

```
1404 0 OUTPUT ACTIVATE SurveyDataCleaningOutput.  
1404 0 M> OUTPUT ACTIVATE SurveyDataCleaningOutput.  
1405 0 TITLE 'Survey Data Cleaning'.  
1405 0 M> TITLE 'Survey Data Cleaning'.
```

Survey Data Cleaning

```
1406 0
1406 0 M>
1407 0 GET DATA
1407 0 M> GET DATA
1408 0 /TYPE=TXT /FILE='D:\DATA ANALYSIS\M Survey and Likert Scale Analysis
\Survey Data Cleaning\dataset.csv' /ENCODING='UTF8' /D
      ELCASE=LINE /DELIMITERS=', ' /QUALIFIER='"'
1408 0 M> /TYPE=TXT /FILE='D:\DATA ANALYSIS\M Survey and Likert Scale Analy
sis\Survey Data Cleaning\dataset.csv' /ENCODING='UTF8'
      /DELCASE=LINE /DELIMITERS=', ' /QUALIFIER='"'
1409 0 /ARRANGEMENT=DELIMITED /FIRSTCASE=2
1409 0 M> /ARRANGEMENT=DELIMITED /FIRSTCASE=2
1410 0 /VARIABLES=
1410 0 M> /VARIABLES=
1411 0 school A8 sex A8 age F8.0 address A8 famsize A8 Pstatus A8 Medu F8.0
Fedu F8.0
1411 0 M> school A8 sex A8 age F8.0 address A8 famsize A8 Pstatus A8 Medu F
8.0 Fedu F8.0
1412 0 Mjob A16 Fjob A16 reason A16 guardian A16 traveltime F8.0 studytime
F8.0 failures F8.0
1412 0 M> Mjob A16 Fjob A16 reason A16 guardian A16 traveltime F8.0 studyti
me F8.0 failures F8.0
1413 0 schoolsup A8 famsup A8 paid A8 activities A8 nursery A8 higher A8 in
ternet A8 romantic A8
1413 0 M> schoolsup A8 famsup A8 paid A8 activities A8 nursery A8 higher A8
internet A8 romantic A8
1414 0 famrel F8.0 freetime F8.0 goout F8.0 Dalc F8.0 Walc F8.0 health F8.0
absences F8.0 G1 F8.0 G2 F8.0 G3 F8.0.
1414 0 M> famrel F8.0 freetime F8.0 goout F8.0 Dalc F8.0 Walc F8.0 health F
8.0 absences F8.0 G1 F8.0 G2 F8.0 G3 F8.0.
1415 0 DATASET NAME SurveyDataCleaningData.
1415 0 M> DATASET NAME SurveyDataCleaningData.
1416 0 COMPUTE subject_id=$CASENUM.
1416 0 M> COMPUTE subject_id=$CASENUM.
1417 0 COMPUTE Dalc_R=6-Dalc.
1417 0 M> COMPUTE Dalc_R=6-Dalc.
1418 0 COMPUTE Walc_R=6-Walc.
1418 0 M> COMPUTE Walc_R=6-Walc.
```

Survey Data Cleaning

```
1419 0  COMPUTE scale_score=MEAN.6(famrel,freetime,goout,Dalc_R,Walc_R,health
).
1419 0 M>  COMPUTE scale_score=MEAN.6(famrel,freetime,goout,Dalc_R,Walc_R,hea
lth).
1420 0  RECODE sex ('F'=1) ('M'=2) INTO sex_code.
1420 0 M>  RECODE sex ('F'=1) ('M'=2) INTO sex_code.
1421 0  RECODE school ('GP'=1) ('MS'=2) INTO school_code.
1421 0 M>  RECODE school ('GP'=1) ('MS'=2) INTO school_code.
1422 0  RECODE higher internet ('no'=0) ('yes'=1) INTO higher_code internet_c
ode.
1422 0 M>  RECODE higher internet ('no'=0) ('yes'=1) INTO higher_code interne
t_code.
1423 0  EXECUTE.
1423 0 M>  EXECUTE.
1424 0
1424 0 M>
1425 0  * BEGIN METHOD-SPECIFIC SPSS WORKFLOW.
1425 0 M>  * BEGIN METHOD-SPECIFIC SPSS WORKFLOW.
1426 0  TITLE 'Survey Data Cleaning'.
1426 0 M>  TITLE 'Survey Data Cleaning'.
```

Survey Data Cleaning

```
1427 0 SUBTITLE 'Duplicate missing domain and numeric-range validation'.  
1427 0 M> SUBTITLE 'Duplicate missing domain and numeric-range validation'.
```

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

```
1428  0  COMPUTE category_violation=(NOT ANY(school,'GP','MS') OR NOT ANY(sex,
'F','M') OR NOT ANY(address,'R','U') OR NOT ANY(famsiz
e,'LE3','GT3') OR NOT ANY(Pstatus,'A','T')).
1428  0 M> COMPUTE category_violation=(NOT ANY(school,'GP','MS') OR NOT ANY(s
ex,'F','M') OR NOT ANY(address,'R','U') OR NOT ANY(fam
size,'LE3','GT3') OR NOT ANY(Pstatus,'A','T')).
1429  0  COMPUTE range_violation=(age<15 OR age>22 OR traveltime<1 OR travelti
me>4 OR studytime<1 OR studytime>4 OR failures<0 OR fa
ilures>3 OR famrel<1 OR famrel>5 OR freetime<1 OR freetime>5 OR go
out<1 OR goout>5 OR Dalc<1 OR Dalc>5 OR Walc<1 OR Walc
>5 OR health<1 OR health>5 OR absences<0 OR G1<0 OR G1>20 OR G2<0
OR G2>20 OR G3<0 OR G3>20).
1429  0 M> COMPUTE range_violation=(age<15 OR age>22 OR traveltime<1 OR trave
ltime>4 OR studytime<1 OR studytime>4 OR failures<0 OR
failures>3 OR famrel<1 OR famrel>5 OR freetime<1 OR freetime>5
OR goout<1 OR goout>5 OR Dalc<1 OR Dalc>5 OR Walc<1 O
R Walc>5 OR health<1 OR health>5 OR absences<0 OR G1<0 OR G1>20
OR G2<0 OR G2>20 OR G3<0 OR G3>20).
1430  0  COUNT string_missing=school sex address famsize Pstatus Mjob Fjob rea
son guardian schoolsup famsup paid activities nursery
higher internet romantic ('').
1430  0 M> COUNT string_missing=school sex address famsize Pstatus Mjob Fjob
reason guardian schoolsup famsup paid activities nurse
ry higher internet romantic ('').
1431  0  COUNT numeric_missing=age Medu Fedu traveltime studytime failures fam
rel freetime goout Dalc Walc health absences G1 G2 G3
(MISSING).
1431  0 M> COUNT numeric_missing=age Medu Fedu traveltime studytime failures
famrel freetime goout Dalc Walc health absences G1 G2
G3 (MISSING).
1432  0  COMPUTE missing_cells=string_missing+numeric_missing.
1432  0 M> COMPUTE missing_cells=string_missing+numeric_missing.
1433  0  FREQUENCIES VARIABLES=category_violation range_violation missing_cell
s.
1433  0 M> FREQUENCIES VARIABLES=category_violation range_violation missing_c
ells.
```

Frequencies

Survey Data Cleaning

Duplicate missing domain and numeric-range validation

Notes

Output Created		18-JUL-2026 13:22:36
Comments		
Input	Data	D:\DATA ANALYSIS\IM Survey and Likert Scale Analysis\Survey Data Cleaning\dataset.csv
	Active Dataset	SurveyDataCleaningData
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	649
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES=category_violation range_violation missing_cells.
Resources	Processor Time	00:00:00.03
	Elapsed Time	00:00:00.03

[SurveyDataCleaningData]

Statistics

		category_violation	range_violation	missing_cells
N	Valid	649	649	649
	Missing	0	0	0

Frequency Table

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

category_violation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	649	100.0	100.0	100.0

range_violation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	649	100.0	100.0	100.0

missing_cells

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	649	100.0	100.0	100.0

```

1434  0  DATASET COPY DistinctSurveyRows.
1434  0 M> DATASET COPY DistinctSurveyRows.
1435  0  DATASET ACTIVATE DistinctSurveyRows.
1435  0 M> DATASET ACTIVATE DistinctSurveyRows.
1436  0  AGGREGATE OUTFILE=* /BREAK=school sex age address famsize Pstatus Med
u Fedu Mjob Fjob reason guardian traveltime studytime
      failures schoolsup famsup paid activities nursery higher internet
romantic famrel freetime goout Dalc Walc health absenc
      es G1 G2 G3 /duplicate_count=N.
1436  0 M> AGGREGATE OUTFILE=* /BREAK=school sex age address famsize Pstatus
Medu Fedu Mjob Fjob reason guardian traveltime studyti
      me failures schoolsup famsup paid activities nursery higher int
ernet romantic famrel freetime goout Dalc Walc health
      absences G1 G2 G3 /duplicate_count=N.
1437  0  FREQUENCIES VARIABLES=duplicate_count /STATISTICS=SUM MINIMUM MAXIMUM
.
1437  0 M> FREQUENCIES VARIABLES=duplicate_count /STATISTICS=SUM MINIMUM MAXI
MUM.

```

Frequencies

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

Notes

Output Created		18-JUL-2026 13:22:36
Comments		
Input	File Label	Aggregated File
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	649
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES=duplicate_count /STATISTICS=SUM MINIMUM MAXIMUM.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

Statistics

duplicate_count

N	Valid	649
	Missing	0
Minimum		1
Maximum		1
Sum		649

duplicate_count

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	649	100.0	100.0	100.0

1438 0 ECHO 'Only exact duplicate rows are collapsed; valid categories and numeric values are not altered'.

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

```
1438 0 M> ECHO 'Only exact duplicate rows are collapsed; valid categories and
numeric values are not altered'.
Only exact duplicate rows are collapsed; valid categories and numeric values are
not altered
1439 0 * END METHOD-SPECIFIC SPSS WORKFLOW.
1439 0 M> * END METHOD-SPECIFIC SPSS WORKFLOW.
1440 0
1440 0 M>
1441 0 SAVE OUTFILE='D:\DATA ANALYSIS\M Survey and Likert Scale Analysis\Survey
Data Cleaning\SPSS_Output\sav\survey-data-cleaning
_spss_ready_data.sav' /COMPRESSED.
1441 0 M> SAVE OUTFILE='D:\DATA ANALYSIS\M Survey and Likert Scale Analysis\Survey
Data Cleaning\SPSS_Output\sav\survey-data-cleaning_spss_ready_data.sav' /COMPRESSED.

1442 0
1442 0 M>
1443 0 ECHO '=====
====='.
1443 0 M> ECHO '=====
====='.
=====
1444 0 ECHO 'SURVEY DATA CLEANING'.
1444 0 M> ECHO 'SURVEY DATA CLEANING'.
SURVEY DATA CLEANING
1445 0 ECHO 'Method type: data_cleaning_workflow'.
1445 0 M> ECHO 'Method type: data_cleaning_workflow'.
Method type: data_cleaning_workflow
1446 0 ECHO 'Design: Deterministic audit of duplicate rows, missing cells, categorical
domains, numeric ranges, and final retained
rows.'.
1446 0 M> ECHO 'Design: Deterministic audit of duplicate rows, missing cells
, categorical domains, numeric ranges, and final retained
rows.'.
Design: Deterministic audit of duplicate rows, missing cells, categorical domains,
numeric ranges, and final retained rows.
1447 0 ECHO 'Inference scope: No hypothesis test is required; data-quality rules
determine whether rows or cells require correction.'.
Inference scope: No hypothesis test is required; data-quality rules determine
whether rows or cells require correction.
```

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

1447 0 M> ECHO 'Inference scope: No hypothesis test is required; data-quality rules determine whether rows or cells require correction.'

Inference scope: No hypothesis test is required; data-quality rules determine whether rows or cells require correction.

1448 0 ECHO 'Formula: duplicate_i=I(row_i appeared earlier); range violation=I(x<min or x>max); clean rows=distinct valid rows.'

1448 0 M> ECHO 'Formula: duplicate_i=I(row_i appeared earlier); range violation=I(x<min or x>max); clean rows=distinct valid rows.'

Formula: duplicate_i=I(row_i appeared earlier); range violation=I(x<min or x>max); clean rows=distinct valid rows.

1449 0 ECHO 'Variables: school, sex, age, address, famsize, Pstatus, Medu, Fedu, Mjob, Fjob, reason, guardian, traveltime, studytime, failures, schoolsup, famsup, paid, activities, nursery, higher, internet, romantic, famrel, freetime, goout, Dalc, Walc, health, absences, G1, G2, G3'.

1449 0 M> ECHO 'Variables: school, sex, age, address, famsize, Pstatus, Medu, Fedu, Mjob, Fjob, reason, guardian, traveltime, studytime, failures, schoolsup, famsup, paid, activities, nursery, higher, internet, romantic, famrel, freetime, goout, Dalc, Walc, health, absences, G1, G2, G3'.

Variables: school, sex, age, address, famsize, Pstatus, Medu, Fedu, Mjob, Fjob, reason, guardian, traveltime, studytime, failures, s

1450 0 ECHO 'Rows analyzed: 649'.

1450 0 M> ECHO 'Rows analyzed: 649'.

Rows analyzed: 649

1451 0 ECHO 'category_violations = 0'.

1451 0 M> ECHO 'category_violations = 0'.

category_violations = 0

1452 0 ECHO 'clean_rows = 649'.

1452 0 M> ECHO 'clean_rows = 649'.

clean_rows = 649

1453 0 ECHO 'duplicate_rows = 0'.

1453 0 M> ECHO 'duplicate_rows = 0'.

duplicate_rows = 0

1454 0 ECHO 'input_rows = 649'.

1454 0 M> ECHO 'input_rows = 649'.

input_rows = 649

Survey Data Cleaning
Duplicate missing domain and numeric-range validation

```
1455 0 ECHO 'missing_cells = 0'.
1455 0 M> ECHO 'missing_cells = 0'.
missing_cells = 0
1456 0 ECHO 'range_violations = 0'.
1456 0 M> ECHO 'range_violations = 0'.
range_violations = 0
1457 0 ECHO 'rows_removed = 0'.
1457 0 M> ECHO 'rows_removed = 0'.
rows_removed = 0
1458 0 ECHO 'Cross-check status: pass'.
1458 0 M> ECHO 'Cross-check status: pass'.
Cross-check status: pass
1459 0 ECHO '=====
====='.
1459 0 M> ECHO '=====
====='.
=====
1460 0
1460 0 M>
1461 0 OUTPUT SAVE OUTFILE='D:\DATA ANALYSIS\M Survey and Likert Scale Analy
sis\Survey Data Cleaning\SPSS_Output\spv\survey-data-c
leaning_spss_output.spv'.
1461 0 M> OUTPUT SAVE OUTFILE='D:\DATA ANALYSIS\M Survey and Likert Scale An
alysis\Survey Data Cleaning\SPSS_Output\spv\survey-dat
a-cleaning_spss_output.spv'.
1462 0 OUTPUT EXPORT /CONTENTS EXPORT=ALL LAYERS=PRINTSETTING MODELVIEWS=PRI
NTSETTING
1462 0 M> OUTPUT EXPORT /CONTENTS EXPORT=ALL LAYERS=PRINTSETTING MODELVIEWS=
PRINTSETTING
```