

DM04

Streamlining Data Extraction with SAS Made Easy with the Creation of a Centralized Code Repository

Gregory P. Fagan
Consultant

Abstract: *Most ad-hoc, clinical or pharmaceutical SAS programmers struggle to create code used to pull data from a large centralized data warehouse. The struggle is partly due because of the hardship used to define disease states and/or therapeutic classes specific to a particular study. Our solution was to develop a centralized code repository for the definition of disease states and therapeutic classes for our research. Our catalog was developed and managed using Microsoft ACCESS and Visual Basic for Applications. We use macro driven Proc SQL to upload data to our 100+gb data warehouse. The purpose was to develop reusable, macro driven code to reduce redundant programming, shorten the length of code and minimize data entry error while maintaining disease states and therapeutic classes outside of SAS. The results of the catalog made data extraction from our IBM Informix Data warehouse more efficient thus alleviating the burden of ICD9 entry in SAS code.*