

# Yıllık Gelirdeki Farklılıklara Sebep Olan Faktörler

Başak Aydın\*

## Özet

Yıllık geliri etkileyen ve değişikliklere sebep olan birçok faktör bulunmaktadır. Bunlar; cinsiyet, ırk, yaş, eğitim durumu gibi faktörlerdir. Bu çalışmada hem yıllık hem aylık geliri göz önünde bulundurduğumuzda geliri etkileyen yaş faktörü yadsınamaz bir gerçektir. Bu projede UCI Machine Learning Repository sitesinden alınan resmi veriler kullanılarak yıllık geliri etkileyen faktörlerin sonuçlarının tahmin edilip yorumlanması yönünde bir çalışma yapılmıştır. Veri seti üzerinde grafik ve analiz için ufak değişiklikler yapılmıştır.

## 1 Giriş

Yapmış olduğum bu çalışmada araştırma sorum yıllık gelirdeki değişimleri etkileyen faktörler nelerdir sorusuna dayanmaktadır. Bu sorudan yola çıkarak elde ettiğim sonuçların gelir üzerindeki etkisini incelemek ve veriyi analiz etmek projemin konusunu oluşturmaktadır. Araştırmama dair kullandığım veri setini UCI Machine Learning Repository sitesi üzerinden aldım. Bu veri seti nüfus sayımı verilerine dayanarak yıllık gelirin 50.000\$'ı aşmış olmadığını incelemektedir. Veri setim toplam 14 değişken ve 48.842 gözlem içermektedir. Değişkenler age, workclass, fnlwgt, education, education-num, marital-status, occupation, relationship, race, sex, capital gain, capital loss, hours per week, native country'den oluşmaktadır.

### 1.1 Çalışmanın Amacı

Bu çalışmada yaş, etnik grup, cinsiyet, eğitim, medeni hali gibi değişkenlerine göre gelirin farklılık gösterip göstermediği incelenmiştir.

Birçok ülke, kurum ve kuruluşlar iş dünyasında ayrım yapmaksızın fırsat eşitliğine yönelik çeşitli hükümlerde bulunsalar da iş hayatında dolayısıyla ücretlerde çok bariz farklılıklar bulunmaktadır.

Bu farklılıkların hangi değişkenlerden kaynaklandığını tespit ederek, yorumlayıp, bir bakış açısı oluşturup, çözümler sunabilmek bu çalışmanın amacıdır.

---

\*Öğrenci Numarası, [Github Repo](#)

## 1.2 Literatür

İş gücü piyasasında cinsiyet,etnik köken,ırk,yaş,din gibi birçok faktörlerden kaynaklanan ayrımcılığın ABD,Çin,Güney Kore ve bunun gibi birçok güçlü devletlerde bariz bir sorun olduğunu görüyoruz.Ayrımcılık özellikle iş ortamında ücret,iş alma ve terfi konularında yoğunlaşıyor.Bu tarz ayrımcılıklar işgücü piyasalarının gelişimini ve işleyişini aynı zamanda geliri etkileyen en temel sorunlardan birisidir.Kişisel ücret farklılıklarında aynı birimde çalışan işçilerin yaş ve cinsiyetine göre çok büyük farklılıkların olduğu bariz.Eşit işe eşit ücret sağlanması gerektiği halde bu göz ardı ediliyor.Bunun en iyi örneği kadın işçiye aynı işi yapması için erkek işçiden daha az ücret ödenmesi durumudur.Toplumsal cinsiyet eşitsizliği ya da bir başka deyişle cinsiyete dayalı işbölümü temelinde kadınlar erkekler ile aynı eğitim düzeyine ve mesleki kıdeme sahip olsalar bile sadece kadın olmalarından kaynaklı cinsiyetçi ücret ayırımına maruz kalmaktadır.Bunu yıllık gelire vurduğumuzda kadınlardaki yıllık gelir erkeklerdekine göre daha düşük bir meblağ olarak istatistiklere yansıyor. Seher ve KUTANİS (2016)

İş gücü piyasasında cinsiyet dışında medeni durum,yaş,eğitim düzeyi,mesleki tecrübe ve verimlilik gibi iş gücünün özelliğinden kaynaklanan ücret/gelir farklılıkları da bulunmaktadır.Sektör özellikleri ve bölgeler arası ücret farklılıklarından kaynaklanan yıllık gelir farklılıkları da bulunmaktadır.İşgücü piyasalarında işleri ve ücretleri farklılaştıran bir başka unsur da firma ölçeğidir. Hemen her yerde ve zamanda gözlenen bir durum büyük firmalarda ücretlerin küçük firmalara oranla daha yüksek olmasıdır. Büyük firmalarda ücretlerin yüksek olması işleri birbirinden heterojen kılarak ücretlerin farklılaşmasına yol açar. MAKTAV (2019)

Etkin Ücret Teorisine göre işçiyi denetlemenin ve performansını ölçmenin zor olduğu, işçinin hatalı üretim yapmasının maliyetli olduğu, işin belirli bir süre spesifik işyerinde eğitim gerektirdiği durumlarda işverenlerin istekli olarak piyasa denge ücretinin üstünde ücret (etkin ücret ödemesi) verdikleri söylenebilir. Etkin ücret ödemelerinin işleri birbirinden farklı (heterojen) hale getirerek emek piyasalarında ücret farklılıklarına neden olabileceğini belirtmek gerekir. Benzeri özelliklere sahip iki işçiden birisi etkin ücret uygulamasının yapıldığı firma veya işkolunda çalışıyorsa diğerinden daha yüksek ücret alabilmektedir. Blackaby ve Murphy (1991) Bir diğer gelirdeki farklılıkların nedeni ise ırktır.Siyah ve beyaz ayrımı geçmişten bugüne devam eden bir sorundur.Siyah erkeklerin hala birçok işyerinde aynı meslekte çalışan beyaz erkeklerden daha düşük bir ücret aldığını istatistikler birçok kez göstermiştir.Bunun yanı sıra gelirdeki farklılıklar ırksal eğitim eşitsizliği ile de doğru orantılı.Örneğin ABD'deki istatistiklere göre beyaz ailelerin %39'u lisans eğitimi veya yüksek eğitim mezunu iken siyahi ve hispanik ailelerde bu oran %23 ve %17'dir.Irklar arasındaki eğitim durumu farkı iş hayatında daha az ücretle çalışma ya da tercih edilmeme gibi durumlara yol açıyor. Stolzenberg (1975)

Sonuç olarak iş gücü piyasasında karşılaştığımız birçok değişken faktör gelirdeki farklılıkların neyden,hangi değişkenlerden kaynaklandığını anlamamıza ve analiz etmemize yardımcı olmaktadır.

## 2 Veri

Bu çalışmada Barry Becker tarafından 1994 yılında yapılan Nüfus sayımı veritabanından alınan veriler kullanılmıştır. Bu veri seti UCI Machine Learning Repository sitesinde oluşturulmuş bir veri setidir. Veri seti yıllık gelirin bahsedilen yıllarda 42 farklı ülkedeki farklılıklarını gösteren verileri sağlar. Kullanılan veri seti üzerinde grafik analizi için NA gözlemleri kaldırarak ve veri üzerinde ufak düzenlemelerle analiz yapılmıştır.

Tablo 1, seçili yıllık gelirlerin 42 ülke için ortalama değerlerini, minimum, medyan ile maksimum değerlerini ve standart sapmalarını göstermektedir.

Tablo 1’de görüldüğü üzere daha öncede bahsedilen ülkelerdeki çalışanların yıllık gelirini incelerken arasındaki ilişkiyi analiz ettiğimiz “age” (yaş) değişkeni ortalama 38,58’i vermektedir. Özellikle bu yaş ortalamasındaki bireylerin bilgi birikimi ve tecrübe bakımından ileri seviyede olduklarını görüyoruz. Aynı zamanda ortalama yaş 38-39 civarındayken yıllık gelir maksimum 70.000 dolara kadar çıkmaktadır. Bu aykırı gibi gözüken değerler aslında bakıldığında belli bir yaş grubunun yıllık gelir bakımından daha fazla kazandığını, bu bağlamda “age” (yaş) değişkeninin yıllık geliri etkileyen bir faktör olduğunu açıkça ortaya koymaktadır.

Tablo 1: Özet İstatistikler

	Ortalama	Std.Sap	Min	Medyan	Mak
Age	38.58	13.64	17.00	37.00	90.00
capital gain	7663.83	422.97	1055.00	7688.00	15024.00
hours per week	40.44	12.35	1.00	40.00	99.00
yearly income	43024.94	700.03	37000.00	43000.00	70000.00

## 3 Yöntem ve Veri Analizi

Şekil 1’de yıllık geliri etkileyen yaş faktörü üzerinden oluşturulan grafikte 30-50 yaş grubu bireylerin yıllık gelirinin daha fazla olduğu görülmüştür. Bu konudaki Doğu Üniversitesi Dergisinde yayımlanan “Sosyoekonomik Belirleyenler” başlıklı yazıda da değinildiği gibi yaş değişkeni yıllık gelir üzerinde önemli bir etkiye sahip olduğu görülmüştür. @birdal2019tuketici

ANOVA (Analysis of Variance), birçok grup arasındaki farkların istatistiksel olarak anlamlı olup olmadığını değerlendirmek için kullanılan bir istatistiksel modeldir. Genellikle, gruplar arasındaki ortalamaların farklı olup olmadığını belirlemek için kullanılır. ANOVA modeli, bir bağımlı değişkenin (genellikle sürekli bir değişken) birden fazla bağımsız değişken tarafından etkilendiği durumlar için uygundur. Bu model, gruplar arasındaki farklılıkları, grup içindeki değişkenliklerle karşılaştırarak değerlendirir. Varyans analizinin uygulanabilmesi için örneklem ana kütesinin sahip olması gereken bazı özellikler bulunmaktadır. Her bir grubun normal dağılım sergilemesi, örneklerin bağımsız ve varyanslarının eşit olması istenmektedir. ANOVA analizinde kullanılan F testi hipotezi şöyle oluşturulur (Birdal Şenoğlu (2020)):

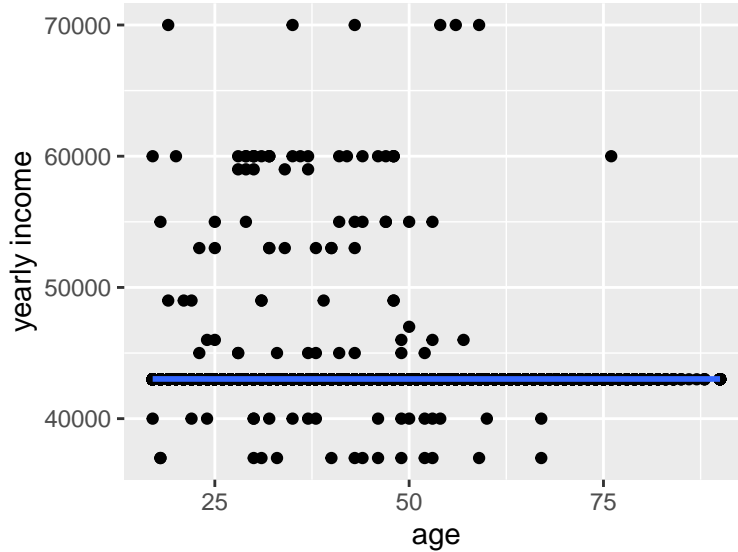
$$H_0 : \mu_0 = \mu_2 = \mu_3 = \dots = \mu_N$$

Bu denklem bize ortalamalar arasında fark olmadığını göstermektedir. 1 ise en az iki ortalama arasında fark olduğunu gösterir. Sonrasında test istatistiği gereklidir. Varyans analizinde temel amaç, ikiden fazla örnek için örnek ortalamalarının genel ortalamadan sapmalarının kareler toplamını, sapmaya yol açan etkileri ayırmak ve analiz etmektir. İki faktörlü varyans analizi modeli şu şekilde gösterilmektedir:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

Burada  $\mu$ , genel ortalamayı;  $\alpha_i$  ilk faktör etkisini;  $\beta_j$  ikinci faktörün etkisini ve  $\alpha\beta_{ij}$  ortak etkiyi ifade ederken,  $\epsilon_{ijk}$  hata terimini açıklamaktadır.

```
veri %>%
ggplot(aes(x = Age, y = `yearly income`)) +
  geom_point() +
  geom_smooth() +
  scale_x_continuous("age") +
  scale_y_continuous("yearly income")
```



Şekil 1: Yıllık Gelir ve Yaş Değişkeni Grafiği

## 4 Sonuç

Bu yıllık geliri etkileyen sebeplerin başında gelen yaş faktörü aslında insanların yaşamları boyunca bilgi birikimlerinin artması, yaptıkları işlerde ustalaşmasından dolayı yıllık geliri

etkileyen faktör olarak süregelenmektedir. Bu çalışmada 42 ülkenin yapılan nüfus sayımı veritabanı verileriyle yıllık geliri etkileyen faktörleri tahmin eden en uygun ANOVA modelinin ortaya konması hedeflenmiştir. Bu doğrultuda UCI Machine Learning Repository adlı siteden alınan veriler kullanılmıştır. 2 adet grubu olması tek yönlü ANOVA modelinin kullanılmasına karar verilmesine sebep olmuştur. Analizde bu iki grup kıyaslanarak aralarındaki fark detaylandırılmak için Tukey analizinden yardım alınmıştır. Yıllık gelir ile ilgili farklı eğilimler olabilir. Bunlar çalışanların yaş aldıkça şirketler tarafından tecrübe ve bilgi birikimi bakımından zengin bulunarak daha yüksek bir maaş verilmesi ya da bulundukları şirkette çalıştıkları yıl arttıkça dolayısıyla çalışanların yaşları da artar, bu da aylık maaşlara ve daha sonrasında da yıllık gelire yansiyarak farklılıkları oluşturabilir. Bu unsurlar bize yıllık geliri etkileyen faktör olan yaşın bize yıllık geliri nasıl etkilediği yönünde bilgi vererek analizi tamamlamamıza yardımcı olur.

**Kaynakça bölümü Rmarkdown tarafından otomatik olarak oluşturulmaktadır. Taslak dosyada Kaynakça kısmında herhangi bir değişikliğe gerek yoktur.**

***Taslakta bu cümleden sonra yer alan hiçbir şey silinmemelidir.***

## 5 Kaynakça

- Blackaby, D. H. ve Murphy, P. D. (1991). Industry characteristics and inter-regional wage differences. *Scottish Journal of Political Economy*, 38(2), 142-161.
- MAKTAV, D. S. (2019). İŞGÜCÜ PİYASASINDA CİNSİYET AYRIMCILIĞININ ÜCRET POLİTİKASINA YANSIMASI VE BİRLEŞMİŞ MİLLETLER ANTLAŞMALARINDA-Kİ YERİ. *Uluslararası Sağlık Yönetimi ve Stratejileri Araştırma Dergisi*, 5(2), 104-117.
- Seher, U. ve KUTANİS, R. Ö. (2016). İşgücü piyasalarında ayrımcılığın kaynakları. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (35), 359-372.
- Stolzenberg, R. M. (1975). Education, occupation, and wage differences between white and black men. *American Journal of Sociology*, 81(2), 299-323.