

จัดทำโดย นายใจชนน์สก์ อดุลรุ่งเรือง 61070319

Task 1: อธิบายที่มาของ api ทั้งหมดที่ใช้

ข้อมูลที่นำมาใช้ในงานนี้ <https://edw-opendata.moi.go.th/api/request/list> โดยขออุปกรณ์สำหรับการค้นหาข้อมูลและคุณภาพงานค้านางงานค้านายความใส่ใจในการนี้จะเป็นไปตามที่ระบุไว้ในหน้าเว็บไซต์นี้

- ชื่อยอดผู้แต่งงานและแยกเพศ จำนวนคุณภาพ พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_POPULATION_MARRIAGE.
- สถิติการหย่าและแยกเพศ จำนวนคุณภาพ พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_POPULATION_DIVORCE.
- จำนวนและสัดส่วนประชากรจากการเมืองตามวัย (วันเด็ก, วันแรงงาน, วันสุกราช) และเพศ พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_POPULATION_AGE_GROUP.
- จำนวนผู้มีงานทำที่ถูกหักในห่วงงานในรอบแพนดะและการบันทึก จำนวนคุณภาพ เพศ เป็นรายภาค และจังหวัด พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_LABOUR_FORMAL_AND_INFORMAL_EMPLOYMENT.
- จำนวนการเกิดมีชีวิตร่วมเพศและการเกิดมีชีวิตร่วมเพศพยาบาลคุณภาพ พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_HEALTH_LIVEBIRTHS.
- จำนวนการตายเชิงเพศและการตายไปร่องพยาบาลคุณภาพ เป็นรายภาคและจังหวัด พ.ศ. 2554 - 2561.
http://edw-opendata.moi.go.th/services/apiV_ODS_HEALTH_DEATHS.

เป็นการติดตั้ง urllib3 requests

```
In [1]: pip install urllib3 requests
      pip install --upgrade certifi
Requirement already satisfied: urllib3 in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (1.25.9)
Requirement already satisfied: requests in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (2.24.0)
Requirement already satisfied: certifi>=2017.4.17 in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (from requests) (2020.6.20)
Requirement already satisfied: chardet<4,>=3.0.2 in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (from requests) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (from requests) (2.10)
Requirement already up-to-date: certifi in /Users/kontup/opt/anaconda3/lib/python3.8/site-packages (2020.6.20)
```

การimport Packages ที่ใช้ทั้งหมด

```
In [2]: import urllib
import urllib.request
import json
import requests
import pandas as pd
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
```

Body api ที่ใช้ในการส่งร้องขอข้อมูล

```
In [3]: body={'user_type': 'ประชาชน', 'user_name': 'ใจชนน์ อดุลรุ่งเรือง', 'user_use':1}
```

Task2 ส่วนในการสร้างFunctionในการตักข้อมูลapi และ นำข้อมูล saveเป็นfile.json

```
In [70]: def create_file(data,name_file):#เป็นfunctionในการsaveข้อมูลจากapi ที่ตั้งชื่อเป็นfile.json
    with open('data'+name_file+'.json','w') as file:#เป็นการกำหนดชื่อไฟล์ที่ต้องการจะบันทึก
        print('create '+data+name_file+'.json '+'Completed')
        return (json.dumps(data, file,ensure_ascii=False,indent = 4, sort_keys=True))#เป็นการบันทึกที่ต้องบันทึกไฟล์
def api(api_link,name_file,list_column):#เป็น Function ในการตั้งชื่อ
    lis = []#list ในการเก็บรายชื่อของ api ที่ต้องการจะรับให้ได้เพื่อทำการนัดหมายไปรับข้อมูลในไฟล์
    url = api_link+'&page=1' #เป็น URL ของ api ที่ต้องการจะรับ ที่ต้องการจะรับข้อมูลที่ต้องบันทึกเข้าไปในไฟล์
    response = requests.get(url,headers={'User-Agent': 'Mozilla/5.0'})#เป็นการรับข้อมูลที่ต้องบันทึก
    pastebin_url = response.text#สามารถดูรายละเอียดของ URL ที่ต้องบันทึก
    data_json.loads(pastebin_url[4:])#เป็นการตัดตัวอักษรที่ไม่ต้องการ
    for p in range(len(data['data_object'])["last_page"]-1):#เป็นการloop ซึ่งมีอยู่ 40 ชุดข้อมูล
        for p > 1: #เป็นการตัดตัวอักษรที่ไม่ต้องการ
            if url == api_link+'&page=' + str(p):#เป็นการตัดตัวอักษรที่ต้องบันทึก
                response = requests.get(url,headers={'User-Agent': 'Mozilla/5.0'})#เป็นการรับข้อมูลที่ต้องบันทึก
                pastebin_url = response.text#สามารถดูรายละเอียดของ URL ที่ต้องบันทึก
                data_json.loads(pastebin_url[4:])#เป็นการตัดตัวอักษรที่ไม่ต้องการ
                for d in data['data_object'][p]:#loop ที่ต้องบันทึก
                    for dic in range(len(list_column)):#Loop ที่ต้องบันทึกในแต่ละแบบ
                        for col in list_column[dic]:#loop ที่ต้องบันทึกในแต่ละคอลัมน์
                            dic_data[list_column[dic]] = d[list_column[dic]]#เป็นการบันทึกข้อมูลที่ต้องบันทึกในแต่ละคอลัมน์
                            lis.append(dic_data)#บันทึกข้อมูลที่ต้องบันทึก
            else:
                print('load Data Completed')
    create_file(lis,name_file)#เป็นการสร้างไฟล์ใน function สร้าง file.json
```

Task3 เป็นการนำlink api มาใส่แล้วแปลงเป็นfile.Json.

```
In [72]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_POPULATION_DIVORCE'#เป็นlink api
list_name='DIVORCE'#เป็นชื่อไฟล์ name
list_column = ['year','region_name_th','province_name_th','divorce_total']#เป็นการเลือกคอลัมน์ที่จะนำมาใช้
api(list_api,list_name,list_column)
load Data Completed
create data/DIVORCE.json Completed

In [52]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_POPULATION_MARRIAGE'#เป็นlink api
list_name = 'MARRIAGE '#เป็นชื่อไฟล์ name
list_column = ['year','region_name_th','province_name_th','marriage_total']#เป็นการเลือกคอลัมน์ที่จะนำมาใช้
api(list_api,list_name,list_column)
load Data Completed
create data/MARRIAGE.json Completed

In [51]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_POPULATION_AGE_GROUP'#เป็นlink api
list_name = 'AGE_GROUP'#เป็นชื่อไฟล์ name
list_column = ['year','region_name_th','province_name_th','age_group_desc_th','population_total','gender_name']#เป็นการเลือกคอลัมน์ที่จะนำมาใช้
api(list_api,list_name,list_column)
load Data Completed
create data/AGE_GROUP.json Completed

In [50]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_LABOUR_FORMAL_AND_INFORMAL_EMPLOYMENT'#เป็นlink api
list_name = 'INFORMAL_EMPLOYMENT '#เป็นชื่อไฟล์ name
list_column = ['year','region_name_th','province_name_th','age_group_desc_th','labour_total','gender_name_th']#เป็นการเลือกคอลัมน์ที่จะนำมาใช้
api(list_api,list_name,list_column)
load Data Completed
create data/INFORMAL_EMPLOYMENT.json Completed

In [49]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_HEALTH_LIVEBIRTHS'#เป็นlink api
list_name= 'LIVE '#เป็นชื่อไฟล์ name
list_column = ['year','region_name_th','province_name_th','birth_death_type_desc_th','births_total']#เป็นการเลือกคอลัมน์ที่จะนำมาใช้
api(list_api,list_name,list_column)
```

```

load Data Completed
create data/LIVE.json Completed

In [48]: list_api='http://edw-opendata.moi.go.th/services/api/V_ODS_HEALTH_DEATHS'#เพิ่ม link api
list_name= 'DEATHS' #เป็นการเพิ่มชื่อ file name
list_column = ['year','region_name_th','province_name_th','birth_death_type_desc_th','deaths_total']#เป็นการเพิ่ม column ที่จะดึงมา
api(list_api,list_name,list_column)

load Data Completed
create data/DEATHS.json Completed

```

Task 4: เป็นการโหลดข้อมูลแล้วนำมาระบบ Dataframe , cleaning , combine the data and filter

create Dataframe

```
In [73]: df_DIVORCE = pd.read_json('data/DIVORCE.json')
df_DIVORCE
```

```
Out[73]:
      divorce_total province_name_th region_name_th year
0        15919.0    กรุงเทพมหานคร          ภาคกลาง 2553
1        15919.0    กรุงเทพมหานคร          ภาคกลาง 2553
2        3325.0     สุนทรปราการ          ภาคกลาง 2553
3        3325.0     สุนทรปราการ          ภาคกลาง 2553
4        2134.0       นนทบุรี          ภาคกลาง 2553
...
1535      360.0       มีลักษณ์          ภาคใต้ชายแดน 2562
1536      411.0       ยะลา          ภาคใต้ชายแดน 2562
1537      411.0       ยะลา          ภาคใต้ชายแดน 2562
1538      380.0       นาทวีสัมภាឍ ภาคใต้ชายแดน 2562
1539      380.0       นาทวีสัมภាឍ ภาคใต้ชายแดน 2562
```

1540 rows x 4 columns

```
In [6]: df_MARRIAGE =pd.read_json('data/MARRIAGE.json')
df_MARRIAGE
```

```
Out[6]:
      marriage_total province_name_th region_name_th year
0        39648.0    กรุงเทพมหานคร          ภาคกลาง 2553
1        39648.0    กรุงเทพมหานคร          ภาคกลาง 2553
2        7371.0     สุนทรปราการ          ภาคกลาง 2553
3        7371.0     สุนทรปราการ          ภาคกลาง 2553
4        5076.0       นนทบุรี          ภาคกลาง 2553
...
1535      2930.0       มีลักษณ์          ภาคใต้ชายแดน 2562
1536      2109.0       ยะลา          ภาคใต้ชายแดน 2562
1537      2109.0       ยะลา          ภาคใต้ชายแดน 2562
1538      2777.0       นาทวีสัมภាឍ ภาคใต้ชายแดน 2562
1539      2777.0       นาทวีสัมภាឍ ภาคใต้ชายแดน 2562
```

1540 rows x 4 columns

```
In [7]: df_AGE_GROUP =pd.read_json('data/AGE_GROUP.json')
df_AGE_GROUP
```

```
Out[7]:
      age_group_desc_th gender_name population_total province_name_th region_name_th year
0        รวมเด็ก (0-14 ปี)      ชาย  1016550.0    กรุงเทพมหานคร          ภาคกลาง 2553
1        วัยเดร招商 (15-59 ปี)      ชาย  3624128.0    กรุงเทพมหานคร          ภาคกลาง 2553
2        วัยสูงอายุ (60 ปี及以上)      ชาย  582352.0    กรุงเทพมหานคร          ภาคกลาง 2553
3        จำนวนคุณภาพไม่ได้      ชาย  196106.0    กรุงเทพมหานคร          ภาคกลาง 2553
4        รวมเด็ก (0-14 ปี)      หญิง  969400.0    กรุงเทพมหานคร          ภาคกลาง 2553
...
6155      จำนวนคุณภาพไม่ได้      ชาย  13816.0       นาทวีสัมภ. ภาคใต้ชายแดน 2562
6156      รวมเด็ก (0-14 ปี)      หญิง  194326.0       นาทวีสัมภ. ภาคใต้ชายแดน 2562
6157      วัยเดร招商 (15-59 ปี)      หญิง  507950.0       นาทวีสัมภ. ภาคใต้ชายแดน 2562
6158      วัยสูงอายุ (60 ปี及以上)      หญิง  101348.0       นาทวีสัมภ. ภาคใต้ชายแดน 2562
6159      จำนวนคุณภาพไม่ได้      หญิง  12382.0       นาทวีสัมภ. ภาคใต้ชายแดน 2562
```

6160 rows x 6 columns

```
In [8]: df_INFORMAL_EMPLOYMENT =pd.read_json('data/INFORMAL_EMPLOYMENT.json')
df_INFORMAL_EMPLOYMENT
```

```
Out[8]:
      age_group_desc_th gender_name_th labour_total province_name_th region_name_th year
0        15 - 19          ชาย  22357.0000    กรุงเทพมหานคร          ภาคกลาง 2554
1        15 - 19          ชาย  22357.0000    กรุงเทพมหานคร          ภาคกลาง 2554
2        15 - 19          หญิง  11074.0000    กรุงเทพมหานคร          ภาคกลาง 2554
3        15 - 19          หญิง  11074.0000    กรุงเทพมหานคร          ภาคกลาง 2554
4        20 - 24          ชาย  67345.0000    กรุงเทพมหานคร          ภาคกลาง 2554
...
55435      55 - 59          หญิง  6504.3368       นาทวีสัมภ. ภาคใต้ชายแดน 2562
55436      60 ปี及以上      ชาย  14081.1371       นาทวีสัมภ. ภาคใต้ชายแดน 2562
55437      60 ปี及以上      ชาย  14081.1371       นาทวีสัมภ. ภาคใต้ชายแดน 2562
55438      60 ปี及以上      หญิง  9742.1051       นาทวีสัมภ. ภาคใต้ชายแดน 2562
55439      60 ปี及以上      หญิง  9742.1051       นาทวีสัมภ. ภาคใต้ชายแดน 2562
```

55440 rows x 6 columns

```
In [9]: df_LIVE =pd.read_json('data/LIVE.json')
df_LIVE
```

```
Out[9]:
      birth_death_type_desc_th births_total province_name_th region_name_th year
0        การเกิดใหม่ทั้งหมด 101112.0    กรุงเทพมหานคร          ภาคกลาง 2552
1        การเกิดใหม่ทั้งหมด 101112.0    กรุงเทพมหานคร          ภาคกลาง 2552
2        การเก็บในโรงพยาบาล 94180.0    กรุงเทพมหานคร          ภาคกลาง 2552
3        การเก็บในโรงพยาบาล 94180.0    กรุงเทพมหานคร          ภาคกลาง 2552
4        การเก็บใหม่ทั้งหมด 14375.0     สุนทรปราการ          ภาคกลาง 2552
...
3075      การเก็บในโรงพยาบาล 9358.0       ยะลา          ภาคใต้ชายแดน 2561
3076      การเก็บใหม่ทั้งหมด 12116.0       นาทวีสัมภ. ภาคใต้ชายแดน 2561
3077      การเก็บใหม่ทั้งหมด 12116.0       นาทวีสัมภ. ภาคใต้ชายแดน 2561
3078      การเก็บในโรงพยาบาล 11949.0       นาทวีสัมภ. ภาคใต้ชายแดน 2561
3079      การเก็บในโรงพยาบาล 11949.0       นาทวีสัมภ. ภาคใต้ชายแดน 2561
```

3080 rows x 5 columns

```
In [10]: df_DEATHS = pd.read_json('data/DEATHS.json')
df_DEATHS
```

	birth_death_type_desc_th	deaths_total	province_name_th	region_name_th	year
0	การตายทั่วไป	37490.0	กรุงเทพมหานคร	ภาคกลาง	2552
1	การตายทั่วไป	37490.0	กรุงเทพมหานคร	ภาคกลาง	2552
2	การตายในโรงพยาบาล	26857.0	กรุงเทพมหานคร	ภาคกลาง	2552
3	การตายในโรงพยาบาล	26857.0	กรุงเทพมหานคร	ภาคกลาง	2552
4	การตายทั่วไป	6020.0	สมัยนิยม	ภาคกลาง	2552
...
3075	การตายในโรงพยาบาล	1124.0	เชียง	ภาคใต้ชายแดน	2561
3076	การตายทั่วไป	4619.0	ราชอาณาจ	ภาคใต้ชายแดน	2561
3077	การตายทั่วไป	4619.0	ราชอาณาจ	ภาคใต้ชายแดน	2561
3078	การตายในโรงพยาบาล	1358.0	ราชอาณาจ	ภาคใต้ชายแดน	2561
3079	การตายในโรงพยาบาล	1358.0	ราชอาณาจ	ภาคใต้ชายแดน	2561

3080 rows x 5 columns

หาค่า Missing Values , ทำการ cleaning และ กำจัดข้อมูลที่ซ้ำกัน

```
In [11]: print(df_DIVORCE.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_DIVORCE = df_DIVORCE.fillna(df_DIVORCE.divorce_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ divorce_total
print(df_DIVORCE.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_DIVORCE.drop_duplicates(['divorce_total','province_name_th',
                            'region_name_th','year'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_DIVORCE.reset_index(drop=True, inplace=True)
print(len(df_DIVORCE))
print('Completed')
```

```
divorce_total      2
province_name_th   0
region_name_th     0
year                0
dtype: int64
True    770
False   770
dtype: int64
770
Completed
```

```
In [12]: print(df_MARRIAGE.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_MARRIAGE = df_MARRIAGE.fillna(df_MARRIAGE.marriage_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ marriage_total
print(df_MARRIAGE.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_MARRIAGE.drop_duplicates(['marriage_total','province_name_th',
                             'region_name_th','year'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_MARRIAGE.reset_index(drop=True, inplace=True)
print(len(df_MARRIAGE))
print('Completed')
```

```
marriage_total      2
province_name_th   0
region_name_th     0
year                0
dtype: int64
True    770
False   770
dtype: int64
770
Completed
```

```
In [13]: print(df_AGE_GROUP.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_AGE_GROUP = df_AGE_GROUP.fillna(df_AGE_GROUP.population_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ population_total
print(df_AGE_GROUP.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_AGE_GROUP.drop_duplicates(['population_total','province_name_th',
                               'region_name_th','age_group_desc_th'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_AGE_GROUP.reset_index(drop=True, inplace=True)
print(len(df_AGE_GROUP))
print('Completed')
```

```
age_group_desc_th   0
gender_name          0
population_total     8
province_name_th    0
region_name_th       0
year                 0
dtype: int64
False   6160
dtype: int64
6156
Completed
```

```
In [14]: print(df_INFORMAL_EMPLOYMENT.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_INFORMAL_EMPLOYMENT = df_INFORMAL_EMPLOYMENT.fillna(df_INFORMAL_EMPLOYMENT.labour_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ lab
our_total
print(df_INFORMAL_EMPLOYMENT.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_INFORMAL_EMPLOYMENT.drop_duplicates(['labour_total','province_name_th',
                                         'region_name_th','year','age_group_desc_th'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_INFORMAL_EMPLOYMENT.reset_index(drop=True, inplace=True)
print(len(df_INFORMAL_EMPLOYMENT))
print('Completed')
```

```
age_group_desc_th   0
gender_name          0
labour_total         80
province_name_th    0
region_name_th       0
year                 0
dtype: int64
True    27742
False   27698
dtype: int64
27681
Completed
```

```
In [15]: print(df_LIVE.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_LIVE = df_LIVE.fillna(df_LIVE.births_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ births_total
print(df_LIVE.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_LIVE.drop_duplicates(['births_total','province_name_th',
                         'region_name_th','year','birth_death_type_desc_th'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_LIVE.reset_index(drop=True, inplace=True)
print(len(df_LIVE))
print('Completed')
```

```
birth_death_type_desc_th   0
births_total                 8
province_name_th             0
region_name_th               0
year                          0
dtype: int64
True    1540
False   1540
dtype: int64
1540
Completed
```

```
In [16]: print(df_DEATHS.isnull().sum())#เป็นการตรวจสอบว่ามีค่าว่างหรือไม่
df_DEATHS = df_DEATHS.fillna(df_DEATHS.deaths_total.mean())#ทำการแทนที่ค่าว่างด้วยค่าเฉลี่ยของ deaths_total
print(df_DEATHS.duplicated().value_counts())#เป็นการตรวจสอบว่ามีค่าซ้ำกันหรือไม่ของผลลัพธ์
df_DEATHS.drop_duplicates(['deaths_total','province_name_th',
                           'region_name_th','year','birth_death_type_desc_th'], inplace=True)#เป็นการลบข้อมูลที่ซ้ำกันออกໄປ
df_DEATHS.reset_index(drop=True, inplace=True)
print(len(df_DEATHS))
print('Completed')
```

```
birth_death_type_desc_th   0
A
```

```

deaths_total          8
province_name_th      0
region_name_th        0
year                  0
dtype: int64
True                 1540
False                1540
dtype: int64
1540
Completed

```

ทำการfilter ข้อมูลโดยเราระบุเวลาเป็น 2554-2561มาใช้

```
In [17]: df_DIVORCE=df_DIVORCE[(df_DIVORCE.year>=2554) & (df_DIVORCE.year <2562)]
display(df_DIVORCE)
print('completed')
```

	divorce_total	province_name_th	region_name_th	year
77	15175.0	กรุงเทพมหานคร	ภาคกลาง	2554
78	3638.0	สมทบฯภาค	ภาคกลาง	2554
79	2019.0	นนทบุรี	ภาคกลาง	2554
80	2301.0	ปทุมธานี	ภาคกลาง	2554
81	1825.0	พระนครศรีอยุธยา	ภาคกลาง	2554
...
688	953.0	เชียง	ภาคใต้	2561
689	881.0	พัทลุง	ภาคใต้	2561
690	351.0	ชลบุรี	ภาคใต้เชียงใหม่	2561
691	413.0	ฉะเชิง	ภาคใต้เชียงใหม่	2561
692	380.0	นาเชิงวัง	ภาคใต้เชียงใหม่	2561

616 rows x 4 columns

completed

```
In [18]: df_MARRIAGE=df_MARRIAGE[(df_MARRIAGE.year>=2554) & (df_MARRIAGE.year <2562)]
display(df_MARRIAGE)
print('completed')
```

	marriage_total	province_name_th	region_name_th	year
77	42316.0	กรุงเทพมหานคร	ภาคกลาง	2554
78	8124.0	สมทบฯภาค	ภาคกลาง	2554
79	5470.0	นนทบุรี	ภาคกลาง	2554
80	5454.0	ปทุมธานี	ภาคกลาง	2554
81	4762.0	พระนครศรีอยุธยา	ภาคกลาง	2554
...
688	2366.0	เชียง	ภาคใต้	2561
689	1824.0	พัทลุง	ภาคใต้	2561
690	2160.0	ชลบุรี	ภาคใต้เชียงใหม่	2561
691	2041.0	ฉะเชิง	ภาคใต้เชียงใหม่	2561
692	2349.0	นาเชิงวัง	ภาคใต้เชียงใหม่	2561

616 rows x 4 columns

completed

```
In [19]: df AGE GROUP=df AGE GROUP[(df AGE GROUP.year>=2554) & (df AGE GROUP.year <2562)]
display(df AGE GROUP)
print('completed')
```

	age_group_desc_th	gender_name_th	population_total	province_name_th	region_name_th	year
612	มนต์ (0-14 ปี)	ชาย	977740.0	กรุงเทพมหานคร	ภาคกลาง	2554
613	วัยเดร wan (15-59 ปี)	ชาย	3617428.0	กรุงเทพมหานคร	ภาคกลาง	2554
614	วัยสูงอายุ (60 ปีขึ้นไป)	ชาย	611790.0	กรุงเทพมหานคร	ภาคกลาง	2554
615	ชั่วโมงคุณภาพชีวิตร้า	ชาย	178950.0	กรุงเทพมหานคร	ภาคกลาง	2554
616	วัยเด็ก (0-14 ปี)	หญิง	932982.0	กรุงเทพมหานคร	ภาคกลาง	2554
...
5535	ชั่วโมงคุณภาพชีวิตร้า	ชาย	13944.0	นราธิวาส	ภาคใต้เชียงใหม่	2561
5536	ชั่วโมง (0-14 ปี)	หญิง	195320.0	นราธิวาส	ภาคใต้เชียงใหม่	2561
5537	วัยเดร wan (15-59 ปี)	หญิง	505122.0	นราธิวาส	ภาคใต้เชียงใหม่	2561
5538	วัยสูงอายุ (60 ปีขึ้นไป)	หญิง	98150.0	นราธิวาส	ภาคใต้เชียงใหม่	2561
5539	ชั่วโมงคุณภาพชีวิตร้า	หญิง	12156.0	นราธิวาส	ภาคใต้เชียงใหม่	2561

4928 rows x 6 columns

completed

```
In [20]: df INFORMAL EMPLOYMENT=df INFORMAL EMPLOYMENT[(df INFORMAL EMPLOYMENT.year>=2554)
&(df INFORMAL EMPLOYMENT.year <2562)]
display(df INFORMAL EMPLOYMENT)
print('completed')
```

	age_group_desc_th	gender_name_th	labour_total	province_name_th	region_name_th	year
0	15 - 19	ชาย	22357.0000	กรุงเทพมหานคร	ภาคกลาง	2554
1	15 - 19	หญิง	11074.0000	กรุงเทพมหานคร	ภาคกลาง	2554
2	20 - 24	ชาย	67345.0000	กรุงเทพมหานคร	ภาคกลาง	2554
3	20 - 24	หญิง	51972.0000	กรุงเทพมหานคร	ภาคกลาง	2554
4	25 - 29	ชาย	147811.0000	กรุงเทพมหานคร	ภาคกลาง	2554
...
24600	50 - 54	หญิง	8007.2689	นราธิวาส	ภาคใต้เชียงใหม่	2561
24601	55 - 59	ชาย	6952.1292	นราธิวาส	ภาคใต้เชียงใหม่	2561
24602	55 - 59	หญิง	10229.3570	นราธิวาส	ภาคใต้เชียงใหม่	2561
24603	60 ปีขึ้นไป	ชาย	14853.2150	นราธิวาส	ภาคใต้เชียงใหม่	2561
24604	60 ปีขึ้นไป	หญิง	8862.7232	นราธิวาส	ภาคใต้เชียงใหม่	2561

24605 rows x 6 columns

completed

```
In [21]: df LIVE=df LIVE[(df LIVE.year>=2554) & (df LIVE.year <2562)]
display(df LIVE)
print('completed')
```

	birth_death_type_desc_th	births_total	province_name_th	region_name_th	year
308	การเก็บอัมมังค์แมลง	101919.0	กรุงเทพมหานคร	ภาคกลาง	2554
309	การเก็บอัมมังค์แมลง	91432.0	กรุงเทพมหานคร	ภาคกลาง	2554
310	การเก็บอัมมังค์แมลง	16460.0	สมทบฯภาค	ภาคกลาง	2554
311	การเก็บอัมมังค์แมลง	11942.0	สมทบฯภาค	ภาคกลาง	2554
312	การเก็บอัมมังค์แมลง	11367.0	นนทบุรี	ภาคกลาง	2554
...
1535	การเก็บอัมมังค์แมลง	10643.0	ชลบุรี	ภาคใต้เชียงใหม่	2561

```

1536  การต์เดินทางท่องเที่ยว 10611.0   ฉลาก  ภาคใต้ชายแดน 2561
1537  การต์เดินทางท่องเที่ยว 9358.0   ฉลาก  ภาคใต้ชายแดน 2561
1538  การต์เดินทางท่องเที่ยว 12116.0   นราธิวาส ภาคใต้ชายแดน 2561
1539  การต์เดินทางท่องเที่ยว 11949.0   นราธิวาส ภาคใต้ชายแดน 2561

```

1232 rows × 5 columns

completed

```
In [22]: df_DEATHS=df_DEATHS[(df_DEATHS.year>=2554) & (df_DEATHS.year <2562)]
display(df_DEATHS)
print('completed')
```

	birth_death_type_desc_th	deaths_total	province_name_th	region_name_th	year
308	การตายในโรงพยาบาล	37832.0	กรุงเทพมหานคร	ภาคกลาง	2554
309	การตายในโรงพยาบาล	26793.0	กรุงเทพมหานคร	ภาคกลาง	2554
310	การตายในโรงพยาบาล	6909.0	สมุทรปราการ	ภาคกลาง	2554
311	การตายในโรงพยาบาล	4252.0	สมุทรปราการ	ภาคกลาง	2554
312	การตายในโรงพยาบาล	6206.0	นนทบุรี	ภาคกลาง	2554
...
1535	การตายในโรงพยาบาล	1227.0	ปัตตานี	ภาคใต้ชายแดน	2561
1536	การตายในโรงพยาบาล	3000.0	ฉลาก	ภาคใต้ชายแดน	2561
1537	การตายในโรงพยาบาล	1124.0	ฉลาก	ภาคใต้ชายแดน	2561
1538	การตายในโรงพยาบาล	4619.0	นราธิวาส	ภาคใต้ชายแดน	2561
1539	การตายในโรงพยาบาล	1358.0	นราธิวาส	ภาคใต้ชายแดน	2561

1232 rows × 5 columns

completed

```
In [23]: df_Div_Mar=pd.concat([df_DIVORCE.loc[(df_DIVORCE.year>=2554)&(df_DIVORCE.year<=2561)].groupby('year').sum(),
                           df_MARRIAGE.loc[(df_MARRIAGE.year>=2554)&(df_MARRIAGE.year<=2561)].groupby('year').sum()])
df_Div_Mar
```

```
Out[23]:
      divorce_total  marriage_total
```

year	divorce_total	marriage_total
2554	109312.0	308048.0
2555	111377.0	314338.0
2556	107031.0	295519.0
2557	111810.0	296258.0
2558	117880.0	304392.0
2559	118539.0	307746.0
2560	121617.0	297501.0
2561	144049.0	354221.0

แปลง data type dataframe

```
In [24]: print(df_DIVORCE.info())
df_DIVORCE.year=df_DIVORCE.year.astype(str)
print('Success')

<class 'pandas.core.frame.DataFrame'>
Int64Index: 616 entries, 77 to 692
Data columns (total 4 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   divorce_total    616 non-null    float64
 1   province_name_th 616 non-null    object  
 2   region_name_th   616 non-null    object  
 3   year              616 non-null    int64  
dtypes: float64(1), int64(1), object(2)
memory usage: 24.1+ KB
None
Success
```

```
In [25]: df_MARRIAGE.info()
df_MARRIAGE.year=df_MARRIAGE.year.astype(str)
print('Success')

<class 'pandas.core.frame.DataFrame'>
Int64Index: 616 entries, 77 to 692
Data columns (total 4 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   marriage_total   616 non-null    float64
 1   province_name_th 616 non-null    object  
 2   region_name_th   616 non-null    object  
 3   year              616 non-null    int64  
dtypes: float64(1), int64(1), object(2)
memory usage: 24.1+ KB
Success
```

```
In [26]: df_AGE_GROUP.info()
df_AGE_GROUP.year=df_AGE_GROUP.year.astype(str)
print('Success')

<class 'pandas.core.frame.DataFrame'>
Int64Index: 4928 entries, 612 to 5539
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   age_group_desc_th 4928 non-null    object  
 1   gender_name        4928 non-null    object  
 2   population_total  4928 non-null    float64
 3   province_name_th  4928 non-null    object  
 4   region_name_th    4928 non-null    object  
 5   year              4928 non-null    int64  
dtypes: float64(1), int64(1), object(4)
memory usage: 269.5+ KB
Success
```

```
In [27]: df_INFORMAL_EMPLOYMENT.info()
df_INFORMAL_EMPLOYMENT.year=df_INFORMAL_EMPLOYMENT.year.astype(str)
print('Success')

<class 'pandas.core.frame.DataFrame'>
Int64Index: 24605 entries, 0 to 24604
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   age_group_desc_th 24605 non-null    object  
 1   gender_name        24605 non-null    object  
 2   labour_total       24605 non-null    float64
 3   province_name_th  24605 non-null    object  
 4   region_name_th    24605 non-null    object  
 5   year              24605 non-null    int64  
dtypes: float64(1), int64(1), object(4)
memory usage: 1.3+ MB
Success
```

```
In [28]: df_LIVE.info()
df_LIVE.year=df_LIVE.year.astype(str)
print('Success')

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1232 entries, 308 to 1539
Data columns (total 5 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   age_group_desc_th 1232 non-null    object  
 1   gender_name        1232 non-null    object  
 2   labour_total       1232 non-null    float64
 3   province_name_th  1232 non-null    object  
 4   region_name_th    1232 non-null    object  
 5   year              1232 non-null    int64  
dtypes: float64(1), int64(1), object(4)
memory usage: 1.3+ MB
Success
```

```

0    birth_death_type_desc_th   1232 non-null  object
1    births_total               1232 non-null  float64
2    province_name_th          1232 non-null  object
3    region_name_th           1232 non-null  object
4    year                      1232 non-null  int64
dtypes: float64(1), int64(1), object(3)
memory usage: 57.8+ KB
Success

```

```

In [29]: df_DEATHS.info()
df_DEATHS.year=df_DEATHS.year.astype(str)
print('Success')
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1232 entries, 308 to 1539
Data columns (total 5 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   birth_death_type_desc_th  1232 non-null  object  
 1   deaths_total              1232 non-null  float64 
 2   province_name_th         1232 non-null  object  
 3   region_name_th           1232 non-null  object  
 4   year                      1232 non-null  int64  
dtypes: float64(1), int64(1), object(3)
memory usage: 57.8+ KB
Success

```

คำนวณสเกลสรุป Data frame

```

In [74]: df_statistics=pd.concat([df_DIVORCE.describe(),df_MARRIAGE.describe(),df AGE GROUP.describe(),
                           df_INFORMAL_EMPLOYMENT.describe(),df_LIVE.describe(),df_DEATHS.describe()],axis=1)  
#เป็นการรวมค่าทางสถิติที่ไว้ในรูป
df_statistics
Out[74]:


```

Task 5: วิเคราะห์และสรุปข้อมูลที่ลังแล้วโดยใช้ตัวกรองและแปลงตามความเหมาะสม จากผลลัพธ์ของคุณการตัดความหรือข้อมูลเชิงลึกที่สามารถทำได้เกียวกับชุดข้อมูลนั้น

วิเคราะห์การแต่งงานกับการหย่าลังของประชากรในประเทศไทย

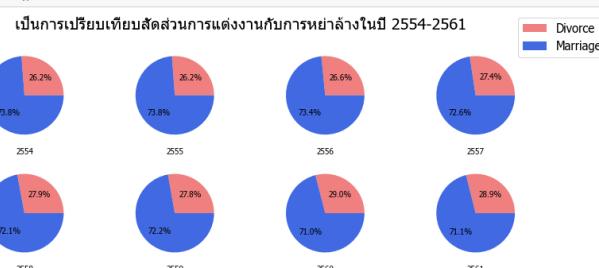
การเปรียบเทียบสัดส่วนการแต่งงานกับการหย่าลัง

-เป็นการplotกราฟข้อมูลโดยนำข้อมูลรวมการแต่งงานแต่เป็นการเปรียบเทียบสัดส่วนกับครบทุกภูมิภาค

```

In [39]: #การตัดเรียงข้อมูลเพื่อมา plot
color = ['lightcoral', 'royalblue']
text = ['Divorce', 'Marriage']
labels = text
#กราฟplotทั้งหมด
fig, axs = plt.subplots(2, 4, figsize=(13,5))
fig.suptitle('เป็นการเรียงต่อไปตามลำดับของจำนวนการหย่าลังในปี 2554-2561', fontsize=20)
#เป็นการเรียงต่อไปตามลำดับของจำนวนการหย่าลังในปี 2554-2561
plt.rcParams['font.family'] = 'tahoma'
fig.legend(labels, loc="upper right", fontsize=15) #เป็นการตั้งชื่อของคุณที่จะมาระดับลึกมากขึ้น
plt.show()

```



จากการเรียงตามลำดับของจำนวนการหย่าลังในปี 2554-2561 จะเห็นได้ว่าต่อการแต่งงานที่มีสัดส่วนที่เปลี่ยนแปลงตามที่กันกันเรื่อยๆเวลาผ่านไปสัดส่วนของการหย่าลังเพิ่มขึ้นและให้กันกันประมาณร้อยละหนึ่งในทุกๆปี

กราฟแสดงการเปลี่ยนตัวแปรการแต่งงานการเมืองต่อละภูมิภาค

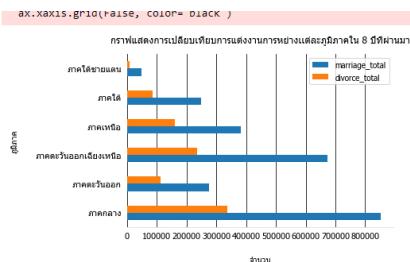
-เป็นการplotกราฟข้อมูลซึ่งกลุ่มตามภูมิภาคต่างๆแล้วหักลมของแต่ละภูมิภาคเพื่อให้รู้ข้อมูลในส่วนของการหย่าลังและการแต่งงานมาใช้ก็จะเป็นกราฟประมวลchart

```

In [40]: #เป็นการเรียงข้อมูล
df=df_MARRIAGE.join(df_DIVORCE[['divorce_total']])
data=df.groupby(['region_name_th']).sum()
# ทำการ Lot กราฟ
ax = data.plot.barh(fontsize=11)
# เป็นการตั้งชื่อของภูมิภาคที่จะต้อง
ax.set_xlabel('ภูมิภาค', labelpad=20)
ax.set_ylabel('ภูมิภาค', labelpad=20)
ax.set_title('จำนวนแต่งงานและจำนวนหย่าลังในปี 2554-2561', pad=15)
plt.rcParams['font.family'] = 'tahoma' #เป็นการตั้งชื่อให้กับกราฟโดยทันที
# ตั้งค่า axis
ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.spines['left'].set_visible(False)
ax.spines['bottom'].set_color('#000000')
ax.tick_params(bottom=False, left=False)
ax.set_axisbelow(True)
ax.yaxis.grid(False)
ax.xaxis.grid(False, color='black')
plt.show()

```

<ipython-input-40-95d82b7bdeba>:19: UserWarning: First parameter to grid() is false, but line properties are supplied. The grid will be enabled.



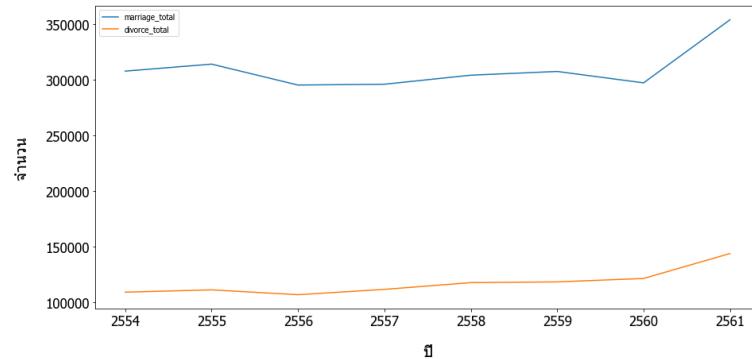
จากการที่ตัวบันทึกแสดงจำนวนการแต่งงานและการหย่าร้างของแต่ละวิสิตนาจะเห็นได้ว่าภาคกลางมีจำนวนการแต่งงานเป็นอันดับ 1 และของมาเป็นภาคตะวันออกเฉียงใต้เป็นอันดับ 2 และของภาคตะวันออกเฉียงเหนือเป็นอันดับ 3 สำหรับตัวบันทึกที่บันทึกไว้ในแต่ละวิสิตนาจะเห็นได้ว่าภาคกลางมีจำนวนการแต่งงานมากที่สุดและภาคตะวันออกเฉียงใต้เป็นอันดับ 2 และภาคตะวันออกเฉียงเหนือเป็นอันดับ 3 สำหรับตัวบันทึกที่บันทึกไว้ในแต่ละวิสิตนาจะเห็นได้ว่าภาคกลางมีจำนวนการแต่งงานมากที่สุดและภาคตะวันออกเฉียงใต้เป็นอันดับ 2 และภาคตะวันออกเฉียงเหนือเป็นอันดับ 3

แสดงการแนวโน้มของการแต่งงานและการหย่าร้าง

-เป็นการทำกราฟแสดงแนวโน้มของการแต่งงานและการหย่าร้างโดยใช้ Line Chart เพื่อที่จะมาใช้ในการคุณ化ใน

```
In [41]: #grafik แนวโน้ม
df=df_MARAGE.join(df_DIVORCE[['divorce_total']])
fig, ax = plt.subplots(figsize=(15,7))
#subplot แรก
df.groupby('year').sum().plot(ax=ax,fontsize=17)
#ไม่กรอกหัวข้อ
ax.set_xlabel('ปี', labelpad=20,fontsize=20)
ax.set_ylabel('จำนวน', labelpad=20,fontsize=20)
ax.set_title('แนวโน้มของการแต่งงานและการหย่าร้าง', pad=15,fontsize=17)
plt.rcParams['font.family']='tahoma' #เป็นการทำให้กราฟดูคมชัดมากขึ้น
```

กราฟแสดงการแนวโน้มของการแต่งงานและการหย่าร้าง



จากการที่จะแสดงผลแบบโน้มของการหย่าและการแต่งงานจะทำให้เราเห็นว่าเมืองท่าฯในปีการหย่าและการแต่งงานมีแนวโน้มที่เพิ่มขึ้น

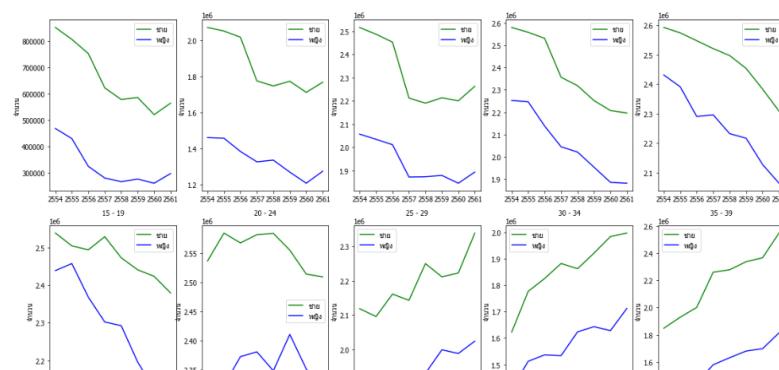
อาชญากรรมในตลาด

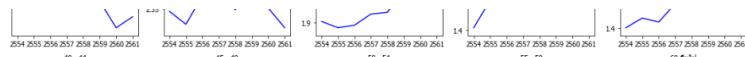
จิเคราะห์แนวโน้มเชื้อมูลของทุกกลุ่มอายุที่ทำงานในและนอกระบบ

-โดยจะทำการหาผลรวมและแยกของแต่ละกลุ่มช่วงอายุในแต่ละปีและแบ่งแยกเพื่อศึกษาเป็นรายและถึงเพื่อมาระบบ plot line chart

```
In [42]: #เป็นการทำกราฟดูอ้อมกว้าง
fig, axes = plt.subplots(2,5,figsize=(19,10))
x=0
y=-1
for i in df_INFORMAL_EMPLOYMENT['age_group_desc_th'].unique():
    if y < 4:
        y+=1
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),
                     df_INFORMAL_EMPLOYMENT[(df_INFORMAL_EMPLOYMENT['gender_name_th']=='ชาย')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='g',label='ชาย')
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),df_INFORMAL_EMPLOYMENT \
                     [(df_INFORMAL_EMPLOYMENT['gender_name_th']=='หญิง')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='b',label='หญิง')
        axes[x,y].set_xlabel(1, labelpad=10,fontsize=10)
        axes[x,y].set_ylabel('จำนวน', labelpad=2,fontsize=10)
        axes[x,y].legend()
    elif y>4:
        y=0
        x+=1
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),
                     df_INFORMAL_EMPLOYMENT[(df_INFORMAL_EMPLOYMENT['gender_name_th']=='ชาย')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='g',label='ชาย')
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),df_INFORMAL_EMPLOYMENT \
                     [(df_INFORMAL_EMPLOYMENT['gender_name_th']=='หญิง')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='b',label='หญิง')
        axes[x,y].set_xlabel(1, labelpad=10,fontsize=10)
        axes[x,y].set_ylabel('จำนวน', labelpad=2,fontsize=10)
        axes[x,y].legend()
    else:
        y=0
        x+=1
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),
                     df_INFORMAL_EMPLOYMENT[(df_INFORMAL_EMPLOYMENT['gender_name_th']=='ชาย')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='g',label='ชาย')
        axes[x,y].plot(df_INFORMAL_EMPLOYMENT['year'].unique(),df_INFORMAL_EMPLOYMENT \
                     [(df_INFORMAL_EMPLOYMENT['gender_name_th']=='หญิง')&
                     (df_INFORMAL_EMPLOYMENT['age_group_desc_th']==i)].groupby('year').sum(),color='b',label='หญิง')
        axes[x,y].set_xlabel(1, labelpad=10,fontsize=10)
        axes[x,y].set_ylabel('จำนวน', labelpad=2,fontsize=10)
        axes[x,y].legend()
    #เป็นในส่วนของการทดสอบทางการแพทย์
    fig.suptitle('กราฟแสดงแนวโน้มเชื้อมูลของทุกกลุ่มอายุที่ทำงานในและนอกระบบตั้งแต่ปี 2554-2561', fontsize=20)
plt.legend()
plt.show()
```

กราฟแสดงแนวโน้มเชื้อมูลของทุกกลุ่มอายุที่ทำงานในและนอกระบบตั้งแต่ปี 2554-2561





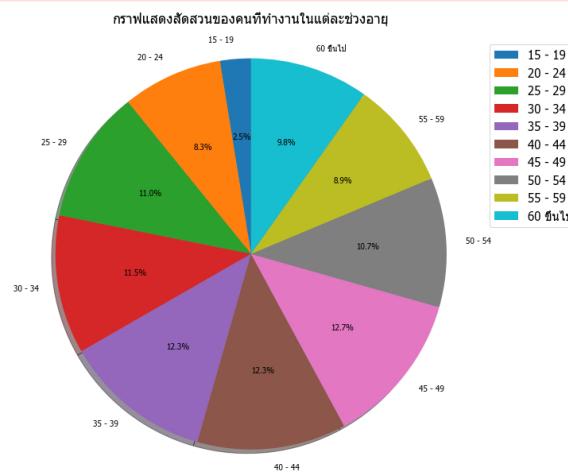
จากกราฟด้านบนจะเห็นได้ว่าจะมีจำนวนที่เล็กๆ หายใจขณะที่เพิ่มขึ้นมาทางเพศช่วงของวัยเด็กและแพร่กระจายอย่างรวดเร็วที่อายุ 50 ปีขึ้นไปมีแนวโน้มที่เพิ่มมากขึ้นอย่างต่อเนื่อง

สัดส่วนของคนที่ทำงานในแต่ละช่วงอายุ

- ในการ plot การไฟเขียวที่เรา弄มานั้นก็จะใช้แบบวงกลมโดยจะแบ่งตามช่วงอายุของคนที่ทำงานโดยจะมาใช้ในการเปรียบเทียบหัวเราะในแต่ละช่วงอายุ มีสัดส่วนเป็นอย่างไร

```
In [43]: # กรณีที่ไม่มีข้อมูลมาใช้ plot
fig1, ax1 = plt.subplots(figsize=(15, 10), subplot_kw=dict(aspect="equal"))
data = df_INFORMAL_EMPLOYMENT.groupby('age_group_desc_th').sum()
ax1.pie(data, labels=data.index, autopct='%.1f%%', shadow=True, startangle=90)
# เป็นการรีบกู้ด้วยการลบค่าที่มีค่าน้อย x และจะแสดงออกอีกครั้ง
ax1.legend(fontsize=15)
plt.rcParams['font.family'] = 'tahoma'
ax1.set_title("กราฟแสดงสัดส่วนของคนที่ทำงานในแต่ละช่วงอายุ", pad=20, fontsize=17)
ax1.axis('equal')
plt.show()

<ipython-input-43-8c5b5f049073>:4: MatplotlibDeprecationWarning: Non-1D inputs to pie() are currently squeeze(), but this behavior is deprecated since 3.1 and will be removed in 3.3; pass a 1D array instead.
    ax1.pie(data, labels=data.index, autopct='%.1f%%', shadow=True, startangle=90)
```



เมื่อเรียบเรียงข้อมูลจากสัดส่วนผู้สูงอายุจะเห็นได้ว่าตอนนี้ก้มความชุลของผู้ที่ทำงานจะลดลงในช่วงของรัฐบาลคนโภสมานะในนี้เป็นผู้สูงอายุมากที่สุดซึ่งจะมีเกิดขึ้นตามดังนี้

- 1 รายเดือน 0-18
- 2 รายปีในภูมิภาคต่ำหรือจังหวัด 18-35
- 3 รายปีในภูมิภาคกลางหรือจังหวัด 35-60
- 4 รายปีในภูมิภาคกลางหรือจังหวัด 60 ปีขึ้นไป

โดยที่มาเก็บมาจากงาน <https://sites.google.com/site/karceryteibtokhngmnusy/home/kar-cery-teibto-laea-phathnakar-khxng-phuhly>

การวิเคราะห์การเกิดการตาย

การเปรียบเทียบจำนวนการตายกับการเกิดในแต่ละปี

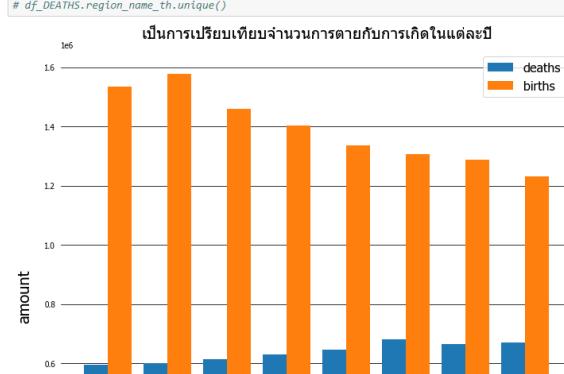
- ลองมาหาภาพหัวรุนแรงแต่เป็นรูปการเกิดและการตายท่าไร มาหาการ plot bar chart เมื่อเรียงลำดับในแต่ละปี

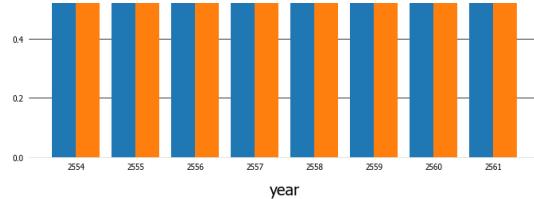
```
In [44]: # กรณีที่ไม่มีข้อมูล
x=np.arange(len(df_DEATHS.year.unique()))
fig, ax = plt.subplots(figsize=(10,10))
bar_width = 0.4
# เป็นกราฟ Lot ที่มีข้อมูลต่อปี plot เมื่อันกันต่อห้องแมลงเม็ดที่เลือกที่การซื้อขายคนให้เช่นนี้จะดูว่ามีคนมาก
b1 = ax.bar(x,df_DEATHS.groupby('year').deaths_total.sum(),width=bar_width, label='deaths')
b2 = ax.bar(x + bar_width, df_LIVE.groupby('year').births_total.sum(), width=bar_width, label='births')
# เป็นกราฟรากฐานของกราฟนี้และกราฟนี้จะถูกกลับมาไว้ในภาพรวม
ax.set_xticks(x + bar_width / 2)
ax.set_xticklabels(df_DEATHS.year.unique())
# เป็นกราฟของกราฟนี้
ax.legend(fontsize=16)

# Axis styling
ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.spines['left'].set_visible(False)
ax.spines['bottom'].set_color('#000000')
ax.tick_params(bottom=False, left=False)
ax.set_axisbelow(True)
ax.yaxis.grid(True, color='black')
ax.xaxis.grid(False)

# เป็นกราฟที่มีข้อมูลและอธิบายไว้ในภาพกราฟนี้
ax.set_xlabel('year', labelpad=15, fontsize=20)
ax.set_ylabel('amount', labelpad=15, fontsize=20)
ax.set_title('เปรียบเทียบจำนวนการตายกับการเกิดในแต่ละปี', pad=15, fontsize=20)

fig.tight_layout()
# df_DEATHS.region_name_th.unique()
```





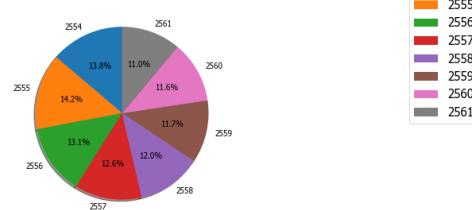
จากการเฝ้าระวังได้ว่าในแต่ละปีจะมีการเกิดและเสียการตายเพิ่มขึ้นเรื่อยๆทำให้ลดจำนวนที่เดินทางไปและจะมีจำนวนการเกิดมากกว่าการตายโดยที่เราจะสังเกตุเห็นได้ว่าการเกิดมีจำนวนที่ลดลงแต่การตายมีจำนวนเพิ่มมากขึ้น

การเบรย์บันเด็ลส่วนการเกิดในแต่ละปีตั้งแต่ปี 2554-2561

```
In [45]: # การสร้างรูปแบบพื้นที่บันเด็ล
color = ['lightcoral', 'royalblue']
marriage_born=df_LIVE.groupby('year').sum()
#grafik lot ที่มีปี
fig, axs = plt.subplots(figsize=(13,5))
fig.suptitle('บันเด็ลที่บ้านเป็นที่บันเด็ลของคนที่เกิดในแต่ละปี(ปี 2554-2561)', fontsize=20)
axs.pie(marriage_born,labels=marriage_born.index,autopt='%1.1f%%',shadow=True, startangle=90)
fig.legend(marriage_born.index, loc='upper right', fontsize=15)#บันเด็ลของคนที่บ้านเป็นที่บันเด็ลของคนที่บ้าน
plt.show()

<ipython-input-45-7ccbf8df889>:7: MatplotlibDeprecationWarning: Non-1D inputs to pie() are currently squeeze(d, but this behavior is deprecated since 3.1 and will be removed in 3.3; pass a 1D array instead.
    axs.pie(marriage_born,labels=marriage_born.index,autopt='%1.1f%%',shadow=True, startangle=90)
```

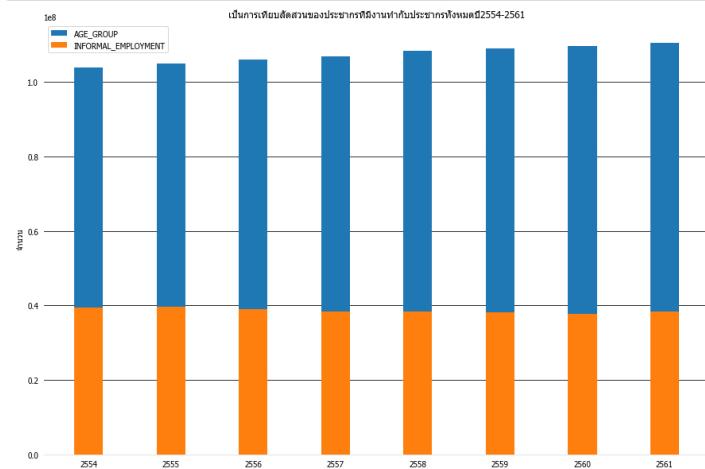
บันเด็ลการเบรย์บันเด็ลส่วนการเกิดในแต่ละปีตั้งแต่ปี 2554-2561



จากการเฝ้าระวังเห็นได้ว่าในแต่ละปีจะมีการเกิดมากกว่าการตายเพิ่มขึ้นเรื่อยๆ

การวิเคราะห์ประชากรทั้งหมด

```
In [46]: # กรณีบันเด็ล
width = 0.35 # ความกว้างของบันเด็ล
fig, ax = plt.subplots(figsize=(15, 10))
x=df_AGE_GROUP.year.unique()
y1=df_AGE_GROUP.loc[(df_AGE_GROUP.age_group_desc_th != 'มนต์เด็ก (0-14 ปี)')].groupby('year').sum().population_total
y2=df_INFORMAL_EMPLOYMENT.groupby('year').sum().labour_total
#บันเด็ลของผู้คน
ax.bar(x,y1,width, label='AGE_GROUP')
ax_bar(x, y2, width, label='INFORMAL_EMPLOYMENT')
#หัวเรื่องคือบันเด็ลของผู้คน
ax.set_xlabel('ปี')
ax.set_title('บันเด็ลส่วนของประชากรที่มีงานทำกับประชากรทั้งหมดปี 2554-2561')
# Axis styling.
ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.spines['left'].set_visible(False)
ax.spines['bottom'].set_color('#DDDDDD')
ax.tick_params(bottom=False, left=False)
ax.set_axisbelow(True)
ax.xaxis.grid(True, color='black')
ax.xaxis.grid(False)
# บันเด็ลของคน
ax.legend()
plt.show()
```



จากการเฝ้าระวังได้ว่าในแต่ละปีจะมีการเพิ่มจำนวนประชากรที่มีงานทำกับประชากรทั้งหมดเพิ่มขึ้นเรื่อยๆ ทั้งนี้มาจากการประกอบอาชีวศึกษาและเศรษฐกิจที่ดีขึ้น แต่ในส่วนของประชากรที่ไม่มีงานทำกับประชากรทั้งหมดลดลงเล็กน้อย

Summary

ข้อมูลที่ได้จากการวิเคราะห์กลุ่มประชากรในประเทศไทยจากที่ทำการวิเคราะห์ในแต่ละปีจะแสดงผลลัพธ์ที่ได้รับได้ดังนี้

- ในส่วนของจำนวนส่วนบุคคลที่มีภาระทางการตายเพิ่มมากขึ้นเช่นเดียวกับการวิเคราะห์ในแต่ละปีจะมีจำนวนที่เดินทางไปและเดินทางกลับคืนเพิ่มขึ้นเรื่อยๆ
- ในส่วนของจำนวนบุคคลที่มีภาระทางการตายเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับจำนวนบุคคลในประเทศไทยโดยเป็นปัจจัยหลักที่ทำให้เกิดการย้ายถิ่นฐาน
- ในส่วนของจำนวนบุคคลที่มีภาระทางการตายเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับจำนวนบุคคลในประเทศไทยโดยเป็นปัจจัยหลักที่ทำให้เกิดการย้ายถิ่นฐาน
- ในส่วนของจำนวนบุคคลที่มีภาระทางการตายเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับจำนวนบุคคลในประเทศไทยโดยเป็นปัจจัยหลักที่ทำให้เกิดการย้ายถิ่นฐาน
- ในส่วนของจำนวนบุคคลที่มีภาระทางการตายเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับจำนวนบุคคลในประเทศไทยโดยเป็นปัจจัยหลักที่ทำให้เกิดการย้ายถิ่นฐาน
- ในส่วนของจำนวนบุคคลที่มีภาระทางการตายเพิ่มขึ้นเรื่อยๆ เช่นเดียวกับจำนวนบุคคลในประเทศไทยโดยเป็นปัจจัยหลักที่ทำให้เกิดการย้ายถิ่นฐาน