Міністерство освіти і науки України

Національний університет “Львівська політехніка”



Звіт

про виконання лабораторної роботи № 3

iз курсу “Аналітичні та нереляційні бази даних”

**Виконав**:

ст. гр. ІР-42, ІКТА

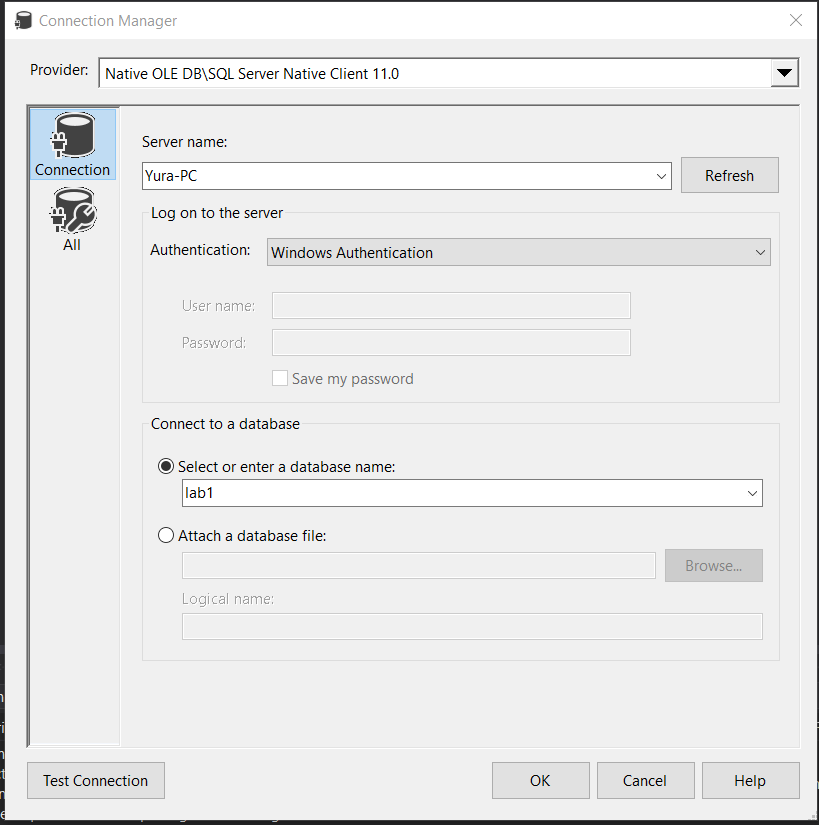
Панейко Юрій

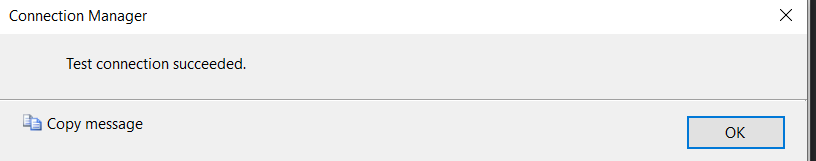
**Прийняв**:

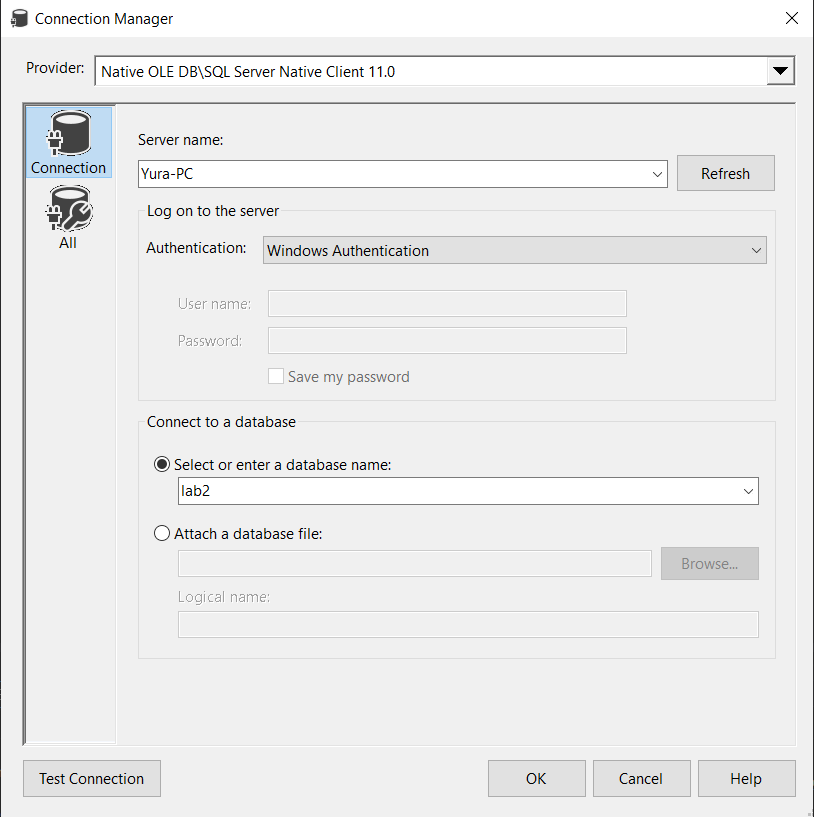
Верес З.Є.

Львів – 2020

Для початку ми створюємо конекшн менеджери для прямого зв’язку з обидвома базами даних.

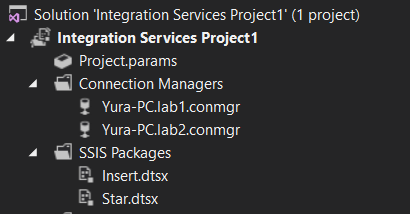




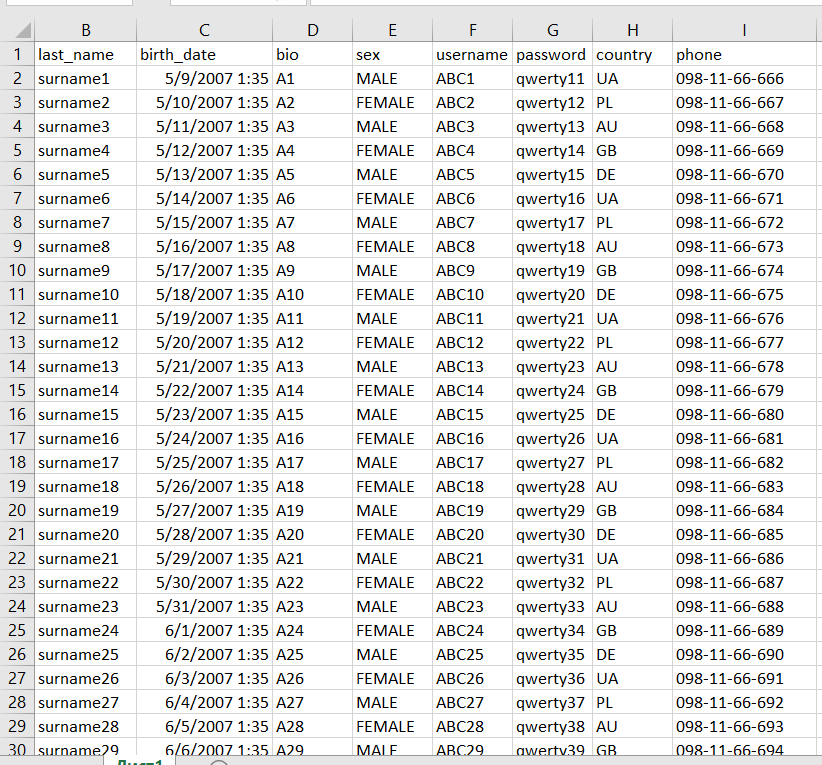




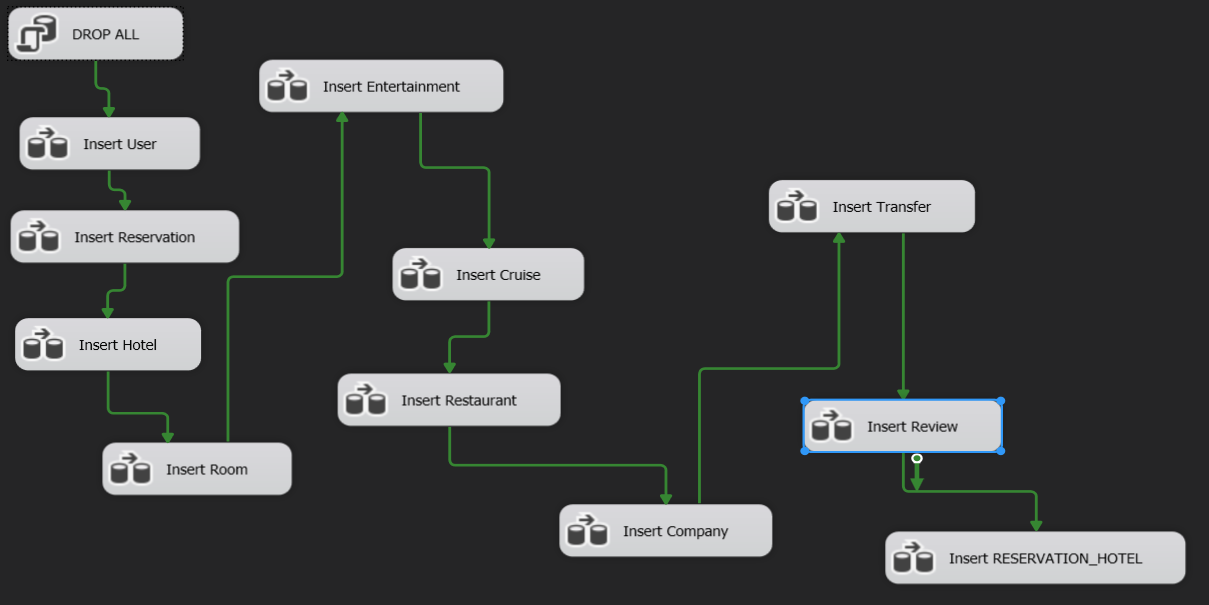
Загальна структура проекту:



У першому SSIS пакеті – ми заповнюємо першу БД на основі даних, що були попередньо занесені у відповідні Excel таблиці.

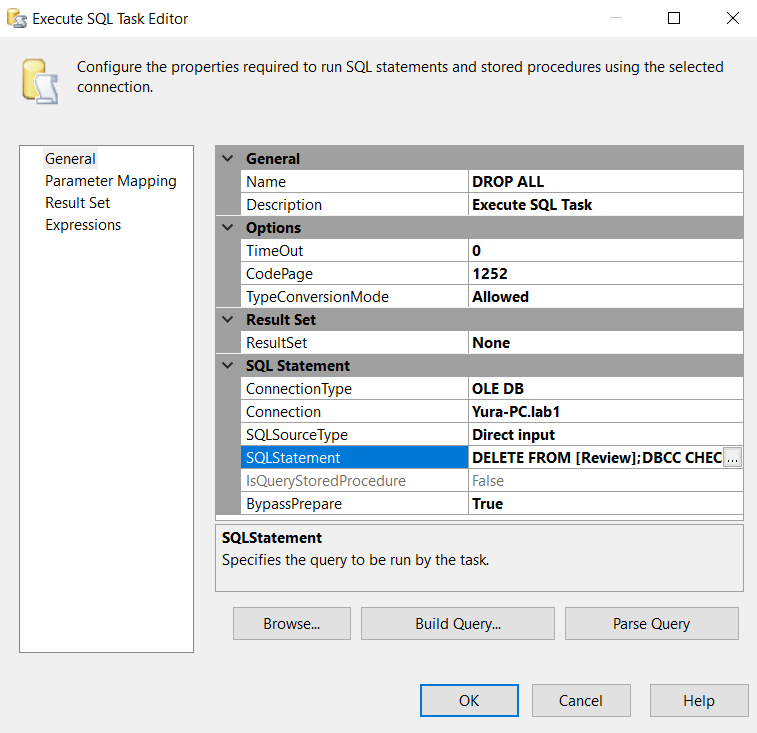


Послідовність виконання логіки:

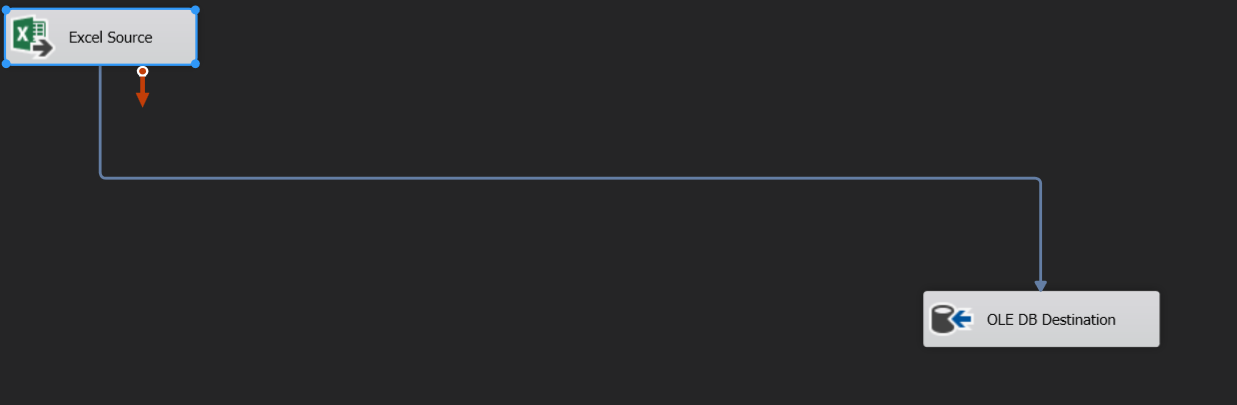


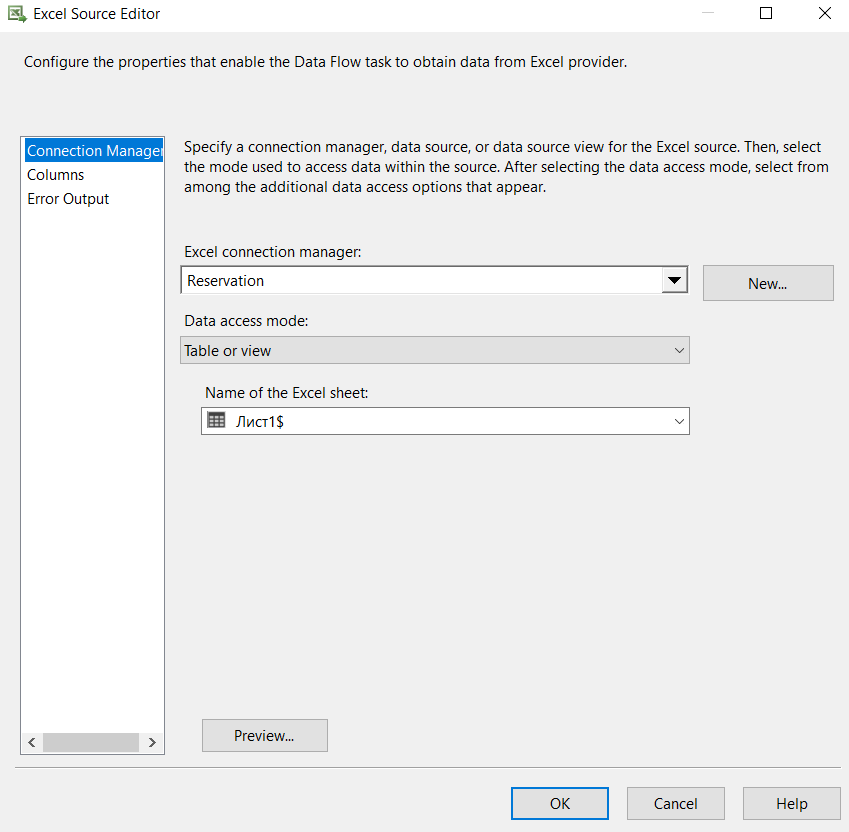
Заради збереження цілісності індексів та унікальних ключів у таблицях – перед запуском виконується команда, що очищає усі таблиці і занулює PKs.

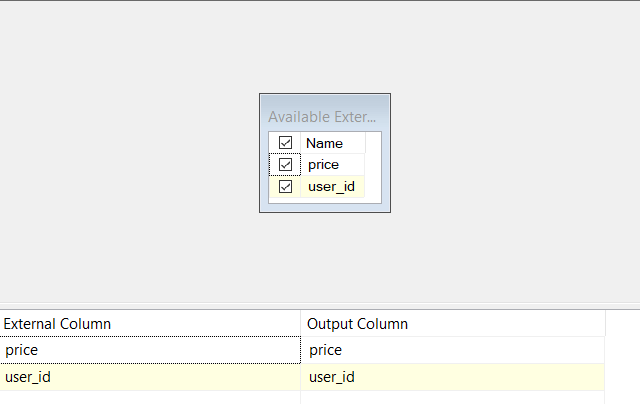
DELETE FROM [Review];  
DBCC CHECKIDENT('Review', RESEED, 0);  
  
DELETE FROM [RESERVATION\_HOTEL];  
  
DELETE FROM [Cruise];  
DBCC CHECKIDENT('Cruise', RESEED, 0);  
  
DELETE FROM [Entertainment];  
DBCC CHECKIDENT('Entertainment', RESEED, 0);  
  
DELETE FROM [Restaurant];  
DBCC CHECKIDENT('Restaurant', RESEED, 0);  
  
DELETE FROM [Room];  
DBCC CHECKIDENT('Room', RESEED, 0);  
  
DELETE FROM [Hotel];  
DBCC CHECKIDENT('Hotel', RESEED, 0);  
  
DELETE FROM [Transfer];  
DBCC CHECKIDENT('Transfer', RESEED, 0);  
  
DELETE FROM [TransportationCompany];  
DBCC CHECKIDENT('TransportationCompany', RESEED, 0);  
  
DELETE FROM [Reservation];  
DBCC CHECKIDENT('Reservation', RESEED, 0);  
  
DELETE FROM [TripAdvisorUser];  
DBCC CHECKIDENT('TripAdvisorUser', RESEED, 0);

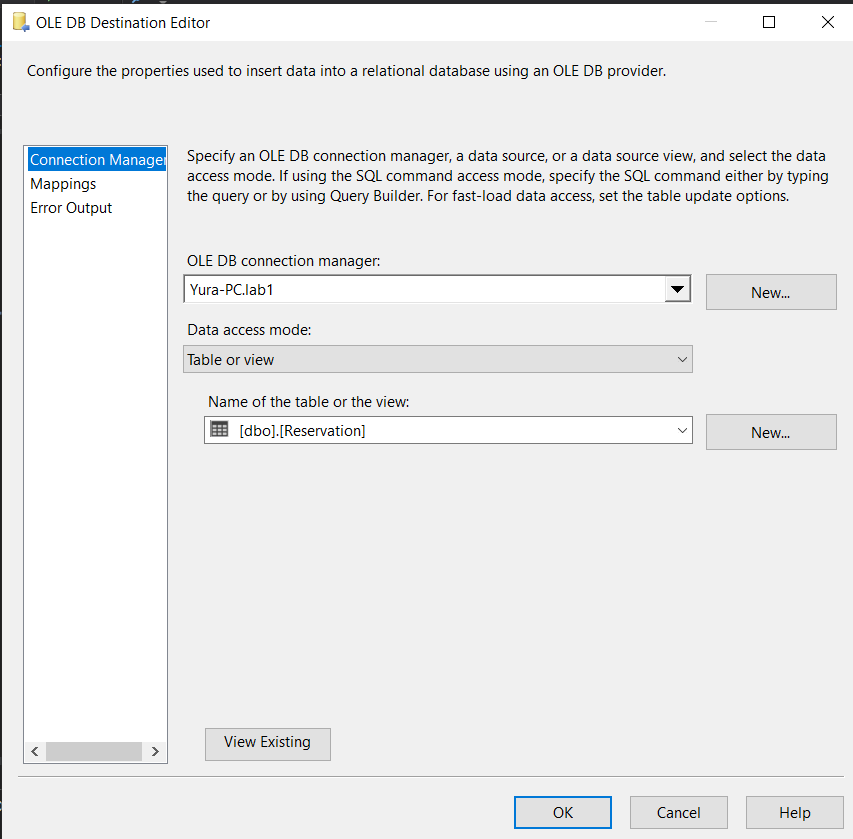


Загальна структура кожного інсерту – витягування даних із ексель файлу (перша колонка містить назви колонок у таблиці для зручнішого маппінгу) і послідовне занесення даних у відповідну таблицю.

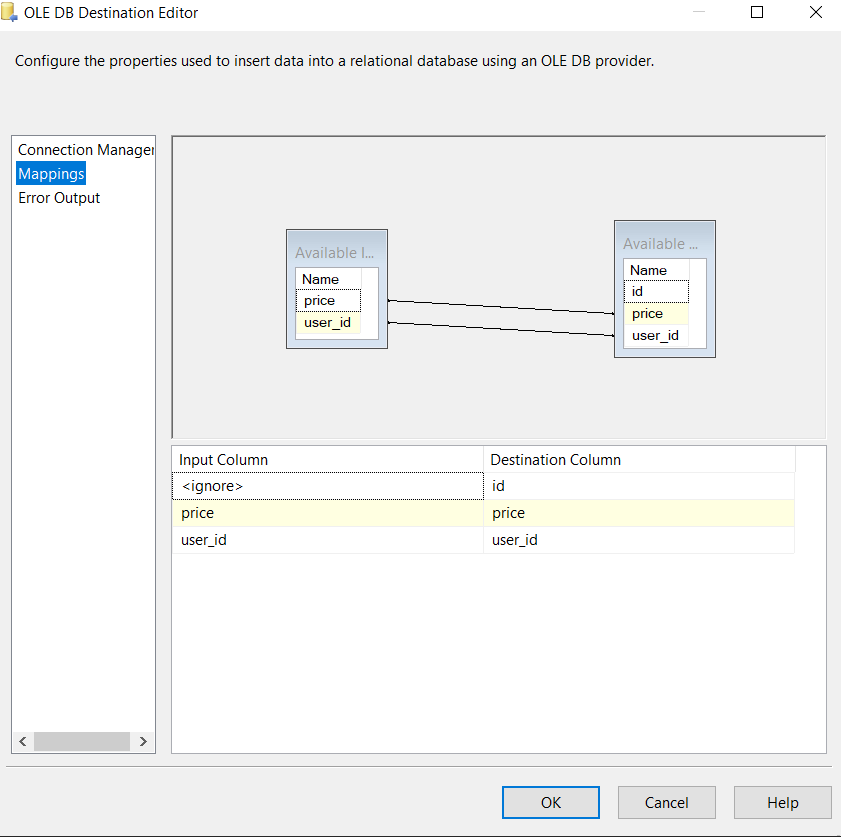






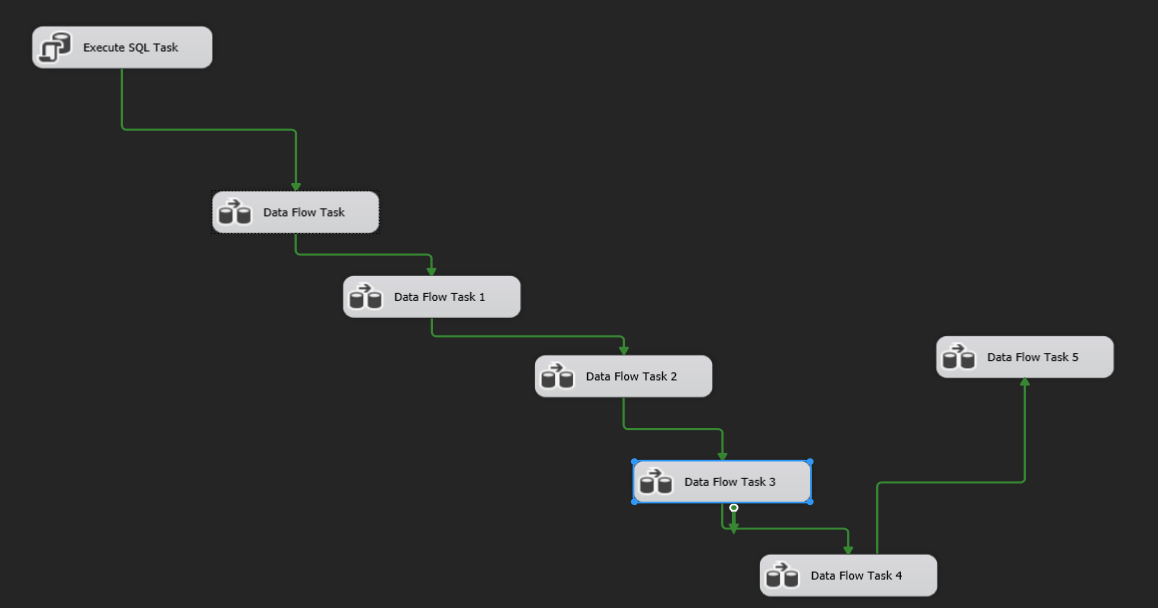


Оскільки айдішка у нас автоінкрементується, то вона не міститься у ексельках та ігнорується як поле в принципі.

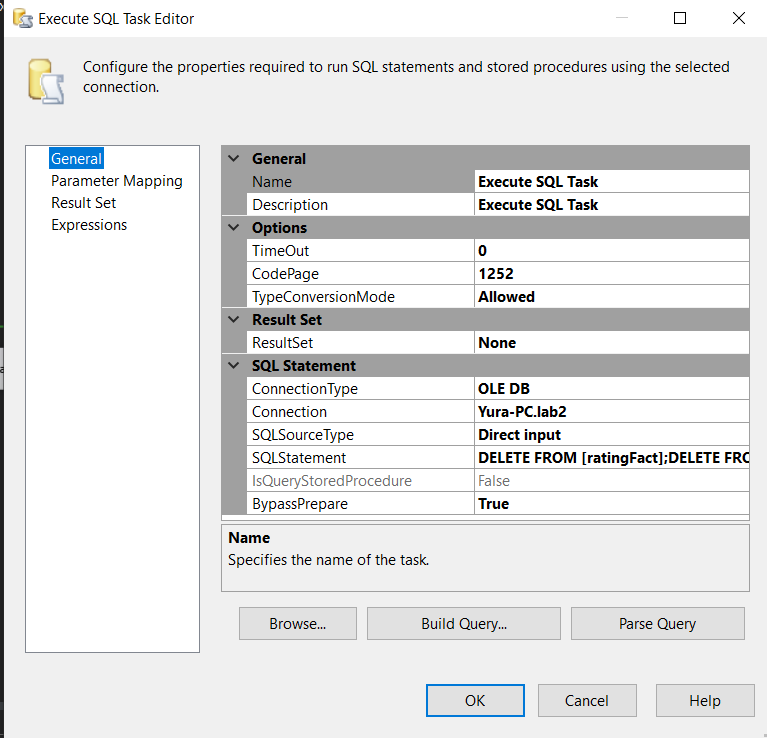




Другий SSIS пакет створений для заповнення БД у вигляді зірки на основі даних із 1 БД.

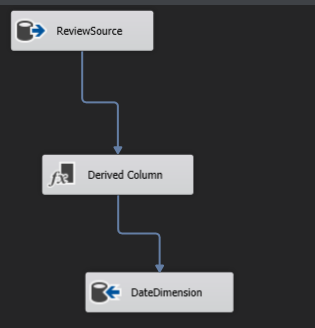


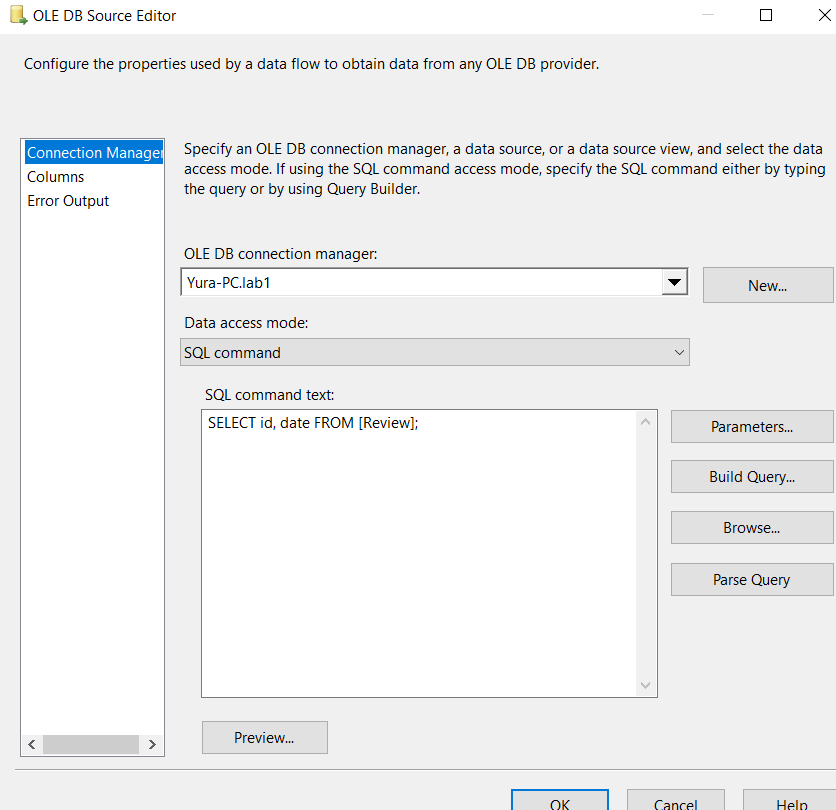
По аналогії із попереднім пакетом – першим кроком ми видаляємо усі дані із наших таблиць.

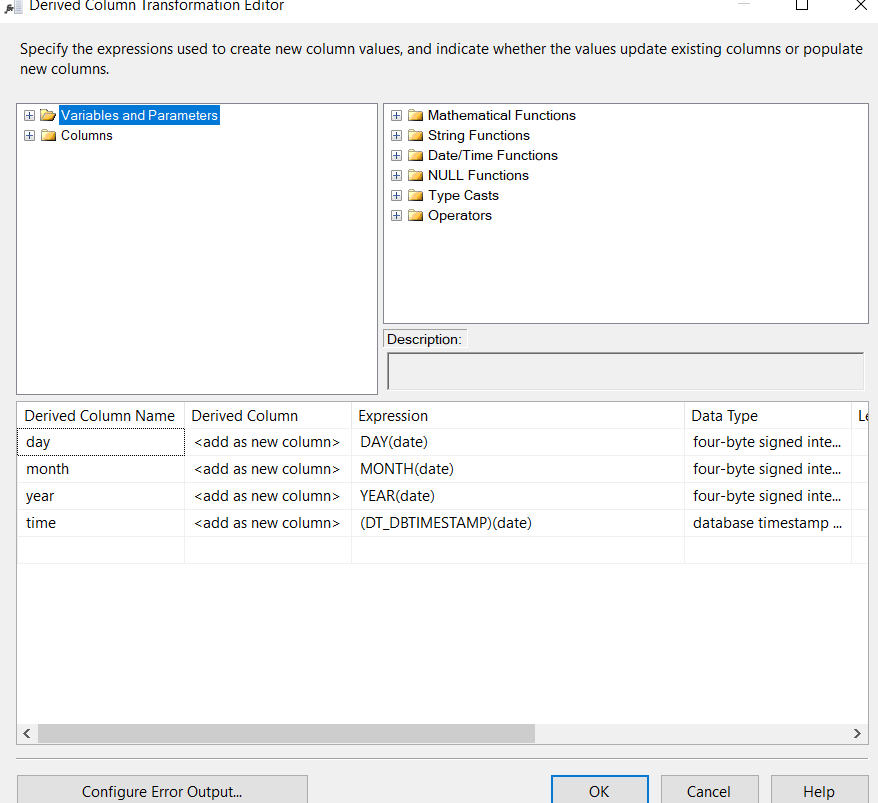


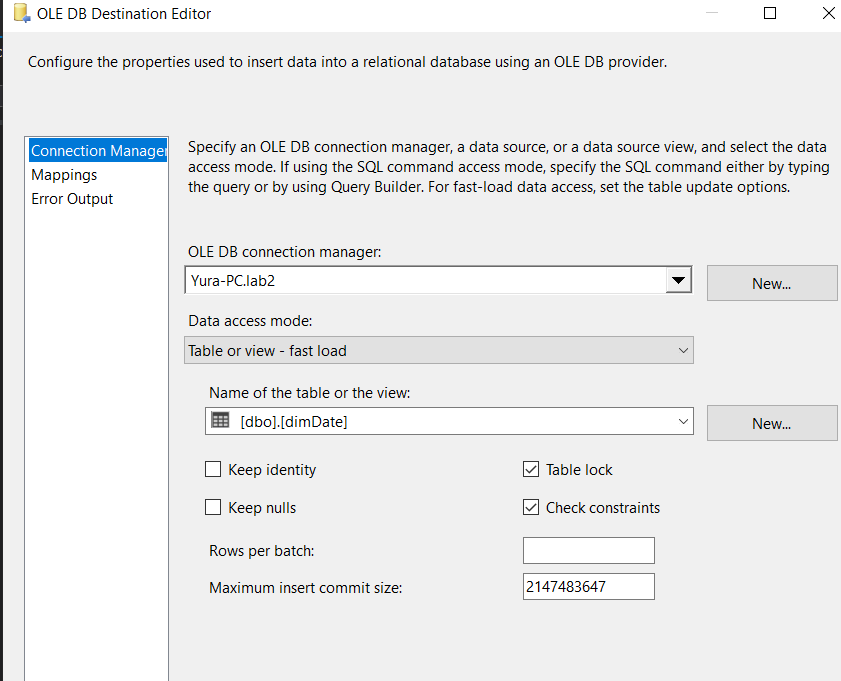
DELETE FROM [ratingFact];  
DELETE FROM [dimDate];  
DELETE FROM [dimLocation];  
DELETE FROM [dimReservation];  
DELETE FROM [dimReview];  
DELETE FROM [dimUser];

Типовий вигляд кожного дата флоу таску – із джерела даних (таблиця із 1 БД) ми отримуємо наші дані, групуємо чи перетворюємо їх і заносимо у відповідний діменшн у 2 БД.









Для агрегації даних і заповнення статистичної таблиці фактів – ми використовуємо даний запит , обраховуючи основні, а також допоміжні (середній рейтинг готелю та загальна сума резервацій) поля.

use lab1;  
select TripAdvisorUser.id as user\_id,  
 Reservation.id as reservation\_id,  
 Review.id as review\_id,  
 Review.hotel\_id as location\_id,  
 Review.id as date\_id,  
 SUM(Reservation.price) as summary\_income,  
 AVG(Hotel.rating) as avg\_hotel\_rating  
from Reservation  
 Left OUTER JOIN RESERVATION\_HOTEL  
 ON Reservation.id = RESERVATION\_HOTEL.reservation\_id  
 RIGHT JOIN Review  
 on Review.reservation\_id = Reservation.id  
 FULL JOIN TripAdvisorUser  
 on Review.user\_id = TripAdvisorUser.id  
 JOIN Hotel  
 on Review.hotel\_id = Hotel.id  
where Review.hotel\_id IS NOT NULL  
group by TripAdvisorUser.id, Reservation.id, Review.id, Review.hotel\_id  
order by summary\_income asc;

