

### **Veri analiz için neden önemli?**

Araştırmacıların ilgilendikleri araştırma sorularına cevap bulabilmeleri için ihtiyaç duydukları kaynağın başında veri gelmektedir. Araştırma sorusunun cevaplanabilmesi için (tabi ki ilgili araştırma uygulamalı bir araştırma ise) verinin elde edilmesi, özelliğine bağlı sınıflandırılarak uygun yöntemin seçilmesi ve ardından analizin gerçekleştirilmesi gerekmektedir. Veri kaynakları kamu kurumları, kooperatifler, özel sektör tarafından toplanıp paylaşılabilir gibi, araştırmacılar da ihtiyaç duydukları verileri toplayabilirler. Tarımsal destekler özelinde bahsedilecek olursa örneğin tarımsal desteklerin üretim üzerindeki etkisini analiz eden bir araştırmacının sahip olması gereken başlıca verileri tarımsal destekler ve üretim verileridir. Ardından araştırma sorularına bağlı çeşitli birçok veri (yaş, parsel sayısı, yıllara bağlı değişimler, hane halkının sahip olduğu çocuğun eğitim düzeyi gibi) analiz içerisinde dâhil edilerek konunun diğer yönleri araştırılabilir.

### **Verilerin Sınıflandırılması**

Veriler kendi içerisinde çeşitli sınıflara ayrılmaktadır. Analiz yönteminin seçiminde ve dolayısıyla politika etkisinin ölçümünde verinin hangi sınıfa ait olduğunun bilinmesi önem arz etmektedir. Bu yüzden öncelikle verilerin hangi sınıflara ait olduğu bilinmelidir.

Veriler analizde kullanılacak değişken açısından düşünüldüğünde sahip oldukları özelliğe göre sınıflandırılabilir. Bu sınıflandırmada aslında veriye bağlı oluşturulan analiz değişkenlerinin sınıflandırılması söz konusudur. Değişken özelliğine göre şeklinde ifade edebileceğimiz bu sınıflandırma başlıca üç sınıfa ayrılabilir. Birinci sınıfta evet/hayır veya olumlu/olumsuz gibi ikili yapıları olan veriler vardır. Tarımsal destekler özelinde üreticinin gelecek dönemde tarıma devam edip etmemesi şeklindeki bir veri ikili yapıya örnek verilebilir. Burada üretici üretime devam edebilir veya etmeyebilir, sadece iki durum söz konusudur. İkinci sınıflandırmada ilk sınıflandırmadan farklı olarak verideki kategori sayısı artar. Üreticinin ekim alanını artırması, azaltması veya aynı düzeyde üretim yapması seçeneklerinden oluşan bir veri seti bu kategorili sınıflandırmaya örnektir. Üçüncü sınıflandırmada ise veri kategorik özellik göstermek yerine süreklilik özelliği gösterir. Örneğin üreticinin gelirinin alabileceği birçok değer olması bu değişkeni süreklilik sınıfına dâhil eder.

Üstteki sınıflandırmanın yanı sıra veriler toplanma düzeylerine göre makro ve mikro olarak ikiye ayrılabilir. Tarım özelinde söylemek gerekirse, Türkiye'nin buğday üretimi makro düzeyde bir veri iken, belirli bir köyde belirli bir üreticinin buğday üretimi mikro düzeyde bir veriyi ifade eder. Verilerin makro ve mikro düzeyde olması kullanılacak analiz seçiminde önemlidir. Geçmiş yıllarda makro düzeyde veriler ile politika analizi gerçekleştirilirken, günümüzde mikro düzeyde veriler daha ayrıntılı bilgiler vermesi açısından daha büyük önem arz etmekte ve sıklıkla araştırmalarda kullanılmaktadır.

Veriler son olarak analiz açısından yatay kesit, zaman serisi ve panel veri olarak üç sınıfa ayrılabilir. Yatay kesit veriler basitçe tek bir yılın verileri olarak tanımlanabilir. Örneğin Türkiye'nin 2015 yılında şehir düzeyinde mısır üretim verisi bir yatay kesit veridir. Yatay kesit verilerin en önemlileri belirli bir dönemde belirli bir bölgede toplanan anket verileridir. Zaman serileri ise genellikle uzun bir zaman dönemini içeren verileridir. Türkiye'nin 1950-2010 yılları arası pamuk üretim verileri zaman serilerine bir örnektir. Panel veriler ise en yalın anlatımı ile yatay kesit ile zaman serisi verilerinin birleştirilmiş haline benzemektedir. Belirli bir bölge veya üreticilerin zaman dönemleri dikkate alınarak elde edilen verileri panel verileri oluşturur. Adana İlinde 2008-2015 yılları arası mısır yetiştiren üreticilerin üretim verileri panel veridir. Günümüzde gelişen veri toplama teknik ve imkânlarına bağlı olarak panel verilerin kullanımı artmıştır. Araştırmacılar ve politika yapımcılar için daha kapsamlı ve çeşitli sorulara cevap verebilmesi nedeniyle günümüzde en fazla tercih edilen veri grubunu oluşturmaktadır.

### **Uluslararası Veri Örnekleri**

Uluslararası bilim camiasında tarımsal desteklerin analizinde sıklıkla kullanılan veri setleri ABD'de ARMS, AB'de FADN'dir. Açılımı Tarımsal Kaynak Yönetim Anketi olan ARMS, ABD'nin tarım ile ilgili en önemli veri setlerinden ve özellikle tarım politikalarının analizinde kullanılan başlıca kaynaklardan biridir. Ülkeyi temsil edecek şekilde seçilen belirli düzeyde üreticiler ile her yıl anket gerçekleştirilmekte ve üreticilerin finansal durumları, kaynak kullanım durumları, demografik bilgileri gibi alanlarda bilgi toplanmaktadır. Açılımı Çiftlik Muhasebe Veri Ağı olan FADN ise AB kaynaklı yapılan politika analizlerinde kullanılan başlıca veri kaynağıdır. Çiftlik Muhasebe Veri Ağı'nda her ülke ortak belirlenen yöntem ve yaklaşımlara göre ülke içi verilerini toplamakta ve ardından veriler ağ içerisinde birleştirilmektedir. Temsili bir şekilde belirli bir sayıda üreticilerden toplanan verilerde ekim alanları, hayvan varlıkları, işgücü durumları, satışlar, üretim masrafları, alınan destekler gibi veriler bulunmaktadır. Her iki veri tabanında da mikro düzeyde veriler karşılıklı anlaşmalar ile araştırmacılara verilmektedir. Ayrıca mikro düzeydeki verilerden oluşturulan özet makro verilerde kamu ile paylaşılmaktadır. AB ve ABD'de belirtilen bu iki ana veri kaynağı dışında kooperatiflerinde topladığı ve tarımsal destekler kapsamında kullanılan veri kaynakları bulunmaktadır. Belirtilen ülkeler dışındaki ülkelerde (örneğin Çin'de) dönemsel olarak toplanan hanehalkı anketleri de tarımsal desteklerin etkisi konusunda kullanılabilir. Araştırmacılar doğrudan kamu kurumlarının

## Veri

topladığı verileri kullanılabilceği gibi, bu verilerin dışında veya bu verilere ek olarak yıllık düzeyde gerçekleştirdikleri saha anketleri ile de belirledikleri araştırma sorularına yanıt aramaktadırlar.

### **Türkiye’de Veri Örnekleri**

Türkiye’de özellikle Gıda, Tarım ve Hayvancılık Bakanlığı tarafından oluşturulan önemli veri kaynakları vardır. Bu kaynakların başında üreticilerin destekleme politikasından yararlanmaları için kayıtlı olmaları gereken Çiftçi Kayıt Sistemi (ÇKS) vardır. Bu sistemde üreticilerin isimleri, yaşları, köyleri, TC kimlik numaraları, ekim alanları, almış oldukları destek verileri bulunmaktadır. Uzaktan algılama ile üreticilerin arazi ve parsellerinin belirlenmesi, hayvan küpelerinin kayıt altına alınması, genç çiftçi veri tabanları gibi araçlarla da son dönemde önemli düzeyde veri toplanmaktadır. Ancak içlerinde ÇKS’nin temel olduğu söylenebilir. Böylelikle ÇKS politika etkisi açısından önemli bir başlangıç noktası olabilir.

ÇKS’nin üreticilerin kişisel bilgilerini içermesi nedeniyle paylaşımında çeşitli yasal kısıtlar bulunmaktadır. Bu kısıtlar ülke içerisinde tarımsal desteklerin etkisinin incelenmesi konusunu zorlaştırmaktadır. Örneğin işletme düzeyinde bir etki analizi gerçekleştirilebilmesi için araştırmacının hangi üreticinin ne kadar ekim alanına sahip olduğunu ve ne kadar destek aldığını bilmesi gerekir. Üreticilerin kimlik numaraları bilindiği durumda hangi üreticinin ne kadar ekim alanına sahip olduğu veya destek aldığı çok rahatlıkla bilinebilir. Buradaki sorun kimlik numaralarının üreticilerin kişisel bilgisi olması ve dolayısıyla paylaşımının yasal kısıtları olmasıdır. Ancak bu basit ikilem (hem kişisel bilgilerin korunması hem de politika etkisinin gerçekleştirilmesi) basit birkaç yol ile çözülebilir. Birinci yol araştırmacılar ile paylaşımında üretici kimlik bilgileri yerine üreticiler için oluşturulmuş eşsiz numaralar verilebilir. Böylelikle üreticilerin kişisel bilgileri paylaşılmaz ve araştırmacılar her bir üreticinin kendine özgü numarası ile işletme düzeyinde analiz yapabilirler. Ayrıca bakanlık ve araştırmacılar arasında oluşturulan ikili işbirlikleri kapsamında da (ABD ve AB’de ARMS ve FADN kapsamında olduğu gibi) ilgili bilgiler sadece araştırma kapsamında kullanılmak üzere paylaşılabilir. Bu paylaşımına bağlı olarak araştırmacılar ilgili araştırma sorularına cevap bulmaları için gerekli veriye sahip olabilir. Ayrıca Bakanlık tarafından öncelikli alan olarak belirlenen tarım politikalarının etki analizi konusunda çalışmalar yapılabilir ve gelecek politika belirleme sürecine katkı sağlanabilir. Son olarak Türkiye İstatistik Kurumu’nun yayımlanmış olduğu ve paylaştığı mikro veriler benzeri bir yaklaşım Gıda Tarım ve Hayvancılık Bakanlığı için de geçerli olabilir. Böylelikle arazi parsellerinden, alınan desteklere kadar birçok veri, kişisel bilgiler gizlenerek araştırmacılar ile paylaşılabilir.