

IT-CS 2022 - Bei 100 Mrd. Clickstream-Events & Backend-Daten den Überblick wahren? Mit SQL geht's - Eine Coding-Challenge

Challenge 1

Jedesmal, wenn ein User ein Produkt in einem der Otto Group online Shops anschaut oder eine Produktliste angezeigt bekommt, wird ein Eintrag in der Tabelle `oghub-analytics.it_cs_2022.display_product_details` (4,295,128,753 Zeilen) bzw `oghub-analytics.it_cs_2022.display_product_list` (61,376,318,277 Zeilen) erzeugt.

Berechne für alle Produkte die Anzahl der eindeutigen Nutzer, die sich das Produkt auf einer Detailseite/Produktliste am 01.05.2021 und 02.05.2021 angeschaut haben.

So könnte der Output einer Lösung aussehen:

productID	date	viewCount
EC0101-S0I0S0RT	2021-05-01 00:00:00 UTC	524
EC0101-759068328	2021-05-01 00:00:00 UTC	202
EC0101-1160556334	2021-05-01 00:00:00 UTC	3404
EC0101-1222277937	2021-05-02 00:00:00 UTC	689