Data Quality Report: Comparison of ship and ship

No time restriction. All available data were analysed

(c) Universitätsklinikum Erlangen

10. März 2022

Inhaltsverzeichnis

1 Data Map

1.1 Target Data System

There is no data to display.

1.2 Source Data System

There is no data to display.

2 Completeness Checks

2.1 Validation

Completeness checks (validation) evaluate the ETL (extract transform load) jobs. They compare for each variable the exact matching of the number of distinct values, the number of valid values (=n), and the number of missing values between the source data system and the target data system.

Variable	Check Distincts	Check Valids	Check Missings
age	passed	passed	passed
cholesterol	passed	passed	passed
contraception	passed	passed	passed
dbp1	passed	passed	passed
dbp2	passed	passed	passed
dev_bp	passed	passed	passed
dev_length	passed	passed	passed
dev_weight	passed	passed	passed
diab_age	passed	passed	passed
diab_known	passed	passed	passed
exdate	passed	passed	passed
family	passed	passed	passed
hdl	passed	passed	passed
height	passed	passed	passed
id	passed	passed	passed
income	passed	passed	passed
ldl	passed	passed	passed
myocard	passed	passed	passed
obs_bp	passed	passed	passed
obs_int	passed	passed	passed
obs_soma	passed	passed	passed
sbp1	passed	passed	passed
sbp2	passed	passed	passed
school	passed	passed	passed
sex	passed	passed	passed
smoking	passed	passed	passed
stroke	passed	passed	passed
waist	passed	passed	passed
weight	passed	passed	passed

2.2 Verification

Variable	Missings [%] (source)	Missings [%] (target)
age	0	0
cholesterol	0.7	0.7
contraception	0	0
dbp1	0.09	0.09
dbp2	0.28	0.28
dev_bp	0	0
dev_length	0	0
dev_weight	0	0
diab_age	0	0
diab_known	0	0
exdate	0	0
family	0	0
hdl	0.74	0.74
height	0.14	0.14
id	0	0
income	0	0
ldl	1.3	1.3
myocard	0	0
obs_bp	0	0
obs_int	0	0
obs_soma	0	0
sbp1	0.09	0.09
sbp2	0.28	0.28
school	0	0
sex	0	0
smoking	0	0
stroke	0	0
waist	0.28	0.28
weight	0.19	0.19

3 Conformance Checks

3.1 Value Conformance

Value conformance checks (verification) compare for each variable the values of each data system to predefined constraints. Those constraints can be defined for each variable and data system individually in the metadata repository (MDR).

Variable	Check Source Data	Check Target Data
age	failed	failed
cholesterol	passed	passed
contraception	failed	failed
dbp1	failed	failed
dbp2	passed	passed
dev_bp	failed	failed
dev_length	failed	failed
dev_weight	failed	failed
diab_known	failed	failed
exdate	passed	passed
family	failed	failed
hdl	passed	passed
height	passed	passed
id	passed	passed
income	failed	failed
ldl	passed	passed
myocard	failed	failed
obs_bp	failed	failed
obs_int	failed	failed
obs_soma	failed	failed
sbp1	failed	failed
sbp2	failed	failed
school	failed	failed
sex	passed	passed
smoking	failed	failed
stroke	failed	failed
waist	passed	passed
weight	passed	passed
a_present_and_b_vv	failed	failed
a_present_and_b_levels_vl	failed	failed

4 Detailed Descriptive Analysis

4.1 age

Description for dataelement 'age'

4.1.1 Representation in source data system

- Variable: age
- Table: ship_data.csv

Overview:

- Variable name: age
- Variable type: integer

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 64

Results:

Mean	49.86
Minimum	19
Median	50
Maximum	82
SD	16.19
Negativ	0
Zero	0
Positive	2154
OutLo	0
OutHi	0
Variance	261.95
Range	63

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
20	Inf

- Extrem values are not conform with constraints.

4.1.2 Representation in target data system

- Variable: age
- Table: ship_data.csv

Overview:

- Variable name: age
- Variable type: integer
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 64

Results:

Mean	49.86
Minimum	19
Median	50
Maximum	82
SD	16.19
Negativ	0
Zero	0
Positive	2154
OutLo	0
OutHi	0
Variance	261.95
Range	63

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
20	Inf

- Extrem values are not conform with constraints.

4.2   cholesterol

Description for dataelement 'cholesterol'

4.2.1   Representation in source data system

- Variable: cholesterol
- Table: ship_data.csv

Overview:

- Variable name: cholesterol
- Variable type: float

    - n: 2154
    - Valid values: 2139
    - Missing values: 15
    - Distinct values: 1668

Results:

Mean	5.76
Minimum	2.67
Median	5.68
Maximum	12.12
SD	1.2
Negativ	0
Zero	0
Positive	2139
OutLo	0
OutHi	23
Variance	1.43
Range	9.45

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.2.2 Representation in target data system

- Variable: cholesterol
- Table: ship_data.csv

Overview:

- Variable name: cholesterol
- Variable type: float
    - n: 2154
    - Valid values: 2139
    - Missing values: 15
    - Distinct values: 1668

Results:

Mean	5.76
Minimum	2.67
Median	5.68
Maximum	12.12
SD	1.2
Negativ	0
Zero	0
Positive	2139
OutLo	0
OutHi	23
Variance	1.43
Range	9.45

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.3   contraception

Description for dataelement 'contraception':

1 = yes; 2 = no

4.3.1   Representation in source data system

- Variable: contraception
- Table: ship_data.csv

Overview:

- Variable name: contraception
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 4

Results:

contraception	Freq	% Valid
99801	1264	58.682
1	584	27.112
2	302	14.020
99900	4	0.186

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99801
  99900

4.3.2 Representation in target data system

- Variable: contraception
- Table: ship_data.csv

Overview:

- Variable name: contraception
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 4

Results:

contraception	Freq	% Valid
99801	1264	58.682
1	584	27.112
2	302	14.020
99900	4	0.186

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99801
  99900

4.4 dbp1

Description for dataelement 'dbp1'

4.4.1 Representation in source data system

- Variable: dbp1
- Table: ship_data.csv

Overview:

- Variable name: dbp1
- Variable type: integer

    - n: 2154
    - Valid values: 2152
    - Missing values: 2
    - Distinct values: 75

Results:

Mean	84.52
Minimum	53
Median	84
Maximum	198
SD	11.81
Negativ	0
Zero	0
Positive	2152
OutLo	0
OutHi	17
Variance	139.52
Range	145

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
40	160

- Extrem values are not conform with constraints.

4.4.2 Representation in target data system

- Variable: dbp1
- Table: ship_data.csv

Overview:

- Variable name: dbp1
- Variable type: integer
    - n: 2154
    - Valid values: 2152
    - Missing values: 2
    - Distinct values: 75

Results:

Mean	84.52
Minimum	53
Median	84
Maximum	198
SD	11.81
Negativ	0
Zero	0
Positive	2152
OutLo	0
OutHi	17
Variance	139.52
Range	145

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
40	160

- Extrem values are not conform with constraints.

4.5   dbp2

Description for dataelement 'dbp2'

4.5.1   Representation in source data system

- Variable: dbp2
- Table: ship_data.csv

Overview:

- Variable name: dbp2
- Variable type: integer

  - n: 2154
  - Valid values: 2148
  - Missing values: 6
  - Distinct values: 71

Results:

Mean	83.52
Minimum	55
Median	83
Maximum	151
SD	11.52
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	17
Variance	132.67
Range	96

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
40	160

4.5.2 Representation in target data system

- Variable: dbp2
- Table: ship_data.csv

Overview:

- Variable name: dbp2
- Variable type: integer

  - n: 2154
  - Valid values: 2148
  - Missing values: 6
  - Distinct values: 71

Results:

Mean	83.52
Minimum	55
Median	83
Maximum	151
SD	11.52
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	17
Variance	132.67
Range	96

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
40	160

4.6 dev_bp

Description for dataelement 'dev_bp':

1 = Dev_01; 2 = Dev_02; 3 = Dev_03; 4 = Dev_04; 5 = Dev_05; 6 = Dev_06; 7 = Dev_07; 8 = Dev_08; 9 = Dev_09; 10 = Dev_10; 11 = Dev_11; 12 = Dev_12; 13 = Dev_13; 14 = Dev_14; 15 = Dev_15; 16 = Dev_16; 17 = Dev_17; 18 = Dev_18; 19 = Dev_19; 20 = Dev_20; 21 = Dev_21; 22 = Dev_22; 23 = Dev_23; 24 = Dev_24; 25 = Dev_25

4.6.1 Representation in source data system

- Variable: dev_bp
- Table: ship_data.csv

Overview:

- Variable name: dev_bp
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 12

Results:

dev_bp	Freq	% Valid
10	444	20.613
9	381	17.688
15	265	12.303
18	257	11.931
16	235	10.910
20	214	9.935
17	156	7.242
22	109	5.060
19	88	4.085
99902	3	0.139
7	1	0.046
14	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.6.2 Representation in target data system

- Variable: dev_bp
- Table: ship_data.csv

Overview:

- Variable name: dev_bp
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 12

Results:

dev_bp	Freq	% Valid
10	444	20.613
9	381	17.688
15	265	12.303
18	257	11.931
16	235	10.910
20	214	9.935
17	156	7.242
22	109	5.060
19	88	4.085
99902	3	0.139
7	1	0.046
14	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.7 dev_length

Description for dataelement 'dev_length':

1 = Dev_01; 2 = Dev_02; 3 = Dev_03; 4 = Dev_04; 5 = Dev_05; 6 = Dev_06; 7 = Dev_07; 8 = Dev_08; 9 = Dev_09; 10 = Dev_10; 11 = Dev_11; 12 = Dev_12; 13 = Dev_13; 14 = Dev_14; 15 = Dev_15; 16 = Dev_16; 17 = Dev_17; 18 = Dev_18; 19 = Dev_19; 20 = Dev_20; 21 = Dev_21; 22 = Dev_22; 23 = Dev_23; 24 = Dev_24; 25 = Dev_25

4.7.1 Representation in source data system

- Variable: dev_length
- Table: ship_data.csv

Overview:

- Variable name: dev_length
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 4

Results:

dev_length	Freq	% Valid
3	1088	50.511
11	904	41.968
4	160	7.428
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.7.2 Representation in target data system

- Variable: dev_length
- Table: ship_data.csv

Overview:

- Variable name: dev_length
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 4

Results:

dev_length	Freq	% Valid
3	1088	50.511
11	904	41.968
4	160	7.428
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.8 dev_weight

Description for dataelement 'dev_weight':

1 = Dev_01; 2 = Dev_02; 3 = Dev_03; 4 = Dev_04; 5 = Dev_05; 6 = Dev_06; 7 = Dev_07; 8 = Dev_08; 9 = Dev_09; 10 = Dev_10; 11 = Dev_11; 12 = Dev_12; 13 = Dev_13; 14 = Dev_14; 15 = Dev_15; 16 = Dev_16; 17 = Dev_17; 18 = Dev_18; 19 = Dev_19; 20 = Dev_20; 21 = Dev_21; 22 = Dev_22; 23 = Dev_23; 24 = Dev_24; 25 = Dev_25

4.8.1 Representation in source data system

- Variable: dev_weight
- Table: ship_data.csv

Overview:

- Variable name: dev_weight
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 4

Results:

dev_weight	Freq	% Valid
1	1226	56.917
11	904	41.968
2	22	1.021
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.8.2 Representation in target data system

- Variable: dev_weight
- Table: ship_data.csv

Overview:

- Variable name: dev_weight
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 4

Results:

dev_weight	Freq	% Valid
1	1226	56.917
11	904	41.968
2	22	1.021
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99902

4.9 diab_age

Description for dataelement 'diab_age'

4.9.1 Representation in source data system

- Variable: diab_age
- Table: ship_data.csv

Overview:

- Variable name: diab_age
- Variable type: integer
  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 49

Results:

Mean	91790.05
Minimum	12
Median	99801
Maximum	99900
SD	27115.85
Negativ	0
Zero	0
Positive	2154
OutLo	173
OutHi	7
Variance	735269205.07
Range	99888

4.9.2 Representation in target data system

- Variable: diab_age
- Table: ship_data.csv

Overview:

- Variable name: diab_age
- Variable type: integer
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 49

Results:

Mean	91790.05
Minimum	12
Median	99801
Maximum	99900
SD	27115.85
Negativ	0
Zero	0
Positive	2154
OutLo	173
OutHi	7
Variance	735269205.07
Range	99888

4.10 diab_known

Description for dataelement 'diab_known':

0 = no; 1 = yes

4.10.1 Representation in source data system

- Variable: diab_known
- Table: ship_data.csv

Overview:

- Variable name: diab_known
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 3

Results:

diab_known	Freq	% Valid
0	1946	90.344
1	201	9.331
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1 '
- Levels that are not conform with the value set:
  99900

4.10.2 Representation in target data system

- Variable: diab_known
- Table: ship_data.csv

Overview:

- Variable name: diab_known
- Variable type: enumerated
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 3

Results:

diab_known	Freq	% Valid
0	1946	90.344
1	201	9.331
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1 '
- Levels that are not conform with the value set:
  99900

4.11 exdate

Description for dataelement 'exdate'

4.11.1 Representation in source data system

- Variable: exdate
- Table: ship_data.csv

Overview:

- Variable name: exdate
- Variable type: datetime

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 688

Results:

Min.	1997-10-16
1st Qu.	1998-08-18
Median	1999-03-26
Mean	1999-04-07
3rd Qu.	1999-11-04
Max.	2001-05-19

Value conformance:

- Conformance check: passed
- Constraining values/rules: 'No future dates allowed.'

4.11.2   Representation in target data system

- Variable: exdate
- Table: ship_data.csv

Overview:

- Variable name: exdate
- Variable type: datetime
  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 688

Results:

Min.	1997-10-16
1st Qu.	1998-08-18
Median	1999-03-26
Mean	1999-04-07
3rd Qu.	1999-11-04
Max.	2001-05-19

Value conformance:

- Conformance check: passed
- Constraining values/rules: 'No future dates allowed.'

4.12   family

Description for dataelement 'family':

1 = married; 2 = married (living apart); 3 = single (never married); 4 = divorced; 5 = widowed

4.12.1   Representation in source data system

- Variable: family
- Table: ship_data.csv

Overview:

- Variable name: family
- Variable type: enumerated
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 7

Results:

family	Freq	% Valid
1	1352	62.767
3	394	18.292
5	161	7.474
4	147	6.825
99914	60	2.786
2	33	1.532
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5 '
- Levels that are not conform with the value set:
  99900
  99914

4.12.2   Representation in target data system

- Variable: family
- Table: ship_data.csv

Overview:

- Variable name: family
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 7

Results:

family	Freq	% Valid
1	1352	62.767
3	394	18.292
5	161	7.474
4	147	6.825
99914	60	2.786
2	33	1.532
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5 '
- Levels that are not conform with the value set:
  99900
  99914

4.13 hdl

Description for dataelement 'hdl'

4.13.1 Representation in source data system

- Variable: hdl
- Table: ship_data.csv

Overview:

- Variable name: hdl
- Variable type: float

  - n: 2154
  - Valid values: 2138
  - Missing values: 16
  - Distinct values: 1111

Results:

Mean	1.45
Minimum	0.42
Median	1.39
Maximum	7.2
SD	0.44
Negativ	0
Zero	0
Positive	2138
OutLo	0
OutHi	33
Variance	0.19
Range	6.78

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.13.2 Representation in target data system

- Variable: hdl
- Table: ship_data.csv

Overview:

- Variable name: hdl
- Variable type: float

    - n: 2154
    - Valid values: 2138
    - Missing values: 16
    - Distinct values: 1111

Results:

Mean	1.45
Minimum	0.42
Median	1.39
Maximum	7.2
SD	0.44
Negativ	0
Zero	0
Positive	2138
OutLo	0
OutHi	33
Variance	0.19
Range	6.78

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.14   height

Description for dataelement 'height'

4.14.1   Representation in source data system

- Variable: height
- Table: ship_data.csv

Overview:

- Variable name: height
- Variable type: integer

    - n: 2154
    - Valid values: 2151
    - Missing values: 3
    - Distinct values: 52

Results:

Mean	168.22
Minimum	144
Median	168
Maximum	198
SD	9.25
Negativ	0
Zero	0
Positive	2151
OutLo	0
OutHi	1
Variance	85.48
Range	54

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
80	230

4.14.2 Representation in target data system

- Variable: height
- Table: ship_data.csv

Overview:

- Variable name: height
- Variable type: integer
    - n: 2154
    - Valid values: 2151
    - Missing values: 3
    - Distinct values: 52

Results:

Mean	168.22
Minimum	144
Median	168
Maximum	198
SD	9.25
Negativ	0
Zero	0
Positive	2151
OutLo	0
OutHi	1
Variance	85.48
Range	54

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
80	230

4.15 id

Description for dataelement 'id'

4.15.1 Representation in source data system

- Variable: id
- Table: ship_data.csv

Overview:

- Variable name: id
- Variable type: string
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 2154

Results:

id	Freq	% Valid
3861	1	0.046
6506	1	0.046
6096	1	0.046
6674	1	0.046
6490	1	0.046
5366	1	0.046
5735	1	0.046
4031	1	0.046
3578	1	0.046
4807	1	0.046
5450	1	0.046
4763	1	0.046
4146	1	0.046
7096	1	0.046
3726	1	0.046
5459	1	0.046
5161	1	0.046
5562	1	0.046
4111	1	0.046
6392	1	0.046
5553	1	0.046
5404	1	0.046
3673	1	0.046
6808	1	0.046
5760	1	0.046

Value conformance:

- Conformance check: passed
- Constraining values/rules: ' [1]{1,5}$ '

---

[1][:digit:]

4.15.2 Representation in target data system

- Variable: id
- Table: ship_data.csv

Overview:

- Variable name: id
- Variable type: string
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 2154

Results:

id	Freq	% Valid
3861	1	0.046
6506	1	0.046
6096	1	0.046
6674	1	0.046
6490	1	0.046
5366	1	0.046
5735	1	0.046
4031	1	0.046
3578	1	0.046
4807	1	0.046
5450	1	0.046
4763	1	0.046
4146	1	0.046
7096	1	0.046
3726	1	0.046
5459	1	0.046
5161	1	0.046
5562	1	0.046
4111	1	0.046
6392	1	0.046
5553	1	0.046
5404	1	0.046
3673	1	0.046
6808	1	0.046
5760	1	0.046

Value conformance:

- Conformance check: passed
- Constraining values/rules: ' [2]{1,5}$ '

---

[2][:digit:]

4.16 income

Description for dataelement 'income':

1 = < 2000; 2 = 2000 ï¿½ 4000; 3 = > 4000

4.16.1 Representation in source data system

- Variable: income
- Table: ship_data.csv

Overview:

- Variable name: income
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 5

Results:

income	Freq	% Valid
2	992	46.054
1	602	27.948
3	444	20.613
99901	100	4.643
99900	16	0.743

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3 '
- Levels that are not conform with the value set:
  99900
  99901

4.16.2 Representation in target data system

- Variable: income
- Table: ship_data.csv

Overview:

- Variable name: income
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 5

Results:

income	Freq	% Valid
2	992	46.054
1	602	27.948
3	444	20.613
99901	100	4.643
99900	16	0.743

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3 '
- Levels that are not conform with the value set:
  99900
  99901

4.17 ldl

Description for dataelement 'ldl'

4.17.1 Representation in source data system

- Variable: ldl
- Table: ship_data.csv

Overview:

- Variable name: ldl
- Variable type: float

  - n: 2154
  - Valid values: 2126
  - Missing values: 28
  - Distinct values: 2125

Results:

Mean	3.58
Minimum	0.7
Median	3.52
Maximum	9.24
SD	1.13
Negativ	0
Zero	0
Positive	2126
OutLo	0
OutHi	21
Variance	1.28
Range	8.54

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.17.2 Representation in target data system

- Variable: ldl
- Table: ship_data.csv

Overview:

- Variable name: ldl
- Variable type: float
    - n: 2154
    - Valid values: 2126
    - Missing values: 28
    - Distinct values: 2125

Results:

Mean	3.58
Minimum	0.7
Median	3.52
Maximum	9.24
SD	1.13
Negativ	0
Zero	0
Positive	2126
OutLo	0
OutHi	21
Variance	1.28
Range	8.54

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
0	Inf

4.18 myocard

Description for dataelement 'myocard':

1 = yes; 2 = no

4.18.1 Representation in source data system

- Variable: myocard
- Table: ship_data.csv

Overview:

- Variable name: myocard
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 5

Results:

myocard	Freq	% Valid
2	2015	93.547
1	65	3.018
99914	60	2.786
99901	7	0.325
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99900
  99901
  99914

4.18.2 Representation in target data system

- Variable: myocard
- Table: ship_data.csv

Overview:

- Variable name: myocard
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 5

Results:

myocard	Freq	% Valid
2	2015	93.547
1	65	3.018
99914	60	2.786
99901	7	0.325
99900	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99900
  99901
  99914

4.19 obs_bp

Description for dataelement 'obs_bp':

1 = Obs_01; 2 = Obs_02; 3 = Obs_03; 4 = Obs_04; 5 = Obs_05; 6 = Obs_06; 7 = Obs_07; 8 = Obs_08; 9 = Obs_09; 10 = Obs_10; 11 = Obs_11; 12 = Obs_12; 13 = Obs_13; 14 = Obs_14; 15 = Obs_15; 16 = Obs_16; 17 = Obs_17; 18 = Obs_18; 19 = Obs_19; 20 = Obs_20

4.19.1 Representation in source data system

- Variable: obs_bp
- Table: ship_data.csv

Overview:

- Variable name: obs_bp
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 11

Results:

obs_bp	Freq	% Valid
4	763	35.422
7	477	22.145
9	384	17.827
3	330	15.320
5	149	6.917
8	41	1.903
1	4	0.186
99902	3	0.139
11	1	0.046
18	1	0.046
16	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 '
- Levels that are not conform with the value set:
  99902

4.19.2   Representation in target data system

- Variable: obs_bp
- Table: ship_data.csv

Overview:

- Variable name: obs_bp
- Variable type: enumerated

  – n: 2154
  – Valid values: 2154
  – Missing values: 0
  – Distinct values: 11

Results:

obs_bp	Freq	% Valid
4	763	35.422
7	477	22.145
9	384	17.827
3	330	15.320
5	149	6.917
8	41	1.903
1	4	0.186
99902	3	0.139
11	1	0.046
18	1	0.046
16	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 '
- Levels that are not conform with the value set:
  99902

4.20 obs_int

Description for dataelement 'obs_int':

1 = Obs_01; 2 = Obs_02; 3 = Obs_03; 4 = Obs_04; 5 = Obs_05; 6 = Obs_06; 7 = Obs_07; 8 = Obs_08; 9 = Obs_09; 10 = Obs_10; 11 = Obs_11; 12 = Obs_12; 13 = Obs_13; 14 = Obs_14; 15 = Obs_15; 16 = Obs_16; 17 = Obs_17; 18 = Obs_18; 19 = Obs_19; 20 = Obs_20 ; 21 = Obs_21; 22 = Obs_22; 23 = Obs_23; 24 = Obs_24; 25 = Obs_25

4.20.1 Representation in source data system

- Variable: obs_int
- Table: ship_data.csv

Overview:

- Variable name: obs_int
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 14

Results:

obs_int	Freq	% Valid
2	600	27.855
1	396	18.384
11	349	16.202
22	277	12.860
3	235	10.910
12	122	5.664
10	78	3.621
23	67	3.110
4	11	0.511
5	11	0.511
7	5	0.232
99900	1	0.046
25	1	0.046
13	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99900

4.20.2 Representation in target data system

- Variable: obs_int
- Table: ship_data.csv

Overview:

- Variable name: obs_int
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 14

Results:

obs_int	Freq	% Valid
2	600	27.855
1	396	18.384
11	349	16.202
22	277	12.860
3	235	10.910
12	122	5.664
10	78	3.621
23	67	3.110
4	11	0.511
5	11	0.511
7	5	0.232
99900	1	0.046
25	1	0.046
13	1	0.046

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 '
- Levels that are not conform with the value set:
  99900

4.21 obs_soma

Description for dataelement 'obs_soma':

1 = Obs_01; 2 = Obs_02; 3 = Obs_03; 4 = Obs_04; 5 = Obs_05; 6 = Obs_06; 7 = Obs_07; 8 = Obs_08; 9 = Obs_09; 10 = Obs_10

4.21.1 Representation in source data system

- Variable: obs_soma
- Table: ship_data.csv

Overview:

- Variable name: obs_soma
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 9

Results:

obs_soma	Freq	% Valid
4	645	29.944
7	483	22.423
9	388	18.013
5	320	14.856
3	280	12.999
8	31	1.439
1	3	0.139
2	2	0.093
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 '
- Levels that are not conform with the value set:
  99902

4.21.2 Representation in target data system

- Variable: obs_soma
- Table: ship_data.csv

Overview:

- Variable name: obs_soma
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 9

Results:

obs_soma	Freq	% Valid
4	645	29.944
7	483	22.423
9	388	18.013
5	320	14.856
3	280	12.999
8	31	1.439
1	3	0.139
2	2	0.093
99902	2	0.093

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 '
- Levels that are not conform with the value set:
  99902

4.22   sbp1

Description for dataelement 'sbp1'

4.22.1   Representation in source data system

- Variable: sbp1
- Table: ship_data.csv

Overview:

- Variable name: sbp1
- Variable type: integer

  - n: 2154
  - Valid values: 2152
  - Missing values: 2
  - Distinct values: 125

Results:

Mean	139.01
Minimum	83
Median	137
Maximum	253
SD	22.5
Negativ	0
Zero	0
Positive	2152
OutLo	0
OutHi	23
Variance	506.09
Range	170

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
80	200

- Extrem values are not conform with constraints.

4.22.2   Representation in target data system

- Variable: sbp1
- Table: ship_data.csv

Overview:

- Variable name: sbp1
- Variable type: integer
    - n: 2154
    - Valid values: 2152
    - Missing values: 2
    - Distinct values: 125

Results:

Mean	139.01
Minimum	83
Median	137
Maximum	253
SD	22.5
Negativ	0
Zero	0
Positive	2152
OutLo	0
OutHi	23
Variance	506.09
Range	170

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
80	200

- Extrem values are not conform with constraints.

4.23   sbp2

Description for dataelement 'sbp2'

4.23.1   Representation in source data system

- Variable: sbp2
- Table: ship_data.csv

Overview:

- Variable name: sbp2
- Variable type: integer

    - n: 2154
    - Valid values: 2148
    - Missing values: 6
    - Distinct values: 123

Results:

Mean	136.4
Minimum	83
Median	134
Maximum	258
SD	21.86
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	24
Variance	477.76
Range	175

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
80	200

- Extrem values are not conform with constraints.

4.23.2   Representation in target data system

- Variable: sbp2
- Table: ship_data.csv

Overview:

- Variable name: sbp2
- Variable type: integer
    - n: 2154
    - Valid values: 2148
    - Missing values: 6
    - Distinct values: 123

Results:

Mean	136.4
Minimum	83
Median	134
Maximum	258
SD	21.86
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	24
Variance	477.76
Range	175

Value conformance:

- Conformance check: failed
- Constraining values/rules:

min	max
80	200

- Extrem values are not conform with constraints.

4.24 school

Description for dataelement 'school':

0 = none; 1 = lower secondary; 2 = secondary; 3 = upper secondary

4.24.1 Representation in source data system

- Variable: school
- Table: ship_data.csv

Overview:

- Variable name: school
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 7

Results:

school	Freq	% Valid
2	981	45.543
0	779	36.165
1	256	11.885
99914	60	2.786
99900	53	2.461
3	18	0.836
4	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1, 2, 3 '
- Levels that are not conform with the value set:
  4
  99900
  99914

4.24.2   Representation in target data system

- Variable: school
- Table: ship_data.csv

Overview:

- Variable name: school
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 7

Results:

school	Freq	% Valid
2	981	45.543
0	779	36.165
1	256	11.885
99914	60	2.786
99900	53	2.461
3	18	0.836
4	7	0.325

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1, 2, 3 '
- Levels that are not conform with the value set:
  4
  99900
  99914

4.25   sex

Description for dataelement 'sex':

1 = males; 2 = females

4.25.1   Representation in source data system

- Variable: sex
- Table: ship_data.csv

Overview:

- Variable name: sex
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 2

Results:

sex	Freq	% Valid
2	1167	54.178
1	987	45.822

Value conformance:

- Conformance check: passed
- Constraining values/rules: ' 1, 2 '

4.25.2   Representation in target data system

- Variable: sex
- Table: ship_data.csv

Overview:

- Variable name: sex
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 2

Results:

sex	Freq	% Valid
2	1167	54.178
1	987	45.822

Value conformance:

- Conformance check: passed
- Constraining values/rules: ' 1, 2 '

4.26 smoking

Description for dataelement 'smoking':

0 = nonsmoker; 1 = former smoker; 2 = smoker

4.26.1 Representation in source data system

- Variable: smoking
- Table: ship_data.csv

Overview:

- Variable name: smoking
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 5

Results:

smoking	Freq	% Valid
0	733	34.030
1	697	32.358
2	656	30.455
99914	60	2.786
99900	8	0.371

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1, 2 '
- Levels that are not conform with the value set:
  99900
  99914

4.26.2   Representation in target data system

- Variable: smoking
- Table: ship_data.csv

Overview:

- Variable name: smoking
- Variable type: enumerated
    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 5

Results:

smoking	Freq	% Valid
0	733	34.030
1	697	32.358
2	656	30.455
99914	60	2.786
99900	8	0.371

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 0, 1, 2 '
- Levels that are not conform with the value set:
  99900
  99914

4.27 stroke

Description for dataelement 'stroke':

1 = yes; 2 = no

4.27.1 Representation in source data system

- Variable: stroke
- Table: ship_data.csv

Overview:

- Variable name: stroke
- Variable type: enumerated

  - n: 2154
  - Valid values: 2154
  - Missing values: 0
  - Distinct values: 5

Results:

stroke	Freq	% Valid
2	2028	94.150
99914	60	2.786
1	56	2.600
99900	7	0.325
99901	3	0.139

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99900
  99901
  99914

4.27.2   Representation in target data system

- Variable: stroke
- Table: ship_data.csv

Overview:

- Variable name: stroke
- Variable type: enumerated

    - n: 2154
    - Valid values: 2154
    - Missing values: 0
    - Distinct values: 5

Results:

stroke	Freq	% Valid
2	2028	94.150
99914	60	2.786
1	56	2.600
99900	7	0.325
99901	3	0.139

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1, 2 '
- Levels that are not conform with the value set:
  99900
  99901
  99914

4.28 waist

Description for dataelement 'waist'

4.28.1 Representation in source data system

- Variable: waist
- Table: ship_data.csv

Overview:

- Variable name: waist
- Variable type: float

  - n: 2154
  - Valid values: 2148
  - Missing values: 6
  - Distinct values: 2102

Results:

Mean	89.21
Minimum	49.27
Median	89.52
Maximum	145.77
SD	13.82
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	6
Variance	191.09
Range	96.5

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
30	Inf

4.28.2   Representation in target data system

- Variable: waist
- Table: ship_data.csv

Overview:

- Variable name: waist
- Variable type: float

    - n: 2154
    - Valid values: 2148
    - Missing values: 6
    - Distinct values: 2102

Results:

Mean	89.21
Minimum	49.27
Median	89.52
Maximum	145.77
SD	13.82
Negativ	0
Zero	0
Positive	2148
OutLo	0
OutHi	6
Variance	191.09
Range	96.5

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
30	Inf

4.29 weight

Description for dataelement 'weight'

4.29.1 Representation in source data system

- Variable: weight
- Table: ship_data.csv

Overview:

- Variable name: weight
- Variable type: float

    - n: 2154
    - Valid values: 2150
    - Missing values: 4
    - Distinct values: 1763

Results:

Mean	77.63
Minimum	42.6
Median	77.04
Maximum	144.44
SD	15.08
Negativ	0
Zero	0
Positive	2150
OutLo	0
OutHi	17
Variance	227.34
Range	101.84

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
30	250

4.29.2   Representation in target data system

- Variable: weight
- Table: ship_data.csv

Overview:

- Variable name: weight
- Variable type: float

  - n: 2154
  - Valid values: 2150
  - Missing values: 4
  - Distinct values: 1763

Results:

Mean	77.63
Minimum	42.6
Median	77.04
Maximum	144.44
SD	15.08
Negativ	0
Zero	0
Positive	2150
OutLo	0
OutHi	17
Variance	227.34
Range	101.84

Value conformance:

- Conformance check: passed
- Constraining values/rules:

min	max
30	250

5 Plausibility Checks

5.1 Atemporal Plausibility

5.1.1 A_present_and_B_vv

Contracept in males.

5.1.1.1 Representation in source data system

- Variable 1: sex
- Variable 2: contraception
- Filter criterion variable 2 (regex): ^1$
- Join criterion: id

Overview:

No data available for reporting

Results:

sex	Freq	% Valid
2	572	97.945
1	12	2.055

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1 '
- Levels that are not conform with the value set:
  2

5.1.1.2 Representation in target data system

- Variable 1: sex
- Variable 2: contraception
- Filter criterion variable 2 (regex): ^1$
- Join criterion: id

Overview:

No data available for reporting

Results:

sex	Freq	% Valid
2	572	97.945
1	12	2.055

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1 '
- Levels that are not conform with the value set:
  2

5.1.2 A__present__and__B__levels__vl

Diab age but no diab.

5.1.2.1 Representation in source data system

- Variable 1: diab_known
- Variable 2: diab_age
- Filter criterion variable 2 (regex): ^(?!.*(99900|99901|99801)).$
- Join criterion: id

Overview:

No data available for reporting

Results:

diab__known	Freq	% Valid
1	138	79.769
0	35	20.231

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1 '
- Levels that are not conform with the value set:
  0

5.1.2.2 Representation in target data system

- Variable 1: diab_known
- Variable 2: diab_age
- Filter criterion variable 2 (regex): ^(?!.*(99900|99901|99801)).$
- Join criterion: id

Overview:

No data available for reporting

Results:

diab__known	Freq	% Valid
1	138	79.769
0	35	20.231

Value conformance:

- Conformance check: failed
- Constraining values/rules: ' 1 '
- Levels that are not conform with the value set:
  0

5.2 Uniqueness Plausibility

5.2.1 sex

With each distinct value of 'id', only one value of 'sex' may be associated.

5.2.1.1 Representation in source data system

- Plausibility check: passed
- Message: No duplicate occurrences of id found in association with sex.

5.2.1.2 Representation in target data system

- Plausibility check: passed
- Message: No duplicate occurrences of id found in association with sex.

6 Appendix

6.1 R-Package Version 'DQAstats'

```
## [1] '0.2.6.9003'
```

6.2 R-Package Version 'DIZutils'

```
## [1] '0.0.15.9002'
```

6.3 R-Package Version 'DIZtools'

```
## [1] '0.0.4'

##
## ## SQL Statments

##
## ### Source Data System

##
## ### Target Data System
```