In [1] : import pandas as pd

In [2] : df =
pd.read_csv('https://query.data.world/s/f2yvvjxauuscnxmasmho4n4sptwlui')

In [3] : df.head ()

DBN	Grade	Year	Demographic	Number Tested	Mean Scale Score	Num Level 1	Pct Level 1	Num Level 2	Pct Level 2	Num Level 3	Pct Level 3	Num Level 4	Pct Level 4	Num Level 3 and 4	Pct Level 3 and 4
0	01M015	3	2006	All Students	39	667	2	5.1	11	28.2	20	51.3	6	15.4	26	66.7
1	01M015	3	2007	All Students	31	672	2	6.5	3	9.7	22	71	4	12.9	26	83.9
2	01M015	3	2008	All Students	37	668	0	0	6	16.2	29	78.4	2	5.4	31	83.8
3	01M015	3	2009	All Students	33	668	0	0	4	12.1	28	84.8	1	3	29	87.9
4	01M015	3	2010	All Students	26	677	6	23.1	12	46.2	6	23.1	2	7.7	8	30.8

In [4] : df.tail ()

DBN	Grade	Year	Demographic	Number Tested	Mean Scale Score	Num Level 1	Pct Level 1	Num Level 2	Pct Level 2	Num Level 3	Pct Level 3	Num Level 4	Pct Level 4	Num Level 3 and 4	Pct Level 3 and 4	
33456	75X723	All Grades	2008	All Students	211	NaN	139	65.9	46	21.8	26	12.3	0	0	26	12.3
33457	75X723	All Grades	2009	All Students	209	NaN	88	42.1	87	41.6	30	14.4	4	1.9	34	16.3
33458	75X723	All Grades	2010	All Students	242	NaN	157	64.9	75	31	10	4.1	0	0	10	4.1
33459	75X723	All Grades	2011	All Students	229	617	153	66.8	67	29.3	8	3.5	1	0.4	9	3.9
33460	75X723	All Grades	2012	All Students	213	620	142	66.7	67	31.5	4	1.9	0	0	4	1.9

In [5] : df.info ()

```
<class 'pandas.core.frame.DataFrame'> RangeIndex: 33461 entries, 0 to 33460 Data columns (total 16 columns): DBN 33461 non-null object Grade 33461 non-null object Year 33461 non-null int64 Demographic 33461 non-null object Number Tested 33461 non-null int64 Mean Scale Score 28106 non-null object Num Level 1 33461 non-null object Pct Level 1 33461 non-null object Num Level 2 33461 non-null object Pct Level 2 33461 non-null object Num Level 3 33461 non-null object Pct Level 3 33461 non-null object Num Level 4 33461 non-null object Pct Level 4 33461 non-null
```

```
object Num Level 3 and 4 33461 non-null object Pct Level 3 and 4
33461 non-null object dtypes: int64(2), object(14) memory usage:
4.1+ MB
```

In [6] : df.describe ()

	Year	Number Tested
count	33461.000000	33461.000000
mean	2009.066137	172.754132
std	1.991814	197.130818
min	2006.000000	1.000000
25%	2007.000000	65.000000
50%	2009.000000	106.000000
75%	2011.000000	204.000000
max	2012.000000	2282.000000

In [7] : df. df.groupby('Num Level 1').mean ()

	Year	Number Tested
Num Level 1		
0	2009.098901	90.608370
1	2009.097332	100.287671
10	2009.176471	149.746524
100	2008.642857	658.071429

Num Level 1	Year	Number Tested
101	2007.769231	486.307692
...	...	...
96	2007.583333	470.833333
97	2009.222222	594.333333
98	2009.187500	557.937500
99	2008.722222	646.111111
s	2008.951318	2.531440

269 rows × 2 columns

In [8] : import matplotlib.pyplot as plt

In [9] : x = (26, 31, 33, 37, 39)

print (x)

```
(26, 31, 33, 37, 39)
```

In [10] : y = (677, 672, 668, 668, 667)

print (y)

```
(677, 672, 668, 668, 667)
```

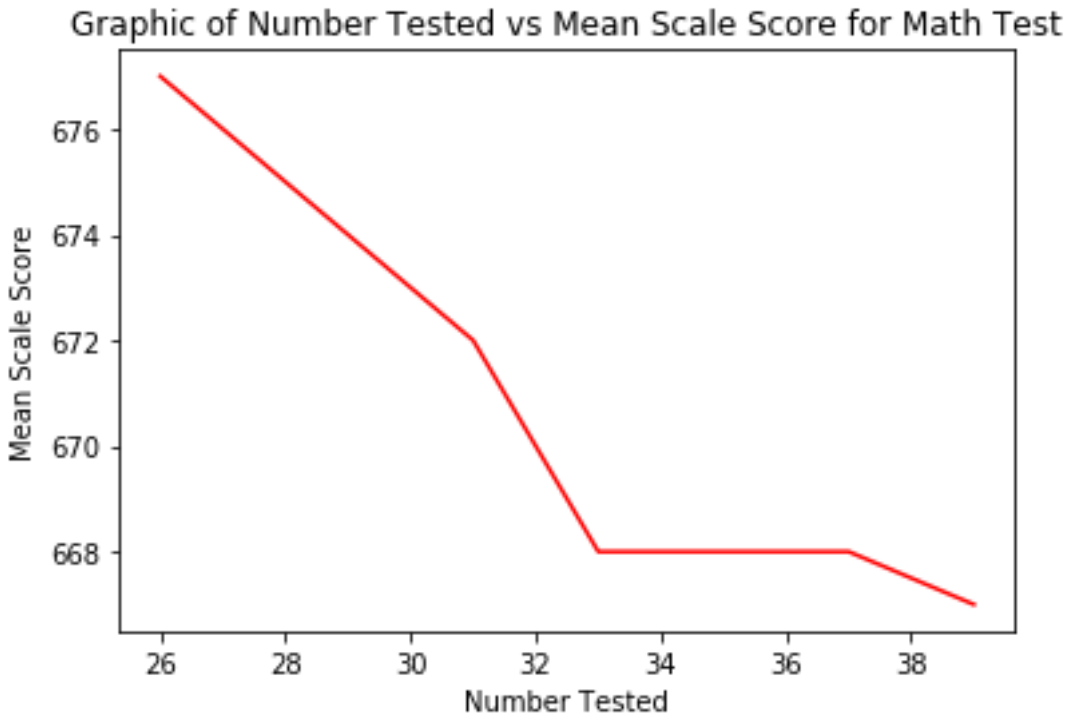
In [11] : plt.plot(x,y, '-', color='red')

plt.title('Graphic of Number Tested vs Mean Scale Score for Math Test')

plt.ylabel('Mean Scale Score')

plt.xlabel('Number Tested')

```
Text(0.5, 0, 'Number Tested')
```



A. Penjelasan Fungsi

1. Import pandas as pd
- agar dapat membaca file csv, txt, tsv dan lain-lain
2. df = pd.read_csv('https://query.data.world/s/f2yvvjxauuscnxmasmho4n4sptwlui')
- memasukkan data berupa link csv
3. df.head ()
- fungsi head tanpa argumen menghasilkan lima baris pertama data dari frame data
4. df.tail ()
- fungsi tail tanpa argumen menghasilkan lima baris terakhir data dari frame data
5. df.info ()
- utuk mengetahui informasi lebih dari data set
6. df.describe ()
- menampilkan data set lebih spesifik dan singkat
7. df. df.groupby('Num Level 1').mean ()

- untuk mencari nilai mean (rata-rata) dari kolom Num Level 1

8. import matplotlib.pyplot as plt

- agar dapat menggunakan fungsi plot

9. x = (26, 31, 33, 37, 39)

print (x)

- Input nilai x

10. y = (677, 672, 668, 668, 667)

print (y)

- Input nilai y

11. plt.plot(x,y, '-', color='red')

plt.title('Graphic of Number Tested vs Mean Scale Score for Math Test')

plt.ylabel('Mean Scale Score')

plt.xlabel('Number Tested')

- Input plot grafik (x,y)
- Memberi judul grafik 'Graphic of Number Tested vs Mean Scale Score for Math Test’
- Memberi nama sumbu y 'Mean Scale Score'
- Memberi nama sumbu x 'Number Tested'

Grafik berbentuk garis tidak stabil karena data bukan merupakan fungsi matematika. Data merupakan hasil riset rata – rata nilai matematika siswa dengan jumlah pengambilan tes ke-berapa. Data tersebut diambil karena dirasa cukup mudah dalam pengaplikasian fungsi iris dan plotting grafik.