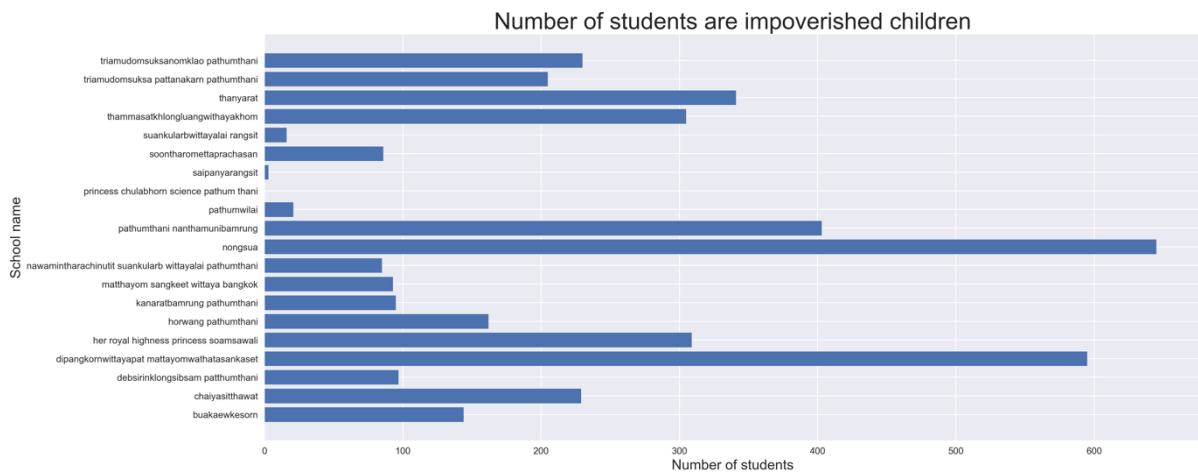


Data Journey

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>)

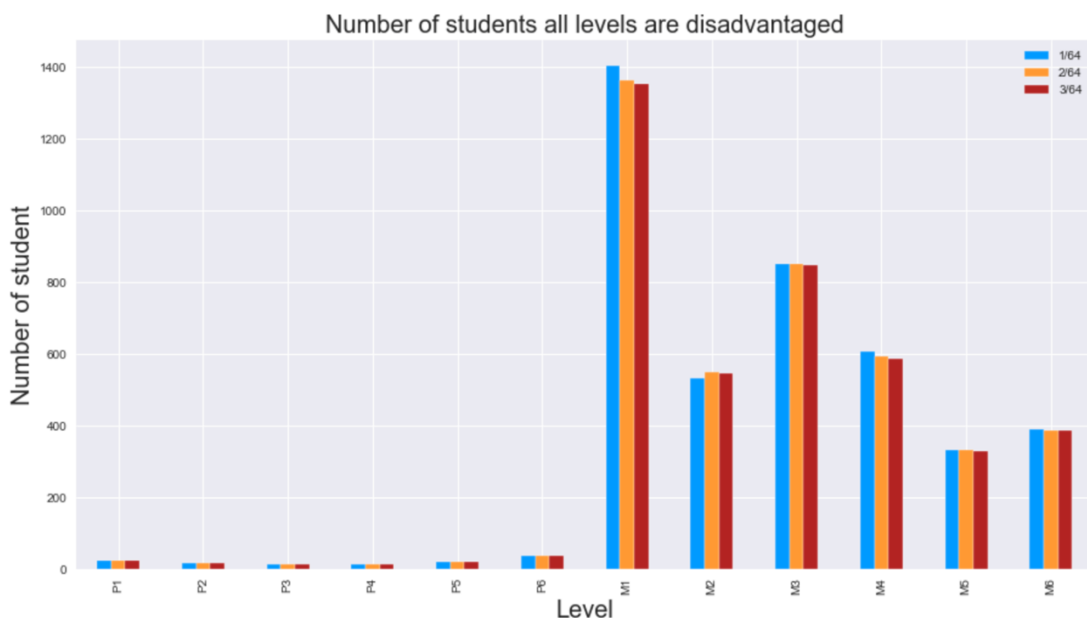
ชุดข้อมูลของจำนวนนักเรียนด้อยโอกาส

1. แสดงแผนภาพของผู้ด้อยโอกาสประเภท “เด็กยากจน(impoverished children)” ทุกระดับชั้น ของแต่ละโรงเรียน



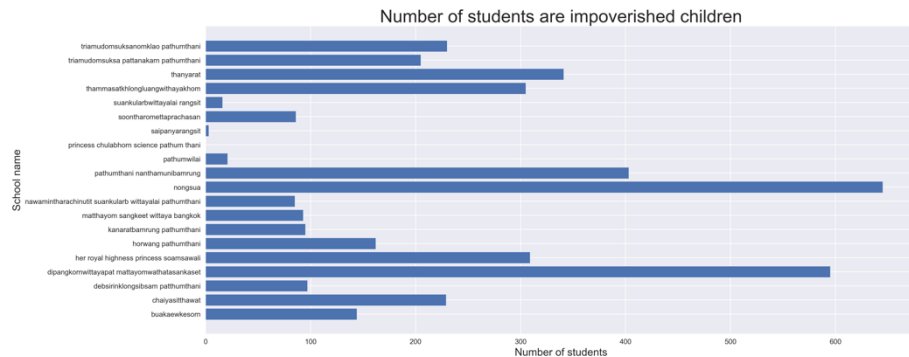
2. แสดงแผนภาพของจำนวนนักเรียนทุกระดับชั้นในแต่ละเทอม ของทุกโรงเรียน ทุกประเภทที่ด้อยโอกาส

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



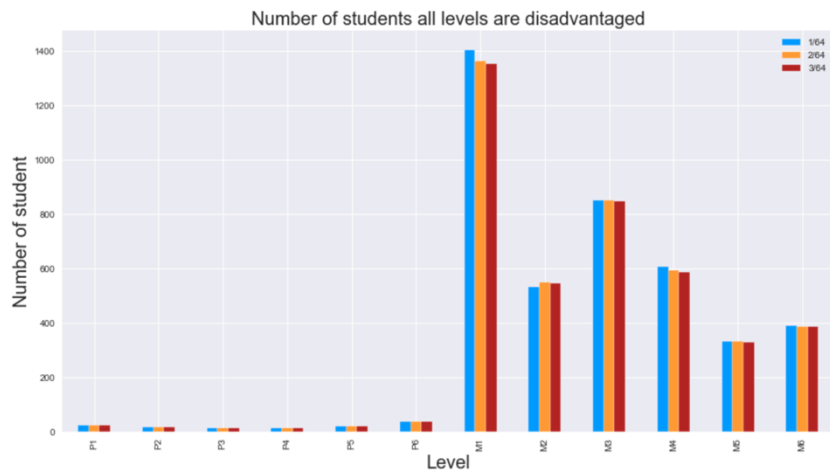
สมมติฐานเพิ่มเติม

จากแผนภาพ 1. จำนวนของนักเรียนผู้ด้อยโอกาสประเภท “เด็กยากจน(impoverished children)” ทุกระดับชั้นของแต่ละโรงเรียน สามารถสรุปได้หรือไม่ ว่าทุกโรงเรียนของประเภทเด็กยากจน **ส่วนใหญ่**เป็นเด็กนักเรียนระดับชั้น “มัธยมศึกษา 1 (M1)” จากแผนภาพที่ 2.



แผนภาพ 1.

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



แผนภาพ 2.

หมายเหตุแนะคำตอบของสมมติเพิ่มเติม

ยกตัวอย่าง รร. หนองเสือวิทยาคม(nongsua) มีจำนวนนักเรียนด้อยโอกาสประเภทยากจน เฉลี่ยทุกระดับชั้นอยู่ที่ 645 คน

1. อาจจะจริงที่ส่วนใหญ่เป็นนักเรียนระดับชั้น M1 เพราะ แผนภาพที่ 2. มีนักเรียน M1 อยู่จำนวนมากจริง ๆ
2. อาจจะไม่จริง

มีขั้นตอนดังนี้

สร้างแผนภาพ ก

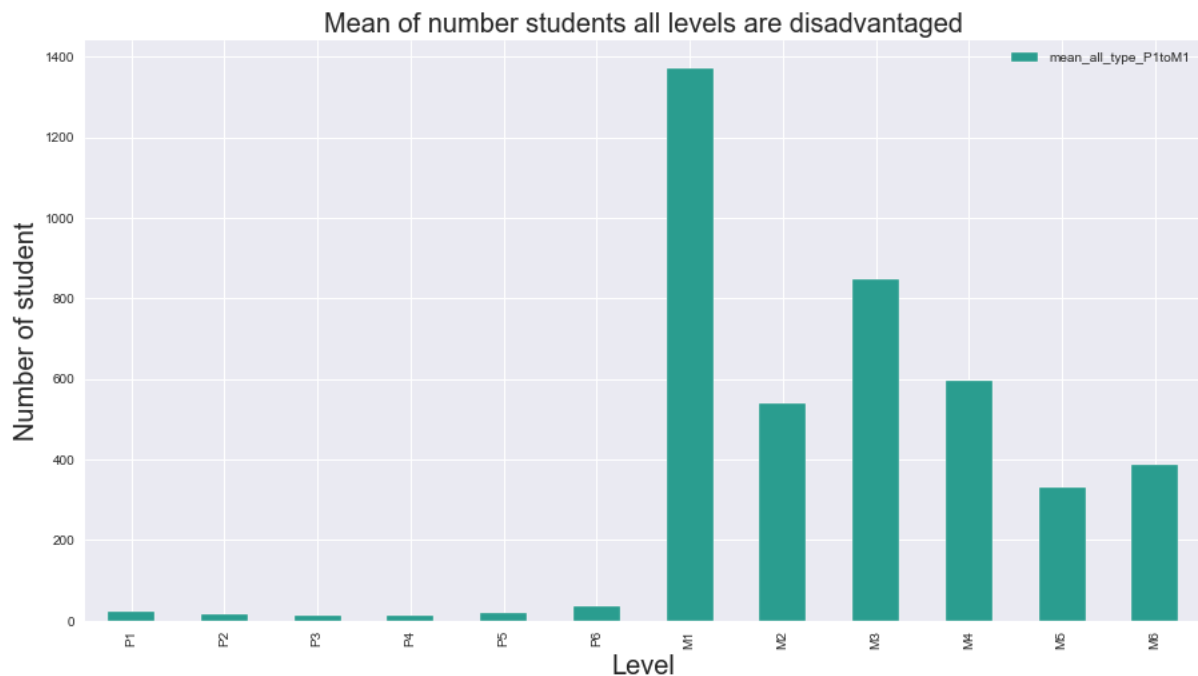
```
In [28]: 1 #หาค่าเฉลี่ยจำนวนนักเรียนทั้งสามเทอมของทุกประเภทที่ค้อยโอกาส แยกแต่ละระดับชั้น ของทุกโรงเรียน
2 df_all_level_c2['mean_all_type_P1toM1'] = df_all_level_c2[['1/64','2/64','3/64']].agg(['mean'], axis=1).fillna(0)
3 df_all_level_c3 = df_all_level_c2.loc[:,['mean_all_type_P1toM1']].round()
4 df_all_level_c3
```

```
Out[28]:
```

mean_all_type_P1toM1	
level	
P1	22.0
P2	17.0
P3	12.0
P4	12.0
P5	20.0
P6	36.0
M1	1372.0
M2	540.0
M3	848.0
M4	594.0
M5	330.0
M6	386.0

```
In [29]: 1 #แผนภาพจำนวนนักเรียนเฉลี่ยทั้งสามเทอมของทุกประเภทที่ค้อยโอกาส แยกแต่ละระดับชั้น ของทุกโรงเรียน
2 df_all_level_c3.plot(kind='bar',figsize=(15,8),rot=90,color=['#2a9d8f'])#width=0.3,color='b')
3 plt.title('Mean of number students all levels are disadvantaged',size=20)
4 plt.xlabel("Level",size=20)
5 plt.ylabel("Number of student",size=20)
```

```
Out[29]: Text(0, 0.5, 'Number of student')
```



สร้างแผนภาพ ข

สร้าง DataFrame ของ studenttype ประเภท impoverished children ของแต่ละระดับชั้น

```
In [30]: 1 #สร้าง DataFrame studenttype ประเภท impoverished children ของแต่ละระดับชั้น
2 df_c = pd.DataFrame([
3     ['impoverished children', 'M1'],
4     ['impoverished children', 'M2'],
5     ['impoverished children', 'M3'],
6     ['impoverished children', 'M4'],
7     ['impoverished children', 'M5'],
8     ['impoverished children', 'M6'],
9     ['impoverished children', 'P1'],
10    ['impoverished children', 'P2'],
11    ['impoverished children', 'P3'],
12    ['impoverished children', 'P4'],
13    ['impoverished children', 'P5'],
14    ['impoverished children', 'P6']],
15    columns=['studenttype', 'level'],
16    index=['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', '10', '11'])
17 df_c
```

	studenttype	level
0	impoverished children	M1
1	impoverished children	M2
2	impoverished children	M3
3	impoverished children	M4
4	impoverished children	M5
5	impoverished children	M6
6	impoverished children	P1
7	impoverished children	P2
8	impoverished children	P3
9	impoverished children	P4
10	impoverished children	P5
11	impoverished children	P6

```
In [34]: 1 #รวมข้อมูลเข้าด้วยกันด้วย merge
2 df_left_h1 = pd.merge(df_c, df1_t_x2, left_on='level', right_on='level', how='left').fillna(0)
3 df_left_h2 = pd.merge(df_left_h1, df1_t_y2, left_on='level', right_on='level', how='left').fillna(0)
4 df_left_h3 = pd.merge(df_left_h2, df1_t_z2, left_on='level', right_on='level', how='left').fillna(0)
5 df_left_h4 = df_left_h3.rename(columns = {'total_x': '1/64', 'total_y': '2/64', 'total': '3/64'})
6 df_left_h4
```

```
Out[34]:
```

	studenttype	level	1/64	2/64	3/64
0	impoverished children	M1	1394	1351	1341
1	impoverished children	M2	514	531	528
2	impoverished children	M3	818	821	817
3	impoverished children	M4	585	571	566
4	impoverished children	M5	301	299	297
5	impoverished children	M6	368	365	365
6	impoverished children	P1	22	22	22
7	impoverished children	P2	17	17	17
8	impoverished children	P3	12	12	12
9	impoverished children	P4	12	12	12
10	impoverished children	P5	20	20	20
11	impoverished children	P6	36	36	36

```
In [36]: 1 #หาค่าเฉลี่ย studenttype ประเภท impoverished children ของแต่ละระดับชั้น ทั้ง 3 เทอม
2 df_all_level_M1_4['mean_all_type_P1toM1'] = df_all_level_M1_4[['1/64','2/64','3/64']].agg(['mean'], axis=1).fillna(0)
3 df_all_level_M1_4 = df_all_level_M1_4.loc[:,['studenttype','level','mean_all_type_P1toM1']].round(0)
4 df_all_level_M1_4
```

```
Out[36]:
```

	studenttype	level	mean_all_type_P1toM1
0	impoverished children	P1	22.0
1	impoverished children	P2	17.0
2	impoverished children	P3	12.0
3	impoverished children	P4	12.0
4	impoverished children	P5	20.0
5	impoverished children	P6	36.0
6	impoverished children	M1	1362.0
7	impoverished children	M2	524.0
8	impoverished children	M3	819.0
9	impoverished children	M4	574.0
10	impoverished children	M5	299.0
11	impoverished children	M6	366.0

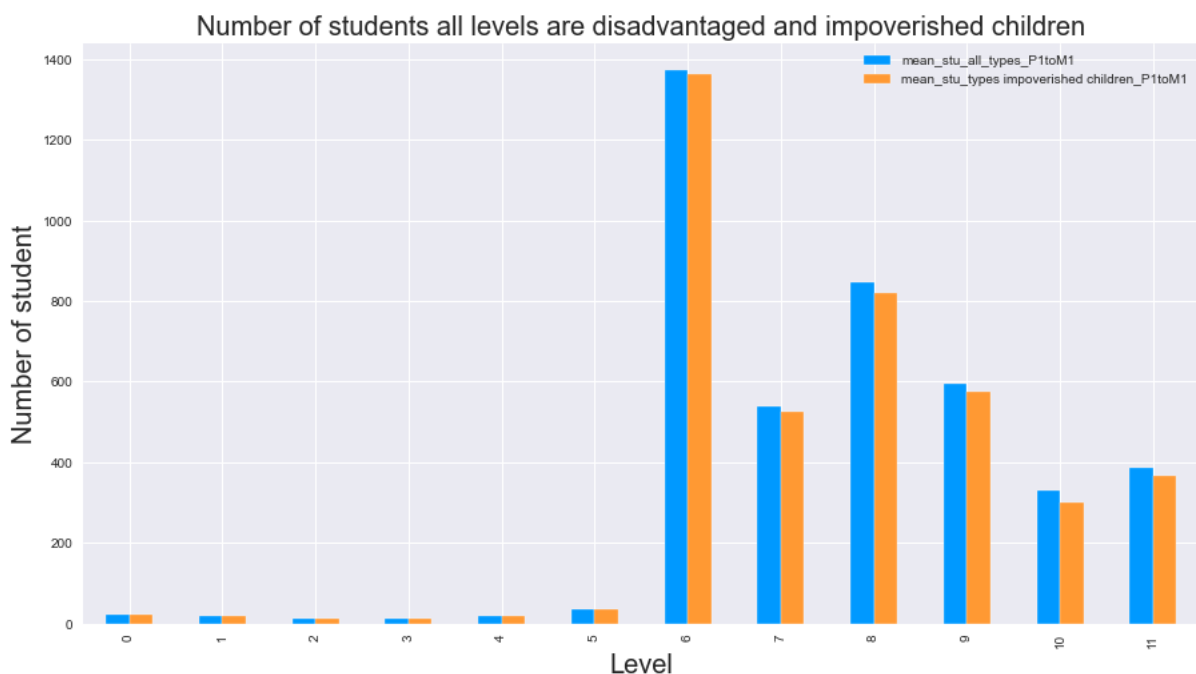
```
In [37]: 1 #รวมแผนภาพ
2 df_hypo_1 = pd.merge(df_all_level_c3,df_all_level_M1_4, left_on='level', right_on='level', how='left' ).fillna(0)
3 df_hypo_11 = df_hypo_1.rename(columns = {'mean_all_type_P1toM1_x':'mean_stu_all_types_P1toM1','mean_all_type_P1toM1': 'mean_all_type_P1toM1'})
4 df_hypo_111 = df_hypo_11.loc[:,['studenttype','level','mean_stu_all_types_P1toM1','mean_stu_types impoverished children_P1toM1']]
5 df_hypo_111
```

```
Out[37]:
```

	studenttype	level	mean_stu_all_types_P1toM1	mean_stu_types impoverished children_P1toM1
0	impoverished children	P1	22.0	22.0
1	impoverished children	P2	17.0	17.0
2	impoverished children	P3	12.0	12.0
3	impoverished children	P4	12.0	12.0
4	impoverished children	P5	20.0	20.0
5	impoverished children	P6	36.0	36.0
6	impoverished children	M1	1372.0	1362.0
7	impoverished children	M2	540.0	524.0
8	impoverished children	M3	848.0	819.0
9	impoverished children	M4	594.0	574.0
10	impoverished children	M5	330.0	299.0
11	impoverished children	M6	386.0	366.0

```
In [38]: 1 #แสดงแผนภาพค่าเฉลี่ย studenttype ของทุกประเภทที่ด้อยโอกาส และประเภท impoverished children ของแต่ละระดับชั้น ทั้ง 3 เทอม
2 df_hypo_111.plot(kind='bar',figsize=(15,8),rot=90,color=['#0099FF','#FF9933','#B22222'],width=0.3,color='b')
3 plt.title('Number of students all levels are disadvantaged and impoverished children',size=20)
4 plt.xlabel("Level",size=20)
5 plt.ylabel("Number of student",size=20)
```

```
Out[38]: Text(0, 0.5, 'Number of student')
```



ตรวจสอบสมมติฐานว่า ส่วนใหญ่ที่ด้อยโอกาสประเภท impoverished children เป็นนักเรียนระดับชั้น M1

```
In [39]: 1 #รวมประเภท impoverished children ระดับชั้น M1 ทุกโรงเรียนของเทอม 1
2 df_check_a0 = df1_1.loc[:, ['schoolname', 'yearterm', 'studenttype', 'level', 'total']]
3 df_check_a0_1 = df_check_a0[(df_check_a0['studenttype'] == 'impoverished children')].sort_values(['total'], ascending=True)
4 df_check_a0_2 = df_check_a0_1[(df_check_a0_1['level'] == 'M1')].sort_values(['total'], ascending=False).reset_index()
5 #df_check_a0_2
6
7 #รวมประเภท impoverished children ระดับชั้น M1 ทุกโรงเรียนของเทอม 2
8 df_check_a1 = df2_1.loc[:, ['schoolname', 'yearterm', 'studenttype', 'level', 'total']]
9 df_check_a1_1 = df_check_a1[(df_check_a1['studenttype'] == 'impoverished children')].sort_values(['total'], ascending=True)
10 df_check_a1_2 = df_check_a1_1[(df_check_a1_1['level'] == 'M1')].sort_values(['total'], ascending=False).reset_index()
11 #df_check_a1_2
12
13 #รวมประเภท impoverished children ระดับชั้น M1 ทุกโรงเรียนของเทอม 3
14 df_check_a2 = df3_1.loc[:, ['schoolname', 'yearterm', 'studenttype', 'level', 'total']]
15 df_check_a2_1 = df_check_a2[(df_check_a2['studenttype'] == 'impoverished children')].sort_values(['total'], ascending=True)
16 df_check_a2_2 = df_check_a2_1[(df_check_a2_1['level'] == 'M1')].sort_values(['total'], ascending=False).reset_index()
17 #df_check_a2_2
18
19 print(df_check_a0_2)
20 print(df_check_a1_2)
21 print(df_check_a2_2)
```

	index	schoolname	yearterm	\
0	402	nongsua	2564/1	
1	186	dipangkornwittayapat mattayomwathasankaset	2564/1	
2	354	thanyarat	2564/1	
3	90	pathumthani nanthamunibamrung	2564/1	

```
In [40]: 1 #หาค่าเฉลี่ยประเภท impoverished children ระดับชั้น M1 ทุกโรงเรียนทั้ง 3 เทอม
2 df_h_h1 = pd.merge(df_check_a0_2, df_check_a1_2, left_on='schoolname', right_on='schoolname', how='left')
3 df_h_h2 = pd.merge(df_h_h1, df_check_a2_2, left_on='schoolname', right_on='schoolname', how='left')
4 df_h_h3 = df_h_h2.drop(['index_x', 'yearterm_x', 'index_y', 'yearterm_y', 'studenttype_y', 'level_y', 'yearterm', 'studenttype_x'])
5 df_h_h4 = df_h_h3.rename(columns={'studenttype_x': 'studenttype', 'level_x': 'level', 'total_x': 'total_1/64', 'total_y': 'total_2/64'})
6 df_h_h5 = pd.merge(df_h_h4, df_h_h4, left_on='schoolname', right_on='schoolname', how='left')
7 df_h_h6 = df_h_h5.sort_values(['total_1/64'], ascending=False).reset_index(drop=False)
8 df_h_h7 = df_h_h6.drop(['index'], axis=1)
9 df_h_h8 = df_h_h7.fillna(method='ffill')
10 df_h_h8['mean_type impoverished children_M1_all sc'] = df_h_h8[['total_1/64', 'total_2/64', 'total_3/64']].agg(['mean'])
11 df_h_h9 = df_h_h8.loc[:, ['schoolname', 'studenttype', 'level', 'mean_type impoverished children_M1_all sc']].round()
12 df_h_h9
```

	schoolname	studenttype	level	mean_type impoverished children_M1_all sc
0	nongsua	impoverished children	M1	267.0
1	dipangkornwittayapat mattayomwathasankaset	impoverished children	M1	166.0
2	thanyarat	impoverished children	M1	139.0
3	pathumthani nanthamunibamrung	impoverished children	M1	136.0
4	triamudomsuksanomklao pathumthani	impoverished children	M1	110.0
5	thammasatkhlongluangwithayakhom	impoverished children	M1	101.0
6	her royal highness princess soamsawali	impoverished children	M1	79.0
7	chaiyasithawat	impoverished children	M1	73.0
8	debsirinklongsam pathumthani	impoverished children	M1	60.0
9	horwang pathumthani	impoverished children	M1	52.0
10	pathumwilai	impoverished children	M1	16.0
11	buakaewkesorn	impoverished children	M1	31.0
12	matthayom sangkeet wittaya bangkok	impoverished children	M1	29.0
13	nawamintharachinitit suankularb wittayalai pat...	impoverished children	M1	26.0
14	soontharomettaprachasan	impoverished children	M1	24.0
15	kanaratbamrung pathumthani	impoverished children	M1	23.0
16	triamudomsuksa pattanakarn pathumthani	impoverished children	M1	21.0
17	suankularbwittayalai rangsit	impoverished children	M1	8.0
18	saipanyarangsit	impoverished children	M1	0.0
19	princess chulabhorn science pathum thani	impoverished children	M1	0.0

```

In [42]: 1 #สร้างแผนภาพ
2 df_E = pd.DataFrame(df_h_h9)
3 plt.figure(figsize=[18,8],dpi=300)
4 plt.bar(df_E['schoolname'],df_E['mean_type impoverished children_M1_all sc'])
5 plt.style.use("seaborn-deep")
6 plt.title('Number of students level "M1" are impoverished children',size=25)
7 plt.xlabel("School name",size=15)
8 plt.ylabel("Number of students",size=15)
9 plt.xticks(rotation=90)
10 plt.show()

```

