

Data Warehouse design assumptions

Paulina Jędrzejewska, Federico Inserra

All the decisions have been made to minimize the cost of KPIs obtaining queries. The costs have been obtained from the “Execution plan” in DBeaver tool. For costs estimation we used filtering criteria listed in the quiz on LearSQL. We also kept focus on the fact that FH, FC, ADOSS and ADOSU are queried more frequently than the other KPIs.

We decided to create B-tree indexes on AIRCRAFTUTILIZATION(AIRCRAFTID), LOGBOOKREPORTING(AIRCRAFTID) and LOGBOOKREPORTING(PERSONID) (B-tree indexes was better for our case than Bitmap indexes - when comparing costs of execution). All the costs are listed in SQL with queries.

We are not using indexes on people, months and aircrafts dimension tables as indexes should only be used on large tables. We haven't noticed significant cost bestenefits after applying indexes on the TemporalDimension table.

When executing nested queries, filtering is applied as early as possible (in the lowest nested query possible) to reduce the number of processed data.

All the KPI queries and index creation queries can be found in the file **Queries.sql**.