

Propostas de Trabalho Final

Plano A

Contextualização

Em análise de dados qualitativos, pesquisadores costumam identificar tópicos recorrentes nos dados e criar classificações para estes tópicos. Uma técnica exploratória utilizada para esse processo é o *pile sorting*, utilizado tanto em *cultural domain analysis* (ver [este link](#)) quanto em análise de textos (Bernard & Ryan, 2010: 63-64). Essa técnica envolve pedir para algumas pessoas agruparem itens (podem ser trechos de texto, imagens, objetos, etc) em pilhas de coisas similares e criarem critérios de classificação para essas pilhas/categorias. Diversas técnicas de análise de dados podem ser utilizadas para analisar concordância entre as pessoas (Bernard & Ryan, 2010: 178-185).

A função

A função realizará uma *Non-Metric Multidimensional Scaling* (NMDS) com os resultados de uma atividade de *pile sorting* (realizada uma ou mais vezes por diferentes pessoas) e construirá um ou mais gráficos de apresentação dos resultados da NMDS. Os pontos nesses gráficos poderão ser diferenciados (cores e/ou símbolos diferentes) de acordo com uma classificação de um dos *pile sorts* (um gráfico por classificação).

Entrada de dados: Deve ser uma lista de *dataframes* (um *dataframe* para cada realização do *pile sorting*). Esses *dataframes* devem estar arranjados de modo que cada coluna corresponde a uma pilha/categoria e as linhas, aos itens de cada categoria. Como as categorias podem ter números diferentes de itens, podem aparecer células a serem ignoradas quando a função for correr.

Argumentos:

- Entrada do objeto de dados.
- Definição dos *character strings* a serem ignorados nos *dataframes* (ver acima).
- Definição da NMDS a ser realizada (isoNMS do pacote MASS ou monoNMS do pacote vegan).
- Classificação a ser utilizada para diferenciar os pontos nos gráficos. Deve ser um vetor com as posições do objeto de entrada de dados onde estão as classificações.
- Vetor de cores para pintar os pontos no gráfico.
- Vetor de número de símbolos para plotar os pontos no gráfico.
- Outros argumentos das funções isoNMS ou monoNMS.

Objeto de saída: o objeto de saída será uma lista com os resultados da NMDS: um *dataframe* com os *scores* finais de cada item e o valor do *stress* encontrado.

Gráficos gerados: *scatterplots* com o número de dimensões especificado para a NMDS (normalmente 2 ou 3). Cada ponto corresponde a um item, e estes poderão ser diferenciados (cores e/ou símbolos diferentes) de acordo com o que foi pedido nos argumentos da função.

Passos:

1. Construção de matriz de similaridade item por item para cada realização do *pile sorting*. O valor alocado para a célula i,j será 1 quando o item i e o item j foram colocados na mesma pilha/categoria (mesma coluna do *dataframe*) e 0 quando foram colocados em diferentes pilhas.
2. Construção de matriz de similaridade S item por item geral. O valor da célula i,j será a proporção (de 0 a 1) das realizações de *pile sorting* em que os itens i e j foram alocados para uma mesma pilha (mesma coluna do *dataframe*). Esse valor será 1 quando os itens i e j apareceram na mesma pilha em todas as realizações e 0 quando não foram colocados nenhuma vez na mesma pilha.
3. Construção da matriz de dissimilaridade D item por item geral. Os valores dessa matriz serão calculados por $D_{i,j}=1-S_{i,j}$.
4. Realização da NMDS a partir de D . As funções possíveis serão a isoMDS do pacote MASS ou a monoMDS do pacote vegan.
5. Construção do gráfico do resultado da NMDS, com os pontos diferenciados de acordo com os argumentos.

Exemplo de uso

A função poderá ser utilizada para averiguar concordância entre diferentes pessoas que realizaram uma atividade de *pile sorting* de 50 trechos de textos relacionados a causas para a lacuna pesquisa-prática na área ambiental retirados da literatura. Cada pessoa terá produzido uma classificação individual de tipos de causas existentes, e essas classificações poderão ser utilizadas na função para se criar um gráfico de duas dimensões, com um ponto por trecho, no qual trechos que foram considerados similares (colocados na mesma categoria) por um maior número de pessoas aparecerão mais perto no espaço das coordenadas (NMDS). A partir da visualização desse gráfico, pode-se perceber se há realmente algum tipo de causa que surge como consenso (grupo claro de trechos juntos). Pode-se repetir o gráfico para cada pessoa, pintando os pontos de acordo com a classificação da pessoa, sendo possível assim averiguar onde houve diferenças entre as pessoas.

Referências

Bernard, H.R. and G.W. Ryan 2010. *Analyzing Qualitative Data: Systematic Approaches*. Los Angeles: Sage. 451p.

Maravilha! Pode mandar bala no plano A, Di! Não tenho nada a modificar, retirar ou acrescentar (sim, no final você veio parar comigo!).

—[Luísa Novara](#)

Plano B

A função irá realizar cálculos sobre o uso da água em domicílios para auxiliar no planejamento doméstico em caso de rodízio (oficial ou não). Os principais argumentos da função serão as entradas

de dados para os seguintes aspectos: número de dias com água, número dias sem água, volume da caixa d'água e consumo médio diário de água.

Alternativamente, será possível também fazer o cálculo sem a entrada do consumo médio diário de água. Nesse caso, a própria função irá calcular o consumo médio diário de água baseado em uma série de parâmetros (retirados de [simulação da Sabesp](#)) que poderão ser inseridos como argumentos da função. São eles: número de residentes no domicílio, abertura da torneira do quintal (1/2 volta, 1 volta ou abertura total), tempo de abertura da torneira do quintal, número de vezes ao dia em que a torneira do quintal é usada, abertura da torneira da pia da cozinha, tempo de abertura da pia da cozinha, número de vezes ao dia em que a torneira da cozinha é aberta, número de vezes em que a descarga é acionada por dia, abertura da torneira do banheiro, tempo de abertura da torneira do banheiro, número de vezes ao dia em que a torneira do banheiro é usada, abertura do chuveiro, tempo de duração do banho, número de banhos por dia, abertura da torneira do tanque, tempo de abertura do tanque, número de vezes ao dia em que o tanque é aberto e número de vezes que a máquina de lavar é utilizada ao dia. O gasto médio de água que varia de acordo com esses parâmetros será retirado da [simulação da Sabesp](#). Como *default* desses parâmetros será colocado uma estimativa para uma pessoa. A função então realizará os seguintes cálculos:

1. Caso não tenha sido entrado o consumo médio diário, estimativa do consumo médio diário de água para o domicílio: soma das multiplicações entre os parâmetros listados e o consumo de água retirado do site da Sabesp para cada parâmetro, multiplicado pelo número de residentes.
2. Tempo em dias de duração da água da caixa d'água: divisão do tamanho da caixa d'água pela estimativa do consumo médio diário de água.
3. Caso a duração da água da caixa d'água seja maior do que o tempo sem água, a função retorna um aviso de que o domicílio não sofrerá com falta de água decorrente do rodízio.
4. Caso a duração da água da caixa d'água seja menor do que o tempo sem água, cálculo do volume de água necessário para completar esses dias sem água: multiplicação do consumo médio diário estimado pelo tempo sem água da caixa d'água. A função retorna um aviso de que será necessário comprar uma caixa d'água maior ou estocar água no período com água.

A função retornará, além dos avisos em tela, um objeto de saída de dados na forma de uma lista com os resultados de todos os cálculos realizados.

Gostei também, mas a primeira já está no <3 e vai ser super útil pro seu trabalho.

—[Luísa Novara](#)

From:
<http://ecor.ib.usp.br/> - ecoR

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2015:alunos:trabalho_final:diana.garcia:propostas_de_trabalho_final

Last update: 2020/08/12 06:04