SALARY\_EXAM.MWX

**회귀 분석: 연봉 대 근무년수, 근무평점**

**회귀 방정식**

|  |  |  |
| --- | --- | --- |
| 연봉 | = | 347.6 + 10.08 근무년수 + 28.75 근무평점 |

**계수**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **항** | **계수** | **SE 계수** | **T-값** | **P-값** | **VIF** |
| 상수 | 347.6 | 26.5 | 13.10 | 0.000 |  |
| 근무년수 | 10.08 | 2.12 | 4.74 | 0.000 | 1.58 |
| 근무평점 | 28.75 | 5.69 | 5.06 | 0.000 | 1.58 |

**모형 요약**

|  |  |  |  |
| --- | --- | --- | --- |
| **S** | **R-제곱** | **R-제곱(수정)** | **R-제곱(예측)** |
| 86.0395 | 74.00% | 72.79% | 69.87% |

**분산 분석**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **출처** | **DF** | **Adj SS** | **Adj MS** | **F-값** | **P-값** |
| 회귀 | 2 | 906109 | 453054 | 61.20 | 0.000 |
| 근무년수 | 1 | 166674 | 166674 | 22.51 | 0.000 |
| 근무평점 | 1 | 189258 | 189258 | 25.57 | 0.000 |
| 오차 | 43 | 318320 | 7403 |  |  |
| 적합성 결여 | 37 | 308122 | 8328 | 4.90 | 0.027 |
| 순수 오차 | 6 | 10198 | 1700 |  |  |
| 총계 | 45 | 1224428 |  |  |  |

**비정상적 관측치에 대한 적합치 및 진단**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **관측** | **연봉** | **적합치** | **잔차** | **표준화 잔차** |  |
| 1 | 585.0 | 787.8 | -202.8 | -2.44 | R |
| 2 | 494.0 | 757.6 | -263.6 | -3.17 | R |
| 40 | 1039.0 | 856.9 | 182.1 | 2.23 | R |

R  큰 잔차

SALARY\_EXAM.MWX

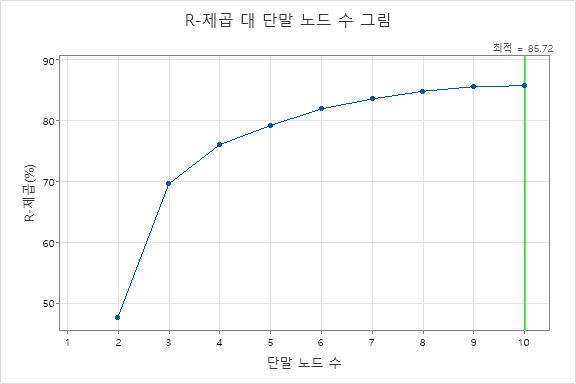
**10 노드 CART® 회귀: 연봉 대 근무년수, 근무평점**

**방법**

|  |  |
| --- | --- |
| 노드 분할 | 최소 제곱 오차 |
| 최적 트리 | 최대 R-제곱 |
| 사용된 행 | 46 |

**반응 정보**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **평균** | **표준 편차** | **최소값** | **Q1** | **중위수** | **Q3** | **최대값** |
| 597.348 | 164.953 | 355 | 493.5 | 548.5 | 693 | 1039 |



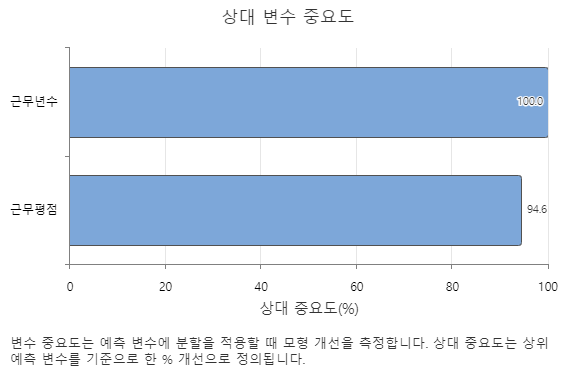
**최적 트리 다이어그램**

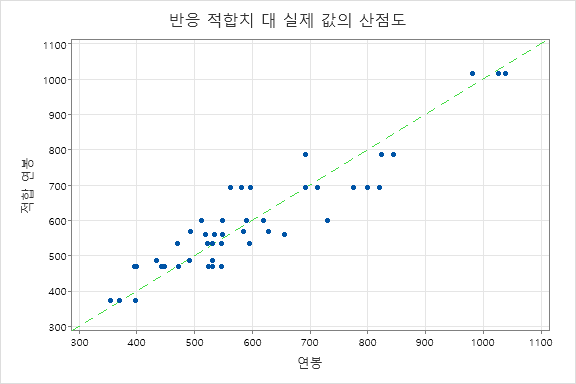
텍스트, 도표, 평면도, 개략도이(가) 표시된 사진

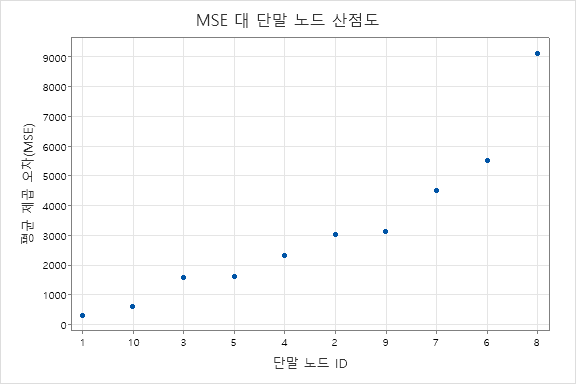
자동 생성된 설명

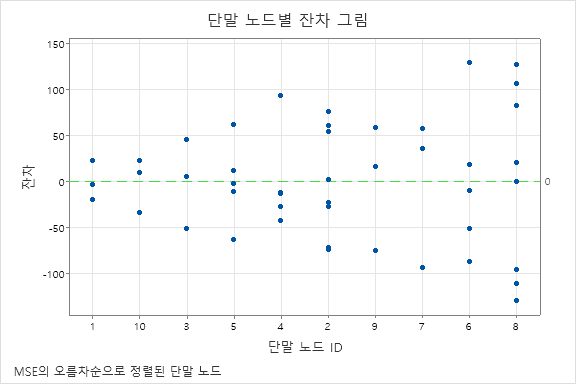
**모형 요약**

|  |  |
| --- | --- |
| 전체 예측 변수 | 2 |
| 중요 예측 변수 | 2 |
| 단말 노드 수 | 10 |
| 최소 단말 노드 크기 | 3 |
| R-제곱 | 85.72% |
| 루트 평균 제곱 오차(RMSE) | 61.6631 |
| 평균 제곱 오차(MSE) | 3802.3370 |
| 평균 절대 편차(MAD) | 48.9558 |
| 평균 절대 백분율 오차(MAPE) | 0.0840 |









SALARY\_EXAM.MWX

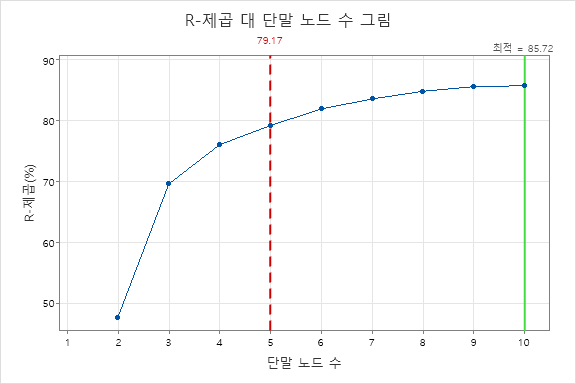
**5 노드 대립 CART® 회귀: 연봉 대 근무년수, 근무평점**

**방법**

|  |  |
| --- | --- |
| 노드 분할 | 최소 제곱 오차 |
| 최적 트리 | 최대 R-제곱 |
| 사용된 행 | 46 |

**반응 정보**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **평균** | **표준 편차** | **최소값** | **Q1** | **중위수** | **Q3** | **최대값** |
| 597.348 | 164.953 | 355 | 493.5 | 548.5 | 693 | 1039 |



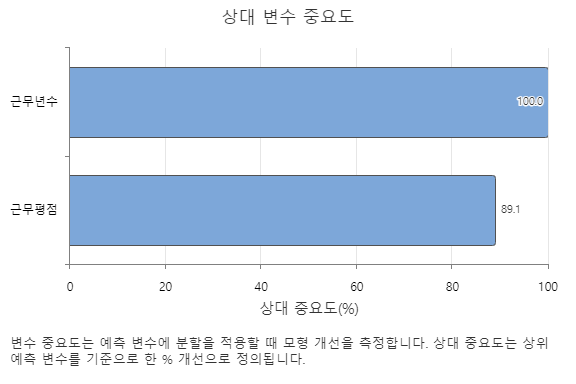
**대립 트리 다이어그램**

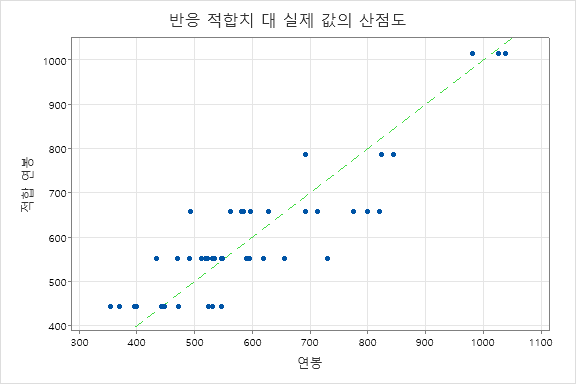
텍스트, 도표, 라인, 평면도이(가) 표시된 사진

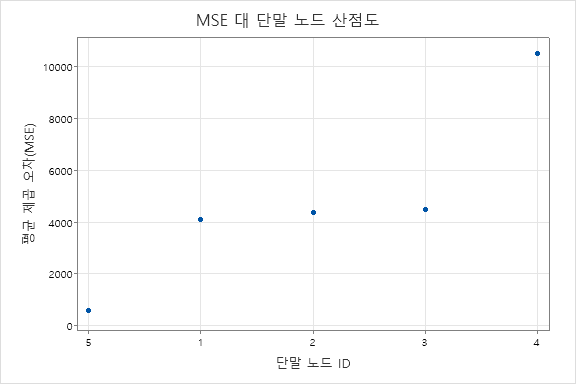
자동 생성된 설명

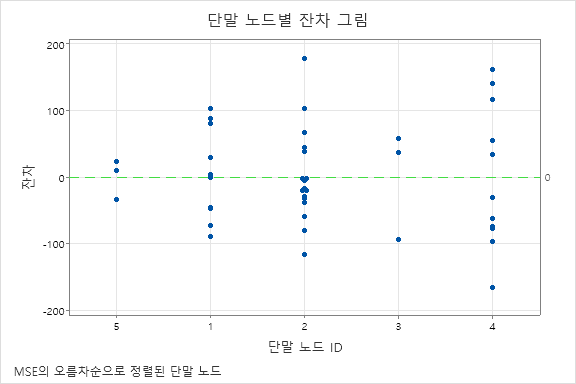
**모형 요약**

|  |  |
| --- | --- |
| 전체 예측 변수 | 2 |
| 중요 예측 변수 | 2 |
| 단말 노드 수 | 5 |
| 최소 단말 노드 크기 | 3 |
| R-제곱 | 79.17% |
| 루트 평균 제곱 오차(RMSE) | 74.4557 |
| 평균 제곱 오차(MSE) | 5543.6459 |
| 평균 절대 편차(MAD) | 59.4209 |
| 평균 절대 백분율 오차(MAPE) | 0.1042 |









**연봉이 높은 집단 & 연봉이 낮은 집단**

(노드 10 분할 기준)연봉이 최고 낮은 집단은 근무년수가 10.5년 이하 이며, 근무평점이 0.5 이하의 집단 이며 전체 평균은 374.667 이었으며, 연봉이 최고 높은 집단은 근무연수가 10.5 년 이상 근무 평점이 9.5 이상인 집단이며 전체평균은 1015.67 로 가장 낮은 집단에 비해 약 2.7배 이상이상높은 것으로 관측되었다.