

18.11.2022

# Veri Görselleştirme

## Hafta 5: Oranların Görselleştirilmesi

# Giriş

Pratikte bir grubun veya miktarın nasıl her biri bütünün bir oranını temsil eden ayrı ayrı parçalara ayrıldığını göstermek gerekebilir.

Örneğin:

- bir topluluktaki kadın ve erkeklerin oranı,
- bir seçimde farklı siyasi partilere oy veren seçmenlerin oranı,
- şirketlerin pazar payları vb.

Bu gibi oranların görselleştirilmesinde en sık kullanılan grafik türleri: ...

# Giriş

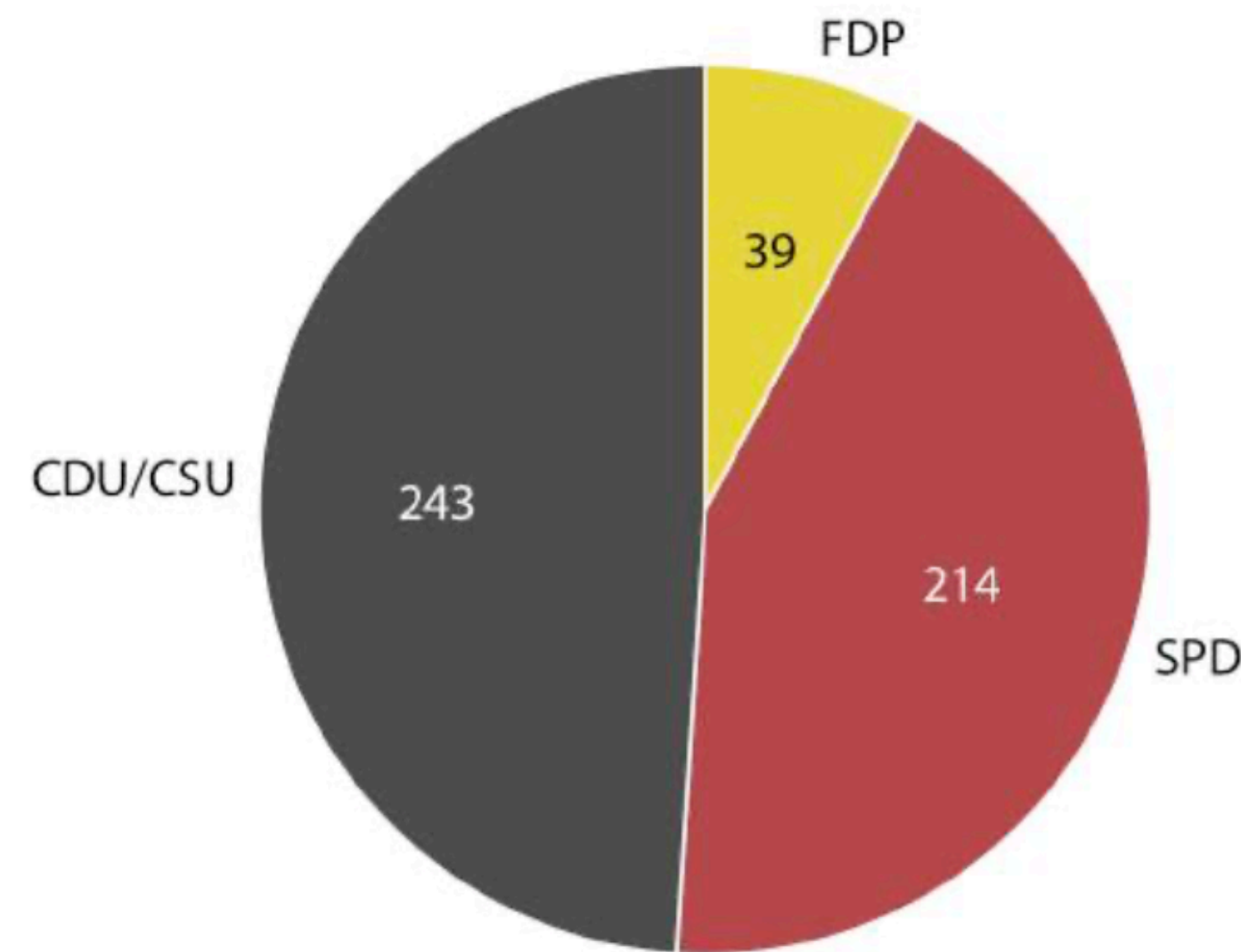
Pratikte bir grubun veya miktarın nasıl her biri bütünün bir oranını temsil eden ayrı ayrı parçalara ayrıldığını göstermek gerekebilir. Örneğin:

- bir topluluktaki kadın ve erkeklerin oranı,
- bir seçimde farklı siyasi partilere oy veren seçmenlerin oranı,
- şirketlerin pazar payları vb.

Bu gibi oranların görselleştirilmesinde en sık kullanılan grafik türleri, **pasta (pie-chart)** ve **çubuk (bar-plot)** grafikleridir.

# Pasta grafiği (Pie-chart)

Pasta grafiği, bir daireyi her dilimin alanı temsil ettiği ve genel toplama orantılı olacak şekilde dilimlere ayırır.



*Figure 10-1. Party composition of the eighth German Bundestag, 1976–1980, visualized as a pie chart. This visualization highlights that the ruling coalition of SPD and FDP had a small majority over the opposition CDU/CSU. Data source: Wikipedia.*

# Pasta grafiği (Pie-chart)

Bu tür veriler, yatay ya da dikey ekseninde yığılmış çubuk grafik kullanarak da görselleştirilebilir.

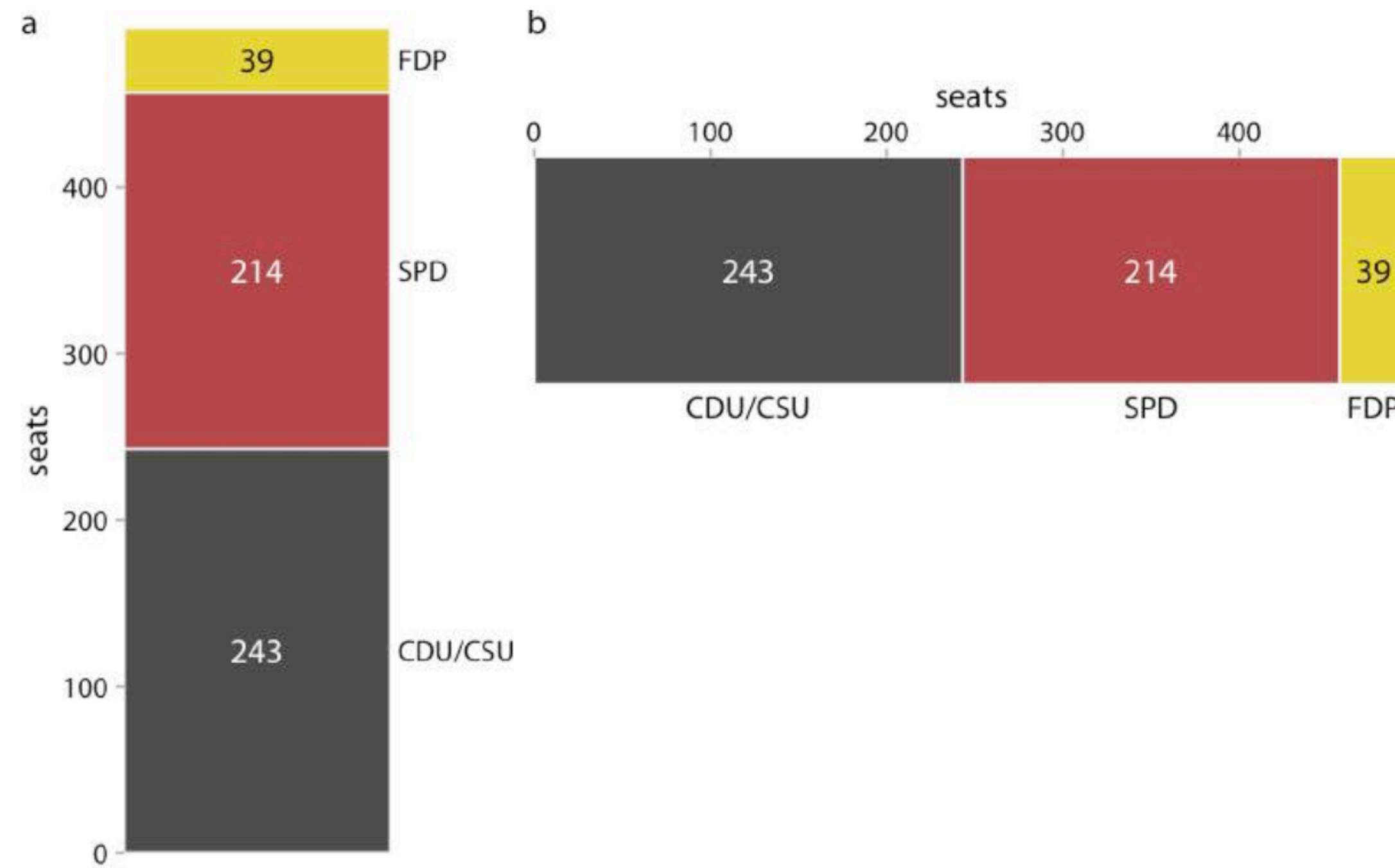


Figure 10-2. Party composition of the eighth German Bundestag, 1976–1980, visualized as stacked bars. (a) Bars stacked vertically. (b) Bars stacked horizontally. It is not immediately obvious that SPD and FDP jointly had more seats than CDU/CSU. Data source: Wikipedia.

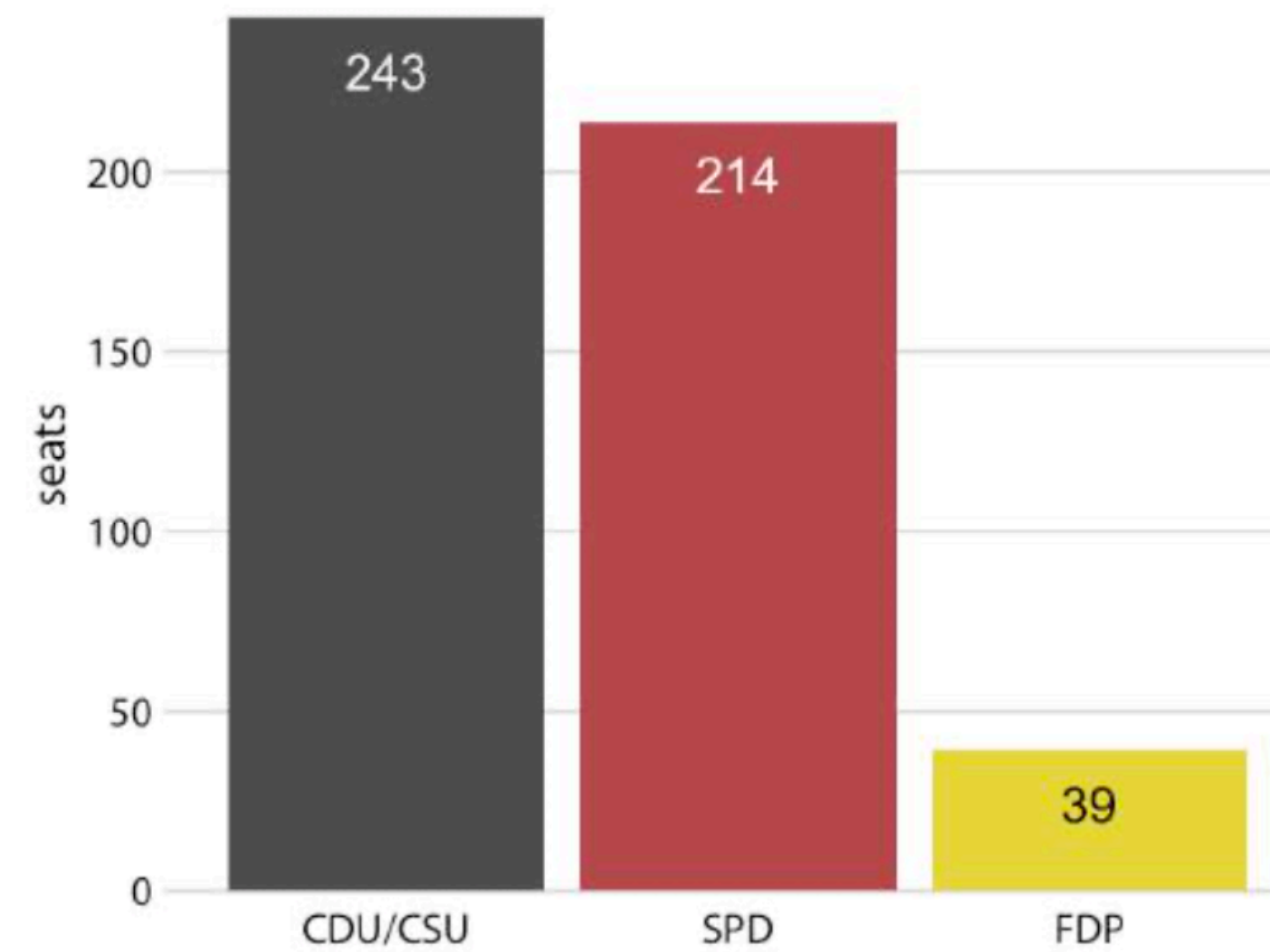
# Pasta grafiđi (Pie-chart)

Pasta grafiđi ve yığılmıř ubuk grafiđinden hangisinin daha iyi oran grselleřtirme aracı olduđu ok sık tartıřılan bir konudur.

Pasta grafikleri genellikle **yarım, te bir ya da drtte bir gibi basit oranları grselleřtirirken** daha iyi alıřır.

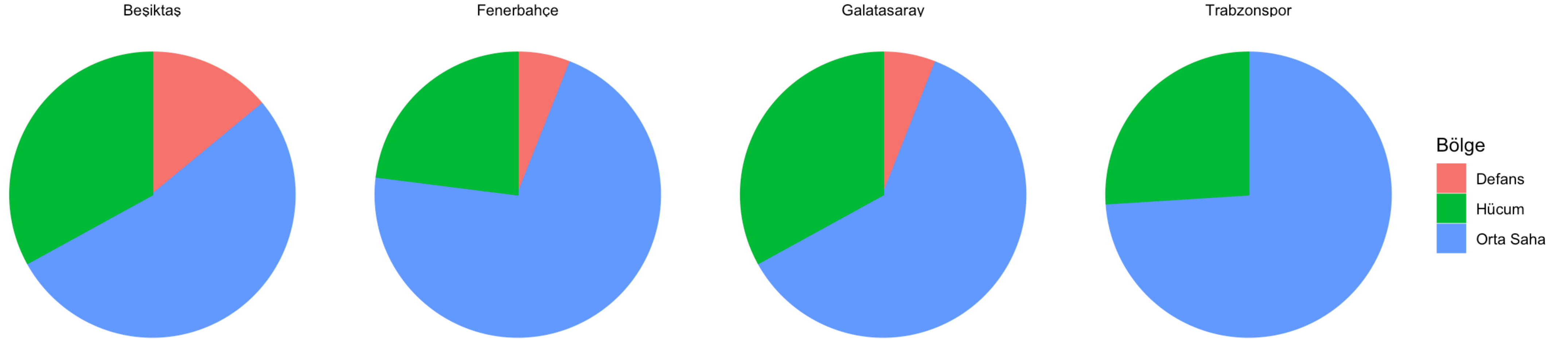
**ok sayıda grup olduđunda** iyi alıřmazlar. Bu gibi durumlarda ubuk grafikleri daha iyi bir alternatiftir.

# Çubuk grafikleri



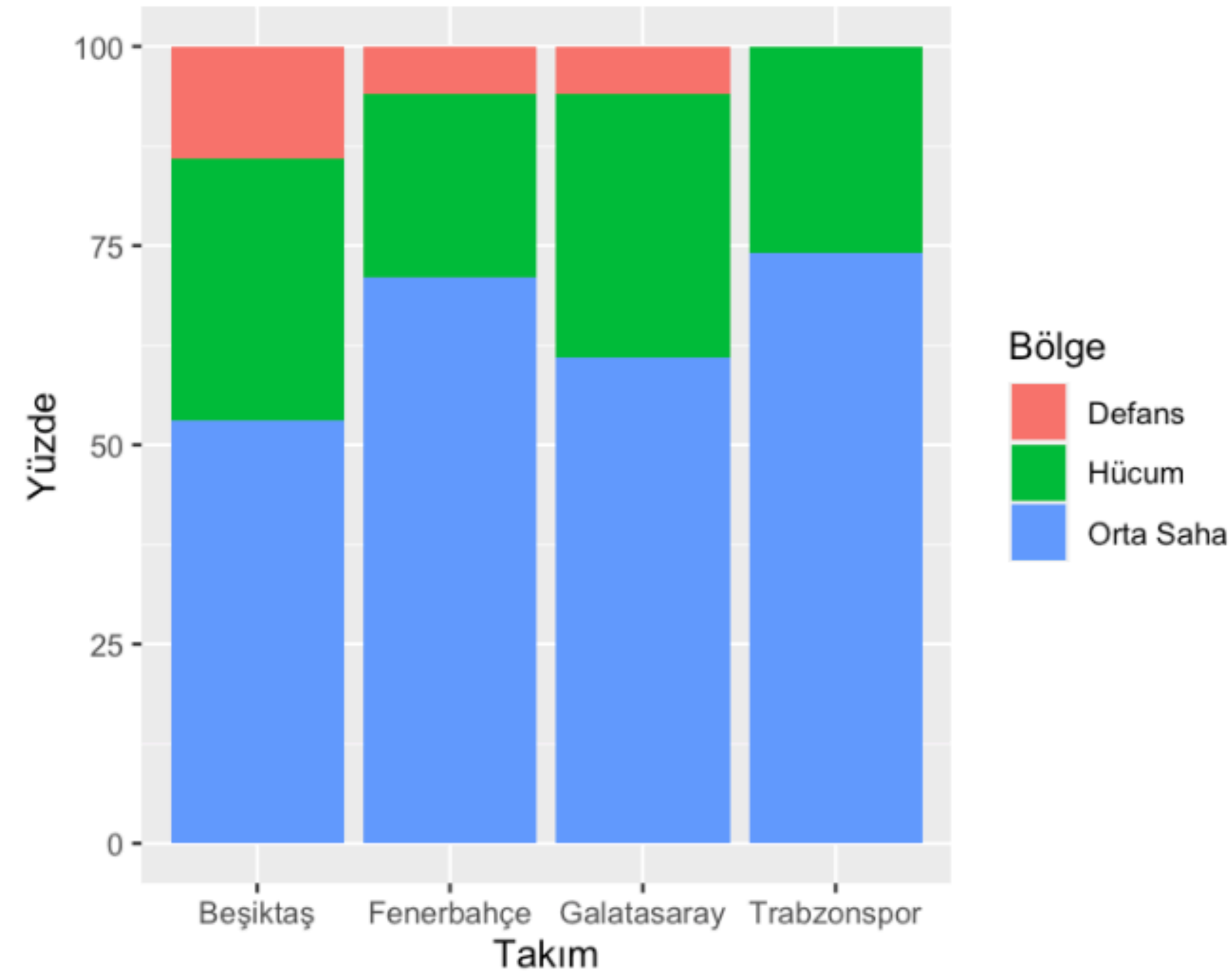
*Figure 10-3. Party composition of the eighth German Bundestag, 1976–1980, visualized as side-by-side bars. As in [Figure 10-2](#), it is not immediately obvious that SPD and FDP jointly had more seats than CDU/CSU. Data source: Wikipedia.*

# Pasta grafiđi kullanımdaki problemler

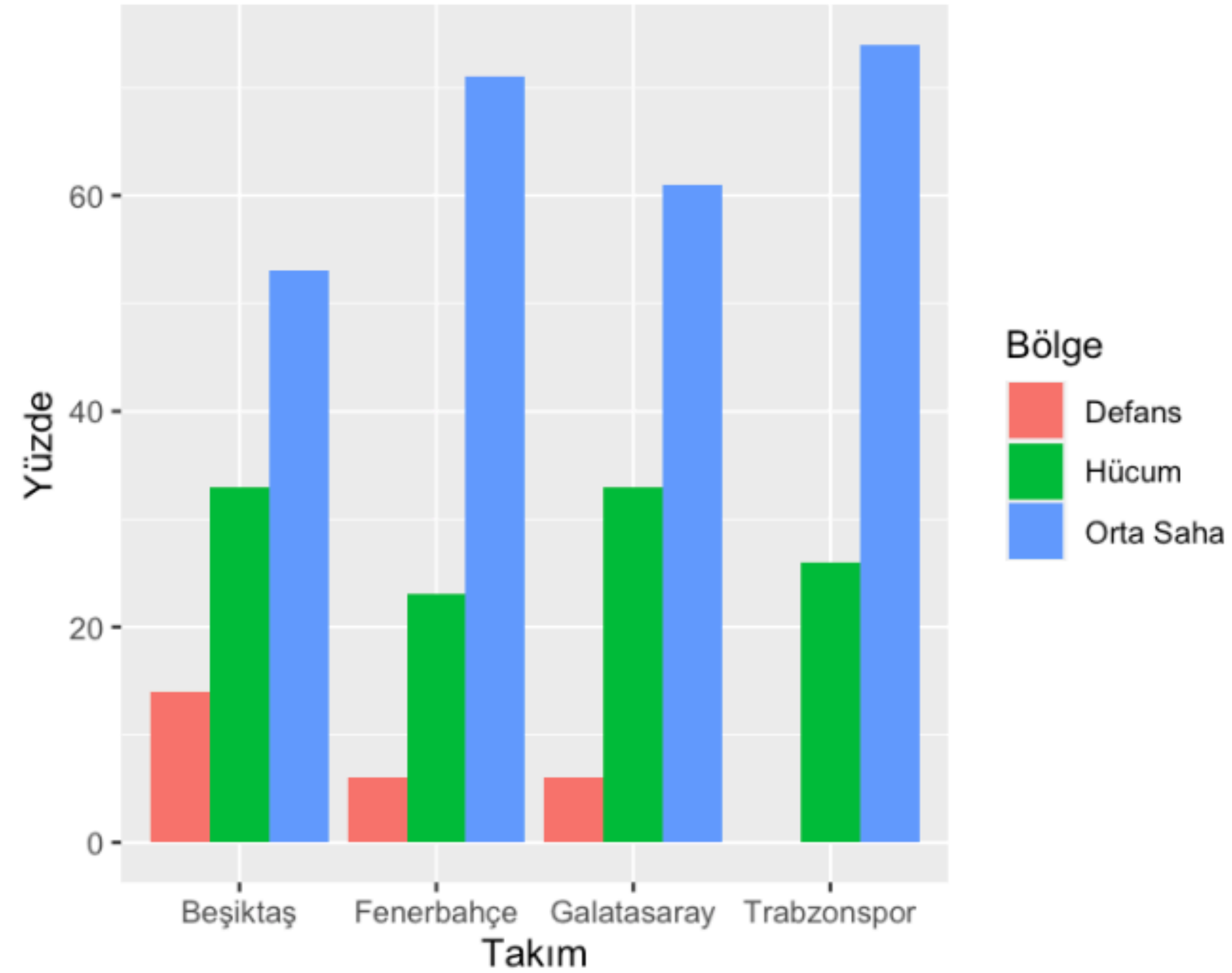




# Yığılmış çubuk grafiğinin kullanımdaki problemler

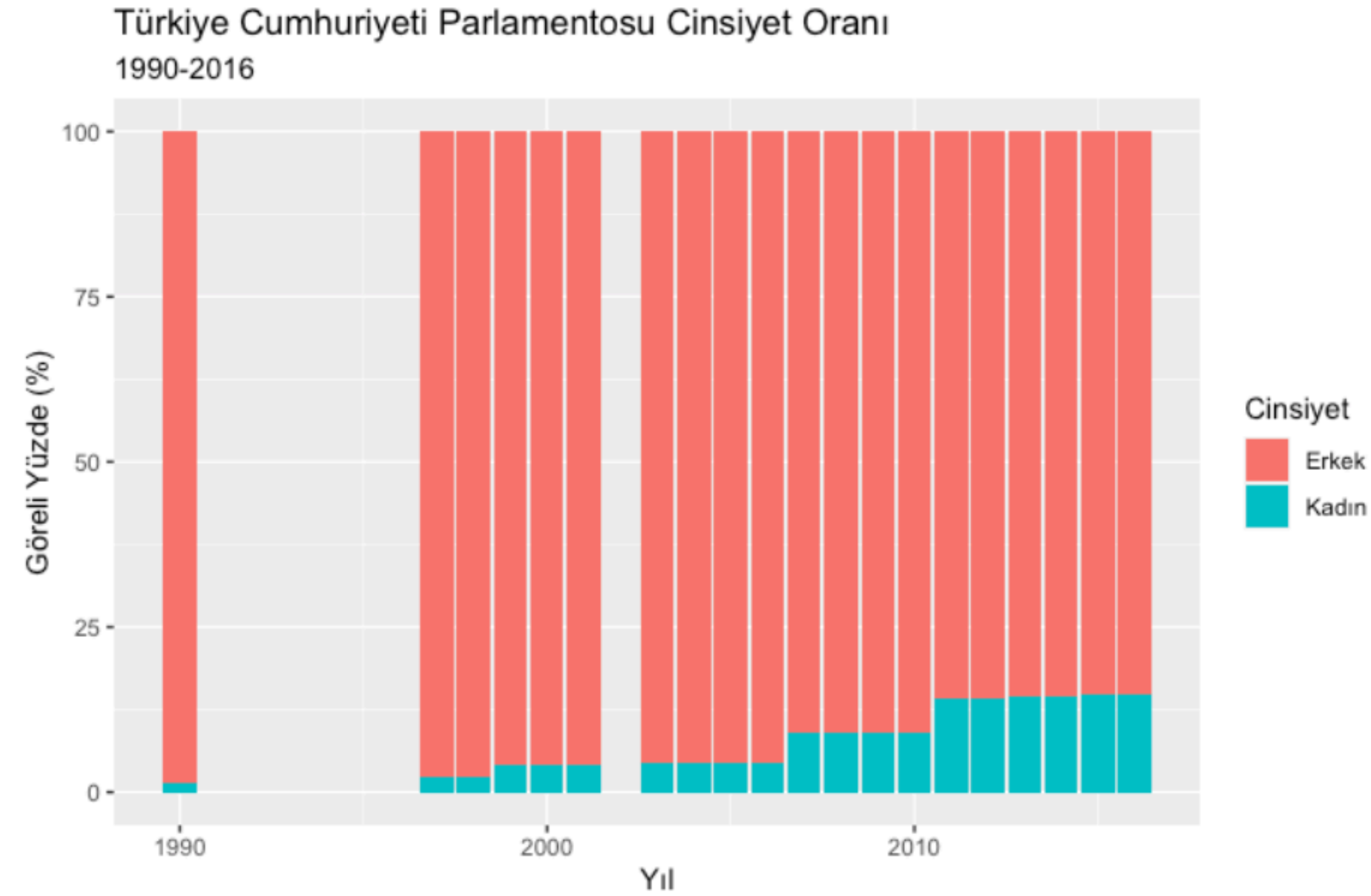


# Yığılmış çubuk grafiğinin kullanımdaki problemler

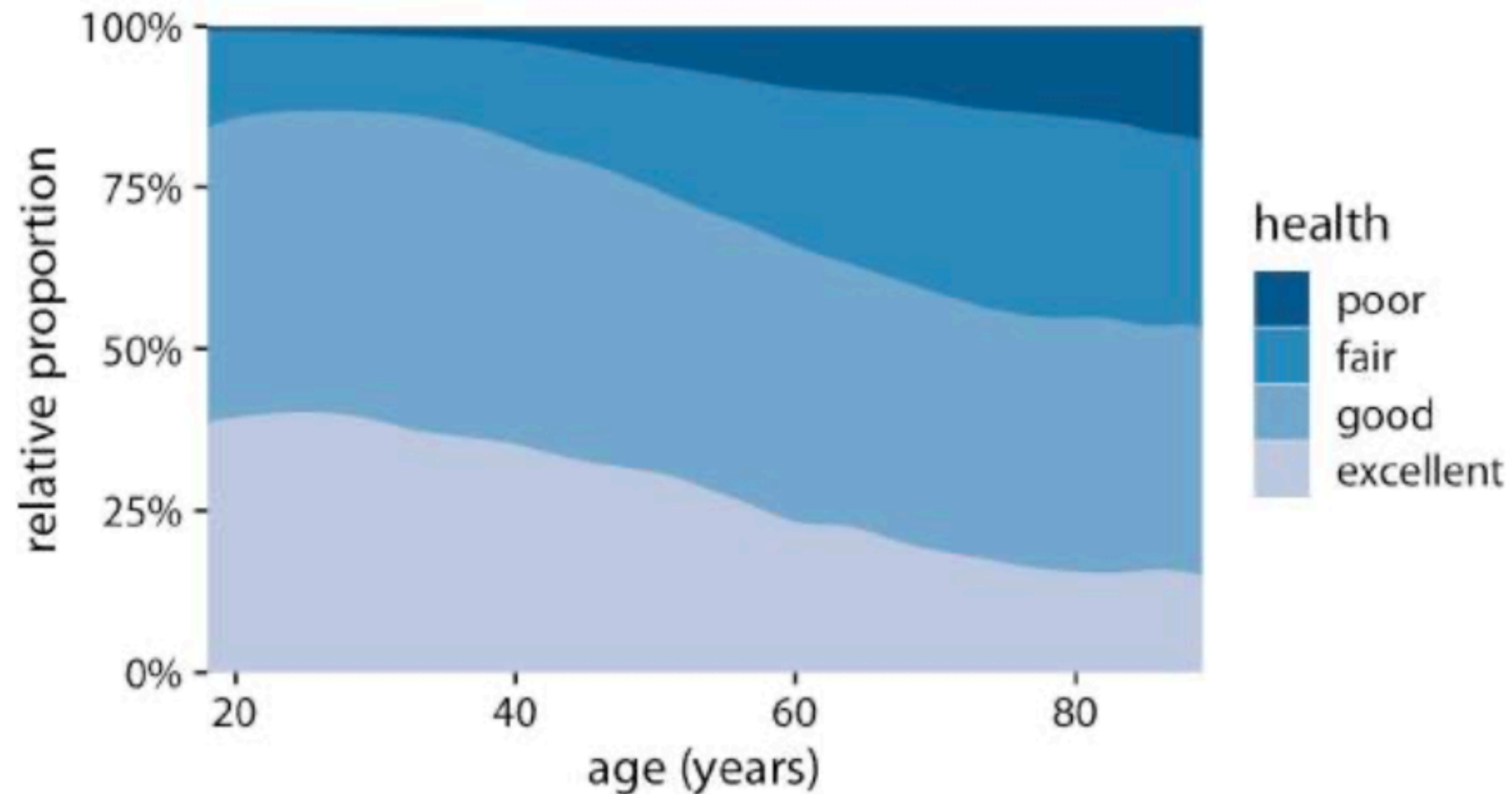


# Yığılmış çubuk grafiğinin kullanımdaki problemler

Yalnızca iki grubun yer aldığı oran görselleştirmelerinde yığılmış çubuk grafikleri iyi şekilde çalışırlar.



# Yığılmış yoğunluk grafiğinin kullanımdaki problemler



*Figure 10-8. Health status by age. Data source: General Social Survey (GSS).*

# Bir bütünün parçası olarak oranların görselleştirilmesi

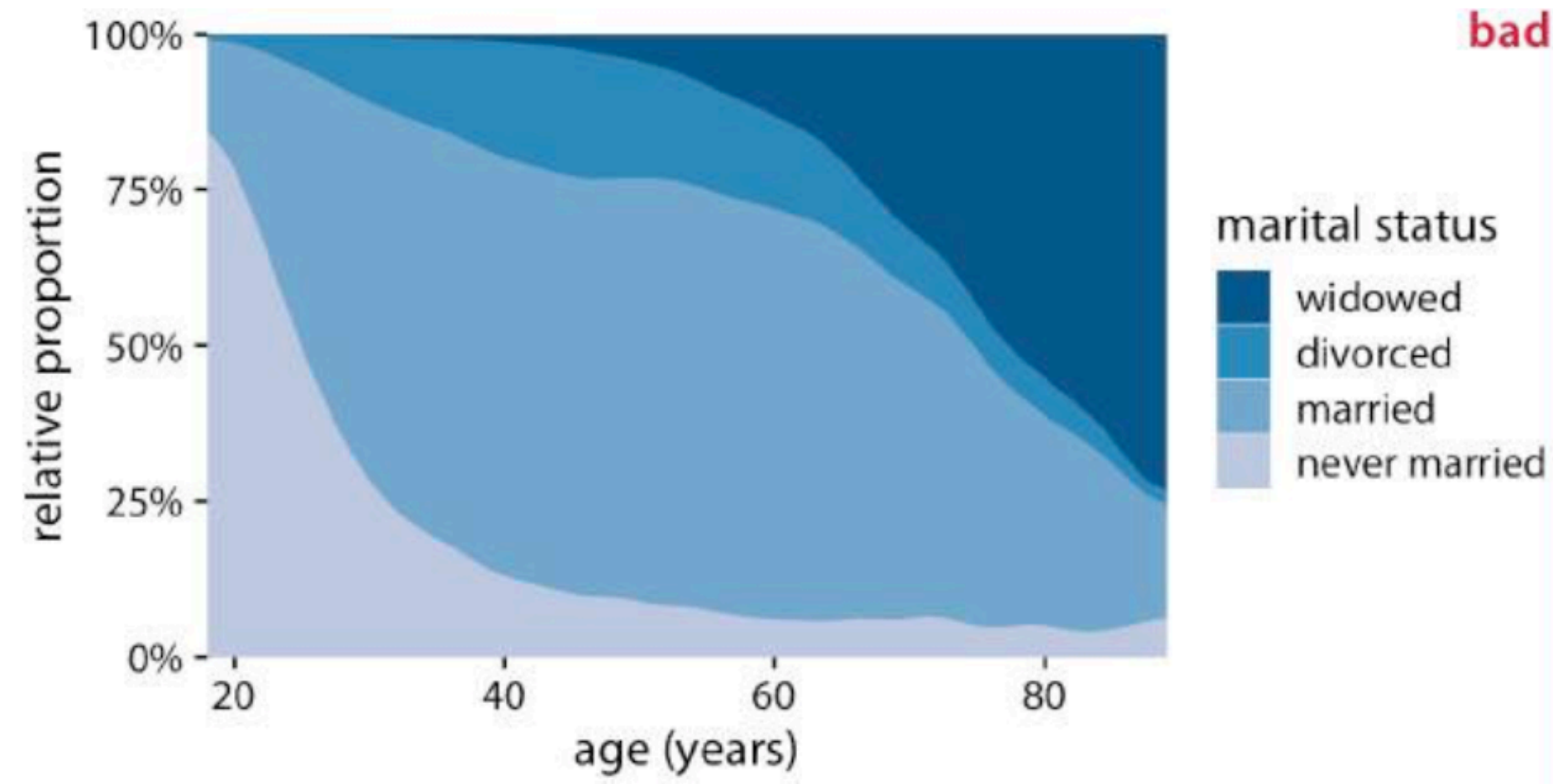
Yan yana (side by side) grafiklerin bazı sorunları olduğu için bütünü parçalara bölerek görselleştirme yoluna gidebiliriz.



*Figure 10-11. Marital status by age, shown as proportion of the total number of people in the survey. The colored areas show the density estimates of the ages of people with the respective marital status, and the gray areas show the overall age distribution. Data source: GSS.*

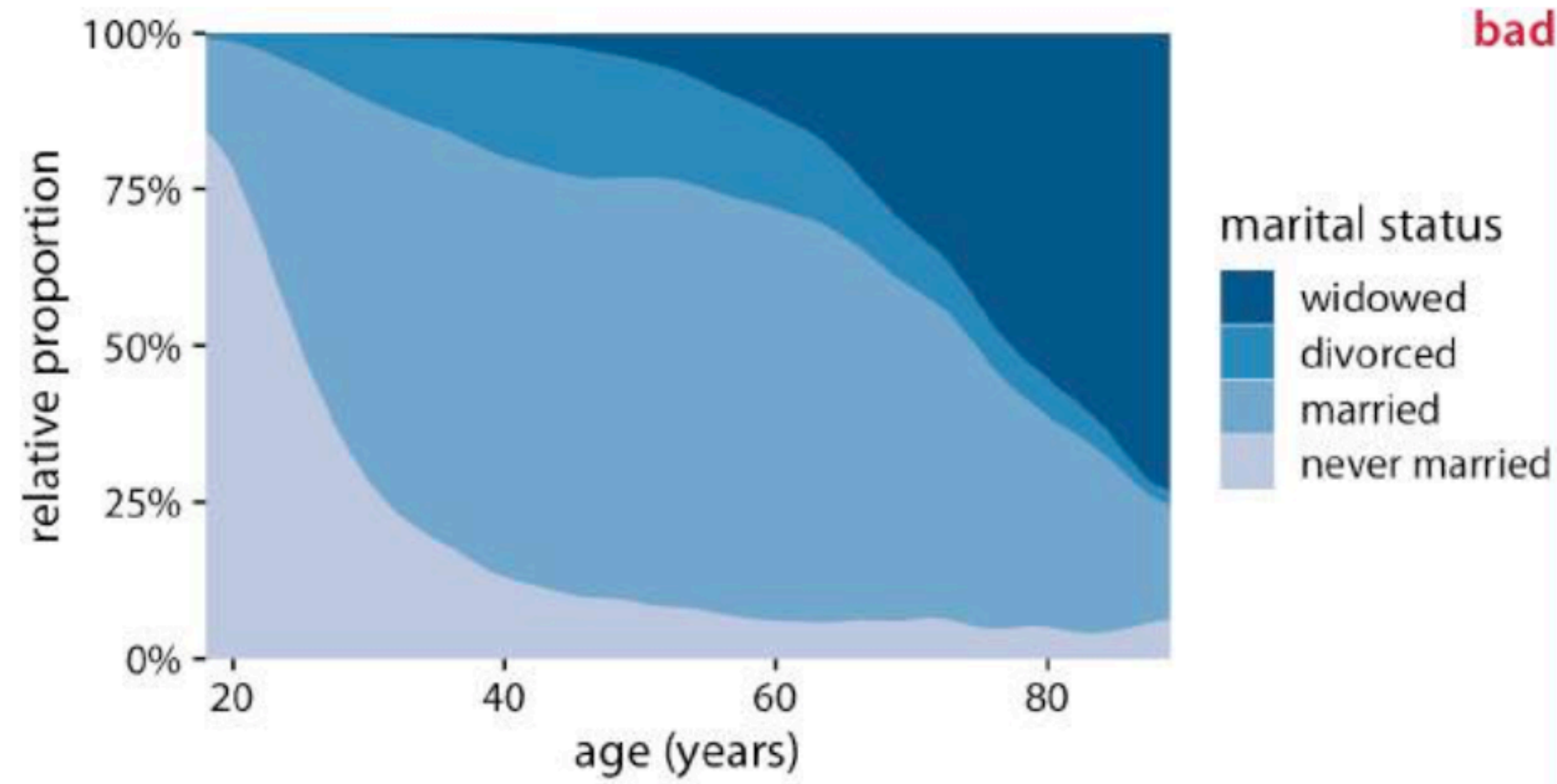


# Bütünü parçalayarak görselleştirmede karşılaşılan sorunlar



*Figure 10-10. Marital status by age. To simplify the figure, I have removed a small number of cases that report as separated. I have labeled this figure as “bad” because the frequency of people who have never been married or are widowed changes so drastically with age that the age distributions of married and divorced people are highly distorted and difficult to interpret. Data source: GSS.*

# Bütünü parçalayarak görselleştirmede karşılaşılan sorunlar



Bu grafiği parçalara ayıralım.

# Bütünü parçalayarak görselleştirmede karşılaşılan sorunlar

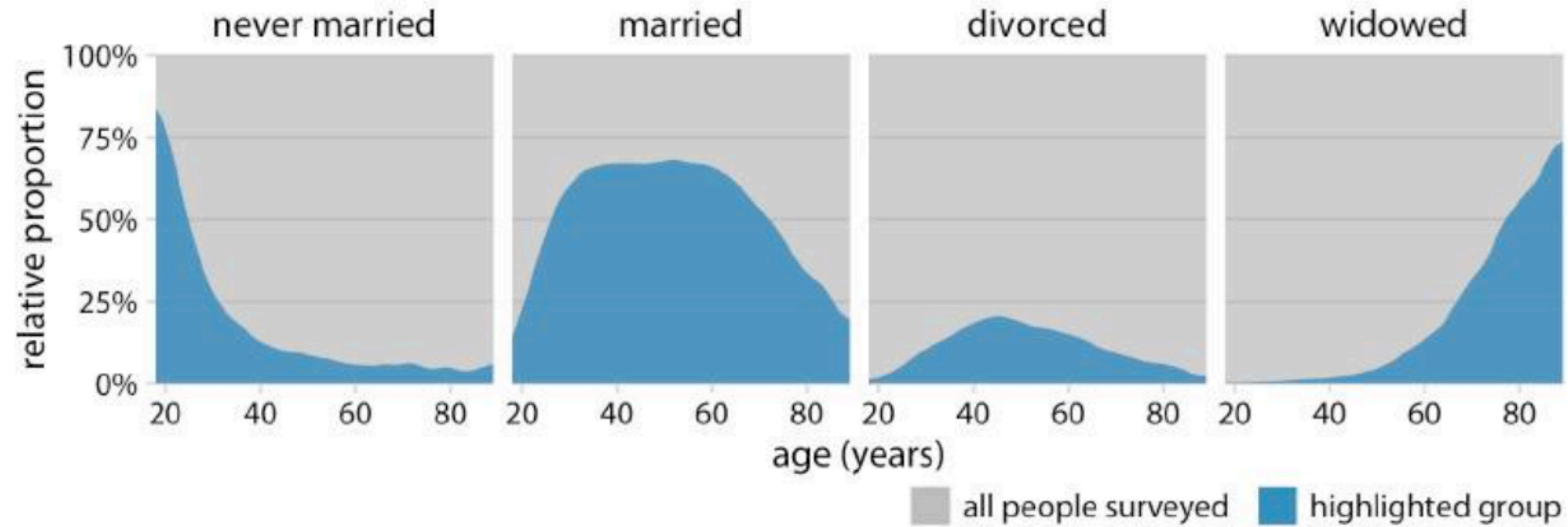


Figure 10-11. Marital status by age, shown as proportion of the total number of people in the survey. The colored areas show the density estimates of the ages of people with the respective marital status, and the gray areas show the overall age distribution. Data source: GSS.

Burada da sorun herhangi bir zaman noktasında grupların görelî oranlarını karşılaştırmaktaki zorluktur.



# Bütünü parçalayarak görselleştirmede karşılaşılan sorunlar



*Figure 10-12. Marital status by age, shown as proportion of the total number of people in the survey. The areas colored in blue show the percent of people at the given age with the respective status, and the areas colored in gray show the percent of people with all other marital statuses. Data source: GSS.*

# İç-içe (nested) oranların görselleştirilmesi

- Bu bölüme kadar bir veri setinin tek bir kategorik değişken üzerinden parçalara ayrıldığı senaryolar üzerinde konuştuk. Aynı anda birden fazla kategorik değişken ile veri setini incelemek ve kırılımlara ayrılması gerekebilir.
- Örneğin, bir mecliste partilerin aldıkları oy oranlarını hem partilere hem de adayların cinsiyetine göre incelenmesi gerekebilir. Bu gibi örnekler iç içe oranlar olarak adlandırılır.
- İç içe oranları görselleştirmek için mozaik grafikleri, ağaç haritaları ve paralel kümeler kullanılabilir.

# Mozaik grafikleri

Bu grafik türü, iki kategorik değişken ve kategorik değişkenin düzeylerine karşılık gelen oranları temsil eden büyüklükte karesel alanlardan oluşur.

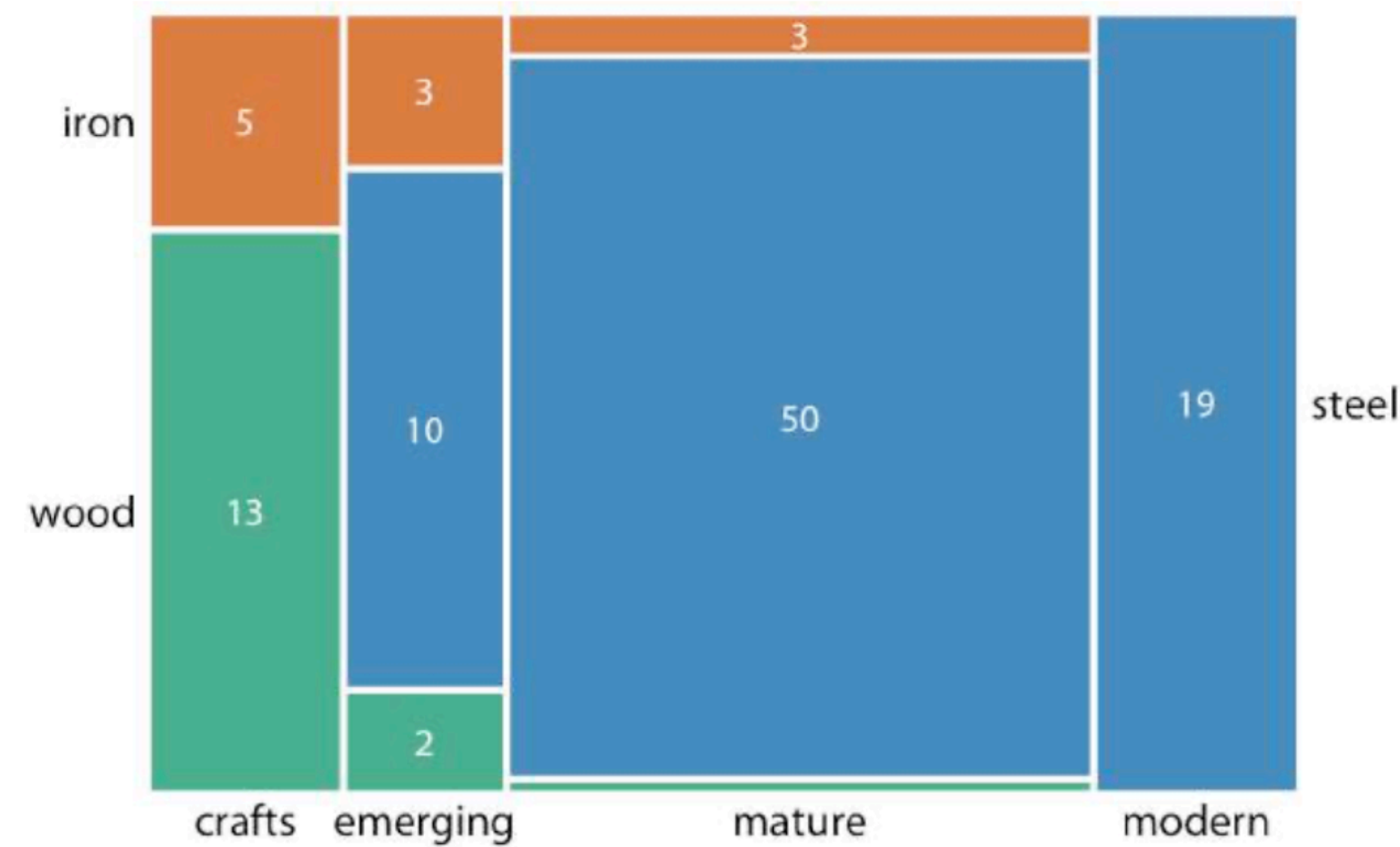
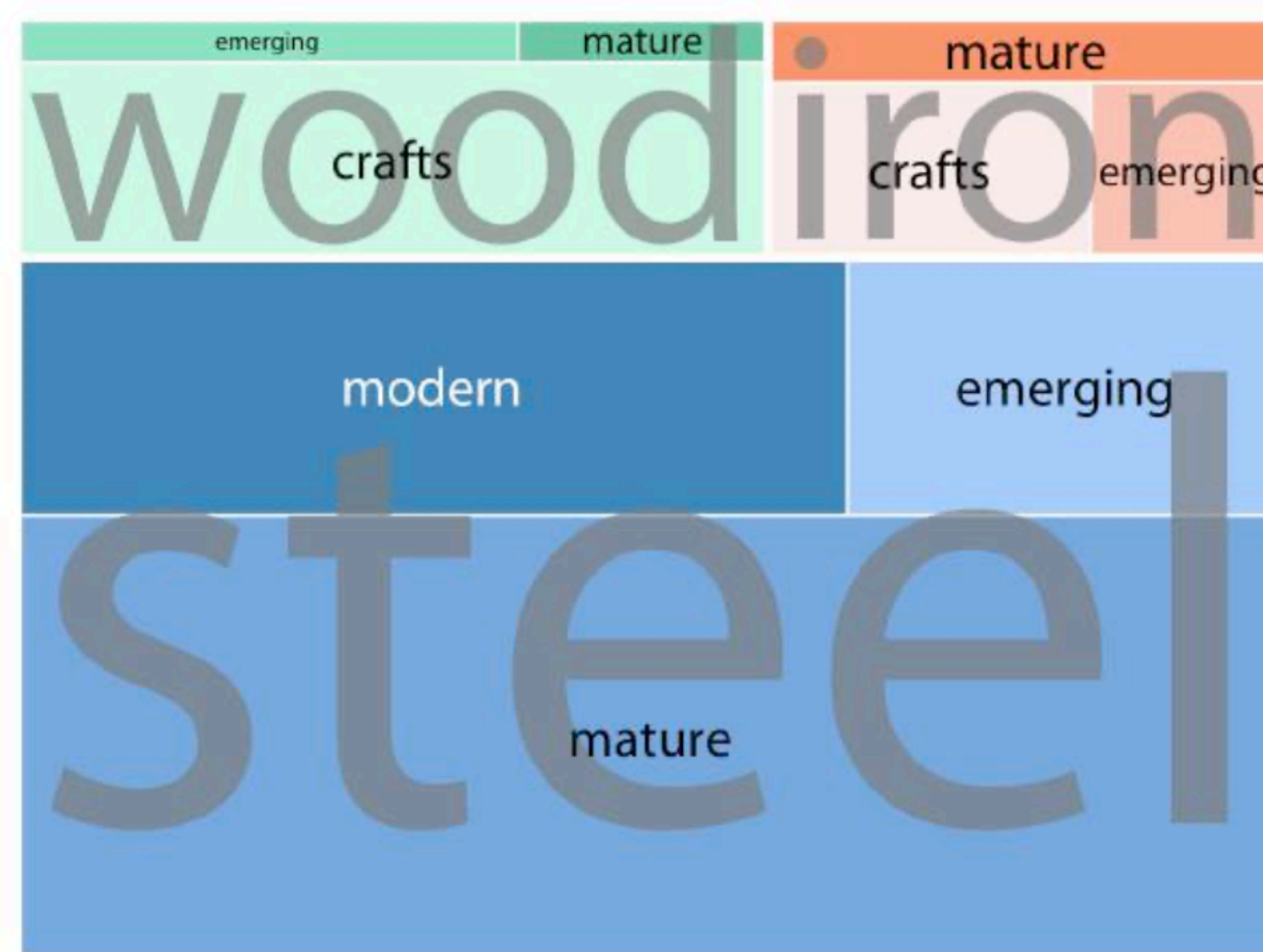


Figure 11-3. Breakdown of bridges in Pittsburgh by construction material (steel, wood, iron) and by era of construction (crafts, emerging, mature, modern), shown as a mosaic plot. The widths of each rectangle are proportional to the number of bridges constructed in that era, and the heights are proportional to the number of bridges constructed from that material. Numbers represent the counts of bridges within each category. Data source: Yoram Reich and Steven J. Fenves.

# Ağaç haritaları (treemaps)

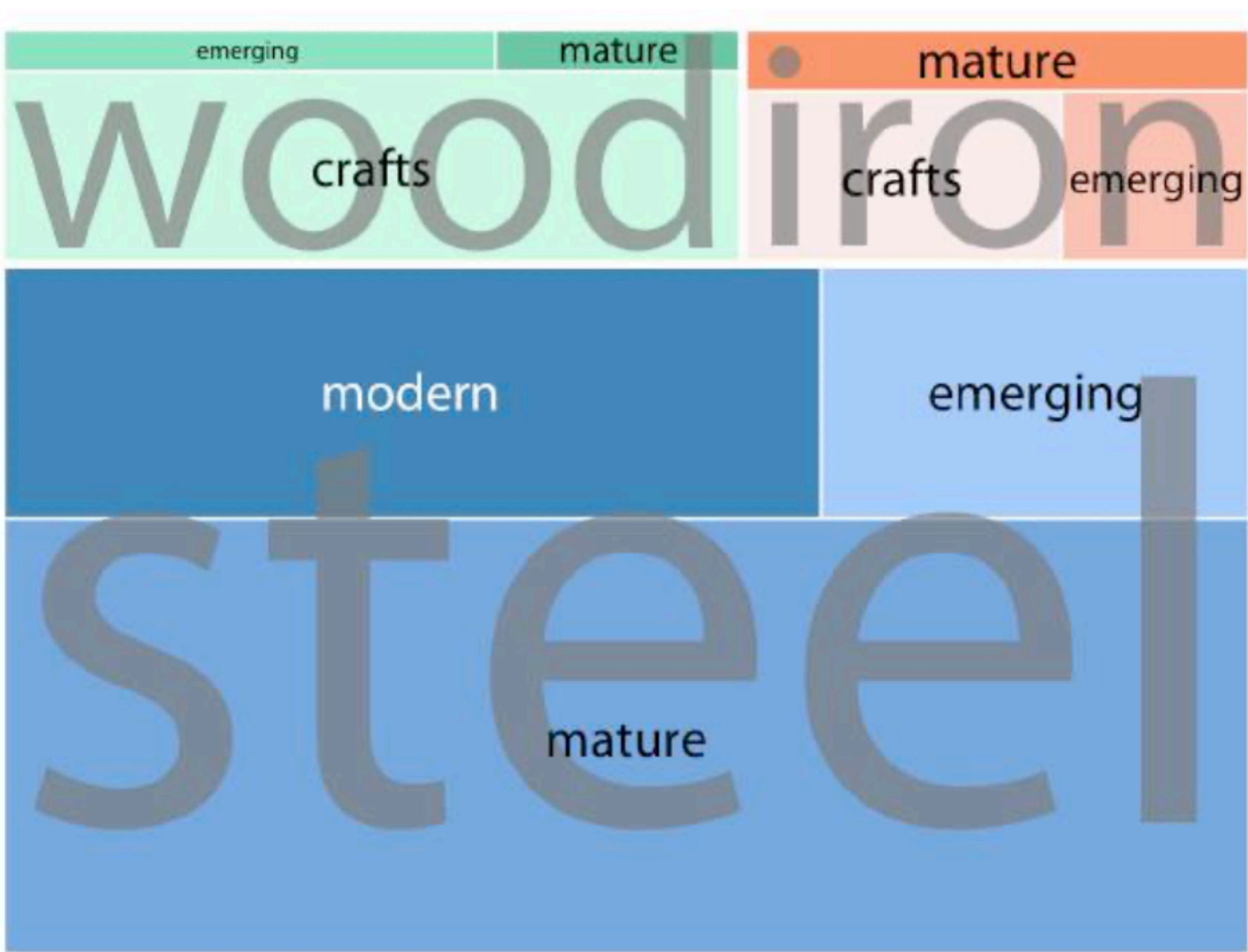
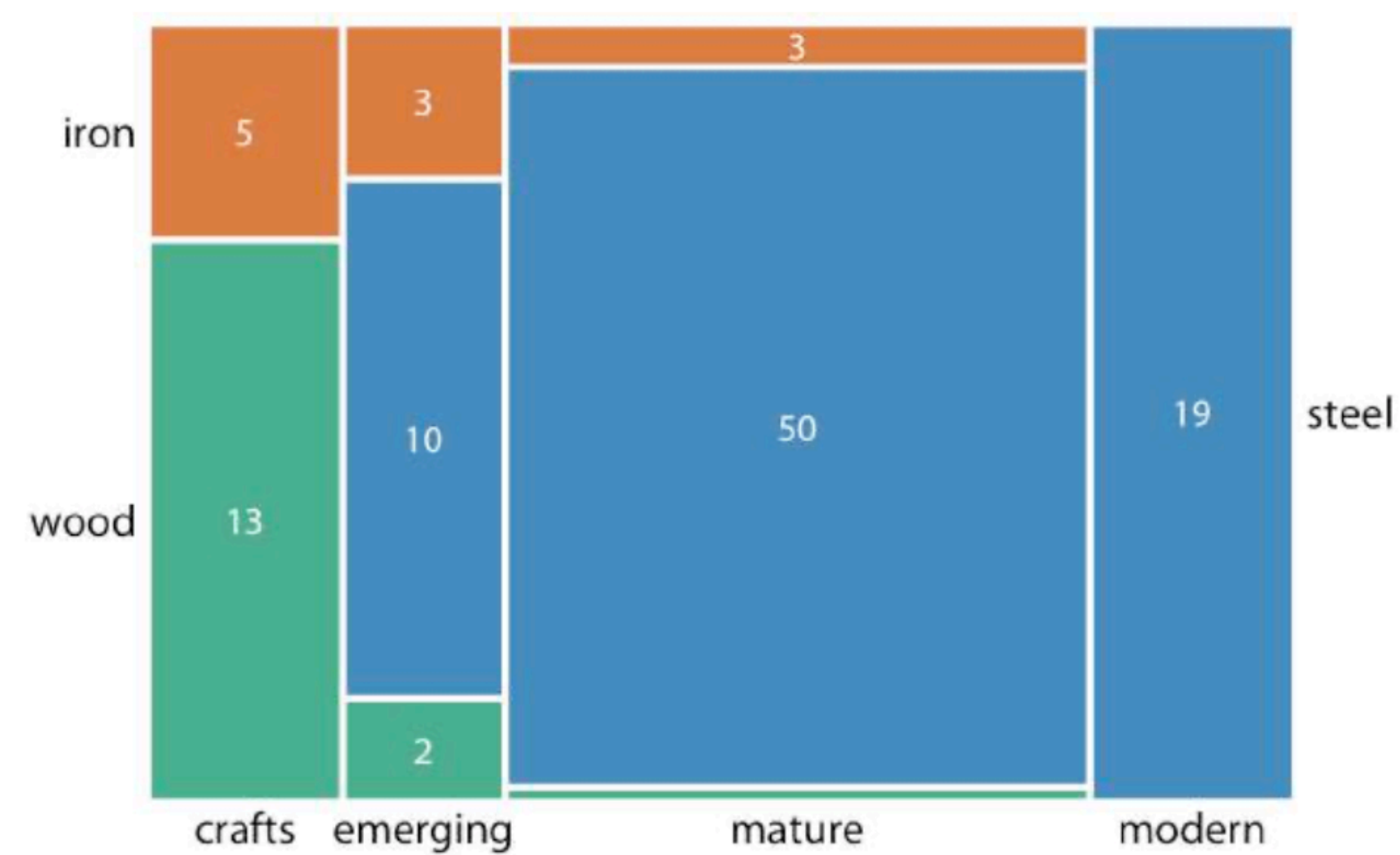
Mozaik grafiklerinden farklı olarak küçük dikdörtgenleri büyük olanın içine yerleştirme yöntemiyle ağaç haritaları (treemap) grafiği oluşturulur.



*Figure 11-4. Breakdown of bridges in Pittsburgh by construction material (steel, wood, iron) and by era of construction (crafts, emerging, mature, modern), shown as a tree-map. The area of each rectangle is proportional to the number of bridges of that type. Data source: Yoram Reich and Steven J. Fenves.*



# Mozaik grafiği vs. ağaç haritaları



# Mozaik grafiđi vs. ağaç haritaları

## Vurgu noktaları farklıdır:

Mozaik grafiđi zanaat çağından modern çađa yapı malzemesi kullanımındaki zamansal evrimi vurgularken, ağaç haritası çelik, demir ve ahşaptan imal edilen köprü sayılarını vurgulamaktadır.

# Sankey diyagramı

Pasta, mozaik grafikleri ve ağaç haritaları, kategorik değişkenin düzey sayısına arttıkça okunması zorlaşabilir. Bu durumda Sankey diyagramı kullanılabilir.

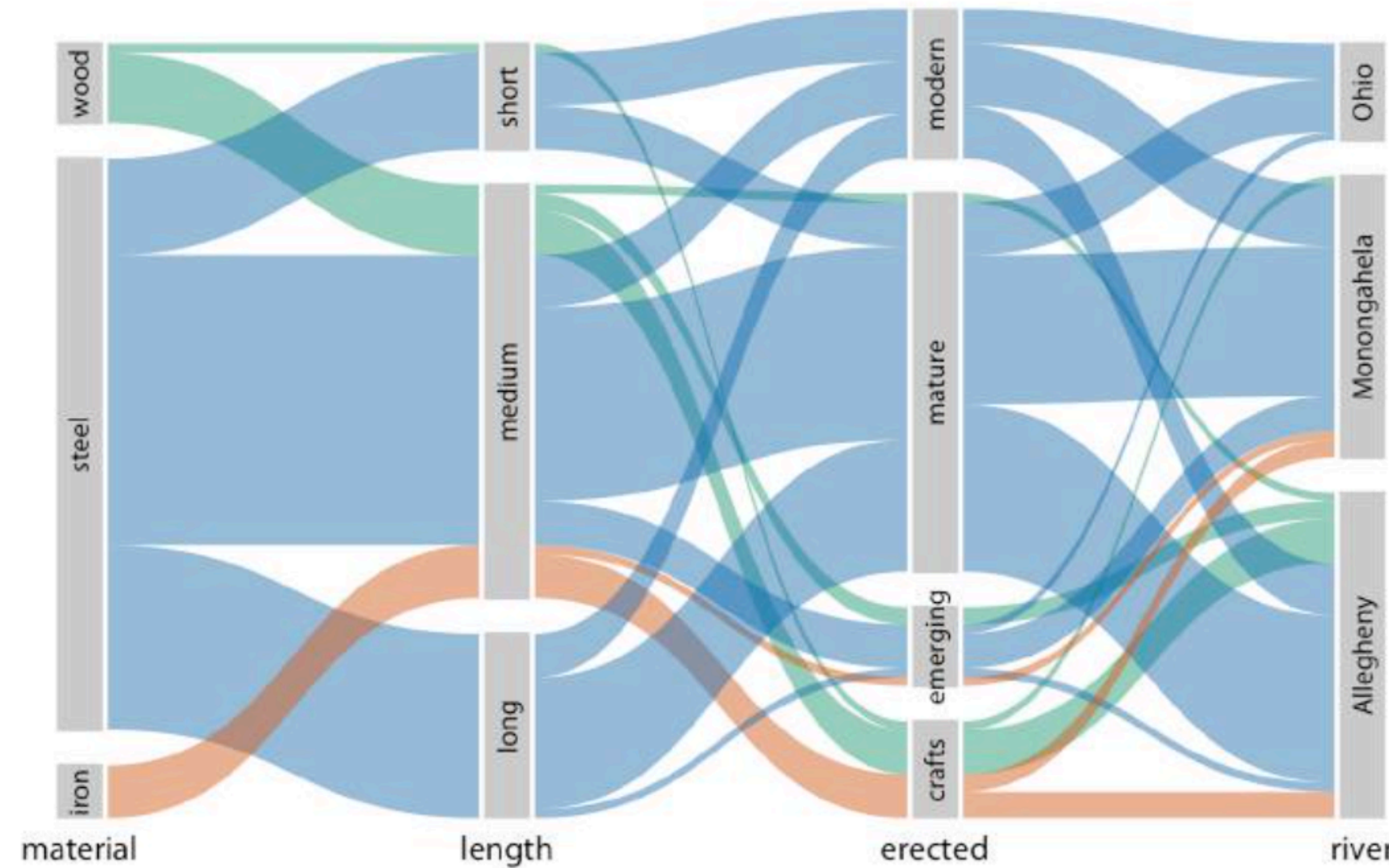


Figure 11-8. Breakdown of bridges in Pittsburgh by construction material, length, era of construction, and the river they span, shown as a parallel sets plot. The coloring of the bands highlights the construction material of the different bridges. Data source: Yoram Reich and Steven J. Fenves.



# Kaynak

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

