

1 Introduction

This document is the report for the data warehouse part of the lecture *DWO*. Before the *SSIS* and *SSAS* project can be used you need to ensure that the database with name *MDE_DWO_WS1617* exists on your local SQL Server instance and has proper permissions set for *SSAS* services user.

The source in the *SSIS* project needs to be modified to point to the proper *.mdb file.

2 Integration Services

This section deals with the *SSIS Integration Services* part of the hands on. The image 1 shows the implemented *Control Flow*. The container *Database Preparation Sequence* holds the components for drop create the warehouse tables which was useful during the development.

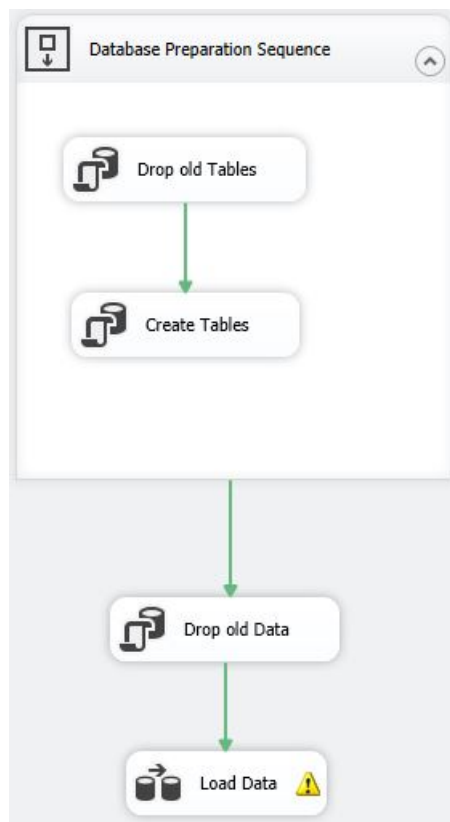


Abbildung 1: Integration Service Control Flow

The following listing describes the components in the control flow:

- *Drop old Tables* component drops the existing warehouse tables.
- *Create tables* component creates the warehouse tables via an native SQL statement.
- *Drop old Data* component drops the old data before data load.
- *Load Data* component holds the data flow for data cleanup and restructure.

The yellow warn icon is caused by data type changes between the raw data and the prepared data. the data type change is caused by the change of the length of some *nvarchar* typed columns.

The image shows the implemented *Data Flow* for the data preparation of the raw data of table *exam* and *student_group* which will be represented by the tables *exam* and *student* in the warehouse.

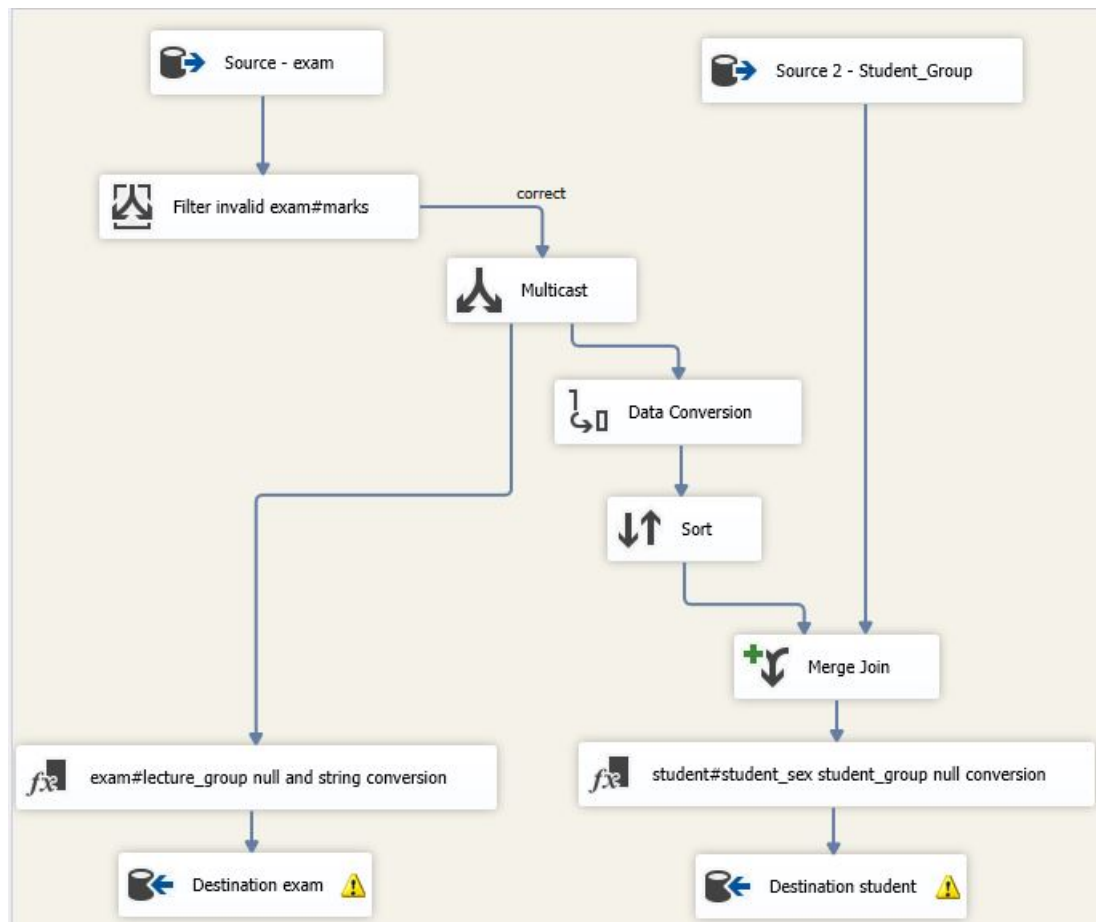


Abbildung 2: Integration Service Data Flow for *exam*

The exam input stream gets duplicated for the *student* stream because there are some students in the *exam* table which are not present in the *student_group* table, therefore the missing students need to be captured from the cleaned up *exam* stream. There are also some students in the *student_group* table, which have no exam entries, but they are considered to be worthless for the further analysis. The *student_code* column gets interpreted as a *DT_WSTR* and in the *student* stream as a *four-byte-signed-integer*. Therefore a cast was necessary to get the proper type for the *Merg Join* component.

The yellow warn icon is caused by some data type changes of some columns. For instance *student_sex* was changed from *nvarchar(2)* to *nchar(1)*.

The result are the two tables *exam* and *student* where the *student* table will only contain students which have *exam* entries. Students not present in the *student_group* table before, will get the *student_group* set to *U* which stands for unknown. The exam table was cleaned of the attribute *student_sex* which is part of the table *student*.

The image shows the implemented *Data Flow* for the data preparation of the raw data of table *lehrer_prod* and *teacher_group* which will be represented by the tables *teacher* in the warehouse.

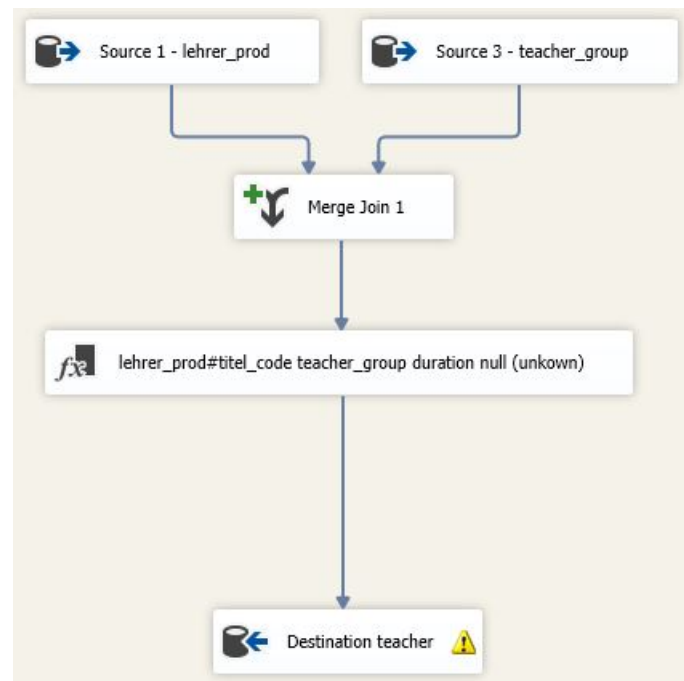


Abbildung 3: Integration Service Data Flow for *lehrer_prod* and *teacher_group*

The *teacher_group* and *lehrer_prod* tables get merged to the *teacher* table. If teachers are missing in the *teacher_group* table or the *typ* column in the *teacher_group* table is null, then they get assigned to the group *unknown*. The duration gets converted to a string of the form *Duration 1* or to *unknown* if no duration is present.

The yellow warn icon is caused by some data type changes of some columns. For instance *teacher_sex* was changed from *nvarchar(2)* to *nchar(1)*.

The result is the table *teacher* which holds all teachers present in the *exam* table with all their attributes.

3 Analysis Services

This section deals with the *SSAS Analysis Services* part of the hands on.

3.1 Implemented Project

This section deals with the implemented project for the analysis part of the hands on.

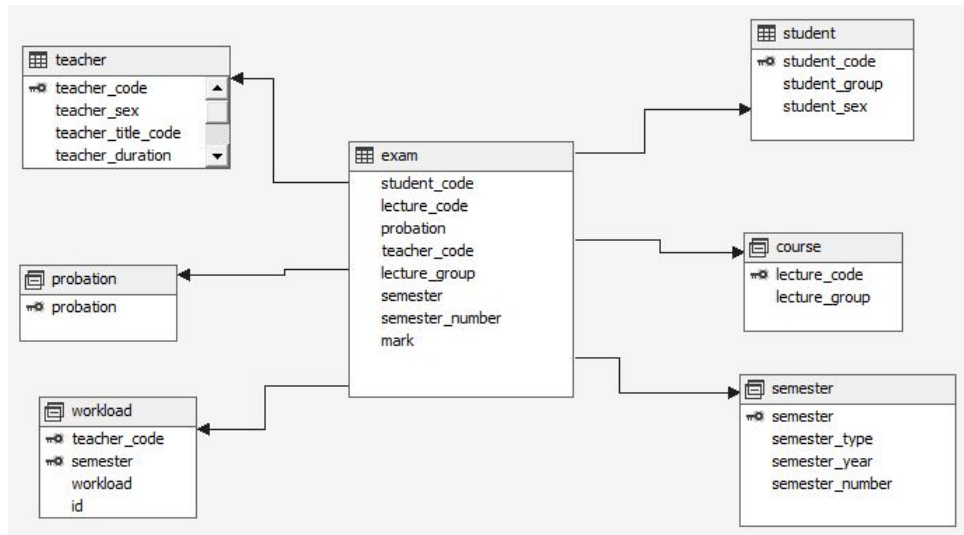


Abbildung 4: Data source view of the warehouse database

The tables *workload*, *probation*, *course* and *semester* were create as named queries.

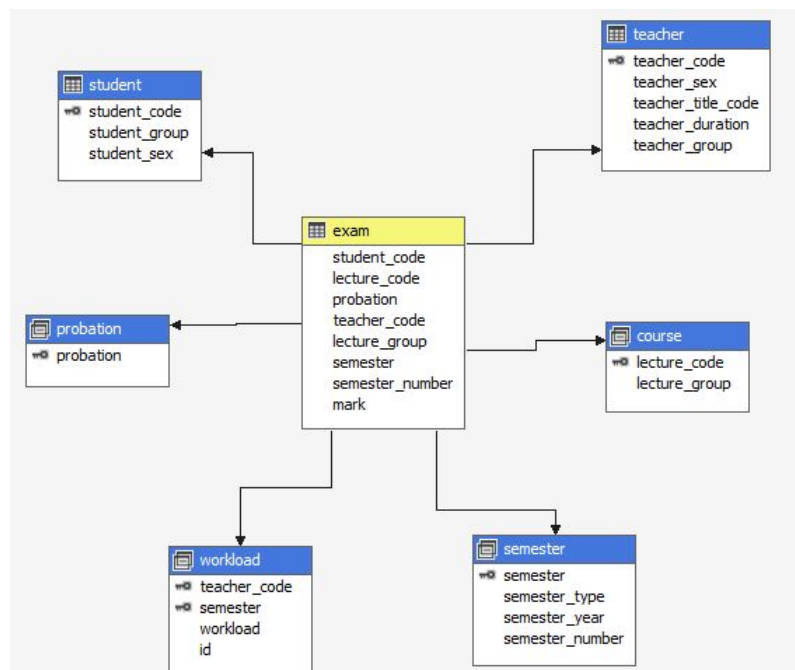


Abbildung 5: Implemented cube

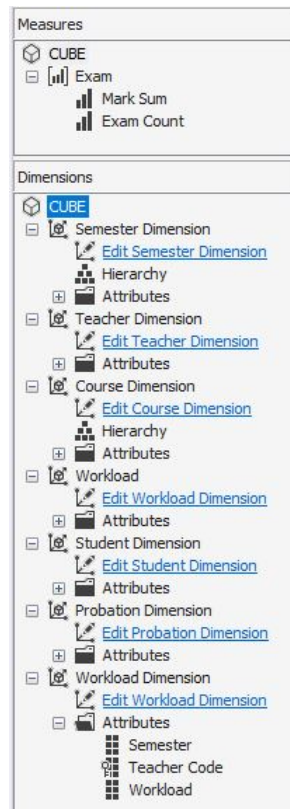


Abbildung 6: Cube measures and dimensions

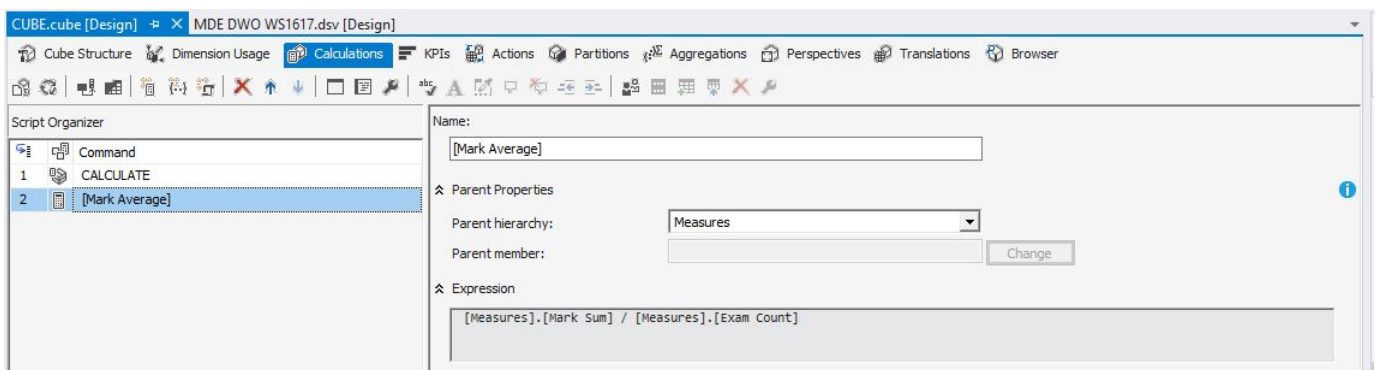


Abbildung 7: Cube measures and dimensions

3.2 Reports

The following images are showing the result of the made reports based on the implemented cube.

Zeilenbeschriftungen ▼	Mark Average
Duration 1	2.075107296
Duration 2	2.00390625
Duration 3	1.756819055
Duration 4	2.35611742
Duration 5	1.927128517
Duration 6	2.154867936
Duration 7	2.239262424
Duration 8	2.547920596
Gesamtergebnis	2.188585512

Abbildung 8: Report Teachers seniority

Zeilenbeschriftungen ▼	Mark Average
⊕ Group 0	1.913610356
⊕ Group 1	2.323254671
⊕ Group 10	1.980667838
⊕ Group 2	2.139311044
⊕ Group 3	2.500674521
⊕ Group 4	2.096781979
⊖ unknown	2.208333333
E1	2.320754717
E10	2.738461538
E11	2.056603774
E12	1.981132075
E13	1
E2	1.886792453
E3	2.754385965
E4	1.928571429
E5	3.283333333
E6	1
E7	3.283333333
E8	1.529411765
E9	2.490566038
Gesamtergebnis	2.188585512

Abbildung 9: Report Course type

Zeilenbeschriftungen ▼	Mark Average
004005007	2.516605166
004011016	1.896943852
005	2.55361596
006	2.273162939
006006	1.807017544
006007	2.12147651
006022019	2.311111111
008	1.946564885
010	1
011016	3.153846154
011016019	3
012	1.500745156
015	1.874166667
016	2.558139535
019	2.377622378
020006	1.666666667
020019022	2.078431373
022	2.068216816
022007	2.325
022019	2.347075671
022019013	2.444444444
022020	1.729166667
039	2.272256729
039019	3.203252033
041	2.445783133
042	2.153846154
042019	2.526466654
043	2.484705882
502	2.363867684
999	2.268914255
999011016	1.693877551
999013	1.204724409
999016	2.333333333
999022	1.969924812
unknown	1.65870098
Gesamtergebnis	2.188585512

Abbildung 10: Report Teacher job title

Zeilenbeschriftungen	Mark Average
M	2.18729535
W	2.212223472
Gesamtergebnis	2.188585512

Abbildung 11: Report Teacher sex

Zeilenbeschriftungen	Mark Average
1	2.142065268
2	3.341477603
3	3.810945274
Gesamtergebnis	2.188585512

Abbildung 12: Report Probation

Zeilenbeschriftungen	Mark Average
1	2
10	3.371794872
100	2.927710843
101	3.131147541
102	2.19089317
103	2.95
104	3.047619048
105	2.790948276

Abbildung 13: Report teacher code (truncated)

Zeilenbeschriftungen	Mark Average
SS	2.167948768
00	2.113297669
01	2.086235489
95	2.394039735
96	2.359360301
97	2.209196355
98	2.329795496
99	2.137848432
WS	2.208973073
00/01	2.087050263
01/02	1.354545455
94/95	2.689986283
95/96	2.399470899
96/97	2.269874477
97/98	2.272251309
98/99	2.324797844
99/00	2.100312256
Gesamtergebnis	2.188585512

Abbildung 14: Report over time (semesters)

Zeilenbeschriftungen	Mark Average
M	2.221655786
A	2.112563544
B	1.915895062
C	2.157303371
D	1.862068966
E	2.679012346
F	2.115384615
U	2.235813652
W	2.091409355
A	2.039215686
B	1.956321839
C	2.440677966
E	1.294117647
F	1.514285714
U	2.103614458
Gesamtergebnis	2.188585512

Abbildung 15: Report student and student group

Zeilenbeschriftungen	Mark Average
SS00	2.113297669
102 - 127	1.9566787
131 - 156	1.927209705
164 - 195	2.355029586
198 - 220	2.430952381
2 - 28	2
263 - 287	2.254681648
29 - 48	2.202398801
292 - 328	1.737201365
342 - 350	2.01754386
49 - 84	2.406494961
85 - 101	1.900516796
SS01	2.086235489
SS95	2.394039735
SS96	2.359360301
SS97	2.209196355
SS98	2.329795496
SS99	2.137848432
WS00/01	2.087050263
WS01/02	1.354545455
WS94/95	2.689986283
WS95/96	2.399470899
WS96/97	2.269874477
WS97/98	2.272251309
WS98/99	2.324797844
WS99/00	2.100312256
Gesamtergebnis	2.188585512

Abbildung 16: Report workload over semesters

Zeilenbeschriftungen	Mark Average
M	2.18729535
M	2.223328827
W	2.083775858
W	2.212223472
M	2.194090909
W	2.29764454
Gesamtergebnis	2.188585512

Abbildung 17: Report teacher and student sex