**1.Aşama – Veri Toplama**

Çevre bakanlığı portalında halka açık sunulan hava kalitesi izleme istasyonları bulunmaktadır. Burada veri bankası > istasyon veri indirme kısmında tipi=Hava Kalitesi, il=Balikesir, istasyon=Balikesir ve son olarak hangi tarih aralığındaki veriler çekilmek istendiği ( son 5 ayı tercih edildi ), hangi özelliklerin olması gerektiği belirtildi. 7 tane özellik seçildi. Bunlar;

1. **PM10 ( µg/m³ )**
2. **SO2 ( µg/m³ )**
3. **NO2 ( µg/m³ )**
4. **O3 ( µg/m³ )**
5. **CO ( µg/m³ )**

Bu kirleticiler seçildi. Havada asılı kalması hava kalitesinin değişimine sebep olur.

1. **Hava Sıcaklığı ( °C )**

Isı birikimi, hava kalitesini etkiler.

1. **Rüzgar Yönü ( Derece )**

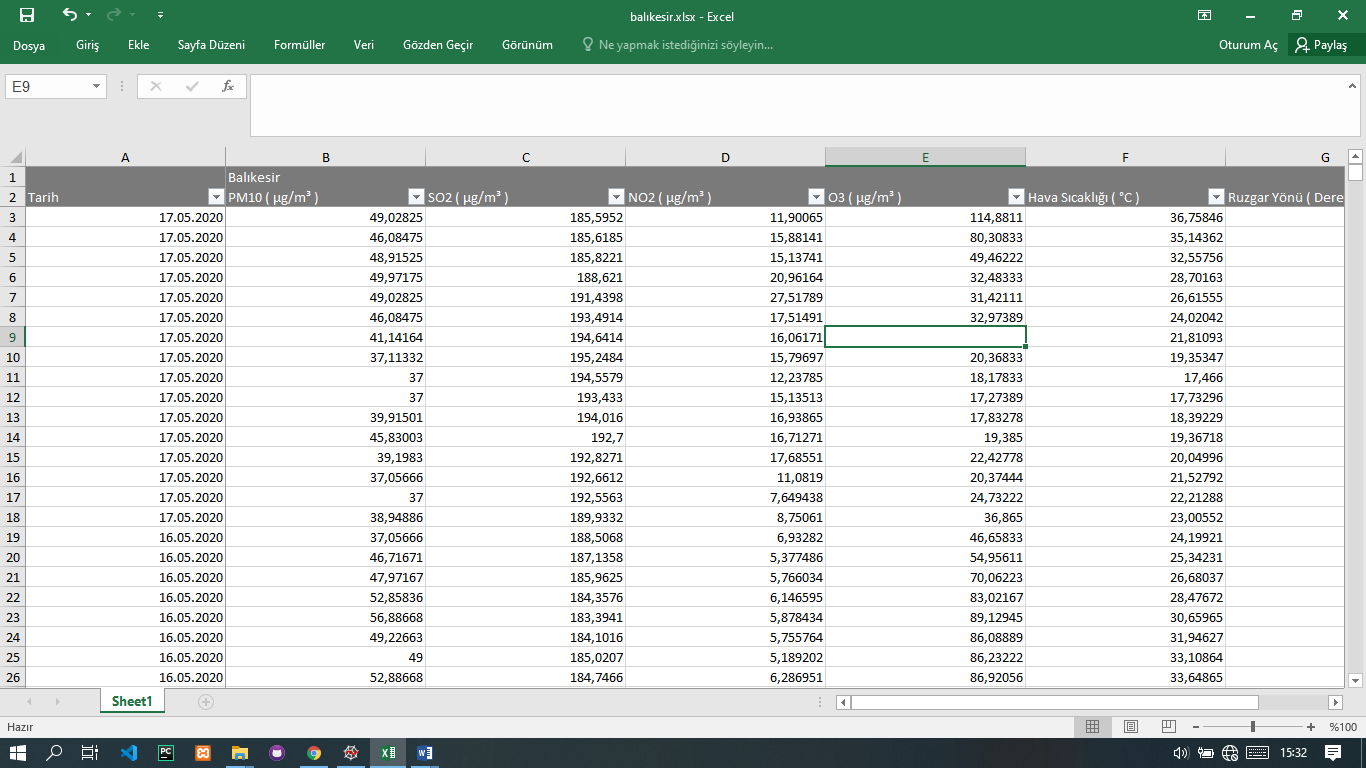
Farklı yönde olması, gökyüzünde kirleticilerin birikmesiyle hava kalitesi bozulur.

1. **Enverziyon**

Yerden yükseldikçe sıcaklık artar. Normalin dışında bir değer olması halinde hava kalitesini olumsuz etkiler.

Bu özelliklerin seçilmesi ile sorgulanan veriseti **Şekil\_1’** deki gibi oluşturulup indirildi. Veri seti 723 satır, 7 sütundan oluşmaktadır. CO ve Enverziyon özellikleri hava kalitesi etkilemesine rağmen eklenmemiş şekilde indi. Verisetinde bulunan **6 özellik** ile Hava Kalite indeksi hesaplanarak **”index”** sütunu oluşturulması hedefleniyor. Sınıflandırma işlemi bu sütuna göre yapılacaktır. Veri setinde eksik verilerin geldiği de gözlemlenmiştir.

**Verisetinin Görünümü**



Şekil\_1