**Rencana Tugas Mandiri Ke-3**

**Mata Kuliah Analisis Data Eksploratif**



**Mata Kuliah Analisis Data Eksploratif**

**Statistika Deskriptif dan Grouping Dataset**

**KELAS A**

**DISUSUN OLEH :**

AHMAD ARDHY ANSYAH ( 20083010008 )

**DOSEN PENGAMPU:**

TRESNA MAULANA FAHRUDIN,/ S.S.T., MT ( 20219930501200 )

PROGRAM STUDI SAINS DATA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”

JAWA TIMUR

2022

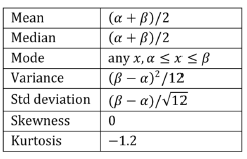
1. Statistika Deskriptif (Distribusi Uniform, Distribusi Normal, Skewness, dan Kurtosis)
   1. Distribusi Uniform

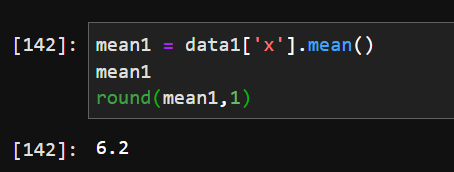
Berapa persentase probabilitas *x* < 5 berdasarkan persamaan dalam distribusi uniform dengan menggunakan kode *script* pemrograman Python secara *scratch*

* + 1. Carilah nilai a dan b terlebih dahulu

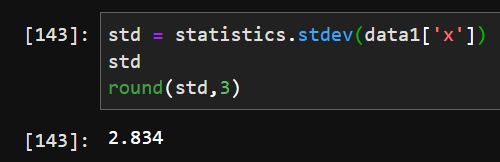
Sebelum menghitung a dan b, dibutuhkan beberapa langkah yang akan digunakan sebagai pembantu untuk menemukan a dan b, yaitu menghitung mean, standar deviasi, a+b, dan b-a dari data dengan bekal rumus berikut

* + 1. Gunakan persamaan *mean* dan *standard deviation* sebagai bekal untuk mendapatkan nilai *a* dan *b* sesuai dengan sifat dasar statistika dalam distribusi uniform berikut

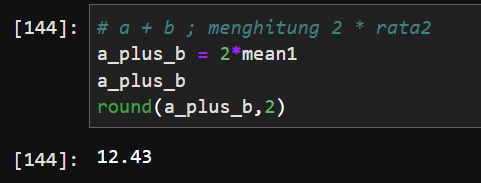




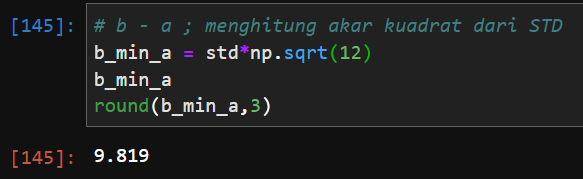
Menghitung mean dari data



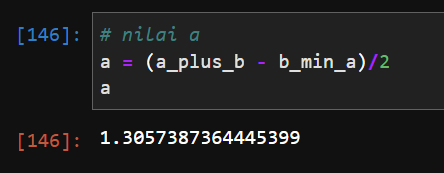
Menghitung standar deviasi dari data

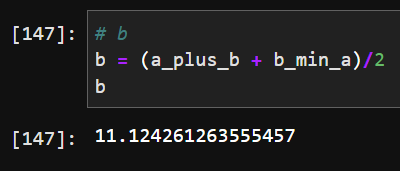


Menghitung a + b



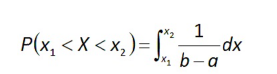
Setelah didapatkan nilai a+b dan b-a, dari situ kita dapat menentukan nilai a dan b dengan rumus

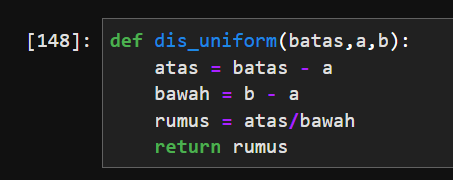
 Didapatkan nilai a nya

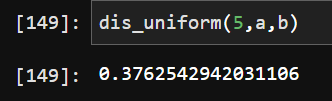


Didapatkan nilai b nya. Setelah semua variabel didapatkan baru kita buat sebuah fungsi untuk menghitung distribusi uniform.

Selanjutnya hitung probabilitas distribusi uniform berdasarkan nilai *a* dan *b* yang sudah didapatkan menggunakan persamaan berikut

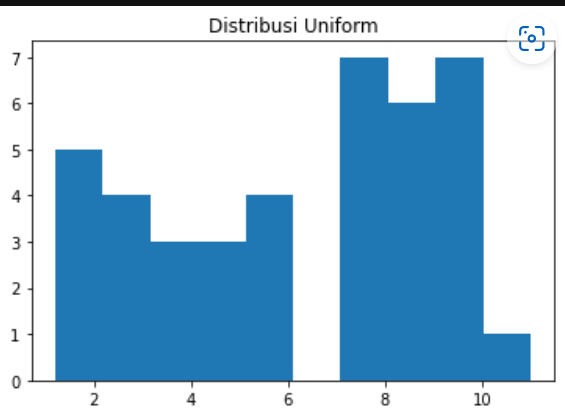






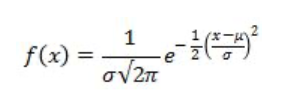
Didapatkan hasil dari P(X<5) dari data ph ini adalah 0.37625 atau dapat disebut juga Probabilitas dari data ini yang bernilai kurang dari (<) 5 adalah **37.625%.**

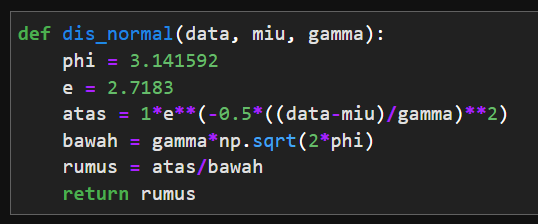
Visualisasi **DISTRIBUSI UNIFORM** dari data ph.csv ini adalah



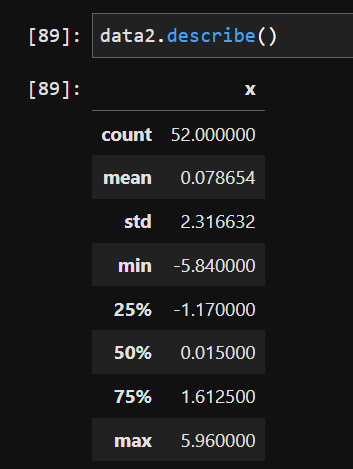
Artinya data ini BERDISTRIBUSI UNIFORM

* 1. Distribusi Normal
     1. Hitunglah persamaan *f(x)* dari nilai x di atas berdasarkan persamaan dalam distribusi normal dengan menggunakan kode *script* pemrograman Python secara *scratch.*
     2. Gunakan persamaan distribusi normal sebagai berikut:

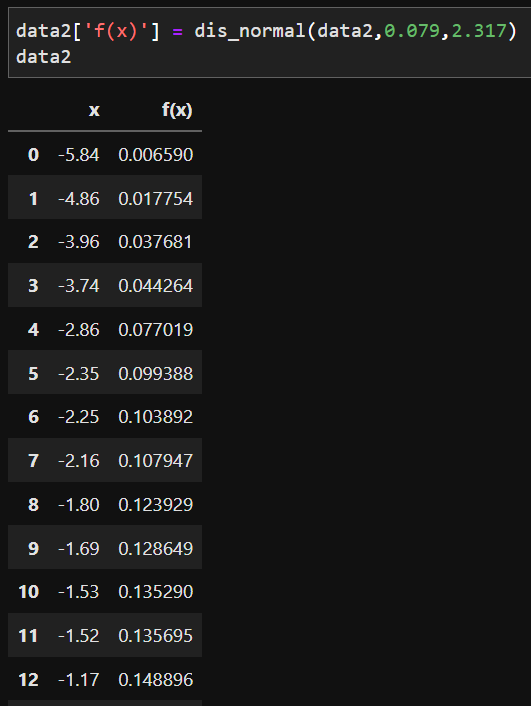
**

 Data adalah data yang akan dihitung untuk mengetes apakah data berdistribusi normal atau tidak.

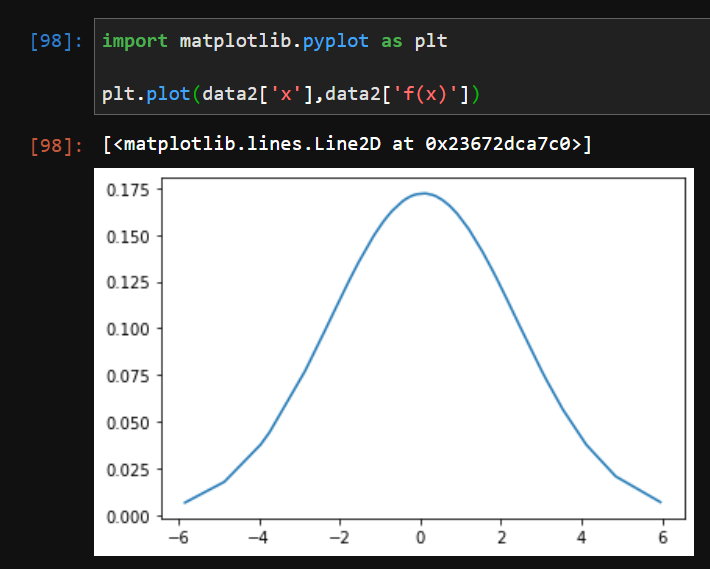
Miu adalah nilai rata – rata dari data, dan gamma adalah nilai standar deviasi dari data



Hasil distribusi normal dari data dimasukkan ke dalam nilai f(x) di dalam tabel data



Visualisasikan antara nilai x dan f(x) menggunakan diagram garis, periksalah apakah dataset berdistribusi normal atau tidak. Jelaskan beserta alasannya.



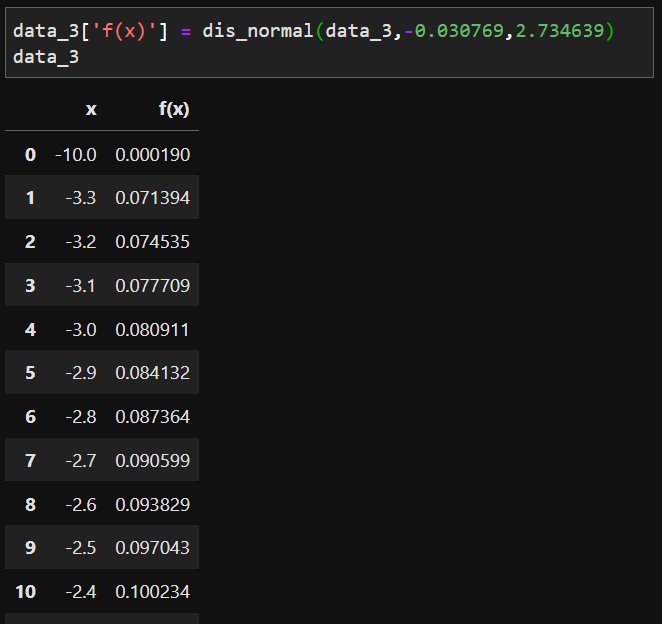
Dari hasil visualisasi didapatkan hasil bahwa grafik berdistribusi normal. Sebab terlihat dari grafik tersebut pas di tengah – tengah data.

* 1. Skewness
     1. Hitunglah persamaan f(x) dari nilai x di atas berdasarkan persamaan dalam distribusi normal dengan menggunakan kode script pemrograman Python secara scratch

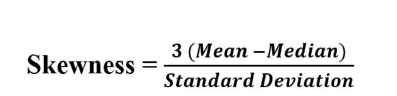
Sebelum menghitung distribusi normal diperlukan nilai miu (rata-rata data) dan gamma (standar deviasi dari data).

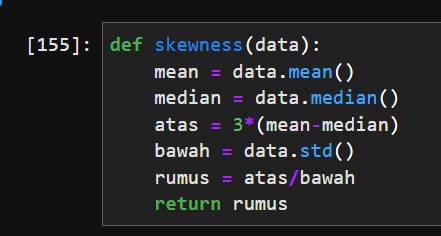


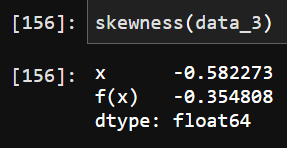
Setelah itu baru hitung distribusi normal dari data, dengan memanggil fungsi yang sudah dibuat sebelumnya dan memasukkan input-an miu dan gamma. Hasilnya akan dimasukkan ke tabel data f(x)



Setelah itu baru buat sebuah fungsi untuk menghitung Distribusi Skewness nya dengan bekal rumus berikut :



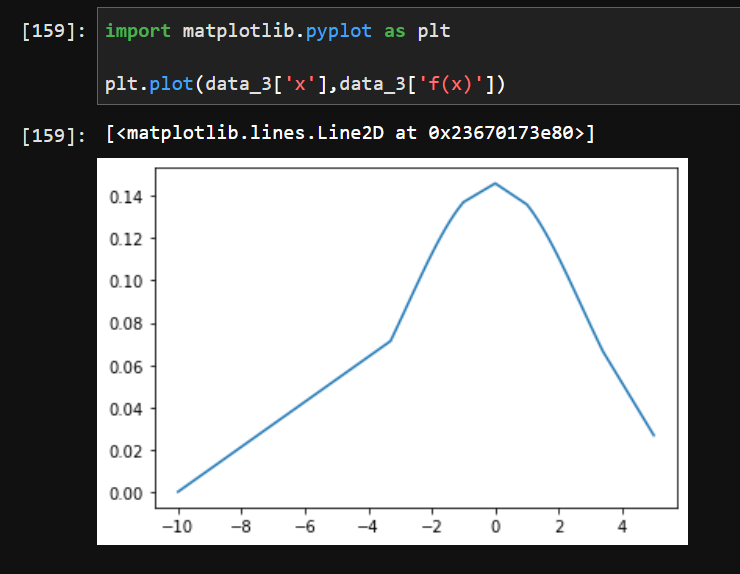




Hasil dari distribusi Skewness nya adalah -0.582273,

* + 1. Apakah ada kecondongan distribusi data, ke arah positif atau ke arah negatif ? Jelaskan beserta alasannya.

Dari hasil pengujian dan visualisasi grafik distribusi Skewness ini adalah cenderung condong ke negatif. Karena didapatkan hasil Mean<Median<Modus.

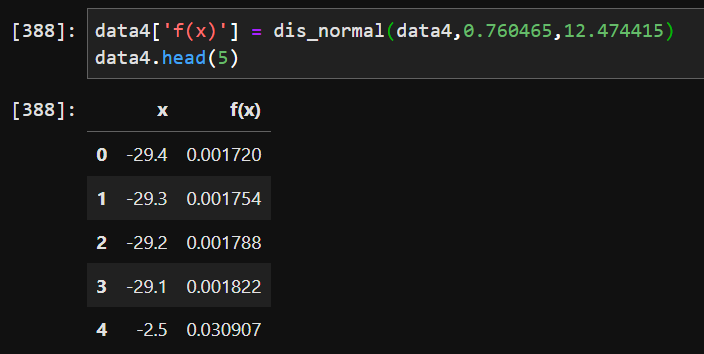


* 1. Kurtosis
     1. Hitunglah persamaan f(x) dari nilai x di atas berdasarkan persamaan dalam distribusi normal dengan menggunakan kode script pemrograman Python secara scratch

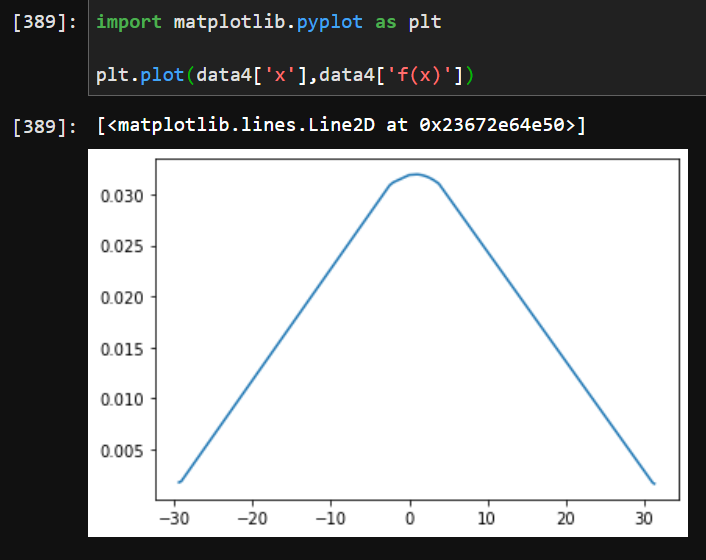
Sebelum menghitung Distribusi Normal dari data, langkah yang diperlukan sebelumnya adalah mengetahui Mean, Standar deviasi dari data.



Setelah itu baru dihitung nilai Distribusi normal nya, dan hasilnya akan dimasukkan ke dalam f(x).



Setelah itu visualisasikan antara nilai x dan f(x)

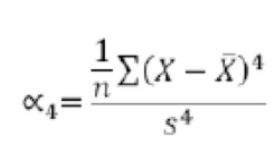


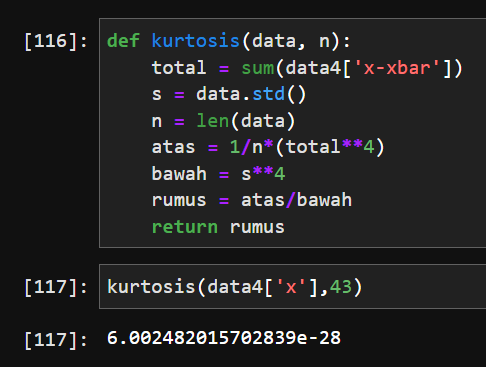
Dari hasil visualisasi didapatkan bentuk grafik seperti itu yang mana grafik tersebut lebih mengarah ke playtkurtic, sebab grafik berbentuk gemuk dan ekornya sangat tipis dibandingkan dengan distribusi normal.

Setelah itu baru kita hitung nilai distribusi Kurtosis nya, dengan membuat sebuah fungsi dengan pedoman rumus

Sebelum itu dibutuhkan data hasil hitungan antara x-xbar, kita buat sebuah columns baru di dalamnya.







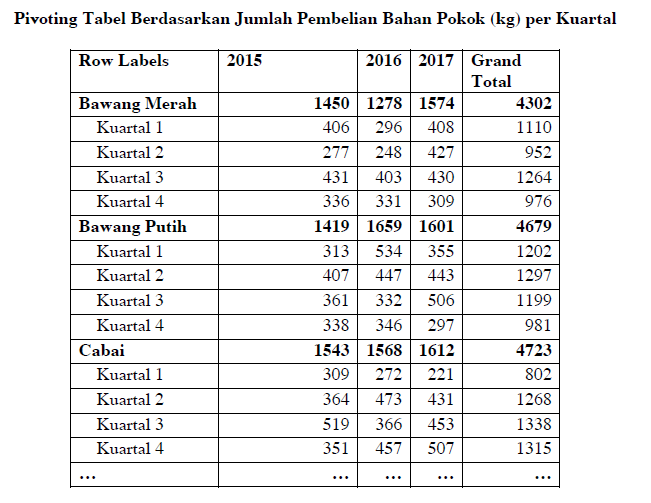
Didapatkan hasil nilai Distribusi Kurtosis dari data ini adalah 6.0024e-28, yang mana hasil ini akurat dengan hasil visualisasi sebelumnya.

1. Grouping Dataset : Pivoting Table

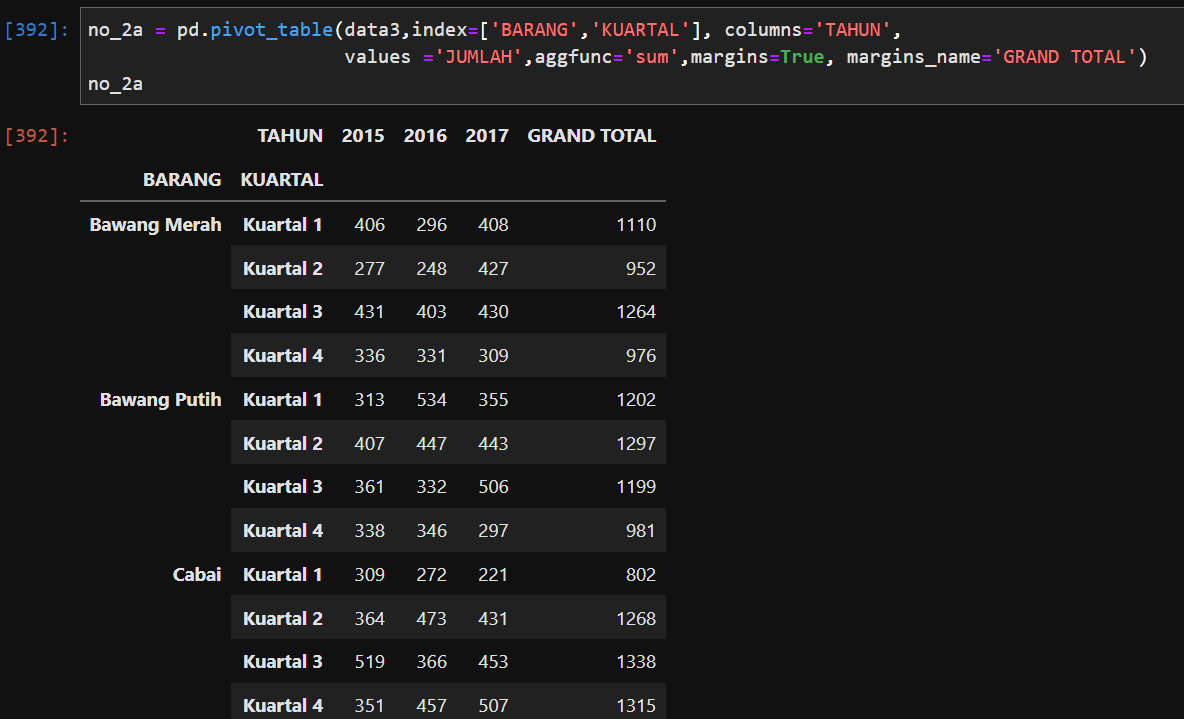
Dataset awal



Soal 2a : Lakukanlah pivoting tabel menggunakan kode *script* pemrograman Python pembelian bahan pokok makanan yang dikirimkan oleh supplier pada kuartal tertentu menjadi tabel berikut:

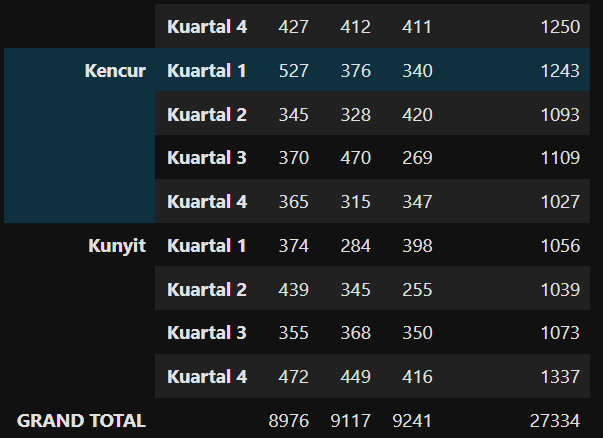


Cara membuat pivot table tersebut menggunakan library Pandas Pivot Table, menggunakan kode script berikut



Maksud dari kode script ini adalah pada bagian awal adalah dengan memanggil data yang akan dipakai, datanya yaitu data3. Lalu index adalah data yang akan dibuat menjadi group, yaitu data barang dan kuartal. Lalu columns adalah data yang akan dibuat menjadi bentuk columns, yaitu data TAHUN. Lalu values adalah data yang akan dihitung dalam columns, yaitu data JUMLAH. Lalu aggfunc adalah sebuah agregasi akan dilakukan apakah pada nilai values sebelumnya, pada case kali ini akan dilakukan penjumlahan dari setiap kuartal, berarti di-inputkan ‘sum’ pada aggfunc. Lalu saya ingin membuat TOTAL Penjumlahan dari data, yaitu dengan melakukan margins dengan margins\_name nya yaitu ‘GRAND TOTAL’.

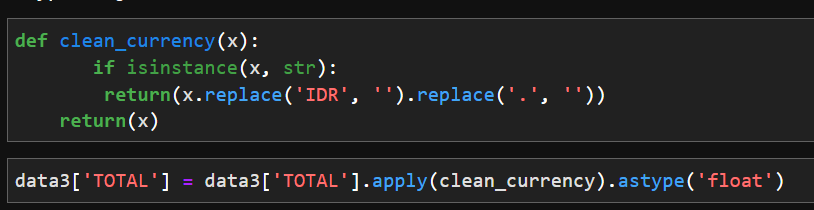
Begini hasil Pivot Table nya

Soal 2b : Pivoting Tabel Berdasarkan Total Pembelian Bahan Pokok (Rupiah) per Kuartal

Pada kasus ini caranya hampir sama dengan langkah pada soal sebelumnya, namun perlu dilakukan beberapa proses untuk mempersiapkan data.

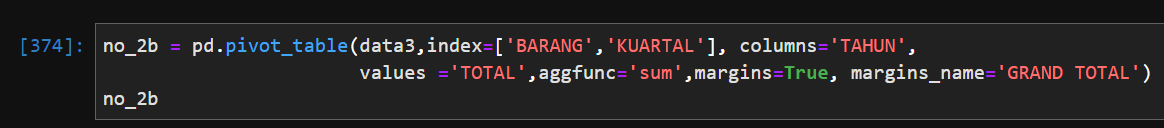
Proses nya yaitu menghilangkan kata IDR di data TOTAL, cara nya dengan membuat sebuah fungsi untuk menghilangkan kata IDR tersebut, dan merubah type datanya menjadi float agar dapat dilakukan proses perhitungan



Begini hasilnya setelah dilakukan preprocess data



Setelah itu baru buat Pivot Table nya :



Maksud dari code script tersebut adalah, menggunakan library Pandas Pivot Table.. Lalu index adalah data yang akan dibuat menjadi group, yaitu data barang dan kuartal. Lalu columns adalah data yang akan dibuat menjadi bentuk columns, yaitu data TAHUN. Lalu values adalah data yang akan dihitung dalam columns, yaitu data TOTAL. Lalu aggfunc adalah sebuah agregasi akan dilakukan apakah pada nilai values sebelumnya, pada case kali ini akan dilakukan penjumlahan dari setiap kuartal, berarti di-inputkan ‘sum’ pada aggfunc. Lalu saya ingin membuat Total Penjumlahan dari data, yaitu dengan melakukan margins dengan margins\_name nya yaitu ‘GRAND TOTAL’.

Begini hasil Pivot Table nya

