Eseza SAS project
Smoking Cessation Study Analysis

Overview

This project implements comprehensive data processing and analysis for a smoking cessation study using SAS programming. At its core, the analysis focuses on participant motivation scores and treatment outcomes, with particular attention to data quality and automated reporting for publication purposes.

Project Description

The primary focus of this analysis is the processing of a 10-item motivation scale from a smoking cessation study. The project handles complex data cleaning requirements, including the conversion of placeholder missing values (coded as '9') to proper SAS missing values, and implements specific validation rules for scale completion. A key feature is the careful handling of reverse-scored items (7, 8, and 10) to ensure consistent interpretation of motivation scores.

Technical Implementation

The analysis is entirely implemented in SAS, with careful attention to reproducibility and documentation. The code processes the motivation scale data according to strict validation rules, requiring at least 90% completion for scale validity. The SCMSmean score calculation incorporates these requirements, ensuring accurate representation of participant motivation levels.

The project generates three publication-ready tables that summarize key study outcomes. Table 1 presents baseline characteristics across treatment groups, including both categorical and continuous variables. Table 2 focuses on smoking cessation outcomes at different time points, while Table 3 provides a detailed breakdown of adverse events by treatment group and study center.

Significance and Applications

This analysis serves as a crucial component of clinical research reporting, providing automated and reproducible methods for generating publication-quality results. The implementation ensures consistency in data handling and presentation, making it valuable for clinical trial reporting and research publication support. The automated nature of the analysis allows for easy replication with different datasets while maintaining consistent formatting and analytical approaches.

Technical Requirements and Output

The project requires SAS statistical software and works with the kpsmokingnew dataset. All outputs are generated as formatted text files that meet publication standards. The code includes comprehensive documentation and can be easily adapted for use with different data sources while maintaining the same analytical rigor and output quality.

This implementation demonstrates sophisticated approaches to clinical data analysis, combining careful attention to data quality with efficient automated reporting capabilities. The project serves as a model for implementing complex data processing requirements while maintaining high standards of reproducibility and documentation.