**Azure POC 2 (Azure Data Factory)**

Scenario – We have some data coming from REST API. Transform the data into proper format and store into Azure Data Lake Storage account.

API - <https://reqres.in/api/users>

API data -

{"page": 1,"per\_page": 6,"total": 12,"total\_pages": 2,"data": [{"id": 1,"email": "george.bluth@reqres.in","first\_name": "George","last\_name": "Bluth","avatar": "https://reqres.in/img/faces/1-image.jpg"},

{"id": 2,"email": "janet.weaver@reqres.in","first\_name": "Janet","last\_name": "Weaver","avatar": "https://reqres.in/img/faces/2-image.jpg"},

{"id": 3,"email": "emma.wong@reqres.in","first\_name": "Emma","last\_name": "Wong","avatar": "https://reqres.in/img/faces/3-image.jpg"},

{"id": 4,"email": "eve.holt@reqres.in","first\_name": "Eve","last\_name": "Holt","avatar": "https://reqres.in/img/faces/4-image.jpg"},

{"id": 5,"email": "charles.morris@reqres.in","first\_name": "Charles","last\_name": "Morris","avatar": "https://reqres.in/img/faces/5-image.jpg"},

{"id": 6,"email": "tracey.ramos@reqres.in","first\_name": "Tracey","last\_name": "Ramos","avatar": "https://reqres.in/img/faces/6-image.jpg"}],

"support": {"url": "https://contentcaddy.io?utm\_source=reqres&utm\_medium=json&utm\_campaign=referral","text": "Tired of writing endless social media content? Let Content Caddy generate it for you."}}

Storage account



1. Create a container to store the API data.

A screenshot of a computer

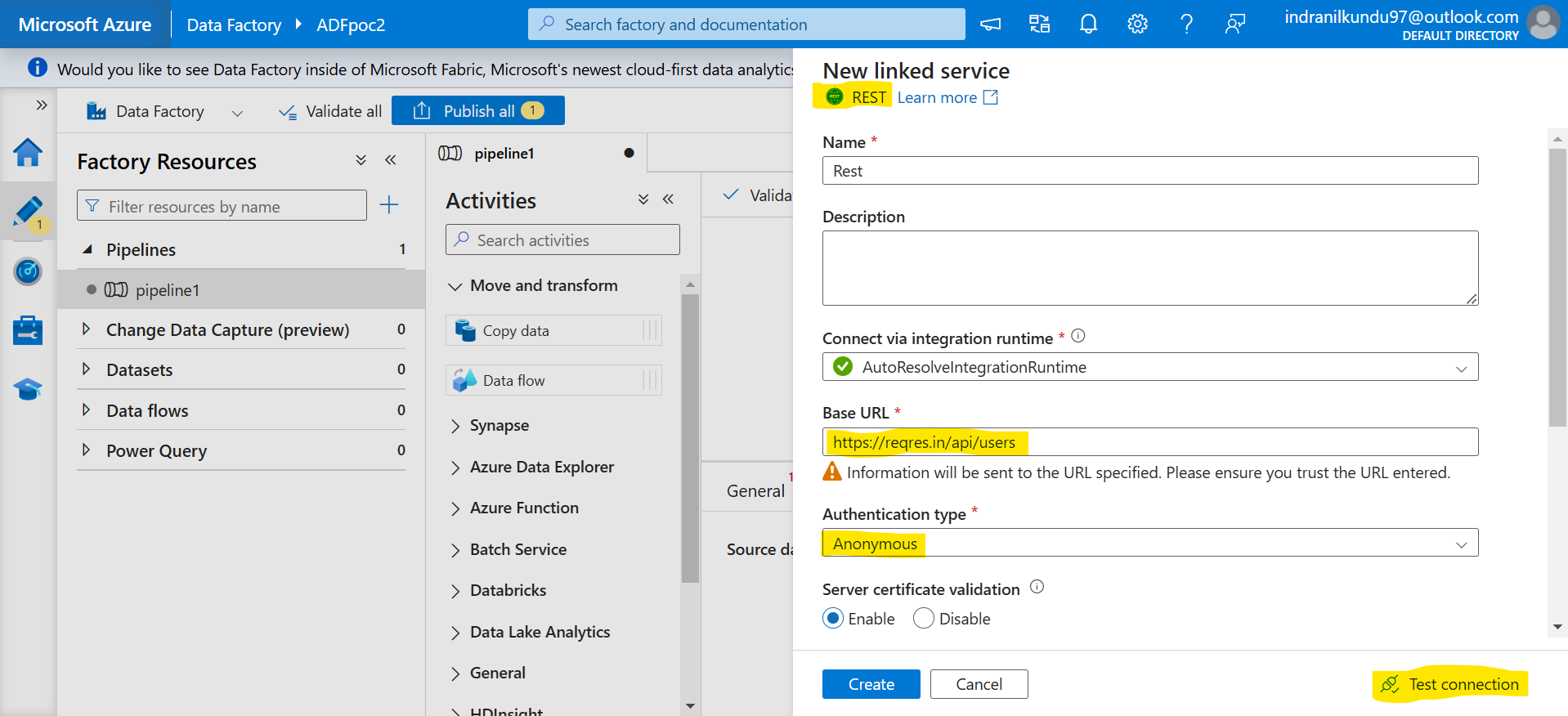
Description automatically generated

1. Create a pipeline in azure data factory and select Copy data activities.

A computer screen shot of a computer screen

Description automatically generated

In Source select the Source dataset as REST.

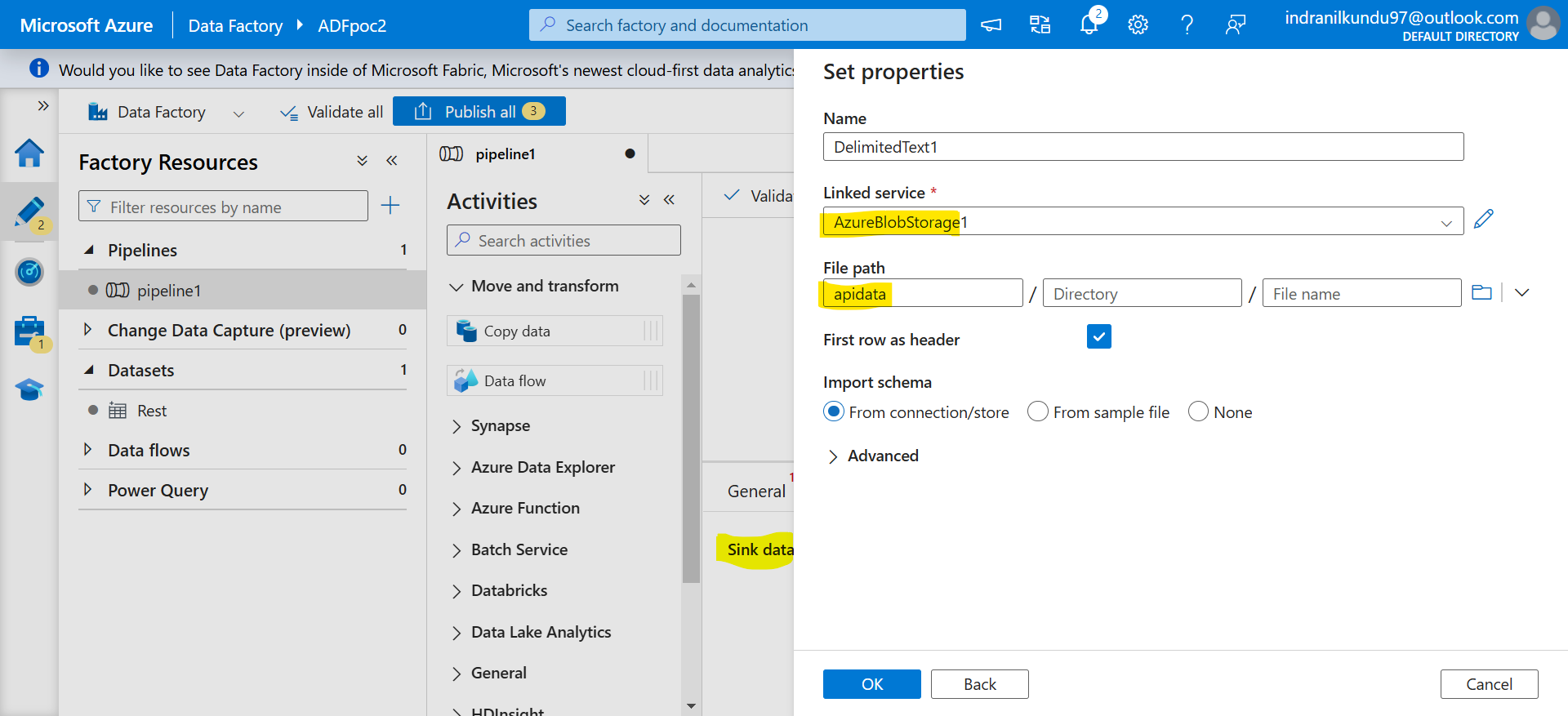


A screenshot of a computer

Description automatically generated

You can click on Preview data to see the data is coming or not.

In Sink select Sink dataset as Blob storage & format CSV type and select the folder file path where you want to store the data.



Go to Mapping -> select Collection reference as $[‘data’] -> select Import schemas -> select which columns you want to keep and delete others.

A screenshot of a computer

Description automatically generated

1. Click on Validate all -> Publish all -> Trigger Now to start the pipeline -> Monitor the pipeline.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. Check the container that the required data is copied or not.

