SAS® Fuzzy Matching Programming Techniques

Announcing a Popular 3-hour Instructor-led course on
Tuesday, February 20th, 2018; 1:00 pm – 4:00 pm !

Location: Holiday Inn Express Suites Belmont, 1650 El Camino Real, Belmont, California 94002

Course Overview

Data comes in all forms, shapes, sizes and complexities. Stored in files and data sets, SAS® users across industries know all too well that data can be, and often is, problematic and plagued with a variety of issues. One issue, in particular, stems from the complexities associated with linking, or matching, observations from two or more data sets where a common unique identifier, referred to as the key, is not reliable or does not exist. But, what happens when a unique identifier such as a subscriber name, mailing address, and/or email address is misspelled, has one or more characters transposed, or is partially and/or incorrectly recorded? This 3-hour instructor-led course introduces attendees with what fuzzy matching is, the variety of data issues users have to deal with, popular data cleaning and organization techniques, the application of functions such as SUBSTR, SCAN, INDEX, COMPRESS, COMPBL, COMPARE, TRANSLATE, CAT, SOUNDEX (for phonetic matching), SPEDIS, COMPLEV, COMPGED, and an assortment of programming techniques used to resolve key identifier issues and to successfully merge, join and match less than perfect or messy data.

By attending this course, attendees will acquire important skills about:

- ✓ Traditional merge and join techniques
- ✓ Popular data cleaning techniques
- ✓ Resolving key identifier issues
- ✓ Using SUBSTR, SCAN, and INDEX functions
- ✓ Using COMPRESS, COMPBL, and COMPARE functions
- ✓ Using TRANSLATE and CAT functions
- ✓ Comparison of popular fuzzy matching techniques
- ✓ Using SOUNDEX fuzzy matching function
- ✓ Using SPEDIS fuzzy matching function
- ✓ Using COMPLEX fuzzy matching function
- ✓ Using COMPGED fuzzy matching function

SAS® Fuzzy Matching Programming Techniques

Who should attend?

SAS® Fuzzy Matching Programming Techniques is a 3-hour instructor-led course designed for SAS users possessing a minimum of 1-year programming experience. Knowledge of basic DATA, PROC step, merge and/or join programming techniques is encouraged, but not required.

Course Topics

Merge and Join Techniques
- ✓ Exploring data relationships
- ✓ Merge techniques when key(s) are reliable
- ✓ Join techniques when key(s) are reliable

The Fuzzy Matching Process Explained
- ✓ Match-Merge when key(s) are unreliable
- ✓ Join when key(s) are unreliable
- ✓ Fuzzy matching explained
- ✓ Steps to successful fuzzy matching
- ✓ Comparison of fuzzy matching techniques

Fuzzy Matching SOUNDEX Algorithm
- ✓ How the SOUNDEX algorithm works

- ✓ The derived SOUNDEX algorithm codes

Fuzzy Matching SPEDIS Function
- ✓ How the SPEDIS function works
- ✓ SPEDIS function operations and costs

Fuzzy Matching COMPLEV Function
- ✓ How the COMPLEV function works
- ✓ COMPLEV function required and optional arguments
- ✓ COMPLEV function operations

Fuzzy Matching COMPGED Function
- ✓ How the COMPGED function works
- ✓ COMPGED function required and optional arguments
- ✓ COMPGED function operations

SAS Application Developer, Programmer, Educator and Author
Kirk Paul Lafler

Kirk Paul Lafler is an entrepreneur and founder of Software Intelligence Corporation. Kirk has worked with SAS software since 1979 as a consultant, application developer, programmer, SAS solutions provider, data analyst, data manager, infrastructure specialist, performance tuner, educator and author. As a SAS Certified professional, mentor, educator at Software Intelligence Corporation, and advisor and SAS programming adjunct professor at the University of California San Diego Extension, Kirk has taught SAS courses, seminars, webinars and workshops to thousands of users around the world.

Kirk has written six books including Google® Search Complete (Odyssey Press. 2014) and PROC SQL: Beyond the Basics Using SAS, Second Edition (SAS Press. 2013); hundreds of papers and articles on a variety of SAS topics; selected as an Invited speaker and educator at SAS International, regional, special-interest, local, and in-house user group conferences; and is the recipient of 25 "Best" contributed paper, hands-on workshop (HOW), and poster awards.

Registration Information and Form

If you're looking for the most cost-effective and up-to-date training, Kirk Paul Lafler and Software Intelligence Corporation courses, seminars and hands-on workshops prepare you for tomorrow, today. As a world leader in the design and development of high-impact static and interactive dashboards and scorecards, Kirk possesses nearly 40 years' experience developing SAS programming techniques. This course provides users with the latest SAS software features, cutting-edge techniques and know-how – information that will help you achieve greater success and put you ahead of the competition. Kirk invites all SAS users to attend this popular instructor-led course, and expand your programming knowledge using the SAS software. This instructor-led course is taught using a tutorial, step-by-step example, approach.

To Register

1. Register one person per form. Make additional copies if necessary.
2. Print or type all requested information on each form to process registration.
3. RSVP the names of each attendee by sending an email to KirkLafler@cs.com no later than Monday, February 19th, 2018.
4. Bring registration form along with payment in the form of a business or personal check, cashier's check, or money order. (**Note:** Credit Cards cannot be accepted for this training event!)
5. **Training Location:** Holiday Inn Express Suites Belmont, 1650 El Camino Real, Belmont, California 94002.
6. **Make payment for $189 to:** Software Intelligence Corporation.

Attendee Information (Please enter all requested information, below)

❑ **SAS® Fuzzy Matching Programming Techniques (3-hours – $189)**
Tuesday, February 20th, 2018; 1:00 pm – 4:00 pm

(Location: Holiday Inn Express Suites Belmont, 1650 El Camino Real, Belmont, California 94002)

❑ Dr. ❑ Professor ❑ Mr. ❑ Mrs. ❑ Ms.

Name _____

Organization _____

Department _____

Address #1 _____

Address #2 _____

City _____ State _____ Zip _____

Phone _____

E-mail _____

Cancellations and Substitutions

Upon registering, attendee(s) are considered successfully enrolled in the program. Should the attendee(s) decide to cancel their enrollment; a cancellation fee will be charged. Cancellations must be submitted in writing. Cancellations 3 days or more in advance of the course start date will be given a complete refund, minus a $50 processing fee. Cancellations received less than 3 days before the start of the course will be subject to a $50 cancellation fee. "No shows" for a course is not considered a cancellation – payment will not be refunded. "No shows" will be sent a copy of the course notes in the form of a PDF document. Please note that attendee substitutions are permissible as long as it is communicated within 3 days prior to the start of the course.

Software Intelligence Corporation provides attendees with cost-effective, comprehensive, and effective technical knowledge. We offer our products, support, education, and consulting, in 10 countries. Copyright © 2015 – 2018, Software Intelligence Corporation. All rights reserved.

SAS and other SAS products are the registered trademarks of SAS Institute Inc., Cary, NC, USA.
Excel, Word, MS-SQL Server and Access are the registered trademarks of Microsoft Corporation.