

PROJETO 1

(ENCONTRO #06)

DISPARIDADES SALARIAIS ENTRE HOMENS E MULHERES

Tratar indivíduos de forma diferente por pertencerem a um ou outro grupo que foge a determinados padrões como cor da pele, opção sexual, religião, gênero, origem social ou qualquer outra marca que se impõe aos indivíduos, pode ser considerado discriminação.

Em particular, as disparidades salariais entre homens e mulheres refletem a discriminação e as desigualdades no mercado de trabalho que, na prática, afetam, sobretudo, as mulheres¹.

Esse assunto ainda é largamente estudado e discutido em diversos países. No Brasil, a Consolidação das Leis do Trabalho (CLT) já proíbe a diferença salarial entre homens e mulheres que exercem o mesmo tipo de atividade, no entanto, muitas empresas ainda resistem em cumprir a exigência².



Problema:

Neste projeto, o objetivo é analisar se é possível encontrar indícios de discriminação salarial entre homens e mulheres.

Para estabelecer um pouco de homogeneidade, considere apenas os indivíduos que trabalhavam em apenas um emprego, com carteira assinada, empregado e “pessoa de referência” da família.

Base de dados:

A base de dados <pes2013P1> é um filtro obtido da PNAD 2013 que elimina muitas variáveis, entretanto preserva todos 362.555 indivíduos. As variáveis descritas na base de dados são:

UF: estado brasileiro ao qual pertence o indivíduo;

V0302: gênero;

V8005: em anos;

¹ http://ec.europa.eu/justice/gender-equality/gender-pay-gap/index_pt.htm

² <http://www12.senado.leg.br/noticias/materias/2015/03/13/empresas-podem-sofrer-multa-por-discriminacao-salarial-contramulheres>

V0402: reflete a condição da pessoa na família. Entre as várias categorias descritas no dicionário da PNAD, a categoria “pessoa de referência” significa a pessoa responsável pela maior fonte de recursos da residência. Essa categoria era antes considerada pelo IBGE como “chefe de família”;

V0404: descreve cor ou raça do indivíduo;

V9001: indica se indivíduo trabalhava na semana da pesquisa;

V9005: número de trabalhos que o indivíduo tinha na semana da pesquisa;

V9029: posição na ocupação no trabalho principal;

V9032: descreve o setor do emprego principal: privado ou público;

V9042: indica se indivíduo tinha carteira assinada no trabalho principal;

V9611: número de anos no trabalho principal;

V4803: anos de estudo;

V4707: horas habitualmente trabalhadas por semana;

V4745: nível de instrução;

V4718: rendimento mensal no trabalho principal. Para diminuir a grande dispersão dessa variável, é comum analisá-la também aplicando o logaritmo natural (ou neperiano). Pesquise sobre os aspectos positivos dessa transformação na variável renda ou salário de um indivíduo. Trabalhadores

Regras:

1. Fazer em DUPLA.
2. Elaborar um relatório que contenha as análises que contribuíram para as conclusões.
3. O relatório deve ser feito em IPython Notebook (com células de texto formatadas em Markdown), Word ou outro editor de texto e deve conter, pelo menos, as seguintes seções:
 - a. Introdução,
 - b. Explicação das análises e resultados,
 - c. Conclusão.

Caso o grupo opte por entregar em IPython Notebook, a estrutura do documento deve facilitar a clareza e a compreensão dos argumentos e da linha de raciocínio. Não deve haver excesso de impressões de variáveis e dumps de tabelas. Neste caso aconselhamos fazer a análise e depois limpar seu IPython Notebook e melhorar o texto.

4. Deve ser entregue impresso.
5. O arquivo com os comandos de programação no Python (pode ser em .py ou ipynb) deve ser entregue via Blackboard. Será disponibilizado um link para enviar esse arquivo.

Cronograma:

Turma B e C:

DATA	EM SALA
31/08	Entrega do projeto Escolha das duplas Entendendo problema e base de dados para traçar possíveis análises
01/09 e 02/09	Trabalhar fora da aula– explorando a base de dados
03/09	Explorando base de dados
04/09 a 10/09	Trabalhar fora da aula – explorando a base de dados e iniciar e finalizar relatório
11/09	Entrega impressa do relatório impreterivelmente até as 14:00 Enviar arquivo com programação via Blackboard – link irá expirar às 9:30 da manhã.