LENGTH

# Maybe remove outliers ?

Removed six outliers which had a Cook’s distance score of above > .1. These verdicts were extremely long.

ECLI:NL:RVS:2020:2706

ECLI:NL:RVS:2020:2053

ECLI:NL:RVS:2020:2439

ECLI:NL:CBB:2020:1

ECLI:NL:RVS:2020:1769

ECLI:NL:RVS:2020:3147

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .640a | .409 | .409 | 8800.632 | 1.181 |
| a. Predictors: (Constant), law\_references, experience\_voorzitter\_days, experience\_griffier\_days, unfounded, district\_court, ecli\_references, single\_judge, first\_instance | | | | | |
| b. Dependent Variable: verdict\_length | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 265937585042.201 | 8 | 33242198130.275 | 429.202 | .000b |
| Residual | 383537962542.374 | 4952 | 77451123.292 |  |  |
| Total | 649475547584.575 | 4960 |  |  |  |
| a. Dependent Variable: verdict\_length | | | | | | |
| b. Predictors: (Constant), law\_references, experience\_voorzitter\_days, experience\_griffier\_days, unfounded, district\_court, ecli\_references, single\_judge, first\_instance | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| B | Std. Error | Beta | Tolerance | VIF |
| 1 | (Constant) | 13570.519 | 562.122 |  | 24.142 | .000 |  |  |
| experience\_voorzitter\_days | -.088 | .037 | -.027 | -2.381 | .017 | .963 | 1.039 |
| single\_judge | -7489.289 | 329.293 | -.268 | -22.744 | .000 | .862 | 1.161 |
| experience\_griffier\_days | .649 | .051 | .143 | 12.794 | .000 | .961 | 1.041 |
| district\_court | -866.967 | 323.030 | -.038 | -2.684 | .007 | .602 | 1.661 |
| first\_instance | 958.989 | 377.281 | .035 | 2.542 | .011 | .636 | 1.573 |
| unfounded | -1599.066 | 277.223 | -.064 | -5.768 | .000 | .974 | 1.027 |
| ecli\_references | 1666.818 | 48.465 | .400 | 34.393 | .000 | .884 | 1.132 |
| law\_references | 695.976 | 37.421 | .212 | 18.599 | .000 | .915 | 1.093 |
| a. Dependent Variable: verdict\_length | | | | | | | | |

# NORMALITY

df

# Values of residuals are independent

Values of residuals are independent (positively correlated) (Durbin-Watson 1.12). See <https://www.open.ac.uk/socialsciences/spsstutorial/files/tutorials/assumptions.pdf> <https://wiki.uva.nl/methodologiewinkel/index.php/Onafhankelijkheid>

# COLLINEARITY

(all VIF under 10)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| B | Std. Error | Beta | Tolerance | VIF |
| 1 | (Constant) | 13570.519 | 562.122 |  | 24.142 | .000 |  |  |
| experience\_voorzitter\_days | -.088 | .037 | -.027 | -2.381 | .017 | .963 | 1.039 |
| single\_judge | -7489.289 | 329.293 | -.268 | -22.744 | .000 | .862 | 1.161 |
| experience\_griffier\_days | .649 | .051 | .143 | 12.794 | .000 | .961 | 1.041 |
| district\_court | -866.967 | 323.030 | -.038 | -2.684 | .007 | .602 | 1.661 |
| first\_instance | 958.989 | 377.281 | .035 | 2.542 | .011 | .636 | 1.573 |
| unfounded | -1599.066 | 277.223 | -.064 | -5.768 | .000 | .974 | 1.027 |
| ecli\_references | 1666.818 | 48.465 | .400 | 34.393 | .000 | .884 | 1.132 |
| law\_references | 695.976 | 37.421 | .212 | 18.599 | .000 | .915 | 1.093 |
| a. Dependent Variable: verdict\_length | | | | | | | | |

# The variance of the residuals is constant.

Checked with scatterplot; looked good

# The values of the residuals are normally distributed.

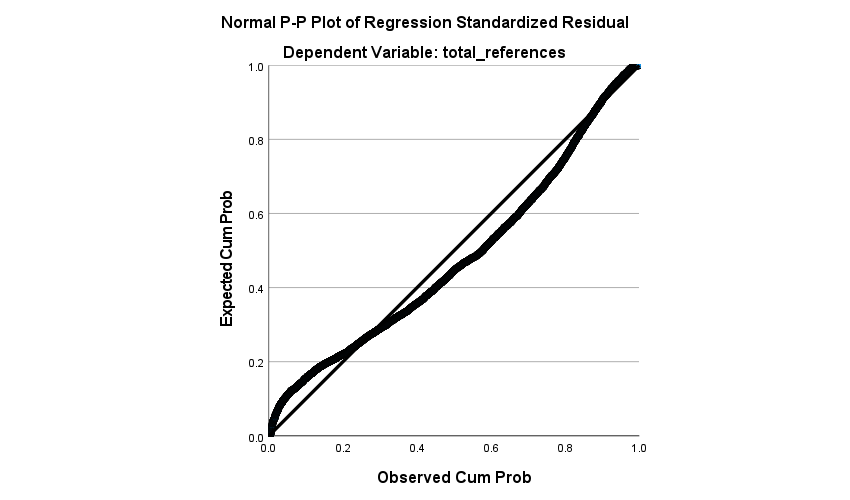
PP plot was not really nice

TOTAL REFERENCES

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .538a | .290 | .289 | 4.20815 | 1.900 |
| a. Predictors: (Constant), unfounded, experience\_voorzitter\_days, experience\_griffier\_days, district\_court, verdict\_length, single\_judge, first\_instance | | | | | |
| b. Dependent Variable: total\_references | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 35809.127 | 7 | 5115.590 | 288.878 | .000b |
| Residual | 87710.225 | 4953 | 17.709 |  |  |
| Total | 123519.352 | 4960 |  |  |  |
| a. Dependent Variable: total\_references | | | | | | |
| b. Predictors: (Constant), unfounded, experience\_voorzitter\_days, experience\_griffier\_days, district\_court, verdict\_length, single\_judge, first\_instance | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| B | Std. Error | Beta | Tolerance | VIF |
| 1 | (Constant) | 1.768 | .282 |  | 6.260 | .000 |  |  |
| experience\_voorzitter\_days | 1.584E-5 | .000 | .011 | .894 | .372 | .962 | 1.039 |
| single\_judge | .829 | .165 | .068 | 5.021 | .000 | .784 | 1.275 |
| experience\_griffier\_days | .000 | .000 | -.097 | -7.885 | .000 | .943 | 1.060 |
| verdict\_length | .000 | .000 | .546 | 41.479 | .000 | .828 | 1.208 |
| district\_court | -.703 | .154 | -.070 | -4.573 | .000 | .608 | 1.645 |
| first\_instance | 1.005 | .180 | .084 | 5.597 | .000 | .642 | 1.558 |
| unfounded | -.332 | .133 | -.030 | -2.502 | .012 | .974 | 1.027 |
| a. Dependent Variable: total\_references | | | | | | | | |



SIMILARITY

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .419a | .176 | .174 | .075125500890717 | 1.970 |
| a. Predictors: (Constant), law\_references, experience\_voorzitter\_days, experience\_griffier\_days, unfounded, district\_court, ecli\_references, single\_judge, first\_instance, verdict\_length | | | | | |
| b. Dependent Variable: similarity\_score | | | | | |

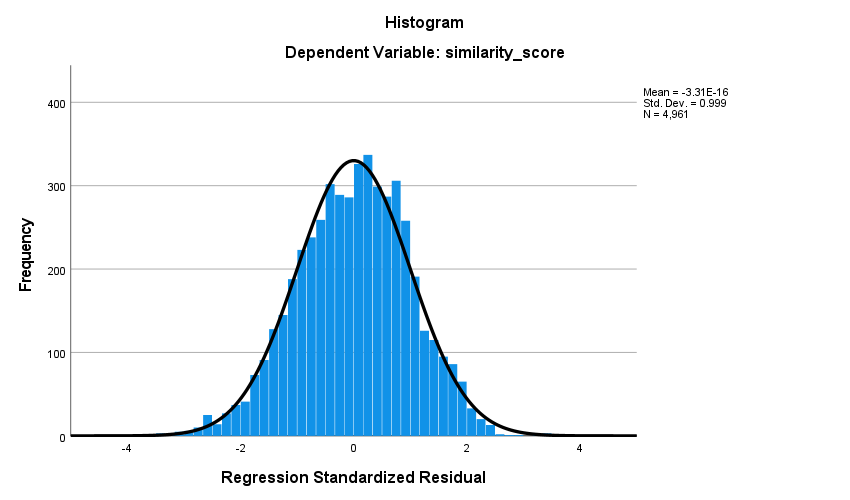
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ANOVAa** | | | | | | |
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 5.964 | 9 | .663 | 117.416 | .000b |
| Residual | 27.943 | 4951 | .006 |  |  |
| Total | 33.907 | 4960 |  |  |  |
| a. Dependent Variable: similarity\_score | | | | | | |
| b. Predictors: (Constant), law\_references, experience\_voorzitter\_days, experience\_griffier\_days, unfounded, district\_court, ecli\_references, single\_judge, first\_instance, verdict\_length | | | | | | |

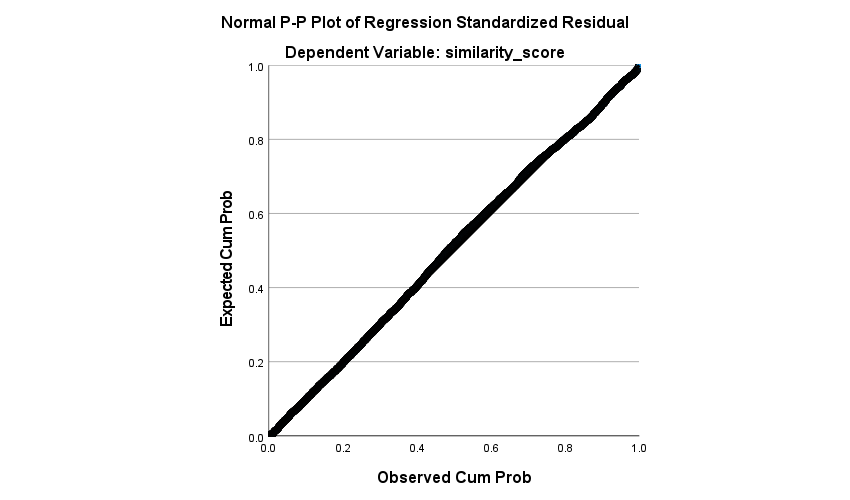
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Collinearity Statistics | |
| B | Std. Error | Beta | Tolerance | VIF |
| 1 | (Constant) | .730 | .005 |  | 143.931 | .000 |  |  |
| experience\_voorzitter\_days | 6.417E-7 | .000 | .027 | 2.027 | .043 | .961 | 1.040 |
| single\_judge | .030 | .003 | .150 | 10.241 | .000 | .780 | 1.282 |
| experience\_griffier\_days | -5.650E-6 | .000 | -.172 | -12.839 | .000 | .930 | 1.075 |
| verdict\_length | -1.251E-6 | .000 | -.173 | -10.309 | .000 | .591 | 1.693 |
| district\_court | -.049 | .003 | -.295 | -17.729 | .000 | .601 | 1.663 |
| first\_instance | .058 | .003 | .289 | 17.857 | .000 | .635 | 1.575 |
| unfounded | .007 | .002 | .039 | 2.983 | .003 | .967 | 1.034 |
| ecli\_references | .003 | .000 | .099 | 6.506 | .000 | .713 | 1.402 |
| law\_references | .001 | .000 | .035 | 2.474 | .013 | .856 | 1.169 |
| a. Dependent Variable: similarity\_score | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Collinearity Diagnosticsa** | | | | | | | | | | | | | |
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | | | | | | | | |
| (Constant) | experience\_voorzitter\_days | single\_judge | experience\_griffier\_days | verdict\_length | district\_court | first\_instance | unfounded | ecli\_references | law\_references |
| 1 | 1 | 6.993 | 1.000 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .00 | .01 |
| 2 | .915 | 2.764 | .00 | .00 | .02 | .00 | .05 | .05 | .01 | .01 | .21 | .05 |
| 3 | .526 | 3.646 | .00 | .04 | .00 | .17 | .00 | .16 | .03 | .05 | .07 | .04 |
| 4 | .412 | 4.118 | .00 | .00 | .00 | .01 | .00 | .00 | .00 | .05 | .34 | .65 |
| 5 | .366 | 4.372 | .00 | .02 | .02 | .37 | .05 | .08 | .00 | .19 | .00 | .08 |
| 6 | .251 | 5.276 | .00 | .47 | .00 | .11 | .01 | .01 | .01 | .30 | .01 | .08 |
| 7 | .221 | 5.627 | .00 | .00 | .16 | .15 | .39 | .03 | .01 | .19 | .27 | .02 |
| 8 | .162 | 6.566 | .01 | .16 | .43 | .06 | .29 | .18 | .00 | .08 | .10 | .07 |
| 9 | .118 | 7.713 | .00 | .06 | .07 | .03 | .08 | .49 | .78 | .00 | .00 | .00 |
| 10 | .036 | 13.927 | .99 | .25 | .30 | .10 | .12 | .01 | .15 | .13 | .00 | .01 |
| a. Dependent Variable: similarity\_score | | | | | | | | | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Residuals Statisticsa** | | | | | |
|  | Minimum | Maximum | Mean | Std. Deviation | N |
| Predicted Value | .604036211967468 | .857774972915649 | .753824306859579 | .034676166947866 | 4961 |
| Std. Predicted Value | -4.320 | 2.998 | .000 | 1.000 | 4961 |
| Standard Error of Predicted Value | .002 | .017 | .003 | .001 | 4961 |
| Adjusted Predicted Value | .592791378498077 | .857920825481415 | .753817865913656 | .034712525833599 | 4961 |
| Residual | -.285129904747009 | .264351963996887 | .000000000000000 | .075057311727554 | 4961 |
| Std. Residual | -3.795 | 3.519 | .000 | .999 | 4961 |
| Stud. Residual | -3.799 | 3.529 | .000 | 1.000 | 4961 |
| Deleted Residual | -.285701662302017 | .265910744667053 | .000006440945921 | .075224862709227 | 4961 |
| Stud. Deleted Residual | -3.804 | 3.533 | .000 | 1.000 | 4961 |
| Mahal. Distance | 1.565 | 260.790 | 8.998 | 9.417 | 4961 |
| Cook's Distance | .000 | .043 | .000 | .001 | 4961 |
| Centered Leverage Value | .000 | .053 | .002 | .002 | 4961 |
| a. Dependent Variable: similarity\_score | | | | | |

**Charts**





EXPLORATION OF SIMILARITY SCORE

Not normal (sig .000 for Shapiro wilk)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Descriptives** | | | | |
|  | | | Statistic | Std. Error |
| similarity\_score | Mean | | .753808004211223 | .001173331554485 |
| 95% Confidence Interval for Mean | Lower Bound | .751507755984634 |  |
| Upper Bound | .756108252437811 |  |
| 5% Trimmed Mean | | .754554671348988 |  |
| Median | | .752601087093353 |  |
| Variance | | .007 |  |
| Std. Deviation | | .082692825292367 |  |
| Minimum | | .489177942276001 |  |
| Maximum | | .962371230125427 |  |
| Range | | .473193287849426 |  |
| Interquartile Range | | .119961440563202 |  |
| Skewness | | -.070 | .035 |
| Kurtosis | | -.493 | .069 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Tests of Normality** | | | | | | |
|  | Kolmogorov-Smirnova | | | Shapiro-Wilk | | |
| Statistic | df | Sig. | Statistic | df | Sig. |
| similarity\_score | .026 | 4967 | .000 | .994 | 4967 | .000 |
| a. Lilliefors Significance Correction | | | | | | |

