

## PROJETO 1

ANÁLISE DESCRITIVA

### PERFIL DOS DOMICÍLIOS BRASILEIROS AO LONGO DE ALGUNS ANOS

#### PNAD:

<http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2015/default.shtm>

A Pesquisa Nacional por Amostra de Domicílios - PNAD investiga anualmente, de forma permanente, características gerais da população, de educação, trabalho, rendimento e habitação e outras, com periodicidade variável, de acordo com as necessidades de informação para o País, como as características sobre migração, fecundidade, nupcialidade, saúde, segurança alimentar, entre outros temas. O levantamento dessas estatísticas constitui, ao longo dos 48 anos de realização da pesquisa, um importante instrumento para formulação, validação e avaliação de políticas orientadas para o desenvolvimento socioeconômico e a melhoria das condições de vida no Brasil.

Em particular, a PNAD considera que uma **unidade domiciliar** é o domicílio particular ou uma unidade de habitação (quarto, apartamento etc.) em domicílio coletivo, podendo esse estar localizado em área urbana ou rural.

A utilização do dicionário é importante para localizar o nome e significado de cada variável considerada em cada ano de pesquisa.

Caso queira saber mais detalhes com relação às modificações de cada PNAD comparada ao ano anterior, utilize o link:

<https://dl.dropboxusercontent.com/u/65763596/projeto1%202017s1/PNAD%20DOM%202015/Notas%20Metodol%C3%B3gicas%20Pesquisa%20B%C3%A1sica%20e%20TIC%20PNAD%202015.docx>

## PERFIL DOS DOMICÍLIOS BRASILEIROS AO LONGO DE ALGUNS ANOS

### PROBLEMA:

Esse projeto tem como proposta levantar possíveis *semelhanças, diferença e/ou melhorias* nas diversas características socioeconômicas que podem descrever os domicílios brasileiros.

Nesse caso, você deverá escolher qual ou quais características dos domicílios pretende estudar em um particular ano e, de forma comparativa, entre pelo menos dois anos de levantamento da PNAD.

Obrigatoriamente, sua análise deverá considerar um misto de variáveis quantitativas e qualitativas e usar as diversas ferramentas estatísticas vistas até o momento na disciplina.

**Importante!** Se preocupe não só em fazer um filtro das variáveis que julgar serem relevantes para diferenciar os domicílios (colunas do *dataframe*), mas também avaliar se há a necessidade ou não de ser feito um filtro dos domicílios que deseja analisar (linhas do *dataframe*). Por exemplo, se deseja comparar domicílios apenas de áreas urbanas, deve antes excluir todas as linhas referentes aos domicílios localizados em área não urbana, e assim por diante. A importância desse “filtro de linhas” é avaliar, para o seu particular problema, se a grande heterogeneidade em muitas características domiciliares pode dificultar na análise dos resultados que serão obtidos.

Por fim, escolha as ferramentas estatísticas mais adequadas para explorar nas suas análises. Seu projeto deve conter um *mix* entre tabela de frequências, gráfico de pizza (ou setor), gráfico de barras, histograma, boxplot, gráfico de dispersão, medidas resumo e ajuste de retas. É válido ressaltar que a escolha dessas ferramentas deve ser norteadas pela adequação aos tipos de variáveis e ao questionamento do problema. Cada recurso estatístico utilizado no seu projeto deve deixar sempre claro o que se deseja investigar e qual a interpretação que se faz dos resultados obtidos.

**Links de algumas notícias do IBGE que podem motivar um perfil de domicílios que deseja investigar:**

<http://g1.globo.com/brasil/noticia/2011/04/numero-de-moradores-por-domicilio-cai-132-em-10-anos-diz-ibge.html>

<http://g1.globo.com/tecnologia/noticia/2015/04/mais-de-50-de-domicilios-brasileiros-tem- apenas-tv-de-tubo-diz-ibge.html>

<http://g1.globo.com/tecnologia/noticia/2016/11/n-de-casas-com-computador-cai-pela-1-vez-no-brasil-diz-ibge.html>

<http://g1.globo.com/economia-e-negocios/noticia/2010/09/ibge-domicilio-alugado-cresce-mais-que-casa-propria.html>

Entre várias outras matérias que pode encontrar no site <http://g1.globo.com/>

## **BASE DE DADOS (LINKS PARA DOWNLOAD):**

### **PNAD DOMICÍLIO 2015**

[https://www.dropbox.com/sh/8f5f0yp65ysbmxc/AABn0UStE\\_0Os5zdMFxEk6hWa?dl=0](https://www.dropbox.com/sh/8f5f0yp65ysbmxc/AABn0UStE_0Os5zdMFxEk6hWa?dl=0)

### **PNAD DOMICÍLIO 2014**

<https://www.dropbox.com/sh/iwbxrlqccq2xw3sp/AADnG7GGCWFGM4RP2wch7mCba?dl=0>

### **PNAD DOMICÍLIO 2013**

[https://www.dropbox.com/sh/nomfwnm8qjbq47a/AAB\\_R2OojMpPdFIYrC6b8v\\_ma?dl=0](https://www.dropbox.com/sh/nomfwnm8qjbq47a/AAB_R2OojMpPdFIYrC6b8v_ma?dl=0)

### **PNAD DOMICÍLIO 2011**

<https://www.dropbox.com/sh/mmwxlucipj14x0k/AADh6a67FLxfnk1sVqKJ-82-a?dl=0>

### **PNAD DOMICÍLIO 2009**

<https://www.dropbox.com/sh/zt8qnfne4uxlcbw/AAAf0TS3L3WWhFqfBBoh2I1va?dl=0>

## **REGRAS:**

1. O projeto 1 é estritamente INDIVIDUAL.
2. Elaborar um relatório que contenha as análises que contribuíram para as conclusões de acordo com seu projeto inicial. Use o **arquivo layout** disponibilizado na pasta Projeto1 do github:
3. O relatório deve ser feito em IPython Notebook (com células de texto formatadas em Markdown) e deve conter, pelo menos, as seguintes seções:
  - a. Introdução,
  - b. Análises e explicação dos resultados,
  - c. Conclusão.

**A estrutura do documento deve ser clara e de fácil compreensão da linha de raciocínio. Nesse caso, não deve haver excesso de impressões não discutidas de variáveis e *dataframe*. Aconselhamos fazer a análise geral e, após finalizada, salve com outro nome, limpe seu IPython Notebook apenas com os resultados relevantes e melhore seu texto.**

4. Seu projeto deve ser adicionado no seu github dentro de uma pasta chamada Projeto1.
5. O arquivo DEVE ser com extensão .ipynb

## **CRONOGRAMA:**

### **Turmas A, B e C**

<b>DATA</b>	<b>EM SALA – Aula Studio</b>
23/02	Entendendo problema e base de dados para traçar possíveis análises Explorando base de dados e construção dos <i>labels</i> das variáveis categóricas de acordo com dicionário da PNAD
02/03	Explorando base de dados e/ou Construção do relatório
<b>03/03</b>	Fazer git push em seu Github até 23:59.

## RUBRICS DE AVALIAÇÃO DO OBJETIVO DE APRENDIZADO

Objetivo de aprendizado	Insatisfatório (I)	Em desenvolvimento (D)	Essencial (C)	Proficiente (B)	Avançado (A)
<b>Elaborar análises exploratórias de dados (univariadas e multivariadas), utilizando ferramentas estatísticas e computacionais adequadas.</b>	Não consegue trabalhar com bases de dados de forma proficiente. Apresenta problemas com arquivos, formatos de arquivos ou não tem habilidades básicas de filtragem de dados.	<p>Consegue abrir e filtrar dados de forma básica, mas não consegue conectar os dados à prática de análise.</p> <p>O notebook é praticamente uma coleção de comandos do panda e gráficos, com pouca análise e interpretação.</p> <p>Tem dificuldade para lidar com valores inválidos e não consegue eliminá-los, corrompendo a análise.</p>	<p>Identifica em uma base de dados as variáveis pertinentes à resposta de uma pergunta e é capaz de filtrá-las, cruzá-las e processá-las para embasar sua análise.</p> <p>Utiliza pelo menos cinco variáveis mistas entre qualitativas e quantitativas e emprega os recursos adequados a aquelas variáveis de forma pertinente.</p>	<p>Apresenta os comportamentos da rubrica C para 7 variáveis mais diversificação excelente de recursos de acordo com o tipo de variável (qualitativa ou quantitativa).</p> <p>Analisa ocorrências conjuntas com recursos adequados ao tipo de variável (crosstab, gráficos de barras, pizza, histograma e boxplot e gráfico de dispersão e ajuste de retas) respeitando o tipo de variável, e interpreta</p>	<p>Executa os comportamentos da rubrica B de maneira excepcional e incorpora aspecto histórico, analisando a PNAD 2015 e mais uma.</p>

---

				corretamente o que as ocorrências conjuntas significam em função dos objetivos da pesquisa. Comunica os resultados com bastante clareza, deixando claro que pergunta está respondendo.	
--	--	--	--	--	--