

21.10.2022

Veri Görselleştirme

Hafta 3: Miktarların Görselleştirilmesi

© Mustafa Çavuş, Ph.D.

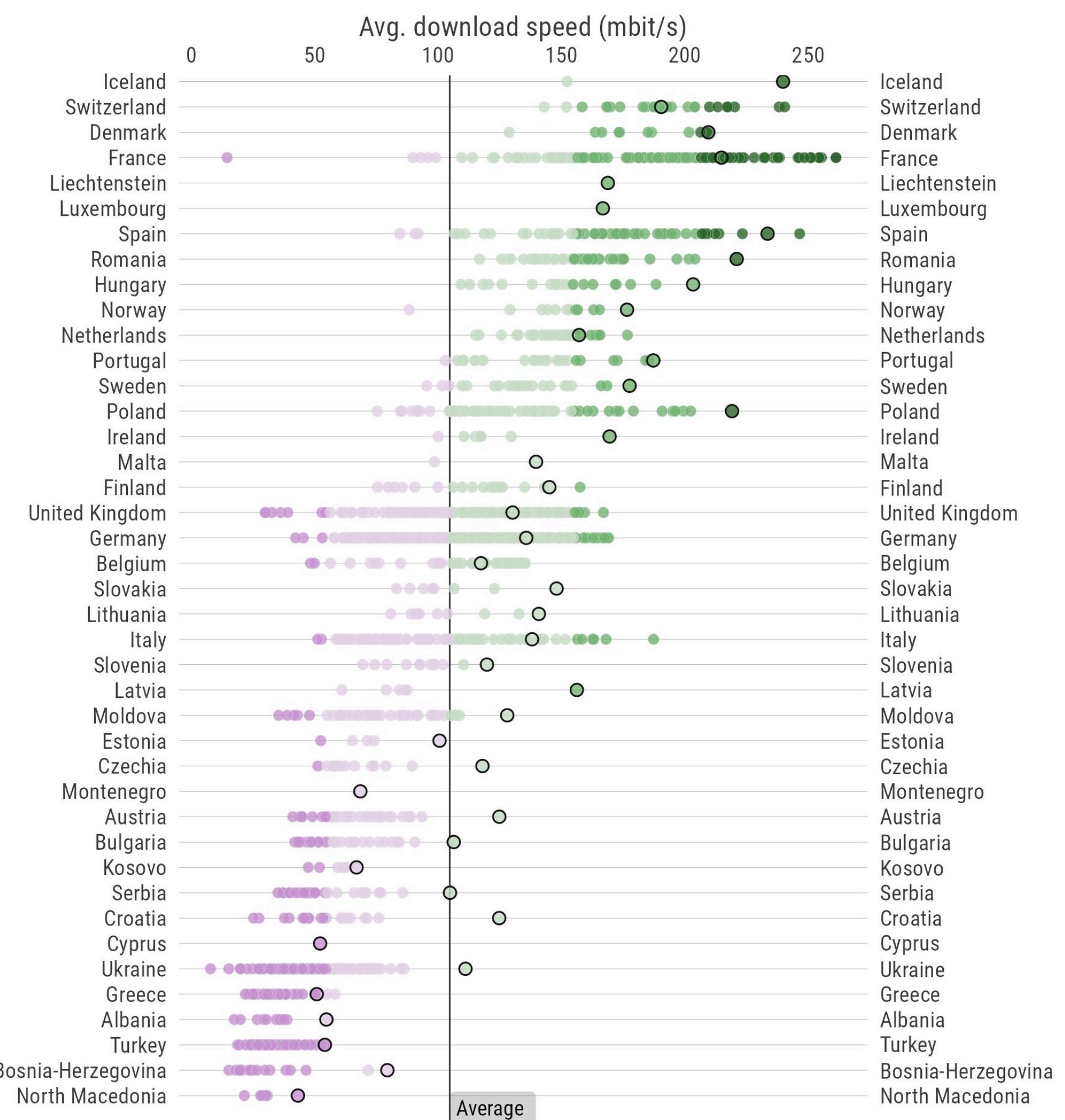
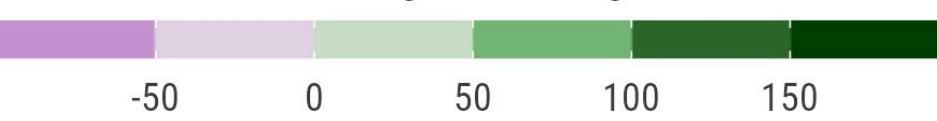
Bir önceki hafta...

Quality of Internet Speed in European Regions

Average download speed in European regions at regional level (NUTS-3) based on Speedtest measurements on fixed and mobile (2022 Q2).

○ marks the capital or the region containing the capital city

Deviation from average of all regions



Source: European Data Journalism Network, Ookla Global Fixed and Mobile Network Performance Maps.

Visualisation: Ansgar Wolsing (Inspiration: Maarten Lambrechts, 'Why Budapest, Warsaw and Lithuania split themselves in two')

Giriş

Miktar görselleştirme çalışmalarında:

- Çubuk grafiği (bar-plot)
- Gruplandırılmış çubuk grafiği (grouped bar-plot)
- Yığılmış çubuk grafiği (stacked bar-plot)

kullanılabilir. Çubuk grafiklerine alternatif olarak, nokta grafikleri (dot-plot) ve ısı haritaları (heatmap) da kullanılabilir.

1. Çubuk Grafikleri

1. Çubuk Grafikleri

Çubuk grafiklerinde sık karşılaşılan iki sorunlardan biri, etiketler arasında herhangi bir mantıksal sıralama düzeni olmadığı durumlarda çubukların rastgele sıralanmasıdır. Grafiğin daha kolay yorumlanabilmesi için çubukların artan ya da azalan sıradada düzenlenmesi gerekmektedir.

1. Çubuk Grafikleri

Diğer sorun ise, eksenlerde yer alan etiket isimlerinin yatayda çok fazla yer kapması, hatta üst üste binmesi nedeniyle okunamaz halde olabilir. Bu durumda en etkili çözüm eksenlerin ters çevrilmesidir.

2. Gruplandırılmış Çubuk Grafikleri

Miktar görselleştirirken **veri setinde birden fazla kategorik değişken olması durumunda** gruplandırılmış çubuk grafikleri kullanılabilir.

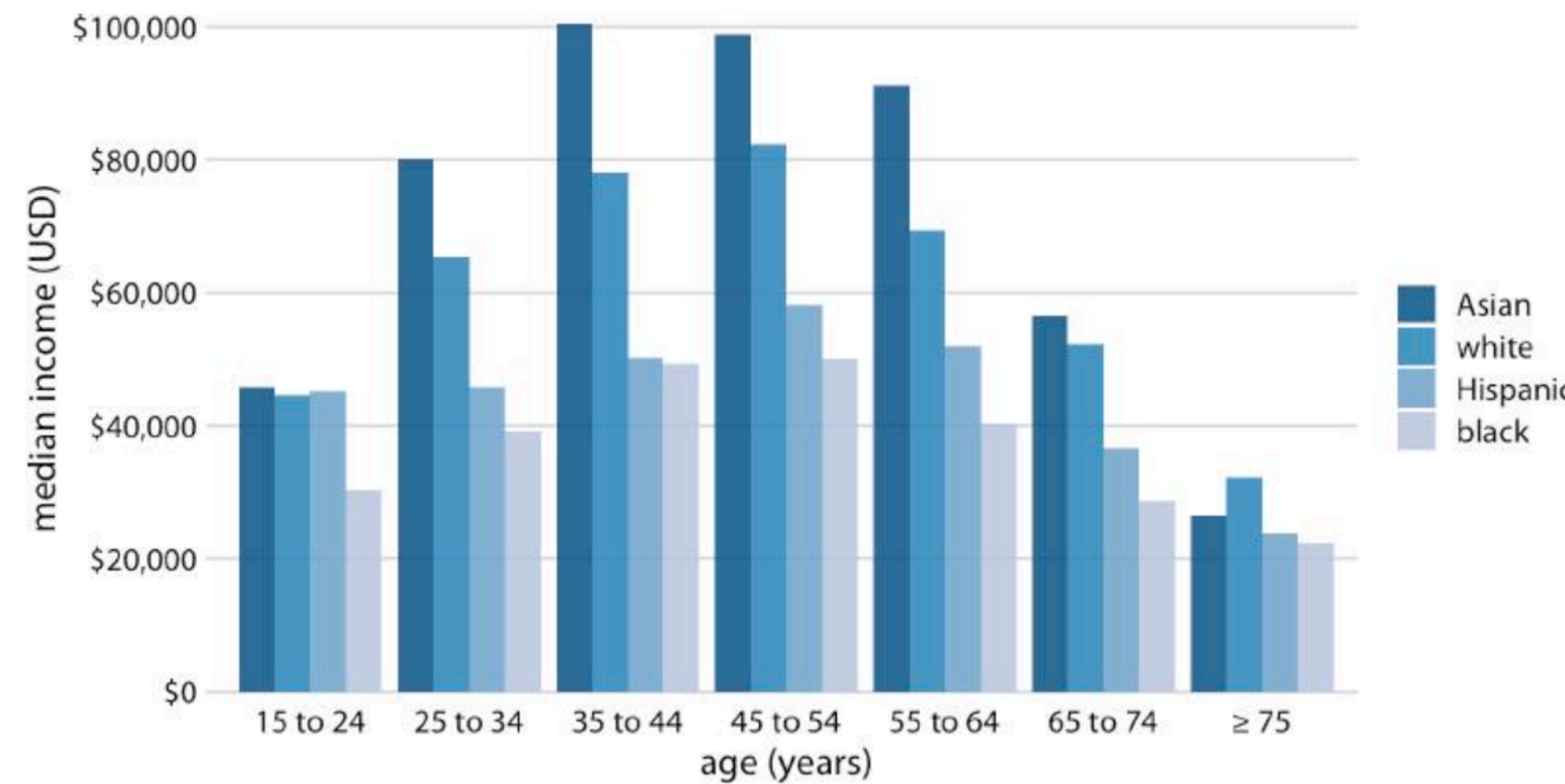


Figure 6-7. 2016 median US annual household income versus age group and race. Age groups are shown along the x axis, and for each age group there are four bars, corresponding to the median income of Asian, white, Hispanic, and black people, respectively.
Data source: US Census Bureau.

2. Gruplandırılmış Çubuk Grafikleri

Bir önceki grafik her ne kadar doğru bir şekilde hazırlanmış olsa da ırkları birbiriyle karşılaştırırken okunması zor bir grafiktir.

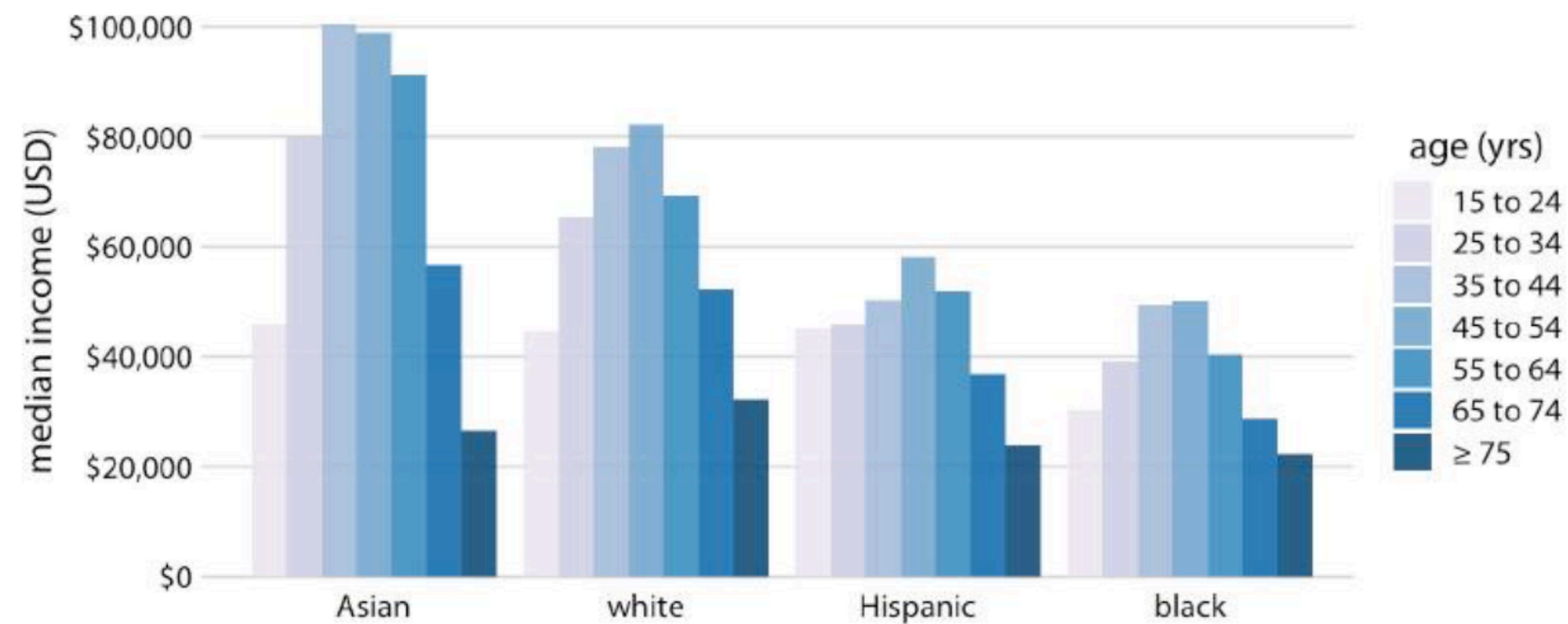


Figure 6-8. 2016 median US annual household income versus age group and race. In contrast to [Figure 6-7](#), now race is shown along the x axis, and for each race we show seven bars according to the seven age groups. Data source: US Census Bureau.

2. Gruplandırılmış Çubuk Grafikleri

Bu gibi durumlarda, kolaylık olması açısından gruplandırılmış çubuk grafiği yerine, dört panelli bir çubuk grafik oluşturulabilir.

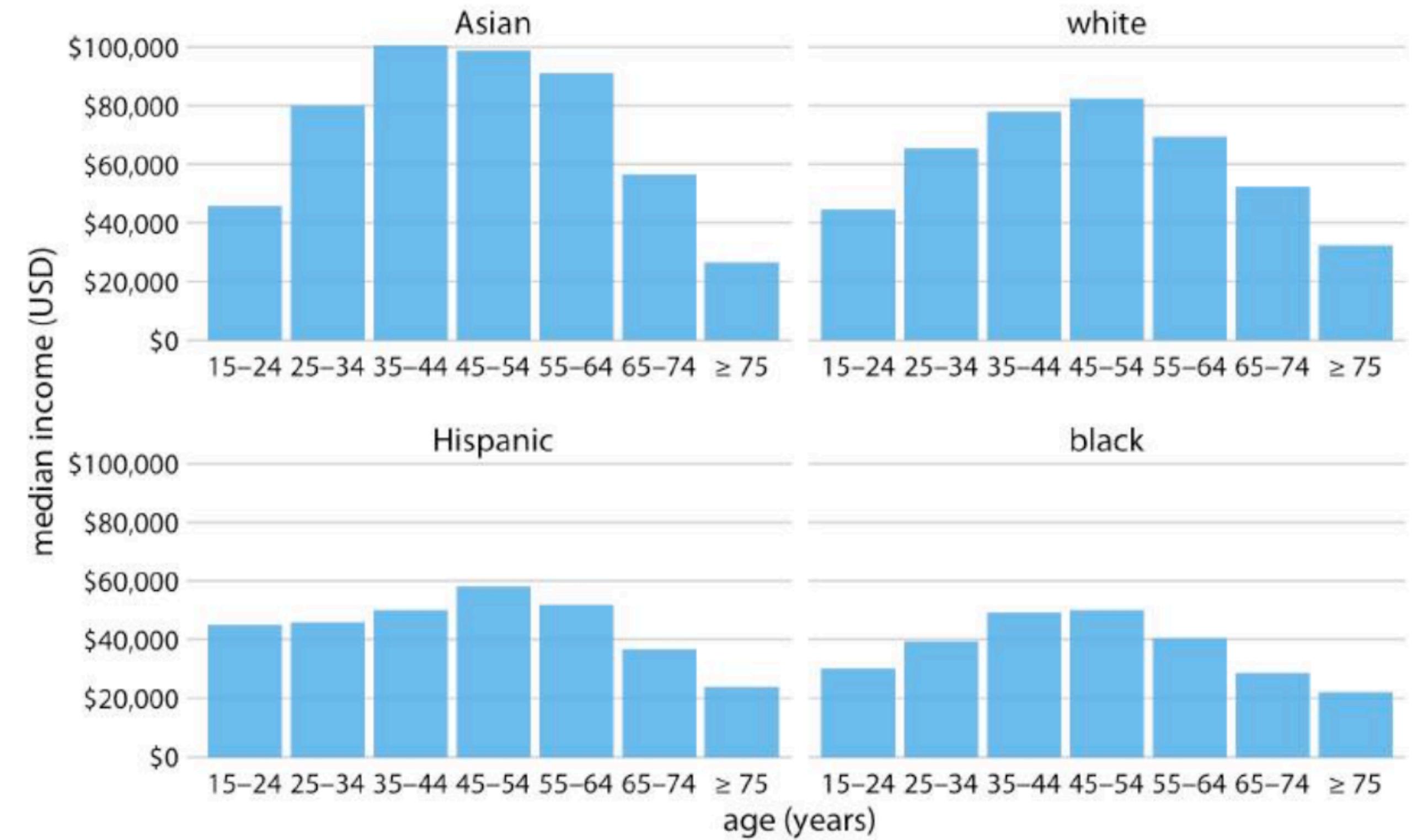


Figure 6-9. 2016 median US annual household income versus age group and race. Instead of displaying this data as a grouped bar plot, as in Figures 6-7 and 6-8, we now show the data as four separate regular bar plots. This choice has the advantage that we don't need to encode either categorical variable by bar color. Data source: US Census Bureau.

3. Yığılmış Çubuk Grafikleri

Gruplandırılmış çubuk grafikleri toplamları karşılaştırmak istediğimizde kullanışlı değildirler. Bu durumda yığılmış çubuk grafikleri kullanılabilir.

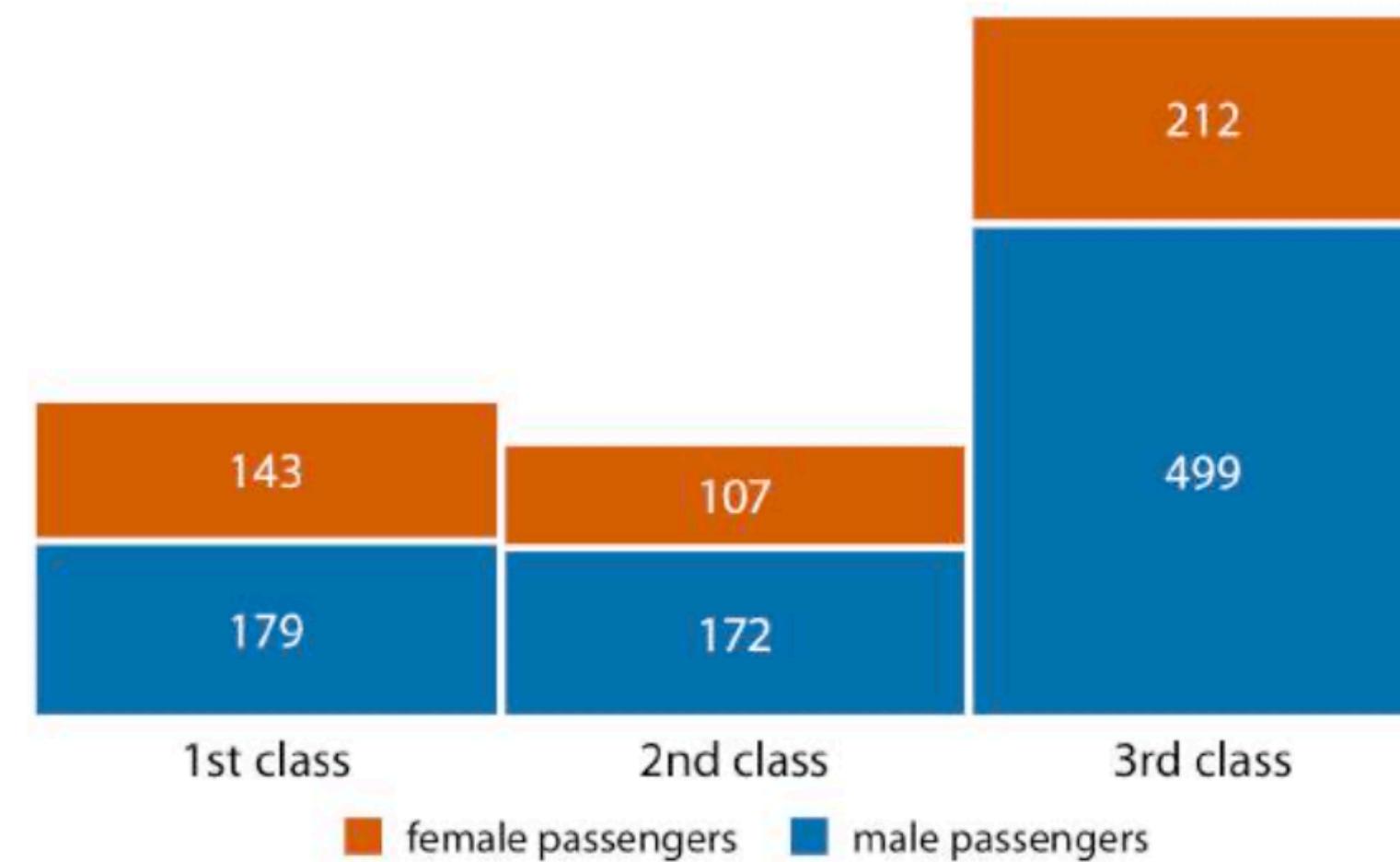


Figure 6-10. Numbers of female and male passengers on the Titanic traveling in 1st, 2nd, and 3rd class. Data source: Encyclopedia Titanica.

Ancak bu tür grafiklerde kategorik değişkenin düzey sayısı 4 ve üzeri olduğunda grafiği okumak zorlaşmaktadır. Bu nedenle düzey sayısının 2 veya 3 olması durumunda tercih edilmelidirler.

4. Nokta Grafikleri

Çubuk grafiklerinin en önemli dezavantajı, temsil ettikleri miktarlar ile orantılı olabilmeleri için sıfır noktasından başlamaları gerekliliğidir. Bu durum düzey sayısı ve miktar arttığında grafiklerin okunmasını zorlaştırabilir. Bu gibi durumlarda nokta grafikleri iyi bir alternatiftir.

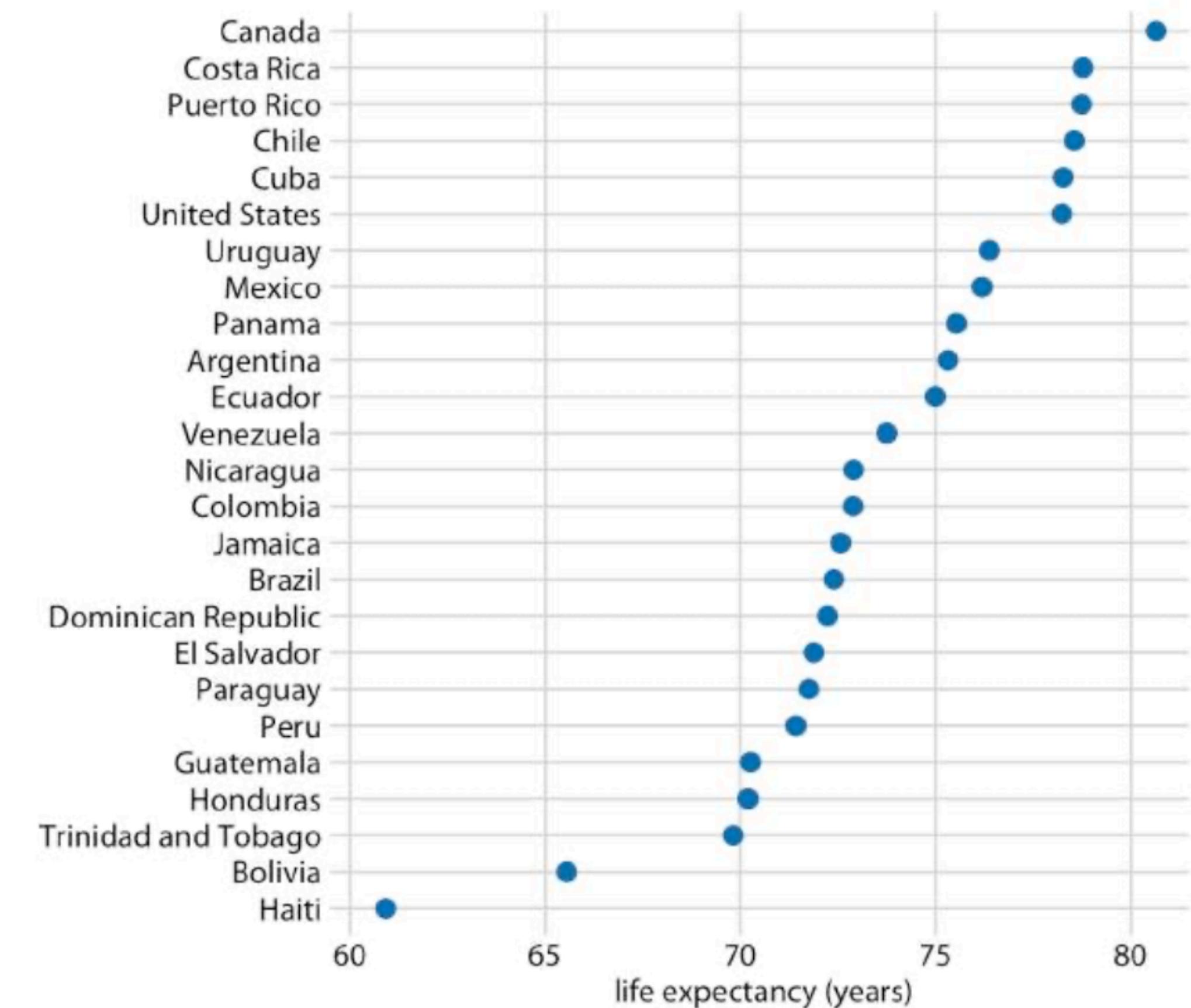


Figure 6-11. Life expectancies of countries in the Americas, for the year 2007. Data source: Gapminder.

4. Nokta Grafikleri

Eğer nokta grafik yerine çubuk grafiği kullanmakta ısrar etseydik?

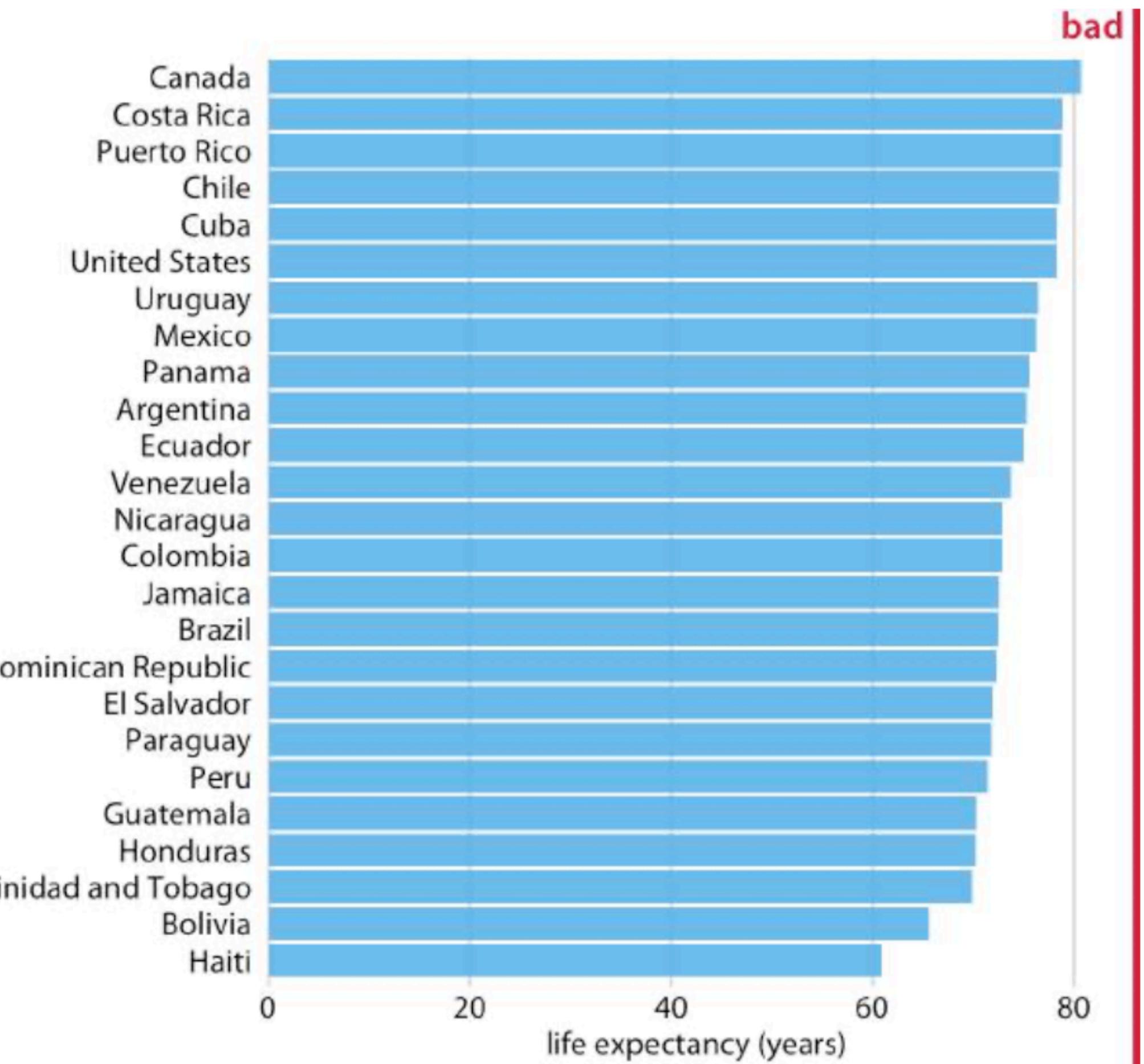


Figure 6-12. Life expectancies of countries in the Americas, for the year 2007, shown as bars. This dataset is not suitable for being visualized with bars. The bars are too long and they draw attention away from the key feature of the data, the differences in life expectancy among the different countries. Data source: Gapminder.

5. Isı Haritaları

Veri setinde iki kategorik değişken bulunması durumunda gruplandırılmış ya da yiğilmiş çubuk grafikleri kullanılabilir. Ancak kategorik değişkenlerin düzey sayısı arttığında bu grafiklerin okunması zorlaşır. Bu gibi durumlarda ısı haritası iyi bir alternatiftir.

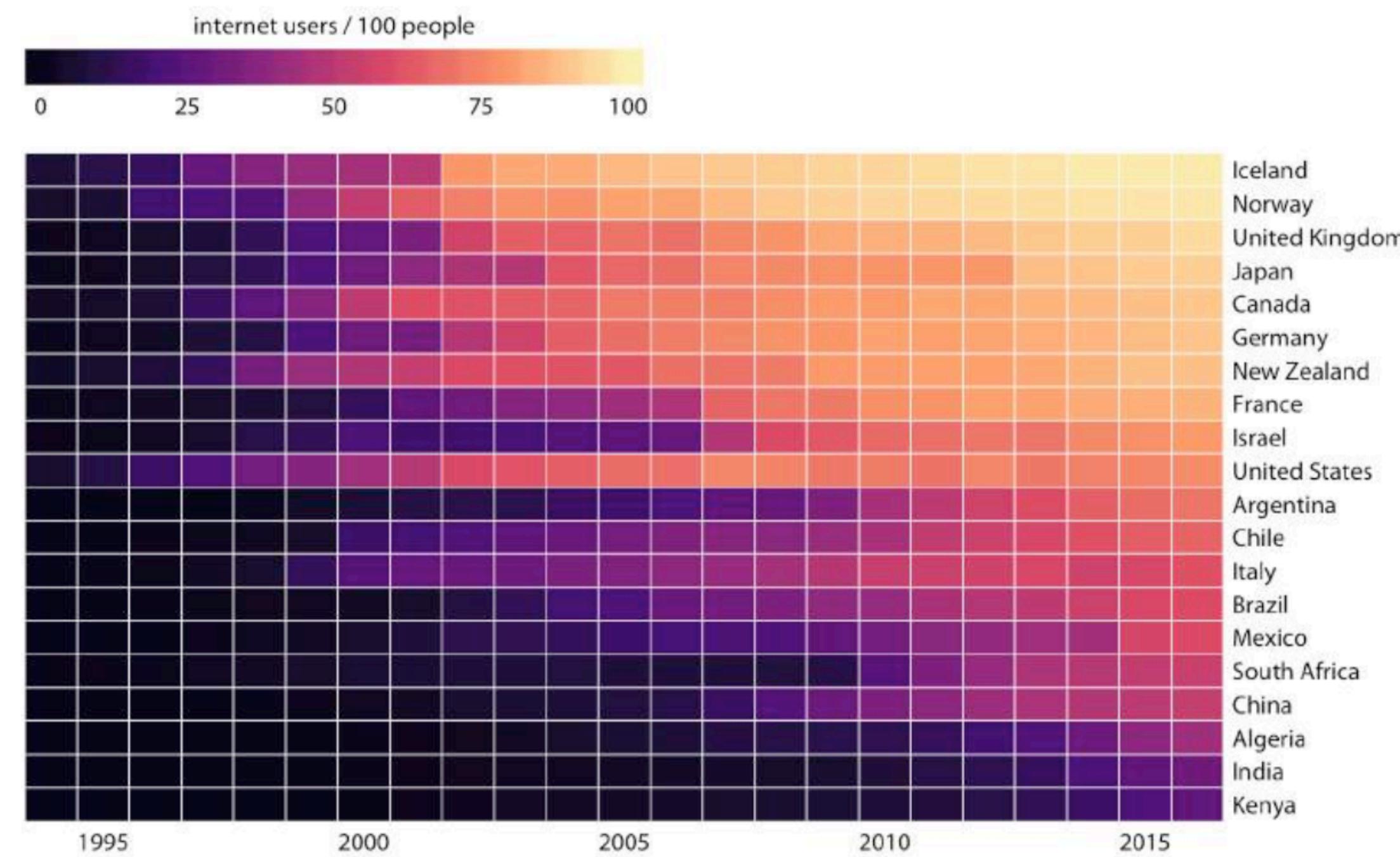


Figure 6-14. Internet adoption over time, for select countries. Color represents the percent of internet users for the respective country and year. Countries were ordered by percent internet users in 2016. Data source: World Bank.

5. Isı Haritaları

Eğer veri setinde yer alan kategorik değişkenin düzeyleri arasında mantıksal bir ilişki yoksa (sıralı değilse) grafiği daha kolay okuyabilmek için düzeylere karşılık gelen sürekli değişkenin aldığı değerlere göre artan ya da azalan düzende sıralanması, kolay anlaşılabilirlik açısından önemlidir. Isı haritasında ise bu düzen renk yoğunluklarının geçişine uygun olarak yapılmalıdır.

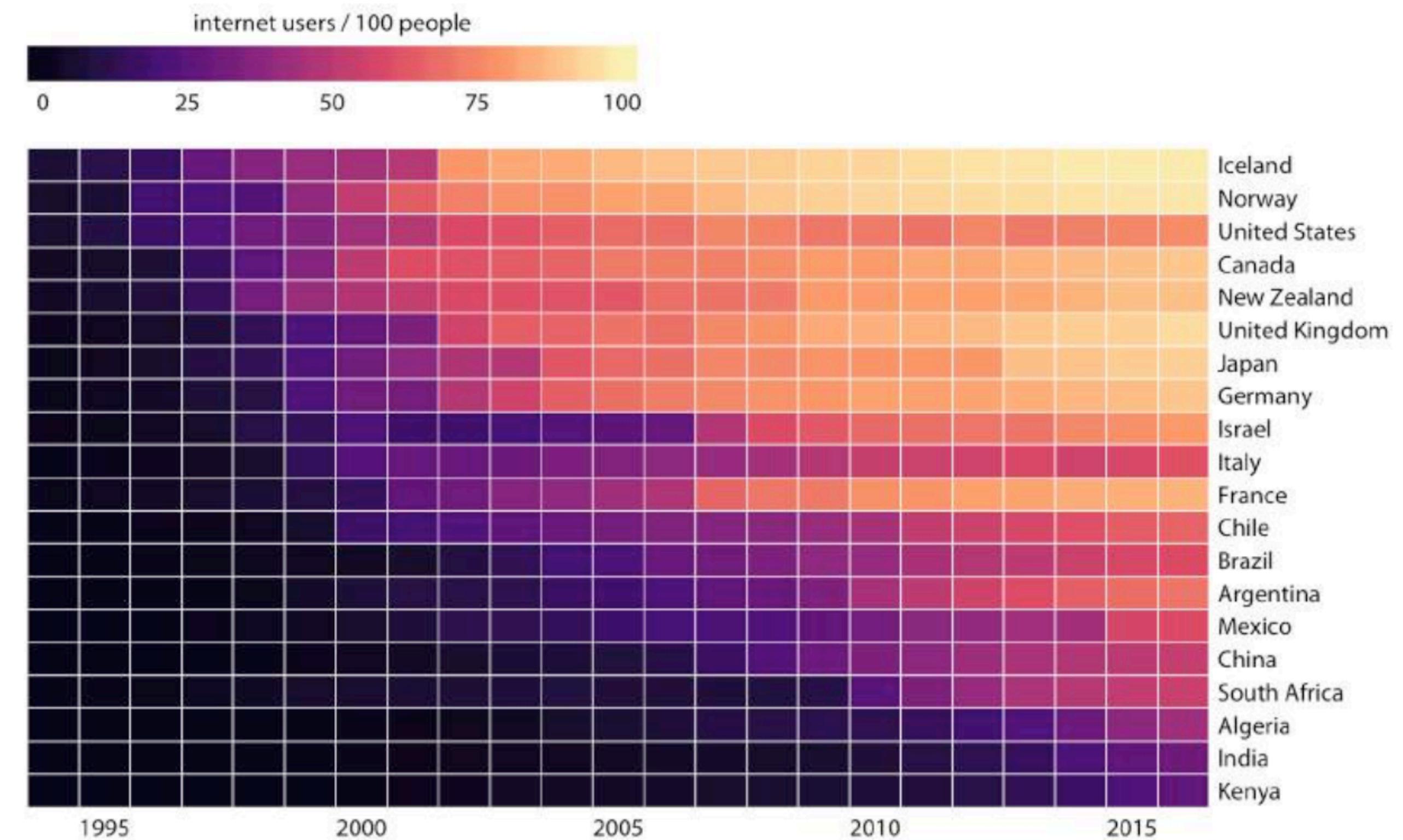


Figure 6-15. Internet adoption over time, for select countries. Countries were ordered by the year in which their internet usage first exceeded 20%. Data source: World Bank.

Kaynak

Bu derste yer alan not ve görseller, Claus O. Wilke'nin “Fundamentals of Data Visualization” isimli kitabından derlenmiştir.

