

---

## **e-Learning - Advanced Concepts**



**BASE 3 SYSTEMS  
THE LOW BARN  
BEAMSLEY  
SKIPTON  
NORTH YORKSHIRE  
BD23 6HJ  
TEL +44 (0)1756 718080  
FAX +44 (0)1756 718087  
E-MAIL [ADMIN@BASE3.COM](mailto:ADMIN@BASE3.COM)**

---



---

SAS and all other SAS Inc. product and service names are registered trademarks of SAS Inc. in the USA and other countries. ® indicates registration in the USA.

All other brand and product names are registered trademarks of their respective owners.

SAS Inc. has not authorized, sponsored, endorsed or approved this course and is not responsible for its content.

Copyright © 2019 by Base 3 Systems Limited, The Low Barn, Beamsley, Skipton, North Yorkshire, BD23 6HJ. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, Base 3 Systems Limited.

Base 3 Systems Limited registered in England no. 3268508. Registered Office: Stuart House 15/17 North Park Road, Harrogate, North Yorkshire, HG1 5PG

Created : 30 January 2019



## Course Aims

After attending this course, attendees should have an understanding of efficiency techniques that may be employed, and the balance between efficiency and maintainability of SAS code, enabling streamlining of existing code and development of new code that meets efficiency requirements. Pattern matching is introduced in the form of Perl regular expressions and Hash objects are explored as an alternative method to combine data. Delegates would gain knowledge of how to utilise the SAS system tables as well as running and debugging simple batch files.

## Duration

2.5 days

## Required Knowledge

Delegates should have attended or have prior knowledge of base programming, SQL and macro

## Advanced Concepts – Course Content

### 1. Performance Measurement Concepts

#### 1.1 Performance Measurement Concepts

- System Resources – CPU, storage space, memory, I/O (input/output), real & elapsed time

#### 1.2 Performance Tuning

- Performance Trade-offs & Tuning

#### 1.3 Timer Statistics

- Available Timer Statistics
- System Options for Timer Statistics

#### 1.4 Baselineing

- Baselineing Programs

#### 1.5 Program Code: Economy & Maintainability

- Program Economy and Maintainability
- Restricting Options

### 2. Efficient Data Storage

#### 2.1 Data Storage

- How is SAS Data Stored - Pages and Buffers
- Specifying Page Size
- The BUFSIZE option & Performance Effects

#### 2.2 Variables

- Character Variables
- Numeric Variables
- Reducing the Length of Variables
- Performance Issues

#### 2.2 Exercises

#### 2.2 Solutions

---

## 2.3 Data Compression

- What is Data Compression?
- Compressed Data vs. Uncompressed Data
- The COMPRESS Option
- Performance Effect of Compressing Data

### 2.3 Exercises

### 2.3 Solutions

## 2.4 Data Step Views

- What are Data Views?
- Creating views
- When to Use Views
- Stored Compiled DATA Step Programs

### 2.4 Exercises

### 2.4 Solutions

## 3. Indexing

### 3.1 Indexing Explained

- What is an Index?
- What are Indexes Used For?
- Types of Indexes

### 3.2 Creating & Maintaining Indexes

- Index Documentation
- Guidelines for Defining Indexes
- Creating Indexes
- Maintaining Indexes
- Deleting Indexes

### 3.2 Exercises

### 3.2 Solutions

### 3.3 Processing Using Indexes

- The MSGLEVEL System Option
- How SAS Uses an Index
- Processing an Indexed Data Set
- Implicit Indexed Access
- Examining & Updating Centiles
- Explicit Index Access – IDXWHERE & IDXNAME

### 3.3 Exercises

### 3.3 Solutions

### 3.4 Overhead & Performance Issues

- Overhead & Performance Issues

### 3.5 The SPDE & Efficient Indexing

- SPDE Advantages & Disadvantages
- Efficient Indexing with SPDE

## 4. Efficient Data Access

### 4.1 Memory & I/O

- The Relationship Between Memory & I/O
- The BUFNO Option

---

- The SASFILE statement

- 4.1 Exercises

- 4.1 Solutions

- 4.2 Managing the WORK Library

- The SAS WORK library

- Insufficient WORK Library Space

- Compressing and Deleting Work Data Sets

- Making Use of the USER Library

- 4.2 Exercises

- 4.2 Solutions

- 4.3 Accessing Only Necessary Data

- Random Access

- The POINT & KEY Statement Options

- Reading & Writing Only Variables of Interest

- Conditionally Reading & Writing Observations

- Implicit Indexed Access

- 4.4 Sorting Data

- Eliminating Duplicate Observations

- Sorting Large Data Sets

- Avoiding Sorting Data

- 4.4 Exercises

- 4.4 Solutions

- 4.5 SAS Function Compiler Procedure

- THE PROC FCMP Procedure – Creating user-defined functions

- Applying Created Functions

- Defining Character Arguments

- 4.5 Exercises

- 4.5 Solutions

- 5. Perl Regular Expressions & Data Step Component Objects

- 5.1 Perl Regular Expressions

- Uses of Pattern Matching

- Basics of Perl Regular Expressions

- 5.2 PRX Functions and CALL Routines

- Capture Buffers

- PRX Functions

- PRX CALL Routines

- 5.2 Exercises

- 5.2 Solutions

- 5.3 Data Step Component Interface & Declaring a Hash Object

- Declaring a Hash Object

- Defining Lookup Keys

- Inserting Data

- 5.3 Exercises

- 5.3 Solutions

- 5.4 The Hash Object

- Creating an Output Data Set

- Loading a Hash Object with a SAS Data Set
- Hash Objects vs Data Step Merges & SQL Joins

5.4 Exercises

5.4 Solutions

#### 5.5 Hash Iterator Object

- Declaring a Hash Iterator Object
- Navigating a Hash Iterator Object

### 6. Efficient Data Modification & Combination

#### 6.1 Modifying Data In-Place

- Modifying Data In-Place
- The SET & UPDATE Statements

#### 6.2 Updating Data with a Transaction Data Set

- Master & Transaction Data Sets
- Transaction Data Sets & The UPDATE Statement
- Transaction Data Sets & The MODIFY Statement
- The \_IORC\_ Automatic Variable

6.2 Exercises

6.2 Solutions

#### 6.3 Re-Dimensioning Data

- Rotating Data with Arrays
- Rotating Data with the Transpose Procedure

6.3 Exercises

6.3 Solutions

#### 6.4 Combining Data

- Concatenating Data with The APPEND Procedure
- Concatenating Raw Data Files
- Merging Data with Indexed Access

6.4 Exercises

6.4 Solutions

#### 6.5 Using Formats in Place of Merges

- Creating Formats with SAS Data Sets – The CNTLIN Method
- Using Formats as an Alternative to Merges

### 7. Metadata, XML & Batch Execution

#### 7.1 Metadata Servers & Proc Metalib

- Metadata Servers & PROC METALIB

#### 7.2 SASHelp Views

- What Views are Available?
- Generating Macro Variables from Dictionary Tables

7.2 Exercises

7.2 Solutions

#### 7.3 XML

- What is XML?
- Elements vs Attributes
- The XML LIBNAME Engine - Importing & Exporting XML in SAS

7.3 Exercises



---

### 7.3 Solutions

#### 7.4 Batch Execution & Checkpoint Mode

- What is a Batch File?
- Creating a Batch File
- Checkpoint & Restart Mode

#### 7.4 Exercises

#### 7.4 Solutions