IÇINDEKILER

DATA WARE HOUSE (VERİ AMBARI) KURULUMU	2
DATA WARE HOUSE:	2
AMACIMIZ:	2
DATABASE KURULUMU:	2
SSIS PROJE OLUŞTURMA:	3
PROJE AYARLARI:	3
PROJE CONNECTION MANAGER:	4
SSIS DATA WARE HOUSE KURULUMU:	4
TABLO TEMİZLEME:	5
TABLO AKTARIMI:	6
SLOWLY CHANGING İLE KAYIT AKTARMA:	8
1-DEĞİŞKEN TANILMALA:	8
2- DEĞİŞKENİ SORGUYA EKLEME:	10
3- SLOWLY CHANGING OLUŞTURMA:	11
4- SİLİNEN KAYITLAR:	13
SSIS PROJE KAYDETME-BUILD-DEPLOY İSLEMLERİ:	17

DATA WARE HOUSE (VERİ AMBARI) KURULUMU

DATA WARE HOUSE:

Veri ambarı, iş zekâsı (BI) faaliyetlerine, özellikle de analitiğe olanak tanımak ve bunları desteklemek üzere tasarlanmış bir veri yönetim sistemidir. Veri ambarları yalnızca sorgulama ve analiz amacıyla kurulur ve çoğu zaman geçmişe ait büyük miktarlarda veri içerir. Bir veri ambarındaki veri genellikle uygulama yazılımlarının günlük dosyaları ve işlem uygulama yazılımları gibi çok çeşitli kaynaklardan elde edilir.

Bir veri ambarı, çok sayıda kaynaktan gelen büyük miktardaki verileri merkezi hâle getirir ve birleştirir. Analitik yetenekleri kurumların karar vermeyi geliştirmek için verilerinden değerli iş iç görüleri elde etmelerine olanak tanır. Zaman içinde, veri bilimcileri ve iş analistleri için paha biçilmez bir tarihi kayıt oluşturur. Veri ambarı, bu özellikleri sayesinde bir kuruluşun "tek doğruluk kaynağı" olarak görülebilir.

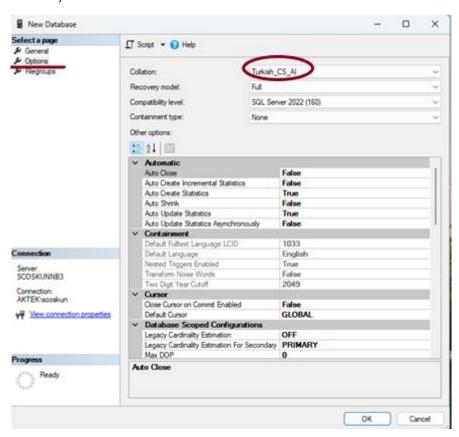
AMACIMIZ:

DWH ı kurma amacımız şubelerimizdeki verileri tek bir çatı altında toplayıp, bu veriler üzerinden raporlarımızı çalıştırmak ve bu sayede canlı database'i işlem yoğunluğuna tutmamış olup sadece raporlar için kullanabileceğimiz bir database yapmaktır.

Bu işlemler için taşıma, değiştirme ve silme işlemlerini SSIS yardımcılığıyla yapacağız

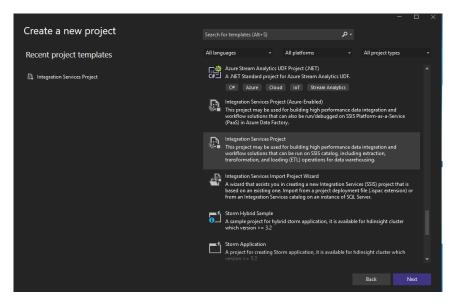
DATABASE KURULUMU:

SQL tarafında oluşturacağımız database için önemli olan hususlardan bir tanesi, DWH için kuracağımız database dili ile ana database dilinin aynı olması gerekmektedir. Bu ayarı New Database işlemi esnasında Options sekmesinde collaction alanında olması gereken dili seçeriz.



SSIS PROJE OLUŞTURMA:

Visual Studio 'da Create a new project or solution seçeneğinden "Integration Services Project" ile yeni projemizi oluştururuz.

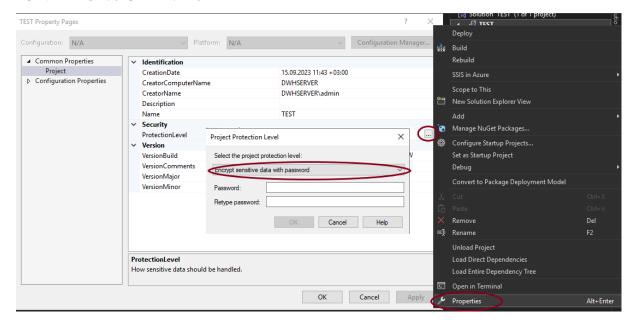


PROJE AYARLARI:

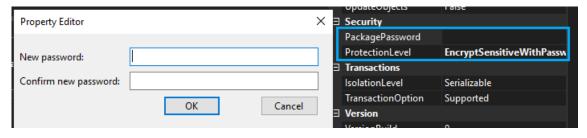
Projemizi birden fazla kişinin kullanacağına varsayarsak şifreli giriş önlemi almak isteyebiliriz.

Bunun için hem projemizin hem de paketimize güvenlik ayarlarımızı değiştirmemiz gerekmektedir.

İlk olarak projemize sağ tıklayarak properties -> ProtectionLevel -> Encrypt sensitive data with password'u seçeriz. Bu özellik ile projemize giriş yapılacak şifreyi belirleriz.



Control flow sayfasına gelerek sağ tık-> properties -> ProtectionLevel -> Encrypt sensitive data with password'u seçeriz PackagePassword ile aynı işlemleri yapmamız gerekmektedir.

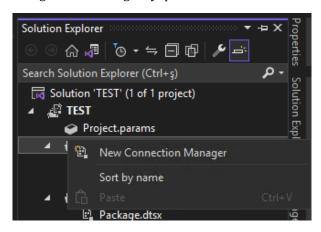


Bu ayarlar ile projemize şifre ile giriş özelliği sağlamaktayız.

PROJE CONNECTION MANAGER:

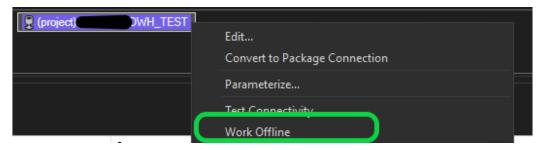
Projemizde kullanacağımız databaseleri tanımlayacağımız alandır.

Solution Explorer'dan Connection managers alanına sağ tık yaparak New connection manager seçilir.



Çıkan seçeneklerde SQL tarafında çalışacağımız için OLE DB seçeriz. New sekmesinden bağlamak istediğimiz databaselerin bilgilerini girdikten sonra connectionlarımız gelecektir. Biz bu alanda ana db miz ve hedefte çalışacağımız db 'yi ekleriz.

Daha sonra connection manager sekmesine gelen bağlantılarımızın ayarlarında work offline işaretli olmaması gerekmektedir.



Bu işlemlerden sonra projemiz aktarım yapısını kurmak için uygun hale gelmiştir.

SSIS DATA WARE HOUSE KURULUMU:

Aktarım için iki adet databasemiz olacaktır.

- Ana db: Canlı işlemlerin yapıldığı kaynak db miz.
- Dwh db: Kaynaktaki verileri yazdıracağımız hedef db miz.

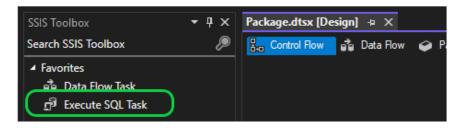
Özet olarak işlem sıralamamız şu şekilde olacaktır.

- 1. Kaynaktan aktaracağımız tablolarımızı belirleriz.
- 2. Tablo temizleme: Control Flow dan Execute SQL Task ile belirlediğimiz tabloları dwh db için truncate cümlelerini yazarız. Bizim izleyeceğimiz yolda büyük tablolarımızı bu truncate cümlelerinin dışında bırakırız bunun sebebi sil ve içeri aktar işleminde içeri aktarım sırasında ana db 'de uzun süre meşgul etmesinden kaçınmak istememizdir.
- 3. Tablo Aktarımı: Aktarım yapacağımız tablolarımızı Data Flow Task içerisinden OLD DB Source ile ana db ye bağlanarak okuma yapacağımız tablomuzu seçeriz. OLE DB Destination ile Dwh db de aktarım yapacağımız tablomuzu belirleriz.
- 4. Büyük Tabloların Aktarımı: Büyük tablolarımızın kaynak db yi çok fazla meşgul etmesini istemediğimizden dolayı bu tabloların aktarımını SSIS Data Flow fonksiyonu olan SLOWLY CHANGING DIMENSION yapısını kullanırız.

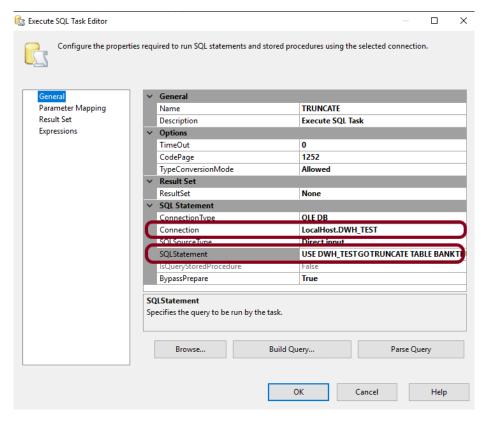
Not: Büyük tabloların aktarımı için direkt olarak temizle ve içeri aktar ile de yapabilirsiniz. Bu işlem sizin alacağınız karara bağlıdır.

TABLO TEMİZLEME:

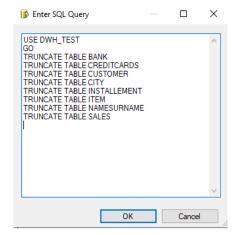
Paketimizi her çalıştırma işleminde ilk olarak ilgili tabloların içerisini temizlemesini istemekteyiz. Bunun için Control Flow içerinden Execute SQL Task 'i kullanırız. Bu Execute SQL Task işlemine <u>TRUNCATE</u> ismini vereceğim.



İlk olarak connection alanından işlem yapacağımız (Dwh db) db 'yi belirleriz. Bu alanda dikkat etmemiz gerekir yanlış bir db seçimi yapmamız ana db üzerinde kayıtların silinmesine neden olacaktır.



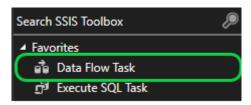
SQL Statement alanından temizleyeceğimiz tabloların SQL komutunu yazacağız.



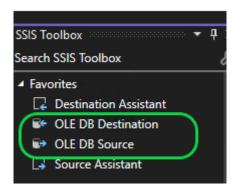
Burada USE ile database belirtme sebebim olası yanlış db seçiminde hata alıp sorgumuzu çalışmasını engellemek amaçlıdır.

TABLO AKTARIMI:

Tablolarımızı kaynaktan hedefe aktarımı yapmak için Data Flow Task kullanırız.

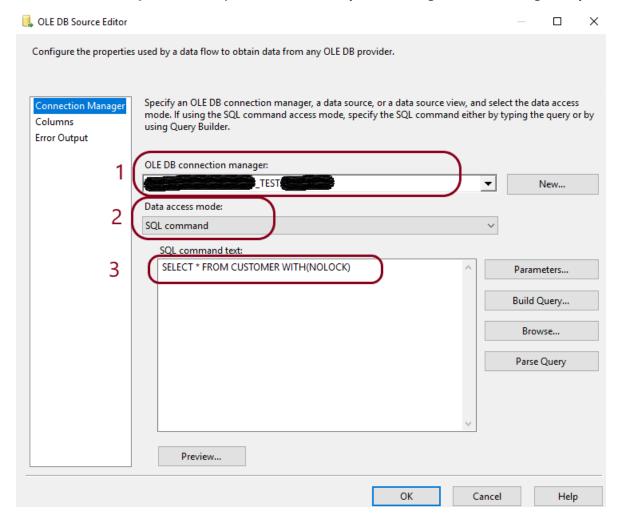


Data Flow Task 'tan kaynaktan veri okumak için OLE DB Source ve hedefe yazdırmak için de OLE DB Destination araçlarını kullanacağız. Data Flow Task için <u>AKTARIM</u> ismini kullanacağım.

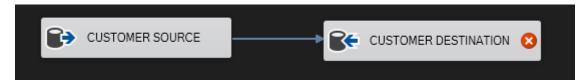


OLE DB Source ile kaynaktaki veriyi okumamızı sağlayacaktır. İçerisine girdiğimizde OLE DB Connection manager alanından kaynak db'i seçeriz.

Data Access mode alanından SQL command seçeriz ve daha sonra kaynaktan alacağımız tablonun sorgusunu yazarız.



Source işlemini bitirdikten sonra Source işleminden çıkan ok ile Destination a bağlarız.

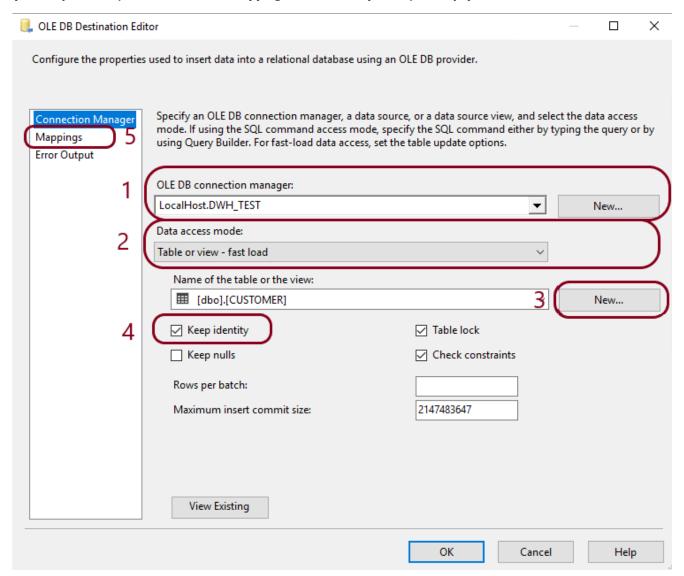


OLE DB Destination kaynaktan okutulan veriyi hedef db ye yazdırmamızı sağlayacaktır. İçerisine girdiğimizde OLE DB Connection manager alanından hedef db 'i seçeriz.

Data Access mode alanında Table or view -fast load seçeneğini seçeriz.

Name of the table or the view alanında eğer aktaracağımız tablo dwh database içerisinde yoksa New butonuyla yeni tablomuzu SSIS aracılığıyla oluştururuz. Eğer tablo DWH db de varsa ilgili tablomuzu seçeriz.

Keep identity alanını işaretledikten sonra mappings alanından mapleme işlemini yaparız.



Bu işlemleri aktarmak istediğimiz her tablo için ayrı ayrı yaparız.

NOT: Büyük tabloları ayırmak istemeyip tüm tabloları bu yapıda işleme almak isteyenler için işlemler buraya kadardır.

16. Sayfa da paketi kaydetme, build ve deploy işlemlerine geçebilirler.

SLOWLY CHANGING ILE KAYIT AKTARMA:

Kayıt sayısının çok olduğu tabloların aktarımında, aktarma işlemlerinin ana db yi uzun süre meşgul edeceği için pek tercih ettiğimiz bir yöntem değildir.

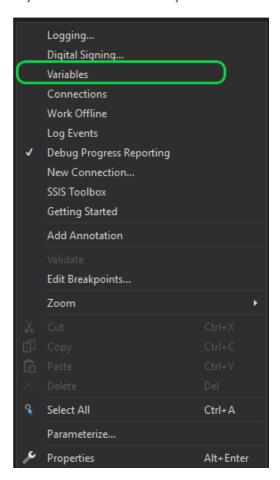
Bu yüzden kaydın çok olduğu tabloları Data Flow Task Fonksiyonu olan SLOWLY CHANGING DIMENSION mimarisini kullanacağız. Bu yapı ana db üzerinde değişen kayıtların güncellenmesi ve yeni oluşan kayıtların da eklenmesini sağlamaktadır.

İlk olarak gerekli kayıtların okunması için OLE DB Source ile ilgili tabloyu okutma işlemi yaparız. Kaynaktan verileri alırken tüm tabloyu okumasını istemediğimiz için where şartımızda dwh db 'de olan son kayıttan sonraki kayıtları getirmesi için bir değişken tanımlamalıyız.

Bu değişken için tanımlayacağımız kolon, tablomuzda ID kolonu veya herhangi bir güncelleme ve ekleme yapıldığında güncel tutulan last update kolonu olabilir. Biz bu yapı için last update kolonunu kullanacağız.

1-DEĞİŞKEN TANILMALA:

Control Flow sayfasına sağ tık ile çıkan seçeneklerden Variables i seçeriz



Açılan sayfada paketimize tanımlayacağımız değişkenimizin ismini ve data type ı belirleriz.

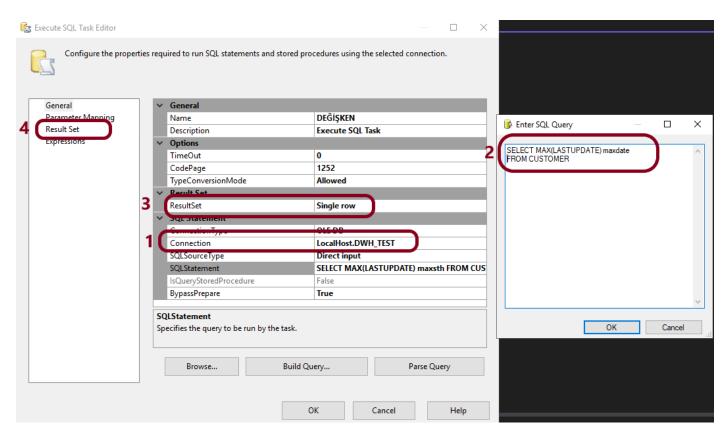


Bu değişkene değer tanımlamak için Execute SQL Task ile DWH db 'ye bağlanarak son ID değeri ya da son last update tarihli kaydı getiren sorguyu yazarız. Bu Execute SQL Task için <u>DEĞİŞKEN</u> ismini kullanacağım.

Sorgumuz tek değer döneceği için Resul set alanında Single Row olarak seçilmelidir.

Bu sorguyu max komutu ile en yüksek değeri getiririm. Sorgumuzda gelen değerin Alias ı mutlaka değişken ile aynı isim olmalıdır.

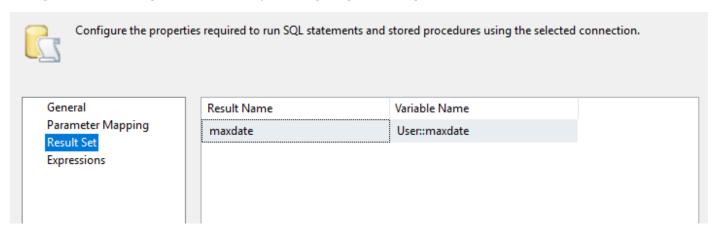
Sorgumuzu ve result set ayarını yaptıktan sonra sol alanda bulunan Result set 'den parametre birleştirme işlemini yaparız.



Result Set alanından Alias için kullandığımız ismi Result Name e yazarız. Variable Name için paketimize tanımladığımız değişkenimizi seçeriz.

Bu işlemler ile paketimizde oluşturduğumuz değişkene yazmış olduğumuz sorgudan gelen değeri getirmektedir.

Bu değer DWH db deki ilgili tablonun son kaydın olduğunu gösteren değerdir.



2- DEĞİŞKENİ SORGUYA EKLEME:

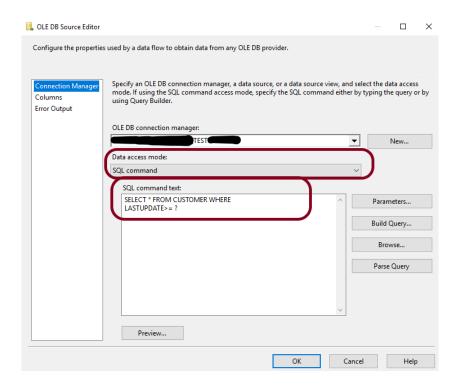
Değişkenlerimizi tanımladıktan sonra SLOWLY CHANGING işlemine özel bir Data Flow Task oluştururuz.

Bu Data Flow Task için **SLOWLY CHANGING** ismini kullanacağım.

OLE DB Source ile kaynaktan veri okuma işlemi yapacağız. Bu işlem için hazırlayacağımız SQL sorgusunda değişkenimizden gelen değer ile where şartı uygulayacağız.

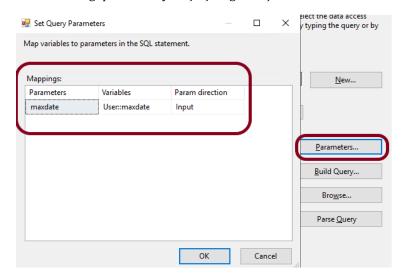
Değişkenimizi sorguya tanımlamak için SSIS de "?" işareti kullanılır.

ÖR: WHERE LASTUPDATE >= ? Bu sorgu ile değişkenimize tanımladığımız DWH db 'de max last update 'den büyük ve eşit değerleri getirmektedir.



Parametremizi tanımlamak için OLE DB Source içerisinde sağ alanda bulunan Parameters... 'e tıklarız.

Açılan pencerede Parametre adını ve hangi parametreyle eşleşeceğini seçeriz.

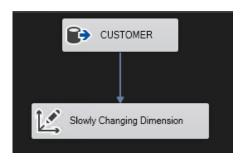


Bu işlemleri yaptığımızda artık ana kaynaktan bizde var olmayan ya da güncellenmiş kayıtları getirmiş olacağız bu sayede kaynaktan daha az veri okuma işlemi yapmaktayız.

3- SLOWLY CHANGING OLUŞTURMA:

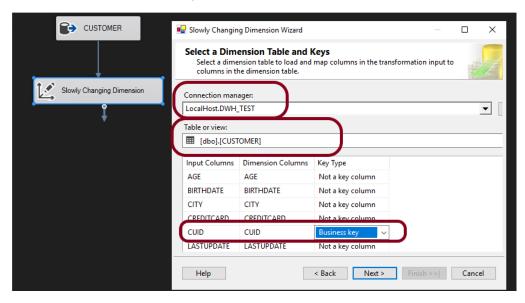
Data Flow içerisinden SSIS Toolbox menüsünde yer alan Slowly Changing Dimension yapısını kullanacağız.

OLE DB Source ile yaptığımız kaynaktan veri okuma işlemini Slowly Changing işlemimize bağlarız.

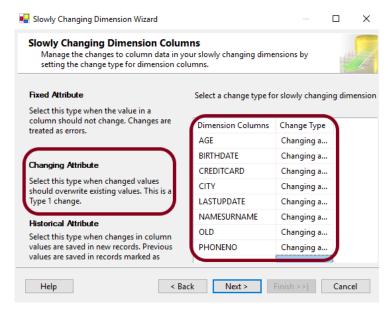


Slowly Changing içerisine girdiğimizde Source ile okuttuğumuz veriyi hangi veri ile karşılaştıracağımızı seçeriz. Bu yüzden Dwh db 'yi ve ilgili tablomu seçerim. Eşlemede kullanmak için en az bir tane business key seçmemiz gerekir.

İşlemlerimiz tamamlandıktan sonra Next butonuyla devam ederiz.

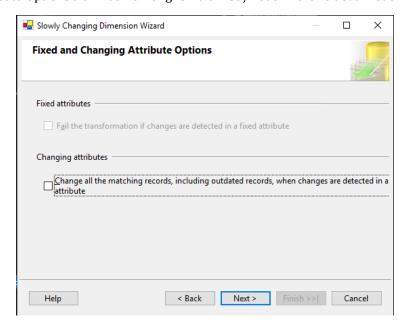


Dimension Columns ekranından Changing Attribute özelliğini kullanacağız. Bu özellik tanımladığımız kolonda kaynakta verinin değişmesi sonucu hedef db (dwh db) 'de değişiklikleri ilgili kolonda uygulamasıdır.

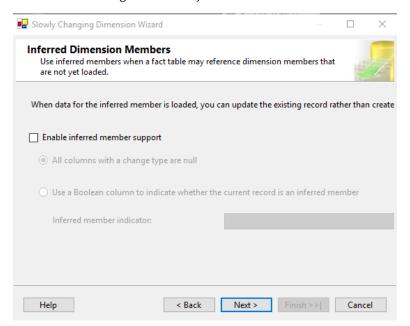


İşlemlerimiz tamamlanınca Next butonuyla devam ederiz.

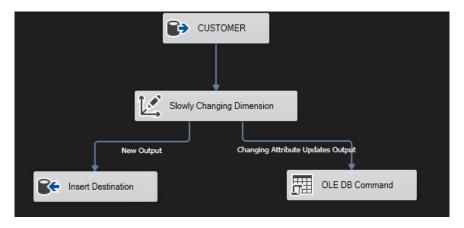
Fixed and Changing Attribute Options alanında herhangi bir alanı seçmeden next ile devam ederiz.



Inferred Dimension Members alanında herhangi bir alanı seçmeden next ile devam ederiz.



İşlemlerimiz bittiğinde Finish seçeneği ile sonlarız.



Slowly Changing Dimension için işlemlerimiz buraya kadar bu yapı sayesinde kaynaktaki değişen ve eklenen kayıtları hedef db 'de güncelleme ve ekleme işlemlerini yapacaktır. Ancak kaynakta silinen kayıtlar için herhangi bir işlem yapmamaktadır bunun için farklı bir yapı kuracağız.

4- SİLİNEN KAYITLAR:

Kaynakta silinen kayıtları hedef db miz olan DWH db 'de silme işlemi için şu adımları uygulayacağız.

- 1- Kaynakta silinen kayıtları trigger ile deleted tablosuna yazdıracağız.
- 2- DWH db 'de sadece bu kayıtlar için bir tablo oluşturacağız.
- 3- Kaynaktaki deleted tablosundaki bilgileri DWH db de oluşturduğumuz tabloya aktaracağız.
- 4- Bu kayıtları DWH db de sileceğiz.

1. Silinen Kayıtlar Tablosu:

Kaynak db de silinen kayıtların ID bilgisinin tutulduğu bir tablo oluştururuz.

```
CREATE TABLE [dbo].[CUSTOMER_DELETED](
        [CUID] nvarchar(36) NOT NULL
)
```

Bizim için önemli olan alan CUID alanıdır bu alan tekil olması gerekmektedir.

Silinen kayıtları trigger ile tablomuza yazdıralım.

```
CREATE TRIGGER [dbo].[Deleted_Customer] on [dbo].[CUSTOMER]

AFTER DELETE
AS BEGIN

DECLARE

@Cuid nvarchar(36)
select @Cuid=CUID
from deleted

insert into [CUSTOMER_DELETED] (CUID) values (@Cuid)
```

END

Trigger ile silinen kayıtlarımızı artık deleted tablomuza yazdırıyoruz.

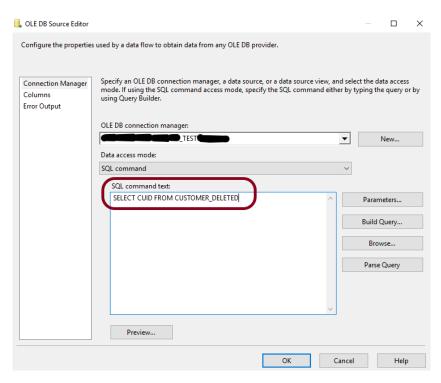
2.DWH db 'de Deleted Tablosu:

Hedef DB miz olan DWH db 'ye bu kayıtları aktarmak için bir tablo oluştururuz.

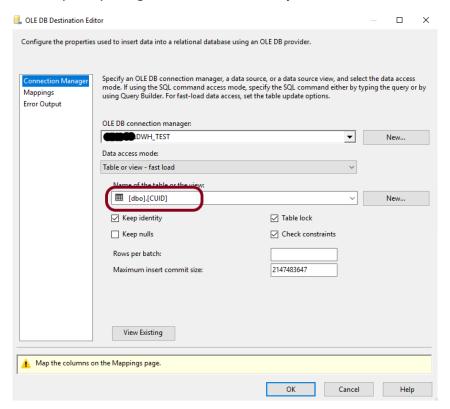
3. Silinen Kayıtları Aktarma:

SSIS paketimizde <u>AKTARIM</u> adı verdiğimiz küçük tabloların aktarımı için oluşturduğumuz Data Flow Task içerisine gelerek yeni bir Source ve Destination oluştururuz.

Source işleminde kaynaktan veri okumak için CUSTOMER_DELETED tablomdan sadece ilgili alanım olan CUID alanını okuturum.



Destination işlemimde DWH db için oluşturduğumuz deleted tablomuza yazdırırız.



Mapleme işlemimiz bittikten sonra artık silinen kayıtların tekil anahtarlarını DWH db 'ye aktarmış olduk.

4. Kaynakta Silinen Kayıtları DWH db 'de Silme:

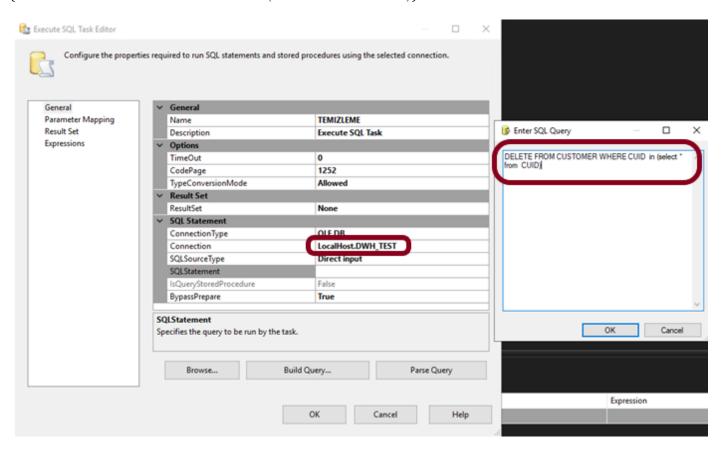
Kaynaktan hedefe aktardığımız silinen kayıtların tekil bilgileri ile DWH db üzerinden bu kayıtları silme işlemi yapan sorgumuzu yazarız.

Bunun için SSIS 'de Data Flow Task üzerinden Execute SQL Task kullanacağız. Bu işleme <u>TEMIZLEME</u> adı vereceğim.

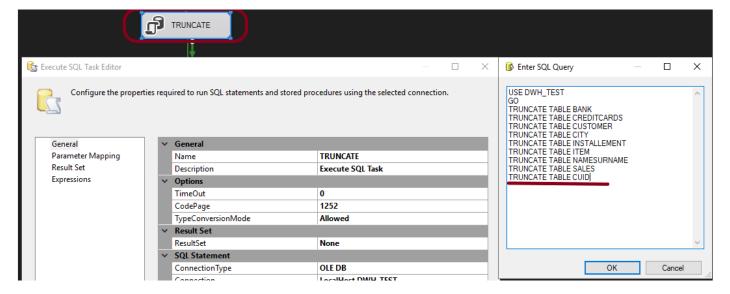
Connection olarak hedef db olan DWH db 'yi seçerim.

SQL Statement alanına silme işlemi için gerekli SQL kodumu yazarım.

(DELETE FROM CUSTOMER WHERE CUID in (select * from CUID))

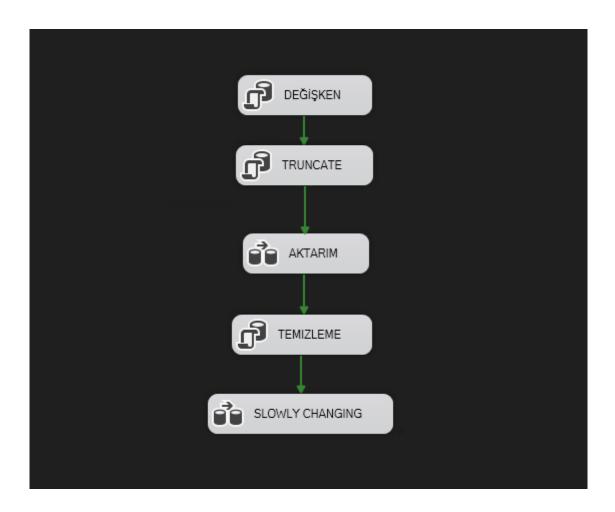


Bu işlemimiz bittiğinde DWH db de oluşturduğumuz tabloların temizlenmesi için <u>TRUNCATE</u> adı verdiğimiz Execute SQL Task in içerisine deleted tablosunu sorgumuzun içerisine ekleriz.



SON:

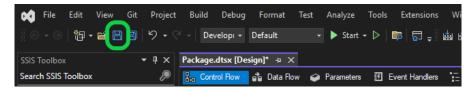
SSIS 'de DWH kurulumu için yapacağımız işlemler buraya kadardır. Kurmuş olduğumuz bu yapıda kaynaktan verileri alıp hedef db 'ye yazdırdık büyük tablolarımız için SLOWLY CHANGING mimarisi kullanarak değişen ve eklenen kayıtları getiririz. Silinen kayıtlar için Kanyak db de trigger ile silinenlerin tekil anahtarı olan ID bilgisini bir tabloda tutarak Hedef db ye yazdırırız ve bu kayıtların hedefte silinmesini sağlayan sorgumuzu yazarız.



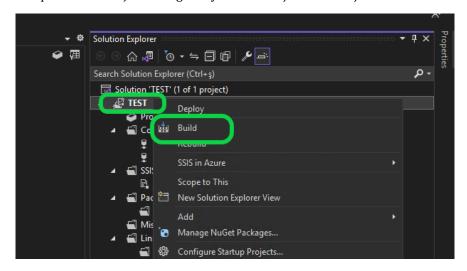
Yapımızın son hali bu şekilde olacaktır.

SSIS PROJE KAYDETME-BUILD-DEPLOY İŞLEMLERİ:

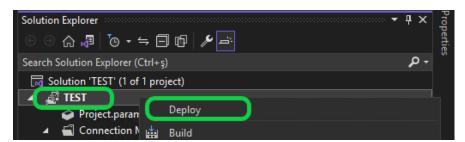
SSIS projemizi kaydetmiş için CTRL+S ya da sağ üst kısımda bulunan kaydetme butonu ile kaydedebiliriz.



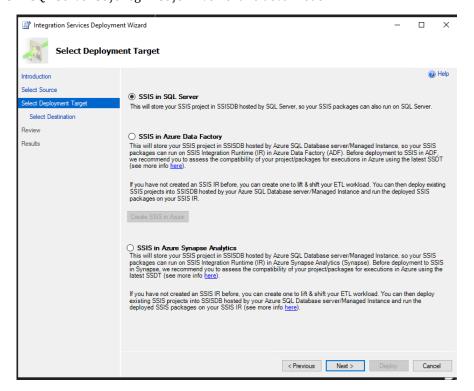
Build işlem için Solution Explorer'dan Projemize sağ tıklayarak build işlemimizi seçeriz.



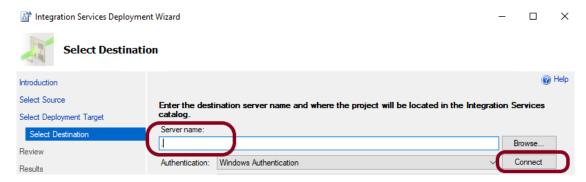
Deploy işlemi için Solution Explorer'dan Projemize sağ tıklayarak deploy işlemimizi seçeriz.



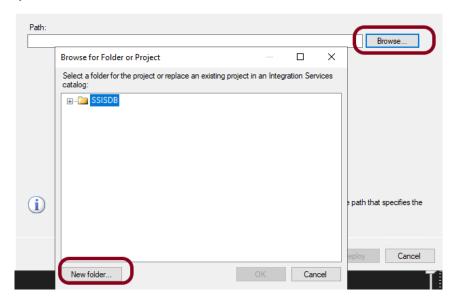
Açılan pencerede SSIS in SQL Server seçeneğini seçeriz ve next ile devam ederiz.



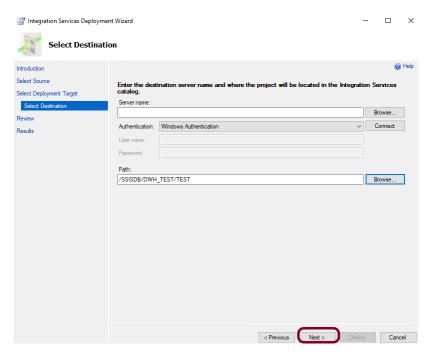
Select Destination sayfasında Server name alanına ". " ya da server name bilgimi yazdıktan sonra connect işlemi yaparız.



Path bölümünden browse 'a tıklayarak deploy işlemini yapacağımız klasörümüzü seçeriz eğer klasörümüz yoksa New folder ile yeni klasör oluştururuz.



Klasörümüzü seçtikten sonra next ile devam ederiz.



Çıkan pencerede deploy u seçerek projemizi deploy etmiş oluruz.

Bu işlemleri yapma amacımız SQL üzerinden SSIS projemizle ilgili oluşturacağımız joblar 'da projemizi kullanmamızı sağlamaktadır.