

Índice

[Índice 2](#_Toc1812428286)

[Objetivo: 3](#_Toc129753178)

[Introdução 3](#_Toc544199158)

[Gráficos 4](#_Toc1544921719)

[Análise 5](#_Toc1063461336)

[Conclusão 6](#_Toc1989680893)

## Objetivo:

Levantar projeções entre os usuários e sessões.

## Introdução

Esse relátorio tem como a finalidade de apresentar/ discutir os dados recebidos, basededados.xls, mas precisasadamente iremos utilizar os dados de usuários e sessões. Para a fabricação desse relátorios foi nescessário: conhecimentos sobre estátistica

programação em python, e a utilização fremeworks que facilitam a produção de dados para análise como numpy, pandas e jypter.

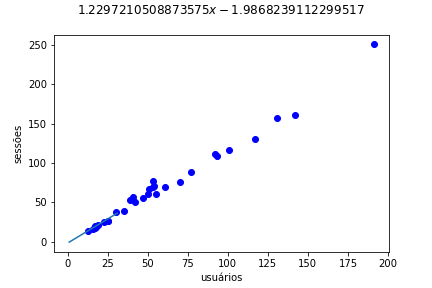
Nossa base de dados fornece 5 informações, sendo que obseva-se que há dois tipos de usuários, ambos relacionam -se com o número de sessões, portanto o estudo estatítico obrigatóriamente tem que levar esse fato em consideração. Mesmo que na conclusão não abordemos tudo que será discutido. Uma vez que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

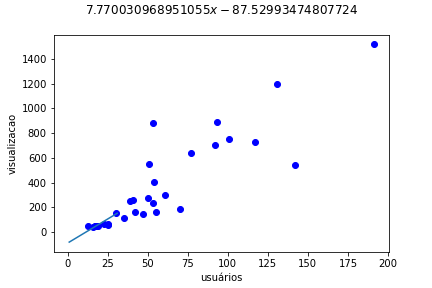
Nesse trabalho para realizar as projeções, utilizamos curvas levantadas pelo metódo dos mínimos quadrados. O metódo consiste em levantar uma curva que tenda a ser equidistante dos pontos obtídos. Além disso a distáncia entre os pontos e reta devem ser a menor possível. Esperamos que esta curva mantenha 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 precissão, 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 sej grande o suficiente.

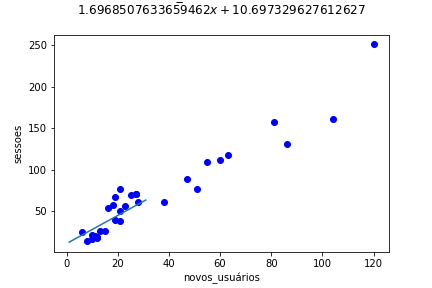
Para medir quão próximo nossos dados estão da curva levantada, usamos o fator de correlação. Um fator que varia 0 a 1. Sendo quando mais preciso nosso modelo mais próximo de 1 esse fator estará, em contra partida quanto mais próximo de 0 mais imprecisso(distante dos dados) nosso modelo estará.

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

## Gráficos







## Análise

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

Ao levantar o grafíco de usuários e sessões, atráves do metódo minímos quadrados. Verifica - se que para uma curva linear, aqui sub entede-se que curva linear como equação 1 grau, obteve o valor de correlação de 99,19%. Curva está 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 verifica -se que essa curva posssui o fator de correlação de 65,96%, em relação a uma equação linear. Tendo um erro percentual em reação 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 afrimar que os usuários possui uma parte mais sgnificativa no número de sessões do que os novos usuários.

Conclusão

Obtivermos como esperado uma curva de projeção em relação á sessões, Tendo a seguinte equação de projeção

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