

3. 양적 자료의 그래프

양적 자료인 경우에 다음과 같은 그래프를 그려 분석한다.

- 줄기와 잎 그림
- 히스토그램 -
- 상대도수 - 도수분포다각형

3.1 줄기와 잎 그림

👉 생각열기	<p>요사이 미세먼지가 자주 발생해 우리 생활에 불편을 주고 있다. 과연 한 달 중 며칠이나 미세먼지가 심하게 발생하는지 살펴보기 위해 서울의 미세먼지 농도를 조사한 자료가 다음과 같다.</p> <p>(자료 3.1) 2021년 2월 서울의 일평균 초미세먼지농도 ($\mu\text{g}/\text{m}^3$) (환경부 대기환경정보 http://www.airkorea.or.kr 참조)</p> <table><tr><td>39</td><td>18</td><td>20</td><td>22</td><td>16</td><td>44</td><td>59</td><td>18</td><td>16</td><td>23</td></tr><tr><td>53</td><td>76</td><td>77</td><td>76</td><td>37</td><td>15</td><td>13</td><td>17</td><td>24</td><td>42</td></tr><tr><td>46</td><td>30</td><td>18</td><td>25</td><td>34</td><td>24</td><td>11</td><td>14</td><td></td><td></td></tr></table>	39	18	20	22	16	44	59	18	16	23	53	76	77	76	37	15	13	17	24	42	46	30	18	25	34	24	11	14		
39	18	20	22	16	44	59	18	16	23																						
53	76	77	76	37	15	13	17	24	42																						
46	30	18	25	34	24	11	14																								
탐구	<p>1) 미세먼지농도 자료수가 28개나 있는데 전반적인 자료의 분포를 어떻게 쉽게 표현할 수 있을까?</p> <p>2) 우리나라는 미세먼지 농도가 $36(\mu\text{g}/\text{m}^3)$을 넘으면 ‘나쁨’으로 으로 평가하는데 2월 중 며칠이나 있을까?</p>																														

- 위의 예에서 미세먼지농도를 39, 18, 20, ... 과 같이 측정하였는데 이와 같이 자료를 수량으로 나타낸 것을 **변량**이라 한다.
- (자료 3.1)과 같이 숫자로 된 자료는 십진법을 사용하기 때문에 각각의 십 자릿수에 해당하는 자료를 모아 다음과 같은 표로 정리할 수 있다. 즉, 첫 번째 자료 39는 십 자릿수가 ‘3’ 이므로 이 자료를 세 번째 행에 적고, 그 다음 18은 십 자릿수가 ‘1’이므로 첫 번째 행에 적는다. 같은 방법으로 모든 자료를 정리하면 [표 3.1]과 같다.

[표 3.1] 미세먼지농도를 십의 자릿수에 대해 정리한 자료


십의 자릿수	자료
1	18 16 18 16 15 13 17 18 11 14
2	20 22 23 24 25 24
3	39 37 30 34
4	44 42 46
5	59 53
6	
7	76 77 76

- [표 3.1]에서 각각의 행(십 자릿수)은 ‘ $10\mu\text{g}/\text{m}^3$ 이상 $20\mu\text{g}/\text{m}^3$ 미만’, ‘20이상 ~ 30미만’, ... ‘70이상 ~ 80미만’과 같은 구간을 의미한다. 이 표에서 각각의 행에 나타난 자료의 일 자릿수만을 오름차순으로 정렬한 [표 3.2]를 **줄기와 잎 그림**이라 부른다. 줄기와 잎 그림에서는 십 자릿수를 나무의 ‘줄기’, 일 자릿수를 ‘잎’으로 부른다.

[표 3.2] 미세먼지농도의 일 자릿수를 오름차 순으로 정리한 줄기와 잎 그림

줄기 (십 자릿수)	잎 (일 자릿수)
1	1 3 4 5 6 6 7 8 8 8
2	0 2 3 4 4 5
3	0 4 7 9
4	2 4 6
5	3 9
6	
7	6 6 7


- [표 3.2]와 같은 줄기와 잎 그림을 관찰하면 미세먼지 농도가 ' $10\mu\text{g}/\text{m}^3$ 이상 $20\mu\text{g}/\text{m}^3$ 미만'인 날이 제일 많고, 그 다음은 ' $20\mu\text{g}/\text{m}^3$ 이상 $30\mu\text{g}/\text{m}^3$ 미만'임을 쉽게 알 수 있다. 그리고 자료가 오름차순으로 정렬이 되어있어 미세먼지농도가 '나뭇잎'인 $36\mu\text{g}/\text{m}^3$ 이상인 날을 쉽게 세어 볼 수 있는데 28일중에서 미세먼지농도가 '나뭇잎'인 수준이 10일이나 되어 심각한 공해 문제임을 잘 살펴볼수 있다.
- 자료가 많을 경우 이와 같이 수작업으로 줄기와 잎 그림을 그리는 것은 시간도 많이 걸리고 쉽지 않다. 『eStat』 소프트웨어를 이용하여 줄기와 잎 그림을 그려보자.

실습 3.1	『eStat』을 이용하여 (자료 3.1) 미세먼지 농도에 대한 줄기와 잎 그림을 그려보자.																
풀이	<ul style="list-style-type: none"> • 왼쪽의 QR을 이용해 『eStatH』 메뉴에서 '줄기와 잎 그림'을 선택하면 <그림 3.1>과 같은 창이 나타난다. • '자료 입력'에 미세먼지농도 자료를 입력하고 (전자책에서 자료를 복사하여 붙여넣기를 해도 됨) '주 제목'에 원하는 그림 제목을 입력한다. • [실행] 버튼을 클릭하면 <그림 3.1> 아래와 같은 줄기와 잎 그림이 나타난다. 																
	<div data-bbox="582 1496 1369 1937"> <p>줄기와 잎 그림 메뉴</p> <p>자료 입력: 39 18 20 22 16 44 59 18 16 23 53 76 77 76 37 15 13 17 24 42 46 30 18</p> <p>주 제목: 2021년 2월 서울의 미세먼지 농도</p> <p>실행 ** 최대 줄기 수 ≤ 30 **</p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> <p style="text-align: center;">2021년 2월 서울의 미세먼지 농도</p> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">줄기</th> <th style="text-align: left;">잎</th> </tr> </thead> <tbody> <tr><td>1</td><td>1 3 4 5 6 6 7 8 8 8</td></tr> <tr><td>2</td><td>0 2 3 4 4 5</td></tr> <tr><td>3</td><td>0 4 7 9</td></tr> <tr><td>4</td><td>2 4 6</td></tr> <tr><td>5</td><td>3 9</td></tr> <tr><td>6</td><td></td></tr> <tr><td>7</td><td>6 6 7</td></tr> </tbody> </table> </div> </div> <p style="text-align: center;"><그림 3.1> 미세먼지 농도의 줄기와 잎 그림</p>	줄기	잎	1	1 3 4 5 6 6 7 8 8 8	2	0 2 3 4 4 5	3	0 4 7 9	4	2 4 6	5	3 9	6		7	6 6 7
줄기	잎																
1	1 3 4 5 6 6 7 8 8 8																
2	0 2 3 4 4 5																
3	0 4 7 9																
4	2 4 6																
5	3 9																
6																	
7	6 6 7																

- 자료가 세 자릿수 이상이거나 소수점이 있는 자료는 마지막 자릿수를 앞으로 하고 그 앞의 숫자들을 앞으로 하여 줄기와 잎 그림을 그릴 수 있다.

<p>실습 3.2</p>	<p>우리나라의 2월 서울의 일별 최저기온을 조사하나 다음과 같다. 『eStat』을 이용하여 일별 최저기온에 대한 줄기와 잎 그림을 그려보자.</p> <p>(자료 3.2) 2021년 2월 서울의 일별 최저기온 (섭씨 도) (기상청)</p> <table border="1"> <tr> <td>-2.3</td><td>-8.2</td><td>-9.4</td><td>-7.4</td><td>-4.4</td><td>4.3</td><td>-2.6</td><td>5.4</td><td>-6.1</td><td>-1.5</td> </tr> <tr> <td>1.3</td><td>0.6</td><td>1.0</td><td>6.4</td><td>-5.2</td><td>-7.0</td><td>-10.4</td><td>-10.6</td><td>-7.1</td><td>5.5</td> </tr> <tr> <td>4.7</td><td>0.4</td><td>-3.1</td><td>-3.0</td><td>0.7</td><td>0.5</td><td>4.3</td><td>3.2</td><td></td><td></td> </tr> </table>	-2.3	-8.2	-9.4	-7.4	-4.4	4.3	-2.6	5.4	-6.1	-1.5	1.3	0.6	1.0	6.4	-5.2	-7.0	-10.4	-10.6	-7.1	5.5	4.7	0.4	-3.1	-3.0	0.7	0.5	4.3	3.2								
-2.3	-8.2	-9.4	-7.4	-4.4	4.3	-2.6	5.4	-6.1	-1.5																												
1.3	0.6	1.0	6.4	-5.2	-7.0	-10.4	-10.6	-7.1	5.5																												
4.7	0.4	-3.1	-3.0	0.7	0.5	4.3	3.2																														
<p>풀이</p>	<ul style="list-style-type: none"> • 왼쪽의 QR을 이용하여 나타나는 『eStatH』 메뉴에서 ‘줄기와 잎 그림’을 선택하면 <그림 3.2>와 같은 자료입력창이 나타난다. • ‘자료 입력’에 일별 최저기온 자료를 입력하고 ‘주 제목’에 원하는 그림 제목을 입력한다. • [실행] 버튼을 클릭하면 <그림 3.2>와 같은 줄기와 잎 그림이 나타난다. • 온도는 소수점과 음수가 있어 마지막 숫자를 앞으로 하여 줄기와 잎 그림을 그린다. <div data-bbox="571 1115 1362 1731"> <p>줄기와 잎 그림 메뉴</p> <p>자료 입력: -2.3 -8.2 -9.4 -7.4 -4.4 4.3 -2.6 5.4 -6.1 -1.5 1.3 0.6 1.0 6.4 -5.2 -7.0 -10.4 -10.6 -</p> <p>주 제목: 2021년 2월 서울의 일별 최저기온</p> <p>실행 ** 최대 줄기 수 ≤ 30 **</p> <div style="border: 1px solid black; padding: 10px; margin-top: 10px;"> <p style="text-align: center;">2021년 2월 서울의 일별 최저기온</p> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">줄기</th> <th style="text-align: left;">잎</th> </tr> </thead> <tbody> <tr><td>-10</td><td>6 4</td></tr> <tr><td>-9</td><td>4</td></tr> <tr><td>-8</td><td>2</td></tr> <tr><td>-7</td><td>4 1</td></tr> <tr><td>-6</td><td>0 1</td></tr> <tr><td>-5</td><td>2</td></tr> <tr><td>-4</td><td>4</td></tr> <tr><td>-3</td><td>1</td></tr> <tr><td>-2</td><td>0 6 3</td></tr> <tr><td>-1</td><td>5</td></tr> <tr><td>0</td><td>4 5 6 7</td></tr> <tr><td>1</td><td>0 3</td></tr> <tr><td>2</td><td></td></tr> <tr><td>3</td><td>2</td></tr> <tr><td>4</td><td>3 3 7</td></tr> <tr><td>5</td><td>4 5</td></tr> <tr><td>6</td><td>4</td></tr> </tbody> </table> </div> <p style="text-align: center;"><그림 3.2> 서울의 일별 최저기온의 줄기와 잎 그림</p> </div>	줄기	잎	-10	6 4	-9	4	-8	2	-7	4 1	-6	0 1	-5	2	-4	4	-3	1	-2	0 6 3	-1	5	0	4 5 6 7	1	0 3	2		3	2	4	3 3 7	5	4 5	6	4
줄기	잎																																				
-10	6 4																																				
-9	4																																				
-8	2																																				
-7	4 1																																				
-6	0 1																																				
-5	2																																				
-4	4																																				
-3	1																																				
-2	0 6 3																																				
-1	5																																				
0	4 5 6 7																																				
1	0 3																																				
2																																					
3	2																																				
4	3 3 7																																				
5	4 5																																				
6	4																																				


과제 3.1




다음은 2016년 현재 서울의 25개 행정구별 자전거 전용 도로 길이에 대한 자료이다. 『eStat』을 이용하여 줄기와 잎 그림 등을 그려 분석하라.

(자료 3.3) 2019년 서울의 자전거 도로 (단위 km) (서울통계정보시스템)

24	15	23	20	30	24	7	8	7	12	28	27	19	35	41
42	11	8	37	13	20	29	53	93	42					

과제 3.2	다음은 2020년 우리나라를 통과한 태풍의 최대 풍속에 대한 자료이다.
	1) 『eStat』을 이용하여 줄기와 잎 그림 등을 그려라.
	2) 태풍의 최대 풍속이 54m/초 이상이면 초강력 태풍으로 분류한다. 몇 개의 초강력 태풍이 통과하였는지 조사하라.
	(자료 3.4) 2020년 우리나라를 통과한 태풍의 최대 풍속 (단위 m/초) (기상청)
	40 22 21 29 19 22 24 45 49 55 24 27 29 35 19 24 35 40 56 24 21 43 18

3.2 히스토그램 - 도수분포표

 생각열기	<p>한 중학교 2학년 학생들의 몸무게를 조사한 자료가 (자료 3.5)와 같다.</p> <p>(자료 3.5) 한 중학교 2학년 학생들의 몸무게 (kg)</p> <table><tr><td>63</td><td>65</td><td>67</td><td>68</td><td>61</td><td>60</td><td>72</td><td>55</td><td>64</td><td>76</td><td>68</td><td>63</td><td>70</td><td>61</td><td>54</td></tr><tr><td>63</td><td>66</td><td>53</td><td>58</td><td>70</td><td>62</td><td>62</td><td>57</td><td>58</td><td>59</td><td>53</td><td>58</td><td>58</td><td>62</td><td>61</td></tr></table>	63	65	67	68	61	60	72	55	64	76	68	63	70	61	54	63	66	53	58	70	62	62	57	58	59	53	58	58	62	61
63	65	67	68	61	60	72	55	64	76	68	63	70	61	54																	
63	66	53	58	70	62	62	57	58	59	53	58	58	62	61																	
탐구	<p>1) 자료수가 30개나 되는데 학생들 몸무게의 분포를 어떻게 쉽게 그래프로 표현할 수 있을까?</p> <p>2) 학생들 몸무게가 70kg이상 75kg미만인 학생의 수가 얼마나 될까?</p>																														

- 위와 같은 몸무게 자료의 전반적인 분포를 보기 위해서는 앞에서 살펴본 줄기와 잎 그림을 생각할 수 있다. 하지만 십 자리수가 5, 6, 7만 있기 때문에 줄기와 잎그림으로는 세밀한 분포를 살펴보기가 어렵다. 그리고 몸무게 70kg이상 75kg 미만인 학생의 수를 쉽게 알아볼 수 없다. 조사한 자료에서 전반적인 분포나 특정한 정보를 알기 위해서는 자료를 적절히 정리할 필요가 있다.
- [표 3.3]은 몸무게 자료를 50kg에서 시작하여 5kg 간격으로 구간을 정하고 각

구간에 해당하는 학생들 몸무게를 조사하여 정리한 것이다. 이러한 자료 정리에는 줄기와 잎 그림이 유용하다.

[표 3.3] 한 중학교 2학년 학생들 몸무게를 각 구간별로 정리한 자료

몸무게 (kg)	자료	자료수
50 ^{이상} ~ 55 ^{미만}	53 53 54	3
55 ~ 60	55 57 58 58 58 58 59	7
60 ~ 65	60 61 61 61 62 62 62 63 63 63 64	11
65 ~ 70	65 66 67 68 68	5
70 ~ 75	70 70 72	3
75 ~ 80	76	1

- [표 3.3]과 같이 정리된 표를 이용하면 전반적인 자료의 분포가 ‘60kg이상 ~ 65kg 미만’인 학생이 제일 많고 그 다음이 ‘55kg이상 ~ 60kg미만’임을 쉽게 알 수 있다. 그리고 ‘70kg이상 ~ 75kg미만’인 학생의 수가 3명임을 바로 알 수 있다.
- [표 3.3]과 같이 몸무게라는 변량을 일정한 간격으로 나눈 구간을 **계급**, 구간의 너비를 **계급의 크기**, 각 계급에 속하는 변량의 개수를 **도수**라 부르고 이를 정리한 표를 **도수분포표**라고 한다. [표 3.4]는 학생들 몸무게의 도수분포표이다.

[표 3.4] 한 중학교 2학년 학생들
몸무게의 도수분포표

계급 (kg)	도수 (개)
50 ^{이상} ~ 55 ^{미만}	3
55 ~ 60	7
60 ~ 65	11
65 ~ 70	5
70 ~ 75	3
75 ~ 80	1
합계	30

- 각 계급을 대표하는 값으로서 각 계급의 양 끝값의 중간값을 그 계급의 **계급값**이라 한다.

$$\text{계급값} = \frac{\text{계급의 양 끝값의 합}}{2}$$

예를 들어 [표 3.4]의 도수분포표에서 ‘50kg이상 ~ 55kg 미만’인 계급의 계급값은 다음과 같다.

$$\text{계급 '50kg이상 ~ 55kg 미만'의 계급값} = \frac{50 + 55}{2} = 52.5(\text{kg})$$

- 도수분포표에서 각 계급의 도수를 비교하면 전반적인 자료의 분포를 관찰할 수

있다. 하지만 각 계급의 도수가 전체 도수에서 차지하는 비율을 계산하면 더 좋을 수 있다. 전체 도수에 대한 각 계급의 도수의 비율을 그 계급의 **상대도수**라고 한다.

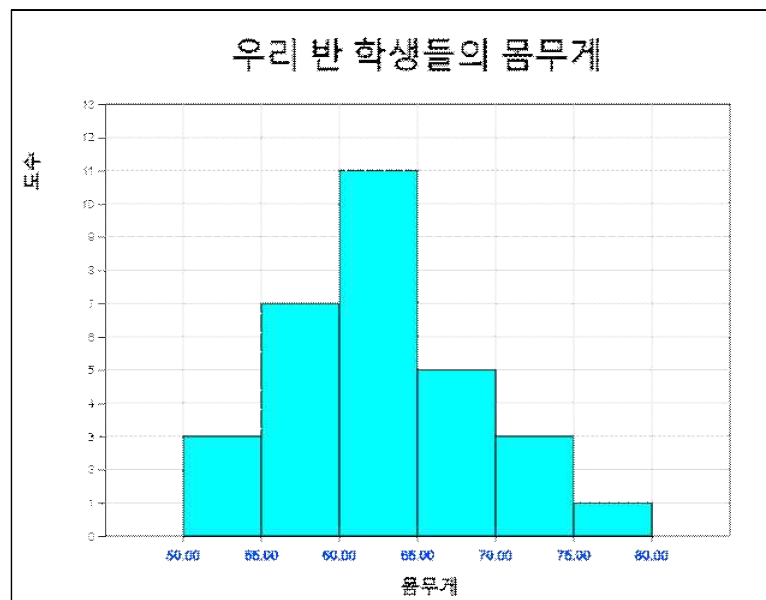
$$\text{계급의 상대도수} = \frac{\text{계급의 도수}}{\text{전체 도수의 합}}$$

- [표 3.5]는 도수분포표에 계급값과 상대도수를 표시한 도수분포표의 변형이다.

[표 3.5] 계급값과 상대도수가 추가된 학생들 몸무게의 도수분포표

계급 (kg)	계급값	도수(개)	상대도수
50 ~ 55	52.5	3	0.10
55 ~ 60	57.5	7	0.23
60 ~ 65	62.5	11	0.37
65 ~ 70	67.5	5	0.17
70 ~ 75	72.5	3	0.10
75 ~ 80	77.5	1	0.03
합계		30	1.00

- 도수분포표는 다음 순서에 따라 그래프로 나타낼 수 있는데 이를 **히스토그램**이라고 한다. <그림 3.3>은 학생들 몸무게에 대한 히스토그램이다.
 - ① 가로축에 각 계급의 끝값을 적는다.
 - ② 세로축에 도수를 적는다.
 - ③ 각 계급에서 계급의 크기를 가로로, 도수를 세로로 하는 직사각형을 그린다.



<그림 3.3> 학생들 몸무게의 히스토그램

- 도수분포표의 계급은 분석자가 정하는 계급의 크기에 따라 여러 가지로 만들 수 있다. (자료 3.5)의 학생들 몸무게 자료를 계급의 크기를 10kg으로 하여 만든 도수분포표는 다음 표와 같다. 이 도수분포표는 10자리수를 이용하여 그린 줄기와 옆 그림의 도수분포표이기도 하다.

[표 3.6] 계급구간을 10kg으로 하여 만든 학생들 몸무게의 도수분포표

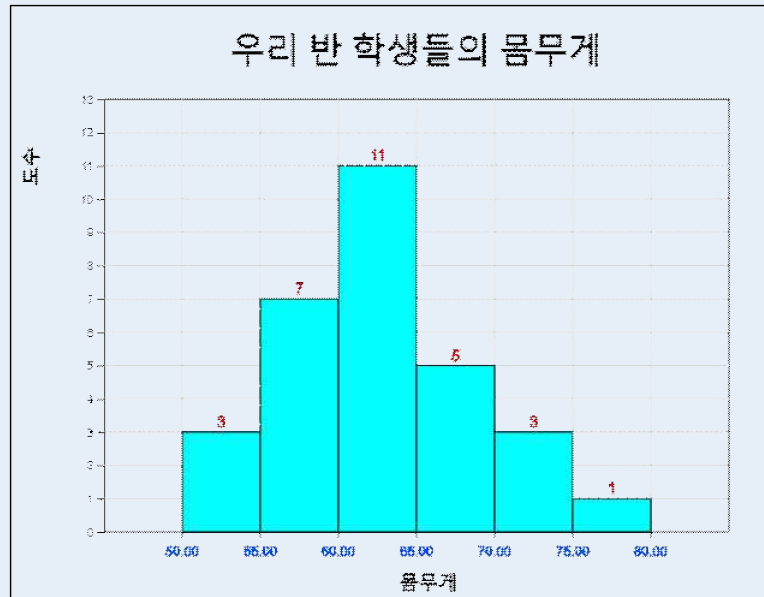
계급 (kg)	도수(개)
50 ^{아래} ~ 60 ^맨	10
60 ~ 70	16
70 ~ 80	4
합계	30

- 자료가 많을 경우 위와 같이 수작업으로 도수분포표와 히스토그램을 그리는 것은 시간도 많이 걸리고 쉽지 않다. 『eStat』 소프트웨어를 이용하여 도수분포표와 히스토그램을 그려보자.

<p>실습 3.3</p>	<p>『eStat』을 이용하여 (자료 3.5) 2학년 학생들의 몸무게에 대한 히스토그램을 그리고 이의 도수분포표를 알아보자.</p>																		
<p>풀이</p> <div data-bbox="343 1444 518 1624"> </div>	<ul style="list-style-type: none"> • 왼쪽의 QR을 이용해 『eStatH』 메뉴에서 ‘히스토그램 - 도수분포표’를 선택하면 <그림 3.4>와 같은 창이 나타난다. • ‘자료 입력’에 학생들의 몸무게 자료를 입력하고 (전자책에서 자료를 복사하여 붙여넣기를 해도 됨) ‘주 제목’에 원하는 그림 제목을 입력한다. • [실행] 버튼을 클릭하면 <그림 3.3>과 같은 히스토그램이 그려진다. <div data-bbox="614 1456 1332 1792"> <p>히스토그램 - 도수분포표 메뉴</p> <p>[자료 입력] 63 65 67 68 61 60 72 55 64 76 68 63 70 61 54 63 66 53 58 70 62 64</p> <table> <tr> <td>자료수</td><td>n</td><td>30</td> <td>최솟값</td><td>min</td><td>53.00</td> </tr> <tr> <td>표본평균</td><td>\bar{x}</td><td>62.23</td> <td>최댓값</td><td>max</td><td>76.00</td> </tr> <tr> <td>표본분산(n-1)</td><td>s^2</td><td>32.12</td> <td>표본표준편차</td><td>s</td><td>5.67</td> </tr> </table> <p>구간시작 <input type="text" value="50"/> ($\leq min$) 세로축 제목: <input type="text" value="도수"/></p> <p>구간너비 <input type="text" value="5"/> 가로축 제목: <input type="text" value="몸무게"/></p> <p>주 제목: <input type="text" value="우리반 학생들의 몸무게"/> 히스토그램 색: <input type="color" value="#00FFFF"/></p> <p><input type="button" value="실행"/></p> </div> <p style="text-align: center;"><그림 3.4> 히스토그램 자료 입력창</p>	자료수	n	30	최솟값	min	53.00	표본평균	\bar{x}	62.23	최댓값	max	76.00	표본분산(n-1)	s^2	32.12	표본표준편차	s	5.67
자료수	n	30	최솟값	min	53.00														
표본평균	\bar{x}	62.23	최댓값	max	76.00														
표본분산(n-1)	s^2	32.12	표본표준편차	s	5.67														

실습 3.3 풀이 (계속)

- 히스토그램 밑에 선택 사항에서 '평균'과 '도수'를 체크하면 <그림 3.5>와 같이 히스토그램 막대위에 도수가 표시된다.



<그림 3.5> 도수가 표시된 히스토그램

- 히스토그램 밑의 선택 사항에서 '도수분포표' 버튼을 클릭하면 <그림 3.6>과 같이 히스토그램의 도수분포표가 표시된다.

히스토그램 도수분포표			
계급구간	계급값	도수	상대도수
50.00 ≤ ~ < 55.00	52.50	3	0.10
55.00 ≤ ~ < 60.00	57.50	7	0.23
60.00 ≤ ~ < 65.00	62.50	11	0.37
65.00 ≤ ~ < 70.00	67.50	5	0.17
70.00 ≤ ~ < 75.00	72.50	3	0.10
75.00 ≤ ~ < 80.00	77.50	1	0.03
합계		30	1.00
계급값 평균	62.67		

<그림 3.6> 히스토그램에 대한 도수분포표

- 도수분포표의 계급 구간의 결정은 분석자가 자료의 최솟값과 최댓값을 살펴보고 결정한다.

실습 3.4

『eStat』을 이용하여 우리나라의 2월 서울의 일별 최저기온([실습 3.2])을 조사한 (자료 3.2)에 대하여 히스토그램을 그려보자.

(자료 3.2) 2021년 2월 서울의 일별 최저기온 (섭씨 도) (기상청)

-2.3 -8.2 -9.4 -7.4 -4.4 4.3 -2.6 5.4 -6.1 -1.5
1.3 0.6 1.0 6.4 -5.2 -7.0 -10.4 -10.6 -7.1 5.5
4.7 0.4 -3.1 -3.0 0.7 0.5 4.3 3.2

풀이

- 왼쪽의 QR을 이용하여 나타나는 『eStatH』 메뉴에서 ‘히스토그램 - 도수분포표’를 선택하면 <그림 3.7>과 같은 자료 입력 창이 나타난다.
- ‘자료 입력’에 일별 최저기온 자료를 입력하면 (전자책에서 자료를 복사하여 붙여넣기를 해도 됨) 즉시 <그림 3.7>과 같이 입력된 자료수 28과 최솟값이 -10.6도이고 최댓값이 6.4도임을 보여준다. 이 정보를 이용하여 구간시작과 구간너비를 정할 수 있다. 여기서는 구간시작을 -15, 구간너비를 5도로 정하였다.
- 원하는 제목을 입력하고 [실행] 버튼을 클릭하면 <그림 3.8>과 같은 히스토그램이 나타난다.



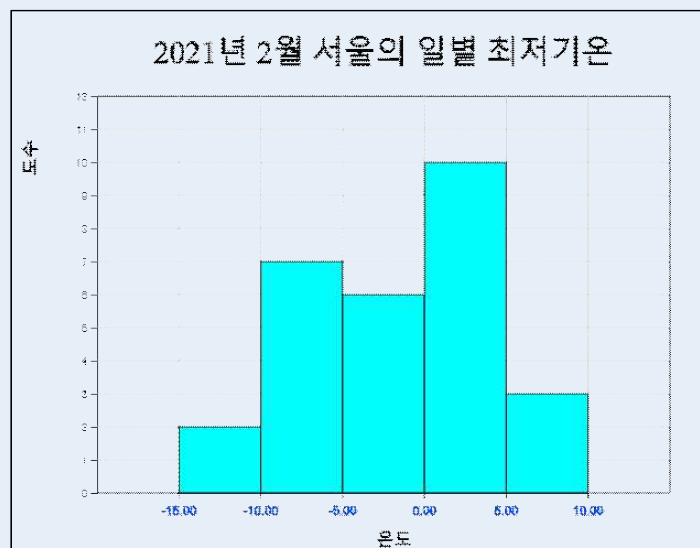
히스토그램 - 도수분포표 메뉴

[자료 입력] -2.3 -8.2 -9.4 -7.4 -4.4 4.3 -2.6 5.4 -6.1 -1.5 1.3 0.6 1.0 6.4 -5.2 -7.0 -10.4 -10.6 -7.1 5.5 4.7 0.4 -3.1 -3.0 0.7 0.5 4.3 3.2

자료수	n	28	최솟값	\min	-10.60
표본평균	\bar{x}	-1.79	최댓값	\max	6.40
표본분산($n-1$)	s^2	27.08	표본표준편차	s	5.20

구간시작 ($\leq \min$) 세로축 제목:
 구간너비 가로축 제목:
 주 제목: 히스토그램 색:

<그림 3.7> 히스토그램의 자료 입력창



<그림 3.8> 서울의 최저기온에 대한 히스토그램

실습 3.4 풀이 (계속)

히스토그램 도수분포표			
계급구간	계급값	도수	상대도수
$-15.00 \leq \sim < -10.00$	-12.50	2	0.07
$-10.00 \leq \sim < -5.00$	-7.50	7	0.25
$-5.00 \leq \sim < 0.00$	-2.50	6	0.21
$0.00 \leq \sim < 5.00$	2.50	10	0.36
$5.00 \leq \sim < 10.00$	7.50	3	0.11
합계		28	1.00
계급값 평균	-1.61		

<그림 3.9> 히스토그램에 대한 도수분포표

과제 3.3



다음은 2016년 현재 서울의 25개 행정구별 자전거 전용 도로 길이에 대한 자료이다 ([과제 3.1]). 『eStat』을 이용하여 히스토그램과 도수분포표를 작성하고 분석하라.

(자료 3.3) 2019년 서울의 자전거 도로 (단위 km) (서울통계정보시스템)

24 15 23 20 30 24 7 8 7 12 28 27 19 35 41 42 11 8 37 13
20 29 53 93 42

과제 3.4

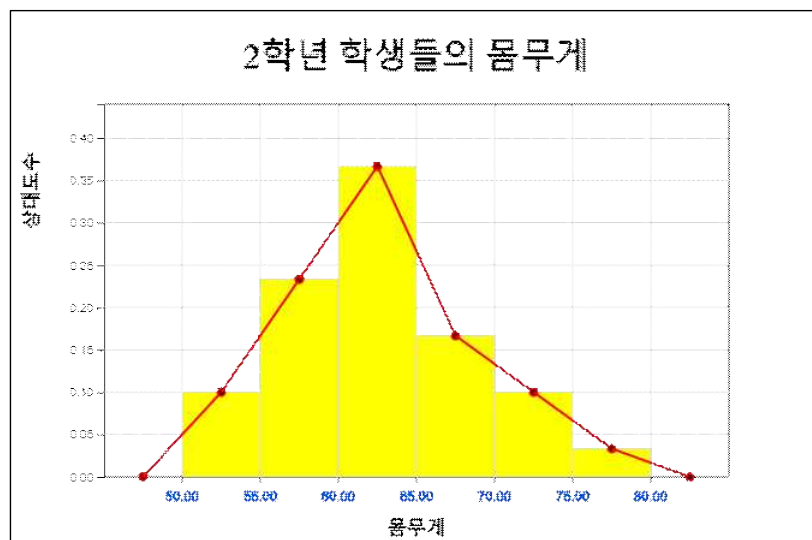


다음은 2020년 우리나라를 통과한 태풍의 최대 풍속에 대한 자료이다 ([과제 3.2]). 『eStat』을 이용하여 히스토그램과 도수분포표를 작성하고 분석하라.

(자료 3.4) 2020년 우리나라를 통과한 태풍의 최대 풍속 (단위 m/초) (기상청)

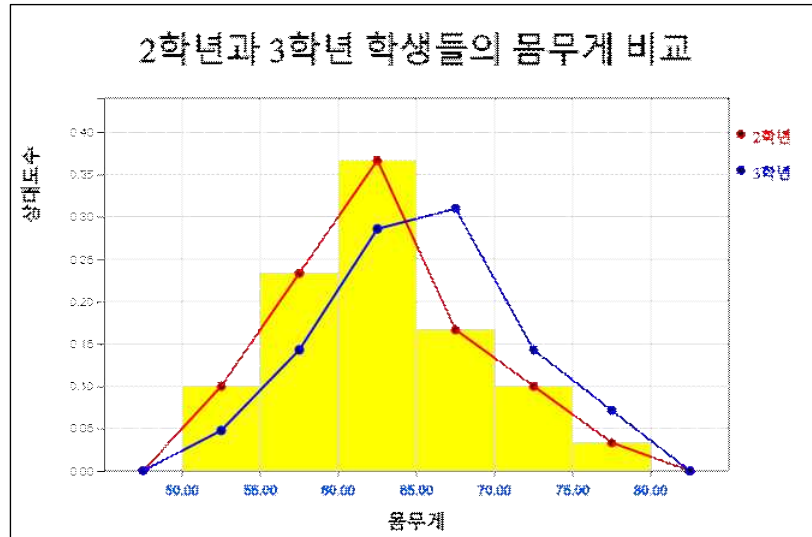
40 22 21 29 19 22 24 45 49 55 24 27 29 35 19 24 35 40 56 24
21 43 18

- 히스토그램을 이용하여 각 계급의 도수에 대하여 꺾은선 모양을 그린 그래프를 **도수분포다각형**이라 부른다. 도수분포다각형을 그리는 방법은 다음과 같다.
 - ① 히스토그램의 각 직사각형의 윗변의 중앙에 점을 찍는다.
 - ② 히스토그램의 양 끝 구간을 도수가 0인 계급이 하나씩 있는 것으로 생각하여 그 가운데 점을 찍는다.
 - ③ 위에서 찍은 점을 선분으로 연결한다.
- 히스토그램은 계급의 도수를 이용하여 그리는 것이 일반적이거나 상대도수를 이용하여 그릴 수 있다. 단지 도수대신 상대도수를 이용하는 것이라 그리는 방법은 같다. 도수분포다각형도 도수 또는 상대도수를 이용하여 그릴 수 있는데 [표 3.7]과 같이 2학년과 3학년 두 그룹에 대한 도수분포를 비교할 때는 각 그룹의 자료수가 다를 수 있어 상대도수를 이용한 도수분포다각형을 그려 비교한다.
- <그림 3.10>은 [표 3.7]에서 2학년 학생들 몸무게의 계급별 상대도수를 이용한 히스토그램과 도수분포다각형이다.



<그림 3.10> 계급의 상대도수를 이용한 히스토그램과 도수분포다각형

- <그림 3.11>은 2학년과 3학년 학생들의 계급별 상대도수를 이용하여 도수분포다각형으로 비교한 것이다.



<그림 3.11> 2학년과 3학년 학생 몸무게의 도수분포다각형 비교

- 자료가 많을 경우 위와 같이 수작업으로 도수분포표와 히스토그램을 그리는 것은 시간도 많이 걸리고 쉽지 않다. 『eStat』 소프트웨어를 이용하여 도수분포표와 히스토그램을 그려보자.

실습 3.5

『eStat』을 이용하여 [표 3.7]의 2학년과 3학년 학생의 몸무게에 대한 히스토그램과 도수분포다각형을 그려보자.

풀이

- 왼쪽의 QR을 이용해 『eStatH』 메뉴에서 ‘도수분포다각형 - 상대도수비교’를 선택하면 <그림 3.12>와 같은 창이 나타난다.
- 원하는 제목을 입력한 후 계급 구간의 왼쪽값을 그림과 같이 입력한 후 ‘도수 1’ 열에 2학년 도수를 입력한다.



도수분포다각형 - 상대도수 비교

주 제목 : 2학년 학생들의 몸무게
 세로축 제목 : 상대도수
 가로축 제목 : 몸무게

계급	2학년	도수 2	상대도수 1	상대도수 2
1 50 ~ 55.00	3		0.100	
2 55 ~ 60.00	7		0.233	
3 60 ~ 65.00	11		0.367	
4 65 ~ 70.00	5		0.167	
5 70 ~ 75.00	3		0.100	
6 75 ~ 80.00	1		0.033	
7				
8				
9				
합계	30		1.000	

[실행]

<그림 3.12> 2학년 몸무게의 도수분포다각형을 위한 자료 입력

- [실행] 버튼을 클릭하면 <그림 3.10>과 같이 2학년 학생들의 히스토그램과 도수분포다각형이 그려진다.

실습 3.5 풀이 (계속)

- 이어서 다음과 같이 3학년 학생들의 몸무게 도수를 추가로 입력하고, 제목을 바꾼 후 [실행] 버튼을 클릭하면 <그림 3.11>과 같은 2학년과 3학년 학생들의 몸무게에 대한 도수분포다각형이 그려진다.

도수분포다각형 - 상대도수 비교 메뉴

주 제목 : 2학년과 3학년 학생들의 몸무게 비교

세로축 제목 : 상대도수

가로축 제목 : 몸무게

계급	2학년	3학년	상대도수 1	상대도수 2
1 50 ≤ < 55.00	3	2	0.100	0.048
2 55 ≤ < 60.00	7	6	0.233	0.143
3 60 ≤ < 65.00	11	12	0.367	0.286
4 65 ≤ < 70.00	5	13	0.167	0.310
5 70 ≤ < 75.00	3	6	0.100	0.143
6 75 ≤ < 80.00	1	3	0.033	0.071
7				
8				
9				
합계	30	42	1.000	1.000

<그림 3.13> 2학년과 3학년 도수분포다각형을 위한 자료 입력

실습 3.6

2021년 우리나라의 연령대별 남녀인구가 다음 표와 같다. 『eStat』을 이용하여 남녀별 도수분포다각형을 그려 비교하라.

[표 3.8] 2021년 우리나라의 연령별 남녀인구 도수분포표 (통계청) (단위: 만명)

계급	도수(개)	
	남자	여자
0 ~ 20	437	411
20 ~ 40	737	659
40 ~ 60	851	827
60 ~ 80	504	557
80 ~ 100	67	132
합계	2596	2586

실습 3.6 풀이

- 왼쪽의 QR을 이용해 『eStatH』 메뉴에서 ‘도수분포다각형 - 상대도수비교’를 선택하면 <그림 3.14>와 같은 창이 나타난다.
- 원하는 제목을 입력한 후 계급 구간의 왼쪽값을 그림과 같이 입력한 후 ‘도수 1’ 열에 남자, ‘도수 2’열에 인구를 입력한다.

도수분포다각형 - 상대도수 비교 메뉴

주 제목 : 2021년 우리나라의 연령별 남녀인구

세로축 제목 : 상대도수

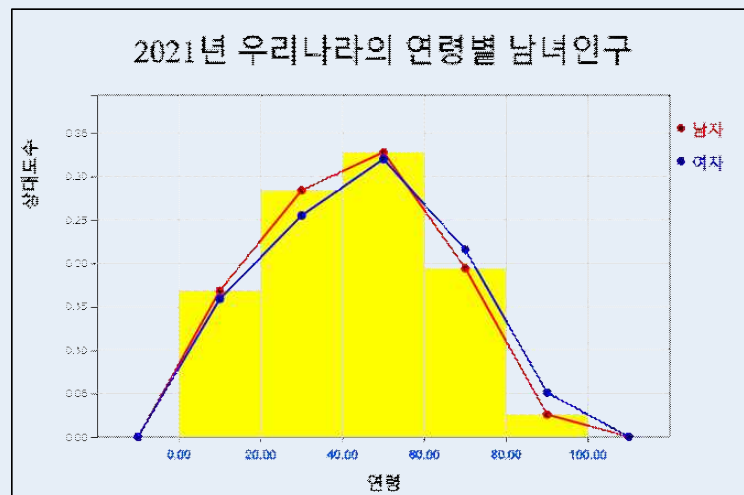
가로축 제목 : 연령

계급	남자	여자	상대도수 1	상대도수 2
1 0 ~ 20.00	437	411	0.168	0.159
2 20 ~ 40.00	737	659	0.284	0.255
3 40 ~ 60.00	851	827	0.328	0.320
4 60 ~ 80.00	504	557	0.194	0.215
5 80 ~ 100.00	67	132	0.026	0.051
6				
7				
8				
9				
합계	2596	2586	1.000	1.000

실행

<그림 3.14> 상대도수다각형을 위한 남녀 인구자료 입력

- [실행] 버튼을 클릭하면 <그림 3.15>와 같이 남녀별 도수분포다각형이 그려진다. 남자의 인구가 60세까지는 여자보다 많다가 60세 이상은 여자의 인구가 더 많음을 쉽게 알 수 있다.



<그림 3.15> 우리나라 인구의 남녀별 상대도수다각형



과제 3.5

다음 표는 한 중학교 남녀 선생님의 나이를 조사한 것이다. 『eStat』을 이용하여 도수분포다각형을 그려 비교하라.



[표 3.9] 한 중학교 남녀 선생님의 나이별 도수분포표

계급	도수(개)	
	남자	여자
20 ^아 ~ 30 ^만	3	2
30 ~ 40	4	6
40 ~ 50	4	4
50 ~ 60	2	3
60 ~ 70	0	2
합계	13	17

과제 3.6

다음 표는 A 중학교와 B 중학교의 학력고사 성적을 비교한 것이다. 『eStat』을 이용하여 도수분포다각형을 그려 비교하라.



[표 3.10] 두 중학교 학력고사 성적의 도수분포표

계급	도수(개)	
	A 중학	B 중학
50 ^아 ~ 60 ^만	2	2
60 ~ 70	5	8
70 ~ 80	20	25
80 ~ 90	23	10
90 ~ 100	10	5
합계	60	50