2022학년도 1학기 기말시험(과제물)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **교과목명** | : | 통계패키지 |
|  | **학번** | : | 202135-368864 |
|  | **성명** | : | 홍원표 |
|  | **연락처** | : | 010-5343-4341 |

- 이하 작성

1. (25점) 다음을 SAS를 이용하여 작성하시오.

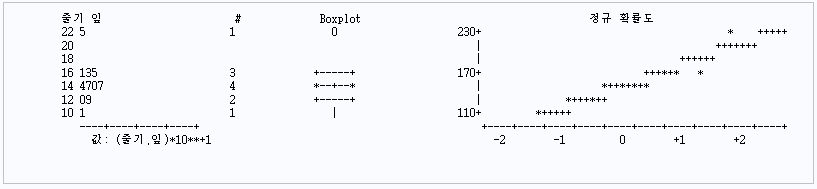
1) 교재 5장 연습문제(p.208) 2, 10번

**(5장)2.** 다음은 상장된 금융사 중 외국기업(F)과 국내기업(K)의 주가이다. 외국기업과 국내기업별로 기술통계량을 구하고 줄기-잎 그림, 상자그림을 그려서 비교 • 분석하라.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| F 120  K 165  K 147  F 144 | K 135  K 161  K 102  K 165 | K 170  F 147  F 235  F 161 | F 139  F 150  F 157  K 173 | K 114  F 175  K 129  F 111 | F 163  K 145  K 129  K 145 |
|  |  |  |  |  |  |

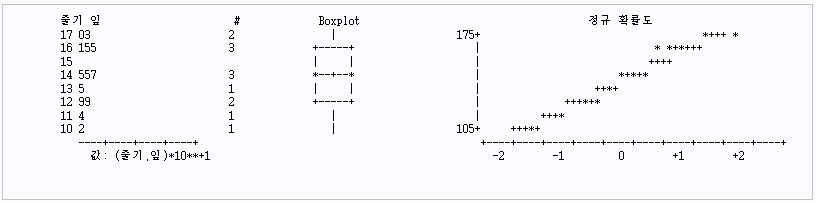


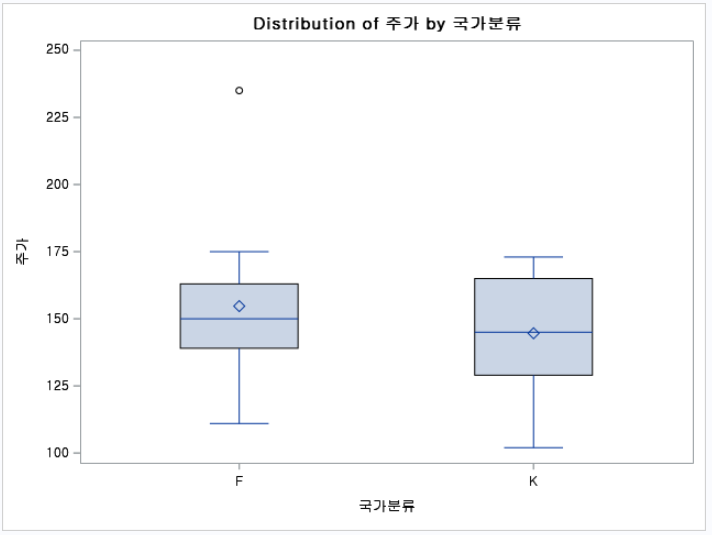




해외 주가 데이터에는 결측치로 보이는 동떨어져 있는 값이 존재한다. 따라서 결측치로 인한 평균값이 높아졌을 것이다. 결측치를 제외하고 평균을 구하면 국내의 평균과 비슷한 수준으로 확인된다.



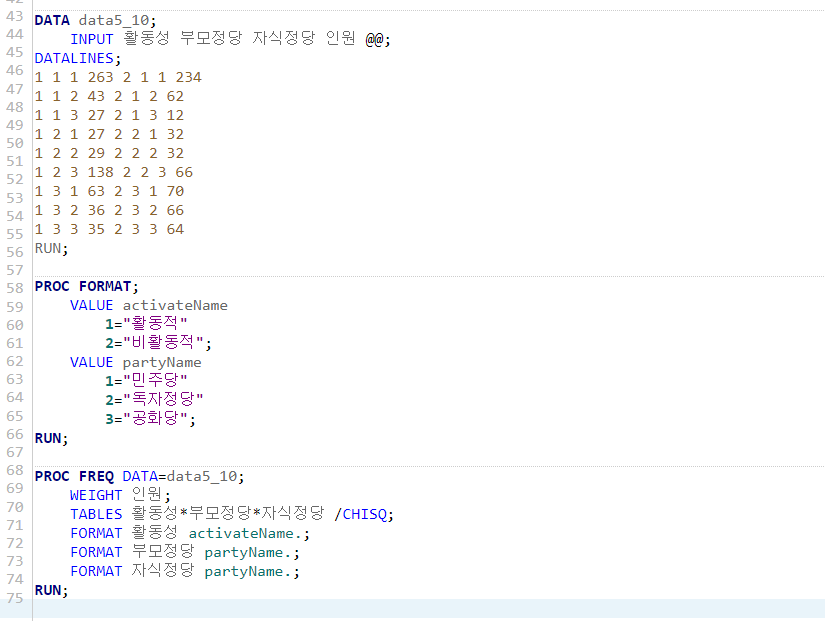


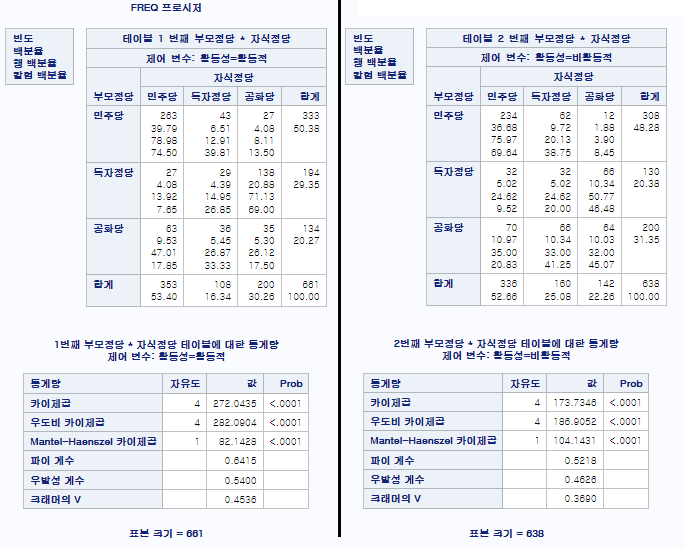


상자그림으로 비교해보면 해외 주가에 결측치를 쉽게 확인할 수 있고 결측치를 이외의 해외와 국내의 주가 분포는 크게 차이가 나지 않아 보인다.

**(5장)10.** 부모와 자식 간 정치 성향의 관련성에 관한 미국 연구자료이다. 부모의 정치활동 수준이 활동적인 경우와 비활동적인 경우 각각에 대해서 분할표를 만든 다음 카이제곱 검정을 시행하라. 어떤 결론을 내릴 수 있는가? 민주당은 진보적, 공화당은 보수적이다.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 부모의 정치활동 | 부모의 정당 | 자식의 정당 | | |
| 민주당 | 독자정당 | 공화당 |
| 활동적 | 민주당  독자정당  공화당 | 263  27  63 | 43  29  36 | 27  138  35 |
| 비활동적 | 민주당  독자정당  공화당 | 234  32  70 | 62  32  66 | 12  66  64 |



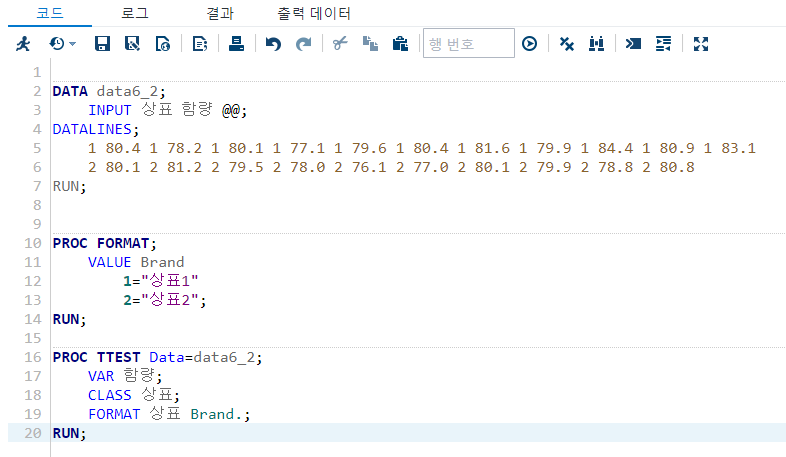


활동적인 부모와 비활동적인 부모 모두에서 부모정당과 자식정당 변수와의 관계는 카이제곱( 통계량의 자유도가 4인 에 대한 *P­*값은 0.0001이므로 유의수준 0.05보다 작기 때문에 귀무가설을 기각하고 대립가설을 채택하여 부모정당과 자식정당의 두 변수는 연관이 있다고 할 수 있다.

2) 교재 6장 연습문제(p.235) 3번, 5번

**(6장)3.** 어떤 화학 약품의 제조에 상표가 다른 두 종류의 원료가 사용되고 있다. 각 원료에서 주성분 A의 함량은 다음과 같다. 단, 함량은 정규분포를 따른다고 가정한다. 이 두 원료의 주성분 A의 함량이 다른지를 분석하라.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 상표1 | 80.4 | 78.2 | 80.1 | 77.1 | 79.6 | 80.4 | 81.6 | 79.9 | 84.4 | 80.9 | 83.1 |
| 상표2 | 80.1 | 81.2 | 79.5 | 78.0 | 76.1 | 77.0 | 80.1 | 79.9 | 78.8 | 80.8 |  |





분산에 대한 검정

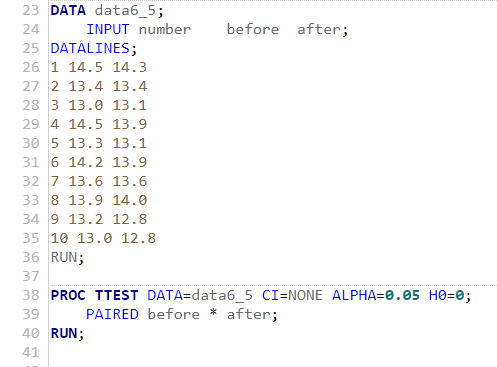
,

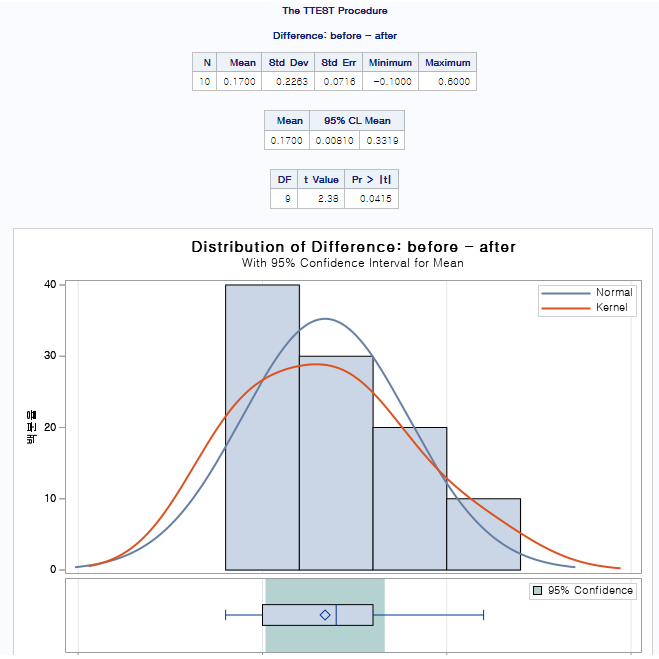
F통계량()에 대한 이 보다 값이 크므로 두 집단의 분산이 동일하다는 가정을 기각할 수 없다는 것을 알 수 있다. 따라서 분산이 동일하다는 가설을 받아들이고 ‘등분산을 가정함’의 결과를 이용한다. 만약 분산이 다를 경우 COCHRAN 통계량을 구해서 검정을 실시해야 한다.

등분산을 가정했을 때, 자유도 19인 검정 통계량 t값은 1.676이고, 이에 대한 유의 확률 으로 유의수준 0.05보다 크므로 이 두 원료의 주성분 A의 함량은 같다는 귀무가설을 기각하지 못한다.

**(6장)5.** 고등학교 육상선수에게 체중감량을 시키면 달리기에 어떤 변화가 일어나는지 조사하기 위하여 어떤 고등학교 육상선수 10명에게 감량훈련(2~4kg 정도)을 시킨 후 100m 달리기를 실시하여 다음의 데이터를 얻었다. 감량전과 후에 달리기 속도에 차이가 있는지 대응 • 비교하라.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 선수 번호 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 감량 전 | 14.5 | 13.4 | 13.0 | 14.5 | 13.3 | 14.2 | 13.6 | 13.9 | 13.2 | 13.0 |
| 감량 후 | 14.3 | 13.4 | 13.1 | 13.9 | 13.1 | 13.9 | 13.6 | 14.0 | 12.8 | 12.8 |





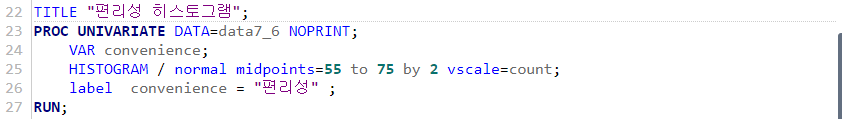
자유도는 9, t값은 2.38이고 P값은 0.0415로 감량전과 후의 속도는 차이가 있다고 볼 수 있다.

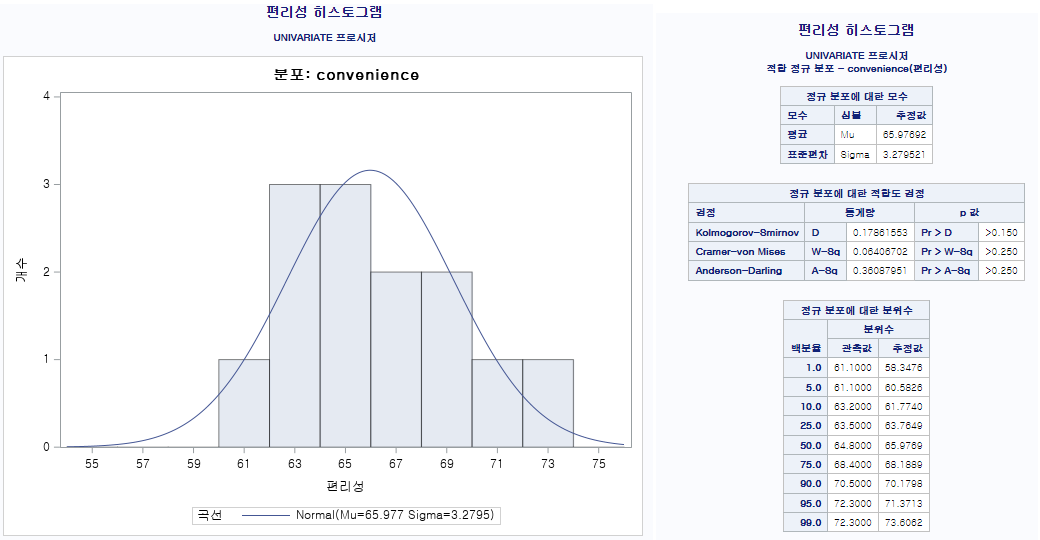
3) 교재 7장 연습문제(p.277) 6번

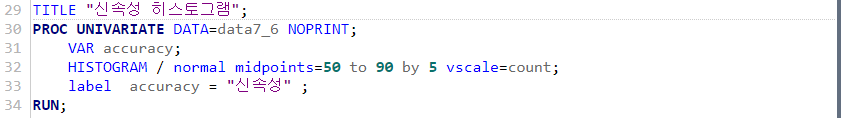
**(7장)6.** 다음은 13개의 시중은행에 대한 편리성, 신속성, 친절, 능률, 쾌적, 자동화 등의 점수를 나타내고 있는 자료이다.

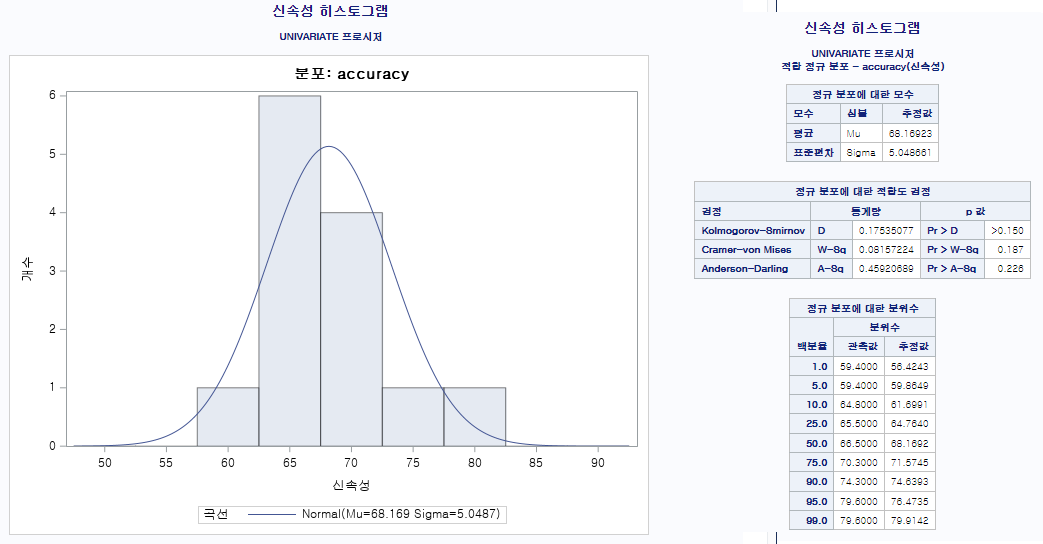
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | convenience | accuracy | kindness | efficiency | pleasant | automatic |
| Kookmin  Enterpr  Boram  Commerce  Seoul  Shinhan  City  Exchange  First  Chohung  Hana  Hanil  House | 70.5  64.8  67.1  61.1  63.4  72.3  64.2  68.4  66.1  63.5  69  63.2  64.1 | 59.4  70.3  79.6  65  66.5  69.1  72  67.5  66.5  65.7  74.3  65.5  64.8 | 63.7  68.6  78.5  65.6  67.9  74.2  71.4  67.3  67.3  64.3  80.5  68.3  67.8 | 54.3  55.2  62.4  54.4  65  60  56.9  51.3  50.7  53.9  63.6  49.8  59.7 | 66.9  68  79.8  64.5  59.7  70.1  72.8  71.3  63.4  61.7  75.7  64.6  65.7 | 62.6  64.1  62.4  63.9  62  68.2  57.8  65.8  63.3  62.7  55.9  59.1  61.8 |

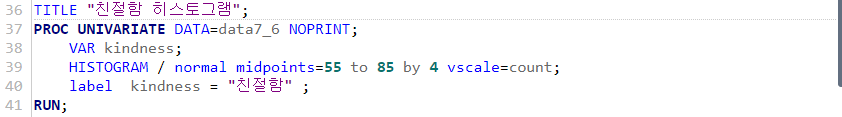
1. 각 변수의 히스토그램을 그리고 설명하라.

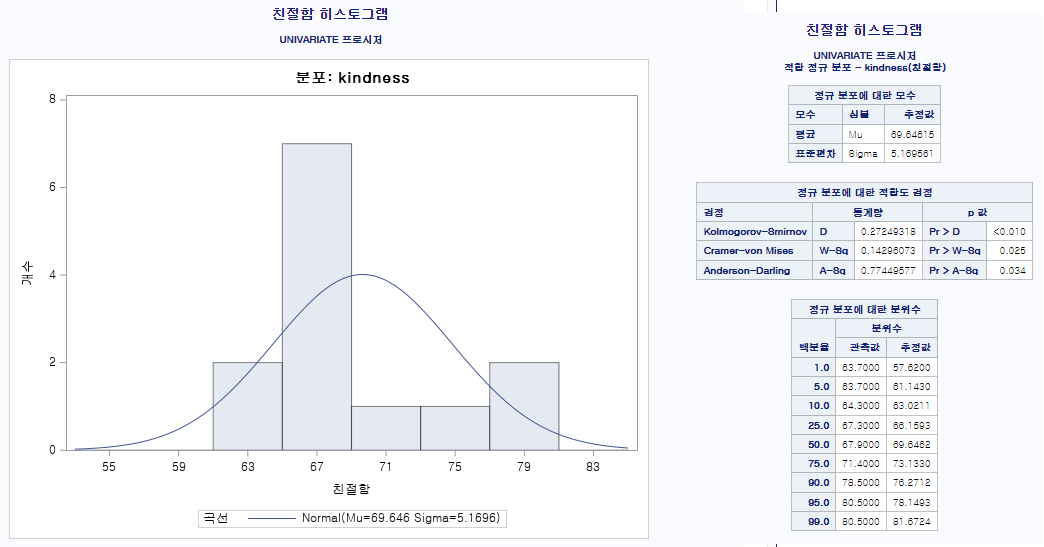


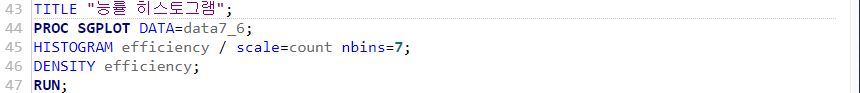


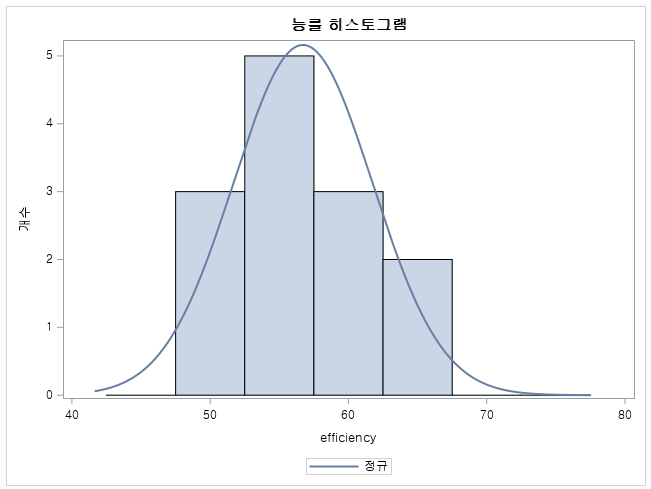




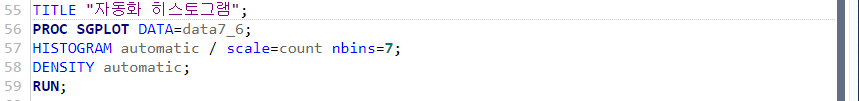


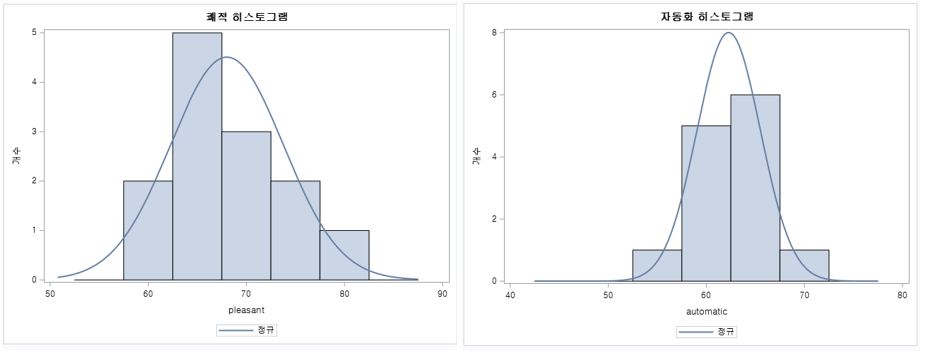






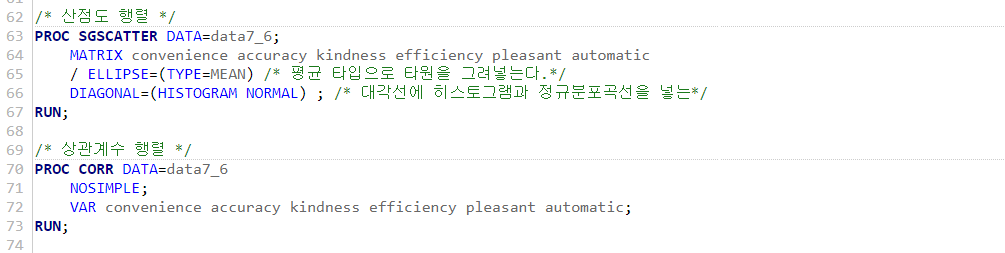


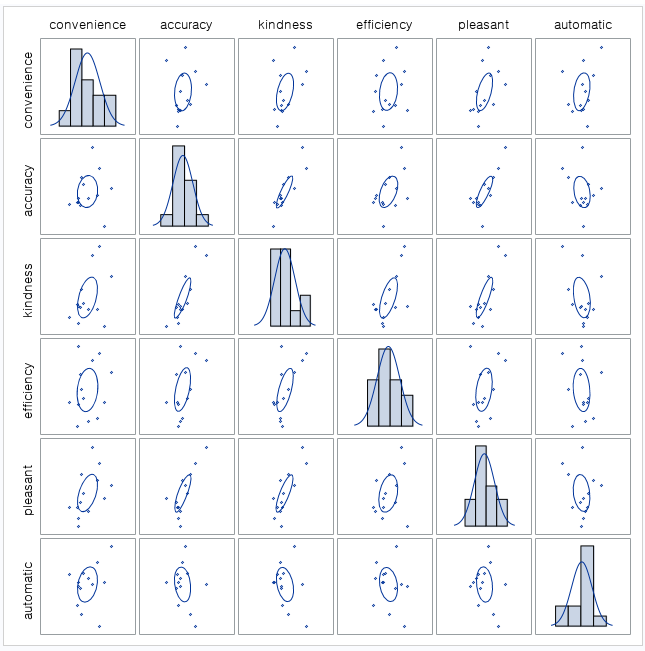




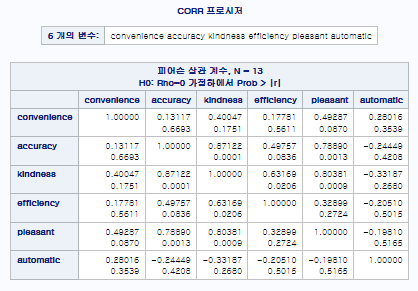
친절 히스트그램을 제외한 편리성, 신속성, 능률, 쾌적, 자동화 등의 전수의 분포는 정규분포에 근접하지만 친절 변수에서는 많은 은행들이 평균보다 낮은 점수를 받고 있다.

1. 산점도 행렬 및 상관계수 행렬을 구하고, 변수들의 관계를 설명하라.





타원이 많이 찌그러진 그림이 있는 변수들의 관계가 선형관계를 쉽게 알 수 있다. 위의 행렬에서 보면 convenience와 automatic은 다른 변수와는 선형관계가 뚜렷이 보이지 않는다. Accuracy\*kindness, accuracy\*pleasant, kindness\*pleasant 변수끼리 선형관계가 뚜렷해 보이고 convenience\* pleasant, kindness\*efficient 변수끼리 선형관계가 약하게 있는 것으로 볼 수 있다.



Accuracy\*kindness, accuracy\*pleasant, kindness\*pleasant의 *P*­값은 0.001 이하로 나타나고 있으며, convenience\* pleasant *P*­값은 0.0206, kindness\*efficient *P*­값은 0.0836으로 나타나고 있다. Accuracy\*kindness, accuracy\*pleasant, kindness\*pleasant, convenience\* pleasant 변수끼리 상관관계가 있다고 할 수 있다.

2. (25점) 다음을 SPSS를 이용하여 작성하시오.

1) 교재 7장 연습문제(p.277) 3번

**(7장)3.** 다음은 어떤 공정에서 생산되는 기계부품의 길이(mm)를 측정한 값이다. 줄기-잎 그림, 히스토그램, 상자그림을 그리고 설명하라

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 16.22  15.93  16.08  16.12  16.00  16.07  15.94  16.26 | 15.75  15.96  16.19  15.70  16.19  16.11  15.74  15.96 | 15.71  15.96  15.78  15.92  15.93  15.75  16.05  15.94 | 15.80  16.12  16.32  15.89  15.91  16.04  16.21  16.08 | 16.25  15.89  15.83  15.95  15.94  16.07  16.02  15.75 | 16.10  15.95  15.96  16.07  16.02  15.75  15.88  15.96 | 16.15  16.01  15.86  16.10  15.80  15.41  15.94  16.17 |

GET

FILE='C:\Users\robinhwp\Documents\7-3.sav'.

DATASET NAME 데이터세트1 WINDOW=FRONT.

EXAMINE VARIABLES=Length

/PLOT BOXPLOT STEMLEAF HISTOGRAM

/COMPARE GROUPS

/STATISTICS DESCRIPTIVES

/CINTERVAL 95

/MISSING LISTWISE

/NOTOTAL.

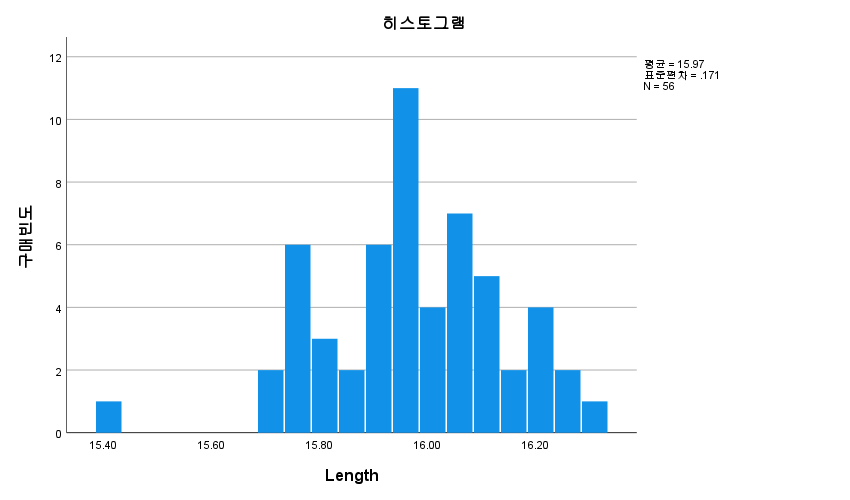
데이터 탐색

|  |  |  |
| --- | --- | --- |
| **노트** | | |
| 작성된 출력결과 | | 29-MAY-2022 07:58:37 |
| 주석 | |  |
| 입력 | 데이터 | C:\Users\robinhwp\Documents\7-3.sav |
| 활성 데이터 세트 | 데이터세트1 |
| 필터 | <없음> |
| 가중 | <없음> |
| 분할 파일 | <없음> |
| 작업 데이터 파일의 행 수 | 56 |
| 결측값 처리 | 결측값 정의 | 종속변수에 대한 사용자 정의 결측값은 누락된 데이터로 처리됩니다. |
| 사용 케이스 | 통계량은 사용된 종속변수나 요인에 대한 결측값이 없는 케이스를 기준으로 결정됩니다. |
| 명령문 | | EXAMINE VARIABLES=Length  /PLOT BOXPLOT STEMLEAF HISTOGRAM  /COMPARE GROUPS  /STATISTICS DESCRIPTIVES  /CINTERVAL 95  /MISSING LISTWISE  /NOTOTAL. |
| 사용된 자원 | 프로세서 시간 | 00:00:00.59 |
| 경과 시간 | 00:00:00.33 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **케이스 처리 요약(O)** | | | | | | |
|  | 케이스(S) | | | | | |
| 유효 | | 결측값 | | 총계 | |
| N(C) | 퍼센트(C) | N(C) | 퍼센트(C) | N(C) | 퍼센트(C) |
| Length | 56 | 100.0% | 0 | 0.0% | 56 | 100.0% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **기술통계** | | | | |
|  | | | 통계 | 표준화 오차 |
| Length | 평균 순위 | | 15.9725 | .02284 |
| 평균의 95% 신뢰구간 | 하한 | 15.9267 |  |
| 상한 | 16.0183 |  |
| 5% 절사평균 | | 15.9762 |  |
| 중위수(D) | | 15.9600 |  |
| 분산(V) | | .029 |  |
| 표준화 편차 | | .17091 |  |
| 최소값(U) | | 15.41 |  |
| 최대값(X) | | 16.32 |  |
| 범위(R) | | .91 |  |
| 사분위수 범위 | | .21 |  |
| 왜도(W) | | -.484 | .319 |
| 첨도(K) | | .886 | .628 |

Length



Length 줄기와 잎그림 도표

구매빈도 Stem & 잎

1.00 Extremes (=<15.41)

8.00 157 . 01455558

7.00 158 . 0036899

15.00 159 . 123344445566666

11.00 160 . 01224577788

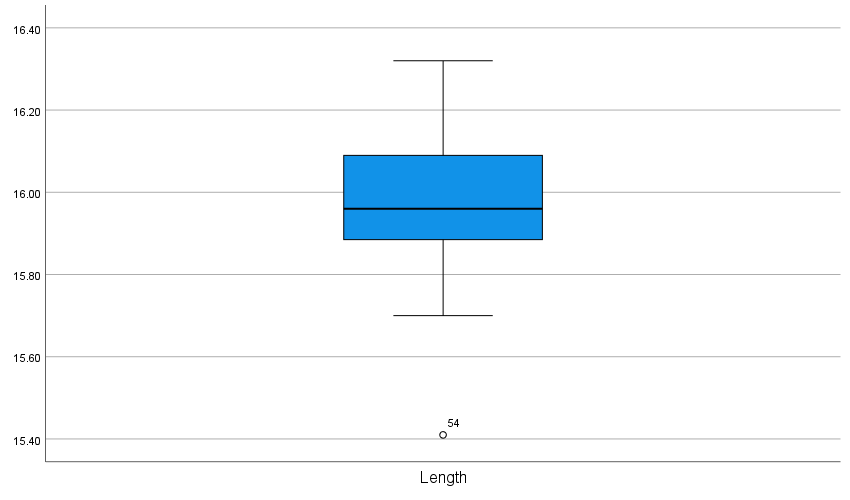
9.00 161 . 001225799

4.00 162 . 1256

1.00 163 . 2

줄기 너비: .10

각 잎: 1 케이스



대부분의 길이는 15.70 ~ 16.32 사이에 값들이 존재하고 결측치로 보이는 15.41 가 존재한다.

평균은 15.97이고 표준편차는 0.171이고 갯수는 56개이다.

2) 교재 8장 연습문제(p.305) 1번, 8번

**(8장)1.** 다음은 동물 25마리로부터 얻은 두 변수에 관한 자료이다. 산점도를 그리고 상관분석과 회귀분석을 하라.

X: arterial calcium deposition(mg/100g)

Y: serum cholesterol(mg/200mL)

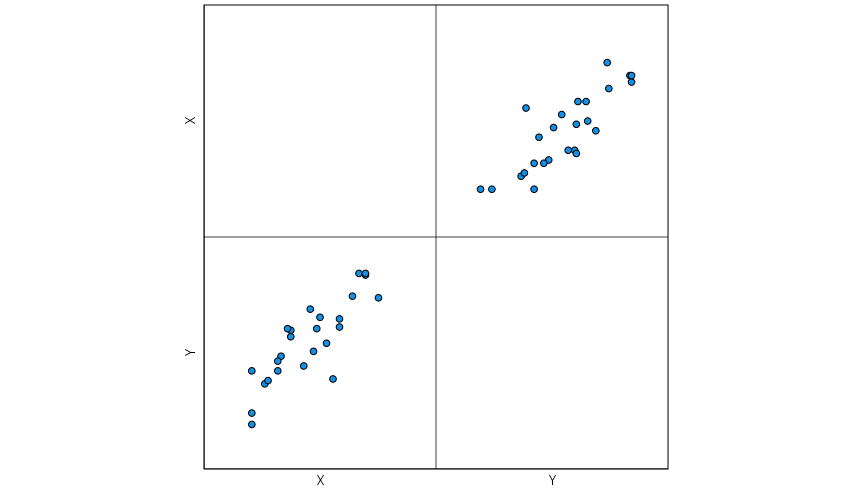
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 번호 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| X  Y | 59  298 | 24  239 | 36  264 | 57  299 | 63  284 | 32  239 | 40  242 | 24  213 | 24  206 | 59  299 | 36  260 | 42  277 | 55  285 |
| 번호 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |  |
| X  Y | 45  272 | 28  231 | 35  265 | 47  256 | 51  271 | 29  233 | 33  248 | 51  266 | 49  234 | 44  265 | 43  251 | 32  245 |  |

GRAPH

/SCATTERPLOT(MATRIX)=X Y

/MISSING=LISTWISE.

그래프



CORRELATIONS

/VARIABLES=X Y

/PRINT=TWOTAIL NOSIG FULL

/MISSING=PAIRWISE.

상관관계

|  |  |  |  |
| --- | --- | --- | --- |
| **상관관계** | | | |
|  | | X | Y |
| X | Pearson 상관 | 1 | .844\*\* |
| 유의확률 (양측) |  | .000 |
| N | 25 | 25 |
| Y | Pearson 상관 | .844\*\* | 1 |
| 유의확률 (양측) | .000 |  |
| N | 25 | 25 |

|  |
| --- |
| \*\*. 상관관계가 0.01 수준에서 유의합니다(양측). |

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Y

/METHOD=ENTER X

/SCATTERPLOT=(\*ZPRED ,\*ZPRED)

/RESIDUALS NORMPROB(ZRESID)

/CASEWISE PLOT(ZRESID) ALL.

회귀

|  |  |  |  |
| --- | --- | --- | --- |
| **입력/제거된 변수a** | | | |
| 모형 | 입력된 변수 | 제거된 변수 | 방법 |
| 1 | Xb | . | 입력 |

|  |
| --- |
| a. 종속변수: Y |
| b. 요청된 모든 변수가 입력되었습니다. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **모형 요약b** | | | | |
| 모형 | R | R 제곱 | 수정된 R 제곱 | 추정값의 표준오차 |
| 1 | .844a | .713 | .701 | 13.79286 |

|  |
| --- |
| a. 예측자: (상수), X |
| b. 종속변수: Y |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| 모형 | | 제곱합 | 자유도 | 평균제곱 | F | 유의확률 |
| 1 | 회귀 | 10875.852 | 1 | 10875.852 | 57.168 | .000b |
| 잔차 | 4375.588 | 23 | 190.243 |  |  |
| 전체 | 15251.440 | 24 |  |  |  |

|  |
| --- |
| a. 종속변수: Y |
| b. 예측자: (상수), X |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **계수a** | | | | | | |
| 모형 | | 비표준화 계수 | | 표준화 계수 | t | 유의확률 |
| B | 표준화 오류 | 베타 |
| 1 | (상수) | 183.424 | 10.201 |  | 17.981 | .000 |
| X | 1.788 | .237 | .844 | 7.561 | .000 |

|  |
| --- |
| a. 종속변수: Y |

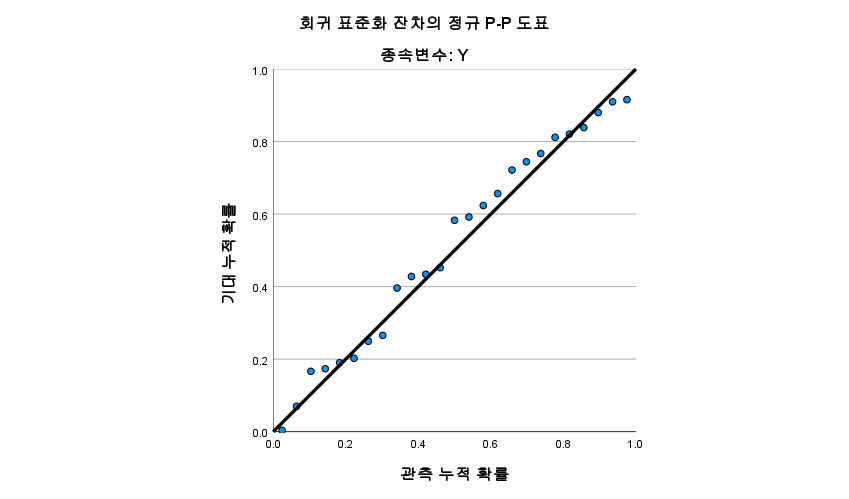
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **케이스별 진단a** | | | | |
| 케이스 번호 | 표준화 잔차 | Y | 예측값 | 잔차 |
| 1 | .657 | 298.00 | 288.9421 | 9.05789 |
| 2 | .917 | 239.00 | 226.3464 | 12.65365 |
| 3 | 1.174 | 264.00 | 247.8078 | 16.19225 |
| 4 | .989 | 299.00 | 285.3652 | 13.63479 |
| 5 | -.877 | 284.00 | 296.0959 | -12.09591 |
| 6 | -.120 | 239.00 | 240.6540 | -1.65395 |
| 7 | -.940 | 242.00 | 254.9616 | -12.96156 |
| 8 | -.968 | 213.00 | 226.3464 | -13.34635 |
| 9 | -1.475 | 206.00 | 226.3464 | -20.34635 |
| 10 | .729 | 299.00 | 288.9421 | 10.05789 |
| 11 | .884 | 260.00 | 247.8078 | 12.19225 |
| 12 | 1.338 | 277.00 | 258.5385 | 18.46154 |
| 13 | .233 | 285.00 | 281.7883 | 3.21169 |
| 14 | .587 | 272.00 | 263.9038 | 8.09619 |
| 15 | -.181 | 231.00 | 233.5002 | -2.50015 |
| 16 | 1.376 | 265.00 | 246.0193 | 18.98070 |
| 17 | -.832 | 256.00 | 267.4807 | -11.48071 |
| 18 | -.264 | 271.00 | 274.6345 | -3.63451 |
| 19 | -.166 | 233.00 | 235.2886 | -2.28860 |
| 20 | .403 | 248.00 | 242.4424 | 5.55760 |
| 21 | -.626 | 266.00 | 274.6345 | -8.63451 |
| 22 | -2.687 | 234.00 | 271.0576 | -37.05761 |
| 23 | .209 | 265.00 | 262.1154 | 2.88464 |
| 24 | -.676 | 251.00 | 260.3269 | -9.32691 |
| 25 | .315 | 245.00 | 240.6540 | 4.34605 |

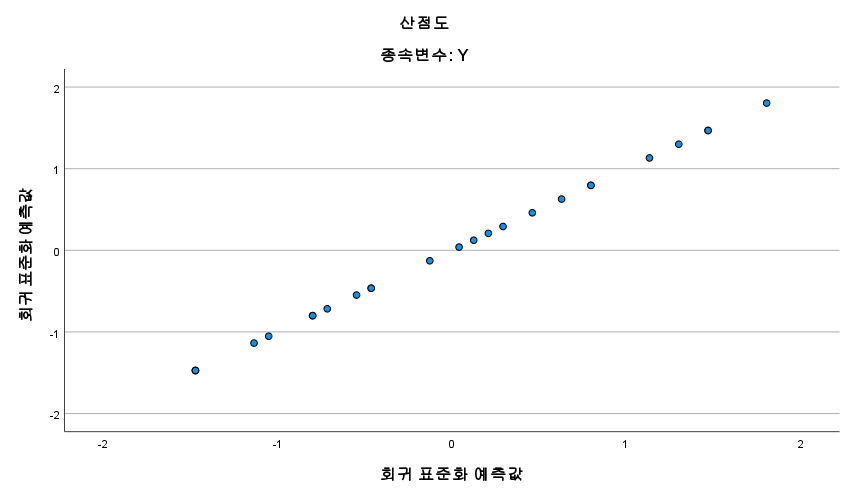
|  |
| --- |
| a. 종속변수: Y |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **잔차 통계량a** | | | | | |
|  | 최소값 | 최대값 | 평균 | 표준화 편차 | N |
| 예측값 | 226.3464 | 296.0959 | 257.6800 | 21.28757 | 25 |
| 잔차 | -37.05761 | 18.98070 | .00000 | 13.50245 | 25 |
| 표준화 예측값 | -1.472 | 1.805 | .000 | 1.000 | 25 |
| 표준화 잔차 | -2.687 | 1.376 | .000 | .979 | 25 |

|  |
| --- |
| a. 종속변수: Y |

차트





산점도를 보면 X와 Y 변수의 관계는 선형성을 가지고 있는 것으로 보여 진다.

상관계수 R=0.884이고 유의확률 P-값은 0.000이고 표본의 수는 25개이다.

요약된 모형의 결정계수 이 0.713로 모형은 설명력이 좋은 편이다.

회귀계수 이고 X값이 1 증가할 때 Y는 1.788만큼 증가한다.

상관분석 결과 유의확률 P-값이 0.000이으로 유의 수준 0.01보다 작기 때문에 귀무가설을 기각하고 대립가설을 지지한다고 볼 수 있고, X는 Y와 선형적으로 관계가 있음을 알 수 있고 상관계수 R=0.844로 84.4% 만큼의 높은 상관관계가 있음을 알 수 있다.

**(8장)8.** x, y변수 간에 다음의 데이터를 얻었다. 두 변수 간의 산점도를 그리고, 적합한 회귀모형을 추정하라.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Y | 20 | 14 | 11 | 10 | 9 | 8 | 8 | 7 | 7 | 6 | 6 | 6 |

GET

FILE='C:\Users\robinhwp\Documents\8-8.sav'.

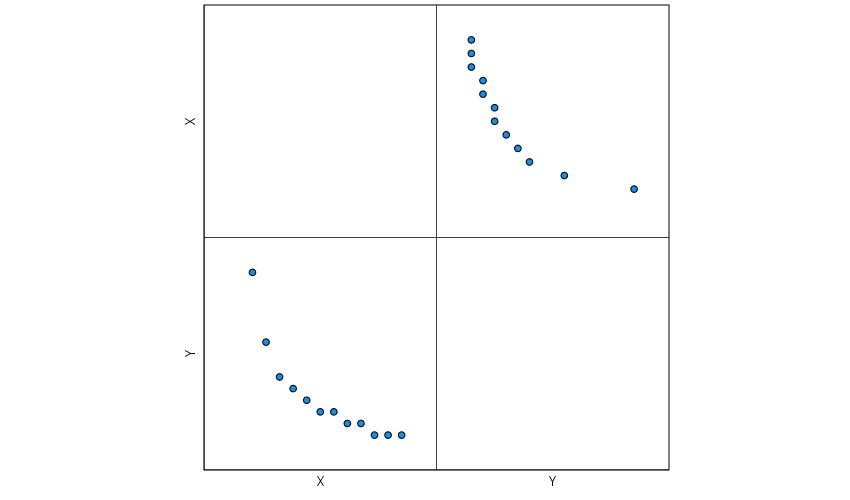
DATASET NAME 데이터세트1 WINDOW=FRONT.

GRAPH

/SCATTERPLOT(MATRIX)=X Y

/MISSING=LISTWISE.

그래프



REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT Y

/METHOD=ENTER X

/SCATTERPLOT=(\*ZPRED ,\*ZPRED)

/RESIDUALS NORMPROB(ZRESID)

/CASEWISE PLOT(ZRESID) ALL.

회귀

|  |  |  |  |
| --- | --- | --- | --- |
| **입력/제거된 변수a** | | | |
| 모형 | 입력된 변수 | 제거된 변수 | 방법 |
| 1 | Xb | . | 입력 |

|  |
| --- |
| a. 종속변수: Y |
| b. 요청된 모든 변수가 입력되었습니다. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **모형 요약b** | | | | |
| 모형 | R | R 제곱 | 수정된 R 제곱 | 추정값의 표준오차 |
| 1 | .863a | .745 | .719 | 2.183 |

|  |
| --- |
| a. 예측자: (상수), X |
| b. 종속변수: Y |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| 모형 | | 제곱합 | 자유도 | 평균제곱 | F | 유의확률 |
| 1 | 회귀 | 139.028 | 1 | 139.028 | 29.184 | .000b |
| 잔차 | 47.639 | 10 | 4.764 |  |  |
| 전체 | 186.667 | 11 |  |  |  |

|  |
| --- |
| a. 종속변수: Y |
| b. 예측자: (상수), X |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **계수a** | | | | | | |
| 모형 | | 비표준화 계수 | | 표준화 계수 | t | 유의확률 |
| B | 표준화 오류 | 베타 |
| 1 | (상수) | 15.742 | 1.343 |  | 11.719 | .000 |
| X | -.986 | .183 | -.863 | -5.402 | .000 |

|  |
| --- |
| a. 종속변수: Y |

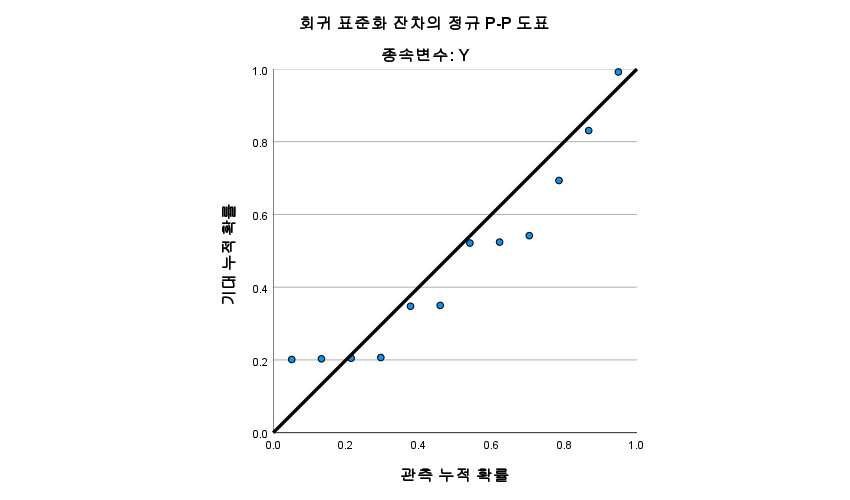
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **케이스별 진단a** | | | | |
| 케이스 번호 | 표준화 잔차 | Y | 예측값 | 잔차 |
| 1 | 2.402 | 20 | 14.76 | 5.244 |
| 2 | .105 | 14 | 13.77 | .230 |
| 3 | -.818 | 11 | 12.78 | -1.784 |
| 4 | -.824 | 10 | 11.80 | -1.798 |
| 5 | -.830 | 9 | 10.81 | -1.812 |
| 6 | -.837 | 8 | 9.83 | -1.826 |
| 7 | -.385 | 8 | 8.84 | -.840 |
| 8 | -.391 | 7 | 7.85 | -.854 |
| 9 | .060 | 7 | 6.87 | .132 |
| 10 | .054 | 6 | 5.88 | .118 |
| 11 | .506 | 6 | 4.90 | 1.104 |
| 12 | .957 | 6 | 3.91 | 2.090 |

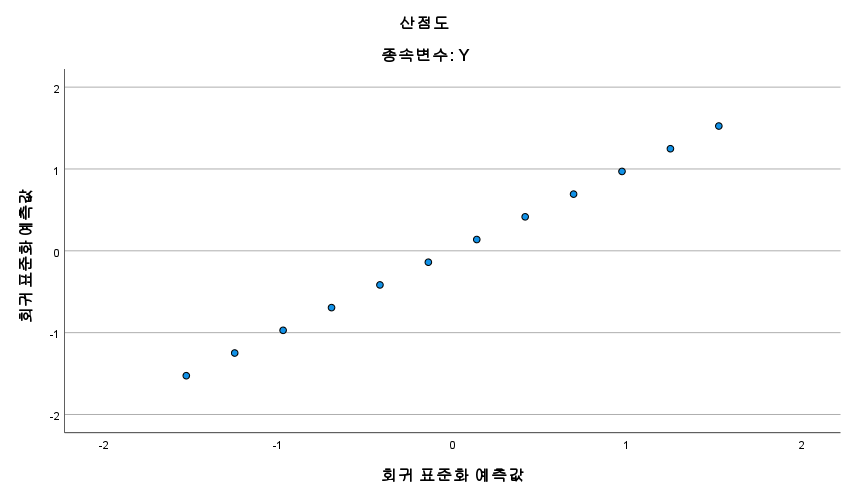
|  |
| --- |
| a. 종속변수: Y |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **잔차 통계량a** | | | | | |
|  | 최소값 | 최대값 | 평균 | 표준화 편차 | N |
| 예측값 | 3.91 | 14.76 | 9.33 | 3.555 | 12 |
| 잔차 | -1.826 | 5.244 | .000 | 2.081 | 12 |
| 표준화 예측값 | -1.525 | 1.525 | .000 | 1.000 | 12 |
| 표준화 잔차 | -.837 | 2.402 | .000 | .953 | 12 |

|  |
| --- |
| a. 종속변수: Y |

차트





산점도를 보면 X와 Y 변수의 관계는 X가 증가함에 따라 Y는 감소하는 선형성을 가지고 있는 것으로 보여 진다.

상관계수 R=0.863이고 유의확률 P-값은 0.000이고 표본의 수는 12개이다.

요약된 모형의 결정계수 이 0.754로 모형은 설명력이 좋은 편이다.

회귀계수 이고 X값이 1 증가할 때 Y는 0.986만큼 감소한다.

상관분석 결과 유의확률 P-값이 0.000이으로 유의 수준 0.01보다 작기 때문에 귀무가설을 기각하고 대립가설을 지지한다고 볼 수 있고, X는 Y와 선형적으로 관계가 있음을 알 수 있고 상관계수 R=0.863으로 86.3% 만큼의 높은 상관관계가 있음을 알 수 있다.

2) 교재 9장 연습문제(p.339) 1번, 6번

**(9장)1.** 유아들을 대상으로 세 가지 읽는 방법을 비교 • 실험하여 다음과 같은 독서평가 점수자료를 얻었다. 세 방법에 통계적으로 유의한 차이가 있는지 검정하라.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 독서방법 | 점수 | | | | |
| B |  | 6 | 9 | 2 | 16 |
| D | 7 | 7 | 12 | 10 | 16 |
| S | 11 | 7 | 4 | 7 | 7 |

ONEWAY 점수 BY 독서방법

/STATISTICS DESCRIPTIVES HOMOGENEITY

/MISSING ANALYSIS

/CRITERIA=CILEVEL(0.95)

/POSTHOC=SCHEFFE LSD ALPHA(0.05).

일원배치 분산분석

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **기술통계** | | | | | | | | |
| 점수 | | | | | | | | |
|  | N | 평균 | 표준편차 | 표준오차 | 평균의 95% 신뢰구간 | | 최소값 | 최대값 |
| 하한 | 상한 |
| B | 4 | 8.25 | 5.909 | 2.955 | -1.15 | 17.65 | 2 | 16 |
| D | 5 | 10.40 | 3.782 | 1.691 | 5.70 | 15.10 | 7 | 16 |
| S | 5 | 7.20 | 2.490 | 1.114 | 4.11 | 10.29 | 4 | 11 |
| 전체 | 14 | 8.64 | 4.050 | 1.082 | 6.30 | 10.98 | 2 | 16 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **분산의 동질성 검정** | | | | | |
|  | | Levene 통계량 | df1 | df2 | CTT 유의확률 |
| 점수 | 평균을 기준으로 합니다. | 1.488 | 2 | 11 | .268 |
| 중위수를 기준으로 합니다. | 1.449 | 2 | 11 | .276 |
| 자유도를 수정한 상태에서 중위수를 기준으로 합니다. | 1.449 | 2 | 8.615 | .287 |
| 절삭평균을 기준으로 합니다. | 1.482 | 2 | 11 | .269 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ANOVA** | | | | | |
| 점수 | | | | | |
|  | 제곱합 | 자유도 | 평균제곱 | F | CTT 유의확률 |
| 집단-간 | 26.464 | 2 | 13.232 | .779 | .482 |
| 집단-내 | 186.750 | 11 | 16.977 |  |  |
| 전체 | 213.214 | 13 |  |  |  |

사후검정

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **다중비교** | | | | | | |
| 종속변수: 점수 | | | | | | |
|  | (I) 독서방법 | (J) 독서방법 | 평균차이(I-J) | 표준오차 | CTT 유의확률 | 95% 신뢰구간 |
| 하한 |
| Scheffe | B | D | -2.150 | 2.764 | .745 | -9.95 |
| S | 1.050 | 2.764 | .931 | -6.75 |
| D | B | 2.150 | 2.764 | .745 | -5.65 |
| S | 3.200 | 2.606 | .493 | -4.15 |
| S | B | -1.050 | 2.764 | .931 | -8.85 |
| D | -3.200 | 2.606 | .493 | -10.55 |
| LSD | B | D | -2.150 | 2.764 | .453 | -8.23 |
| S | 1.050 | 2.764 | .711 | -5.03 |
| D | B | 2.150 | 2.764 | .453 | -3.93 |
| S | 3.200 | 2.606 | .245 | -2.54 |
| S | B | -1.050 | 2.764 | .711 | -7.13 |
| D | -3.200 | 2.606 | .245 | -8.94 |

|  |  |  |  |
| --- | --- | --- | --- |
| **다중비교** | | | |
| 종속변수: 점수 | | | |
|  | (I) 독서방법 | (J) 독서방법 | 95% 신뢰구간 |
| 상한 |
| Scheffe | B | D | 5.65 |
| S | 8.85 |
| D | B | 9.95 |
| S | 10.55 |
| S | B | 6.75 |
| D | 4.15 |
| LSD | B | D | 3.93 |
| S | 7.13 |
| D | B | 8.23 |
| S | 8.94 |
| S | B | 5.03 |
| D | 2.54 |

동질적 부분집합

|  |  |  |  |
| --- | --- | --- | --- |
| **점수** | | | |
|  | 독서방법 | N | 유의수준 = 0.05에 대한 부분집합 |
| 1 |
| Scheffea,b | S | 5 | 7.20 |
| B | 4 | 8.25 |
| D | 5 | 10.40 |
| CTT 유의확률 |  | .519 |

|  |
| --- |
| 동질적 부분집합에 있는 집단에 대한 평균이 표시됩니다. |
| a. 조화평균 표본크기 4.615을(를) 사용합니다. |
| b. 집단 크기가 동일하지 않습니다. 집단 크기의 조화평균이 사용됩니다. I 유형 오차 수준은 보장되지 않습니다. |

분산의 동질성 검정의 유의확률 P-값=0.268로 값이 크기 때문에 분산은 동일하다고 볼 수 있다. 독서방법은 3개이므로 집단간의 자유도는 3-1=2이고 전체 자료 수는 14개이다.

F-값(평균제곱의 비)에 대한 P-값은 0.482이므로 귀무가설을 받아들여 독서방법에 따른 차이가 있다고 볼 수 없다

**(9장)6.** 어떤 기계의 소음을 작게 하려고 모터(motor)의 베어링 부분에 대하여 조립 후의 볼 베어링의 유격(play)을 3수준, 조립 후의 진동을 3수준으로 바꾸어 3회 반복하여 3x3x3=27(회)의 실험을 랜덤하게 행하고, 소음계로 소음을 측정한 결과가 아래와 같다.

[실험조건]

|  |  |  |  |
| --- | --- | --- | --- |
| 볼 베어링의 유격 | | 조립후의 진동: | |
|  |  | |  | |  |
|  | 78  76  75 | | 74  76  75 | | 75  76  77 |
|  | 79  80  80 | | 78  79  80 | | 78  77  79 |
|  | 81  83  83 | | 79  77  75 | | 79  80  81 |

RECODE 유격 진동 (0=1) (5=2) (10=3) (40=1) (110=2) (180=3).

EXECUTE.

(1) 이원배치 분산분석을 실시하고 분석하라.

UNIANOVA 소음 BY 유격 진동

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/PLOT=PROFILE(유격\*진동) TYPE=LINE ERRORBAR=NO MEANREFERENCE=NO YAXIS=AUTO

/CRITERIA=ALPHA(0.05)

/DESIGN=유격 진동 유격\*진동.

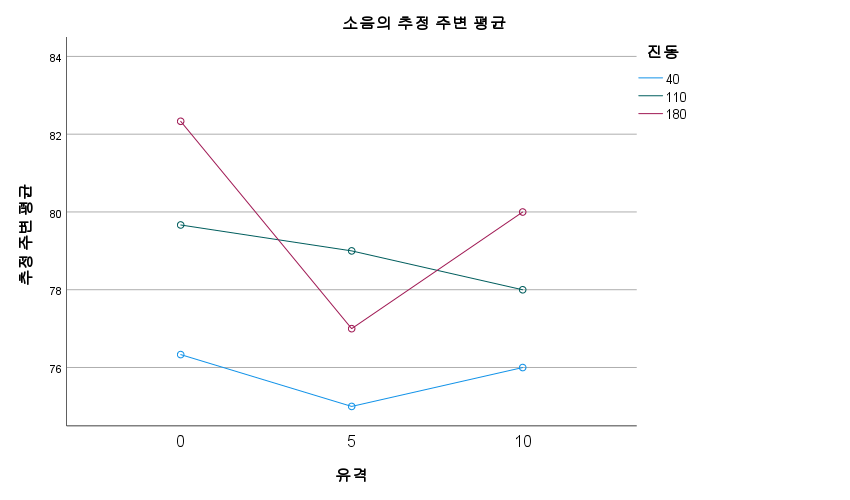
일변량 분산분석

|  |  |  |  |
| --- | --- | --- | --- |
| **개체-간 요인** | | | |
|  | | 값 레이블 | N |
| 유격 | 1 | 0 | 9 |
| 2 | 5 | 9 |
| 3 | 10 | 9 |
| 진동 | 1 | 40 | 9 |
| 2 | 110 | 9 |
| 3 | 180 | 9 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **개체-간 효과 검정** | | | | | |
| 종속변수: 소음 | | | | | |
| 원인 | 제 III 유형 제곱합 | 자유도 | 평균제곱 | 거짓 | 유의확률 |
| 수정된 모형 | 129.407a | 8 | 16.176 | 11.199 | .000 |
| 절편 | 164892.593 | 1 | 164892.593 | 114156.410 | .000 |
| 유격 | 27.185 | 2 | 13.593 | 9.410 | .002 |
| 진동 | 79.407 | 2 | 39.704 | 27.487 | .000 |
| 유격 \* 진동 | 22.815 | 4 | 5.704 | 3.949 | .018 |
| 추정값 | 26.000 | 18 | 1.444 |  |  |
| 전체 | 165048.000 | 27 |  |  |  |
| 수정된 합계 | 155.407 | 26 |  |  |  |

|  |
| --- |
| a. R 제곱 = .833 (수정된 R 제곱 = .758) |

프로파일 도표



개체-간 효과 검정을 보면 유격의 유의확률 P-값=0.002로 유격에 따른 소음에 차이가 있고, 진동의 유의확률 P-값=0.000으로 진동에 따른 소음에도 차이가 있다.

유격\*진동의 P-값=0.018로 교유작용 효과가 있다고 볼 수 있다.

(2) 위의 데이터 중에서 A1B1에서는 78이 없어지고, A3B3에서 80이 없어진 경우에 이원분산분석을 실시하고 분석하라.

UNIANOVA 소음 BY 유격 진동

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/PLOT=PROFILE(유격\*진동) TYPE=LINE ERRORBAR=NO MEANREFERENCE=NO YAXIS=AUTO

/CRITERIA=ALPHA(0.05)

/DESIGN=유격 진동 유격\*진동.

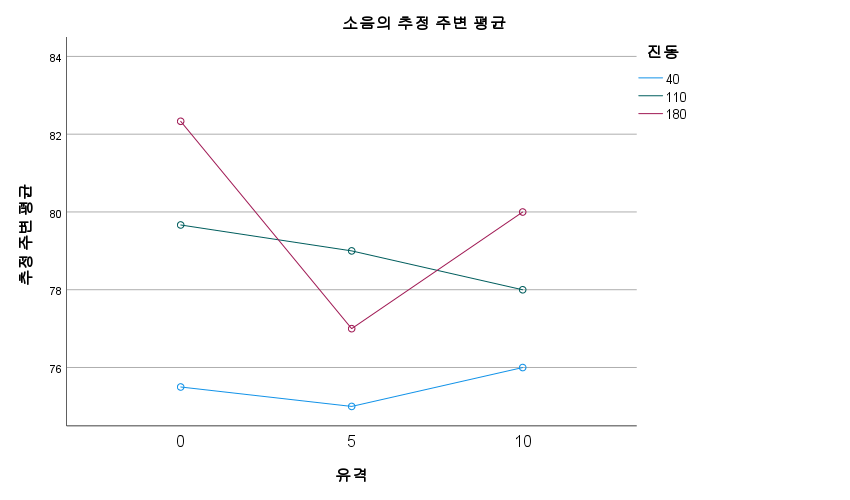
일변량 분산분석

|  |  |  |  |
| --- | --- | --- | --- |
| **개체-간 요인** | | | |
|  | | 값 레이블 | N |
| 유격 | 1 | 0 | 8 |
| 2 | 5 | 9 |
| 3 | 10 | 8 |
| 진동 | 1 | 40 | 8 |
| 2 | 110 | 9 |
| 3 | 180 | 8 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **개체-간 효과 검정** | | | | | |
| 종속변수: 소음 | | | | | |
| 원인 | 제 III 유형 제곱합 | 자유도 | 평균제곱 | 거짓 | 유의확률 |
| 수정된 모형 | 130.007a | 8 | 16.251 | 11.909 | .000 |
| 절편 | 148051.875 | 1 | 148051.875 | 108496.031 | .000 |
| 유격 | 19.500 | 2 | 9.750 | 7.145 | .006 |
| 진동 | 79.465 | 2 | 39.732 | 29.117 | .000 |
| 유격 \* 진동 | 25.797 | 4 | 6.449 | 4.726 | .010 |
| 추정값 | 21.833 | 16 | 1.365 |  |  |
| 전체 | 152564.000 | 25 |  |  |  |
| 수정된 합계 | 151.840 | 24 |  |  |  |

|  |
| --- |
| a. R 제곱 = .856 (수정된 R 제곱 = .784) |

프로파일 도표



A1B1에서는 78이 없어지고, A3B3에서 80이 없어진 경우에 개체-간 효과 검정을 보면 유격의 유의확률 P-값=0.006로 0.004 증가하였고,

유격\*진동의 P-값=0.010으로 0.008 감소한 것으로 볼 수 있다.

A1B1에서는 78이 없어지고, A3B3에서 80이 없어진 경우에도 유격과 진동에 따른 소음은 차이가 있었고 유격\*진동의 교우작업 효과가 있다고 볼 수 있다.

<참고자료>

<https://data-flair.training/blogs/sas-correlation-analysis/>

<https://documentation.sas.com/doc/en/vdmmlcdc/1.0/grstatproc/p0lfzklhx36ylln1t9sssgzuf64m.htm>

-- 통계패키지 기말과제물 끝 --