

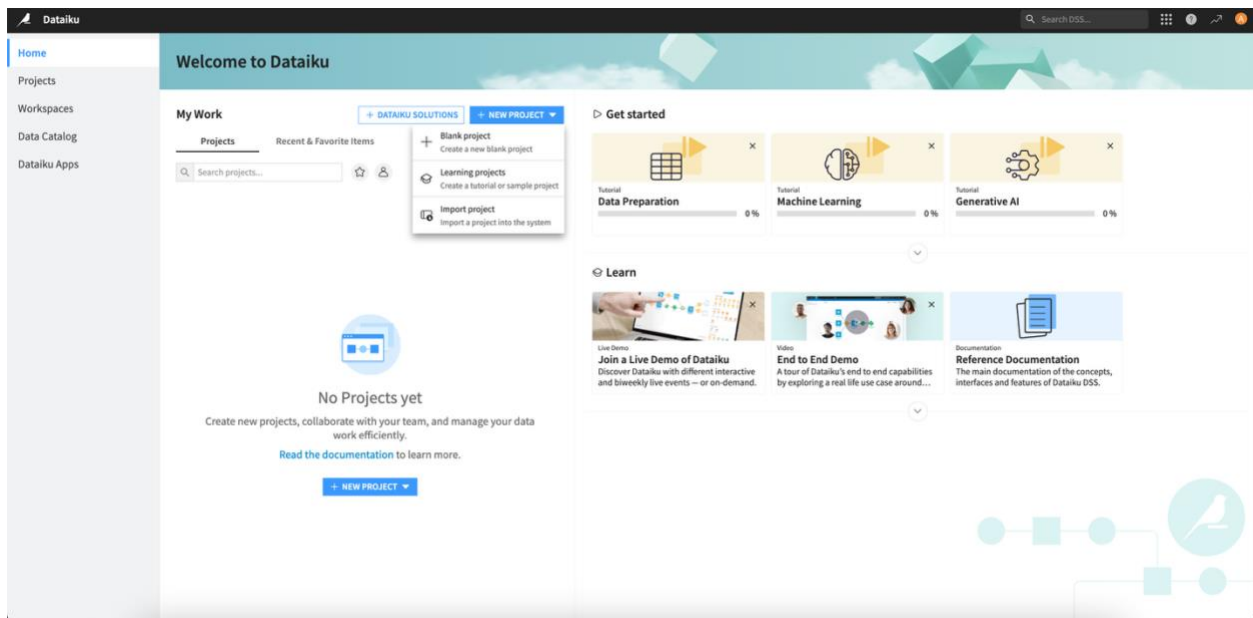
Trần Tấn Phát

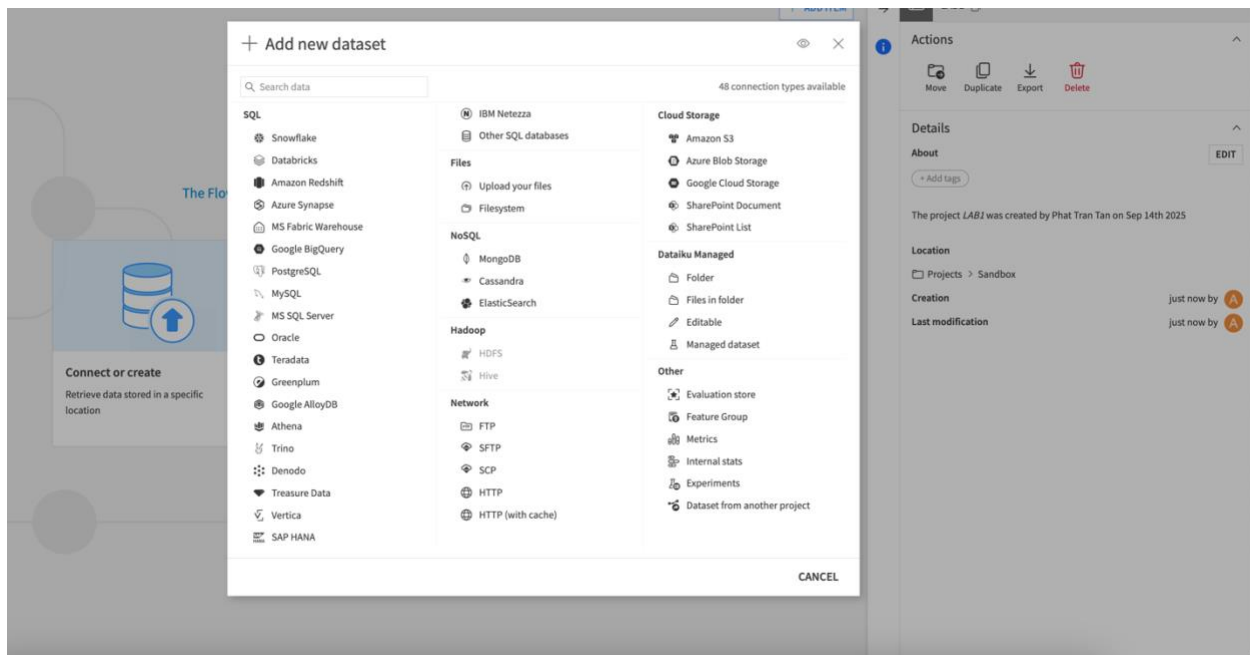
2274802010644

## LAB1: KHÁI NIỆM CƠ BẢN

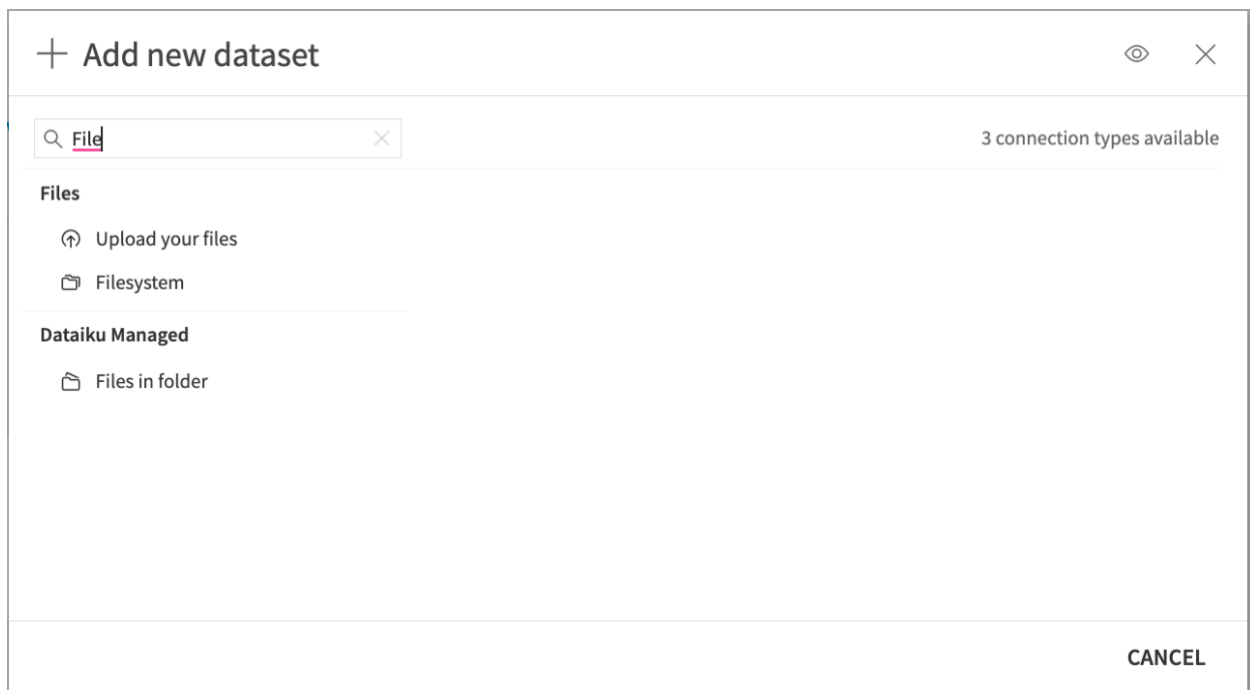
### Bài làm

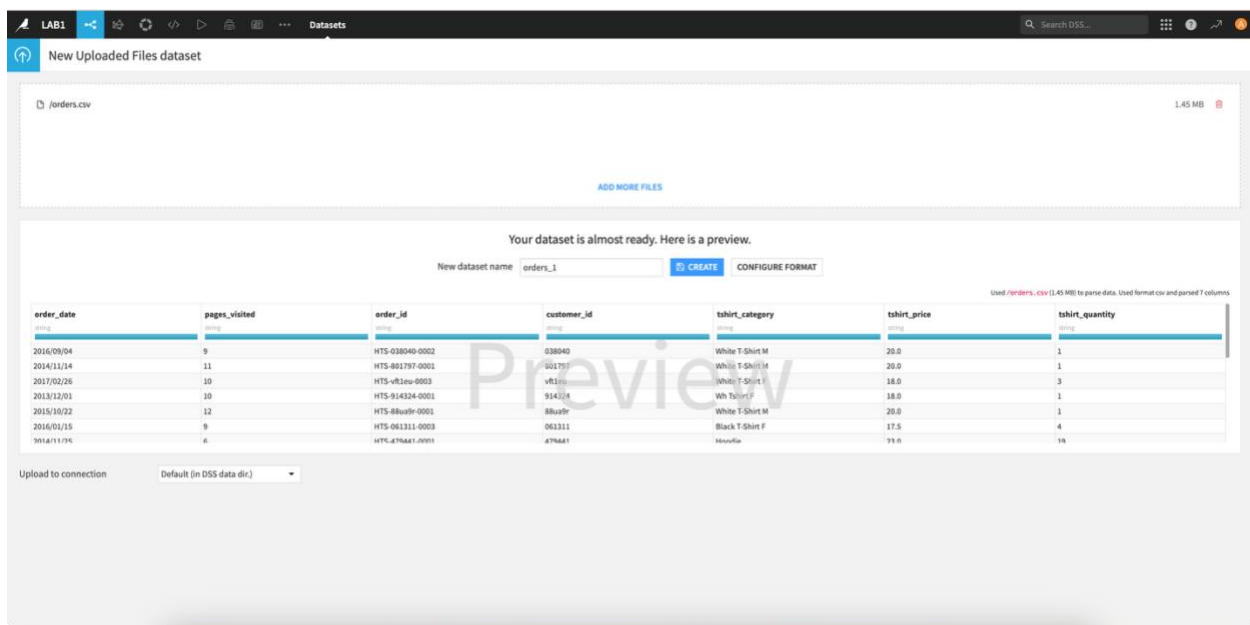
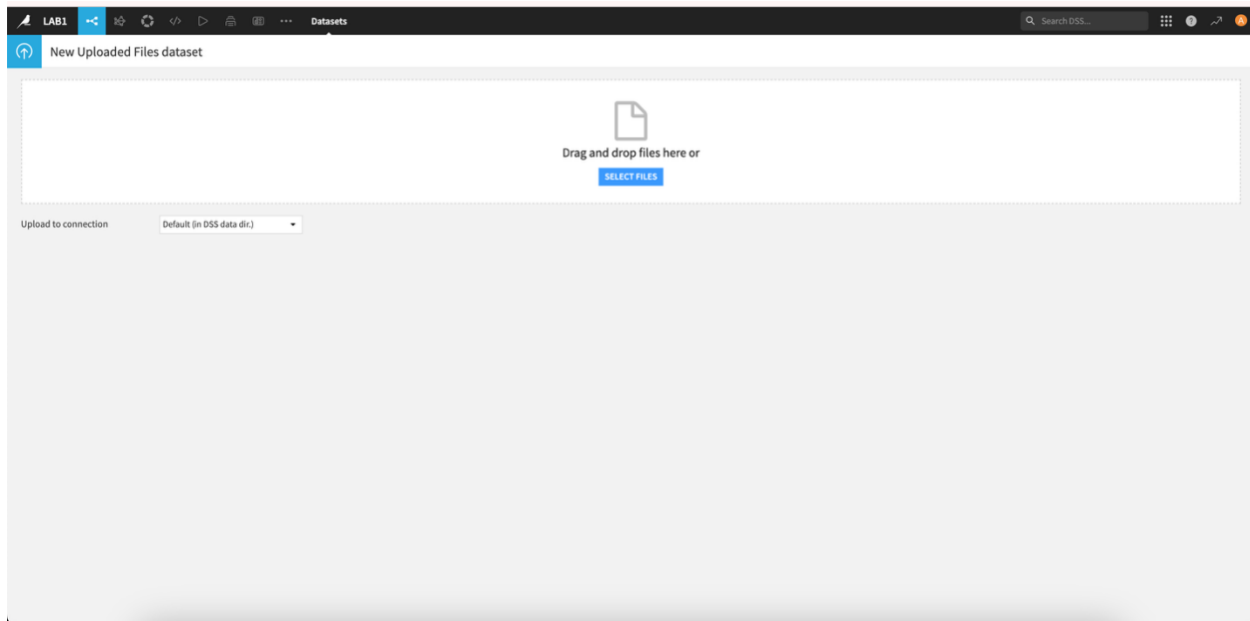
#### 1. Tạo dự án mới





## 2. Đưa dữ liệu vào hệ thống





### 3. Dữ liệu sau khi hoàn thành tạo và xử lý

LAB1

Wh Tshirt F

orders\_1

Explore Charts Statistics Data Quality Metrics History Settings ACTIONS

Sample First 10,000 rows out of not computed

DISPLAY TABLE COLUMNS

16,000 rows 7 columns

order_date	pages_visited	order_id	customer_id	tshirt_category	tshirt_price	tshirt_quantity
string Date only	string Integer	string Text	string Integer	string Text	string Decimal	string Integer
2013/12/01	10	HTS-914324-0001	914324	Wh Tshirt F	18.0	1
2015/10/22	12	HTS-88u9r-0001	88u9r	White T-Shirt M	20.0	1
2016/01/15	9	HTS-061311-0003	061311	Black T-Shirt F	17.5	4
2014/11/25	6	HTS-479441-0001	479441	Hoodie	23.0	19
2014/11/01	10	HTS-352809-0001	352809	Black T-Shirt M	19.0	2
2013/11/03	10	HTS-494332-0001	494332	Wh Tshirt F	18.0	2
2013/10/02	14	HTS-463794-0001	463794	Tennis Shirt	24.0	1
2013/07/23	17	HTS-132885-0001	132885	Wh Tshirt M	20.0	1
2013/11/24	17	HTS-519113-0001	519113	Bl Tshirt F	17.5	1
2015/11/24	15	HTS-j557b-0001	j557b	Hoodie	23.0	3
2015/10/10	8	HTS-252675-0002	252675	White T-Shirt M	20.0	3
2016/08/15	6	HTS-049348-0004	049348	Black T-Shirt F	17.5	1
2015/11/29	15	HTS-heb26i-0001	heb26i	Black T-Shirt F	17.5	1
2015/10/09	10	HTS-989355-0003	989355	Hoodie	23.0	2
2015/03/26	11	HTS-679143-0002	679143	Hoodie	23.0	1
2016/04/24	12	HTS-301362-0001	301362	Black T-Shirt F	17.5	1
2016/01/16	15	HTS-1bd0nt-0002	1bd0nt	White T-Shirt M	20.0	1
2015/01/17	5	HTS-198773-0002	198773	White T-Shirt M	20.0	1
2016/08/14	9	HTS-996602-0005	996602	Hoodie	23.0	3
2014/07/20	11	HTS-476084-0001	476084	White T-Shirt M	20.0	2
2016/05/20	14	HTS-907041-0002	907041	White T-Shirt F	18.0	3
2016/01/31	11	HTS-ws13bp-0002	ws13bp	Hoodie	23.0	5
2014/03/09	12	HTS-399607-0002	399607	White T-Shirt M	20.0	5

#### 4. Chọn 100 Nb Record và lưu kết quả

2017/02/22

Wh Tshirt F

orders\_1

Explore Charts Statistics Data Quality Metrics History Settings ACTIONS

Sample First 10,000 rows out of not computed

DISPLAY TABLE COLUMNS

16,000 rows 7 columns

Sample settings

Sampling method  
First records

Nb. records  
100

Filtering  
☐ Filter the rows of the dataset

Advanced  
☒ Limit memory usage  
Max memory usage (MB)  
100  
☐ Auto-refresh sample  
Automatically refresh the sample when the dataset content changes. The check can be costly for large datasets.

UPDATE SAMPLE

order_date	pages_visited	order_id	customer_id	tshirt_category	tshirt_price	tshirt_quantity
string Date only	string Integer	string Text	string Integer	string Text	string Decimal	string Integer
2015/10/01	10	HTS-591805-0002	591805	Hoodie	23.0	1
2014/11/13	8	HTS-567892-0001	567892	White T-Shirt F	18.0	3
2016/11/20	15	HTS-42j6dm-0001	42j6dm	White T-Shirt M	20.0	1
2015/11/05	12	HTS-616743-0002	616743	White T-Shirt F	18.0	23
2016/07/07	5	HTS-157754-0006	157754	Black T-Shirt M	19.0	3
2016/01/26	12	HTS-29hudg-0001	29hudg	Hoodie	23.0	1
2016/01/30	10	HTS-h7fgy9-0002	h7fgy9	Black T-Shirt M	19.0	1
2017/02/22	13	HTS-294598-0002	294598	White T-Shirt F	18.0	1
2016/01/12	9	HTS-wch13p-0001	wch13p	Hoodie	23.0	1
2015/06/17	11	HTS-030666-0001	030666	Black T-Shirt M	19.0	1
2015/10/10	10	HTS-27jppi-0001	27jppi	Black T-Shirt M	19.0	5
2016/02/26	12	HTS-024475-0003	024475	Tennis Shirt	24.0	13
2015/12/14	10	HTS-873855-0002	873855	Black T-Shirt M	19.0	1
2016/02/26	14	HTS-603814-0002	603814	White T-Shirt F	18.0	2
2014/11/04	11	HTS-332602-0002	332602	Hoodie	23.0	3
2015/01/05	8	HTS-538532-0001	538532	Black T-Shirt F	17.5	1
2016/04/29	9	HTS-rs9rnk-0001	rs9rnk	White T-Shirt M	20.0	1
2014/07/17	8	HTS-749298-0001	749298	Black T-Shirt F	17.5	2
2016/08/03	15	HTS-4nucnh-0002	4nucnh	White T-Shirt M	20.0	1
2014/09/28	14	HTS-437019-0001	437019	Hoodie	23.0	4
2017/02/05	10	HTS-759800-0003	759800	Hoodie	23.0	11
2013/08/22	14	HTS-240612-0001	240612	Bl Tshirt F	17.5	2
2013/10/23	14	HTS-783774-0001	783774	Hoodie	23.0	10

#### 5. Biểu đồ sau khi phân tích dữ liệu

