

Data Journey

ข้อมูลชุดนี้เป็นข้อมูลของจำนวนนักเรียนด้อยโอกาส ปีการศึกษา 2564 เทอม 1 ถึง 3 ในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาปทุมธานี จำนวน 504 rows, 492 rows และ 492 rows ตามลำดับ แต่ละเทอมมี 6 columns โดยตั้งสมมติฐานหรือคำถามที่ผู้จัดทำสงสัย ว่า 1. ในแต่ละประเภทของนักเรียนที่ด้อยโอกาส แบ่งเป็นโรงเรียนใดบ้าง และจำนวนเท่าใด และ 2. จำนวนนักเรียนแต่ละระดับชั้นของทุกโรงเรียน ในแต่ละเทอม มีนักเรียนระดับชั้นไหนที่เป็นผู้ด้อยโอกาสมากที่สุด (แหล่งข้อมูล: <https://data.go.th/dataset/underpri2564>)

มีขั้นตอนดังนี้

1. นำข้อมูลมาดูภาพรวมเพื่อดูรายละเอียด

Import data

```
In [2]: df1 = pd.read_csv('https://raw.githubusercontent.com/napapolsomchatcharoenkal/DADS5001_6420412014/main/stu_64_1.csv')
df2 = pd.read_csv('https://raw.githubusercontent.com/napapolsomchatcharoenkal/DADS5001_6420412014/main/stu_64_2.csv')
df3 = pd.read_csv('https://raw.githubusercontent.com/napapolsomchatcharoenkal/DADS5001_6420412014/main/stu_64_3.csv')
print(df1.info())
print('-----')
print(df2.info())
print('-----')
print(df3.info())
print('-----')
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 504 entries, 0 to 503
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   SchoolName  504 non-null    object
1   YearTerm    504 non-null    object
2   StudentType 504 non-null    object
3   Level       504 non-null    object
4   Total       504 non-null    int64
dtypes: int64(1), object(4)
memory usage: 19.8+ KB
None
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 492 entries, 0 to 491
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   SchoolName  492 non-null    object
1   YearTerm    492 non-null    object
2   StudentType 492 non-null    object
3   Level       492 non-null    object
4   Total       492 non-null    int64
dtypes: int64(1), object(4)
memory usage: 19.3+ KB
None
```

2. Clean Data

Check data และเปลี่ยนชื่อแต่ละ Columns

```
In [3]: #อ่านไฟล์ของข้อมูล และเปลี่ยนชื่อแต่ละ columns
df1_1 = df1.rename(columns={'SchoolName':'schoolname','YearTerm':'yearterm','StudentType':'studenttype','Level':'level','Total':'total'})
df2_1 = df2.rename(columns={'SchoolName':'schoolname','YearTerm':'yearterm','StudentType':'studenttype','Level':'level','Total':'total'})
df3_1 = df3.rename(columns={'SchoolName':'schoolname','YearTerm':'yearterm','StudentType':'studenttype','Level':'level','Total':'total'})
df1_1
```

```
Out [3]:
```

	schoolname	yearterm	studenttype	level	total
0	โรงเรียนปทุมวิไล	2564/1	เด็กยากจน	P1	0
1	โรงเรียนปทุมวิไล	2564/1	เด็กยากจน	P2	0
2	โรงเรียนปทุมวิไล	2564/1	เด็กยากจน	P3	0
3	โรงเรียนปทุมวิไล	2564/1	เด็กยากจน	P4	0
4	โรงเรียนปทุมวิไล	2564/1	เด็กยากจน	P5	0
...
499	โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร	2564/1	เด็กยากจน	M2	23
500	โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร	2564/1	เด็กยากจน	M3	20
501	โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร	2564/1	เด็กยากจน	M4	26
502	โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร	2564/1	เด็กยากจน	M5	0
503	โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร	2564/1	เด็กยากจน	M6	2

504 rows x 5 columns

เปลี่ยนชื่อแต่ละโรงเรียน และแต่ละประเภทของผู้ด้อยโอกาสเป็นภาษาอังกฤษ เพราะว่า Plot ออก มาแล้วทำให้กราฟ error (ทำแบบนี้ทั้งหมด 3 เทอม)

```
In [6]: #เปลี่ยนชื่อโรงเรียนจากภาษาไทยเป็น ภาษาอังกฤษ ของเทอม 1 ปีการศึกษา 2564
df1_1.loc[df1_1['schoolname']=='โรงเรียนปทุมวิไล','schoolname'] = 'pathumwilai'
df1_1.loc[df1_1['schoolname']=='โรงเรียนคณะราษฎรบำรุงปทุมธานี','schoolname'] = 'kanaratbamrung pathumthani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนปทุมธานี นันทมนิบำรุง','schoolname'] = 'pathumthani nanthamunibamrung'
df1_1.loc[df1_1['schoolname']=='โรงเรียนหอวัง ปทุมธานี','schoolname'] = 'horwang pathumthani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนที่ปึงกรวิทยาพัฒน์ (มัยยิมวัดหัดถสธารเกษตร) ในพระราชูปถัมภ์','schoolname'] = 'dipangkornwittayapat mattayomwatha'
df1_1.loc[df1_1['schoolname']=='โรงเรียนสวนกุหลาบวิทยาลัย รังสิต','schoolname'] = 'suankularb Wittayalai rangsit'
df1_1.loc[df1_1['schoolname']=='โรงเรียนธรรมศาสตร์คลองหลวงวิทยาคม','schoolname'] = 'thammasatkhlongluangwithayakhom'
df1_1.loc[df1_1['schoolname']=='โรงเรียนบัวแก้วเกษร','schoolname'] = 'buaekawkesorn'
df1_1.loc[df1_1['schoolname']=='โรงเรียนวราชาทิพย์ดามาศุภวิทยา','schoolname'] = 'her royal highness princess soamsawali'
df1_1.loc[df1_1['schoolname']=='โรงเรียนสุนทรโรจนประเสริฐ','schoolname'] = 'soontharomettaprachasan'
df1_1.loc[df1_1['schoolname']=='โรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ปทุมธานี','schoolname'] = 'princess chulabhorn science pathum thani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนชัยสิทธิ์อาริย์','schoolname'] = 'chaiyasitthawat'
df1_1.loc[df1_1['schoolname']=='โรงเรียนเตรียมอุดมศึกษาน้อมเกล้า ปทุมธานี','schoolname'] = 'triamudomsuksanomkiao pathumthani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนธัญรัตน์','schoolname'] = 'thanyarat'
df1_1.loc[df1_1['schoolname']=='โรงเรียนสายปัญญารังสิต','schoolname'] = 'saipanyarangsit'
df1_1.loc[df1_1['schoolname']=='โรงเรียนหนองเสือวิทยาคม','schoolname'] = 'nongsua'
df1_1.loc[df1_1['schoolname']=='โรงเรียนเตรียมอุดมศึกษาพัฒนาการ ปทุมธานี','schoolname'] = 'triamudomsuksa pattanakarn pathumthani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนเทพศิรินทร์คลองสิบสาม ปทุมธานี','schoolname'] = 'debsirinklongsibsam pathumthani'
df1_1.loc[df1_1['schoolname']=='โรงเรียนนวมินทราชินูทิศ สวนกุหลาบวิทยาลัย ปทุมธานี','schoolname'] = 'navamintharachinutit suankularb wittayalai path'
df1_1.loc[df1_1['schoolname']=='โรงเรียนมัธยมสังคีตวิทยา กรุงเทพมหานคร','schoolname'] = 'matthayom sangkeet wittaya bangkok'

#เปลี่ยนชื่อของประเภทด้อยโอกาสจากภาษาไทยเป็น ภาษาอังกฤษ ของเทอม 1 ปีการศึกษา 2564
df1_1.loc[df1_1['studenttype']=='เด็กยากจน','studenttype'] = 'impoverished children'
df1_1.loc[df1_1['studenttype']=='กำพร้า','studenttype'] = 'orphaned'
df1_1.loc[df1_1['studenttype']=='ทำงานรับผิดชอบตนเองและครอบครัว','studenttype'] = 'work responsibility for yourself and family'
df1_1.loc[df1_1['studenttype']=='มีความด้อยโอกาสมากกว่า 1 ประเภท','studenttype'] = 'more than one disadvantaged'
df1_1.loc[df1_1['studenttype']=='เด็กถูกทอดทิ้ง','studenttype'] = 'abandoned child'
df1_1.loc[df1_1['studenttype']=='ชนกลุ่มน้อย','studenttype'] = 'minority'
df1_1.loc[df1_1['studenttype']=='เด็กในสถานพินิจและคุ้มครองเยาวชน','studenttype'] = 'children in juvenile detention'

df1_2 = df1_1
df1_2
```

Out[6]:

	schoolname	yearterm	studenttype	level	total
0	pathumwilai	2564/1	impoverished children	P1	0
1	pathumwilai	2564/1	impoverished children	P2	0
2	pathumwilai	2564/1	impoverished children	P3	0
3	pathumwilai	2564/1	impoverished children	P4	0
4	pathumwilai	2564/1	impoverished children	P5	0
...
499	matthayom sangkeet wittaya bangkok	2564/1	impoverished children	M2	23
500	matthayom sangkeet wittaya bangkok	2564/1	impoverished children	M3	20
501	matthayom sangkeet wittaya bangkok	2564/1	impoverished children	M4	26
502	matthayom sangkeet wittaya bangkok	2564/1	impoverished children	M5	0
503	matthayom sangkeet wittaya bangkok	2564/1	impoverished childen	M6	2

504 rows x 5 columns

รวมจำนวนนักเรียนทั้งหมดที่ด้อยโอกาสในแต่ละประเภท ของแต่ละโรงเรียน ปีการศึกษา 2564 ทั้ง 3 เทอม (ทำแบบนี้ทั้งหมด 3 เทอม)

```
In [9]: #รวมจำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสแต่ละประเภท ของแต่ละโรงเรียน ในเทอม 1 ปีการศึกษา 2564

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท impoverished children ของแต่ละโรงเรียน
df1_t1 = df1_2[(df1_2['studenttype'] == 'impoverished children')].groupby(['schoolname'])
df1_t11 = df1_t1.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท orphaned ของแต่ละโรงเรียน
df1_t2 = df1_2[(df1_2['studenttype'] == 'orphaned')].groupby(['schoolname'])
df1_t22 = df1_t2.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท work responsibility for yourself and family ของแต่ละโรงเรียน
df1_t3 = df1_2[(df1_2['studenttype'] == 'work responsibility for yourself and family')].groupby(['schoolname'])
df1_t33 = df1_t3.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท more than one disadvantaged ของแต่ละโรงเรียน
df1_t4 = df1_2[(df1_2['studenttype'] == 'more than one disadvantaged')].groupby(['schoolname'])
df1_t44 = df1_t4.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท abandoned child ของแต่ละโรงเรียน
df1_t5 = df1_2[(df1_2['studenttype'] == 'abandoned child')].groupby(['schoolname'])
df1_t55 = df1_t5.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท minority ของแต่ละโรงเรียน
df1_t6 = df1_2[(df1_2['studenttype'] == 'minority')].groupby(['schoolname'])
df1_t66 = df1_t6.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

#จำนวนนักเรียนทั้งหมด ที่ด้อยโอกาสประเภท children in juvenile detention ของแต่ละโรงเรียน
df1_t7 = df1_2[(df1_2['studenttype'] == 'children in juvenile detention')].groupby(['schoolname'])
df1_t77 = df1_t7.sum('total').sort_values(['schoolname'],ascending=[True]).reset_index(drop=False)

df1_t11
```

Out [9]:

	schoolname	total
0	buakaewkesorn	146
1	chaiyasitthawat	210
2	debsirinklongsibsam pathumthani	97
3	dipangkornwittayapat mattayomwathatasankaset	600
4	her royal highness princess soamsawali	313
5	horwang pathumthani	165
6	kanaratbamrung pathumthani	94
7	matthayom sangkeet wittaya bangkok	100
8	nawamintharachinutit suankularb wittayalai pat...	85
9	nongsua	644
10	pathumthani nanthamunibamrung	404
11	pathumwilai	48
12	saipanyarangsit	3
13	soontharomettaprachasan	88
14	suankularbwittayalai rangsit	16
15	thammasatkhloungluangwithayakhom	307
16	thanyarat	342
17	triamudomsuksa pattanakarn pathumthani	206
18	triamudomsuksanomklao pathumthani	231

สร้าง Data Frame ของรายชื่อโรงเรียนทั้งหมด (ปัญหาที่พบ: เหตุผลที่ต้องสร้าง DataFrame รายชื่อโรงเรียน เพราะว่า หลังจากที่ “รวมจำนวนนักเรียนทั้งหมดที่ด้อยโอกาสในแต่ละประเภท ของแต่ละโรงเรียน” แล้ว ทำให้รายชื่อของโรงเรียนหนึ่งหายไป (ภาพด้านบน) เนื่องจากโรงเรียนวิทยาศาสตร์จุฬาภรณราชวิทยาลัย ปทุมธานีนั้นไม่มีผู้ด้อยโอกาส ประเภท “เด็กยากจน(impoverished children)” จึงต้องสร้างรายชื่อโรงเรียน เพื่อที่จะได้ทำ merge ในขั้นตอนถัดไป)

In [12]:

```
#ลบ schoolname ที่ซ้ำกันทั้งหมด
df_aaa = df1_2.drop_duplicates(['schoolname'])
df_aa = df_aaa.sort_values(['schoolname'],ascending=[True]).reset_index(drop=True)
df_a = df_aa.drop(['yearterm','studenttype','level','total'], axis=1)
df_a
```

Out [12]:

	schoolname
0	buakaewkesorn
1	chaiyasitthawat
2	debsirinklongsibsam pathumthani
3	dipangkornwittayapat mattayomwathatasankaset
4	her royal highness princess soamsawali
5	horwang pathumthani
6	kanaratbamrung pathumthani
7	matthayom sangkeet wittaya bangkok
8	nawamintharachinutit suankularb wittayalai pat...
9	nongsua
10	pathumthani nanthamunibamrung
11	pathumwilai
12	princess chulabhorn science pathum thani
13	saipanyarangsit
14	soontharomettaprachasan
15	suankularbwittayalai rangsit
16	thammasatkhloungluangwithayakhom
17	thanyarat
18	triamudomsuksa pattanakarn pathumthani
19	triamudomsuksanomklao pathumthani

หาค่าเฉลี่ยของผู้ด้อยโอกาสแต่ละประเภท ของทุกโรงเรียน

```
In [14]: #รวมจำนวนนักเรียน ของแต่ละโรงเรียน ของแต่ละ Type และหาค่าเฉลี่ย

df_left_t1 = pd.merge(df_a,df1_t11, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t1 = pd.merge(df_left_t1,df2_t11, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t1 = pd.merge(df_left_t1,df3_t11, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t1 = df_left_t1.rename(columns = {'total_x':'1/64','total_y':'2/64','total':'3/64'})
df_left_t1['m_Type1'] = df_left_t1[['1/64','2/64','3/64']] .agg(['mean'], axis=1).fillna(0)
df_mean_type1 = df_left_t1.loc[:,['schoolname','m_Type1']].fillna(0)

df_left_t2 = pd.merge(df_a,df1_t22, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t2 = pd.merge(df_left_t2,df2_t22, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t2 = pd.merge(df_left_t2,df3_t22, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t2 = df_left_t2.rename(columns = {'total_x':'1/64','total_y':'2/64','total':'3/64'})
df_left_t2['m_Type2'] = df_left_t2[['1/64','2/64','3/64']] .agg(['mean'], axis=1).fillna(0)
df_mean_type2 = df_left_t2.loc[:,['schoolname','m_Type2']].fillna(0)

df_left_t3 = pd.merge(df_a,df1_t33, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t3 = pd.merge(df_left_t3,df2_t33, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t3 = pd.merge(df_left_t3,df3_t33, left_on='schoolname', right_on='schoolname', how='left' ).fillna(0)
df_left_t3 = df_left_t3.rename(columns = {'total_x':'1/64','total_y':'2/64','total':'3/64'})
df_left_t3['m_Type3'] = df_left_t3[['1/64','2/64','3/64']] .agg(['mean'], axis=1).fillna(0)
df_mean_type3 = df_left_t3.loc[:,['schoolname','m_Type3']].fillna(0)
```

	schoolname	1/64	2/64	3/64	\		schoolname	m_Type1
0	buakaewkesorn	146.0	144.0	143.0		0	buakaewkesorn	144.333333
1	chaiyasitthawat	210.0	239.0	237.0		1	chaiyasitthawat	228.666667
2	debsirinklongsibsam pathumthani	97.0	97.0	96.0		2	debsirinklongsibsam pathumthani	96.666667
3	dipangkornwittayapat mattayomwathasankaset	600.0	594.0	591.0		3	dipangkornwittayapat mattayomwathasankaset	595.000000
4	her royal highness princess soamsawali	313.0	311.0	302.0		4	her royal highness princess soamsawali	308.666667
5	horwang pathumthani	165.0	161.0	160.0		5	horwang pathumthani	162.000000
6	kanaratbamrung pathumthani	94.0	95.0	95.0		6	kanaratbamrung pathumthani	94.666667
7	matthayom sangkeet wittaya bangkok	100.0	90.0	89.0		7	matthayom sangkeet wittaya bangkok	93.000000
8	nawamintharachinitit suankularb wittayalai pat...	85.0	85.0	85.0		8	nawamintharachinitit suankularb wittayalai pat...	85.000000
9	nongsua	644.0	646.0	646.0		9	nongsua	645.333333
10	pathumthani nanthamunibamrung	404.0	403.0	403.0		10	pathumthani nanthamunibamrung	403.333333
11	pathumwilai	48.0	7.0	7.0		11	pathumwilai	20.666667
12	princess chulabhorn science pathum thani	0.0	0.0	0.0		12	princess chulabhorn science pathum thani	0.000000
13	saipanyarangsit	3.0	3.0	3.0		13	saipanyarangsit	3.000000
14	soontharomettaprachasan	88.0	86.0	84.0		14	soontharomettaprachasan	86.000000
15	suankularbwittayalai rangsit	16.0	16.0	16.0		15	suankularbwittayalai rangsit	16.000000
16	thammasatkhlongluangwithayakhom	307.0	304.0	304.0		16	thammasatkhlongluangwithayakhom	305.000000
17	thanyarat	342.0	342.0	339.0		17	thanyarat	341.000000
18	tri mudomsuksa pattanakarn pathumthani	206.0	205.0	204.0		18	tri mudomsuksa pattanakarn pathumthani	205.000000
19	tri mudomsuksanomklao pathumthani	231.0	229.0	229.0		19	tri mudomsuksanomklao pathumthani	229.666667

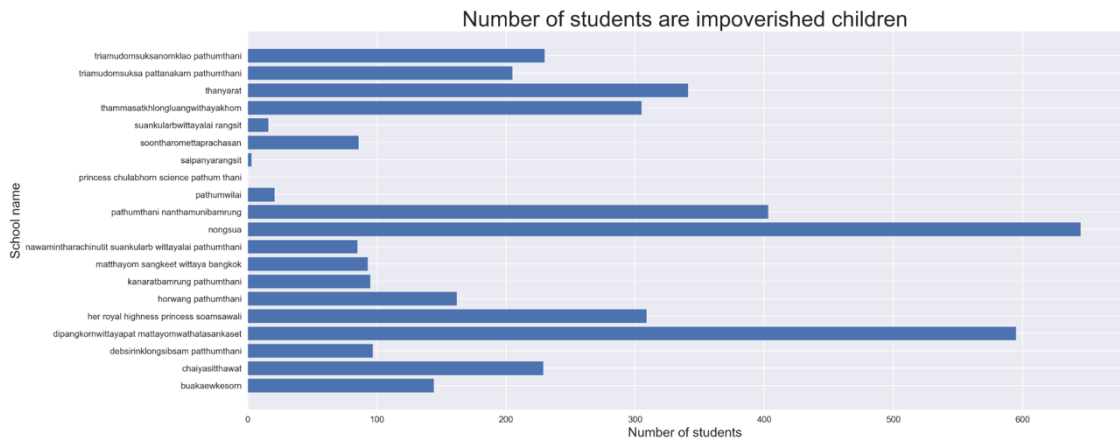
ได้ Data Frame ของทุกโรงเรียน และทุกประเภทของผู้ด้อยโอกาส ดังภาพ (ปัญหาที่พบ: จำนวนประเภทของผู้ด้อยโอกาสประเภทที่ 2 เป็นต้นไปมีน้อย จึงไม่นำมาวิเคราะห์)

Out[15]:	schoolname	impoveryshed children	orphaned	work responsibility for yourself and family	more than one disadvantaged	abandoned child	minority	children in juvenile detention
0	buakaewkesorn	144.0	0.0	0.0	0.0	0.0	0.0	0.0
1	chaiyasitthawat	229.0	0.0	0.0	0.0	0.0	0.0	0.0
2	debsirinklongsibsam pathumthani	97.0	0.0	0.0	0.0	0.0	0.0	0.0
3	dipangkornwittayapat mattayomwathasankaset	595.0	0.0	0.0	0.0	0.0	0.0	18.0
4	her royal highness princess soamsawali	309.0	0.0	0.0	0.0	0.0	0.0	0.0
5	horwang pathumthani	162.0	1.0	6.0	0.0	4.0	1.0	0.0
6	kanaratbamrung pathumthani	95.0	45.0	6.0	1.0	0.0	0.0	0.0
7	matthayom sangkeet wittaya bangkok	93.0	0.0	0.0	0.0	0.0	0.0	0.0
8	nawamintharachinitit suankularb wittayalai pat...	85.0	0.0	0.0	0.0	0.0	0.0	0.0
9	nongsua	645.0	1.0	0.0	1.0	0.0	0.0	0.0
10	pathumthani nanthamunibamrung	403.0	7.0	0.0	1.0	2.0	0.0	0.0
11	pathumwilai	21.0	0.0	0.0	0.0	0.0	0.0	0.0
12	princess chulabhorn science pathum thani	0.0	1.0	0.0	0.0	0.0	0.0	0.0
13	saipanyarangsit	3.0	0.0	0.0	0.0	0.0	0.0	0.0
14	soontharomettaprachasan	86.0	2.0	0.0	0.0	0.0	0.0	0.0
15	suankularbwittayalai rangsit	16.0	0.0	0.0	0.0	0.0	0.0	0.0
16	thammasatkhlongluangwithayakhom	305.0	0.0	1.0	0.0	0.0	0.0	0.0
17	thanyarat	341.0	0.0	1.0	0.0	0.0	0.0	0.0
18	tri mudomsuksa pattanakarn pathumthani	205.0	24.0	1.0	1.0	0.0	0.0	0.0
19	tri mudomsuksanomklao pathumthani	230.0	0.0	1.0	1.0	0.0	0.0	0.0

3. Visualization

แสดงแผนภาพของผู้ด้อยโอกาสประเภท “เด็กยากจน(impoverished children)” ของทุกโรงเรียน

Out[18]: Text(0, 0.5, 'School name')

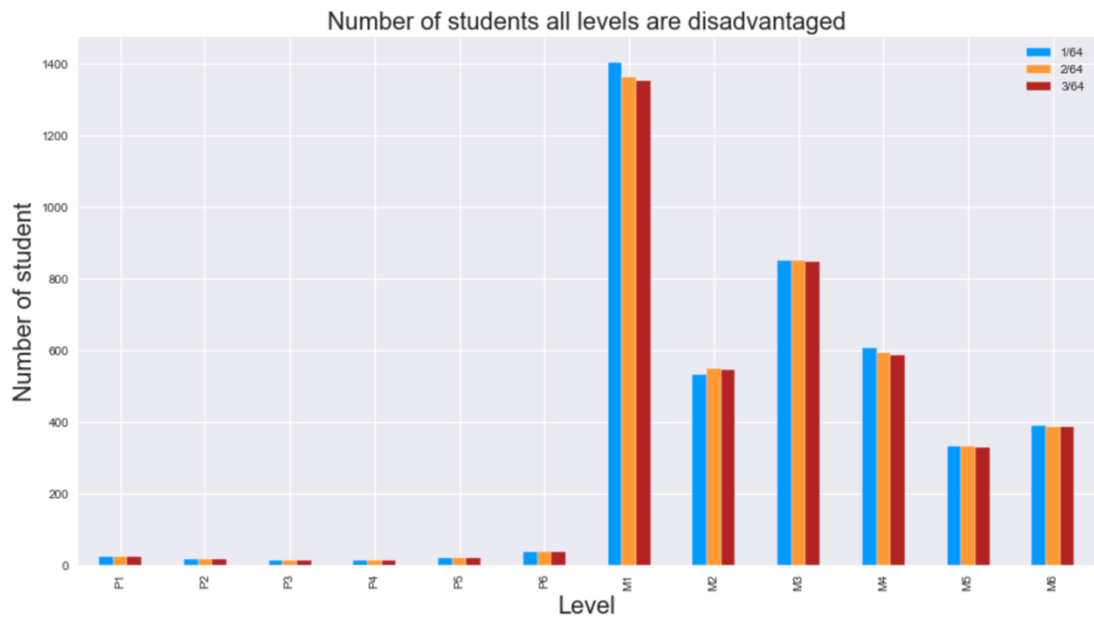


แสดงแผนภาพของจำนวนนักเรียนทุกระดับชั้นในแต่ละเทอม ของทุกโรงเรียน ทุกประเภทที่ด้อยโอกาส

Out[26]: 1/64 2/64 3/64

level			
P1	22	22	22
P2	17	17	17
P3	12	12	12
P4	12	12	12
P5	20	20	20
P6	36	36	36
M1	1404	1361	1351
M2	530	547	544
M3	849	850	846
M4	606	591	586
M5	332	330	328
M6	388	385	385

Out[27]: Text(0, 0.5, 'Number of student')



ปัญหาที่พบ:

1. ในตอนแรกเกิดความสับสน ว่าควรจะแก้ไขชุดข้อมูลของโรงเรียนที่หายไป 1 โรงเรียนหรือไม่ เพราะโรงเรียนที่หายไป ไม่น่าจะมีผลต่อชุดข้อมูล แต่พอนึกถึงขั้นตอนแสดงแผนภาพ จะทำให้เราไม่สามารถตอบได้ว่า เป็นโรงเรียนใดที่หายไป และทำไมถึงหายไป จึงต้องเสียเวลาแก้ Code รวมถึงตัวแปรใหม่หมด หลังจากแก้และ Clean data ได้แล้วพบว่าโรงเรียนนั้น ไม่มีประเภทของผู้ด้อยโอกาสนี้อยู่ จะต้องหาวิธี Clean data ใหม่ เพราะต้องการแสดงแผนภาพของโรงเรียนทั้งหมด
2. หลังจากได้ชุดข้อมูลมาครบทั้ง 20 โรงเรียนแล้ว พบว่าไม่ได้ทำการรวมเข้าทั้ง 3 เทอม ทำให้ตัวเลขของชุดข้อมูลเกิดความผิดพลาด
3. ชุดข้อมูลที่ได้ครบ 3 เทอม เมื่อแสดง Data Frame ทำให้ตัวเลขเกิดความกว้างของชุดข้อมูล เช่น โรงเรียนหนึ่ง ประเภทแรกมี 600 คน ประเภทสองมี 1 คน ทำให้ชุดข้อมูลไม่มีความสัมพันธ์กัน จึงต้องตัดออก

NapapolS_6420412014