-based and filename-specific filtering
- Organized output directory management

Configuration Support:

- Specific filename mapping with custom prefixes
- Regex pattern matching for automated selection
- Global save-all option

Technical Implementation Notes

Error Handling

- Comprehensive `tryCatch` usage for statistical calculations
- Data validation before correlation/regression analysis
- Graceful handling of insufficient data scenarios

Performance Considerations

- Efficient data filtering and grouping operations
- Memory-conscious handling of large time series
- Modular function design for selective execution

File Organization

- Systematic output directory creation
- Method-specific file naming conventions
- Intermediate result preservation for debugging

Data Quality Assurance

- Missing value handling strategies
- Data sufficiency validation
- Visual quality control for extensions

Usage Patterns

Typical Workflow

1. **Data Preparation:** Ensure input data has required columns
2. **Method Selection:** Choose welfare level method (isoelastic/3282)
3. **Analysis Execution:** Run `analyze_welfare_data()` with desired parameters
4. **Result Review:** Examine generated Excel files and visualizations
5. **Quality Control:** Review trend extension plots for data validity

Integration Points

- Designed to work with upstream data preparation modules
- Compatible with multiple welfare quantification approaches
- Extensible for additional welfare methods or species categories

This framework provides a complete analytical pipeline for comparative welfare research, supporting both detailed individual species analysis and high-level aggregate trend evaluation across different methodological approaches.