

Exercício Final

Marcelo Prudente e Rafael Giacomini

12 de abril de 2018

Instruções

O exercício final tem o intuito de consolidar a implementação do uso de dados com o R.

- A apresentação do trabalho final poderá ser feita por meio de script ou de arquivo em pdf do R Markdown.
- No caso do script, lembre-se:
 - de utilizar o “#” para comentar o script separando as respostas
- No caso do R Markdown, utilize os campos de código para apresentar a sua solução.

O exercício pressupõe que o aluno já sabe importar arquivos para o R e quais bibliotecas serão necessárias para a análise.

Prazo: 22 de abril de 2018.

Parte 1

Exercício com a base SIOP

Dados obtidos no painel do orçamento.

- Adeque os nomes do banco siop
- Separe as colunas:
 - de Órgão Orçamentário em `orgao` e `orgao_desc`.
 - de Resultado Primário em `rp` e `rp_desc`
 - de Função em `funcao` e `funcao_desc` Dica: `?separate` ou [<http://tidyr.tidyverse.org/reference/separate.html>]
- Identifique:
 - os valor total e o percentual pagos por órgão;
 - os valores pagos por função;
- Calcule o gasto público identificando o gasto primário e financeiro. Transforme essas informações em gráficos.
 - qual percentual de gasto financeiro e gasto primário?
 - entre as despesas primárias obrigatórias (RP 1), quais as funções com maior gasto ao longo do tempo? Selecione as 5 maiores despesas (ranquear despesas).
 - qual percentual do orçamento é de despesas primária discricionária - compare com os outros tipos de Resultado Primário? O orçamento ficou mais rígido ao longo dos anos?
 - filtre os dados evidenciando as despesas primárias da função 28 - Encargos Especiais. Depois, encontre os maiores gastos por ação. Quais despesas se destacam? Você as considera como despesas genuínas da União?

Parte 2

PNAD Contínua

Instruções

- Para questões 1 a 5: baixe o arquivo da **PNAD Contínua Anual de 2016 da 5ª entrevista**, que contém informações sobre outras formas de trabalho.
- Para questões 6 a 9: baixe o arquivo da **PNAD Contínua Anual de 2017 da 1ª entrevista**, que contém informações sobre rendimento de outras fontes.
- Se não conseguir baixar *on-line*, baixe *off-line* de acordo com as informações contidas na página de **Douglas Braga** para a análise da PNADc com o PNADcIBGE.
- Observe com cuidado as variáveis derivadas.
- Utilize como guia a **publicação do IBGE**
- Diante disso, encontre:
 1. a produção de bens para consumo próprio por sexo e nível de instrução no Brasil.
 2. percentual de pessoas que realizam afazeres domésticos por regiões e segundo o sexo.
 3. média de horas dedicadas pelas pessoas de 14 anos ou mais aos afazeres domésticos por região e segundo o sexo. Também segregue por nível de instrução.
 4. percentual de pessoas que fazem trabalho voluntário segundo sexo.
 5. crie faixas de renda com cortes de R\$1.000,00 e verifique qual a relação entre as faixas de renda o trabalho voluntário.
 6. O percentual de pessoas que recebem BPC-LOAS. Qual a média de idade dessas pessoas?
 7. O percentual de pessoas que recebem Bolsa Família. Segregue as informações segundo o sexo.
 8. O rendimento domiciliar médio *per capita* dos que recebem BPC-LOAS segundo a região.
 9. O rendimento domiciliar médio *per capita* dos que recebem Bolsa Família segundo a região.
- Elabore gráficos para os **itens 8 e 9**.
- Encontre o índice de gini do rendimento mensal efetivo de todos os trabalhos das regiões brasileiras para os anos de 2012, 2015 e 2017. Transforme essas informações em um gráfico de linha.

Parte 3

- Vamos retomar o exercício do arquivo *exercicio_dplyr_avancado.pdf*.

RAIS

- Esse exercício tem por objetivo ampliar a capacidade de analisar dados. Envolve manipulação de datas, uso dos filtros (com operadores lógicos), mesclar dados (join), o uso do `if_else` e a confecção de alguns sumários estatísticos. Vamos fazer algumas inferências com a base RAIS.

Básico - exercício com datas

- Baixe a RAIS
- Peça a estrutura dos dados e verifique se os números (a exemplo do valor das remunerações) foram baixados corretamente.

- Transforme todas as datas de *character* em *date* (se necessário, use o paste)

Dados sobre qualificação e remuneração

- Na base RAIS, identifique quais as profissões que têm a maior qualificação.
- Segregue os dados por sexo e raça
- Qual a remuneração média por raça?
 - E por sexo?
 - E por raça e sexo?

Duplicados e número de empregos

- Identificar os casos duplicados no banco RAIS. Quantos são?
 - Em seguida, filtrar apenas os casos que contenham registros duplicados em um objeto chamado *rais2*. Com isso, você poderá identificar qual indivíduo teve mais de um emprego em um exercício.

```
# dica: utilizar
any(...)
```

Atenção: cada PIS é um caso que pode ter inúmeras observações. Portanto, pense bem na forma de filtrar os dados.

- Responda: qual o número máximo de empregos que um indivíduo teve em um ano
 - faça uma tabela com a contagem de indivíduos por número de empregos por ano.

Quem foi admitido ou demitido

- Com o objeto **rais2** (ou seja, indivíduos que tiveram mais de um emprego no período) encontre:
- os nomes dos empregos que os indivíduos tiveram no período (dica: mescle as duas bases - *rais2* e *cbo*);
- Depois, encontre apenas os casos em que os indivíduos têm dois ou mais vínculos e não foram demitidos no período.
 - Conte o número de casos individuais
 - Quais as profissões mais comuns?
- Encontre os casos em que os indivíduos têm dois ou mais vínculos, porém em algum momento foram demitidos no período. Depois, com o `if_else` indique se:
 - o indivíduo foi admitido no ano de 2015;
 - se foi admitido em 2015 e permaneceu empregado neste emprego em 2015
 - se o emprego com o vínculo ativo em 2015 é anterior a 2015.

Passos Necessários

1. Baixar os dados de forma eficiente.
2. Verificar a estrutura dos dados.
3. Pedir o sumário dos dados
4. Trabalhar com as datas
5. Cruzar dados encontrando as chaves em comum.
6. Fazer as estimativas apontadas
7. Exportar os dados em csv.

Parte 4

Exercício Seguro Demprego

O objetivo desse exercício é identificar se os possíveis beneficiários do seguro desemprego têm direito ao benefício requerido. Assim, devemos identificar se os requerentes estão enquadrados nos requisitos legais. Para o exercício, utilizaremos duas bases de dados do governo com identificações simuladas para o exercício: - RAIS - com identificação simulada; - Seguro Desemprego - com identificação simulada.

Segundo o sítio do **Ministério do Trabalho**, tem direito ao seguro desemprego o trabalhador que:

- Tiver sido dispensado sem justa causa;
- Estiver desempregado, quando do requerimento do benefício;
- Ter recebido salários de pessoa jurídica ou pessoa física equiparada à jurídica (inscrita no CEI) relativos a:
 - pelo menos 12 (doze) meses nos últimos 18 (dezoito) meses imediatamente anteriores à data de dispensa, quando da primeira solicitação;
 - pelo menos 9 (nove) meses nos últimos 12 (doze) meses imediatamente anteriores à data de dispensa, quando da segunda solicitação; e
 - cada um dos 6 (seis) meses imediatamente anteriores à data de dispensa, quando das demais solicitações;
- Não possuir renda própria para o seu sustento e de sua família;
- Não estiver recebendo benefício de prestação continuada da Previdência Social, exceto pensão por morte ou auxílio-acidente.

Para o exercício, vamos flexibilizar esses requisitos: *não possuem renda própria para o sustento da sua família e não estejam recebendo BPC*. Considere a quantidade de reincidências como o número de vezes que o indivíduo pediu o auxílio desemprego anteriormente.

Assim: - Cruze a base RAIS e Seguro Desemprego. - Encontre quantos indivíduos estão aptos a receber o benefício do seguro-desemprego. - Crie e exporte uma tabela contendo quantos indivíduos podem receber o seguro desemprego segregando os dados pelos três níveis de período de dispensa.

Passos Necessários

1. Baixar os dados de forma eficiente.
2. Verificar a estrutura dos dados.
3. Pedir o sumário dos dados
4. Trabalhar com as datas
5. Cruzar dados encontrando as chaves em comum.
6. Fazer as estimativas apontadas
7. Exportar os dados em csv.