Ciência dos Dados

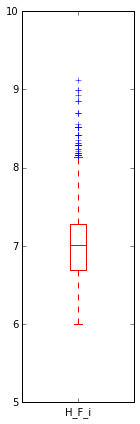
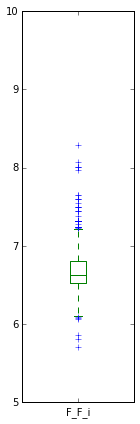
PROJETO 1 – DESIGUALDADE SALARIAL ENTRE GÊNEROS

# Introdução

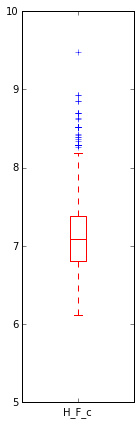
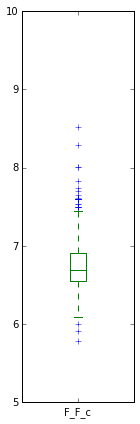
Para o projeto 1 de ciência dos dados, utilizamos como base de dados o arquivo do IBGE (Instituto Brasileiro de Geografia e Estatística) 2013, Para realizarmos a análise proposta, utilizamos uma parcela dos dados fornecidos pela base de dados. O objetivo do projeto é esclarecer se existe uma possível discrepância salarial entre homens e mulheres. Para estabelecer um pouco de homogeneidade, consideramos apenas os indivíduos que trabalhavam em apenas um emprego, com carteira assinada, empregado e “pessoa de referência” da família. (Além de fazer parte da atividade).

# O salário e a instrução

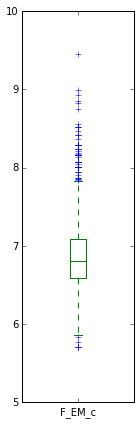
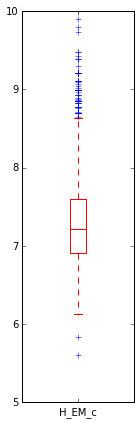
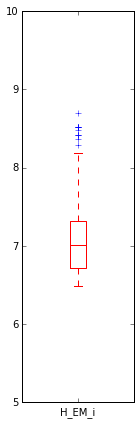
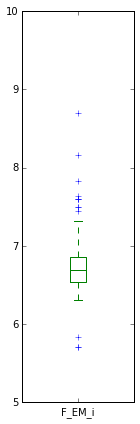
Dados analisados pela Fundação Getúlio Vargas (FGV) indicam que em média o salário de um indivíduo aumenta 15% a cada ano de estudo.

Essa comparação tem como finalidade analisar indivíduos de mesma formação, dessa forma criando uma associação entre a renda e a instrução. Em teoria pessoas com o mesmo nível de instrução deveriam receber o mesmo, proporcionalmente.

Ao Analisar os gráficos que relacionam os dados de salário e com o nível de instrução, o crecimento está mascarado pela escala logarítmica. Contudo, um fator que fica evidente é a discrepâcia entre os valores salarias entre gênero. Em verde nos gráficos está representado o conjunto feminino, em vermelho, masculino.

A partir da leitura destes box-plots os valores de mediana nos gráficos masculinos são, para todos os níveis de instrução, maiores do que as medianas nos femininos. Isso significa que a 50% da amostra masculina, que é faz parte dos dados, está localizada acima dos 50% femininos.

Muitas das medianas estão ainda acima dos terceiros quartis dos dados femininos. O que indica uma discrepância ainda maior.

Após uma análise geral dos cruzamentos gerados a partir dessas variáveis, concluímos que em média, homens ganham mais que mulheres, que possuem mesmo nível de escolaridade.

***Legenda:***

**F**= Mulheres

**H**= Homens

**Sem**= Sem instrução

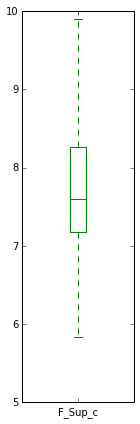
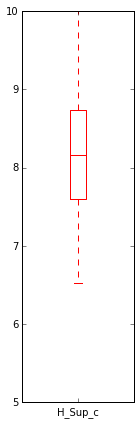
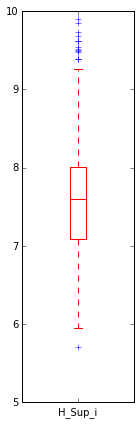
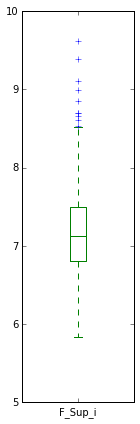
**i** = incompleto

**c**= completo

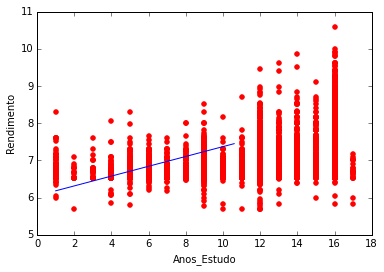
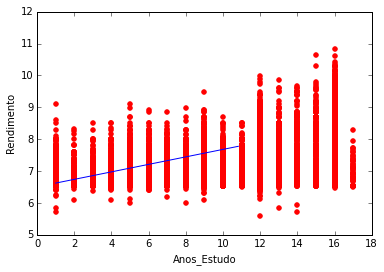
**F**= Ensino Fundamental

**EM**= Ensino Médio

**Sup**= Ensino Superior

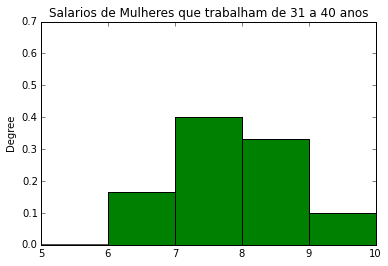
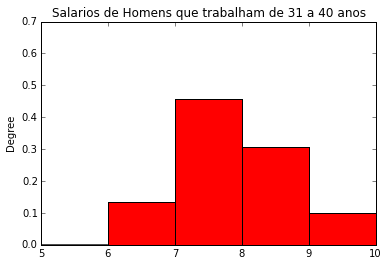
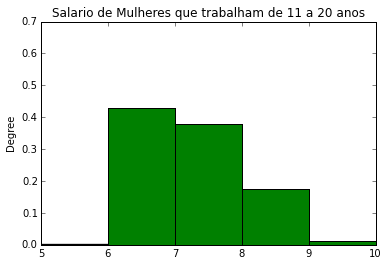
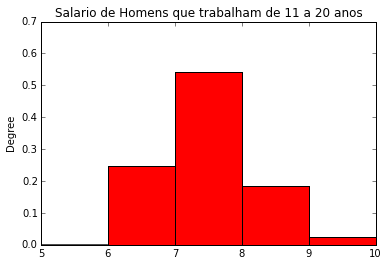
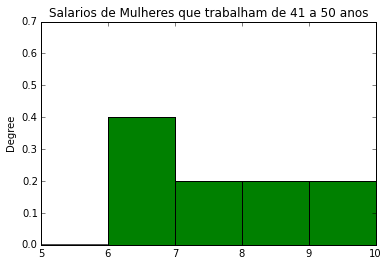
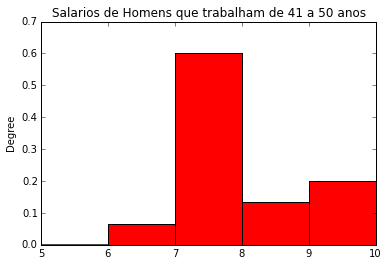
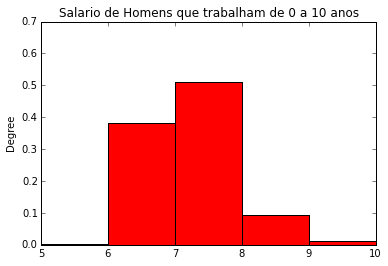
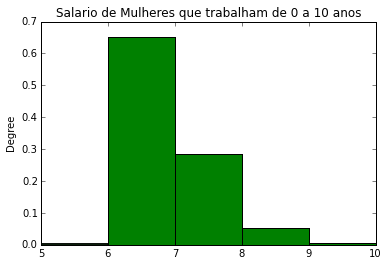


# O salário e anos de estudo

Os anos de estudo demonstram que a renda masculina aumenta com uma derivada menor, ou seja, a taxa de crescimento linear dos salários masculinos é ligeiramente menor do que o feminino. 0,072 para homens e 0,079 para mulheres. Contudo, assim como nos dados de níveis de ensino, o valor inicial da equação é maior entre os homens, o que indica que existe um crescimento maior entre as mulheres, porém, o ponto inicial é menor.

Sendo o masculino igual a 6,5 e o feminino igual a 6,1, lembrando que estamos utilizando uma escala logarítmica, o que resulta em uma discrepância maior

# O salário e os anos de trabalho

A Pesquisa realizada pela FGV também se aplica a esses dados. Da mesma forma que homens e mulher que tem o mesmo tempo de escolaridade deveriam possuir o mesmo salário, uma vez que o gênero não deveria discriminar renda, o salario de homens e mulheres que trabalham há mesma quantidade de anos em uma empresa deveria ser semelhante.Podemos concluir a partir da interpretação dos gráficos produzidos, que o padrão da primeira análise se repete, não na mesma medida. Contudo, homens continuam em média possuindo salários maiores do que os das mulheres.

# Conclusão

Ao analisarmos o contexto além dos dados evidenciados pelo banco de dados podemos concluir que existe uma discrepância salarial entre homens e mulheres, a qual homens recebem salários maiores que os das mulheres, mesmo sendo ilegal, e possuindo diversos fatores associados à cultura e preconceitos, é uma infeliz realidade em nosso país. Após interpretar dados reais, se torna evidente a necessidades de tomar medidas para que esse processo seja cada vez menos presente, e cada vez menos real.