

# Índice

Índice.....	2
Objetivo:.....	3
Introdução.....	3
Gr á ficos.....	4
An á lise.....	5
Conclusão.....	6

## Objetivo:

Levantar projeções entre os usuários e sessões.

## Introdução

Este relatório tem como a finalidade de apresentar/ discutir os dados recebidos, basededados.xls, mas precisadamente iremos utilizar os dados de usuários e sessões. Para a fabricação desse relatório foi necessário: conhecimentos sobre estatística programação em python, e a utilização frameworks que facilitam a produção de dados para análise como numpy, pandas e jupyter.

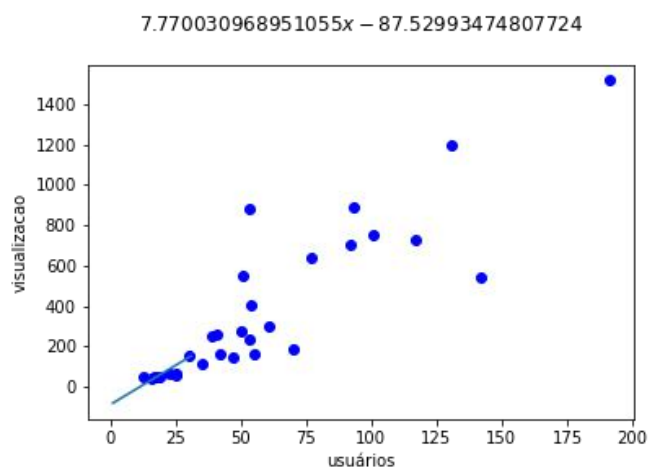
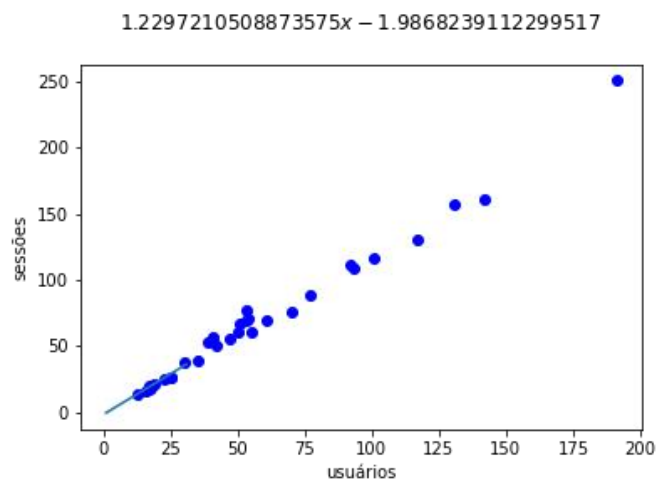
Nossa base de dados fornece 5 informações, sendo que observa-se que há dois tipos de usuários, ambos relacionam-se com o número de sessões, portanto o estudo estatístico obrigatoriamente tem que levar esse fato em consideração. Mesmo que na conclusão não abordemos tudo que será discutido. Uma vez que há necessidade de levantar essas duas curvas principais, no desenvolvimento do trabalho resolvemos levantar todas as curvas condizentes com o assunto principal. Esses levantamentos geraram métodos que vão além de levantar as projeções pedidas, tais métodos podem ser interessantes quando formos lidar com uma base maior.

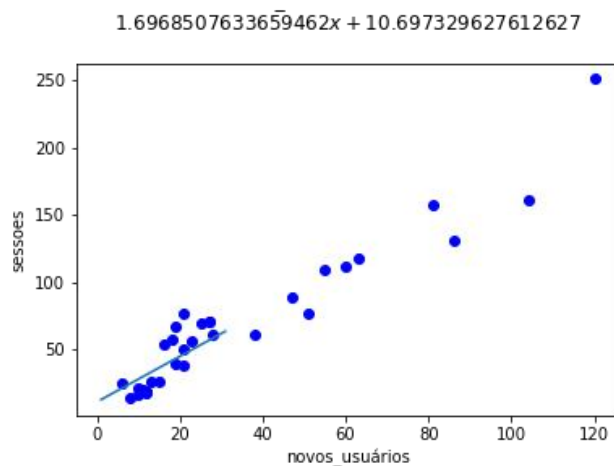
Nesse trabalho para realizar as projeções, utilizamos curvas levantadas pelo método dos mínimos quadrados. O método consiste em levantar uma curva que tenda a ser equidistante dos pontos obtidos. Além disso a distância entre os pontos e reta devem ser a menor possível. Esperamos que esta curva mantenha seu comportamento, mesmo que o número de medidas aumente com o passar do tempo. Futuramente podemos ter  $M$  amostras dos dados, quando isso ocorrer poderemos ter modelos com mais precisão, podendo quantificar melhor as incertezas associadas ao gráfico e quem saiba ter incertezas associadas menores. Com isso seremos capazes de avaliar quais os fatores significativos no alvo das nossas medições. Mas tais capacidade irão surgir a medida que nossa base de dados seja grande o suficiente.

Para medir quão próximo nossos dados estão da curva levantada, usamos o fator de correlação. Um fator que varia 0 a 1. Sendo quando mais preciso nosso modelo mais próximo de 1 esse fator estará, em contrapartida quanto mais próximo de 0 mais impreciso (distante dos dados) nosso modelo estará.

Vale apenas ressaltar que: no ponto de vista da estatística não faz sentido apresentar valores absolutos. Tal fato justifica o levantamento das curvas de outros dados dentro do contexto.

## Gráficos





## Análise

Começamos a análise pelas de usuários e sessões, alvo do estudo e relatório.

Ao levantar o gráfico de usuários e sessões, através do método dos mínimos quadrados. Verificamos que para uma curva linear, aqui subentende-se que curva linear como equação 1 grau, obteve o valor de correlação de 99,19%. Curva esta apresenta o maior fator de correlação medido. Com esse nível de precisão podemos afirmar que a equação da nossa projeção usuário/ sessão é:  $y = 1,2297x - 1,9868$ .

Como a tabela havia campo novos usuários e sessões, vale mencionar a curva levantada em nossa análise. Portanto verificamos que essa curva possui o fator de correlação de 65,96%, em relação a uma equação linear. Tendo um erro percentual em relação ao modelo do nosso objetivo de 50,38%.

Por fato de termos um número N, muito pequeno de medições, poucas coisas podemos afirmar em relação a grandes amostras.

Mas vemos dentro dessa amostra que usuários vs sessões são mais significativos. Não sabemos qual o critério para a definição de usuários/novos usuários, mas podemos afirmar que os usuários possui uma parte mais significativa no número de sessões do que os novos usuários.

## Conclusão

Obtivemos como esperado uma curva de projeção em relação às sessões, Tendo a seguinte equação de projeção

$y = 1,2297x - 1,9868$  com fator de correlação 99,19%.