Azure PDW What is Active

Lately I have had the opportunity to work with performance tuning of queries running on the Azure Parallel Data Warehouse (Azure PDW). This has been interesting in that everything you thought you knew about SQL Server DMV's, writing queries and overall performance tuning is just a little bit different.

My goal was to write a query to show me what is currently active and running on the Azure PDW database.

To start with, I discovered the view called SYS.DM_PDW_EXEC_REQUESTS which contains all kinds of great information to get us started. Specifically it hold information on queries that are currently running or have recently been run or attempted to be run on the Azure PDW database.

Which returned hundreds of rows, and didn't really get me to where I wanted to go.
Next I added a WHERE statement to filter out those queries that were "done". In this case done means that were 'Completed', 'Failed' or 'Cancelled'.

That got me a bit closer to what I was looking for, but it wasn't entirely useful, so I changed the SELECT * to just the columns I was looking for and came up with this query, which produced exactly what I was looking for.

```
SELECT [submit_time],
[start_time],
[request_id],
[command],
[status],
[total_elapsed_time]
FROM SYS.DM_PDW_EXEC_REQUESTS
WHERE [STATUS] NOT IN('Completed', 'Failed', 'Cancelled'); -- Skip those that are done.
```

But every time I run it, it includes my current query observing the results. I know that the query that I am running must be running, therefor I can exclude my current session to see what the real "other stuff" running really is.

Then since the elapsed time in milliseconds is a bit hard to track, I have converted it into seconds.

And there you have it, a quick script to show what is active on the Azure Parallel Data Warehouse.