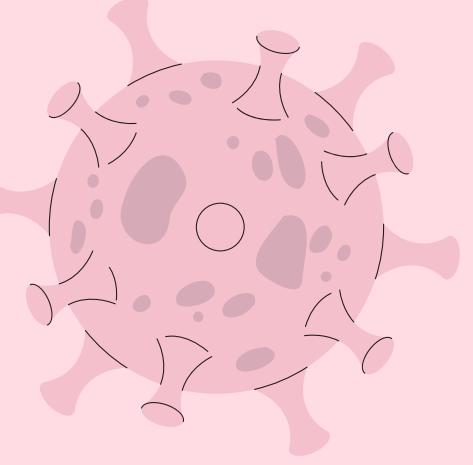


Project

นำเสนอด้วย

กลุ่มสาม สามชาย

โควิด และการต่อต้าน



รายชื่อ

นางสาวชล德拉ดา คำริวงศ์ไพบูลย์ 623020434-3 STAT
นางสาวนวลแพร พนาวัฒนวงศ์ 623020525-0 SI
นายอภิวัฒน์ เหลี่ยมสิงห์ 623020545-4 SI

แบบนำตารางที่นำมาใช้

ตารางรวมผู้ป่วยโควิด-19

No.	announce_date	Notified date	sex	age	Unit	nationality	province_of_isolation	risk	province_of_onset	district_of_onset	Unnamed: 11	
0	1	12/1/20	NaN	หญิง	61.0	ปี	China	กรุงเทพมหานคร	คนต่างด้าวเดินทางมาจากต่างประเทศ	กรุงเทพมหานคร	Nan	Nan
1	2	17/1/20	NaN	หญิง	74.0	ปี	China	กรุงเทพมหานคร	คนต่างด้าวเดินทางมาจากต่างประเทศ	กรุงเทพมหานคร	Nan	Nan
2	3	22/1/20	NaN	หญิง	73.0	ปี	Thailand	นครปฐม	คนต่างด้าวเดินทางมาจากต่างประเทศ	นครปฐม	เมือง	Nan
3	4	22/1/20	NaN	ชาย	68.0	ปี	China	กรุงเทพมหานคร	คนต่างด้าวเดินทางมาจากต่างประเทศ	กรุงเทพมหานคร	Nan	Nan
4	5	24/1/20	NaN	หญิง	66.0	ปี	China	นนทบุรี	คนต่างด้าวเดินทางมาจากต่างประเทศ	กรุงเทพมหานคร	Nan	Nan
...	
433548	3303165	18/3/22	17/3/22	หญิง	Nan	Nan	Nan	กรุงเทพมหานคร	พัฒนาสถาน/เรือนจำ	Nan	Nan	Nan
433549	3303166	18/3/22	17/3/22	หญิง	Nan	Nan	Nan	กรุงเทพมหานคร	พัฒนาสถาน/เรือนจำ	Nan	Nan	Nan
433550	3303167	18/3/22	17/3/22	หญิง	Nan	Nan	Nan	กรุงเทพมหานคร	พัฒนาสถาน/เรือนจำ	Nan	Nan	Nan
433551	3303168	18/3/22	17/3/22	หญิง	Nan	Nan	Nan	สงขลา	พัฒนาสถาน/เรือนจำ	Nan	Nan	Nan
433552	3303169	18/3/22	17/3/22	หญิง	Nan	Nan	Nan	สงขลา	พัฒนาสถาน/เรือนจำ	Nan	Nan	Nan

3303169 rows x 12 columns



ตารางโรงพยาบาลที่ปิดกิจกรรม และโรงพยาบาลที่จดทะเบียนใหม่แต่ละจังหวัด

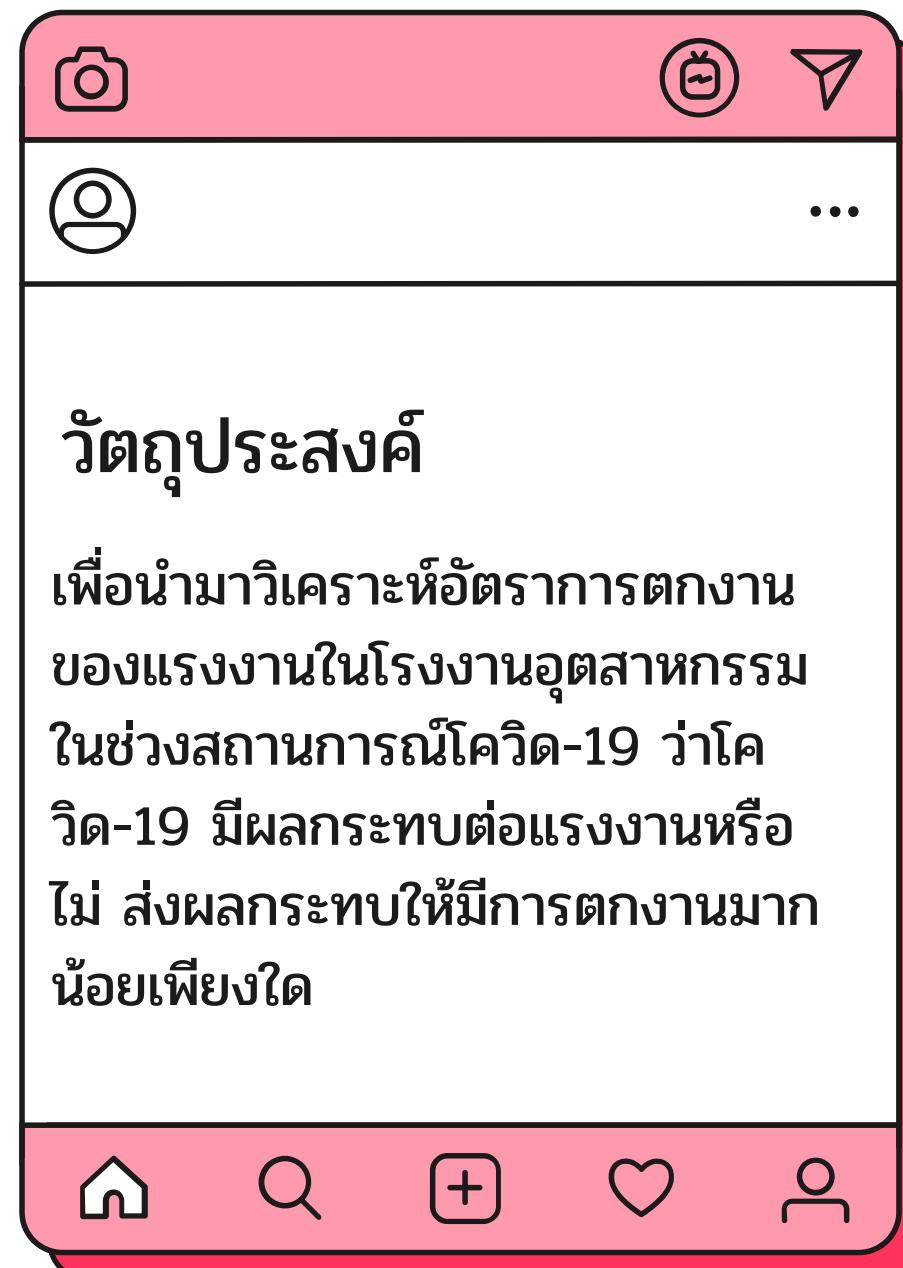
จังหวัด	โรงพยาบาลใหม่	แรงงานชายได้งาน	แรงงานหญิงได้งาน	รวมแรงงานได้งาน	เลิกกิจการ	แรงงานชายต葵งาน	แรงงานหญิงต葵งาน	รวมแรงงานต葵งาน	การคอกงาน	คอกงานผู้ชาย	คอกงานผู้หญิง	
0	กรุงปี	11	231	104	335	11	51	27	78	-257	-180	-77
1	กรุงเทพมหานคร	185	2,583	2,119	4,702	67	1,229	920	2,149	-2,553	-1,354	-1,199
2	กาญจนบุรี	31	325	80	405	5	157	41	198	-207	-168	-39
3	กาฬสินธุ์	26	160	29	189	5	35	8	43	-146	-125	-21
4	กำแพงเพชร	24	218	86	304	3	30	17	47	-257	-188	-69
...
72	อ่านเจริญ	7	46	10	56	-	0	0	-	-56	-46	-10
73	อุดรธานี	21	215	52	267	3	62	21	83	-184	-153	-31
74	อุตรดิตถ์	10	94	58	152	7	25	17	42	-110	-69	-41
75	อุทัยธานี	4	21	0	21	5	21	198	219	198	0	198
76	อุบลราชธานี	39	363	153	516	8	39	2	41	-475	-324	-151

77 rows x 12 columns

ตารางภูมิภาคและจำนวนผู้ติดเชื้อปัจจุบัน

จังหวัด	ภาค	ผู้ติดเชื้อปัจจุบัน	แก้ไข
0	อ่านเจริญ	ตะวันออกเฉียงเหนือ	63
1	บึงกาฬ	ตะวันออกเฉียงเหนือ	111
2	บุรีรัมย์	ตะวันออกเฉียงเหนือ	484
3	ชัยภูมิ	ตะวันออกเฉียงเหนือ	119
4	กาฬสินธุ์	ตะวันออกเฉียงเหนือ	285
...
72	พัชญา	ใต้	94
73	ภูเก็ต	ใต้	357
74	ระนอง	ใต้	232
75	สตูล	ใต้	294
76	ตรัง	ใต้	173

77 rows x 3 columns



ตารางรวมผู้ป่วยโควิด-19 จาก Government Data of Thailand

<https://data.go.th/dataset/covid-19-daily>

ตารางภูมิภาคและจำนวนผู้ติดเชื้อปัจจุบัน จาก วิกิพีเดีย และ กระทรวงสาธารณสุข

<https://th.wikipedia.org/wiki/ภูมิภาคของประเทศไทย> และ

<https://ddc.moph.go.th/covid19-dashboard/>

ตารางงานอุตสาหกรรมที่ปิดกิจการ และงานอุตสาหกรรมที่จดทะเบียนใหม่แต่ละจังหวัด จากรัฐบาล

<https://www.diw.go.th/webdiw/static-fac/?fbclid=IwAROzVelYLrnukaRrrK3CKAOyzMROXw1JoeaimGr6DYoWzdH-crxwrUtePY4>



นำเข้าข้อมูล

ทำการนำเข้าข้อมูลผู้ติดเชื้อโควิดทั้งหมด ตั้งแต่วันแรกจนถึงวันที่ 18 มีนาคม 2565

```
In [1]: import pandas as pd  
import os
```

```
In [2]: from google.colab import drive  
drive.mount('/content/drive')  
path = '/content/drive/My Drive/dataviz_2021_data'
```

Mounted at /content/drive

```
In [3]: %cd /content/drive/MyDrive/dataviz_2021_data
```

/content/drive/MyDrive/dataviz_2021_data

```
In [4]: data_file_path = os.path.join(path,'covid_1.csv') #นำเข้าข้อมูล  
data1 = pd.read_csv(data_file_path)
```

```
In [5]: data_file_path = os.path.join(path,'covid_2.csv') #นำเข้าข้อมูล  
data2 = pd.read_csv(data_file_path)
```

```
In [6]: data_file_path = os.path.join(path,'covid_3.csv') #นำเข้าข้อมูล  
data3 = pd.read_csv(data_file_path)
```

/usr/local/lib/python3.7/dist-packages/IPython/core/interactiveshell.py:2882: DtypeWarning: Columns (4) have mixed types.Specify dtype option on import or set low_memory=False.
exec(code_obj, self.user_global_ns, self.user_ns)

```
In [7]: data_file_path = os.path.join(path,'covid_4.csv') #นำเข้าข้อมูล  
data4 = pd.read_csv(data_file_path)
```

รวมตาราง

ทำการรวมตารางโดยใช้คำสั่ง pd.concat เพื่อรวมตารางผู้ป่วยโควิดเข้าด้วยกันทั้งหมด

In [8]:

```
data5 = pd.concat([data1 , data2 , data3 , data4]) #รวมตาราง  
data5
```

Out[8]:

	No.	announce_date	Notified date	sex	age	Unit	nationality	province_of_isolation
0	1	12/1/20	NaN	หญิง	61.0	ปี	China	กรุงเทพมหานคร
1	2	17/1/20	NaN	หญิง	74.0	ปี	China	กรุงเทพมหานคร
2	3	22/1/20	NaN	หญิง	73.0	ปี	Thailand	นครปฐม
3	4	22/1/20	NaN	ชาย	68.0	ปี	China	กรุงเทพมหานคร
4	5	24/1/20	NaN	หญิง	66.0	ปี	China	นนทบุรี
...
433548	3303165	18/3/22	17/3/22	หญิง	NaN	NaN	NaN	กรุงเทพมหานคร
433549	3303166	18/3/22	17/3/22	หญิง	NaN	NaN	NaN	กรุงเทพมหานคร
433550	3303167	18/3/22	17/3/22	หญิง	NaN	NaN	NaN	กรุงเทพมหานคร
433551	3303168	18/3/22	17/3/22	หญิง	NaN	NaN	NaN	สงขลา
433552	3303169	18/3/22	17/3/22	หญิง	NaN	NaN	NaN	สงขลา

3303169 rows × 12 columns

A

B



ตารางจังหวัด

เลือกคอลัมน์ที่ต้องการใช้

สร้างตารางจังหวัดและผู้ติดเชื้อสะสมปัจจุบัน(18 มีนาคม 2565)
โดยใช้คำสั่ง pd.DataFrame

In [17]:

```
data6 = data5[['province_of_onset']]  
data6
```

Out[17]:

province_of_onset

0 กรุงเทพมหานคร

1 กรุงเทพมหานคร

2 นครปฐม

3 กรุงเทพมหานคร

4 กรุงเทพมหานคร

... ...

433548 NaN

433549 NaN

433550 NaN

433551 NaN

433552 NaN

3303169 rows × 1 columns

ทำการเลือกเอาเฉพาะข้อมูลจังหวัด เพราะเป็นข้อมูลเดียวที่ต้องการใช้ใน
ตารางดังกล่าว

Out[18]:

```
data7= pd.DataFrame.from_records(data,columns=header) ## ตารางจังหวัด  
data7
```

จังหวัด

ภาค

ผู้ติดเชื้อปัจจุบัน

0	สำนักงานเขตใหญ่	ตะวันออกเฉียงเหนือ	63
1	ปีบาก้าฟี	ตะวันออกเฉียงเหนือ	111
2	บูรีรัมย์	ตะวันออกเฉียงเหนือ	484
3	ชัยภูมิ	ตะวันออกเฉียงเหนือ	119
4	กาฬสินธุ์	ตะวันออกเฉียงเหนือ	285
...
72	พัทลุง	ใต้	94
73	ภูเก็ต	ใต้	357
74	ระนอง	ใต้	232
75	สตูล	ใต้	294
76	ตรัง	ใต้	173

77 rows × 3 columns



รวมตาราง

In [19]:

```
data8 = data6.merge(data7, how='left', left_on='province_of_onset', right_on='จังหวัด')  
data8
```

Out[19]:

	province_of_onset	จังหวัด	ภาค	ผู้ติดเชื้อปัจจุบัน
0	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
1	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
2	นครปฐม	นครปฐม	กลาง	642
3	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
4	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
...
3303164	Nan	Nan	Nan	Nan
3303165	Nan	Nan	Nan	Nan
3303166	Nan	Nan	Nan	Nan
3303167	Nan	Nan	Nan	Nan
3303168	Nan	Nan	Nan	Nan

3303169 rows × 4 columns



ใช้คำสั่ง `merge` ในการรวมตาราง โดยทำการใช้คีย์จากตารางด้านซ้าย (`How='left'` ใช้คีย์จากตารางด้านซ้าย `How='right'` ใช้คีย์จากตารางด้านขวา)

โดยเอาข้อมูลตารางซ้ายจนถึง `province_of_onset`
ข้อมูลตารางขวาเริ่มจากจังหวัด

In [32]:

```
print(data8.isnull().any()) ## ตรวจสอบค่า Missing หลังจากการรวมตาราง
print('*'*100)
data9 = data8.dropna() #แก้Missingด้วยการ Dropna
print(data9.isnull().any()) ## ตรวจสอบค่า Missing อีกครั้ง
data9
```

```
province_of_onset      True
จังหวัด                True
ภาค                  True
ผู้ติดเชื้อปัจจุบัน  True
dtype: bool
*****
province_of_onset      False
จังหวัด                False
ภาค                  False
ผู้ติดเชื้อปัจจุบัน  False
dtype: bool
```

Out[32]:

	province_of_onset	จังหวัด	ภาค	ผู้ติดเชื้อปัจจุบัน
0	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
1	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
2	นครปฐม	นครปฐม	กลาง	642
3	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
4	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
...
3303060	นครราชสีมา	นครราชสีมา ตะวันออกเฉียงเหนือ		468
3303062	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
3303063	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
3303064	กรุงเทพมหานคร	กรุงเทพมหานคร	กลาง	3273
3303065	กำแพงเพชร	กำแพงเพชร	กลาง	169

2899473 rows × 4 columns

ตรวจสอบค่า Missing



โดยทำการแก้ค่า Missing ด้วยคำสั่ง dropna
เนื่องจากข้อมูลไม่มีช่วงจังหวัด
หรือไม่ใช่ข้อมูลผู้ติดเชื้อภายในประเทศ



นับจำนวนผู้ติดเชื้อ^{แต่ละจังหวัด}

ทำการนับจำนวนผู้ติดเชื้อแต่ละจังหวัด โดยใช้คำสั่ง
.groupby().size()
และการตั้งชื่อคอลัมน์เป็น^{ยอดจำนวนผู้ติดเชื้อ}

In [41]:

```
data10 = data9.groupby('จังหวัด').size().reset_index(name='ยอดจำนวนผู้ติดเชื้อ')
data10
```

Out[41]:

	จังหวัด	ยอดจำนวนผู้ติดเชื้อ
0	กรุงเทพมหานคร	17631
1	กรุงเทพมหานคร	563732
2	กาญจนบุรี	24319
3	กาฬสินธุ์	18579
4	กำแพงเพชร	13013
...
72	เพชรบุรี	30186
73	เพชรบูรณ์	18350
74	เลย	9922
75	แพร่	5667
76	แม่ฮ่องสอน	4145

77 rows × 2 columns

รวมตาราง

ใช้คำสั่ง .merge ในการรวมตาราง

In [34]:

```
data11 = data10.merge(data7, how='right', left_on='จังหวัด', right_on='จังหวัด')  
data11
```

Out[34]:

	จังหวัด	ยอดจำนวนผู้ติดเชื้อ	ภาค	ผู้ติดเชื้อปัจจุบัน
0	อ่านาจเจริญ	2990	ตะวันออกเฉียงเหนือ	63
1	ปีงกาฬ	4960	ตะวันออกเฉียงเหนือ	111
2	บุรีรัมย์	7812	ตะวันออกเฉียงเหนือ	484
3	ชัยภูมิ	5015	ตะวันออกเฉียงเหนือ	119
4	กาฬสินธุ์	18579	ตะวันออกเฉียงเหนือ	285
...
72	พัทลุง	11206	ใต้	94
73	ภูเก็ต	48894	ใต้	357
74	ระนอง	14683	ใต้	232
75	สตูล	14173	ใต้	294
76	ตรัง	22728	ใต้	173

77 rows × 4 columns



นำเข้าข้อมูล

ทำการนำเข้าข้อมูลตารางในงานอุตสาหกรรมกีปิดกิจการ และในงานอุตสาหกรรมกีจดทะเบียนใหม่แต่ละจังหวัด

```
In [35]:  
data_file_path = os.path.join(path,'การตั้งงาน.csv') #นำเข้าข้อมูล  
data12 = pd.read_csv(data_file_path)
```

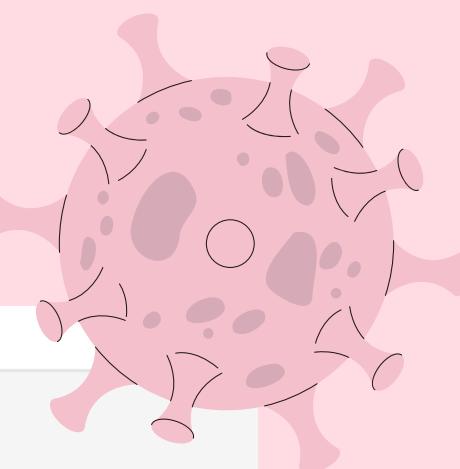
```
In [36]: data12
```

	จังหวัด	โรงงานใหม่	แรงงานชายได้งาน	แรงงานหญิงได้งาน	รวมแรงงานได้งาน	เลิกกิจการ	แรงงานชายตกงาน	แรงงานหญิงตกงาน	รวมแรงงานตกงาน	การตกงาน	ตกงานผู้ชาย	ตกงานผู้หญิง
0	กรุงเทพมหานคร	11	231	104	335	11	51	27	78	-257	-180	-77
1	กรุงเทพมหานคร	185	2,583	2,119	4,702	67	1,229	920	2,149	-2,553	-1,354	-1,200
2	กาญจนบุรี	31	325	80	405	5	157	41	198	-207	-168	-100
3	กาฬสินธุ์	26	160	29	189	5	35	8	43	-146	-125	-100
4	กำแพงเพชร	24	218	86	304	3	30	17	47	-257	-188	-100
...
72	อ่านาจเจริญ	7	46	10	56	-	0	0	-	-56	-46	-46
73	อุดรธานี	21	215	52	267	3	62	21	83	-184	-153	-153
74	อุดรดิตถ์	10	94	58	152	7	25	17	42	-110	-69	-69
75	อุทัยธานี	4	21	0	21	5	21	198	219	198	0	0
76	อุบลราชธานี	39	363	153	516	8	39	2	41	-475	-324	-324

77 rows × 12 columns

รวมตาราง

ใช้คำสั่ง.merge ในการรวมตาราง



In [37]:

```
data13 = data11.merge(data12,how='right',left_on='จังหวัด',right_on='จังหวัด')
data13
```

Out[37]:

	จังหวัด	ยอดจำนวนผู้ติดเชื้อ	ภาค	ผู้ติดเชื้อปัจจุบัน	โรงพยาบาลใหม่	แรงงานชายได้งาน	แรงงานหญิงได้งาน	รวมแรงงานได้งาน	เลิกกิจการ	แรงงานชายตกงาน	แรงงานหญิงตกงาน	รวมแรงงานตกงาน	การตกงาน	ตกงานผู้ชาย	ตกงานผู้หญิง
0	กรุงเทพมหานคร	17631	ใต้	178	11	231	104	335	11	51	27	78	-257	-180	-77
1	กรุงเทพมหานคร	563732	กลาง	3273	185	2,583	2,119	4,702	67	1,229	920	2,149	-2,553	-1,354	-1,199
2	กาญจนบุรี	24319	ตะวันตก	225	31	325	80	405	5	157	41	198	-207	-168	-39
3	กาฬสินธุ์	18579	ตะวันออกเฉียงเหนือ	285	26	160	29	189	5	35	8	43	-146	-125	-21
4	กำแพงเพชร	13013	กลาง	169	24	218	86	304	3	30	17	47	-257	-188	-69
...
72	อำนาจเจริญ	2990	ตะวันออกเฉียงเหนือ	63	7	46	10	56	-	0	0	-	-56	-46	-10
73	อุดรธานี	14260	ตะวันออกเฉียงเหนือ	379	21	215	52	267	3	62	21	83	-184	-153	-31
74	อุตรดิตถ์	9138	เหนือ	40	10	94	58	152	7	25	17	42	-110	-69	-41
75	อุทัยธานี	5710	กลาง	113	4	21	0	21	5	21	198	219	198	0	198
76	อุบลราชธานี	27788	ตะวันออกเฉียงเหนือ	341	39	363	153	516	8	39	2	41	-475	-324	-151

77 rows × 15 columns

Rate



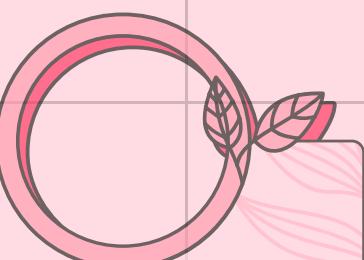
ตรวจสอบค่า Missing

เมื่อไม่มีค่า Missing แล้วจะสามารถนำข้อมูลไปใช้งานได้

In [38]:

```
print(data13.isnull().any()) ## ตรวจสอบค่า Missing หลังจากการดึง  
print('*'*100)
```

จังหวัด	False
ยอดจำนวนผู้ติดเชื้อ	False
ภาค	False
ผู้ติดเชื้อปัจจุบัน	False
โรงพยาบาล	False
แรงงานชายได้งาน	False
แรงงานหญิงได้งาน	False
รวมแรงงานได้งาน	False
เลิกกิจการ	False
แรงงานชายตกงาน	False
แรงงานหญิงตกงาน	False
รวมแรงงานตกงาน	False
การตกงาน	False
ตกงานผู้ชาย	False
ตกงานผู้หญิง	False
dtype: bool	

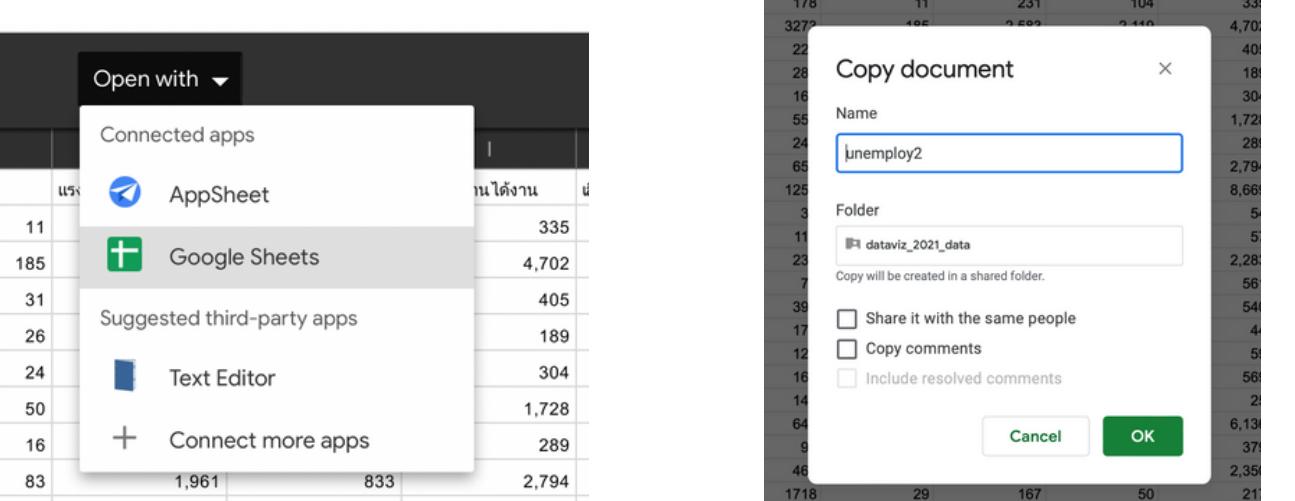
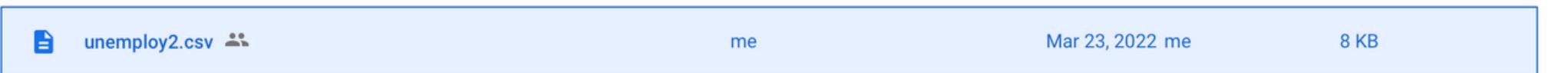


★★★

บันทึกตาราง

```
data13.to_csv('unemploy2.csv', encoding='utf-8')
```

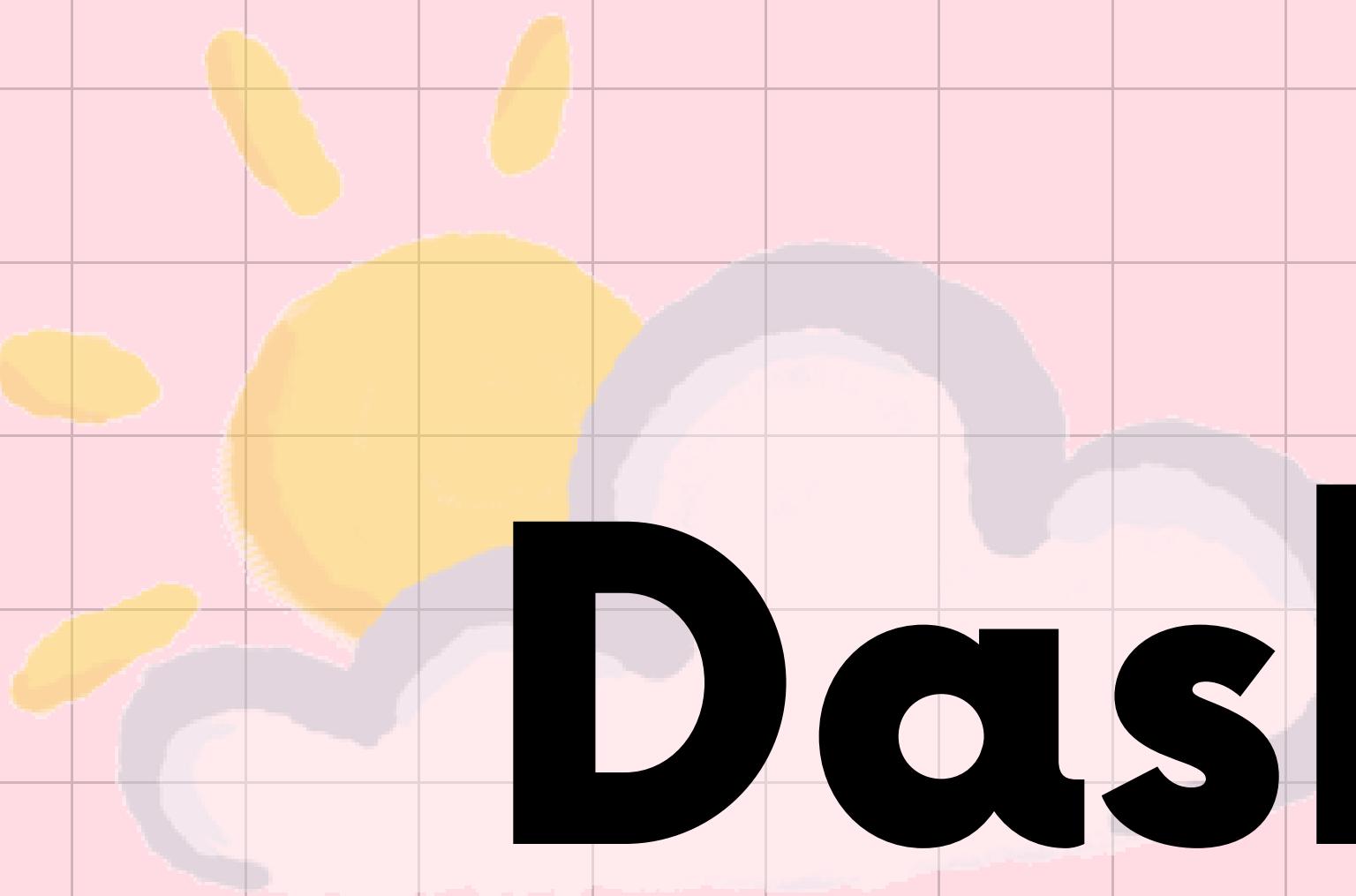
ทำการบันทึกตาราง เป็นไฟล์ CSV ลงในไดฟ์



นำไฟล์ CSV ในไดฟ์มาทำการบันทึกเป็น
Google Sheets ดังภาพ

นำไฟล์เข้า Google Studio โดยนำเข้าผ่านทางการใช้
ไฟล์ Google Sheets





Dashboard