

**KORELASI ANTARA NILAI PISA: MEAN PERFORMANCE ON THE READING SCALE
DENGAN NILAI INDEX HARAPAN HIDUP (KESEHATAN)**

ENIKE DEWINTA SEMBIRING (161402026)

AGNES MANURUNG (161402131)



PROGRAM STUDI TEKNOLOGI INFORMASI

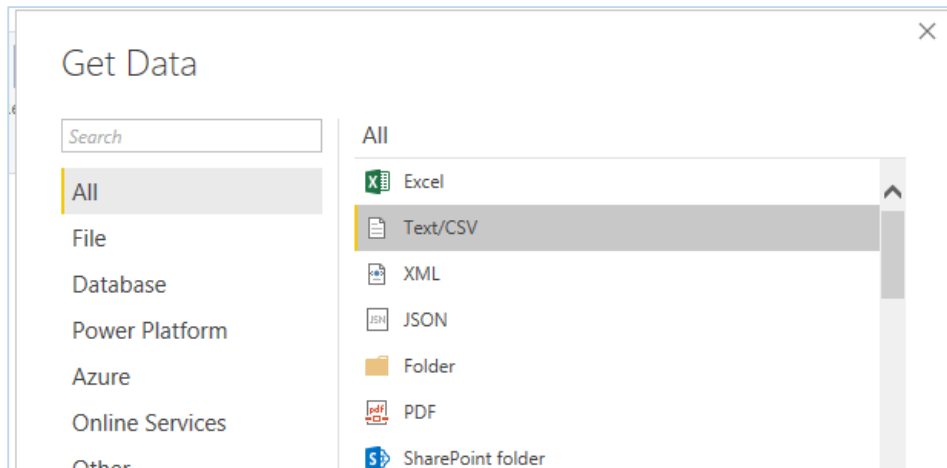
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS SUMATERA UTARA

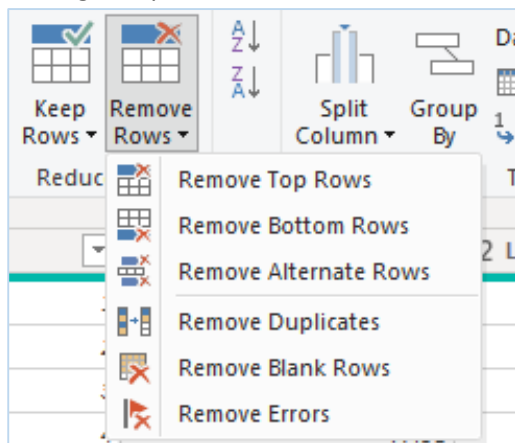
TAHUN AJARAN 2019/2020

A. Menyiapkan Data

1. Masukkan dataset yang diperlukan. Dataset yang dimaksud memiliki format .csv. Masukkan data melalui Tab Home > Get Data > Text/CSV



2. Lakukan pengeditan terhadap dataset '**2016**' yang telah dimasukkan dengan menggunakan fitur **Edit Query > Remove Rows**. Pengeditan dilakukan untuk menghilangkan data yang tidak perlu seperti baris yang berisi data kosong ataupun data error.



3. Kemudian juga pengeditan terhadap dataset '**Pisa Score**'.
 - a. Gunakan baris pertama dari dataset sebagai header dari tabel dan ganti tipe data dari kolom '**2015**' menjadi desimal.
 - b. Hapus kolom yang kosong, baris yang berisi data yang kosong dan error, kemudian hapus beberapa baris terakhir yang tidak diperlukan.
 - c. Ganti nama-nama negara berikut agar menyamakannya dengan nama negara dari tabel '**2016**':
 - (1) Hong Kong SAR, China → Hong Kong
 - (2) Korea, Rep. → South Korea
 - (3) Macedonia, FYR → Macedonia
 - (4) Russian Federation → Russia
 - (5) Slovak Republic → Slovakia
4. Simpan query dan tutup

B. JOIN TABEL

1. Buatlah relasi antara tabel '2016' dengan 'Pisa Score' dengan fitur Manage Relationship. Buatlah relasi baru antar kedua tabel dengan menggunakan 'Country' sebagai key.

Pisa Score

Country Name	Country Code	Series Name	Series Code	2015
Albania	ALB	PISA: Mean performance on the mathematics scale	LO.PISA.MAT	413.157
Algeria	DZA	PISA: Mean performance on the mathematics scale	LO.PISA.MAT	359.6062
Argentina	ARG	PISA: Mean performance on the mathematics scale	LO.PISA.MAT	409.0333

2016

Country	Region	Happiness Rank	Happiness Score	Lower Confidence Interval	Upper Confidence Inter
Denmark	Western Europe	1	7.526	7.46	
Switzerland	Western Europe	2	7.509	7.428	
Iceland	Western Europe	3	7.501	7.333	

Cardinality

Many to one (*:1)

Cross filter direction

Single

☒ Make this relationship active
 ☐ Apply security filter in both directions

☐ Assume referential integrity

2. Buat tabel join dengan menggunakan DAX. Pilih tab Modelling > New Table, lalu ketikkan koding DAX berikut

```
Kore =
SUMMARIZE(FILTER(NATURALINNERJOIN('2016','Pisa Score'),
'Pisa Score'[Series Name]="PISA: Mean performance on the reading scale"),
'2016'[Country],'Pisa Score'[2015],'2016'[Health (Life Expectancy)])
```

Keterangan:

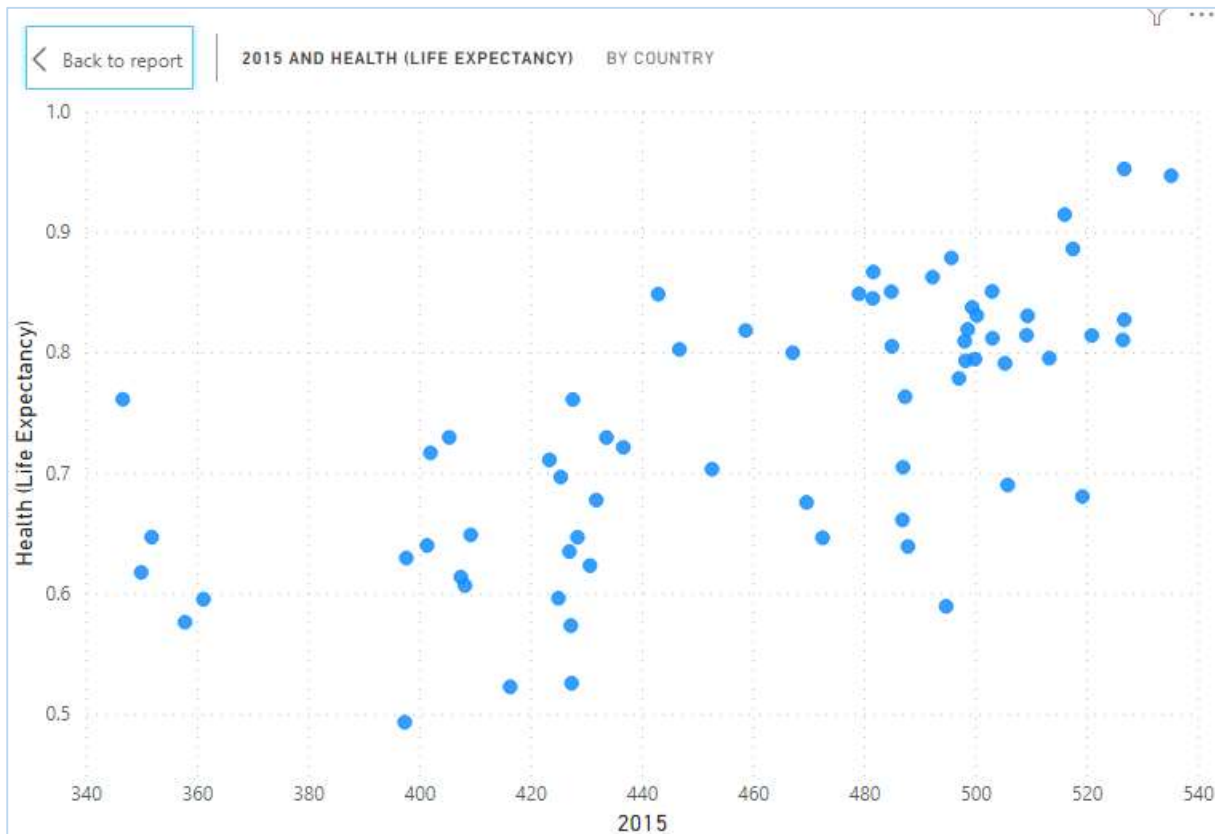
- a. Kore, merupakan nama tabel baru yang akan terbentuk nantinya
- b. NATURALINNERJOIN, memiliki dua parameter yaitu nama-nama tabel yang akan di join. Untuk melakukan join, tabel-tabel tersebut harus memiliki relasi terlebih dahulu
- c. FILTER, memiliki dua parameter yaitu tabel yang akan di filter, serta ekspresi filter yang diinginkan. Dalam hal ini, tabel yang ingin difilter yaitu tabel hasil join dari fungsi NATURALINNERJOIN dan di filter sehingga hasil yang keluar nantinya hanya baris tertentu saja (nilai PISA berdasarkan skala keterampilan membaca)
- d. SUMMARIZE, untuk memunculkan kolom-kolom apa saja yang diinginkan.

Hasil dari fungsi DAX tersebut akan menghasilkan tabel 'Kore' dengan tampilan sebagai berikut:

1 Kore = SUMMARIZE(FILTER(NATURALINNERJOIN('2016','Pisa Score'),		
[Country],'Pisa Score'[2015],'2016'[Health (Life Expectancy)])		
Country	Health (Life Expectancy)	2015
Albania	0.73007	405.2588
Algeria	0.61804	349.8593
Argentina	0.69711	425.3031
Australia	0.8512	502.9006
Austria	0.80565	484.8656
Belgium	0.81959	498.5242
Brazil	0.61415	407.3486
Bulgaria	0.67806	431.7175
Canada	0.8276	526.6678
Chile	0.81883	458.5709
Colombia	0.59659	424.9052
Costa Rica	0.76146	427.4875

C. KORELASI

1. Korelasi bisa dilihat melalui scatter plot. Buatlah scatter plot dengan menggunakan nilai '**Kore[2015]**' sebagai X-axis, '**Kore[Health (Life Expectancy)]**' sebagai Y-axis, dan '**Kore[Country]**' sebagai detail. Scatter plot yang dihasilkan adalah sebagai berikut:



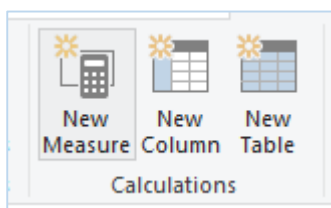
2. Korelasi juga dapat dilihat apakah berbanding lurus (positif), terbalik (negatif), atau tidak memiliki hubungan berdasarkan nilai koefisien korelasi nya. Untuk menentukan koefisien relasi dapat dilakukan dengan menggunakan rumus koefisien korelasi Pearson (Pearson Product Moment).

$$r = \frac{n \sum xy - (\sum x) (\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2) (n \sum y^2 - (\sum y)^2)}}$$

Dimana:

- n = banyaknya pasangan data X dan Y
- $\sum x$ = total jumlah dari variabel X
- $\sum y$ = total jumlah dari variabel Y
- $\sum x^2$ = kuadrat dari total jumlah variabel X
- $\sum y^2$ = kuadrat dari total jumlah variabel Y
- $\sum xy$ = hasil perkalian dari total jumlah variabel X dan variabel Y

3. Buat sebuah measure baru dengan mengaplikasikan rumus Pearson dengan menggunakan DAX.



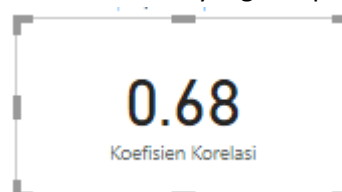
Ketikkan fungsi DAX berikut di tempat yang telah disediakan.

```

Koefisien Korelasi =
VAR Tabel_korelasi =
    FILTER (
        ADDCOLUMNS (
            VALUES ( Kore[Country] ),
            "Nilai_X", CALCULATE ( SUM ( Kore[2015] ) ),
            "Nilai_Y", CALCULATE ( SUM ( Kore[Health (Life Expectancy)] ) )
        ),
        AND (
            NOT ( ISBLANK ( [Nilai_X] ) ),
            NOT ( ISBLANK ( [Nilai_Y] ) )
        )
    )
VAR Count_Items =
    COUNTROWS ( Tabel_korelasi )
VAR Sum_X =
    SUMX ( Tabel_korelasi, [Nilai_X] )
VAR Sum_X2 =
    SUMX ( Tabel_korelasi, [Nilai_X] ^ 2 )
VAR Sum_Y =
    SUMX ( Tabel_korelasi, [Nilai_Y] )
VAR Sum_Y2 =
    SUMX ( Tabel_korelasi, [Nilai_Y] ^ 2 )
VAR Sum_XY =
    SUMX ( Tabel_korelasi, [Nilai_X] * [Nilai_Y] )
VAR Pearson_Numerator =
    Count_Items * Sum_XY - Sum_X * Sum_Y
VAR Pearson_Denominator_X =
    Count_Items * Sum_X2 - Sum_X ^ 2
VAR Pearson_Denominator_Y =
    Count_Items * Sum_Y2 - Sum_Y ^ 2
VAR Pearson_Denominator =
    SQRT ( Pearson_Denominator_X * Pearson_Denominator_Y )
RETURN
    DIVIDE ( Pearson_Numerator, Pearson_Denominator )

```

Tampilkan nilai koefisien korelasi dengan menggunakan Card Visualization dengan measure Koefisien Korelasi sebagai field nya. Berikut adalah nilai koefisien korelasi yang didapat.



D. KESIMPULAN

Berdasarkan scatter plot dan koefisien korelasi yang didapat, dapat dilihat bahwa korelasi antara nilai PISA: Mean performance on the reading scale dengan nilai index harapan hidup (kesehatan) memiliki korelasi linear positif yang bersifat sedang.