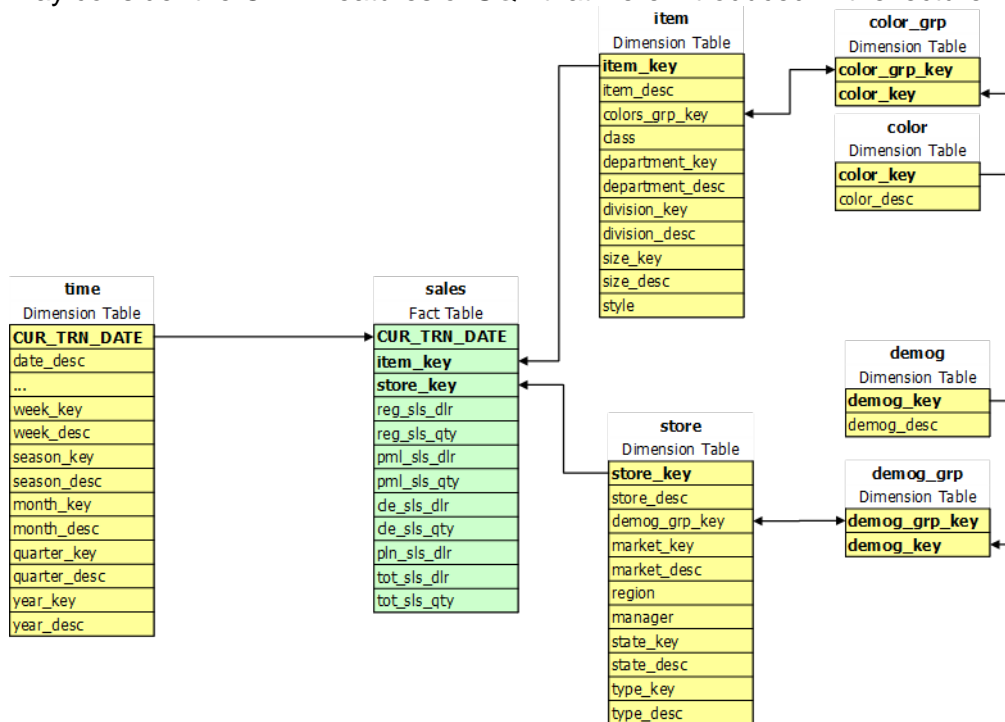


Assignment 1

Please follow the following rules when working on this assignment:

- This assignment is mandatory for all students.
- Work on this assignment in teams of three students. Each team has to prepare its own solution.
- Prepare one result document (PDF) with your solutions for all the tasks listed below.
 - You may write this document in English or German.
 - The PDF has to include the team number, name, study programme and matriculation number for each team member.
 - Document name: A1-Txx-Vy.pdf (where “xx” is the team number and “y” the version).
- Prepare one text file that includes all the queries you prepared for the tasks listed below.
 - This file should only include queries that are executable in ODPS and no further text or comments. Please conclude each query by a semicolon and add an empty line between two queries.
 - Document name: A1-Txx-Vy.txt (where “xx” is the team number and “y” the version).
- Submit the documents in Ilias no later than February 1, 10:00 am. This works as follows:
 - Go to ‘Submit Assignments’ and select Assignment 1
 - One team member has to select ‘Create Team’ and afterwards ‘Manage Team’ to add your team members. Choose ‘Add Users of Current Course’ to add your team members.
 - Note that you can only select individual students and not directly the teams/groups we already have organized in Ilias. Anyway, the teams you are creating to submit your document have to match one of the available teams.
 - Finally click ‘Hand in’ to submit the prepared files.
 - Please make sure that you submit only one document per team and that you add all team members before submitting your document.
- Contact Holger Schwarz for any further questions or use the forum in Ilias.

For this assignment, we use the well-known star schema of trendy4all (see Exercises). Most tables of this schema are available as schema **dw2** in ODPS. For the SQL statements, you may consider the OLAP features of SQL that were introduced in the lecture.



Task 1

Provide the turnover in regular sales (reg_sls_dlr) for the years 1999, 2000 and 2001 as well as the increase (in percent) for each year compared to the previous year. This information should be calculated for each combination of type_desc and department_desc.

Your result document must include:

- a screenshot of the complete result table
- the SQL statement or the sequence of SQL statements that you used to derive the result table (add all statements to the text file as well)
- a short explanation on the idea behind the queries

Task 2

An OLAP tool should provide for each manager a ranking of item classes for items that were sold in stores this manager is responsible for. The ranking should be based on the turnover of sold items of a certain item class (tot_sls_dlr). Provide the SQL statement the OLAP tool could generate to retrieve this requested information from the data warehouse in a single table. This table should include the manager, the item class, the total sales and the ranking position.

Your result document must include:

- the complete result table
- the SQL statement or the sequence of SQL statements that you used to derive the result table (add all statements to the text file as well)
- a short explanation on the idea behind the queries

Task 3

Provide the value of promotional sales (pml_sls_dlr) per day and division. For each day $d1$, the result table should also cover the average of this value for items that were sold in the same year on any day before and including day $d1$. The rows of the result table should look like this:

YEAR_DESC	DATE_DESC	DIVISION_DESC	SUM_PML_SLS_DLR	AVG_UP_TO_DATE
1999	1.12.1999 00:00:00	Athletics	71005.65	71005.65
1999	2.12.1999 00:00:00	Athletics	83763.28	77384.46
1999	3.12.1999 00:00:00	Athletics	61930.80	72233.24
1999	4.12.1999 00:00:00	Athletics	78788.09	73871.95
1999	5.12.1999 00:00:00	Athletics	47727.12	68642.98
1999	6.12.1999 00:00:00	Athletics	60124.71	67223.27
1999	7.12.1999 00:00:00	Athletics	38764.94	63157.79

1999	27.12.1999 00:00:00	Casual Clothing	117129.70	98291.15
1999	28.12.1999 00:00:00	Casual Clothing	83025.15	98624.05
1999	29.12.1999 00:00:00	Casual Clothing	107129.70	98917.35
1999	30.12.1999 00:00:00	Casual Clothing	60636.74	97641.33
1999	31.12.1999 00:00:00	Casual Clothing	97270.41	97629.36
2000	1.12.2000 00:00:00	Athletics	56248.30	56248.30
2000	2.12.2000 00:00:00	Athletics	66338.52	61293.41
2000	3.12.2000 00:00:00	Athletics	41555.95	54714.25

Your result document must include:

- the SQL statement or the sequence of SQL statements that you used to derive the result table (add all statements to the text file as well)
- a short explanation on the idea behind the queries

Task 4

Provide the aggregated number of sold items (reg_sls_qty) for each combination of the dimension attributes

- month_desc, state_desc and style,
- month_desc and state_desc,
- month_desc,
- as well as the overall sum.

Hence, four different grouping combinations have to be considered. Consider only data from the years 2000 and 2001.

Provide a similar query that considers all possible grouping combinations. How many grouping combinations do exist for n dimension attributes.

Your result document must include:

- the two SQL statements that you used to derive the result table (add all statements to the text file as well)
- screenshots of the result tables
- The calculation for the number of grouping combinations