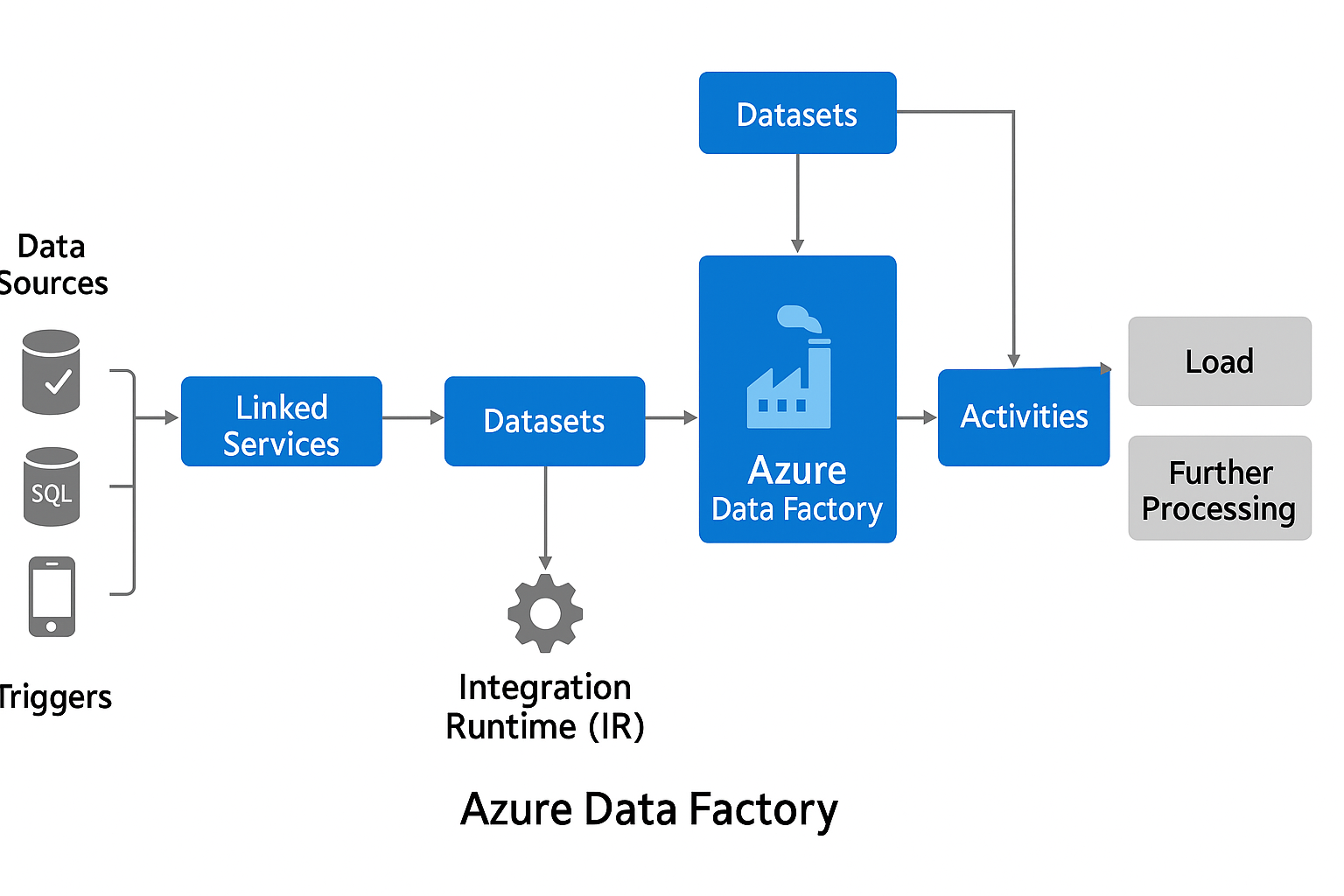
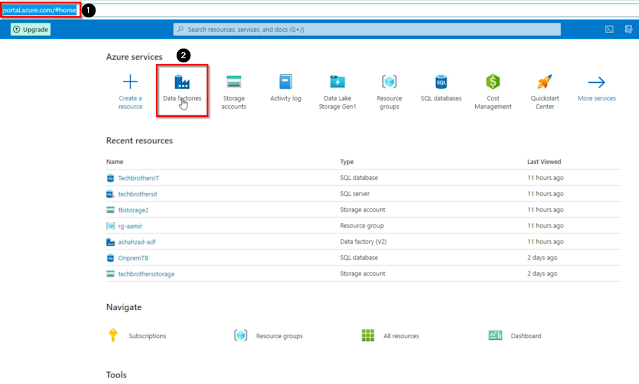
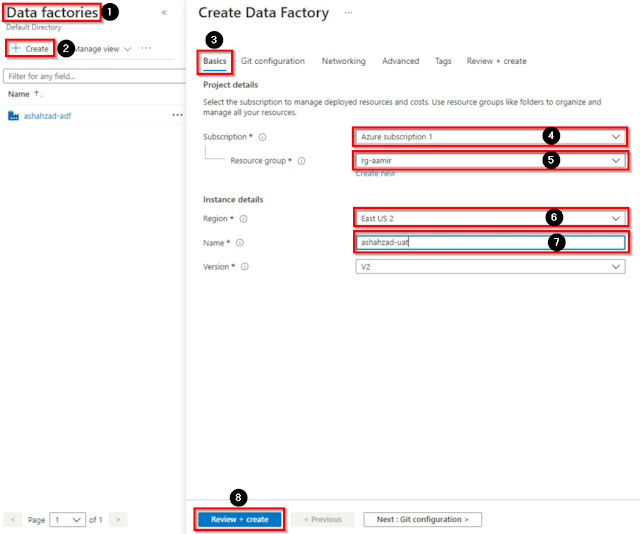
Azure Data Factory is Azure's cloud ETL service (Extract, transform, and load) for scale-out serverless data integration and data transformation. It offers a code-free UI for intuitive authoring and single-pane-of-glass monitoring and management. You can also lift and shift existing SQL Server Integration Service packages to Azure and run them with full compatibility in Azure Data Factory.



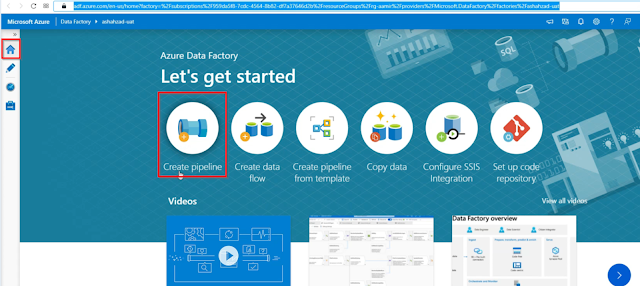
Once we are on our Azure portal, here is a Tab named Data factories, click on that to create a new Azure Data Factory.

[](https://blogger.googleusercontent.com/img/a/AVvXsEgJLc-C8D0SOSP8ssFhPdg66lPQxvpDw-kliMcq6DEfb9Hfe4KzwBwEUALf9gW3f4amzSozWBPRziIGORtr5cL5xPqIe_sH3oHcliHfE6dNdVQjl_R-_9TN4Rj7jKy0UKuVxcC9STTsxGrA2PZ09_pXP8M7kUmK_ejYXdWwi-3bJpYvIUiOVQGzHHXE=s1321)

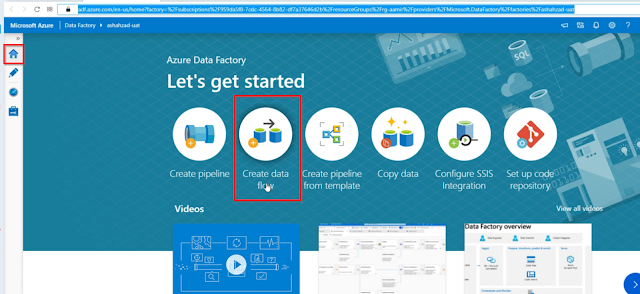
In the data factories, click on the + Create button then, in the basics tab select Azure subscription, then select resource group, then select your region, then give a name to data factory, and then click on the Review + create button then click on Configuration Git later and click on Create.

[](https://blogger.googleusercontent.com/img/a/AVvXsEheEQq9EiYnUWtb52qXBZwaPkm2cSXjCnDnk9nUShH8eRfLLyQa1BWojqcBI3blWjN8M5sCmF0FgMtbeGgkKwiJyUSVnORM845PLNPH6GUH5kVAa8AzkFxQtT6dcUO-6qzPPq20Oe86VjYfgdHLxkPip2394I9knf6OMKcvQgGe4UG-702NvPmdKYAV=s1050)

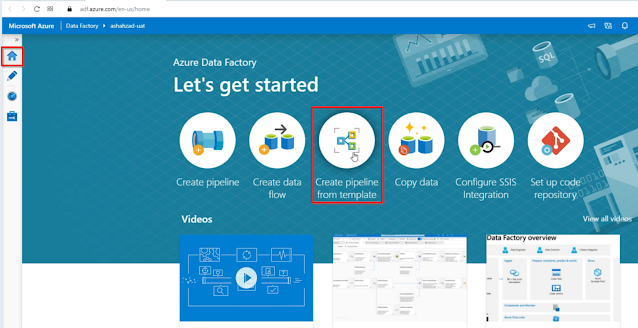
Once our Azure Data Factory is created, go to the resource and open then Azure data factory studio, in the Azure data factory studio we have four major tabs which are 1. Home, 2. Author, 3. Monitor, 4. Manage, in the Home tab we have many more options i.e, Create pipeline, in the pipeline we can create the ETL activities (Extract, transform, and load) A pipeline is a logical grouping of activities that performs a unit of work. Together, the activities in a pipeline perform a task. For example, a pipeline can contain a group of activities that ingests data from an Azure blob and then runs a Hive query on an HDInsight cluster to partition the data Create Data flow Create Pipeline from the template, Copy data, Configure SSIS integration, Set up code repository.

[](https://blogger.googleusercontent.com/img/a/AVvXsEgw_kDXq3SJlhoGatZbt77FzlIxKIC3UUQWJn7jsxnfx5qiDsn4Vl2oD8QoonIm-Z8S4x7wgD5tgrRu-fkSPGw0dss15HXARCkFF12pNJjnCqPHst_YDkPgXL-iqr5geMeI3bUasZfR7Ef3zaV6bfl5Ck7bImH6Ma2aV_7BH0s3dfFveiBr6yy5NykL=s1651)

Data Flow allows you to develop graphical data transformation logic that can be executed as activities within ADF pipelines. The intent of ADF Data Flows is to provide a fully visual experience with no coding required.

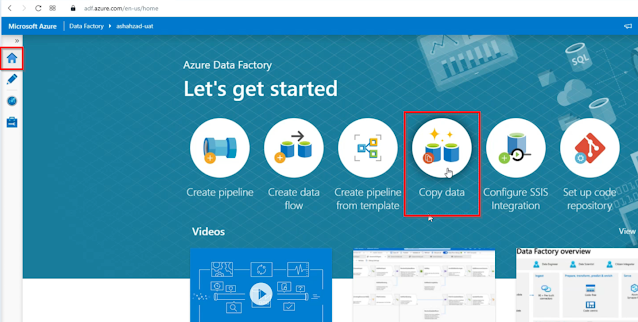
[](https://blogger.googleusercontent.com/img/a/AVvXsEhkg46maZ6jj8DlGNhjyQNBJNPAHFehytMWefeBjmO0MQsM1RBbxtqzZV8WQGuEnaZgXpcNTARAXOgJjL--4mDeoxXeU91T5E0Gd10yTbcpm-f-ZDzbepPL32gi35VJS8fgZhYpcoyNijqcC14lsonkVbYPSlJleh42gH47aitC-3QcoXqnL3zRJfwP=s1616)

In the home Tab, the 3rd option is Create Pipeline from a template, An Azure Data Factory pipeline template is a predefined pipeline that provides you with the ability to create a specific workflow quickly, without the need to spend time in designing and developing the pipeline, using an existing Template Gallery that contains data Copy templates, External Activities templates, data Transformation templates, SQL Server Integration Service template or your custom pipelines templates.

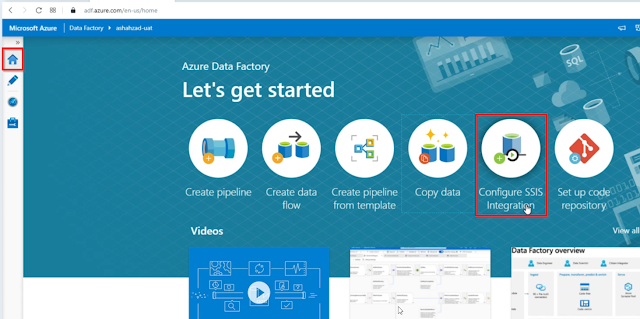
[](https://blogger.googleusercontent.com/img/a/AVvXsEhppR7_ZwQk61_G2ZrpjifSrFAvrITrT8-HoDPj0Hi2pV_qBYEnEjuY3Ny5hj6UtUeU_untQkEJDjZ3yrzRqhbhMmNXwAXYXdv1YsxnVGw6LcQXwMF7xPK1O7eePzFcJ4wEUot6i9bR7wuyNvMEGPYGPlqFftbS2kyqZ9dupPLui7jZoVe03miPK1WN=s1532)

.

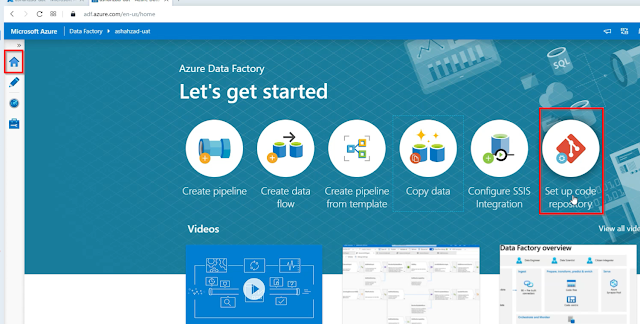
Next is Copy Data, In Azure Data Factory and Synapse pipelines, you can use the Copy activity to copy data among data stores located on-premises and in the cloud. After you copy the data, you can use other activities to further transform and analyze it. You can also use the Copy activity to publish transformation and analysis results for business intelligence and application consumption.

[](https://blogger.googleusercontent.com/img/a/AVvXsEh0u-cEfGyM0nfO48K1cS09PEC11LEGFQOrltHY5jtqxOhVU9W-NVF_l7WPhBDcqJLHy8kD--Kz2bW2w1-95Tl_1ozSzOSFWQcO7SCA3SqvrWCciVeAsYiMqp_pxEPMysZDlgrKBXFDKKh8ODJ7JMEXGHmzFQIu2lT5C38wdyCmtKmbUN9F9xdiDhPJ=s1441)

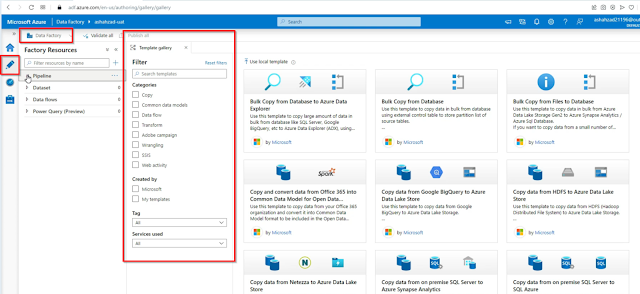
Then we have (Configure SSIS Integration) Data Integration. SSIS performs data integration by combining the data from multiple sources and provides unified data to the users.

[](https://blogger.googleusercontent.com/img/a/AVvXsEjdkXrQhB5yShcyAa1BAYghj2jpwT2q9tyz99M512EAFnDW8V5ukSul4xYROd5gYdpF0HmyJ3bbtrYhBorndsfENst6abzFTvSlSPbmYfh4ZIcAjojAa7cB4JMc5h_g5XBgRm_r0VUvyFDzXSVUORo11ihF5vwEjYYaQJgKgaUeNoqHndog0_61sEdm=s1462)

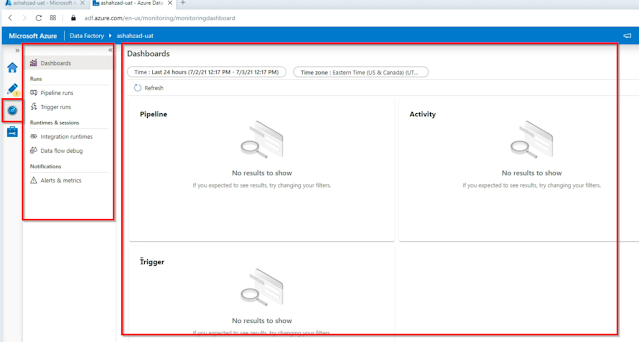
At last, we have Set up a code repository, Your Azure Repos code repository name. Azure Repos projects contain Git repositories to manage your source code as your project grows. You can create a new repository or use an existing repository that's already in your project.

[](https://blogger.googleusercontent.com/img/a/AVvXsEi3MasAufNHvCURGgJ-Mc8Q1mCFYQi6fYRzKcNp3Egrb_Craled7tyVTvUHVGOE5wsZwkZFDNCbaus3Mpzhozxmcfc8GuNq97Ptvoqj5mbtxkMcxaWPqNvJzpT5YUQhdnn61AfnAbYoUsMdYzuoB9kPGDJmkz-poSuCKWRkx3nMrch0uJ9ivgElNmub=s1501)

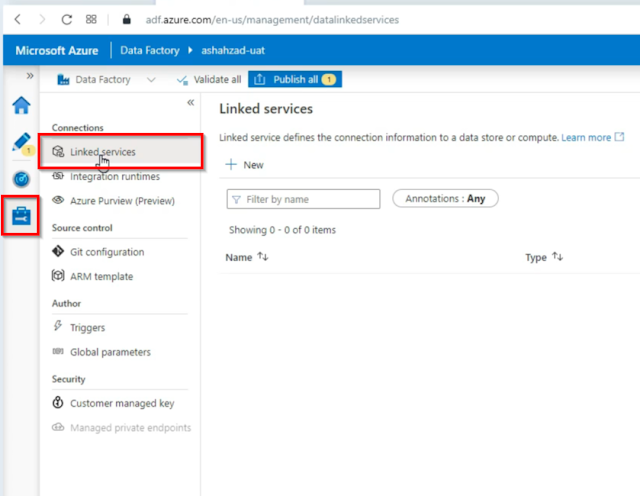
Then we have the second tab which is Author, in this tab we can create pipelines and then we have Dataset, then Data flows, and then power Query, in this tab we can perform tons of activities.

[](https://blogger.googleusercontent.com/img/a/AVvXsEgt53jR_71PX-oFKbAkH_C2oR_hzaNGb2CvfGdlNlRZcdQceD2fZdkTDvPyc7TcozljZxpQPWO6WSww6xQDeba71HMjCZUslqxK8PtM1cg8m_O9cVMllfMwPyVfk1PV7T88jxBHRsylXEMVpvdXuf8mUQjnUSVwSpEJhjL4RgiNJSiyWGIUwTU4m-Eo=s1788)

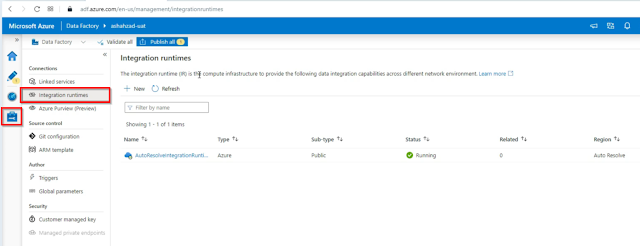
Then we have the 3rd tab which is Monitor, in this tab you can monitor the pipeline that has been executed in your respective azure data factory account can be seen here. Here you have the filter available using which you can filter the pipeline run which you want to monitor. You can [filter](https://azurelib.com/how-do-you-use-filter-activity-in-azure-data-factory/) based on pipeline name, execution time, [triggered](https://azurelib.com/create-triggers-in-adf/)by and many other filter criteria and see the performance of your Pipelines, Activities, and trigers,

[](https://blogger.googleusercontent.com/img/a/AVvXsEi1NeT2Bw6jUF6N2YOsH3cSx2auyTsN5inOgMREd3LTe5-IlS6xnVYgAnanG9M4CnhZ7F0-UfvPd6gEl42UyXDbMNDNBHlYLqzxJuLblagaBhCtVX9Lk8sQpNIwJhbxCCea7DLpVhthK8ilTZHhwfZpSaDipSu780zRIMPJrAS_cDx2WLUy7mb5UZD0=s1446)

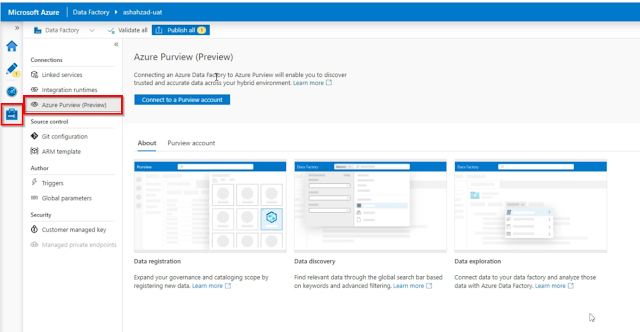
Then we have our 4th and final tab which is Manage, in the manage tab, first of all, we have Linked services, linked services are the connection to your blob or Azure SQL server or any other data source.

[](https://blogger.googleusercontent.com/img/a/AVvXsEiY8kmoWPPoAcxIZNbLY7e7mVnOX-Vn3lR841yr6hIp8JP7W1uy4iDIufQUxL5XvqEtxwR0IOoRgcoXvt9ID9cAk49xK1Sx6ijMo05Efc8AVprJ0JxzKMB0CDSmpd3VLSY0iByxYszBQkGXMjPVPbwm7HKi2QueNcJkpbRj_igfe2rGIWbZTFG56BaO=s839)

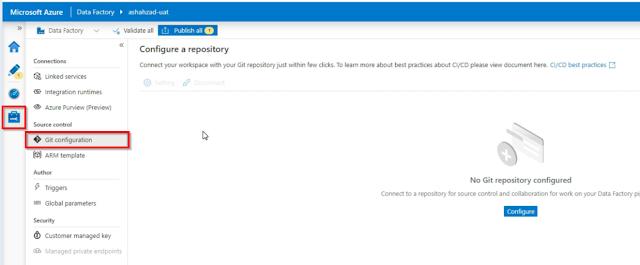
Then we have integration runtime, The Integration Runtime (IR) is the compute infrastructure used by Azure Data Factory and Azure Synapse pipelines to provide the following data integration capabilities across different network environments.

[](https://blogger.googleusercontent.com/img/a/AVvXsEj0Lh1xd8X6nAtISVAap-Qxs0AbrxaXwnVj_0J_v7WVzSpqRe3dwdlh_zivOfGACJyNKFna4uJg3nijvC4mY_5WmZE8incpV4ieU8piwBcY8p_OEZMAWpNGalT--SUzAELD-ZrmG4bnjIK034nmvDjNXbT-JEKm_v8N8PguA3bB4fnwBFRT8iE9jnn4=s1531)

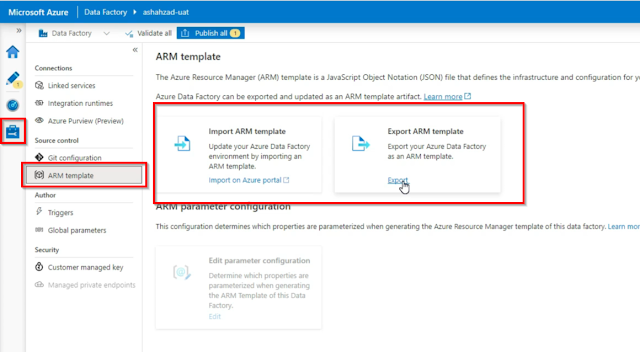
Then we have Azure Purview (Preview), Connection in Azure Data factory, to Azure Purview will enable you to discover trusted and accurate data across your hybrid environment.

[](https://blogger.googleusercontent.com/img/a/AVvXsEhYtVTvRe1pqReCwpN6qx4w25VV81BKvWwdkeIH0nKcEUUqxYhJ0WnOGXkE3VdtL8Z_NZsxibHYy49cmFX9eK78RtMOOMs2Nqh_tt1Y0Td16Q7PjRZP7gpRWoWfiLuBHVxRrTYGLW3ahwBg6nUru8DPTJ17fURBLUQvi01TyVLnDQmIMKjobvNmGQ2D=s1359)

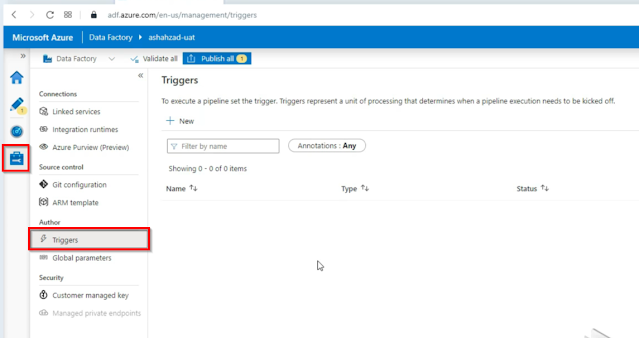
Under the Source control tab, we have Git Configuration, It is a very important tool, when you are going to create your Azure data factory, it is recommended to configure your Github because if in case you deleted your Azure data factory then your all pipelines, datasets and linked services whatever associated with your Azure data factory will be removed, now if you have configured your Github repository, you can easily recover your Azure data factory.

[](https://blogger.googleusercontent.com/img/a/AVvXsEgm-gG7w76052KKeG8vnscf3ngpEOoCDRf8Icrxj1O9YNoFnu1pAEOxgApaahcF94JrS_NuGVoEzb_nzNn3sS7ftUnaGohvKfQmMB9dLpS5pQ7NU2G5tTN-mjeOVy6ad263BP9iU8-oEWLZzUUEJqeg1IRB3YCjRgMlvwYkdQUrJsAGpVhwTZSMBSaJ=s1322)

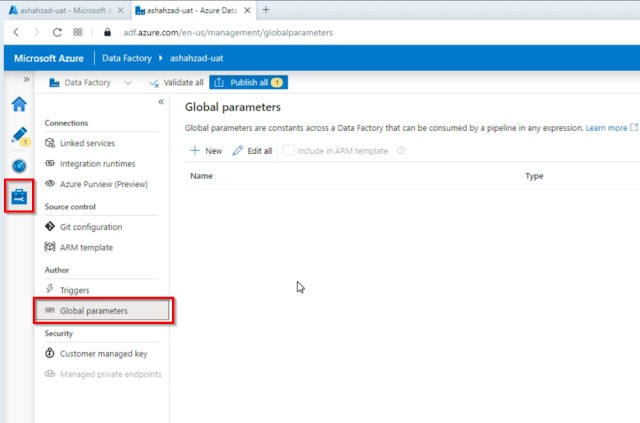
Then we have the ARM template, again it is a very important tool, in this tool you can export (Backup) your entire Data Factory, or you can Import your Backed up data factory.

[](https://blogger.googleusercontent.com/img/a/AVvXsEj7FiSQ0F1I5nxjxm_Z1Hapwr1lfvdvxm5fmDwVINARdulwZLQY51naSIH5oct3kXIRpWJctieIjtgeDunUnY2udYydxkZG_BD4SJ9HS0sBLiAz-Pz8HkJ038RVQ5GUoa9FI2m5SHLZZ1v8rZvJzutMHg5HkgOEDbIyDxwch9SVPHi4mavGfO_hVds1=s1160)

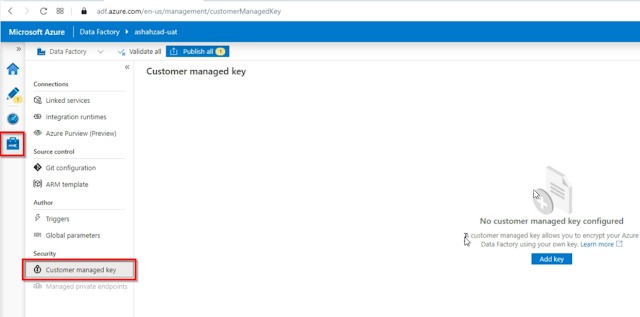
In the Author, we have Triggers, as you know the triggers are the objects that can run your pipelines.

[](https://blogger.googleusercontent.com/img/a/AVvXsEjXKH3qnhqLjwiS8d9XcYfOaWbBuXk-txWm9jCoiRLaZTmQJTXSosR2na8UWUvQB1zFb2M67D49-6_zr_jQyG9KgXcK2ydBFemrJwKmpZdTP4oIBQuYMW-NgVUN2M7KU0926H5H2j2UOQEUL0Y_r_DT8H0raMYa66NFqGz6U1oz_HsDIlOJtorSj3F9=s1145)

Then we have global parameters, Global parameters are constants across a data factory that can be consumed by a pipeline in any expression. They're useful when you have multiple pipelines with identical parameter names and values.

[](https://blogger.googleusercontent.com/img/a/AVvXsEgcJAaVObcL9UcVE60GvKHFcm-LYzJdqDUWEm26UXqINVqjQCTe5KXx_ZbLVB2oivNDPZzkoOuCzx18slyNGFRGbrDyUkUZiuIeiILoiT0Oo-_be74hOHhTeogf-GTnlIyRlhoMdn0w-rybj-JhcNOaOIfaX7eBkWreFwmlY58fJLDJRNuwjMHqaYbb=s1001)

Inside the Security tab, we have the customer-managed key, When you specify a customer-managed key, Data Factory uses both the factory system key and the Customer Managed Key to encrypt customer data. Missing either would result in Deny of Access to data and factory. Azure Key Vault is required to store customer-managed keys.

[](https://blogger.googleusercontent.com/img/a/AVvXsEixtQ07xFKhFTbv0S7aQoegZK4-DhDlfUloAy9eX5kXQd_j8SXPOgobZQqs2WG1SwdwKp8HRnrBqc3sVwe1M0lgzl8N-kK9DEhZt49oOVdwB-VNqkbv48ndz_ARBbTdCsdmUuFDZNWlkVY9_I4LX7Wo_wHwL6vmd7B61DD9GKQH4mKrj12zYW-Cj9Tf=s1243)

