

Практическое занятие 4

Импорт и экспорт данных средствами MS SQL Server 2019

Задание 1. Извлечение данных из представления хранилища и загрузка их в таблицу (Сарка, стр. 92-100)

- Загрузите мастер импорта/экспорта MS SQL Server 2019
- Подключитесь к базе данных AdventureWorks2019;
- В источниках импорта укажите два представления: [Production].[ProductAndDescription] и [Production].[ProductModelInstructions].
- Создайте в том же экземпляре сервера целевую базу данных Import.
- Переместите данные таблиц-источников в целевые таблицы [Production].[ProductAndDescription] и [Production].[ProductModelInstructions] соответственно.

Ответы на вопросы

- Перечислите и опишите источники импорта, предлагаемые мастером.
- Совпадают ли источники импорта и целевые источники в мастере? В чем разница?
- Что такое «сопоставление столбцов» (mappings) и для чего это нужно при импорте?
- Почему результирующий файл мастера называется пакетом (package)?

Задание 2. Извлечение данных из «необработанного файла» и загрузка их в таблицу

- В свободных источниках интернета найдите подходящий массив «открытых данных» (например, открытые данные министерства культуры РФ <https://opendata.mkrf.ru>)
- Изучите описание и метаданные предлагаемых наборов данных, выберите подходящий (минимум очистки данных) и загрузите файл набора данных в формате *.csv
- Подключитесь к базе данных AdventureWorks2019;
- В источниках импорта укажите скачанный файл *.csv, внимательно изучите mappings и структуру получаемой таблицы. Если необходимо, внесите изменения.
- Создайте в том же экземпляре сервера целевую базу данных OpenData.
- Переместите данные из файла *.csv в целевую таблицу в созданной базе данных.

Ответы на вопросы

- Что такое *.csv файл? Какие разделители в нем используются?
- Как узнать кодировку информации внутри файла *.csv?
- Какие проблемы вы обнаружили в выбранном наборе данных при изучении mappings? Перечислите их списком. Как бы вы их решали (где это возможно)?

Код проблемы	Суть проблемы	Вариант решения
pr001	Текстовые значения столбца «Аннотация» содержат несколько запятых. При этом символ запятой используется как разделитель csv. В итоге одно значение переносится на несколько столбцов.	<ol style="list-style-type: none">1. Выбрать другой разделитель для проблемного столбца.2. Убрать столбец, если он не нужен.3. ...

Задание 3. Извлечение данных из документа JSON и загрузка их в таблицу

- В свободных источниках интернета найдите подходящий массив «открытых данных» (например, открытые данные министерства культуры РФ <https://opendata.mkrf.ru>).
- Просмотрите структуру JSON-документа.
- Откройте любую оболочку менеджмента СУБД MS SQL Server.
- Подготовьте таблицу в целевой базе данных OpenData в соответствии со структурой документа JSON.
- Переместите данные из документа JSON в целевую таблицу в созданной базе данных.

Ответы на вопросы

- Для чего нужна функция OPENROWSET(BULK) при работе с документами JSON? Что является результатом ее работы? Покажите на скриншоте.
- Напишите код, позволяющий переместить содержимое JSON документа во временную таблицу MS SQL Server.
- Напишите код транзакции, которая пакетами (SAVEPOINT) перенесет содержимое временной таблицы в постоянную таблицу целевой базы данных OpenData.