

```
569 0 M>  
TITLE "Influence Diagnostics in SPSS".  
570 0 M> TITLE "Influence Diagnostics in SPSS".
```

Influence Diagnostics in SPSS

```
SUBTITLE "G3 regression model".
```

```
571 0 M> SUBTITLE "G3 regression model".
```

Influence Diagnostics in SPSS
G3 regression model

```
572 0 M>
* ----- 2. Import student data -----
573 0 M> * ----- 2. Import student data -----
* This delimiter setting accepts both comma CSV and semicolon CSV.
574 0 M> * This delimiter setting accepts both comma CSV and semicolon CSV.
GET DATA
575 0 M> GET DATA
  /TYPE=TXT
576 0 M>  /TYPE=TXT
  /FILE='D:\low kda score priority basis posts\first post\Influence Diagnostic
s\student-por.csv'
  /ENCODING='UTF8'
577 0 M>  /FILE='D:\low kda score priority basis posts\first post\Influence
Diagnostics\student-por.csv'
  /DELCASE=LINE
578 0 M>  /ENCODING='UTF8'
579 0 M>  /DELCASE=LINE
  /DELIMITERS=',;'
  /QUALIFIER='"'
580 0 M>  /DELIMITERS=',;'
  /ARRANGEMENT=DELIMITED
581 0 M>  /QUALIFIER='"'
582 0 M>  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
583 0 M>  /FIRSTCASE=2
  /IMPORTCASE=ALL
584 0 M>  /IMPORTCASE=ALL
  /VARIABLES=
585 0 M>  /VARIABLES=
  school A20
586 0 M>  school A20
  sex A20
587 0 M>  sex A20
  age F8.0
588 0 M>  age F8.0
  address A20
589 0 M>  address A20
  famsize A20
590 0 M>  famsize A20
```

Influence Diagnostics in SPSS
G3 regression model

Pstatus A20
591 0 M> Pstatus A20
Medu F8.0
592 0 M> Medu F8.0
Fedu F8.0
593 0 M> Fedu F8.0
Mjob A30
594 0 M> Mjob A30
Fjob A30
595 0 M> Fjob A30
reason A30
596 0 M> reason A30
guardian A30
597 0 M> guardian A30
traveltime F8.0
598 0 M> traveltime F8.0
studytime F8.0
599 0 M> studytime F8.0
failures F8.0
600 0 M> failures F8.0
schoolsup A20
601 0 M> schoolsup A20
famsup A20
602 0 M> famsup A20
paid A20
603 0 M> paid A20
activities A20
604 0 M> activities A20
nursery A20
605 0 M> nursery A20
higher A20
606 0 M> higher A20
internet A20
607 0 M> internet A20
romantic A20
608 0 M> romantic A20
famrel F8.0
609 0 M> famrel F8.0
freetime F8.0

Influence Diagnostics in SPSS
G3 regression model

```
610  0 M>  freetime F8.0
        goout F8.0
611  0 M>  goout F8.0
        Dalc F8.0
612  0 M>  Dalc F8.0
        Walc F8.0
613  0 M>  Walc F8.0
        health F8.0
614  0 M>  health F8.0
        absences F8.0
615  0 M>  absences F8.0
        G1 F8.0
616  0 M>  G1 F8.0
        G2 F8.0
617  0 M>  G2 F8.0
        G3 F8.0.
618  0 M>  G3 F8.0.
CACHE.
619  0 M>  CACHE.
EXECUTE.
620  0 M>  EXECUTE.

621  0 M>
DATASET NAME student_imported WINDOW=FRONT.
622  0 M>  DATASET NAME student_imported WINDOW=FRONT.

623  0 M>
* ----- 3. Check import before cleaning -----.
624  0 M>  * ----- 3. Check import before cleaning -----.
TITLE "Import Check Before Cleaning".
625  0 M>  TITLE "Import Check Before Cleaning".
```

Import Check Before Cleaning

```
DATASET ACTIVATE student_imported.
626 0 M> DATASET ACTIVATE student_imported.

627 0 M>
COMPUTE case_id=$CASENUM.
628 0 M> COMPUTE case_id=$CASENUM.
FORMATS case_id (F8.0).
629 0 M> FORMATS case_id (F8.0).
EXECUTE.
630 0 M> EXECUTE.

631 0 M>
* Number of model-missing fields per row.
632 0 M> * Number of model-missing fields per row.
COUNT model_missing_count = G3 G1 G2 studytime failures absences age Medu Fedu
(MISSING).
633 0 M> COUNT model_missing_count = G3 G1 G2 studytime failures absences a
ge Medu Fedu (MISSING).
COMPUTE complete_model_case = (model_missing_count = 0).
634 0 M> COMPUTE complete_model_case = (model_missing_count = 0).
FORMATS model_missing_count complete_model_case (F8.0).
635 0 M> FORMATS model_missing_count complete_model_case (F8.0).
VALUE LABELS complete_model_case 0 "Incomplete model case" 1 "Complete model c
ase".
636 0 M> VALUE LABELS complete_model_case 0 "Incomplete model case" 1 "Comp
lete model case".
EXECUTE.
637 0 M> EXECUTE.

638 0 M>
FREQUENCIES VARIABLES=complete_model_case model_missing_count
639 0 M> FREQUENCIES VARIABLES=complete_model_case model_missing_count
/ORDER=ANALYSIS.
640 0 M> /ORDER=ANALYSIS.
```

Frequencies

[student_imported]

Import Check Before Cleaning

Statistics

		complete_model l_case	model_missing _count
N	Valid	649	649
	Missing	0	0

Frequency Table

complete_model_case

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Complete model case	649	100.0	100.0	100.0

model_missing_count

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	649	100.0	100.0	100.0

```
641 0 M>
DESCRIPTIVES VARIABLES=G3 G1 G2 studytime failures absences age Medu Fedu
642 0 M> DESCRIPTIVES VARIABLES=G3 G1 G2 studytime failures absences age Me
du Fedu
/STATISTICS=MEAN STDDEV MIN MAX.
643 0 M> /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptives

Import Check Before Cleaning

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
G3	649	0	19	11.91	3.231
G1	649	0	19	11.40	2.745
G2	649	0	19	11.57	2.914
studytime	649	1	4	1.93	.830
failures	649	0	3	.22	.593
absences	649	0	32	3.66	4.641
age	649	15	22	16.74	1.218
Medu	649	0	4	2.51	1.135
Fedu	649	0	4	2.31	1.100
Valid N (listwise)	649				

```
644 0 M>
* ----- 4. Keep only complete cases for the influence-diagnostics model -
-----
645 0 M> * ----- 4. Keep only complete cases for the influence-diagnos
tics model -----
USE ALL.
646 0 M> USE ALL.
FILTER OFF.
647 0 M> FILTER OFF.
SELECT IF (complete_model_case = 1).
648 0 M> SELECT IF (complete_model_case = 1).
EXECUTE.
649 0 M> EXECUTE.

650 0 M>
DATASET NAME student_clean WINDOW=FRONT.
651 0 M> DATASET NAME student_clean WINDOW=FRONT.
```

Dataset Name

Warnings

The active dataset will replace the existing dataset named
student_clean.

```
652 0 M>
```

Import Check Before Cleaning

```
TITLE "Cleaned Influence Diagnostics Data".
```

```
653 0 M> TITLE "Cleaned Influence Diagnostics Data".
```

Cleaned Influence Diagnostics Data

SUBTITLE "Complete cases only".

654 0 M> SUBTITLE "Complete cases only".

Cleaned Influence Diagnostics Data
Complete cases only

```
655 0 M>
SAVE TRANSLATE
656 0 M> SAVE TRANSLATE
      /OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\student_por_influence_diagnostics_clean_spss.csv'
      /TYPE=CSV
657 0 M> /OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\student_por_influence_diagnostics
      _clean_spss.csv'
658 0 M> /TYPE=CSV
      /ENCODING='UTF8'
      /REPLACE
659 0 M> /ENCODING='UTF8'
660 0 M> /REPLACE
      /FIELDNAMES
661 0 M> /FIELDNAMES
      /CELLS=VALUES.
662 0 M> /CELLS=VALUES.

663 0 M>
SAVE OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\student_por_influence_diagnostics_clean_spss.sav'
      /COMPRESSED.
664 0 M> SAVE OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\student_por_influence_diagnostics
      ics_clean_spss.sav'
665 0 M> /COMPRESSED.

666 0 M>
* ----- 5. Basic model summaries -----.
667 0 M> * ----- 5. Basic model summaries -----.
TITLE "Influence Diagnostics in SPSS".
668 0 M> TITLE "Influence Diagnostics in SPSS".
```

Influence Diagnostics in SPSS

```
SUBTITLE "Descriptive statistics and correlations".
```

```
669 0 M> SUBTITLE "Descriptive statistics and correlations".
```

Influence Diagnostics in SPSS

Descriptive statistics and correlations

```

670  0 M>
DESCRIPTIVES VARIABLES=G3 G1 G2 studytime failures absences age Medu Fedu
671  0 M>  DESCRIPTIVES VARIABLES=G3 G1 G2 studytime failures absences age Me
du Fedu
      /STATISTICS=MEAN STDDEV MIN MAX.
672  0 M>  /STATISTICS=MEAN STDDEV MIN MAX.

```

Descriptives

[student_clean] D:\low kda score priority basis posts\first post\Influence Dia
gnostics\SPSS\student_por_influence_diagnostics_clean_spss.sav

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
G3	649	0	19	11.91	3.231
G1	649	0	19	11.40	2.745
G2	649	0	19	11.57	2.914
studytime	649	1	4	1.93	.830
failures	649	0	3	.22	.593
absences	649	0	32	3.66	4.641
age	649	15	22	16.74	1.218
Medu	649	0	4	2.51	1.135
Fedu	649	0	4	2.31	1.100
Valid N (listwise)	649				

```

673  0 M>
CORRELATIONS
674  0 M>  CORRELATIONS
      /VARIABLES=G3 G1 G2 studytime failures absences age Medu Fedu
675  0 M>  /VARIABLES=G3 G1 G2 studytime failures absences age Medu Fedu
      /PRINT=TWOTAIL NOSIG
676  0 M>  /PRINT=TWOTAIL NOSIG
      /MISSING=PAIRWISE.
677  0 M>  /MISSING=PAIRWISE.

```

Correlations

Influence Diagnostics in SPSS
Descriptive statistics and correlations

Correlations

		G3	G1	G2	studytime	failures	absences
G3	Pearson Correlation	1	.826**	.919**	.250**	-.393**	-.091*
	Sig. (2-tailed)		.000	.000	.000	.000	.020
	N	649	649	649	649	649	649
G1	Pearson Correlation	.826**	1	.865**	.261**	-.384**	-.147**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	649	649	649	649	649	649
G2	Pearson Correlation	.919**	.865**	1	.240**	-.386**	-.125**
	Sig. (2-tailed)	.000	.000		.000	.000	.001
	N	649	649	649	649	649	649
studytime	Pearson Correlation	.250**	.261**	.240**	1	-.147**	-.118**
	Sig. (2-tailed)	.000	.000	.000		.000	.003
	N	649	649	649	649	649	649
failures	Pearson Correlation	-.393**	-.384**	-.386**	-.147**	1	.123**
	Sig. (2-tailed)	.000	.000	.000	.000		.002
	N	649	649	649	649	649	649
absences	Pearson Correlation	-.091*	-.147**	-.125**	-.118**	.123**	1
	Sig. (2-tailed)	.020	.000	.001	.003	.002	
	N	649	649	649	649	649	649
age	Pearson Correlation	-.107**	-.174**	-.107**	-.008	.320**	.150**
	Sig. (2-tailed)	.007	.000	.006	.831	.000	.000
	N	649	649	649	649	649	649
Medu	Pearson Correlation	.240**	.260**	.264**	.097*	-.172**	-.009
	Sig. (2-tailed)	.000	.000	.000	.013	.000	.827
	N	649	649	649	649	649	649
Fedu	Pearson Correlation	.212**	.218**	.225**	.050	-.166**	.030
	Sig. (2-tailed)	.000	.000	.000	.200	.000	.448
	N	649	649	649	649	649	649

Influence Diagnostics in SPSS
Descriptive statistics and correlations

Correlations

		age	Medu	Fedu
G3	Pearson Correlation	-.107**	.240**	.212**
	Sig. (2-tailed)	.007	.000	.000
	N	649	649	649
G1	Pearson Correlation	-.174**	.260**	.218**
	Sig. (2-tailed)	.000	.000	.000
	N	649	649	649
G2	Pearson Correlation	-.107**	.264**	.225**
	Sig. (2-tailed)	.006	.000	.000
	N	649	649	649
studytime	Pearson Correlation	-.008	.097*	.050
	Sig. (2-tailed)	.831	.013	.200
	N	649	649	649
failures	Pearson Correlation	.320**	-.172**	-.166**
	Sig. (2-tailed)	.000	.000	.000
	N	649	649	649
absences	Pearson Correlation	.150**	-.009	.030
	Sig. (2-tailed)	.000	.827	.448
	N	649	649	649
age	Pearson Correlation	1	-.108**	-.121**
	Sig. (2-tailed)		.006	.002
	N	649	649	649
Medu	Pearson Correlation	-.108**	1	.647**
	Sig. (2-tailed)	.006		.000
	N	649	649	649
Fedu	Pearson Correlation	-.121**	.647**	1
	Sig. (2-tailed)	.002	.000	
	N	649	649	649

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Influence Diagnostics in SPSS
Descriptive statistics and correlations

```
* ----- 6. Regression model and saved influence diagnostics -----.  
679 0 M> * ----- 6. Regression model and saved influence diagnostics -  
-----.  
* p = 9 model parameters including intercept:  
680 0 M> * p = 9 model parameters including intercept:  
* intercept + G1 + G2 + studytime + failures + absences + age + Medu + Fedu.  
681 0 M> * intercept + G1 + G2 + studytime + failures + absences + age + Me  
du + Fedu.  
  
682 0 M>  
TITLE "Regression Model for Influence Diagnostics".  
683 0 M> TITLE "Regression Model for Influence Diagnostics".
```

Regression Model for Influence Diagnostics

SUBTITLE "Dependent variable: G3".

684 0 M> SUBTITLE "Dependent variable: G3".

Regression Model for Influence Diagnostics
Dependent variable: G3

```
685 0 M>
REGRESSION
686 0 M> REGRESSION
      /MISSING LISTWISE
687 0 M> /MISSING LISTWISE
      /STATISTICS COEFF OUTS R ANOVA CHANGE CI(95) COLLIN TOL
688 0 M> /STATISTICS COEFF OUTS R ANOVA CHANGE CI(95) COLLIN TOL
      /DEPENDENT G3
689 0 M> /DEPENDENT G3
      /METHOD=ENTER G1 G2 studytime failures absences age Medu Fedu
690 0 M> /METHOD=ENTER G1 G2 studytime failures absences age Medu Fedu
      /SAVE PRED(fitted_g3)
691 0 M> /SAVE PRED(fitted_g3)
      RESID(residual_g3)
692 0 M> RESID(residual_g3)
      ZRESID(z_residual_g3)
693 0 M> ZRESID(z_residual_g3)
      SRESID(studentized_residual_g3)
694 0 M> SRESID(studentized_residual_g3)
      LEVER(leverage_hat)
695 0 M> LEVER(leverage_hat)
      COOK(cooks_distance)
696 0 M> COOK(cooks_distance)
      DFFIT(dffits)
697 0 M> DFFIT(dffits)
      COVRATIO(covratio)
698 0 M> COVRATIO(covratio)
      SDBETA.
699 0 M> SDBETA.
```

Regression

Regression Model for Influence Diagnostics
 Dependent variable: G3

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	Fedu, absences, studytime, age, G2, failures, Medu, G1 ^b	.	Enter

a. Dependent Variable: G3

b. All requested variables entered.

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics		
					R Square Change	F Change	df1
1	.922 ^a	.851	.849	1.256	.851	456.111	8

Model Summary^b

Model	Change Statistics	
	df2	Sig. F Change
1	640	.000

a. Predictors: (Constant), Fedu, absences, studytime, age, G2, failures, Medu, G1

b. Dependent Variable: G3

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5754.033	8	719.254	456.111	.000 ^b
	Residual	1009.234	640	1.577		
	Total	6763.267	648			

a. Dependent Variable: G3

b. Predictors: (Constant), Fedu, absences, studytime, age, G2, failures, Medu, G1

Regression Model for Influence Diagnostics
Dependent variable: G3

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.501	.774		-.648	.518
	G1	.143	.037	.122	3.910	.000
	G2	.885	.034	.798	25.744	.000
	studytime	.097	.062	.025	1.556	.120
	failures	-.235	.095	-.043	-2.471	.014
	absences	.023	.011	.033	2.085	.038
	age	.023	.044	.009	.520	.604
	Medu	-.045	.058	-.016	-.776	.438
	Fedu	.022	.059	.007	.371	.711

Coefficients^a

Model		95.0% Confidence Interval for B		Collinearity Statistics	
		Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-2.021	1.019		
	G1	.071	.215	.240	4.165
	G2	.817	.952	.243	4.121
	studytime	-.025	.219	.917	1.090
	failures	-.422	-.048	.762	1.312
	absences	.001	.044	.948	1.055
	age	-.063	.108	.860	1.163
	Medu	-.159	.069	.563	1.776
	Fedu	-.094	.139	.572	1.749

a. Dependent Variable: G3

Regression Model for Influence Diagnostics
 Dependent variable: G3

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions		
					G1	G2	studytime
1	1	7.051	1.000	.00	.00	.00	.00
	2	.945	2.731	.00	.00	.00	.00
	3	.570	3.518	.00	.00	.00	.00
	4	.211	5.777	.00	.00	.00	.15
	5	.108	8.064	.00	.01	.01	.83
	6	.062	10.680	.00	.00	.00	.00
	7	.043	12.869	.02	.06	.08	.01
	8	.008	30.135	.00	.87	.88	.00
	9	.002	56.814	.97	.06	.02	.00

Collinearity Diagnostics^a

Model	Dimension	Variance Proportions				
		failures	absences	age	Medu	Fedu
1	1	.00	.01	.00	.00	.00
	2	.61	.05	.00	.00	.00
	3	.12	.88	.00	.00	.00
	4	.01	.00	.00	.13	.19
	5	.00	.02	.00	.02	.02
	6	.00	.00	.00	.84	.76
	7	.21	.04	.03	.00	.01
	8	.00	.00	.00	.00	.00
	9	.04	.01	.96	.00	.01

a. Dependent Variable: G3

Regression Model for Influence Diagnostics
Dependent variable: G3

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.25	19.55	11.91	2.980	649
Std. Predicted Value	-3.911	2.565	.000	1.000	649
Standard Error of Predicted Value	.068	.417	.141	.043	649
Adjusted Predicted Value	.26	19.56	11.91	2.979	649
Residual	-9.051	5.782	.000	1.248	649
Std. Residual	-7.207	4.605	.000	.994	649
Stud. Residual	-7.296	4.625	.000	1.002	649
Deleted Residual	-9.275	5.833	-.001	1.268	649
Stud. Deleted Residual	-7.614	4.700	-.003	1.019	649
Mahal. Distance	.891	70.420	7.988	6.680	649
Cook's Distance	.000	.147	.002	.009	649
Centered Leverage Value	.001	.109	.012	.010	649

a. Dependent Variable: G3

```

700 0 M>
EXECUTE.
701 0 M> EXECUTE.

702 0 M>
* ----- 7. Diagnostic thresholds and flags -----
703 0 M> * ----- 7. Diagnostic thresholds and flags -----
COMPUTE all_cases_for_count=1.
704 0 M> COMPUTE all_cases_for_count=1.
AGGREGATE
705 0 M> AGGREGATE
  /OUTFILE=* MODE=ADDVARIABLES
706 0 M> /OUTFILE=* MODE=ADDVARIABLES
  /BREAK=all_cases_for_count
707 0 M> /BREAK=all_cases_for_count
  /n_cases=N.
708 0 M> /n_cases=N.

709 0 M>

```

Regression Model for Influence Diagnostics
Dependent variable: G3

```
COMPUTE p_parameters=9.
710 0 M> COMPUTE p_parameters=9.
COMPUTE cooks_4_n=4/n_cases.
711 0 M> COMPUTE cooks_4_n=4/n_cases.
COMPUTE cooks_4_np=4/(n_cases-p_parameters).
712 0 M> COMPUTE cooks_4_np=4/(n_cases-p_parameters).
COMPUTE leverage_2p_n=2*p_parameters/n_cases.
713 0 M> COMPUTE leverage_2p_n=2*p_parameters/n_cases.
COMPUTE leverage_3p_n=3*p_parameters/n_cases.
714 0 M> COMPUTE leverage_3p_n=3*p_parameters/n_cases.
COMPUTE dffits_cutoff=2*SQRT(p_parameters/n_cases).
715 0 M> COMPUTE dffits_cutoff=2*SQRT(p_parameters/n_cases).
COMPUTE dfbetas_cutoff=2/SQRT(n_cases).
716 0 M> COMPUTE dfbetas_cutoff=2/SQRT(n_cases).
COMPUTE covratio_lower=1-(3*p_parameters/n_cases).
717 0 M> COMPUTE covratio_lower=1-(3*p_parameters/n_cases).
COMPUTE covratio_upper=1+(3*p_parameters/n_cases).
718 0 M> COMPUTE covratio_upper=1+(3*p_parameters/n_cases).

719 0 M>
COMPUTE flag_cook_4n=(cooks_distance > cooks_4_n).
720 0 M> COMPUTE flag_cook_4n=(cooks_distance > cooks_4_n).
COMPUTE flag_cook_4np=(cooks_distance > cooks_4_np).
721 0 M> COMPUTE flag_cook_4np=(cooks_distance > cooks_4_np).
COMPUTE flag_leverage_2pn=(leverage_hat > leverage_2p_n).
722 0 M> COMPUTE flag_leverage_2pn=(leverage_hat > leverage_2p_n).
COMPUTE flag_leverage_3pn=(leverage_hat > leverage_3p_n).
723 0 M> COMPUTE flag_leverage_3pn=(leverage_hat > leverage_3p_n).
COMPUTE flag_studentized_2=(ABS(studentized_residual_g3) > 2).
724 0 M> COMPUTE flag_studentized_2=(ABS(studentized_residual_g3) > 2).
COMPUTE flag_studentized_3=(ABS(studentized_residual_g3) > 3).
725 0 M> COMPUTE flag_studentized_3=(ABS(studentized_residual_g3) > 3).
COMPUTE flag_dffits=(ABS(dffits) > dffits_cutoff).
726 0 M> COMPUTE flag_dffits=(ABS(dffits) > dffits_cutoff).
COMPUTE flag_covratio=(covratio < covratio_lower OR covratio > covratio_upper)
.
727 0 M> COMPUTE flag_covratio=(covratio < covratio_lower OR covratio > covratio_upper).
```

Regression Model for Influence Diagnostics
Dependent variable: G3

```
728 0 M>
COMPUTE influence_flag_count =
729 0 M> COMPUTE influence_flag_count =
      SUM(flag_cook_4n, flag_cook_4np, flag_leverage_2pn, flag_leverage_3pn,
730 0 M>   SUM(flag_cook_4n, flag_cook_4np, flag_leverage_2pn, flag_leverage
_3pn,
      flag_studentized_2, flag_studentized_3, flag_dffits, flag_covratio).
731 0 M>   flag_studentized_2, flag_studentized_3, flag_dffits, flag_covrati
o).

732 0 M>
COMPUTE likely_influential=(influence_flag_count >= 2).
733 0 M> COMPUTE likely_influential=(influence_flag_count >= 2).
EXECUTE.
734 0 M> EXECUTE.

735 0 M>
VARIABLE LABELS
736 0 M> VARIABLE LABELS
      case_id "Original case index"
      fitted_g3 "Predicted or fitted G3"
737 0 M>   case_id "Original case index"
      residual_g3 "Raw residual for G3"
738 0 M>   fitted_g3 "Predicted or fitted G3"
      z_residual_g3 "Standardized residual for G3"
739 0 M>   residual_g3 "Raw residual for G3"
      studentized_residual_g3 "Studentized residual for G3"
740 0 M>   z_residual_g3 "Standardized residual for G3"
      leverage_hat "Leverage or hat value"
741 0 M>   studentized_residual_g3 "Studentized residual for G3"
      cooks_distance "Cook's Distance"
742 0 M>   leverage_hat "Leverage or hat value"
      dffits "DFFITS"
743 0 M>   cooks_distance "Cook's Distance"
      covratio "COVRATIO"
744 0 M>   dffits "DFFITS"
      cooks_4_n "Cook's Distance threshold: 4/n"
745 0 M>   covratio "COVRATIO"
      cooks_4_np "Cook's Distance threshold: 4/(n-p)"
```

Regression Model for Influence Diagnostics
Dependent variable: G3

```
746 0 M>  cooks_4_n "Cook's Distance threshold: 4/n"
      leverage_2p_n "Leverage threshold: 2p/n"
747 0 M>  cooks_4_np "Cook's Distance threshold: 4/(n-p)"
      leverage_3p_n "Leverage threshold: 3p/n"
748 0 M>  leverage_2p_n "Leverage threshold: 2p/n"
      dffits_cutoff "DFFITs threshold: 2*sqrt(p/n)"
749 0 M>  leverage_3p_n "Leverage threshold: 3p/n"
      dfbetas_cutoff "DFBETAS threshold: 2/sqrt(n)"
750 0 M>  dffits_cutoff "DFFITs threshold: 2*sqrt(p/n)"
      covratio_lower "COVRATIO lower threshold: 1 - 3p/n"
751 0 M>  dfbetas_cutoff "DFBETAS threshold: 2/sqrt(n)"
      covratio_upper "COVRATIO upper threshold: 1 + 3p/n"
752 0 M>  covratio_lower "COVRATIO lower threshold: 1 - 3p/n"
      influence_flag_count "Number of influence diagnostic flags"
753 0 M>  covratio_upper "COVRATIO upper threshold: 1 + 3p/n"
      likely_influential "Likely influential case: 2 or more diagnostic flags".
754 0 M>  influence_flag_count "Number of influence diagnostic flags"
755 0 M>  likely_influential "Likely influential case: 2 or more diagnostic
flags".

756 0 M>
VALUE LABELS
757 0 M>  VALUE LABELS
      flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
758 0 M>  flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
      flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio likely_influ
ential
759 0 M>  flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio l
ikely_influential
      0 "No" 1 "Yes".
760 0 M>  0 "No" 1 "Yes".
EXECUTE.
761 0 M>  EXECUTE.

762 0 M>
* ----- 8. Diagnostic flag summary tables -----
763 0 M>  * ----- 8. Diagnostic flag summary tables -----
TITLE "Influence Diagnostics Flag Summary".
764 0 M>  TITLE "Influence Diagnostics Flag Summary".
```

Regression Model for Influenza Diagnostics
Dependent variable: G3

Influence Diagnostics Flag Summary

SUBTITLE "Common influence thresholds".

765 0 M> SUBTITLE "Common influence thresholds".

Influence Diagnostics Flag Summary
Common influence thresholds

```

766  0 M>
FREQUENCIES VARIABLES=
767  0 M>  FREQUENCIES VARIABLES=
         flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
768  0 M>  flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
         flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio likely_influ
ential
769  0 M>  flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio l
ikely_influential
         /ORDER=ANALYSIS.
770  0 M>  /ORDER=ANALYSIS.

```

Frequencies

Statistics

		flag_cook_4n	flag_cook_4np	flag_leverage_2 pn	flag_leverage_3 pn	flag_studentize d_2
N	Valid	649	649	649	649	649
	Missing	0	0	0	0	0

Statistics

		flag_studentize d_3	flag_dffits	flag_covratio	Likely influential case: 2 or more diagnostic flags
N	Valid	649	649	649	649
	Missing	0	0	0	0

Frequency Table

Influence Diagnostics Flag Summary
Common influence thresholds

flag_cook_4n

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	621	95.7	95.7	95.7
	Yes	28	4.3	4.3	100.0
	Total	649	100.0	100.0	

flag_cook_4np

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	621	95.7	95.7	95.7
	Yes	28	4.3	4.3	100.0
	Total	649	100.0	100.0	

flag_leverage_2pn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	615	94.8	94.8	94.8
	Yes	34	5.2	5.2	100.0
	Total	649	100.0	100.0	

flag_leverage_3pn

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	631	97.2	97.2	97.2
	Yes	18	2.8	2.8	100.0
	Total	649	100.0	100.0	

flag_studentized_2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	631	97.2	97.2	97.2
	Yes	18	2.8	2.8	100.0
	Total	649	100.0	100.0	

Influence Diagnostics Flag Summary
Common influence thresholds

flag_studentized_3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	639	98.5	98.5	98.5
	Yes	10	1.5	1.5	100.0
	Total	649	100.0	100.0	

flag_dffits

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	649	100.0	100.0	100.0

flag_covratio

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	607	93.5	93.5	93.5
	Yes	42	6.5	6.5	100.0
	Total	649	100.0	100.0	

Likely influential case: 2 or more diagnostic flags

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	595	91.7	91.7	91.7
	Yes	54	8.3	8.3	100.0
	Total	649	100.0	100.0	

```

771 0 M>
DESCRIPTIVES VARIABLES=
772 0 M> DESCRIPTIVES VARIABLES=
    fitted_g3 residual_g3 z_residual_g3 studentized_residual_g3 leverage_hat coo
ks_distance dffits covratio
773 0 M> fitted_g3 residual_g3 z_residual_g3 studentized_residual_g3 lever
age_hat cooks_distance dffits covratio
    /STATISTICS=MEAN STDDEV MIN MAX.
774 0 M> /STATISTICS=MEAN STDDEV MIN MAX.

```

Influence Diagnostics Flag Summary
Common influence thresholds

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Predicted or fitted G3	649	.25124	19.55068	11.9060092	2.97987928
Raw residual for G3	649	-9.05080	5.78244	.0000000	1.24798207
Standardized residual for G3	649	-7.20744	4.60474	.0000000	.99380799
Studentized residual for G3	649	-7.29633	4.62464	-.0004196	1.00179810
Leverage or hat value	649	.00137	.10867	.0123267	.01030913
Cook's Distance	649	.00000	.14681	.0018089	.00890335
DFFITs	649	-.22464	.14459	-.0010611	.02379735
COVRATIO	649	.47567	1.12432	1.0157148	.05200999
Valid N (listwise)	649				

```

775  0 M>
* ----- 9. Top influential case tables -----
776  0 M> * ----- 9. Top influential case tables -----
SORT CASES BY cooks_distance(D).
777  0 M> SORT CASES BY cooks_distance(D).
TITLE "Top 20 Cases by Cook's Distance".
778  0 M> TITLE "Top 20 Cases by Cook's Distance".

```

Top 20 Cases by Cook's Distance

```

LIST VARIABLES=case_id G3 fitted_g3 residual_g3 leverage_hat studentized_resid
ual_g3 cooks_distance dffits covratio influence_flag_count
779 0 M> LIST VARIABLES=case_id G3 fitted_g3 residual_g3 leverage_hat stude
ntized_residual_g3 cooks_distance dffits covratio infl
uence_flag_count
/CASES=FROM 1 TO 20.
780 0 M> /CASES=FROM 1 TO 20.

```

List

case_id	G3	fitted_g3	residual_g3	leverage_hat	studentized_residual_g3	cooks_distance
dfits		covratio	influence_flag_count			
164	0	9.05080	-9.05080	.02268	-7.29633	.146
81	-.22464	.47567	5.00			
640	0	7.56366	-7.56366	.01949	-6.08753	.088
45	-.16248	.60562	5.00			
587	0	8.16764	-8.16764	.01278	-6.55122	.069
27	-.11865	.55088	5.00			
641	0	6.83013	-6.83013	.01706	-5.49035	.063
48	-.12945	.66941	5.00			
627	0	5.43656	-5.43656	.02417	-4.38606	.056
41	-.14349	.79092	5.00			
638	0	7.13337	-7.13337	.01304	-5.72241	.053
85	-.10558	.64151	5.00			
173	1	10.03298	-9.03298	.00693	-7.22391	.049
53	-.07716	.47572	5.00			
520	0	7.32802	-7.32802	.00905	-5.86669	.040
94	-.07845	.62327	5.00			
584	0	6.17862	-6.17862	.01183	-4.95345	.036
94	-.08371	.72303	5.00			
280	5	7.63103	-2.63103	.05624	-2.15846	.031
75	-.16135	1.00787	5.00			
62	16	10.21756	5.78244	.00705	4.62464	.020
58	.05009	.75337	5.00			
170	10	7.96184	2.03816	.04655	1.66354	.015
53	.10297	1.02466	4.00			
1	11	9.83267	1.16733	.10867	.98547	.013
37	.14459	1.12432	5.00			

Top 20 Cases by Cook's Distance

	558	10	8.27142	1.72858	.04876	1.41251	.011
74	.09156		1.03830	4.00			
	559	10	12.58495	-2.58495	.02108	-2.08216	.011
15	-.05982		.97607	3.00			
	519	8	5.61659	2.38341	.02456	1.92325	.011
02	.06389		.98843	2.00			
	156	6	7.89824	-1.89824	.03448	-1.53961	.009
84	-.07092		1.01751	3.00			
	573	7	4.81667	2.18333	.02557	1.76272	.009
62	.06085		.99776	2.00			
	326	10	11.63436	-1.63436	.04220	-1.33092	.009
00	-.07475		1.03443	4.00			
	174	10	7.85585	2.14415	.02333	1.72909	.008
47	.05469		.99712	2.00			

Number of cases read: 20 Number of cases listed: 20

```
781 0 M>
SORT CASES BY leverage_hat(D).
782 0 M> SORT CASES BY leverage_hat(D).
TITLE "Top 20 Cases by Leverage".
783 0 M> TITLE "Top 20 Cases by Leverage".
```

Top 20 Cases by Leverage

```
LIST VARIABLES=case_id G3 fitted_g3 residual_g3 leverage_hat cooks_distance in  
fluence_flag_count  
784 0 M> LIST VARIABLES=case_id G3 fitted_g3 residual_g3 leverage_hat cooks  
_distance influence_flag_count  
/CASES=FROM 1 TO 20.  
785 0 M> /CASES=FROM 1 TO 20.
```

List

Top 20 Cases by Leverage

case_id	G3	fitted_g3	residual_g3	leverage_hat	cooks_distance	influen
1	11	9.83267	1.16733	.10867	.01337	
5.00						
611	0	.54130	-.54130	.08302	.00208	
3.00						
598	0	1.10980	-1.10980	.07048	.00726	
5.00						
198	14	14.24742	-.24742	.06442	.00033	
3.00						
213	16	15.61659	.38341	.06298	.00076	
3.00						
280	5	7.63103	-2.63103	.05624	.03175	
5.00						
569	9	9.91392	-.91392	.05612	.00382	
3.00						
604	0	.68168	-.68168	.05345	.00202	
3.00						
441	0	1.03593	-1.03593	.05160	.00448	
3.00						
558	10	8.27142	1.72858	.04876	.01174	
4.00						
564	0	.70369	-.70369	.04819	.00192	
3.00						
544	8	8.48270	-.48270	.04793	.00090	
3.00						
170	10	7.96184	2.03816	.04655	.01553	
4.00						
479	8	6.52350	1.47650	.04545	.00795	
5.00						
408	12	11.61097	.38903	.04485	.00054	
3.00						
414	10	10.29150	-.29150	.04321	.00029	
3.00						

Top 20 Cases by Leverage

79	10	8.46101	1.53899	.04239	.00802
4.00					
326	10	11.63436	-1.63436	.04220	.00900
4.00					
171	11	9.83682	1.16318	.04160	.00449
2.00					
180	8	6.75269	1.24731	.04141	.00514
2.00					

Number of cases read: 20 Number of cases listed: 20

```
786 0 M>
SORT CASES BY studentized_residual_g3(A).
787 0 M> SORT CASES BY studentized_residual_g3(A).
TITLE "Top 20 Negative Studentized Residual Cases".
788 0 M> TITLE "Top 20 Negative Studentized Residual Cases".
```

Top 20 Negative Studentized Residual Cases

```
LIST VARIABLES=case_id G3 fitted_g3 residual_g3 studentized_residual_g3 cooks_
distance influence_flag_count
789 0 M> LIST VARIABLES=case_id G3 fitted_g3 residual_g3 studentized_residu
al_g3 cooks_distance influence_flag_count
/CASES=FROM 1 TO 20.
790 0 M> /CASES=FROM 1 TO 20.
```

List

Top 20 Negative Studentized Residual Cases

	case_id	G3	fitted_g3	residual_g3	studentized_residual_g3	cooks_distance
	164	0	9.05080	-9.05080	-7.29633	.1468
1		5.00				
	173	1	10.03298	-9.03298	-7.22391	.0495
3		5.00				
	587	0	8.16764	-8.16764	-6.55122	.0692
7		5.00				
	640	0	7.56366	-7.56366	-6.08753	.0884
5		5.00				
	520	0	7.32802	-7.32802	-5.86669	.0409
4		5.00				
	638	0	7.13337	-7.13337	-5.72241	.0538
5		5.00				
	641	0	6.83013	-6.83013	-5.49035	.0634
8		5.00				
	584	0	6.17862	-6.17862	-4.95345	.0369
4		5.00				
	627	0	5.43656	-5.43656	-4.38606	.0564
1		5.00				
	63	10	13.42083	-3.42083	-2.73461	.0064
1		4.00				
	647	9	12.40915	-3.40915	-2.72552	.0065
2		4.00				
	280	5	7.63103	-2.63103	-2.15846	.0317
5		5.00				
	559	10	12.58495	-2.58495	-2.08216	.0111
5		3.00				
	642	15	17.20051	-2.20051	-1.76525	.0051
2		.00				
	156	6	7.89824	-1.89824	-1.53961	.0098
4		3.00				
	41	10	11.68366	-1.68366	-1.35108	.0031
4		.00				

Top 20 Negative Studentized Residual Cases

0	326	10	11.63436	-1.63436	-1.33092	.0090
		4.00				
6	285	6	7.63850	-1.63850	-1.31713	.0036
		.00				
3	40	12	13.58205	-1.58205	-1.26734	.0021
		.00				
7	110	14	15.52859	-1.52859	-1.22264	.0014
		.00				

Number of cases read: 20 Number of cases listed: 20

```
791 0 M>
SORT CASES BY studentized_residual_g3(D).
792 0 M> SORT CASES BY studentized_residual_g3(D).
TITLE "Top 20 Positive Studentized Residual Cases".
793 0 M> TITLE "Top 20 Positive Studentized Residual Cases".
```

Top 20 Positive Studentized Residual Cases

```
LIST VARIABLES=case_id G3 fitted_g3 residual_g3 studentized_residual_g3 cooks_
distance influence_flag_count
794 0 M> LIST VARIABLES=case_id G3 fitted_g3 residual_g3 studentized_residu
al_g3 cooks_distance influence_flag_count
/CASES=FROM 1 TO 20.
795 0 M> /CASES=FROM 1 TO 20.
```

List

Top 20 Positive Studentized Residual Cases

case_id	G3	fitted_g3	residual_g3	studentized_residual_g3	cooks_distance
62	16	10.21756	5.78244	4.62464	.0205
8	5.00				
308	14	11.27518	2.72482	2.17966	.0047
8	2.00				
359	15	12.41063	2.58937	2.07314	.0051
8	1.00				
355	15	12.41654	2.58346	2.06434	.0032
5	1.00				
390	12	9.45183	2.54817	2.03750	.0037
9	1.00				
519	8	5.61659	2.38341	1.92325	.0110
2	2.00				
401	17	14.71265	2.28735	1.83163	.0041
6	.00				
643	11	8.78768	2.21232	1.77950	.0071
3	2.00				
573	7	4.81667	2.18333	1.76272	.0096
2	2.00				
366	17	14.82096	2.17904	1.74926	.0055
2	.00				
174	10	7.85585	2.14415	1.72909	.0084
7	2.00				
176	8	5.91991	2.08009	1.67369	.0065
1	2.00				
170	10	7.96184	2.03816	1.66354	.0155
3	4.00				
459	13	10.93764	2.06236	1.66111	.0070
6	2.00				
369	11	9.06808	1.93192	1.55064	.0042
5	.00				
310	14	12.10740	1.89260	1.51658	.0032
1	.00				

Top 20 Positive Studentized Residual Cases

6	323	10	8.16600	1.83400	1.47034	.0032
		.00				
2	588	8	6.20014	1.79986	1.44016	.0022
		.00				
1	365	17	15.21462	1.78538	1.42928	.0024
		.00				
9	393	15	13.21022	1.78978	1.42837	.0009
		.00				

Number of cases read: 20 Number of cases listed: 20

```

796 0 M>
* ----- 10. Save final case-level diagnostic metrics -----
797 0 M> * ----- 10. Save final case-level diagnostic metrics -----
--
SAVE TRANSLATE
798 0 M> SAVE TRANSLATE
   /OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\influence_diagnostics_case_metrics_spss.csv'
   /TYPE=CSV
799 0 M> /OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\influence_diagnostics_case_metrics_spss.csv'
800 0 M> /TYPE=CSV
   /ENCODING='UTF8'
   /REPLACE
801 0 M> /ENCODING='UTF8'
802 0 M> /REPLACE
   /FIELDNAMES
803 0 M> /FIELDNAMES
   /CELLS=VALUES
804 0 M> /CELLS=VALUES
   /KEEP=case_id G3 G1 G2 studytime failures absences age Medu Fedu
805 0 M> /KEEP=case_id G3 G1 G2 studytime failures absences age Medu Fedu
      fitted_g3 residual_g3 z_residual_g3 studentized_residual_g3
806 0 M> fitted_g3 residual_g3 z_residual_g3 studentized_residual_g3
      leverage_hat cooks_distance dffits covratio
807 0 M> leverage_hat cooks_distance dffits covratio
      cooks_4_n cooks_4_np leverage_2p_n leverage_3p_n dffits_cutoff

```

Top 20 Positive Studentized Residual Cases

```
808 0 M>  cooks_4_n cooks_4_np leverage_2p_n leverage_3p_n dffits_cutoff
      dfbetas_cutoff covratio_lower covratio_upper
809 0 M>  dfbetas_cutoff covratio_lower covratio_upper
      flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
810 0 M>  flag_cook_4n flag_cook_4np flag_leverage_2pn flag_leverage_3pn
      flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio
811 0 M>  flag_studentized_2 flag_studentized_3 flag_dffits flag_covratio
      influence_flag_count likely_influential.
812 0 M>  influence_flag_count likely_influential.

813 0 M>
SAVE OUTFILE='D:\low kda score priority basis posts\first post\Influence Diagn
ostics\SPSS\influence_diagnostics_case_metrics_spss.sav'
/COMPRESSED.
814 0 M>  SAVE OUTFILE='D:\low kda score priority basis posts\first post\Inf
luence Diagnostics\SPSS\influence_diagnostics_case_met
      rics_spss.sav'
815 0 M>  /COMPRESSED.

816 0 M>
* ----- 11. Export SPSS Viewer output automatically -----
817 0 M>  * ----- 11. Export SPSS Viewer output automatically -----
-
TITLE "Influence Diagnostics SPSS Run Complete".
818 0 M>  TITLE "Influence Diagnostics SPSS Run Complete".
```

Influence Diagnostics SPSS Run Complete

SUBTITLE "Saved in SPSS folder".

819 0 M> SUBTITLE "Saved in SPSS folder".

Influence Diagnostics SPSS Run Complete
Saved in SPSS folder

```
820 0 M>
OUTPUT EXPORT
821 0 M> OUTPUT EXPORT
      /CONTENTS EXPORT=ALL LAYERS=PRINTSETTING MODELVIEWS=PRINTSETTING
822 0 M> /CONTENTS EXPORT=ALL LAYERS=PRINTSETTING MODELVIEWS=PRINTSETTING
      /PDF DOCUMENTFILE='D:\low kda score priority basis posts\first post\Influenc
e Diagnostics\SPSS\Influence-Diagnostics-SPSS-output.pdf'
      EMBEDBOOKMARKS=YES
823 0 M> /PDF DOCUMENTFILE='D:\low kda score priority basis posts\first po
st\Influence Diagnostics\SPSS\Influence-Diagnostics-SP
      SS-output.pdf'
824 0 M> EMBEDBOOKMARKS=YES
      EMBEDFONTS=YES.
825 0 M> EMBEDFONTS=YES.
```

Output Export

[student_clean] D:\low kda score priority basis posts\first post\Influence Dia
gnostics\SPSS\influence_diagnostics_case_metrics_spss.sav

Export Summary

Viewer	InfluenceDiagnosticsOutput
Document File	D:\low kda score priority basis posts\first post\Influence Diagnostics\SPSS\Influence- Diagnostics-SPSS-output.pdf

```
826 0 M>
OUTPUT SAVE OUTFILE='D:\low kda score priority basis posts\first post\Influenc
e Diagnostics\SPSS\Influence-Diagnostics-SPSS-output.spv'.
827 0 M> OUTPUT SAVE OUTFILE='D:\low kda score priority basis posts\first p
ost\Influence Diagnostics\SPSS\Influence-Diagnostics-S
      PSS-output.spv'.
```

Output Save

Influence Diagnostics SPSS Run Complete
Saved in SPSS folder

Warnings

Cannot write to OUTFILE='D:\low kda score priority basis
posts\first post\Influence Diagnostics\SPSS\Influence-
Diagnostics-SPSS-output.spv'. File is open in another window.

Execution of this command stops.

828 0 M>

* =====.

829 0 M> * =====.

* End of fixed SPSS syntax.

830 0 M> * End of fixed SPSS syntax.

* =====.

831 0 M> * =====.