```
In [1]: import pandas
         import datetime
         import pytz
In [2]: now = datetime.datetime.now(tz=pytz.timezone("Asia/Kuala Lumpur"))
         json file = f"../data/{now:%Y-%m-%d}.json"
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
                                   {'date': '2024-02-05
                                                                          {'image baseurl':
         0
              ads 105113750
                                             02:16:50',
                                                             'https://img.rnudah.com/imag...
                                        'modified_ts':...
                                   {'date': '2024-02-05
                                                                          {'image baseurl':
         1
              ads 105069635
                                             02:16:50',
                                                             'https://img.rnudah.com/imag...
                                        'modified ts':...
                                   {'date': '2024-02-05
                                                                          {'image baseurl':
              ads 105554382
         2
                                             02:15:01',
                                                             'https://img.rnudah.com/imag...
                                        'modified ts':...
                                   {'date': '2024-02-05
                                                                          {'image baseurl':
         3
              ads 105185585
                                             02:13:06'.
                                                             'https://img.rnudah.com/imag...
                                        'modified ts':...
                                   {'date': '2024-02-05
                                                                          {'image baseurl':
              ads 104372697
         4
                                             02:12:55',
                                                             'https://img.rnudah.com/imag...
                                        'modified ts':...
In [4]: pandas.json normalize(cleaned data["attributes"])
```

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	
1	1094	2024- 02-05 11:06:31	1707102391	NaN	sell	Commercial Property	97396	
1	1095	2024- 02-05 11:06:23	1707102383	NaN	let	Room	105487	
1	1096	2024- 02-05 11:06:19	1707102379	ee9a4ae2- 60df-11ee- 94fb- 5f56d832f7f0	sell	Apartment / Condominium	105557	
1	1097	2024- 02-05 11:06:17	1707102377	c8975da3- 79fd-4f41- b80c- 347d11cec992	sell	Apartment / Condominium	105457	
1	1098	2024- 02-05 11:06:14	1707102374	3138a7a3- 6e37-41d4- 94fc- 294ca6e2b5a4	sell	Apartment / Condominium	101764	
1000 55								

1099 rows × 55 columns

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

1099 rows × 57 columns