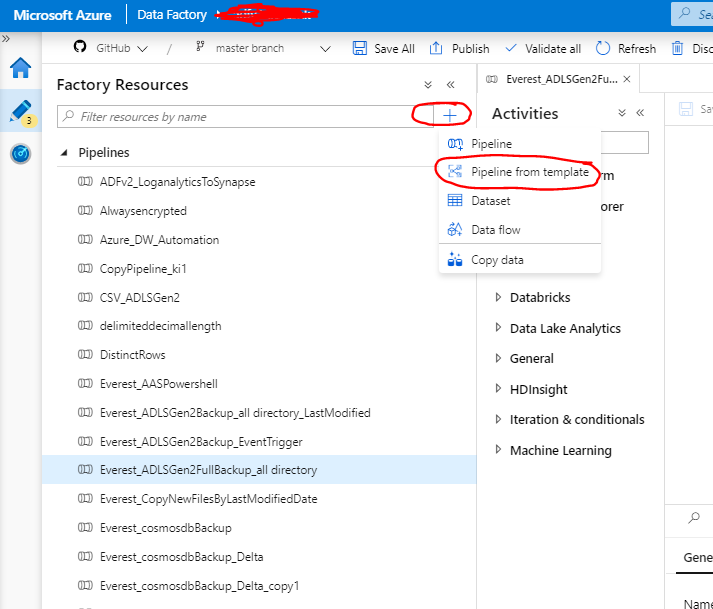
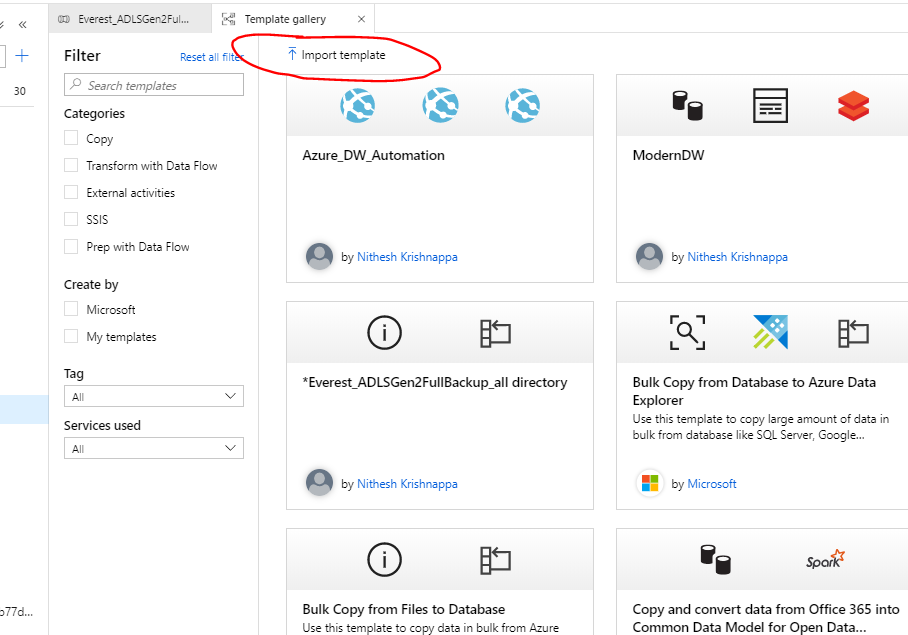
**Deployment Guide**

1. **ADLS gen2 Backup using ADF v2 :**
2. **ADLS gen2 Full backup**

* Prerequisite:
* [Provision](https://docs.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory-portal#create-a-data-factory) ADFv2 workspace if you don’t have one.
* Create source [Linked service](https://docs.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory-portal#create-a-linked-service) to your Source ADLS gen2 Storage in the Data factory.
* Create Sink(Destination) [Linked service](https://docs.microsoft.com/en-us/azure/data-factory/quickstart-create-data-factory-portal#create-a-linked-service) to your Source ADLS gen2 Storage in the Data factory.
* Download the Zip File “ADLSGen2FullBackup\_all directory” from Email
* In your data factory Workspace Click Create pipeline from Template.

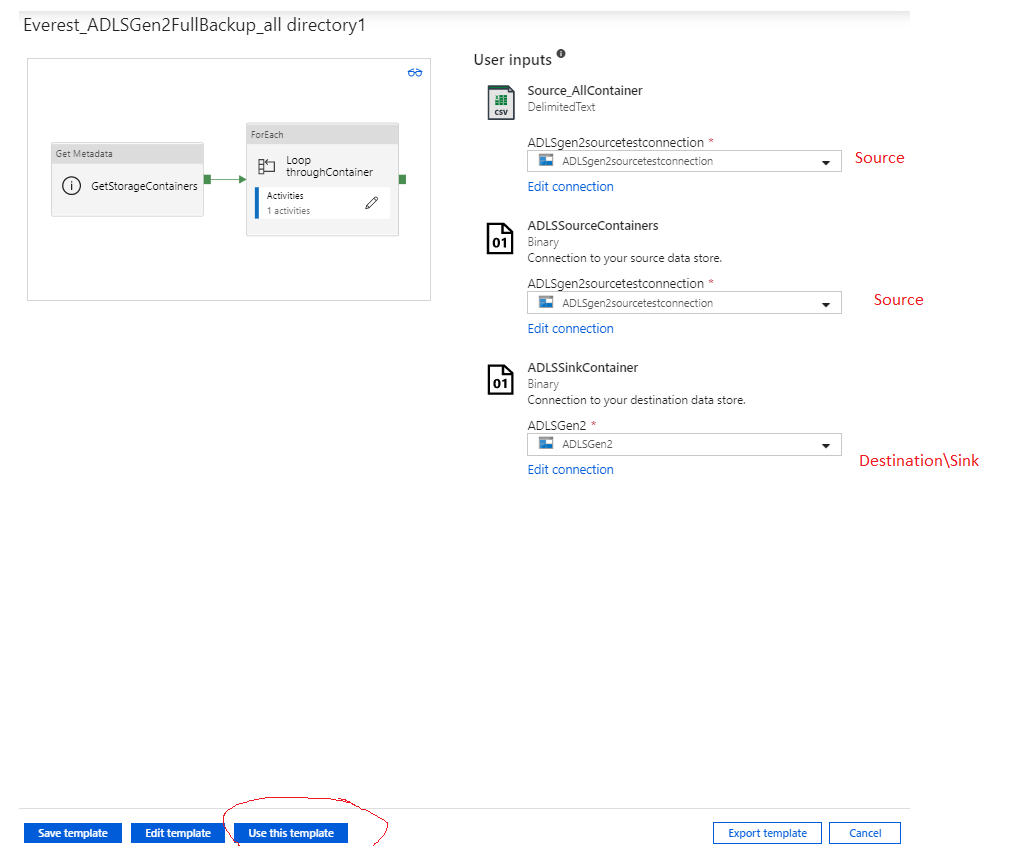


* Once you are in Template gallery click import Template icon.

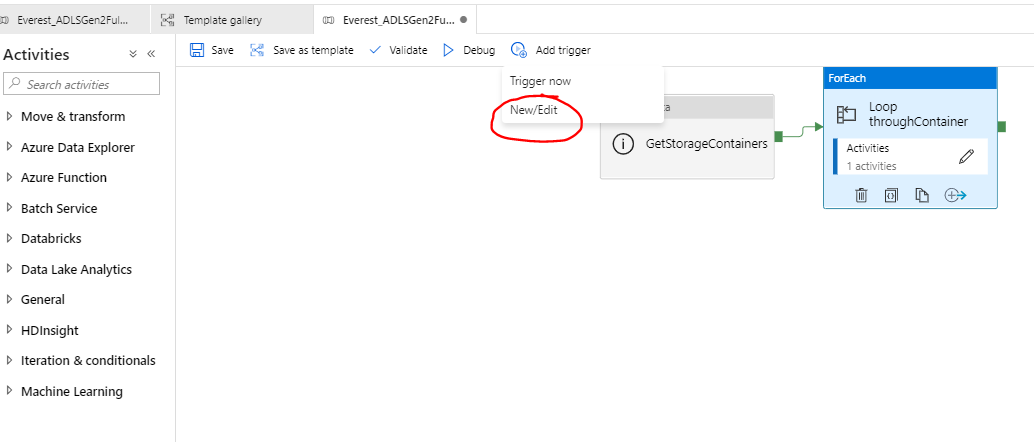


* Browse for the zip file that you downloaded “ADLSGen2FullBackup\_all directory”
* You will see below screen in your ADF. Map Source Linked services that you created in prerequisite to “Source\_Allcontainer” and “ADLSSourceContainers” in the input dropdown.

Map Sink\destination Linked service that you created in prerequisite to “ADLSSinkContainer”. And then click Use this template button.

****

* Your pipeline is imported now and save and publish. There are two activity first activity(GetStorageContainers) will pull all container names for the storage account and Second activity (Loop throughContainer) will loop through each container recursively and uses copy activity inside the for each activity to copy data.
* To Schedule this pipeline click on “Add trigger” button on the top and click new\edit.

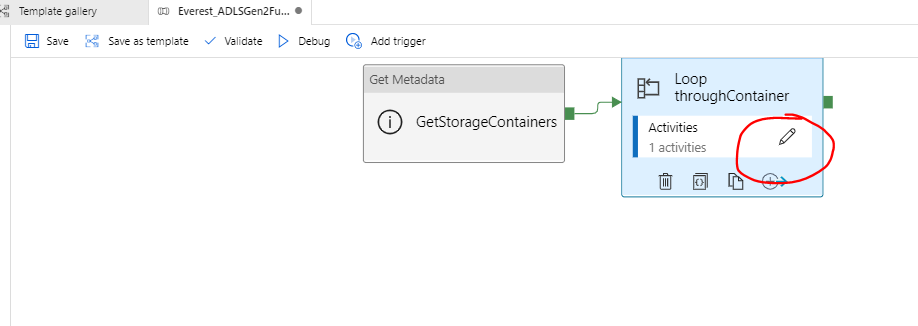


* Please follow below two links to create trigger (Schedule Pipeline) (Scheduled or Tumbling window).

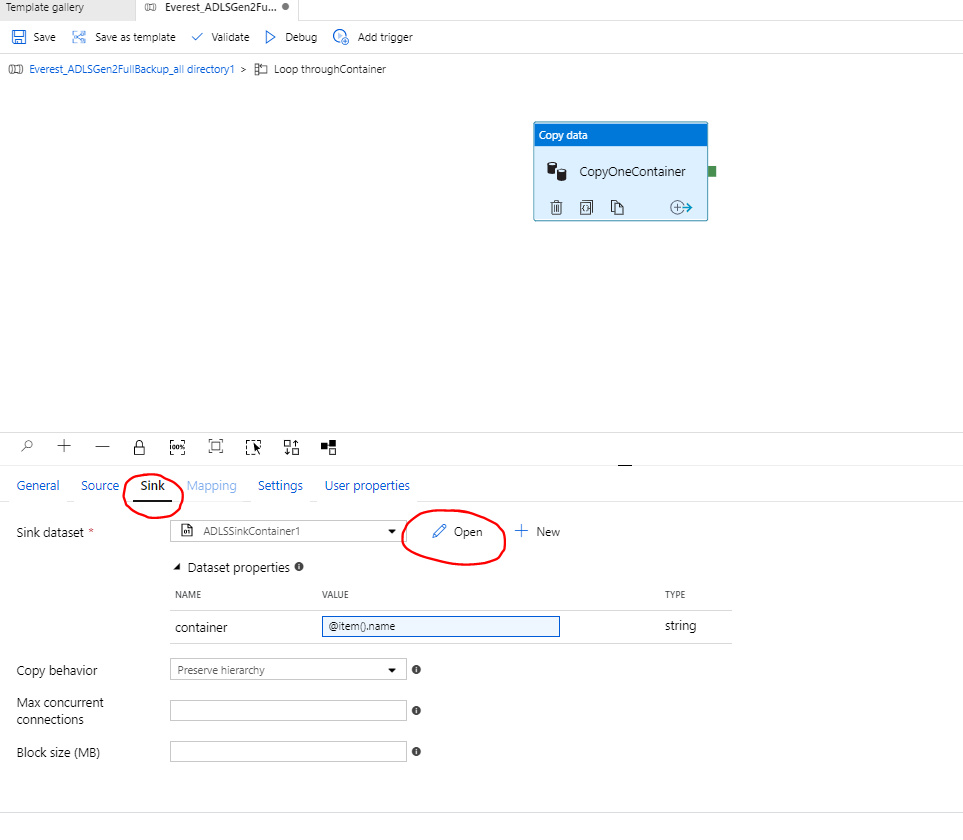
<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-schedule-trigger#data-factory-ui>

<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-tumbling-window-trigger>

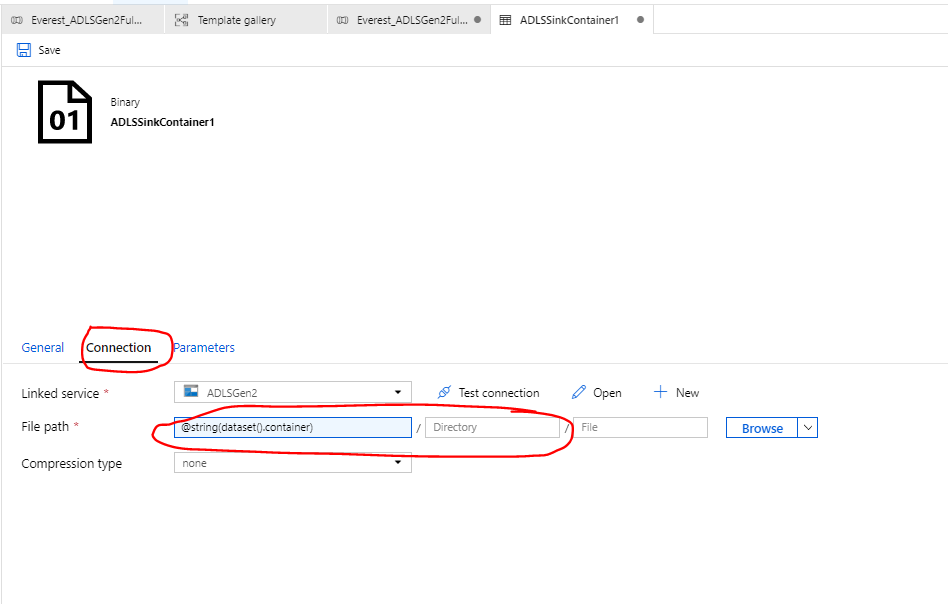
* You can customize the destination path the folder way you want structure the folder. To Customize sink of copy Activity click on pencil button on for each loop activity



Click Copy data Activity and click on Sink tab and click on Open to open the dataset setting



This will open Dataset Tab and click on Connection and you customize file path using expression. In this example I am using simple structure destination will create exactly same container name as source.

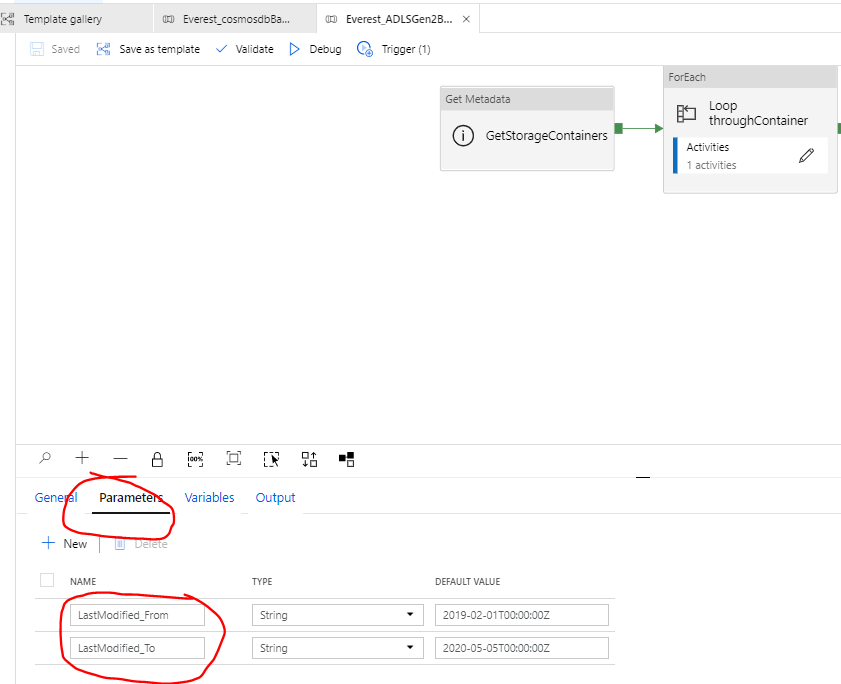


Click one of the box in File Path which will open expression editor, you can [write expression](https://docs.microsoft.com/en-us/azure/data-factory/control-flow-expression-language-functions) that you want .

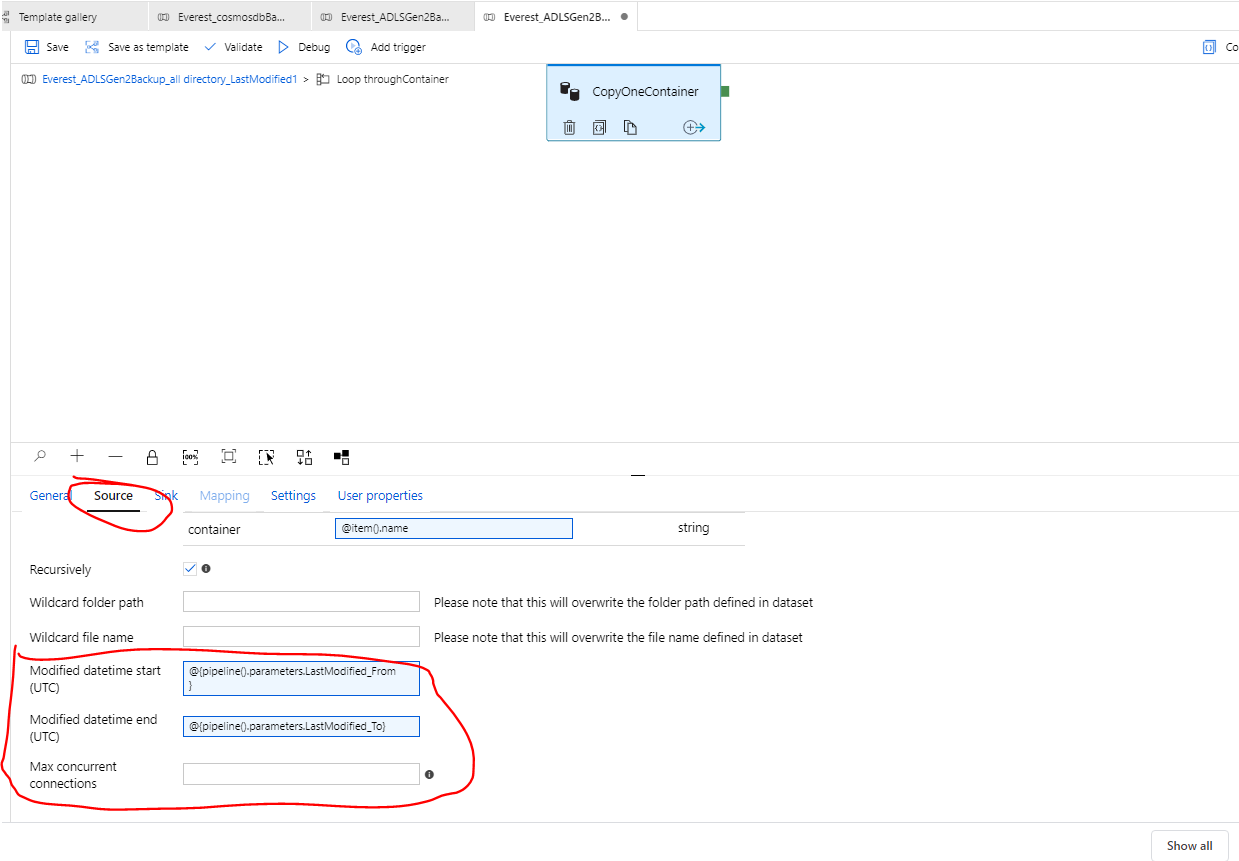
Example : If you want this container to be copied to yyyy\mm\dd format you can use below expression: @concat(formatDateTime(utcnow(),'yyyy/MM/dd'),'/',string(dataset().container))

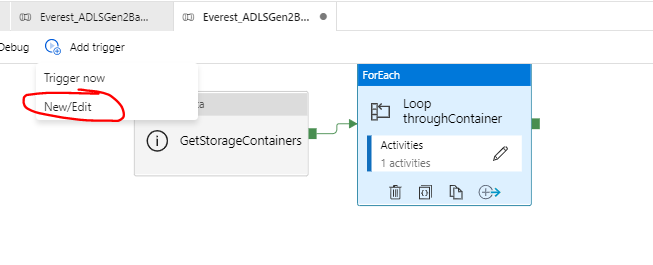
1. **ADLS gen2 Incremental backup:**

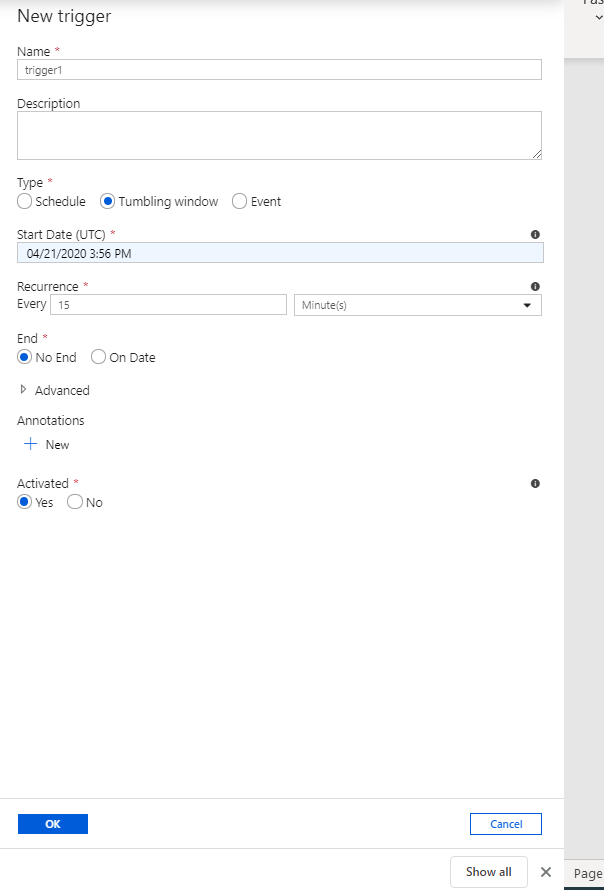
* Download the Template “**ADLSGen2Backup\_all directory\_LastModified**” and follow the same step as full backup template (a) to create linked service and import the template.
* Only difference in this template is we added two parameter to keep track of last modified files.



* If you look at the copy activity source inside for each loop you could see built in functionality to pick files based on last modified date and we are mapping Pipeline parameter to these field .



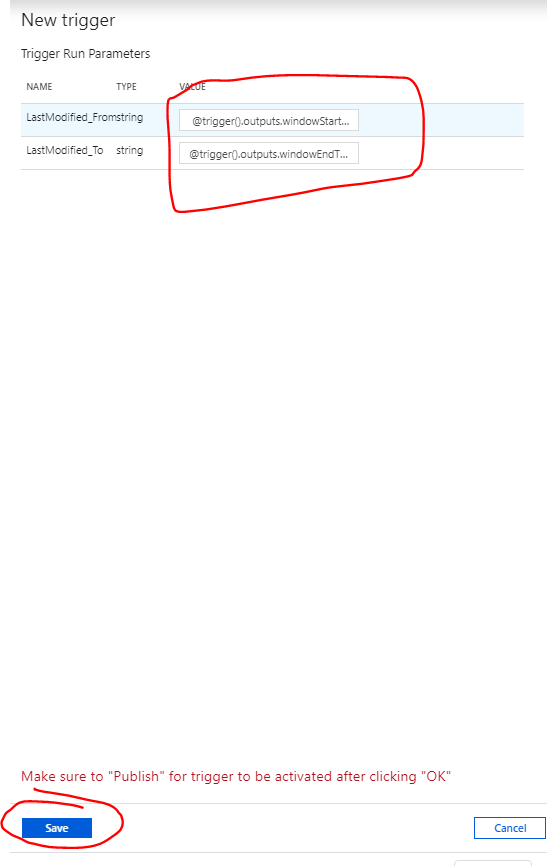
* You need to create trigger
* Next window 
* Next select the appropriate trigger option



* When you click ok it will prompt value for parameter and fill the below value :

LastModified\_From = @trigger().outputs.windowStartTime. It is a system variable from the trigger determining the time when the pipeline was triggered last time.

LastModified\_To = @trigger().outputs.windowEndTime. It is a system variable from the trigger determining the time when the pipeline is triggered this time.



Note : Please refer below link for more details :

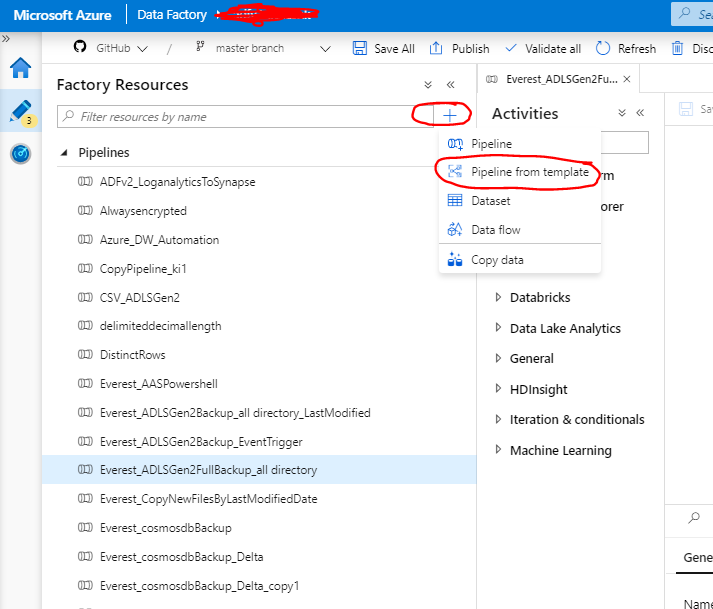
<https://docs.microsoft.com/en-us/azure/data-factory/solution-template-copy-new-files-lastmodifieddate>

1. **Repeat the above adls gen2 deployment method(a and b) for Blob Storage and Azure File Storage:**

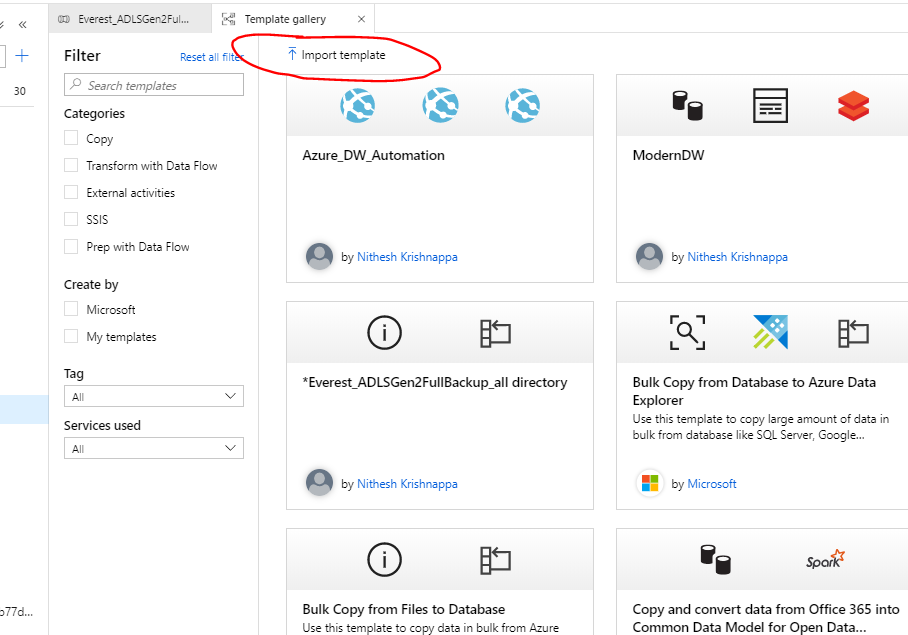
* **Blob Storage Templates:** CreateSource and Sink(destination) [blob linked services](https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-blob-storage) and use below templates to deploy.
* Full backup 🡪 BlobFullBackup\_all.zip
* Incremental backup 🡪 BlobFullBackup\_all\_Incrsemental.zip
* **Azure File Storage:** CreateSource and Sink(destination) [Azure File linked services](https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-file-storage) and use below templates to deploy.
* Full backup 🡪 FilestorageFullBackup\_all.zip
* Incremental Backup 🡪 FilestorageFullBackup\_Incremental.zip

1. **Cosmodb** **Backup :**
2. **SQL API Cosmos:**
3. Full backup :

* Create Cosmos Db [linked service](https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-cosmos-db)
* Create Destination Storage linked service where you want to store the backup.
* Download the Zip File “cosmosdbBackup” from Email
* In your data factory Workspace Click Create pipeline from Template.

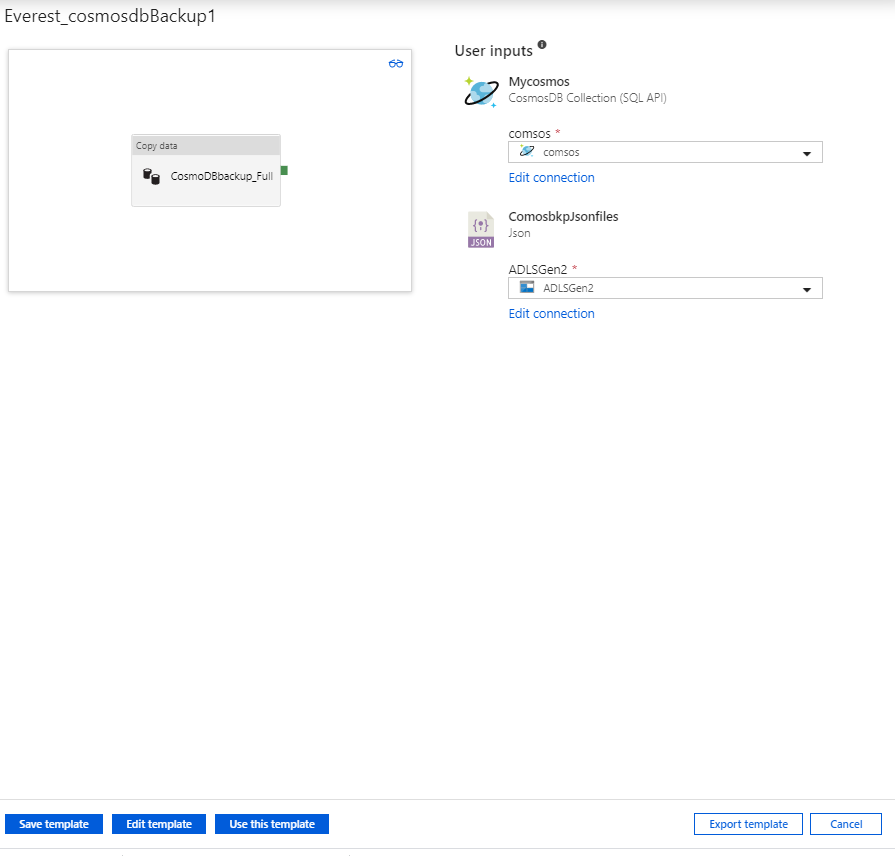


* Once you are in Template gallery click import Template icon.

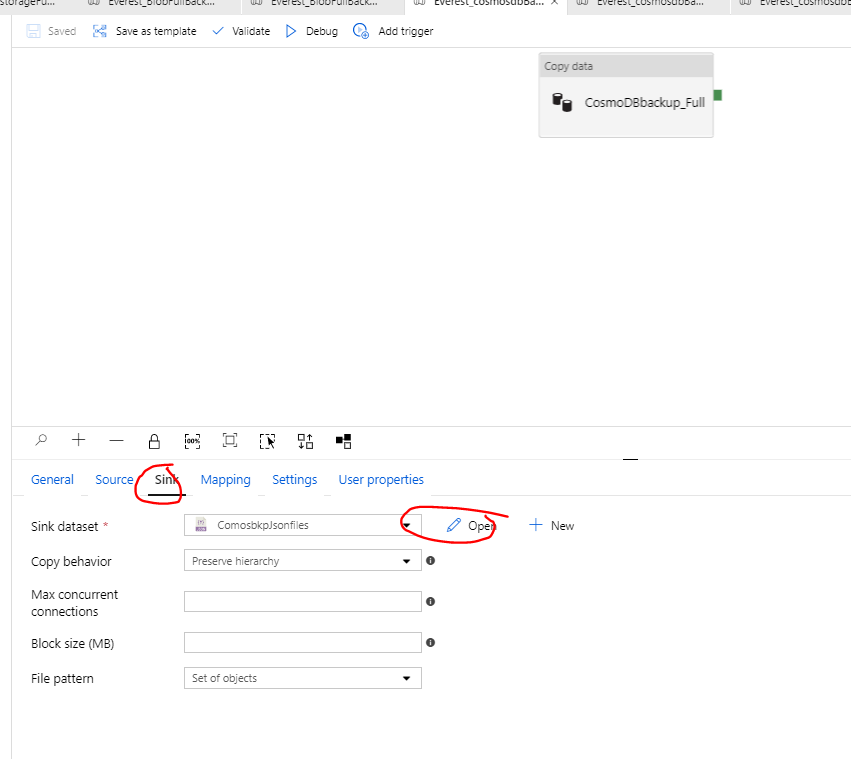


* Browse for the zip file that you downloaded “cosmosdbBackup”
* You will see below screen in your ADF. Map Source Cosmos Linked services to “Cosmos” and

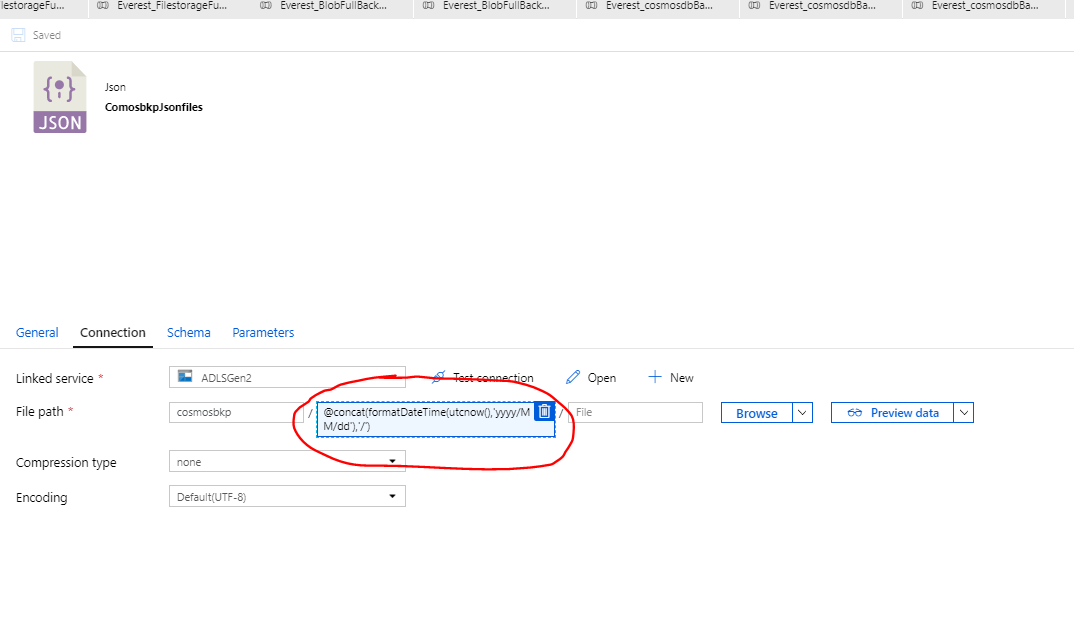
Map Sink\destination Linked service that you created in prerequisite to “ADLSgen2”. And then click Use this template button.



* Once you create pipeline if you want to modify folder structure of backup please check your sink data set by clicking open :



* In Dataset tab you can modify the storage path by writing [expressions](https://docs.microsoft.com/en-us/azure/data-factory/control-flow-expression-language-functions)



* Please follow below two links to create trigger (Schedule Pipeline) (Scheduled or Tumbling window).

<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-schedule-trigger#data-factory-ui>

<https://docs.microsoft.com/en-us/azure/data-factory/how-to-create-tumbling-window-trigger>

* You can use template “cosmosdb\_Restore.Zip” to restore the Backup.

1. **Incremental Cosmosdb Backup(SQL API):**
   * Prerequisite:

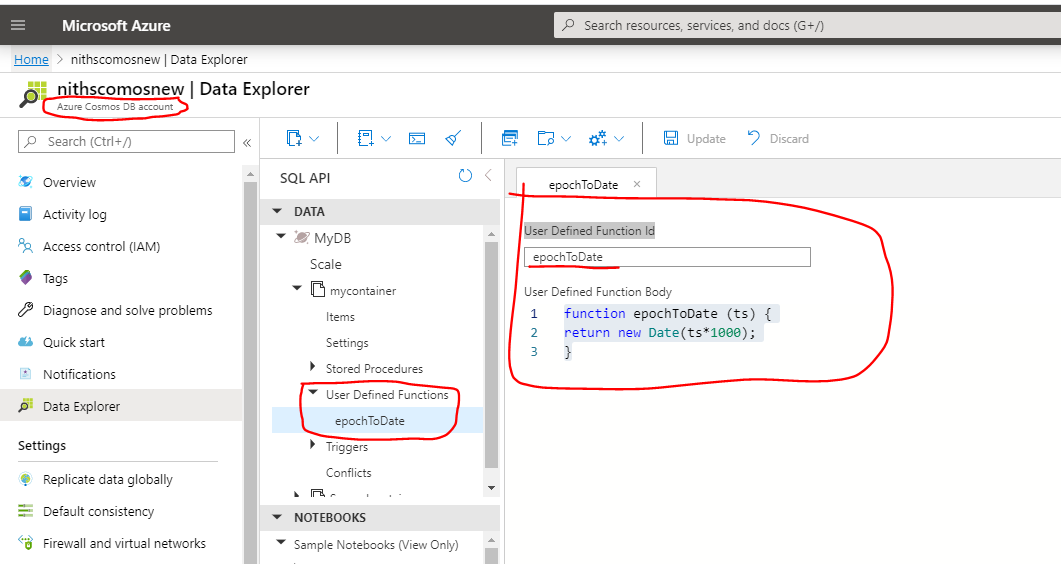
Create below User defined function in cosmosdb Collection :

User Defined Function Id :epochToDate

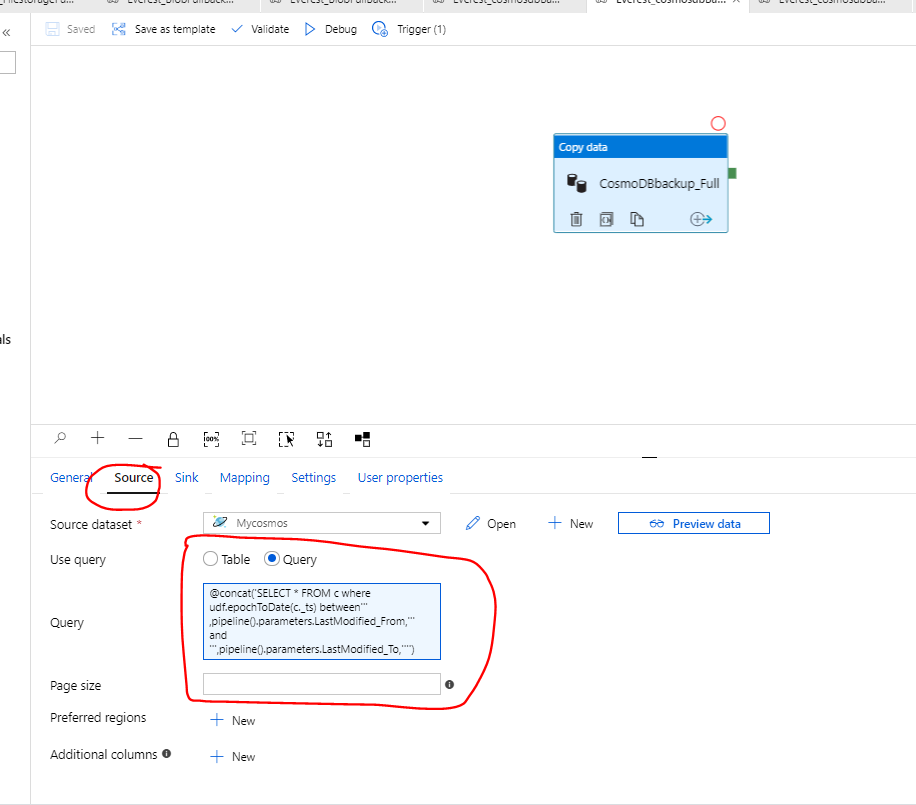
function epochToDate (ts) {

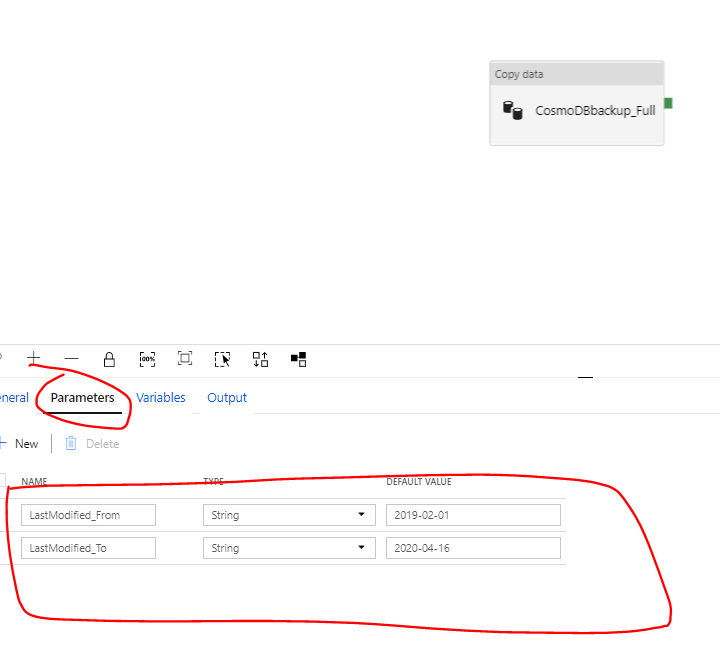
return new Date(ts\*1000);

}

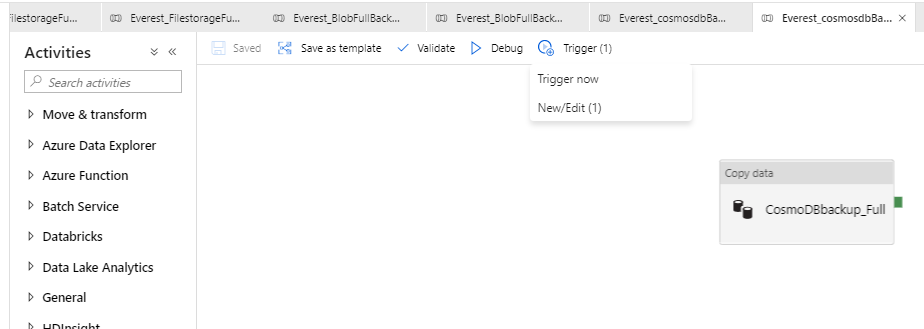


* Use template “cosmosdbBackup\_Delta.zip” , follow the same step as previous one to import the template .
* This template is same as cosmodb Full backup template except this has dynamic query in Source which uses UDF to convert cosmos db Unix timestamp to datetiemformat and has two parameter to get from and to date.

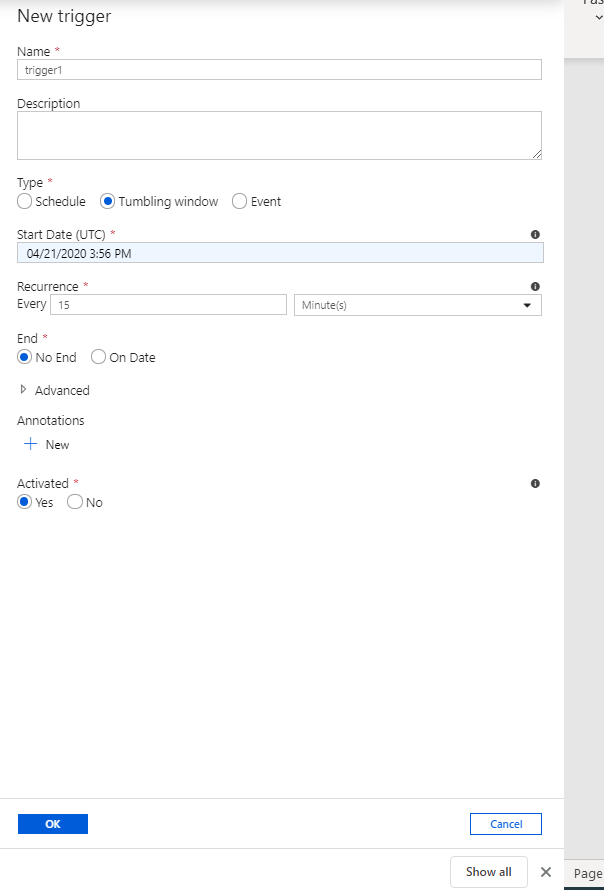




* Next step is to create trigger to automate the incremental load.click on add trigger and then new/edit .



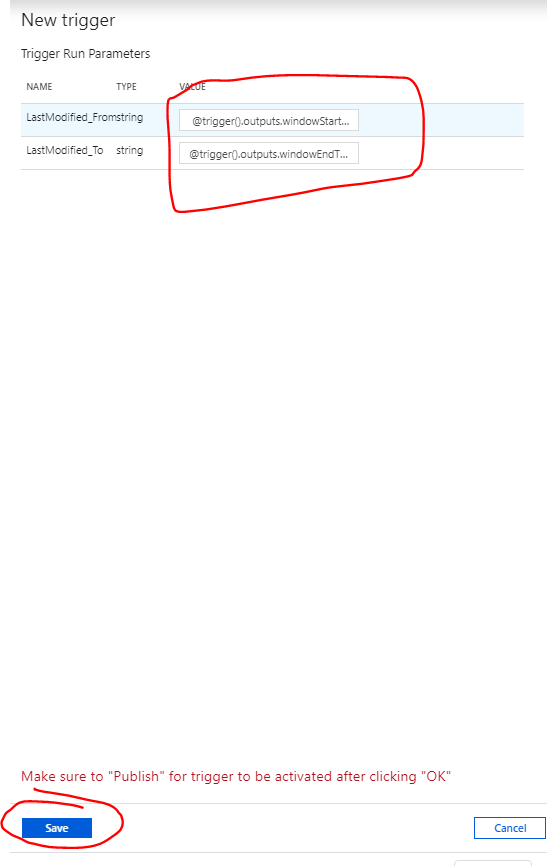
* Next select the appropriate trigger option



* When you click ok it will prompt value for parameter and fill the below value :

LastModified\_From = @trigger().outputs.windowStartTime. It is a system variable from the trigger determining the time when the pipeline was triggered last time.

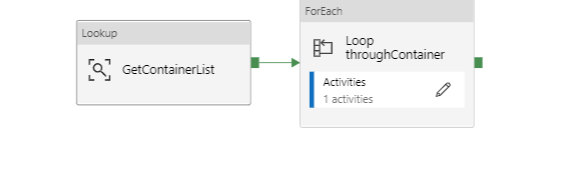
LastModified\_To = @trigger().outputs.windowEndTime. It is a system variable from the trigger determining the time when the pipeline is triggered this time.



Note : Please refer below link for more details :

<https://docs.microsoft.com/en-us/azure/data-factory/solution-template-copy-new-files-lastmodifieddate>

* Optionally If you want to load multiple cosmosDb collection you can use template “cosmosdbBackup\_LoopContainer.zip” . Please follow same step as explained before , only additional step is to have lookup table or log file which has all cosmosdb Container name .



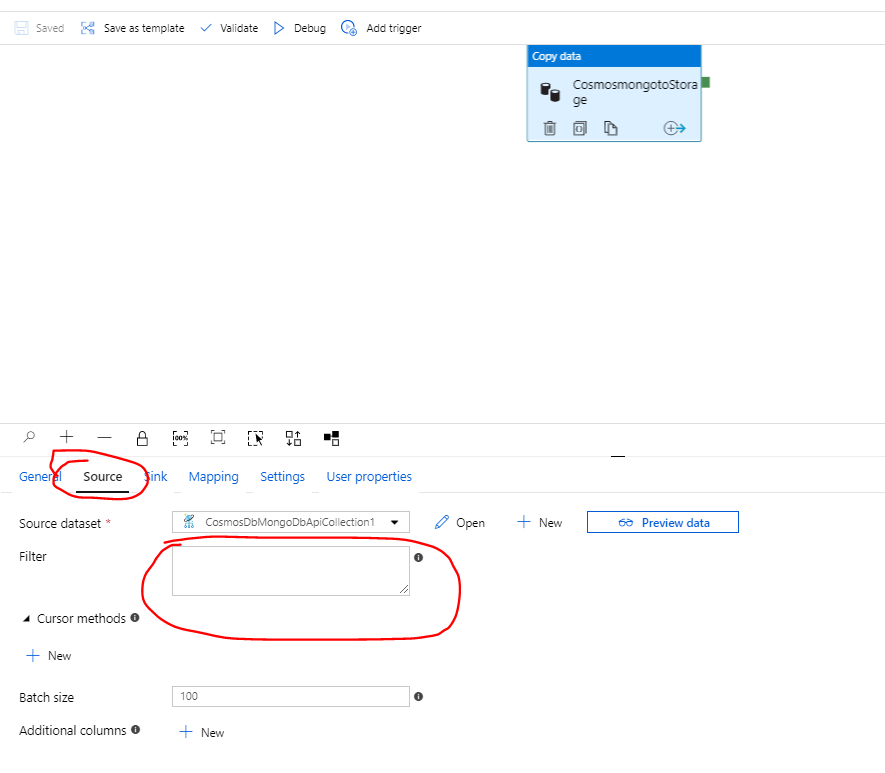
As shown in above image first activity “GetconatinerList” will get all the container name from log files.

1. **Cosmos Mongo API Backup:**

* Create Cosmos Db [MongoAPI linked service](https://docs.microsoft.com/en-us/azure/data-factory/connector-azure-cosmos-db-mongodb-api)
* Create Destination Storage linked service where you want to store the backup.
* Download the Zip File “MongoCosmosDBbkp.zip” and follow the exact step as explained previous section to import the template.
* If you want to restore the backup use template “MongoCosmosDB\_Restore.zip”
* You can use template “MongoCosmosDBbkp.zip” for delta load as well , but you need to have filter criteria (appTimestamp) in cosmos dB to filter

You can apply expression something like below under filter section:

**{"appTimestamp":{$gt: ISODate("@{adddays(utcnow(),-5)}")}}**

**s**