Mini Assignment – week 5 , Analytics and SQL

Shivam Kumar

SO for importing the dataset, I first tried with [HOW TO QUERY NESTED XML DATA IN SNOWFLAKE](https://community.snowflake.com/s/article/HOW-TO-QUERY-NESTED-XML-DATA-IN-SNOWFLAKE) ,this solution, I tried but the splitting of the xml row in the variant type in snowflake didn’t materialize correctly , so I tried using the another method from [6. Loading XML Data Into Snowflake from a Local File System Using COPY Command| Snowflake| VCKLYTech - YouTube](https://www.youtube.com/watch?v=1nCdaJGvmko), But this as well showed me how to do it on the nested xml data,

I just couldn’t find any thing on splitting by attributes of xml data, But I saw a lot on CSV into snowflake so I used convertcsv.com to convert xml to csv as I don’t have experience on how to write code on C++ to parse.

Still due to table like post having different attributes for ‘Questions’ and ‘Answers’ I had some error, but I was able to upload badges table data.

Then I tried making models and was able to write satisfactory queries for all except 5th milestone , I wanted to use a view to check if my approach works but I didn’t have any data to test it against.

I have pushed the models on my github