

Leveraging SQL and SAS for Analysis-Ready Datasets

Shavonne J Standifer, University Health, Kansas City, Mo

Leveraging SQL and SAS for Analysis-Ready Datasets

Abstract

Data professionals often use a combination of various technologies. Effective data management is essential for ensuring high-quality analysis and decision-making. SQL is a powerful language for querying and manipulating relational databases, while SAS offers a suite of advanced tools for data preparation, statistical analysis, and reporting. By integrating these technologies, organizations can improve data management, enhance data integrity, and foster collaboration across teams.

This paper provides a general guide to utilizing SQL and SAS programming to efficiently create, manage, and maintain analysis-ready datasets. Through step-by-step instructions and real-world examples, readers will gain the skills needed to harness the power of structured query language (SQL) and SAS for streamlining data processes within their organizations.

Overview

By following this technical report, readers will:

- Use a step by step process to create a dynamic report.
- Learn best practices for writing SQL queries in SAS using PROC SQL.
- Understand how to use SAS tools for cleaning, validating, and structuring datasets.
- Discover techniques for using PROC Report to create custom reports.
- Gain the ability to maintain analysis-ready datasets that support data-driven decision-making.
- Use SAS ODS and Proc SGPLOT to create data visuals for custom reporting.

This paper serves as a hands-on guide, equipping readers with the knowledge to apply SQL and SAS in real-world scenarios. It provides detailed instructions for creating optimized datasets, updating records, and ensuring data consistency across various platforms.

At the end of this presentation, you will be able to reference this document as a step-by-step manual to implement best practices in data preparation, management, and maintenance within your organization. Whether you are a data analyst, researcher, or IT

professional, mastering these techniques will enable you to work more efficiently and drive impactful results.

Working Outline

This page shows my working outline.

[Section 1: Understanding Analysis Ready Datasets](#)

[Section 2: SAS Libraries](#)

CREATE YOUR OWN VIRTUAL DATA LIBRARY

This section will give examples and a general overview of the SAS LIBNAME statement. Will also include how to create, manage, and apply them in your code.

[Section 3: Proc SQL – Topic Covered](#)

A QUICK REFRESHER

SQL stands for “structured query language”. It is a language to query, analyze, and manipulate data from databases. Today, SQL is one of the most widely used tools in data.

SQL is primarily used for the initial stages of data extraction and manipulation. SAS excels at data transformation, cleaning, and structuring the data into analysis-ready formats for statistical analysis and reporting. This combined approach leverages the strengths of both tools to efficiently prepare data for a variety of analytical tasks.

THIS SECTION DESCRIBES THE TECHNICAL CONCEPTS THAT YOU NEED TO KNOW IN ORDER TO CREATE, MAINTAIN, AND UPDATE YOUR ANALYSIS READY DATASET AND WILL DEMONSTRATE HOW TO DO THE FOLLOWING:

Table Variables : what are they + how to use + how it helps

Table Alias : what are they + how to use + how it helps

Joining Tables : what are they + how to use + how it helps

Join Types

- One to one join
- One to many joins
- Inner join
- Left join
- Right join

Unions

Set Operators

- Union All
- Union
- Intersect
- Except

Querying Tables

Filtering Data

Aggregating Data

Section 4: The Report Procedure:

CREATE DYNAMIC REPORTS USING THE REPORT PROCEDURE

This section will show the user through example how to use PROC REPORT to customize a data report. Will also cover how to update, and ways to share results with team members.

Section 5: More ways to Customize Reports

This section will show the users through example how to use the SAS ODS and PROC SGPlot to create data visualizations that can be updated dynamically.