# **Proje 5: Veri Temizleme ve ETL Süreçleri Tasarımı**

Hazırlayan: Mert Efe Kandemir, Kaan Kaya

Numara: 21290233, 21290436

Teslim: 25.04.2025   
 Github: https://github.com/kayakaan02/-BLM4522-A-Tabanl-Paralel-Da-t-m-Sistemleri-Proje-Videolari/blob/main/proje5.mp4

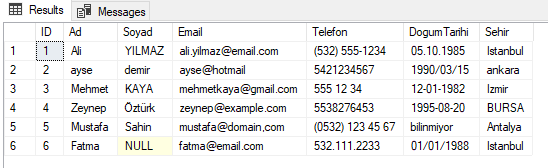
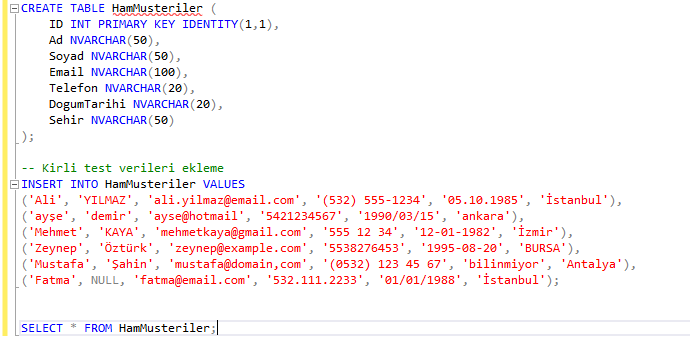
## **1. Giriş**

Bu proje kapsamında, veri temizleme ve ETL (Extract, Transform, Load) süreçleri tasarlanmıştır. Veri analizinde doğru sonuçlar elde etmek için öncelikle verilerin temizlenmesi, tutarsızlıkların giderilmesi ve standart bir formata dönüştürülmesi gerekir. Projemizde, kirli veriler üzerinde bir ETL süreci uygulanarak veri kalitesi artırılmış ve analiz için uygun hale getirilmiştir.

## **2. Veri Temizleme**

Veri temizleme aşamasında, ham verilerdeki eksiklikler, tutarsızlıklar ve formatlarla ilgili sorunlar tespit edilip düzeltilmiştir. İlk olarak örnek bir veri seti oluşturulmuş, ardından bu veri seti üzerinde temizleme işlemleri uygulanmıştır.

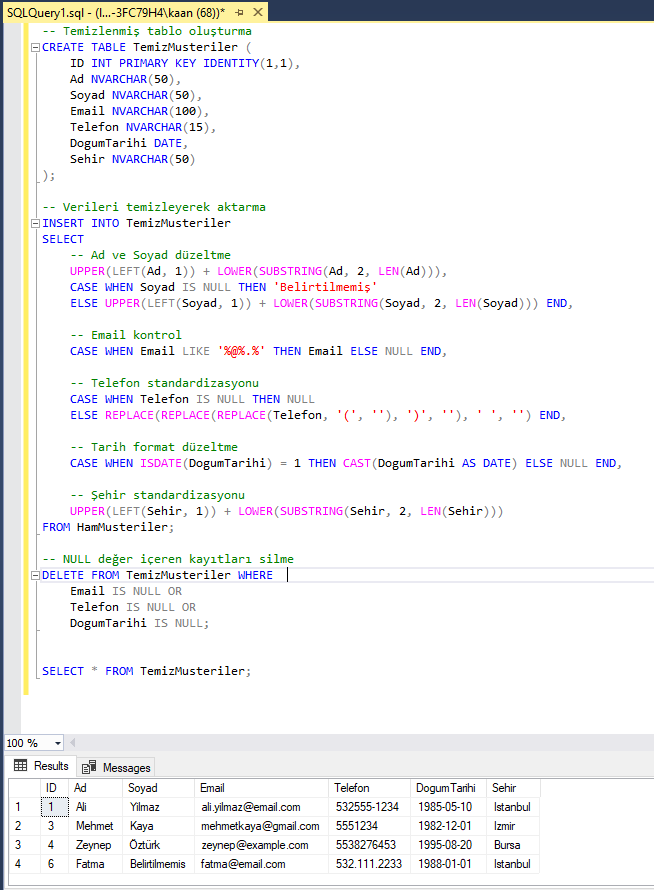
### **Örnek Veri Oluşturma**



### **NULL Değerleri Tespit Etme**

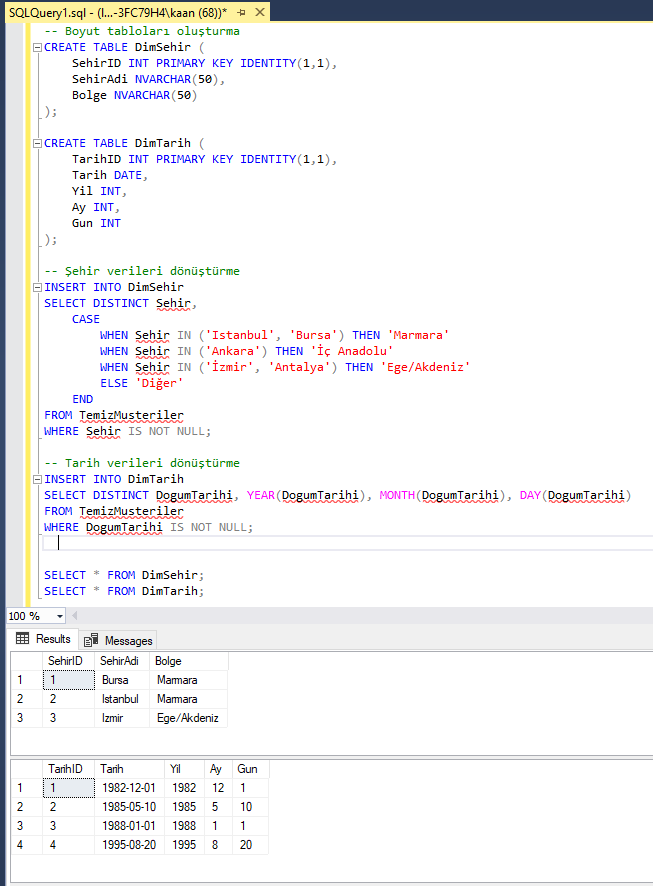
### 

### **Veri Temizleme**



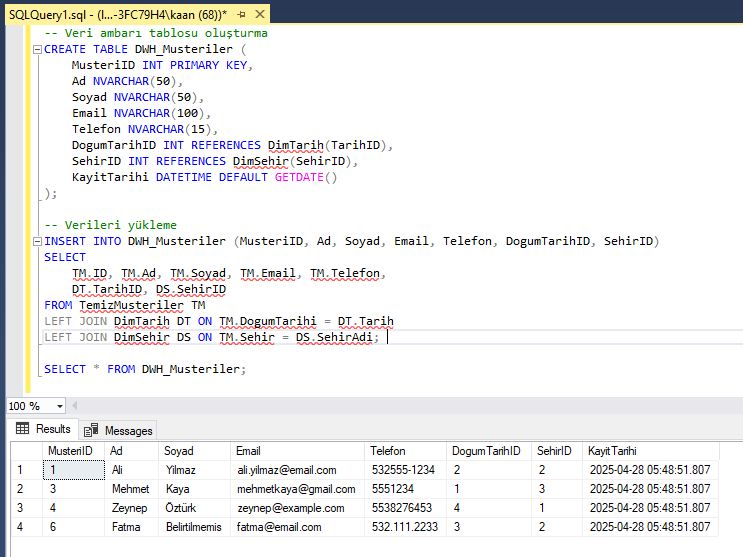
## **3. Veri Dönüştürme**

Veri dönüştürme aşamasında, temizlenmiş verilerin analiz için uygun formatlara dönüştürülmesi gerçekleştirilmiştir. Bu aşamada, boyut tabloları (dimension tables) ve veri küpleri (fact tables) oluşturulmuştur.



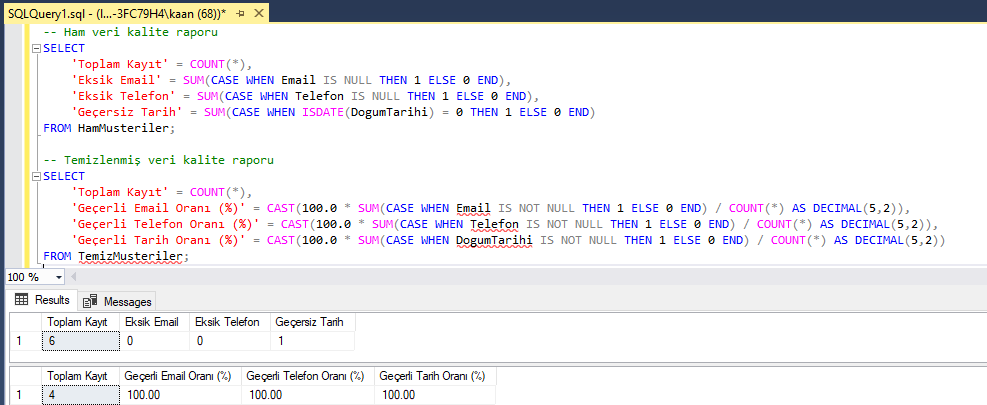
## **4. Veri Yükleme**

Veri yükleme aşamasında, dönüştürülmüş veriler hedef veritabanına yüklenmiştir. Bu projede, boyut tabloları ve veri küpü, OLAP analizleri için uygun bir veri ambarı yapısı oluşturmak üzere kullanılmıştır.



## **5. Veri Kalitesi Raporları**

Veri kalitesi raporları, veri temizleme işlemleri sonrasında verilerin ne kadar iyileştirildiğini göstermektedir. Bu raporlar, veri kalitesi ölçütlerini dikkate alarak hazırlanmıştır.



## **6. Sonuç**

Bu projede, veri temizleme ve ETL süreçleri başarıyla uygulanmıştır. Kirli veriler tespit edilip standart formatlara dönüştürülmüş, boyut tabloları oluşturulmuş ve veri ambarına yüklenmiştir. Sonuçta, analiz için tutarlı ve kaliteli veriler elde edilmiştir.