Stem-and-leaf Display

2016314726 정영준

Stem-and-leaf(줄기 잎 그림)을 만드는 방법에 대해 설명하기전 먼저 줄기 잎 그림에 대해 설명하겠습니다. 줄기 잎 그림은 통계적 자료를 표와 그래프가 혼합된 형태로 나타내는 것 입니다. 변수의 값을 자릿수를 기준으로 나누어서 앞자리를 줄기로 뒷자리를 잎으로 나타내는 그래프입니다. 장점은 자료가 크기 순서대로 나열되기 때문에 사분위 수나 중간값을 쉽게 알 수 있습니다. 분포의 전체적은 모습을 한눈에 확인할 수 있는것도 줄기 잎 그림의 장점입니다. 이상치를 쉽게 찾을 수 있습니다. 하지만 자료의 수가 매우 많은 경우 이를 전부 줄기 잎 그림에 나타내는 것은 힘들고 한눈에 알아볼 수 없게 됩니다. 그리는 방법은 먼저 관측값을 주로 자리수를 기준으로 줄기 부분과 잎부분으로 나눕니다. 일반적으로 줄기는 두 자리수 이상이 될 수 있지만 잎은 한자리수로 설정합니다. 잎을 마지막 한 자리수로 정하고 나머지 앞 자리수를 줄기로 정합니다. 다음으로 줄기를 열을 맞추어 오름차순으로 작은값부터 위에서부터 아래로 나열합니다. 그리고 줄기 옆에 선을 그어 잎과의 경계를 표시합니다. 그리고 각 줄기에 해당하는 잎을 표시한 경계의 우측에 맞추어 작성합니다. 보통 좌측에 작은 값 우측에 큰 값으로 오름차순으로 잎을 나열합니다. 예시로 실험적 자료 변수 시험점수가 90, 88, 82, 76, 72, 72, 68, 62, 60, 54, 52, 51, 50, 49, 48, 43, 38, 27, 22, 18, 14, 13 일 때 이를 줄기 잎 그림으로 나타내보겠습니다. 먼저 모든 자료가 두자리 수로 일의 자리와 십의자리로 나눌 수 있습니다. 십의 자리를 줄기, 일의 자리를로 정하겠습니다. 먼저 줄기를 쭉 나열하면 1, 2, 3, 4, 5, 6, 7, 8, 9가 됩니다. 줄기 1에 해당하는 잎은 3, 4, 8이 있고 줄기 2에 해당하는 잎은 2, 7이 있으며 줄기 3에는 8, 줄기 4에는 8, 9, 줄기 5에는 0, 1, 2, 4, 줄기 6에는 0, 2, 8, 줄기 7에는 2, 2, 6, 줄기 8에는 2, 8, 줄기 9에는 0의 잎을 가집니다. 잎을 각 줄기에 맞는 행에 가로로 나열하고 줄기를 세로로 열 벡터로 오름차순으로 나열하면 다음과 같은 줄기 잎 그림이 완성됩니다.

|  |  |
| --- | --- |
| 1 | 3 4 8 |
| 2 | 2 7 |
| 3 | 8 |
| 4 | 8 9 |
| 5 | 0 1 2 4 |
| 6 | 0 2 8 |
| 7 | 2 2 6 |
| 8 | 2 8 |
| 9 | 0 |

줄기 잎 그림을 보면 최댓값 최솟값을 쉽게 확인할 수 있으며 전체적인 데이터의 분포를 대략 확인이 가능합니다. 이 줄기 잎 그림을 좌측으로 90도 회전시키면 잎의 개수가 y축이 되어 히스토그램이 됩니다. 한눈에 전체적인 데이터의 분포를 확인할 수 있는 줄기 잎 그림은 그리기가 쉽고 간단하지만 자료가 많을 때 사용이 번거롭고 가시성이 떨어진다는 단점이 있습니다.