

Работа с данными в IBM SPSS Modeler

01 Введение в работу с данными

Процесс data mining
Формат файла для анализа
Единица анализа для моделирования
Совместимость данных и метода моделирования

02 Извлечение выборок

Узел Sample
Типы выборок
Простой отбор
Сложный отбор

03 Работа с датами

Чтение данных с датами
Вычисления с датами
Применение одного и того же выражения к нескольким переменным

04 Работа с текстовыми данными

Управление текстовыми данными
Пример работы с текстом

05 Преобразования данных

Применение итоговых статистик с помощью узла Set Globals

Преобразования количественных полей
Категоризация полей
Приложение
Ошибки в выражениях CLEM

06 Работа с последовательностями

Функции CLEM для последовательностей
Режимы Count и State узла Derive
Реструктуризация данных последовательности с помощью узла History

07 Экспорт файлов данных

Что использовать при моделировании: весь стрим, или готовый файл данных?
Поддерживаемые форматы экспорта
Экспорт в файл простой структуры
Экспорт в базу данных

08 Производительность

SQL Pushback
SQL...оптимизация
Порядок узлов
Используйте выборки
Максимальное количество значений
Производительность отдельных узлов



09 Объединения таблиц баз данных

Основные установки объединения файлов данных в Modeler

Внутреннее объединение (Inner Join)

Объединение с несогласующимися записями в одной из таблиц

Объединение с несогласующимися записями в обеих таблицах

Более сложные варианты объединения данных в Modeler

Узлы Sort и Distinct

Объединение более двух таблиц

10 Работа с узлом Statistics Transform

Подсчет количества значений

