**Домашняя работа №12**

**Тема 12: Создание решения ETL**

**Задания**

1. Создайте SSIS пакет с добавлением данных (10 новых пловцов) в таблицу Swimmer из AdventureWorksDW2017.dbo.DimEmployee или из своего сгенерированного файла. Используйте задачу OLE DB.
2. Определите отличие в концепциях построения AdventureWorks2017 и AdventureWorksDW2017 баз. Какая база будет работать быстрее с отчётами? Почему?
3. Создайте SSIS пакет с добавлением данных (10 новых пловцов) в таблицу Swimmer из AdventureWorksDW2017.dbo.DimEmployee или из своего сгенерированного файла. Используйте задачу OLE
4. **Источник "OLE DB AdventureWorksDW2017\_dbo\_DimEmployee "**

Сделала выборку 10 первых уникальных людей, сгенерировала Rank и TrainerId.

SELECT [Number]

,[FirstName]

,[LastName]

,[BirthDate]

,[Gender]

,[Rank] = ABS(CHECKSUM(NewId())) % 7 + 1

,[TrainerId] = FLOOR(RAND(CHECKSUM(NEWID()))\*(1000-1)+1)

FROM

(SELECT DISTINCT TOP (10)

[Number] = [EmployeeNationalIDAlternateKey]

,[FirstName]

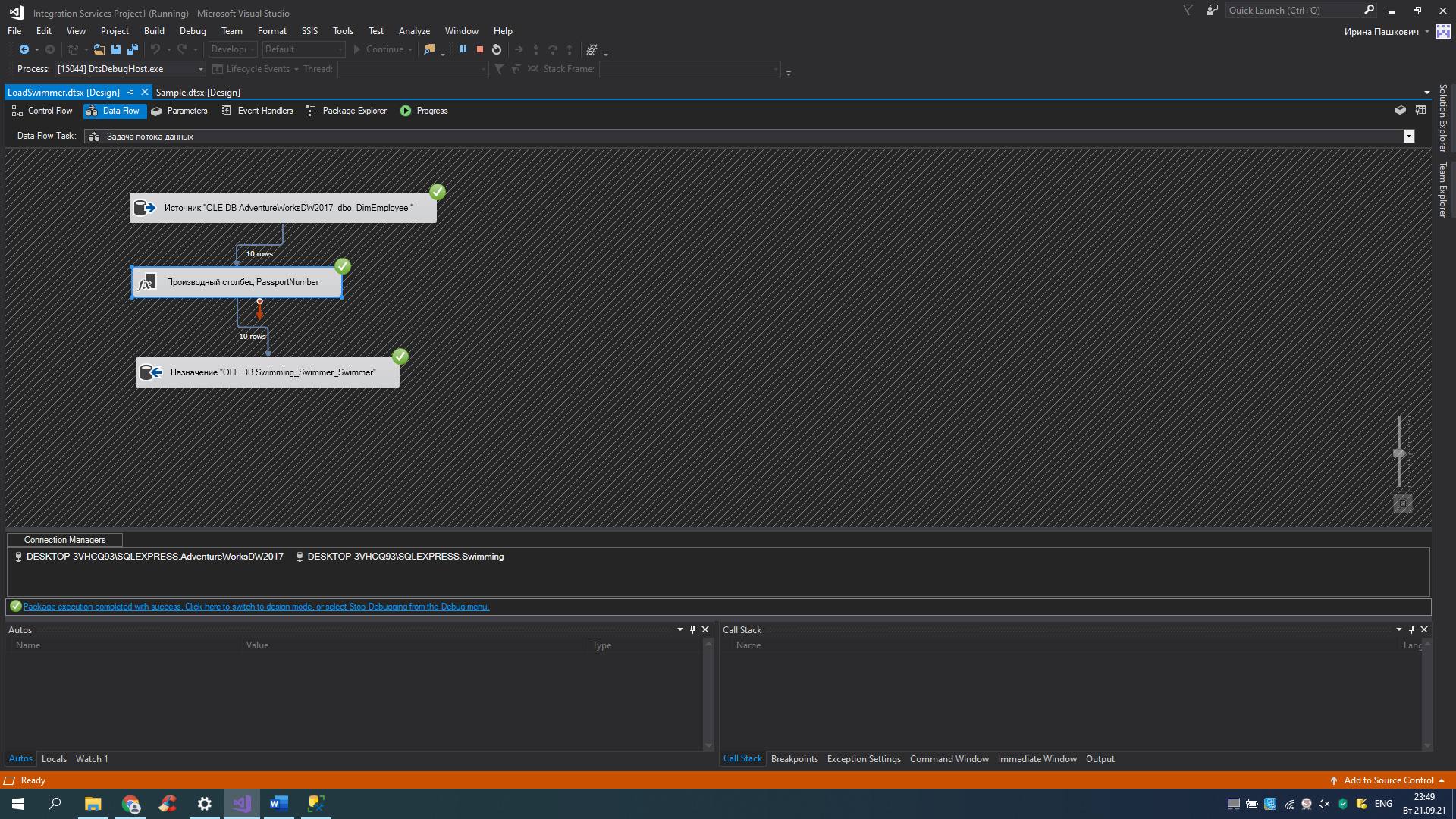
,[LastName]

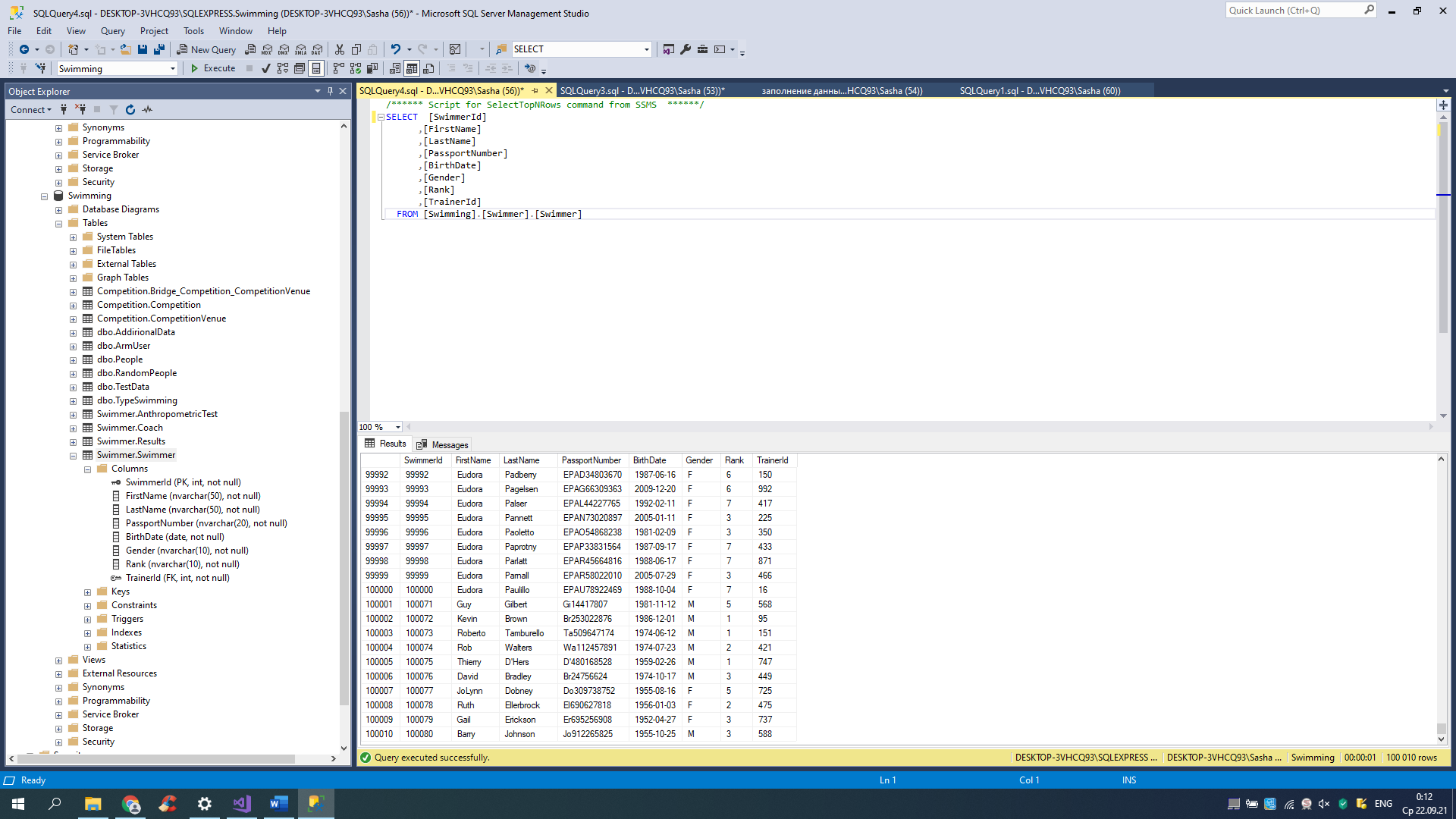
,[BirthDate]

,[Gender]

FROM [AdventureWorksDW2017].[dbo].[DimEmployee]) people

1. Сделала производный столбец PassportNumber к полю Number ( в исходной базе [EmployeeNationalIDAlternateKey]) добавляю 2 первые буквы фамилии.
2. Вставляю данные в назначение Назначение "OLE DB Swimming\_Swimmer\_Swimmer"





1. Определите отличие в концепциях построения AdventureWorks2017 и AdventureWorksDW2017 баз. Какая база будет работать быстрее с отчётами? Почему?

|  |  |
| --- | --- |
| **[AdventureWorks2017]** | **[AdventureWorksDW2017]** |
| Нормализованные данные | Применяется денормализация |
| Модель данных снежинка | Модель данных Звезда |
|  | Есть таблицы фактов и связанные с ними таблицы размерностей |
|  | **Будет работать с отчётами быстрее , т.к. в запросах будет применяться меньше соединений, меньше используется внешних ключей, а значит скорость выполнения запроса выше, отчёт будет строится быстрее** |