```python
In [1]:  import pandas
         import datetime
         import pytz
```

```python
In [2]:  now = datetime.datetime.now(tz=pytz.timezone("Asia/Kuala_Lumpur"))
         json_file = f"./data/{now:%Y-%m-%d}.json"
```

```python
In [3]:  data = pandas.read_json(json_file)
         # remove duplicates
         cleaned_data = data[~data.duplicated(subset="id")]
         cleaned_data.head(5)
```

Out[3]:

	type	id	attributes	links
0	ads	105113750	{'date': '2024-02-05 02:16:50', 'modified_ts':...	{'image_baseurl': 'https://img.rnudah.com/imag...
1	ads	105069635	{'date': '2024-02-05 02:16:50', 'modified_ts':...	{'image_baseurl': 'https://img.rnudah.com/imag...
2	ads	105554382	{'date': '2024-02-05 02:15:01', 'modified_ts':...	{'image_baseurl': 'https://img.rnudah.com/imag...
3	ads	105185585	{'date': '2024-02-05 02:13:06', 'modified_ts':...	{'image_baseurl': 'https://img.rnudah.com/imag...
4	ads	104372697	{'date': '2024-02-05 02:12:55', 'modified_ts':...	{'image_baseurl': 'https://img.rnudah.com/imag...

```python
In [4]:  pandas.json_normalize(cleaned_data["attributes"])
```

Out[4]:

	date	modified_ts	building_id	type_name	category_name	lis
0	2024-02-05 02:16:50	1707070610	03d25073-7dd2-4b22-8c09-8e14ac8ce5b8	let	Apartment / Condominium	105113
1	2024-02-05 02:16:50	1707070610	fa8a5fa4-fc50-4942-87fe-8f1abe605f1e	let	Apartment / Condominium	105069
2	2024-02-05 02:15:01	1707070501	NaN	sell	House	105554
3	2024-02-05 02:13:06	1707070386	NaN	sell	House	105185
4	2024-02-05 02:12:55	1707070375	NaN	sell	House	104372
...	...	...	...	...	...	
1094	2024-02-05 11:06:31	1707102391	NaN	sell	Commercial Property	97396
1095	2024-02-05 11:06:23	1707102383	NaN	let	Room	105487
1096	2024-02-05 11:06:19	1707102379	ee9a4ae2-60df-11ee-94fb-5f56d832f7f0	sell	Apartment / Condominium	105557
1097	2024-02-05 11:06:17	1707102377	c8975da3-79fd-4f41-b80c-347d11cec992	sell	Apartment / Condominium	105457
1098	2024-02-05 11:06:14	1707102374	3138a7a3-6e37-41d4-94fc-294ca6e2b5a4	sell	Apartment / Condominium	101764

1099 rows × 55 columns

In [5]:
```python
merged_data = pandas.concat([pandas.json_normalize(cleaned_data["attributes"
                            pandas.json_normalize(cleaned_data["links"])], a
                            )
merged_data
```

	date	modified_ts	building_id	type_name	category_name	lis
0	2024-02-05 02:16:50	1707070610	03d25073-7dd2-4b22-8c09-8e14ac8ce5b8	let	Apartment / Condominium	105113
1	2024-02-05 02:16:50	1707070610	fa8a5fa4-fc50-4942-87fe-8f1abe605f1e	let	Apartment / Condominium	105069
2	2024-02-05 02:15:01	1707070501	NaN	sell	House	105554
3	2024-02-05 02:13:06	1707070386	NaN	sell	House	105185
4	2024-02-05 02:12:55	1707070375	NaN	sell	House	104372
...	...	...	...	...	...	
1094	2024-02-05 11:06:31	1707102391	NaN	sell	Commercial Property	97396
1095	2024-02-05 11:06:23	1707102383	NaN	let	Room	105487
1096	2024-02-05 11:06:19	1707102379	ee9a4ae2-60df-11ee-94fb-5f56d832f7f0	sell	Apartment / Condominium	105557
1097	2024-02-05 11:06:17	1707102377	c8975da3-79fd-4f41-b80c-347d11cec992	sell	Apartment / Condominium	105457
1098	2024-02-05 11:06:14	1707102374	3138a7a3-6e37-41d4-94fc-294ca6e2b5a4	sell	Apartment / Condominium	101764

1099 rows × 57 columns