**Relatório Atividade 1 – Regressão Linear**

Nome: Wallesson Cavalcante da Silva

Curso: Ciência da Computação

Matrícula: 397670

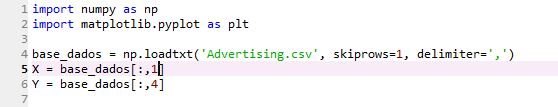
**Considerações iniciais:**

Tive um pouco de dificuldade no inicio por não conhecer a linguagem Python e das bibliotecas que poderia usar. Com isso justifico um pouco a deselegância do código. Para realizar a atividade recorri a pedir ajuda de alguns colegas que estavam mais habituados com a linguagem e com as bibliotecas usadas.

Decidi comentar também como elaborei o algoritmo.

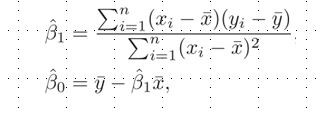
**Implementação da Regressão Linear Simples:**

Nesse trecho do código basicamente carreguei os dados do arquivo para a variável “base\_dados” e fiz a alteração dos valores que X iria receber para os 3 testes manualmente. (Como expliquei, foquei mais em resolver o problema e colher os resultados. Porém acredito que se tivesse pensado mais teria feito algo mais agradável).



A seguir, criei a classe “Regressao\_Linear\_Simples”. Nessa classe tratei a implementação dos cálculos para encontrar os valores dos betas (B0 e B1) que são a chave principal para gerar a função linear que definiria a regressão.

Com isso, minha função “fit” recebe os dados coletados em “X” e “Y” (atualizados da base\_dados), onde realizo uma serie de operações matemáticas simples para encontrar os valores estimados para os betas que se da pela função geral:



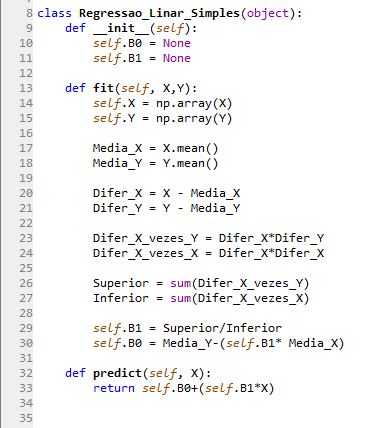
Primeiramente utilizei a função mean para armazenar em Media\_X o valor da media das amostras X e em Media\_Y o valor da media das observações Y.

Logo depois, tendo em vista resolver primeiramente a parte superior da formula supracitada, realizo a subtração das diferenças entre as amostras X com sua media e armazeno em Difer\_X, e das observações Y com sua media e armazeno em Difer\_Y.

Com esse resultado realizo uma multiplicação entre os valores armazenados em Difer\_X e Difer\_Y para obter o resultado do somatório superior da formula onde armazeno esse valor em Difer\_X\_vezes\_Y. E realizo uma multiplicação de Difer\_X por ele mesmo, para obter seu valor que deve ser ao quadrado seguindo a formula, armazenando ele em Difer\_X\_vezes\_X.

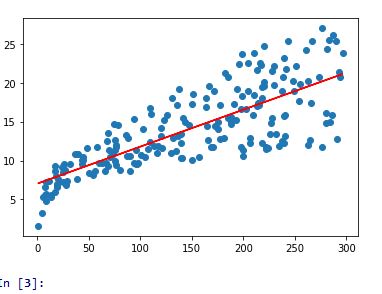
Para terminar, armazeno em “Superior” e “Inferior” o resultado obtido pela função sum aplicado na variável Difer\_X\_vezes\_Y e Difer\_X\_vezes\_X respectivamente.

Com isso, criei a função predict para montar equação linear. Onde B1 recebe Superior/Inferior e BO recebe Media\_Y – (B1 \* Media\_X) .

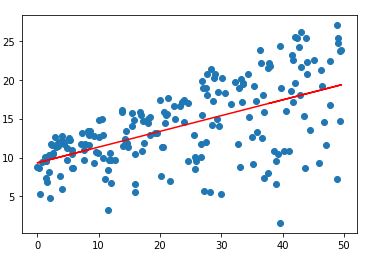


**2) Resultados**

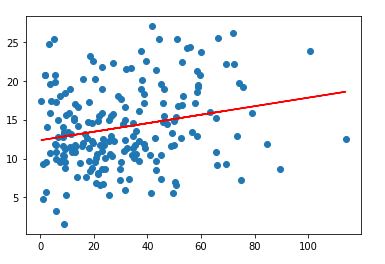
**A/B)**



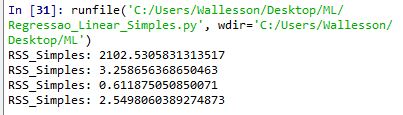
**Resultado obtido com as amostras de dados de TV**



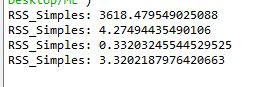
**Resultado obtido com as amostras de dados de Radio**



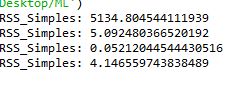
**Resultado obtido com as amostras de dados de Jornal**



**Funções aplicadas às amostras de TV**



**Funções aplicadas às amostras de Radio**



**Funções aplicadas às amostras de Jornal**

**OBS:** Na função RSS deu um valor bastante diferente da dos meus colegas. Acredito que errei na elaboração de alguma formula matemática, porém não encontrei o erro para corrigi-lo. O que estranhei mais foi o fato de as outras darem certo.

**C) Conclusões sobre as medidas de erros e os parâmetros**

Não sei se minha analise está correta, porém observando os dados obtidos e fazendo uma comparação dos gráficos também, percebi que quando é utilizada a amostra de TV o gráfico tende a ter um comportamento mais parecido com a função linear.