

Data Integration

SAS Institute Inc

109004, Москва,
ул. Станиславского, д. 21, стр. 1

+7 (495) 937 4151

www.sas.com/russia

7/1/2010

Перспектива курсов

Курсы по интеграции данных будут интересны специалистам, задействованным в разработке хранилищ данных, в проектировании ETL-процессов, разработке моделей данных. Возможные роли в проектах внедрения и сопровождения системы SAS на предприятии: программисты, разработчики моделей\систем интеграции данных, ИТ-администраторы. На базе ускоренного курса по DI можно подготовиться к сертификации на звание SAS Data Integration Developer и пройти ее в тестовых центрах Thompson Prometric или в УЦ SAS Institute (Москва).

Список курсов, представленных в проспекте

Код курса	Название курса	Версия системы	Длительность курса	Уровень сложности	Стоимость за 1 чел. (руб.)
<u>DIFAST</u>	SAS Data Integration. Ускоренный курс	9.2	3	4	55 860
<u>DIWN</u>	Что нового в SAS 9.2 Data Integration	9.2	1	2	14 480
<u>DIS</u>	Использование SAS Data Inegration Studio для построения витрин данных предприятия	9.1.3	3	3	55 860

Сертификация и курс, который к ней готовит

Название курса/теста	Длит-ть курса/теста	Цена теста за 1 чел
Ускоренный курс по подготовки по Data Inegration	3	
Exam SAS Data Integration Developer	3 часа	5200

Что нового в SAS 9.2 Data Integration

Цель курса

Этот курс для разработчиков систем интеграции данных, имеющих опыт использования SAS Data Integration Studio версии SAS 9.1.3 и желающих понять, что изменилось в приложении версии SAS 9.2.

Описание курса

Изучить как

- Использовать новые особенности для Data Integration Studio
- Работать с новым интерфейсом
- Работать с Job Editor
- Работать с новыми трансформациями

Аудитория, требуемый уровень подготовки

Этот курс для разработчиков систем интеграции данных, имеющих опыт использования SAS Data Integration Studio версии SAS 9.1.3 и желающих понять, как изменилось приложение в новой версии SAS 9.2.

Программа курса:

Обзор основных особенностей SAS Data Integration Studio

- работа в рамках платформы SAS Business Analytics
- основные изменения SAS Data Integration Studio
- основные изменения в Change Management

Особенности нового редактора процессов

- основные изменения
- работа в редакторе процессов
- работа с трансформацией SQL Join
- редактирование процессов
- статистика выполнения процессов
- отчеты по объектам метаданных

Трансформация SAS Code

- работа с трансформацией SAS code
- использование мастера создания трансформаций

Особенности связывания полей (Mapping and Propagation)

- обзор новых возможностей связывания полей
- обзор новых возможностей распространения изменений

Новое в управление качеством данных (Data Quality)

- введение в DataFlux Integration Server
- работа с трансформацией DataFlux IS

DIWN

Длительность: **1**
день

Уровень: **2**

Версия ПО: **9.2**

DIFAST

Длительность: 3 дня

Уровень: 4

Версия ПО: 9.2

Сертификация:

SAS Certified Data
Integration
Developer

SAS Data Integration. Ускоренный курс

Цель курса

Этот интенсивный курс обеспечивает ускоренное обучение пользователей, которые имеют дело с процессами обработки данных, включая их очистку и обогащение, ETL-обработку и объединение.

Описание курса

В курсе изучается как

- Создать витрины/наборы, обеспечивающие работу соответствующих запросов и генерацию отчетов

- Применять задания (jobs) для планировщика

- Управлять заданиями (jobs) для формирования наборов с контролем эффективности

- Внедрить стандарты качества данных в ETL-процессы

- Осуществлять техническое обслуживание средствами базы знаний Knowledge Base

- По-возможности использовать SPDS (Scalable Performance Data Server)

Аудитория, требуемый уровень подготовки

Курс рассчитан на разработчиков, которым требуется изучить процессы преобразования данных в кратчайшие сроки

Этот курс предназначен для тех, кто способен воспринять большой объем информации в сжатые сроки.

Программа курса:

Введение в SAS Data Integration Studio

- определение платформы для SAS Business Analytics
- навигация в SAS Data Integration Studio

Работа под Working Under Change Management

- defining change management
- определение типов репозиториев
- using change management

Построение Витрин

- обзор исходных данных, используемых в курсе
- определение целевых таблиц
- загрузка целевых таблиц

Построение OLAP-кубов

- определения OLAP
- построение OLAP-куба
- просмотр OLAP-куба

Создание отчетов

- обсуждение преобразований discussing reporting transformations
- построение частотного отчета
- создание пользовательских преобразований
- определение хранимых процессов
- задание свойств хранимых процессов
- описание инструментов, использующих хранимые процессы

Дополнительные разделы и преобразования

- задание свойств загрузки и выгрузки
- определение свойств копирования, перемещения и специальных возможностей
- задание свойств анализа взаимодействий
- задание признаков для Loop-преобразований

Отслеживание эффективности в ETL-диаграммах: Основные настройки

- определение важных положений для SAS Data Integration Studio
- отслеживание и измерение эффективности
- работа с промежуточными файлами
- представления или реальные таблицы для ETL-процессов
- преобразования и оптимизация переменных

Отслеживание эффективности в ETL-диаграммах: данные, сортировка и объединения

- определение дополнительных техник для улучшения производительности
- работа с данными, сортировка, объединение

SAS и качество данных

- определение качества данных
- описание методологии обеспечения качества данных
- обзор задач обеспечения качества данных

«Профилирование» данных

- определение «профилирования» данных
- «профилирование» с применением dfPower Studio

Очистка данных: стандартизация

- определение стандартизации
- стандартизация с применением dfPower Studio
- стандартизация с применением SAS Data Integration Studio
- стандартизация с применением SAS

Очистка данных: дополнительная функциональность

- дополнительные разделы обеспечения качества/очистки данных
- описание функциональности dfPower Architect
- упорядочивание шагов для построения Задач (job)
- описание функций SAS Data Quality Server

Интеграция данных (Data Integration)

- согласование и интеграция данных
- определение частей dfPower согласования
- применение dfPower Match и dfPower Architect для создания кода согласования
- описание возможностей SAS Data Integration Studio для целей интеграции данных

DIS

Длительность: 3 дня

Уровень: 3

Версия ПО: 9.1.3

Использование SAS 9.2 Data Integration Studio для построения витрин данных предприятия

Цель курса

Курс знакомит с продуктом SAS Data Integration Studio, ранее известным как SAS ETL Studio.

Описание курса

Слушатели знакомятся с навигацией по интерфейсу SAS Data Integration Studio, изучают, как извлекать, преобразовывать и загружать данные с помощью SAS Data Integration Studio. Дополнительные вопросы включают использование встроенных блоков (плагинов) SAS Data Integration Studio, управление хранилищем данных, организацию внесения изменений, формирование расписаний, использование метаданных.

Аудитория, требуемый уровень подготовки

Данный курс предназначен для разработчиков хранилищ данных, проектировщиков ETL-процессов, разработчиков моделей данных, занимающихся вопросами извлечения, преобразования и загрузки данных в хранилище.

Для успешного обучения на тренинге, слушатели должны быть знакомы с

- языком программирования SAS
- языком программирования SQL (Structured Query Language)

с принципами моделирования данных.

Программа курса:

Обзор платформы SAS 9 Business Intelligence и SAS Data Integration Studio

- обзор Business Intelligence
- потребители информации Business Intelligence
- интерфейс SAS Data Integration Studio

Организация внесений изменений (Change Management)

- понятие организации внесения изменений
- типы репозиториев
- применение организации внесения изменений

Разработка витрин данных для курса

- планирование хранилища данных
- данные курса (компания Orion Star)
- модели данных для Orion Star

Построение витрин данных

- обзор рассматриваемых задач
- описание источников данных
- описание целевых таблиц
- загрузка целевых таблиц

Формирование OLAP-куба

- определение OLAP-куба
- построение OLAP-куба
- анализ данных OLAP-куба с помощью Microsoft Excel

Создание отчетов

- отчет о преобразованиях
- использование мастера Transformation Generator
- SAS Data Integration Studio и сохраняемые процессы

Преобразования для медленно изменяющихся измерений

- определение медленно изменяющихся измерений
- применение задачи SCD Type 2 Loader
- применение задачи Fact Table Lookup
- применение задачи Surrogate Key Generator
- применение задачи Key Effective Date

Планирование процессов

- составление расписания для процессов SAS Data Integration Studio

SAS Data Integration Studio и Data Quality

- SAS и качество данных
- использование задачи Apply Lookup Standardization
- использование задачи Create Match Code
- использование задачи Data Validation

Администрирование SAS Data Integration Studio

- постановка задач для SAS Data Integration Studio
- отладка организации внесения изменений
- безопасность планирования и выполнения (самостоятельно)

Сертификация по разработке баз данных. SAS Data Integration Developer

Длительность экзамена: 3 часа

Пройти сертификацию можно в тестовом центре Thompson Prometric или в Учебном центре SAS Institute (Московское представительство), язык сертификации – английский

Стоимость – 5 200 р. без учета НДС.



Аудитория

Разработчики в области интеграции данных. Для успешного прохождения сертификации необходимо уметь:

- создавать таблицы витрин данных/хранилищ данных для выполнения запросов и составления отчетности;
- планировать выполнение заданий;
- вести задания, выполняющие построение таблиц на основе отслеживания эффективности;
- инкорпорировать стандарты качества данных в потоки ETL;
- обслуживать базу знаний Quality Knowledge Base;
- по возможности, использовать преимущества сервера SAS Scalable Performance Data Server.

Рекомендуемые материалы для подготовки

Рекомендуемая база:

- [SAS Data Integration Studio: Экспресс курс SAS 9.2 \(DIFAST\)](#)

Для расширенного изучения данной тематики в течение более длительного времени доступны такие курсы, как

- [Использование SAS Data Integration Studio для создания витрин данных на основе источников данных предприятия.](#)
- [Использование SAS Data Quality Solution для очистки данных.](#)

Содержание экзамена

Overview

- Define the architecture of the platform for SAS Business Analytics.
- Describe the available interfaces.
- Define the change management feature of SAS Data Integration Studio.
- Discuss the DataFlux Integration Server.

Creating Metadata for Source and Target Data

- Define administrative tasks to be performed for SAS Data Integration Studio
- Describe the New Library Wizard

Use Register Tables wizard to register source data.

Use Register Tables wizard to register metadata for a Microsoft Access database table using ODBC.

Register metadata for a comma-delimited external file

Creating Metadata for Target Data and Jobs

Describe features of the New Tables wizard.

Discuss SAS packages

Discuss importing and exporting of relational metadata.

Discuss components of SQL Join's Designer window.

Investigate mapping and propagation.

Work with performance statistics.

Generate reports on metadata for tables and jobs.

Working with Transformations

Discuss and use the Extract and Summary Statistics transformation.

Discuss and use the Loop transformations.

Investigate where status handling is available.

Discuss and use the Data Validation transformation.

Discuss and use the Rank, Transpose, Append, List and Sort transformations.

Discuss and use the Apply Lookup Standardization transformation.

Working with Tables and the Table Loader Transformation

Discuss reasons to use the Table Loader transformation.

Discuss various load styles provided by the Table Loader transformation.

Discuss various types of keys and how to define in SAS Data Integration Studio.

Discuss indexes and how to define in SAS Data Integration Studio.

Discuss Table Loader options for keys and indexes

Working with Slowly Changing Dimensions

List the functions of the SCD Type 2 transformation.

Define business keys.

Define surrogate and retained keys.

Detect and track changes.

Discuss the Lookup transformation.

Defining Generated Transformations

Define SAS code transformation templates.

Create a custom transformation

Deploying Jobs

Discuss the types of job deployment available for SAS Data Integration Studio Jobs.

Provide an overview of the scheduling process.

Discuss the types of scheduling servers.

Discuss the Schedule Manager in SAS Management Console.

Discuss batch servers.

List the applications that can be used to create and execute stored processes.

Describe deployment of SAS Data Integration Studio jobs as a SAS Stored Process.

Образцы вопросов на экзамене <http://support.sas.com/certify/creds>

По вопросам сертификации обращайтесь по адресу:
cert@rus.sas.com

Тел. +7 495 937 4151