



Get started



Publishe

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.



Nikola Ilic

Follow

Nov 7 · 8 min read · ✨ · 🎧 Listen

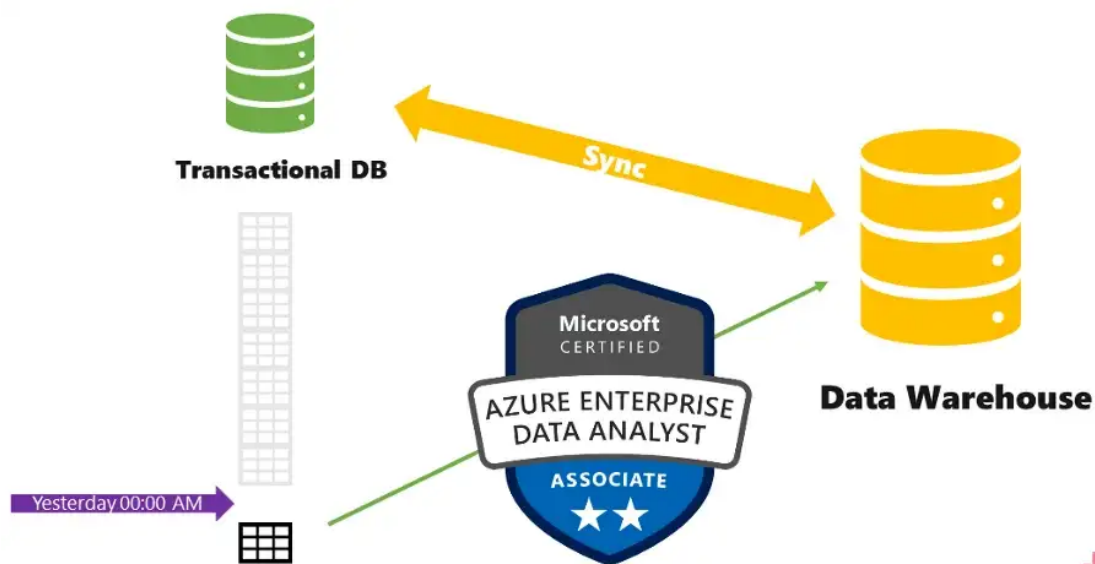


Save



# Incremental refresh in Power BI — everything you need to know

Tired of slow and inefficient data refresh process? Learn how to leverage the Incremental refresh feature to speed up and optimize the data refresh



@DataMozart

Image by author

How many times have you bragged about the following: oh, no, I need to refresh my fact table, but it takes soooo long...Maybe it was happening because you haven't used





Get started

wrong! This  
traditional d

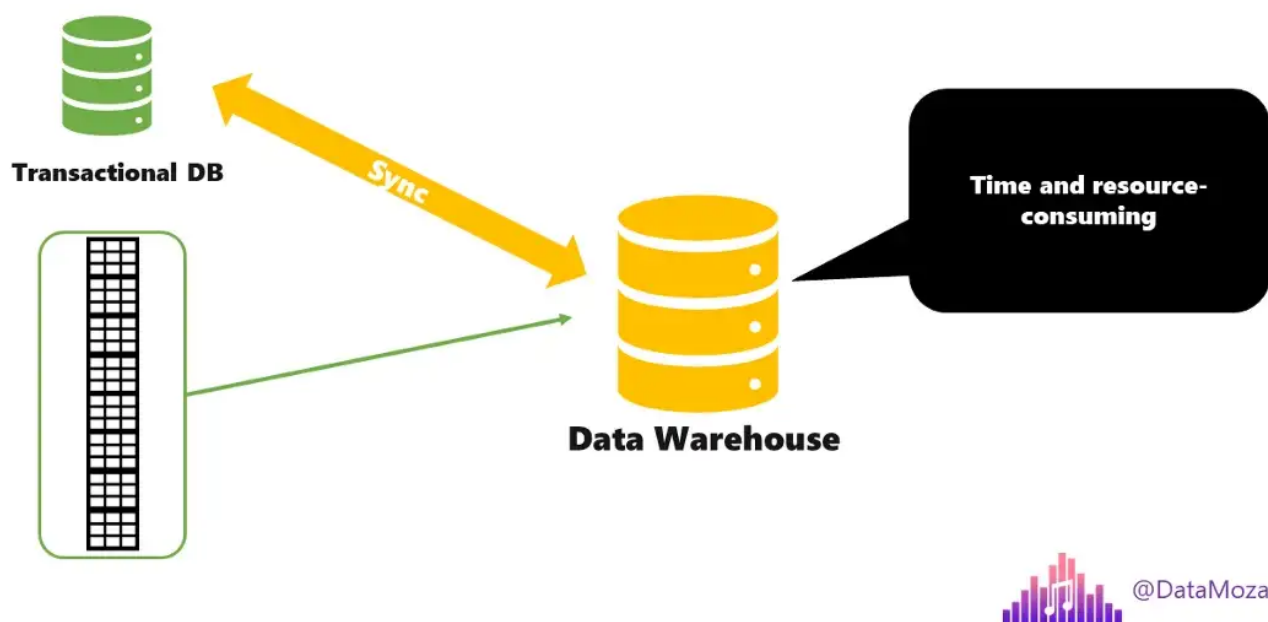
To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

ically every

## Incremental refresh is the process of loading new or updated data, after the previous data loading cycle!

Behind this definition, there are various important aspects to understand. First, why is incremental refresh so important? Well, imagine that you have a huge transactional database, where the main table contains millions of rows. In order to keep your analytic workloads in sync, you need to load the data from that transactional database into the data warehouse. So, imagine what would happen if you load the whole giant table each and every time?! Exactly, it will consume a lot of resources and may also have an impact on the transactional database:

### Why Incremental Refresh?



@DataMozart

Image by author

Therefore, you should check and consume only those records that arrived or changed **AFTER** the last refresh process! This significantly reduces the pressure on the transactional database as only a small chunk of data needs to be handled.





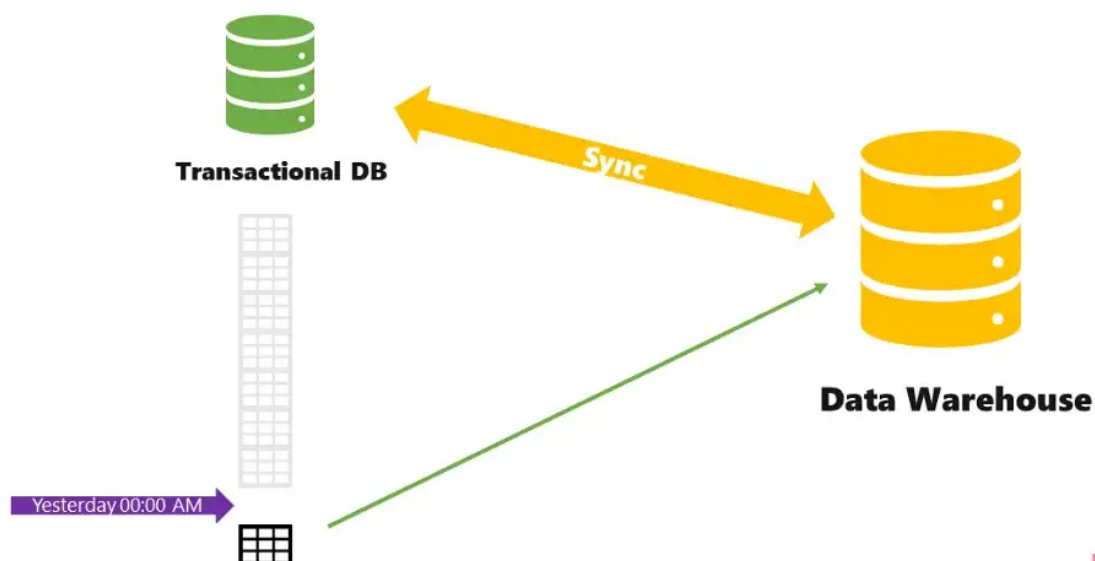
inserted in tl  
only this sm:

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

Get started

Chronize

## Why Incremental Refresh?



@DataMozart

Image by author

### Benefits of Incremental refresh

There are multiple advantages of implementing incremental refresh in your Power BI solution:

- Refreshes are faster — obviously, as you operate with a smaller amount of data, the data loading process will run faster
- Reduced resources consumption — again, handling a smaller chunk of data will help you decrease memory consumption and use other Power BI and data source system resources more efficiently
- Reliable refresh process — if you decide to go “all-in” and against Incremental refresh, it may happen that long-running connections become vulnerable and non-





To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

Get started

## Incremental refresh in Power BI

Once you publish your data model to Power BI Service, each table contains one single partition. That partition contains all the rows, which for large tables data refresh, as we've already explained, may be overwhelming.

When you configure incremental refresh, Power BI will automatically partition your table — one partition will contain data that has to be refreshed frequently, while the other partition will hold rows that are not changing.

In the most simplified way, this is how the workflow should look with the incremental refresh in place:

### Incremental Refresh Workflow

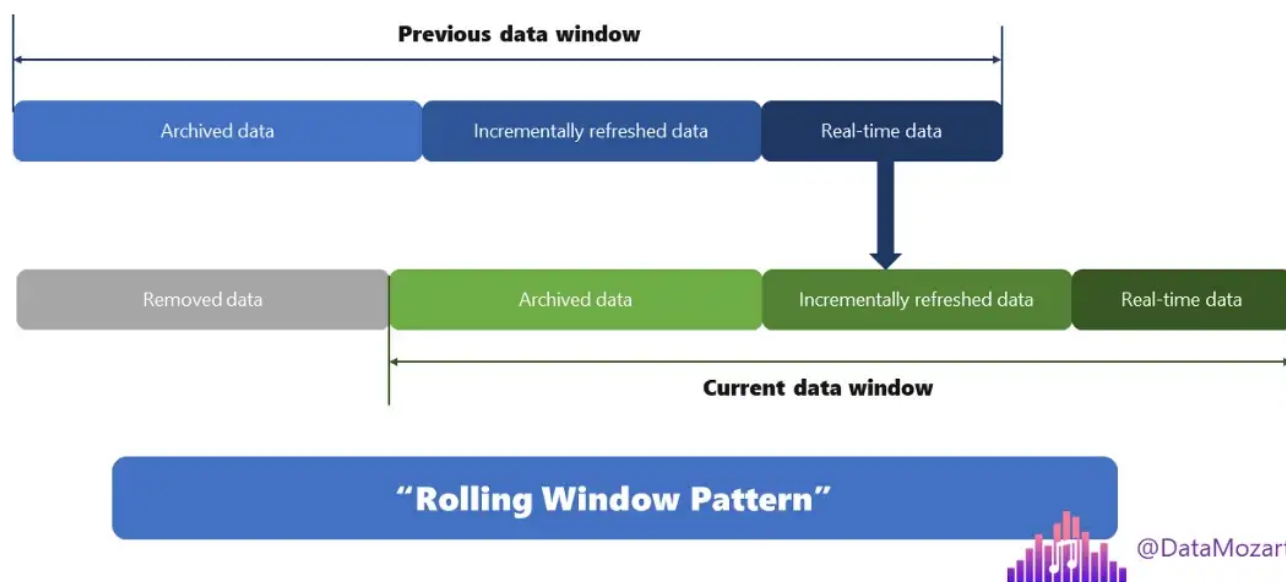


Image by author

As you may notice comparing the previous and current data window, the window is rolling and data that was considered real-time in the previous window, now become part of the incrementally refreshed partition. This is an ongoing process and is known





Get started

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

prerequisite:

- **Date column** — a table to which you want to apply Incremental refresh, must contain a date column, that can be either date/time or integer data type. This is because you need to set up parameters for separating the data between partitions
- **Query folding** — now, you're probably asking yourself: what is query folding? So, let's first explain this one. I've already written about it in more detail in [this article](#), but, in a nutshell, query folding is the ability of Power Query to generate a single SQL query that will be executed on the SQL data source side. Why does query folding matter for the incremental refresh? Well, your date range parameters need to be translated to a WHERE clause in SQL, in order to separate data in relevant partitions. Therefore, without query folding, there is no WHERE clause, no partitioning possible, and consequentially no incremental refresh possible
- **Single data source** — means, all your partitions must query data from a single source

## Wait...There is more!

By leveraging a [Hybrid tables feature](#), you can enhance a data refresh process even more. Essentially, the idea is to set the incremental refresh for the table, but set up the partition with the most recent data in DirectQuery mode, while keeping the older data in the partition that uses Import mode.

This way, you are getting the best of both worlds: blazing-fast performance for analytic queries over older data, and real-time synchronization with the latest data from the original data source. However, at this moment, the Hybrid tables feature is available only with Power BI Premium licenses.

## Setting up Incremental refresh in Power BI

The first step in the configuration process is to define parameters and set their default values to be applied when filtering the data that should be loaded into the Power BI Desktop. You should keep this range short and include only the most recent data (3 days





Get started

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

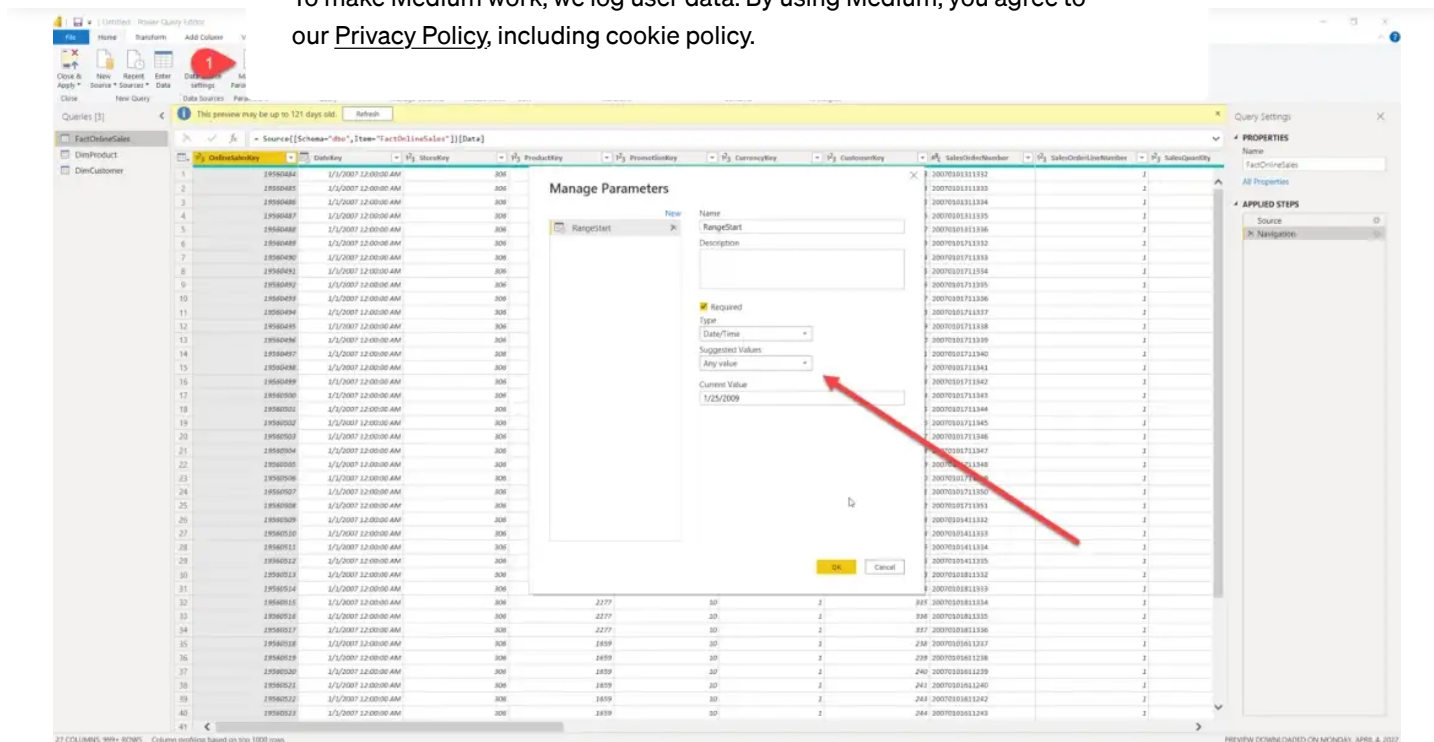


Image by author

Please keep in mind that you should use predefined RangeStart and RangeEnd names for your parameters. That way, Power BI “knows” that these parameters will be used for setting up the Incremental refresh.

The next step is to filter the data, based on our newly created parameters:





Get started

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

Filter Rows

Apply one or more filter conditions to the rows in this table.

Basic Advanced

Keep rows where 'DateKey'

is after or equal to

And Or

Date/Time

Parameter

New Parameter...

OK Cancel

27 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

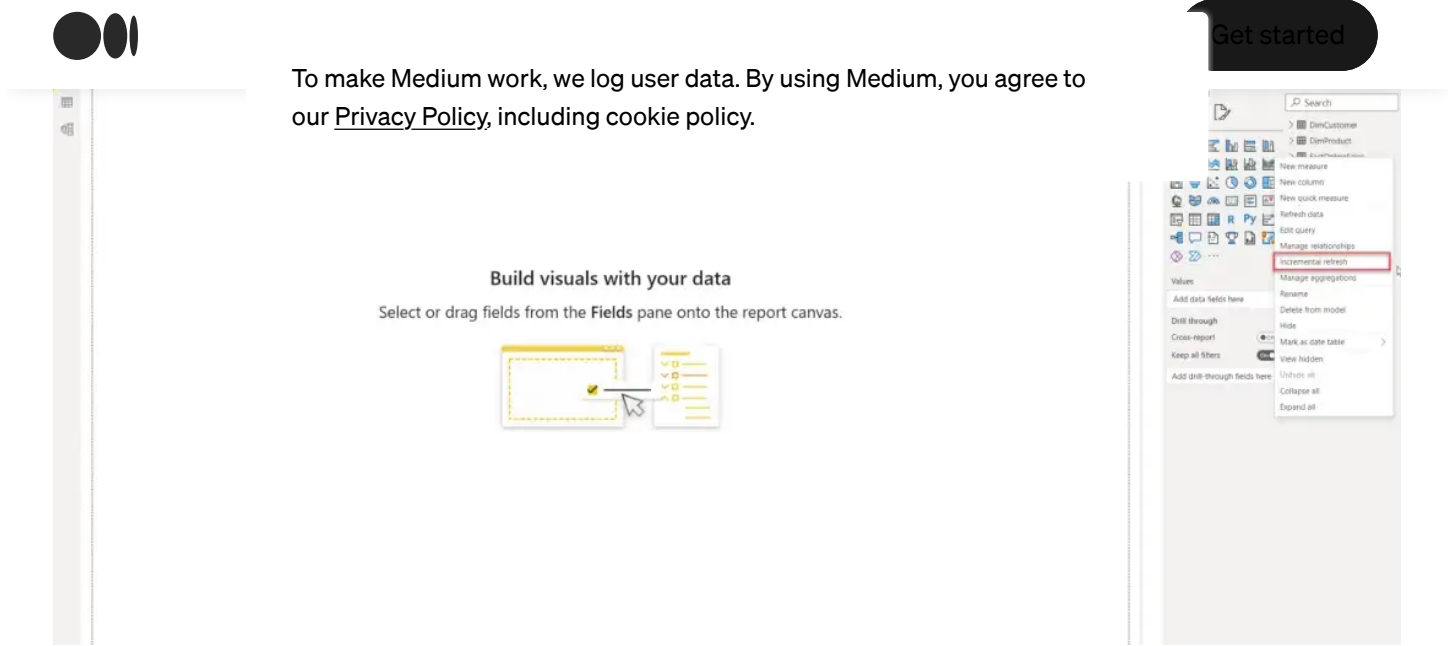
Image by author

Watch out when setting the filter conditions — you should avoid using equality on both parameters (i.e. is after or equal to/is before or equal to), because that can fetch duplicated data and produce incorrect results in your report!

I can now hit Close&Apply in Power Query Editor and you may notice that Power BI will load only a small portion of data — in my case, instead of the original 12.6 million rows, this time only 36.000 rows will be loaded.

Now comes the crucial part of the workflow — we will define the incremental refresh policy for our giant fact table:






Click on three dots next to a table name and select Incremental refresh

Once the Incremental refresh dialog window pops up, there are various options to choose from:





To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

Get started

that, y

### 1. Select table

FactOnlineSales

### 2. Set import and refresh ranges

☒

 Incrementally refresh this table

Archive data starting 

3

Years

 before refresh date

Data imported from 1/1/2019 to 7/30/2022 (inclusive)

Incrementally refresh data starting 

3

Days

 before refresh date

Data will be incrementally refreshed from 7/31/2022 to 8/2/2022 (inclusive)

### 3. Choose optional settings

☐ Get the latest data in real time with DirectQuery (Premium only) [Learn more](#)

☐ Only refresh complete days [Learn more](#)

☐ Detect data changes [Learn more](#)

### 4. Review and apply

Archived

Incremental Refresh

3 years before refresh date

3 days before refresh date

Refresh date

Apply

Cancel

Image by author

After I toggled on Incrementally refresh this table property, I'll define the archive period (3 years in my example) and the period for which my incremental refresh policy will be applied (3 days). The cool thing is that you immediately see which dates will be



[Get started](#)

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

*Only refresh complete days* option comes in very handy if you have your data refresh process scheduled for, let's say, 3 AM (assuming that at that time there is less pressure on the underlying data source), but you don't want to import into Power BI only these records that were inserted in the source system between midnight and 3 AM. With this option enabled, only records for completed days will be loaded.

*Detect data changes* may bring an even greater performance boost to the data refresh process because it enables the processing of only those records that were changed since the previous run. For this option to work, you need to have a specific column in the data source containing the information when the specific record was last time updated (for example, LastUpdated, LoadDate, and so on). This column must be the date/time data type and *should not be the same column that you use for partitioning the fact table* (in our case, it should not be the DateKey column).

Finally, if you scroll down, you can see your incremental refresh rolling window visualized:



11





Get started

To make Medium work, we log user data. By using Medium, you agree to our [Privacy Policy](#), including cookie policy.

that,

## 1. Select table

FactOnlineSales

## 2. Set import and refresh ranges

☒ Incrementally refresh this table

Archive data starting 3 Years before refresh date

Data imported from 1/1/2019 to 7/30/2022 (inclusive)

Incrementally refresh data starting 3 Days before refresh date

Data will be

## 3. Choose data

- ☐ Get the latest data
- ☐ Only the latest data
- ☐ Detect data changes [Learn more](#)

You can see your incremental refresh policy rolling window

## 4. Review and apply

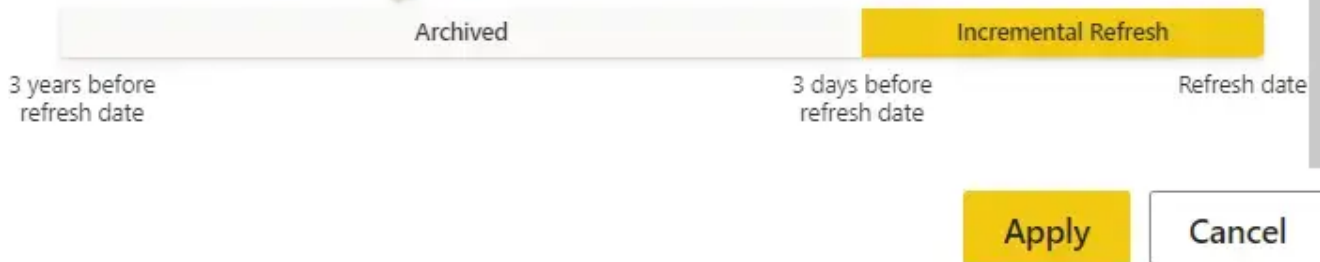


Image by author

Hit Apply and now I'm ready to publish the report and the dataset. Once the dataset is in the





This first dat  
refreshes sh

To make Medium work, we log user data. By using Medium, you agree to  
our [Privacy Policy](#), including cookie policy.

Get started

ut all subsequer  
l in the

incremental refresh policy will be loaded.

## Conclusion

When you're dealing with relatively small data models, you may sneak without implementir  
an Incremental refresh for your data. However, once the data volumes begin to scale,  
Incremental refresh will quickly become your best friend in optimizing the data refresh  
process.

And, it's not only about the time needed for bringing the fresh data into your Power BI  
datasets, but also the resources needed to reprocess the whole gigantic tables.

Thanks for reading!

[Become a member and read every story on Medium!](#)

