

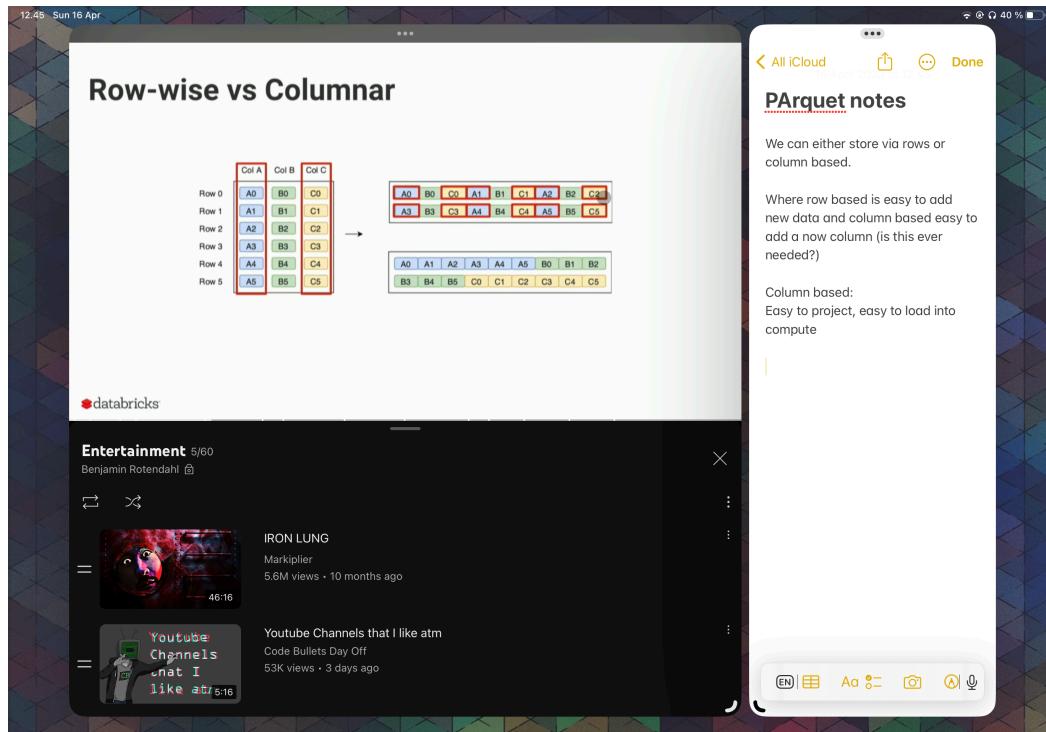
# PArquet notes

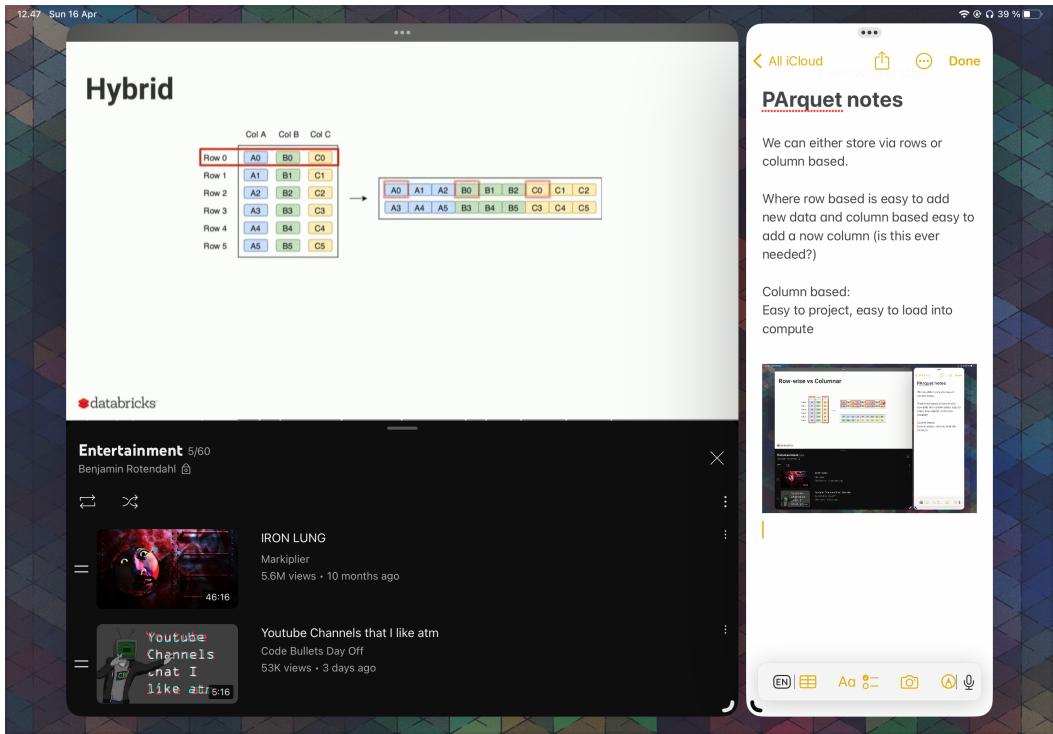
We can either store via rows or column based.

Where row based is easy to add new data and column based easy to add a new column (is this ever needed?)

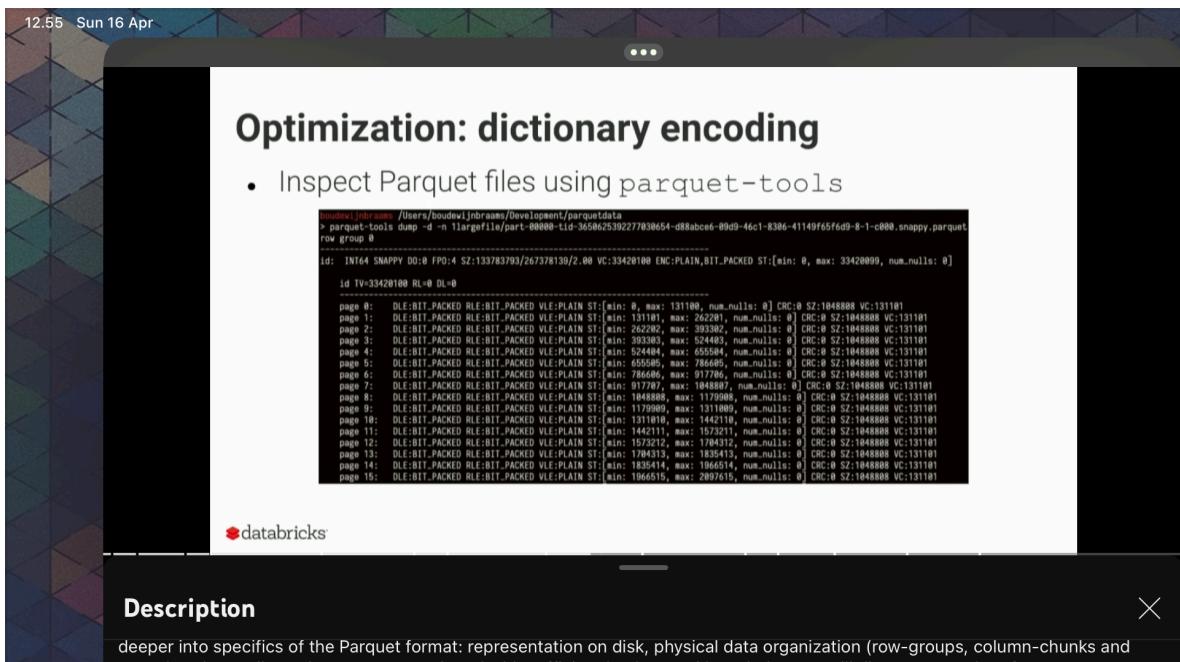
Column based:

Easy to project, easy to load into compute





<https://www.databricks.com/product/data-lakehouse>



[https://youtu.be/1j8SdS7s\\_NY](https://youtu.be/1j8SdS7s_NY)