Nama : Aditya Allaam

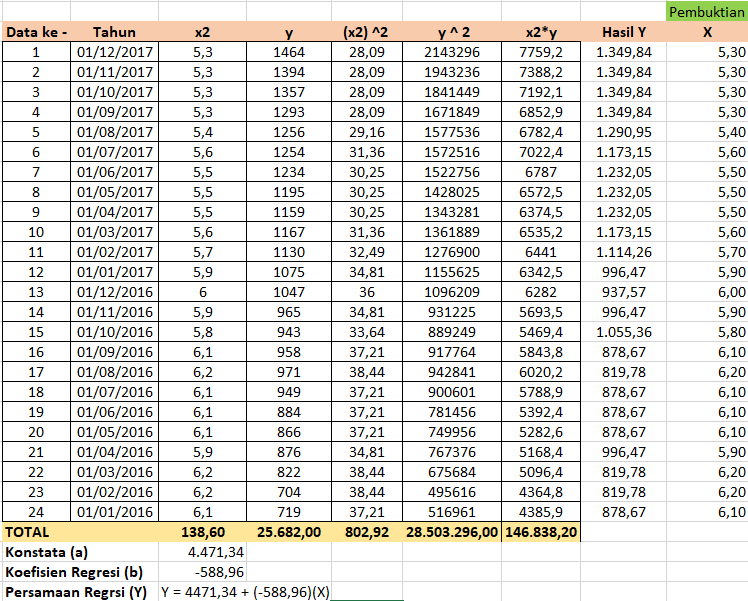
NIM : 11190910000084

Teknik Informatika’19

**Big Data**

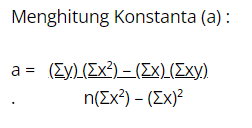
*Semua perhitungan manual terdapat di excel ‘TUGAS 10.xlsx’.*

1. **Hitung nilai linea regresi sederhana dengan nilai x2 dan y**

****

Untuk persamaan regresi memiliki bentuk perhitngan ***Y = a + bx,*** sehingga untuk mencari nilai tersebut, sebagai berikut :

1. Mencari nilai *konstata (a)* dengan menggunakan rumus,

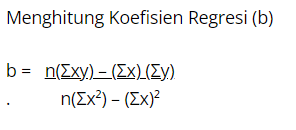


Maka terlihat pada rumus ada beberapa pencarian yang dilakukan,

* Mencari total nilai x , nilai dari variabel bebas
* Mencari total nilai x2, nilai dari variabel bebas
* Mencari total nilai perkalian x.y, nilai perkalian dari variabel bebas dan terikat

Sehingga dilkakan perhitungan untuk *konstata (a)* dan mendapatkan hasil **4.471,34**

1. Mencari nilai koefisien regresi dengan menggunakan rumus,



Maka terlihat pada rumus ada bebeapa pencarian yang dilakukan,

* Mencari total nilai perkalian x.y, nilai perkalian dari variabel bebas dan terikat
* Mencari total nilai x, nilai dari variabel bebas
* Mencari total nilai y, nilai dari variabel terkiat
* Mencari total nilai x2, nilai dari vairabel bebas

Sehingga dilakukan perhitungan untuk koefisien regresi (b) dan mendapatkan hasil

**-588, 96**

1. Mencari nilai Persamaan regresi,dengan menggunakan rumus,

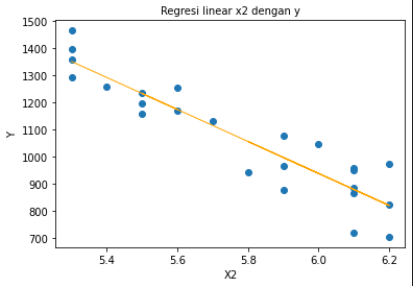


Maka akan didapatkan, ***Y = 4.4471,34 + (-588,96)(X)*** , dan pada perhitungan ini dapat menentukan nilai dari ***Y prediksi***, terlihat pada tabel perhitung Y prediksi pada kolom ***Hasil Y*** selain melakukan perhitungan nilai prediksi dapat dilakukan juga pembuktian nilai variabel bebas apakah sama atau tidak dengan nilai variabel bebas di awal.

Pembuktian nilai, konstanta dan koefisien regresi dengan menggunakan *python*



Visualisasi persebaran data, sebagai berikut,



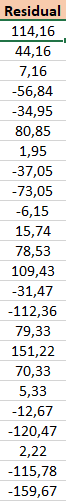
Terlihat pada visualisasi, penyebaran data dan garis lurus naik keatas ke arah kiri, ini dikarenakan nilai dari koefisien regresi (b) bernilai negaif dan garis lurus merupakan nilai dari Y prediksi (Y Hasil)

1. **Akurasi prediksi regresi linear sederhana dengan residual dan RSME,**

* Perhitungan residual dapat dilakukan dengan menggunakan rumus,

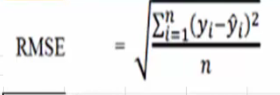
***Residual (e) = Y aktual – Y prediksi***

Sehingga akan mendapatkan hasil,



Pada nilai residual memiliki nilai positif jika data diatas garis lurus jika memiliki nilai negatif berarti data berada dibawah garis lurus.

* Perhitungan nilai RSME dapat dilakukan dengan menggunakan rumus,



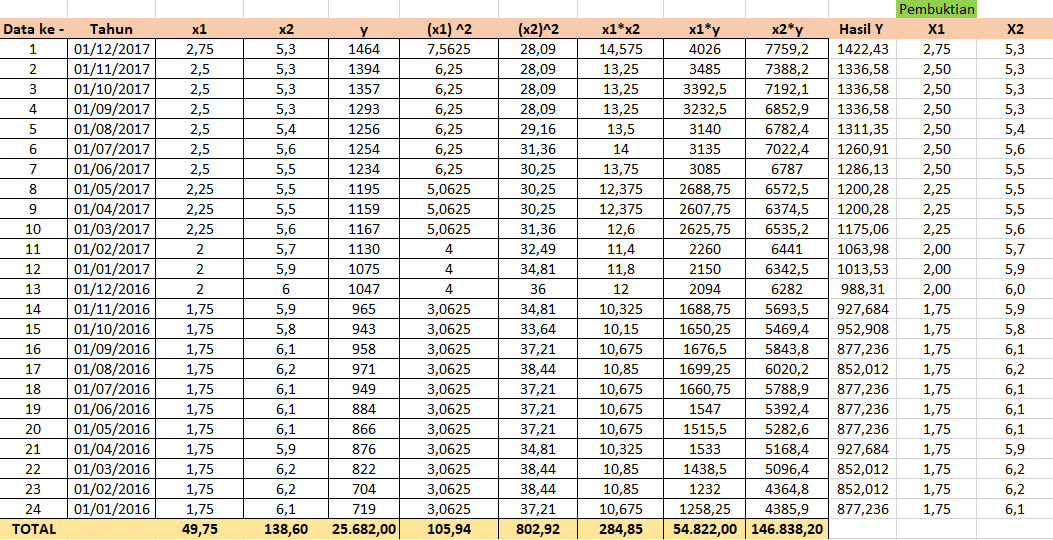
Untuk mengetahui nilai keakurasin pada data jika nilai RMSE atau nilai penyimpangannya semakin kecil maka nilai keakurasian data semakin baik.

Dengan perhitungan tersebut maka nilai RMSE mendapatkan nilai **79,71,**

1. **Kesimpulan**

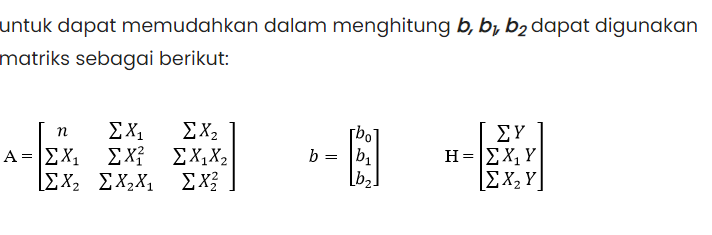
Jadi dapat dilihat bahwa semakin kecil nilai dari **x2** maka akan semakin tinggi nilai dari **Y**, sehingga nilai **x2** mempengaruhi nilai **Y** secara negatif.

1. **Hitung nilai regresi linear bergada dengan nilai x1, x2 dan y**

****

Untuk persaamaan regresi berganda memiliki bentuk perhitungan **Y = a + b1(x1) + b2(x2)….+bn(xn),** sehingga untuk mencari nilai tersebut sebagai berikut :

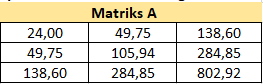
1. Menggunakan bentuk matriks dengan perhitungan,



Maka terlihat pada rumus ada beberapa pencarian yang hasrus dilakukan, sebagai berikut :

* Mencari nilai total dari X1, nilai dari variabel bebas
* Mencari nilai total dari X2, nilai dari variabel bebas
* Mencari nilai total dari X12, nilai dari variabel bebas
* Mencari nilai total dari X22, nilai dari variabel bebas
* Mencari nilai total perkalian X1.X2, nilai perkalian antar variabel bebas

Sehingga nilai tersebut akan dimasukan ke dalam bentuk matriks A, dan mendapatkan hasil sebagai berikut :



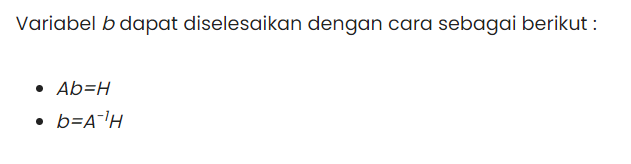
Jika nilai matriks A sudah didapatkan meka akan dilakukan perhitungan untuk mencari matriks H, sebagai berikut :

* Mencari nilai total dari Y, nilai dari variabel terikat
* Mencari nilai perkalian dari X1.Y, nilai dari perkalian variabel bebas dan terikat
* Mencari nilai perkalian dari X2.Y, nilai dari perkalian variabel bebas dan terkiat

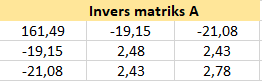
Sehingga nilai tersebut akan dimasukan ke dalam bentuk matriks H, dan mendapatkan hasil sebagai berikut :



1. Sehingga untuk pencarian nilai dari bentuk persamaan regresi berganda, sebagai berikut :



Terlihat pada rumus, bahwa bentuk matriks A, harus dilakukan *invers* sehingga akan mendapatkan nilai sebagai berikut :



Setelah nilai matriks A didaptkan nilai inversenya, lalu dapat dilakukan perkalian antara invers matriks A dengan matriks H, sehingga akan mendapatkan nilai sebagai berikut :



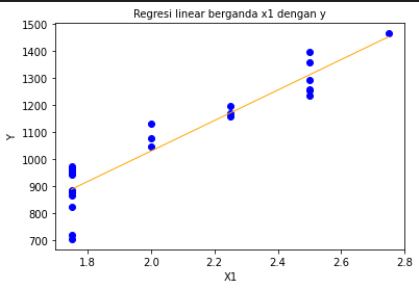
Sehingga pada hasil perhitungan matriks B yang telah didapatkan maka akan mendapatkan bentuk persamaan regresi bergada ***Y = 1814,95 + 343,40(X1) + (-252,24)(X2)*** , dan pada perhitungan ini dapat menentukan nilai dari ***Y prediksi***, terlihat pada tabel perhitung Y prediksi pada kolom ***Hasil Y*** selain melakukan perhitungan nilai prediksi dapat dilakukan juga pembuktian nilai variabel bebas apakah sama atau tidak dengan nilai variabel bebas di awal.

Pembuktian nilai dari konstata dan koefisien regresi berganda dengan pyhton :

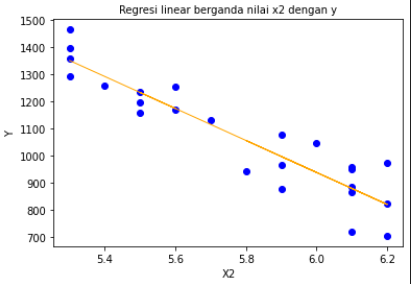


Terdapat perbedaan pada perhitungan manual dengan pyhton, karena pembuatan kode program menggunakan pustaka scikit learn. Penentuan model regresi dengan menambahkan konsep Ordinary Least Square (penentuan model terbaik berdasarkan nilai eror terkecil).

Visualisasi persebaran data, sebagai berikut :



Terlihat pada visualisasi, penyebaran data dan garis lurus naik keatas ke arah kanan, ini dikarenakan nilai dari koefisien regresi (b) bernilai positif dan garis lurus merupakan nilai dari Y prediksi (Y hasil)



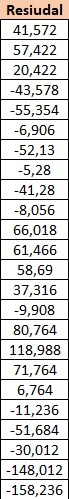
Terlihat pada visualisasi penyebaran data dan garis lurus naik keatas ke arah kiri ini dikarenakan nilai dari koefisien regresi (b) bernilai negatif dan garis lurus merupakan nilai dari dari Y prediksi (Y hasil)

1. **Akurasi prediksi regresi linear berganda dengan residual dan RSME**

* Perhitungan residual dapat dilakukan dengan menggunakan rumus,

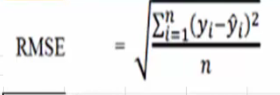
***Residual (e) = Y aktual – Y prediksi***

Sehingga akan mendapatkan hasil,



Pada nilai residual memiliki nilai positif jika data diatas garis lurus jika memiliki nilai negatif berarti data berada dibawah garis lurus.

* Perhitungan nilai RSME dapat dilakukan dengan menggunakan rumus,



Untuk mengetahui nilai keakurasin pada data jika nilai RMSE atau nilai penyimpangannya semakin kecil maka nilai keakurasian data semakin baik.

Dengan perhitungan tersebut maka nilai RMSE mendapatkan nilai **66,0052**

1. **Kesimpulan**

* Jadi dapat dilihat bahwa semakin besar nilai dari **x1** maka akan semakin tinggi nilai dari **Y**, sehingga nilai **x1** mempengaruhi nilai **Y** secara positif.
* Jadi dapat dilihat bahwa semakin kecil nilai dari **x2** maka akan semakin tinggi nilai dari **Y**, sehingga nilai **x2** mempengaruhi nilai **Y** secara negatif.