

Exercício Fixação dplyr

Marcelo Prudente e Rafael Giacomini

27 de março de 2018

Contents

Parte 1 - Seguro Defeso e Bolsa Família	1
Exercício	1
Resultado esperado	2
Passos necessários	2
Parte 2 - Seguro Defeso e Bolsa Família: sua política obteve êxito?	3
Parte 3 - RAIS	3
Básico - exercício com datas	3
Dados sobre qualificação e remuneração	3
Duplicados e número de empregos	3
Quem foi admitido ou demitido	4
Passos Necessários	4
Parte 4	4
Exercício Seguro Demplo	4
Passos Necessários	5

Parte 1 - Seguro Defeso e Bolsa Família

Exercício

Na prática, a análise de dados envolve uma série encadeada de operações. Vamos utilizar o caso do seguro defeso para ilustrar a aplicação da manipulação de dados.

Caso Seguro Defeso - Bolsa Família

Imagine que você está em **janeiro de 2015** e, diante das restrições orçamentárias, lhe é encomendado um estudo para verificar a possibilidade de redução de despesas com programas sociais. Após algumas reuniões, verifica-se um consultor sugere a existência de superposição entre dois programas sociais: o Seguro-Defeso e o Bolsa Família. De fato, após a consulta ao Decreto que regulamenta o Programa Bolsa Família, você observa não haver óbice legal à acumulação dos benefícios.

Decreto 5.209, de setembro de 2004:

Art. 25. As famílias atendidas pelo Programa Bolsa Família permanecerão com os benefícios liberados mensalmente para pagamento, salvo na ocorrência das seguintes situações: I - comprovação de trabalho infantil na família, nos termos da legislação aplicável; II - descumprimento de condicionalidade que acarrete suspensão ou cancelamento dos benefícios concedidos, na forma do § 4o do art. 28;

III - omissão de informações ou prestação de informações falsas para o cadastramento que habilitem indevidamente o declarante e sua família ao recebimento dos benefícios financeiros do Programa Bolsa Família ou dos Programas Remanescentes;

IV - desligamento por ato voluntário do beneficiário ou por determinação judicial; V - alteração cadastral na família, cuja modificação implique a inelegibilidade ao Programa, observado o disposto no art. 21;

VI - ausência de saque dos benefícios financeiros por período superior ao estabelecido pelo Ministério do Desenvolvimento Social e Combate à Fome;

VII - esgotamento do prazo:

a) para ativação dos cartões magnéticos da conta contábil indicada no inciso III do § 12 do art. 2o da Lei no 10.836, de 2004;

b) para revisão de benefícios, na forma do art. 21.

VIII - desligamento em razão de posse do beneficiário do Programa Bolsa Família em cargo eletivo remunerado, de qualquer das três esferas de Governo.

Apesar da ausência dessa restrição no comando legal, não há nenhuma estimativa de quantos indivíduos acumulam recebimento dos dois programas. Tampouco, não há estimativa da possível economia aos cofres públicos que uma possível restrição na concessão desses benefícios pode gerar.

Resultado esperado

Utilizando os seus conhecimentos em manipulação de dados, descubra:

1. Identificar qual tipo de seguro-defeso tem o maior número de beneficiários.
2. Identifique qual o percentual de pescadores beneficiados em relação à população municipal.
3. Quantos indivíduos recebem bolsa família e seguro defeso no período.
4. Em seguida, segregue essa informação por Estado.
5. Estime a economia de recursos que a suspensão do recebimento do Bolsa Família durante o recebimento do seguro defeso poderá gerar.
6. Sumário dos dados do seguro defeso por região: número de beneficiários, média dos benefícios.
7. Exporte os dados para tabelas em formato csv.
 - A tabela 1 deve conter uma coluna com os dados do total de indivíduos beneficiados pelos dois programas e o total por estado. Outra coluna deve incluir a soma dos valores pagos total e por estado.
 - A tabela 2 deve conter os dados do total de economia orçamentária prevista, segregando os dados por região.
 - A tabela 3 deve conter os dados sobre a distribuição dos beneficiários do seguro defeso por tipo de benefício.
 - A tabela 4 deve conter os dados sobre o percentual de beneficiários do seguro defeso por município.

Passos necessários

A realização do exercício envolve os seguintes passos:

1. Encontrar e baixar os dados do Bolsa Família e do Seguro defeso no site do portal da transparência.
 - Qual o tamanho dos arquivos? Qual a melhor estratégia para lê-los?
 - Que tal ler algumas linhas para ter uma ideia das variáveis?
 - Qual o separador?
2. Salvar os dados na pasta de dados do seu computador.
3. Selecionar o diretório de trabalho
4. Baixar os dados de forma eficiente.
5. Verificar a estrutura dos dados.
6. Pedir o sumário dos dados
7. Adequar nomes - baixar função adequar_nomes
 - Lembre-se de como renomear colunas dos data.frames
8. Verificar linhas repetidas.
 - Há observações repetidas em cada uma das tabelas? Quais critérios podem identificar isso.

9. Trabalhar com as datas
 - Pesquise nos dados se a repetição de linhas ocorre por conta das datas.
 - Lembre-se de transformar as datas
10. Limpar os dados
11. Cruzar dados encontrando as chaves em comum.
12. Fazer as estimativas apontadas
13. Compilar as tabelas:
 - `bind_rows` ou `bind_cols`
14. Exportar os dados em csv.

Parte 2 - Seguro Defeso e Bolsa Família: sua política obteve êxito?

- Verifique se a sua indicação de mudança da política deu resultado em janeiro de 2016.
 - Há ainda beneficiários que recebem Bolsa Família e Seguro Defeso concomitantemente?

Parte 3 - 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.