Ana Paula Liboni

Doutoranda em Biologia Vegetal na Universidade Estadual de Campinas (UNICAMP), orientada pelo Prof. Ricardo Ribeiro Rodrigues.

Tese “REMANESCENTES FLORESTAIS DE PAISAGENS INTENSAMENTE FRAGMENTADAS: PAPEL NA CONSERVAÇÃO DA DIVERSIDADE FLORÍSTICA E FUNCIONAL E IMPLICAÇÕES PARA O MANEJO ADAPTATIVO”

exec

Exercícios

Reamostragem e simulação - exe8.1_8.2.r Regressão linear múltipla - exe7b.r Regressão linear simples e análise de covariância - exe107.2_ajuste_e_diag_regressao.r Tutorial testes de significância - tutorial6_testessign.r ANOVA - exe106.2.r

Trabalho Final

Comentários Vitor Rios

Ana Paula, suas propostas estão um pouco confusas. Tente dar uma clarificada, explicando o que é a entrada de dados, o que a função fará e o que ela te retorna, deixando a teoria e a motivação um pouco de lado. A proposta A me parece que é uma junção de várias análises diferentes, que talvez devam ser separadas em funções diferentes. a idéia não é automatizar todas as análises que vc precisa fazer, e sim obter um trecho de código reutilizável que possa ser aplicado em outros casos. A proposta B principalmente precisa de mais trabalho

vrios81@gmail.com

Acho que faltou realmente as descrições dos objetos de entrada e saída da função. Minha sugestão é que faça a função A para retornar avaliação de vários fragmentos ou amostras. Ou seja, que tenha uma entrada com uma coluna especificando a localidade (ou fragmento) e uma segunda definindo a unidade amostral (dentro do fragmento ou localidade). Caso seja apenas uma localidade e fragmento, essas colunas devem conter o mesmo valor (1 ou o que

seja!). Ajuste a proposta com as entradas e saídas da função para darmos uma olhada o quanto antes...

— *Alexandre Adalardo de Oliveira 2016/04/28 18:33*

Proposta A - reformulada 04-05-2016

INDICE DE INTEGRIDADE FLORESTAL BASEADO EM AVALIAÇÃO ECOLÓGICA RÁPIDA: UMA FERRAMENTA PARA GUIAR AÇÕES DE MANEJO EM FRAGMENTOS FLORESTAIS (FLORESTA ESTACIONAL SEMIDECIDUAL - BRASIL)

A Restauração Ecológica no Brasil tem dedicado pouca – ou quase nenhuma – atenção à restauração de fragmentos florestais degradados. Grande ênfase tem sido dada ao restabelecimento de comunidades vegetais nativas em áreas já convertidas para uso alternativo do solo. Em algumas paisagens intensamente perturbadas, as florestas secundárias constituem os únicos habitats florestais, garantindo a provisão de serviços ecossistêmicos como a atividade de polinização nas lavouras, controle de pragas, proteção dos cursos d'agua e armazenamento de carbono, além de serem fontes de propágulos e de indivíduos para fragmentos do entorno e para áreas em processo de restauração. Frente ao atual cenário de degradação ambiental, torna-se iminente a elaboração de protocolos para a avaliação e diagnóstico de remanescentes florestais, com a finalidade de guiar ações de manejo. É necessário que os fragmentos degradados sejam manejados adequadamente, visando à conservação da biodiversidade e à manutenção dos processos ecológicos que garantam a provisão de serviços ecossistêmicos, especialmente em paisagens extremamente alteradas pelas atividades humanas.

Partindo disso, a minha proposta é criar um índice de integridade florestal (IIF) com base em variáveis coletadas em campo por meio de avaliação ecológica rápida (AER) de fragmentos florestais. Restringi o índice para a fisionomia vegetal Floresta Estacional Semidecidual (FES), por ser bastante estudada, abranger grandes áreas no interior do país muito alteradas principalmente em função de práticas agrícolas, e também por já conter inúmeras iniciativas de restauração ecológica.

Com base na literatura, selecionei 10 variáveis relacionadas a estrutura e composição vegetal de fragmentos de FES, que seriam indicadoras da "integridade florestal" e que são coletadas de maneira mais rápida em comparação com os levantamentos fitossociológicos. As variáveis são quantitativas (discretas – contagens) ou qualitativas. Irei explicar a escolha de cada variável mais detalhadamente no help da função.

As variáveis são: árvores mortas em pé, emaranhados de trepadeiras, presença de epífitas vasculares, presença de figueiras (Ficus spp), presença de palmito (Euterpe edulis), eco-unidades, qualidade da regeneração, riqueza de espécies arbustivas e arbóreas nativas, riqueza de espécies exóticas (exceto gramíneas), densidade total de espécies arbustivas e arbóreas nativas (estrato adulto).

Todas as variáveis são fatores com 5 níveis. Para cada nível é atribuído um valor numérico em campo, sendo que o valor 1 se refere ao pior cenário para a variável, e o valor 5 se refere ao melhor cenário para a variável avaliada. Por exemplo: variável "emaranhados de trepadeiras (número de ocorrências):

Nível 5: 0 a 2 emaranhados (recebe valor 5)

Nível 4: 3 a 5 (recebe valor 4)

Nível 3: 6 a 8 (recebe valor 3)

Nível 2: 9 ou 10 (recebe valor 2)

Nível 1: > 10 (recebe valor 1)

Os dados de entrada para a função devem estar em um data frame que tenha as avaliações ecológicas rápidas (AER) nas linhas (observações), e as variáveis nas colunas. O usuário também deve especificar em uma coluna a localidade onde foi feita a AER. É nesta coluna que deve ser especificado qual fragmento será sua área referência (ref). O data frame também deve ter outra coluna especificando a unidade amostral que o usuário irá utilizar (várias observações dentro do mesmo fragmento ou comparando fragmentos.

O protocolo para a avaliação ecológica rápida vai indicar a área de cada transecto para ser feita cada avaliação (10 transectos de 100 m2 (4m x 25m) em cada fragmento florestal, totalizando 0,01 ha por fragmento). Portanto, se o usuário tiver 10 avaliações em um mesmo fragmento, deverá preencher a coluna “unidade amostral” com “1 a 10”. Se o usuário tem somente uma observação por fragmento, deverá repetir os valores de cada variável nas 10 avaliações, e na coluna “unidade amostral” deverá especificar “1”. Exemplo abaixo.

Comentários Vitor Rios

Uma coisa me preocupa: porque repetir a informação unidade 10 vezes? Se você não tem unidades amostrais suficientes, você não pode usar o índice. Repetir a informação altera o resultado e me soa *muito* errado. E se tiver 3 unidades por fragmento? descarta duas? quais? repete duas 3 vezes e uma 4 vezes? vrios81@gmail.com

Ana Paula

Olá, Vitor! Agradeço as considerações! Acho que não consegui explicar direito. A avaliação ecológica rápida pressupõe uma amostragem em uma determinada área do fragmento (vou propor 0.1 hectares divididas em 10 transectos de avaliação), porém isso pode se feito de diferentes maneiras pelo usuário, se ele mantiver a área total de amostragem (0.1 ha). Esse valor único que o usuário pode optar por usar já seria a média das AER que ele fez no fragmento (se dividiu a área em 5, 2 ou 1 amostragem fica a critério dele), por isso essa opção de repetir o valor. Mas pensando melhor, acho que isso realmente é desnecessário, seria mais coerente ele entrar com o valor das observações, independente de quantas tiver. Acho que fiquei presa a um detalhe do método da amostragem e me perdi na hora de passar isso pra função. Você tem razão.

Várias observações por fragmento: Ou somente uma observação por fragmento: O índice de integridade para cada observação é calculado com base na somatória por linha (somatória dos valores obtidos para todas as variáveis na aer). Em seguida é feita a média da coluna de somatórias para obtermos o valor do índice de integridade para o fragmento. Neste ponto a função retornará uma nova coluna no data frame com a somatória das variáveis por linha (AER.soma), uma coluna com o valor do IIF correspondente (IIF.soma), uma coluna com a média da AER para o fragmento

(AER.media) e uma coluna o valor do IIF médio (IIF.media), conforme exemplificado abaixo.

Comentários Vitor Rios

Essa parte não ficou clara. Isso é um objeto que a função retorna quando ela termina de executar ou é parte do cálculo? Me parece que você está tentando simular o cálculo do índice no excel usando o R. Isso fica óbvio na sua figura abaixo com os valores repetidos 10 vezes. Tente pensar em um jeito mais prático de fazer, desapegando da necessidade uma planilha. Mais ainda, se você está fazendo uma média, não é necessário restringir o valor a 10 unidades, ou repetir o valor dez vezes. Simplesmente o calculo com 10 amostras será mais confiável do que o com 1 amostra por representar emlhor a área. vrios81@gmail.com

Ana Paula

Isso foi ilustrativo mesmo, estava explicando como o cálculo do índice é feito. Mas gostaria que a função retornasse como produto um data frame novo com essas colunas. Essas colunas são informativas, dão os resultados parciais por amostra, e juntamente com a parte gráfica podem ser interpretadas pelo usuário para tomar as decisões quanto ao manejo. Vou colocar isso como sendo um produto da função, juntamente com os resumos por área e a parte gráfica. Acha pertinente?

O valor da AER pode variar de 10 a 50 (considerando 10 variáveis com 5 níveis cada). E para os intervalos de somatória dos valores de AER teremos as seguintes classes do índice (IIF):

AER.soma: 43 a 50 –> IIF 5 = excelente

AER.soma: 35 a 42 –> IIF 4 = bom

AER.soma: 27 a 34 –> IIF 3 = regular

AER.soma: 19 a 26 –> IIF 2 = ruim

AER.soma: 10 a 18 –> IIF 1 = muito ruim

Comentários Vitor Rios

Se vocÊ tem os vaores de 10 a 50 , porque transformar numa escala verbal e reduzir a informação a um décimo da original? vrios81@gmail.com

Ana Paula

Simplesmente para traduzir o valor observado para uma linguagem de fácil entendimento para comparação entre fragmentos…. Dizer que um fragmento tem índice de integridade “bom” é diferente de dizer que ele tem valor 35 na avaliação ecológica rápida. É somente para facilitar a linguagem mesmo, e fica muito mais acessível para tratar com as pessoas envolvidas e que não são muito familiarizadas com os números…

Exemplo de saída da função de acordo com opções do data frame: Quando há várias observações no fragmento:

Quando há somente um valor para o fragmento:

Nesta parte da função, se o usuário tiver várias observações por fragmento (unidades 1 a 10) será possível gerar um gráfico comparando o valor observado da variável em cada observação (por exemplo: variável trepadeiras » 4, 4, 3, 2, 2, 4, 4, 3, 4 e 4) com a média daquela variável para a área referência (representada por uma unidade de conservação da mesma fisionomia ou fragmento referência da região; exemplo: 4.2 neste caso). Teremos então um gráfico de dispersão com os valores obtidos para a variável na AER no fragmento analisado (4, 4, 3, 2, 2, 4, 4, 3, 4 e 4) com as barras de desvio padrão, e a linha com o valor médio daquela variável para a área referência (4.2). Como temos 10 variáveis, o painel vai mostrar 10 gráficos por fragmento. Além dos gráficos, a função apresenta um quadro de resumo POR ÁREA contendo valor máximo, mínimo, média, mediana, desvio padrão e variância PARA CADA VARIÁVEL.

Se o usuário tiver 10 fragmentos ou mais, também poderá obter a comparação da AER.media de cada fragmento com a AER.media para a área referência. Teremos então um gráfico de dispersão com os valores AER.media dos fragmentos (exemplo: 32, 41, 37, 49, 23, 29, 40, 28, 44 e 33) com as barras de desvio padrão, e a linha com o valor AER.media da referência (exemplo: 45). Além dos gráficos, a função apresenta um resumo PARA O CONJUNTO DE FRAGMENTOS (valor máximo da AER.media, mínimo, média, mediana, desvio padrão e variância).

Haverá, portanto, um argumento na função para identificar qual é o nível de observação do usuário (se existem várias observações em um mesmo fragmento (coluna unidade de 1 a 10) ou se existe somente uma observação por fragmento (coluna unidade somente com valores 1)).

O índice de integridade florestal proposto não irá substituir os levantamentos fitossociológicos, que são mais criteriosos e que demandam maior experiência botânica do avaliador. Mas penso em sua utilização como ferramenta exploratória, como um primeiro passo na tomada de decisão em relação às ações de manejo em fragmentos florestais degradados da Floresta Estacional Semidecidual.

COMPLICANDO UM POUCO, MINHA DÚVIDA (ainda não tive resposta/análise sobre este ponto): Até agora a minha função só executa se eu tiver um fragmento referência para comparar com os valores das variáveis ou de outros fragmentos. Com base na literatura e na proposta do índice, eu sei que os fragmentos mais “conservados” terão IIF máximo (“excelente”) com a AER.media variando entre 43 e 50. Para locais em que não há dados disponíveis de uma referência (por ausência de unidade de conservação na região por exemplo), pensei em fazer simulações ou reamostragem com reposição para a criação de um cenário “excelente” para comparar com os valores do fragmento que está sendo analisado. Sei que isso é factível, mas não sei se é pertinente ou se somente irá complicar a proposta de forma desnecessária. Vou precisar da opinião/ajuda dos monitores aqui.

Comentários Vitor Rios

O problema aqui é conceitual. A idéia de “fragmento de referência” é parte intríseca do índice, e portanto se você não tem, não pode usar . Existem programas que simulam áreas com cobertura vegetal (mas não outras variáveis) baseados em paisagens reais. Na prática, criar uma simulação adequada e realista das variáveis já seria em si uma proposta de tese

(http://www.sciencedirect.com/science/article/pii/S1574954114001617). Eu vejo duas saídas aqui: adotar um fragmento qualquer do seu conjunto como referência, e aí o seu índice se torna relativo a esse fragmento (menos útil), ou criar um “fragmento ideal”, com todas as variáveis com valor máximo, representando o melhor dos mundos, e todos os seus fragmentos reais seriam comparados com esse fragmento platônico. vrios81@gmail.com

Ana Paula

Acho que vou optar por adotar um fragmento da amostra como referência. Porque na maioria das situações é difícil ter uma área contínua e preservada da mesma fisionomia e na mesma região. Se foram feitas avaliações ecológicas na “referência” de verdade, melhor. Desta forma o índice pode ser aplicado de forma mais abrangente. O problema de comparar com um fragmento platônico que tem todas as variáveis máximas é: para um fragmento ser considerado “excelente” na sua integridade, a sua média das avaliações ecológicas pode variar entre 43 e 50, e isso não significa que o fragmento terá todas as variáveis com valor máximo. Isso soa muito irreal, e acho que dessa forma o índice perde o valor na prática. Agradeço mais uma vez pelas considerações. Vou enviar um e-mail com posteriores dúvidas. Obrigada!

Ana Paula - considerações finais

Olá, Vitor! Você verá na minha função final que eu alterei o meu último argumento. Mudei de ideia sobre a área referência, e as avaliações dos dados passaram a ser feitas em reação ao grupo de fragmentos amostrados. Como sabemos que o melhor cenário para as variáveis é 5, e para o índice de integridade é de 43 a 50, já dá para ter uma ideia da distância que as áreas estão de uma “situação ideal” somente olhando os gráficos e os data frames com dados resumo. Meu argumento 3 faz uma análise das variáveis amostradas, com o eixo padronizado (pra gerar uma proporção do valor máximo), então dá pra ver como as áreas variam… Acho que isso será mais relevante para avaliar as áreas e traz maior importância para o índice como ferramenta para guiar ações de manejo (uma vez que até mesmo áreas “referências” (como Unidades de Conservação, por exemplo) necessitam de manejo em algumas situações!). Obrigada por toda a ajuda.

Proposta B

QUANTIFICAÇÃO DE BIOMASSA E ESTIMATIVA DE ESTOQUE DE CARBONO EM ÁREAS EM PROCESSO DE RESTAURAÇÃO

O objetivo é criar uma função para quantificar a biomassa acima do solo (do fuste, método indireto) e estimar estoque de carbono a partir de um data frame que tem as variáveis na coluna (altura, dap, grupo (pioneiras/não-pioneiras)) e as observações (espécies) nas linhas. Serão utilizadas duas equações alométricas, que consideram ritmos diferentes de crescimento das espécies (pioneiras e não-pioneiras). A função retorna uma coluna no data frame com a quantificação da biomassa, e uma

com a estimativa de carbono. Também é mostrado um summary com a biomassa por grupo (pioneiras e não-pioneiras) e gerado um gráfico boxplot comparando os grupos para a área (um boxplot comparando biomassa e outro comparando estimativa de carbono). A função também irá gerar um boxplot comparando os valores observados para os grupos com valores de biomassa e estimativa de carbono obtidas em áreas de regeneração natural (que não tiveram plantio de mudas) de mesma idade (dados obtidos da literatura). Esses dados serão coletados para diferentes idades de regeneração, para servir como base para o monitoramento de áreas em processo de restauração.

```
##########################################################################################
```

HELP função ii.florestal

```
R Documentation

i.i.florestal
package: nenhum

Índice de Integridade Florestal e Análise Exploratória de Dados Biológicos
de Fragmentos Florestais

Description:

Função que calcula o Índice de Integridade Florestal (IIF) a partir de dados
coletados em fragmentos florestais por meio de Avaliações Ecológicas Rápidas
(AER) e faz uma análise exploratória comparativa dos locais e/ou amostras e
das variáveis coletadas em campo. Os produtos da função são data frames e
gráficos.

Usage:

ii.florestal <- function(data, leveldata=1, scatter.plot=FALSE)

Arguments:

x           Um data frame contendo os valores das variáveis nas colunas e as
observações (Avaliações Ecológicas Rápidas (AER)) nas linhas. As duas
primeiras colunas devem conter o nível da análise (local e amostras por
local) e as demais colunas devem ter os valores das 10 variáveis (somente
dados numéricos).

leveldata       Valor 1 ou 2. Indica se o data frame de entrada apresenta 1
ou mais níveis de observação. Quando leveldata==1 (valor do default) o data
frame apresenta mais de um nível de observação; quando leveldata==2 o data
frame apresenta somente um nível de observação. Este argumento define
produtos diferentes para a função.

scatter.plot        Argumento condicionado ao argumento leveldata==2, que
plota um scatter.plot das variáveis analisadas.
```

Details:
Com base na literatura foram selecionadas 10 variáveis relacionadas a estrutura e composição vegetal de fragmentos de Floresta Estacional Semidecidual, que seriam indicadoras da integridade florestal e que podem ser coletadas em avaliações ecológicas rápidas.
As variáveis são quantitativas (discretas – contagens) ou qualitativas, e são escritas na forma de um fator com 5 níveis (esses são os dados de entrada no data frame). O valor 1 se refere ao pior cenário para a variável, e o valor 5 se refere ao melhor cenário para a variável avaliada. Para cada amostra ou área a função traz a categoria do índice de integridade (muito ruim, ruim, regular, bom, excelente).
As variáveis não são fixas (podem se alterar contanto que mantenham o formato de fator com 5 níveis de mesma amplitude), o que permite flexibilidade na utilização da função. Para esta função foi fixado o número de 10 variáveis.
A função não é executada se houver valores faltantes (NAs) no data frame.
Value:

Se leveldata==1, a função retorna uma lista contendo dois data frames com dados resumo. O primeiro traz o valor da AER.soma e a categoria do ii.florestal para cada amostra; o segundo traz dados resumo para os locais. Também é gerado um gráfico com o valor médio da AER para cada local cuja abline indica a média de AER para o grupo. Este argumento também retorna uma janela gráfica que contém 10 gráficos (1 para cada variável) que permite fazer a análise do comportamento das variáveis nos locais.

Se leveldata==2, a função retorna um data frame com dados resumo para os locais e um gráfico com o valor médio da AER para cada área cuja abline indica a média de AER para o conjunto de locais.

Se scatter.plot==TRUE, a função plota um gráfico scatterplot juntamente com um data frame que contém dados resumo para os locais (idem descrito no argumento acima).

Warning:

Caso o leveldata seja um número diferente de 1 e 2, a função retorna uma mensagem de erro e não executa.

Caso seja indicado leveldata==1 e scatter.plot==TRUE, a função retorna uma mensagem de erro dizendo que só é possível obter o scatterplot com leveldata==2.
Author(s):

Ana Paula Liboni
apliboni@yahoo.com.br

References:

Variáveis selecionadas para testar a função:(1)árvores mortas em pé,

```
(2)emaranhados de trepadeiras, (3)presença de epífitas vasculares,
(4)presença de figueiras (Ficus spp), (5)presença de palmito (Euterpe
edulis), (6)eco-unidades, (7)qualidade da regeneração, (8)riqueza de
espécies arbustivas e arbóreas nativas, (9)riqueza de espécies exóticas
(exceto gramíneas), (10)densidade total de espécies arbustivas e arbóreas
nativas (estrato adulto).
Medeiros & Torezan (2013). Evaluating the ecological integrity of Atlantic
forest remnants by using rapid ecological assessment. Environmental
monitoring and assessment, 185(5), p. 4373-82.
DOI 10.1007/s10661-012-2875-7

Examples:

Simulação de dados para testar função
vetor.v1 = (rep(c(1:5),30))
vetor.v2 = (rep(c(1:5),30))
vetor.v3 = (rep(c(1:5),30))
vetor.v4 = (rep(c(1:5),30))
vetor.v5 = (rep(c(1:5),30))
vetor.v6 = (rep(c(1:5),30))
vetor.v7 = (rep(c(1:5),30))
vetor.v8 = (rep(c(1:5),30))
vetor.v9 = (rep(c(1:5),30))
vetor.v10 = (rep(c(1:5),30))

v1 = sample(vetor.v1, replace=T)
v2 = sample(vetor.v2, replace=T)
v3 = sample(vetor.v3, replace=T)
v4 = sample(vetor.v4, replace=T)
v5 = sample(vetor.v5, replace=T)
v6 = sample(vetor.v6, replace=T)
v7 = sample(vetor.v7, replace=T)
v8 = sample(vetor.v8, replace=T)
v9 = sample(vetor.v9, replace=T)
v10 = sample(vetor.v10, replace=T)

amostras = rep(c(1:10),15)
local = rep(c(1:15), each=10)

aer <- data.frame(local,amostras,v1,v2,v3,v4,v5,v6,v7,v8,v9,v10)
```

```
##############################################################
#################################################
```

Código da função ii.florestal

```
ii.florestal <- function(data, leveldata=1, scatter.plot=FALSE)
{
  if(!leveldata==1 & !leveldata==2) #a função não aceita valores de
```

```
leveldata diferentes de 1 e 2 e exibe mensagem no console
    {
      stop("Unexpected value for LEVELDATA. Only values 1 or 2 are
accepted")
  }
  else{}
  #data frame 1: retorna valores AER.soma e as classes do índice de
integridade (iif.soma) correspondentes
  aer <- data
  dataf.1 <- aer
  dataf.1$AER.soma <- apply(aer[,3:ncol(aer)], MARGIN=1 ,FUN=sum) #soma dos
valores das variáveis por observação, acrescenta uma nova coluna no data
frame (AER.soma)
  dataf.1$iif.soma <- rep(NA, times=length(dataf.1$amostras)) #insere uma
coluna vazia no data frame para ser preenchida com os valores de iif
  #inserindo classes do índice de integridade para cada amostra (coluna
iif.soma) conforme os intervalos abaixo
  dataf.1[dataf.1$AER.soma > 42 & dataf.1$AER.soma <= 50,14] <- "excelente"
  dataf.1[dataf.1$AER.soma > 34 & dataf.1$AER.soma <= 42,14] <- "bom"
  dataf.1[dataf.1$AER.soma > 26 & dataf.1$AER.soma <= 34,14] <- "regular"
  dataf.1[dataf.1$AER.soma > 18 & dataf.1$AER.soma <= 26,14] <- "ruim"
  dataf.1[dataf.1$AER.soma <= 18,14] <- "muito ruim"
  if(leveldata==1)
  {
    if(scatter.plot==TRUE) #a função não gera um scatter.plot quando
leveldata=1 e exibe mensagem no console
      {
        stop("Unexpected leveldata value. Only leveldata==2 returns
scatter.plot")
    }
    else{}
    #data frame 2: resumo das AER por local
    dataf.1$local <- as.factor(dataf.1$local) #tranformação da coluna local
para fator
    dataf.2 <- data.frame(levels(dataf.1$local)) #criando um data frame com
uma coluna com o mesmo número de níveis de "local"
    colnames(dataf.2) <- "local" #mudando o nome da coluna criada
    dataf.2$aer.min <- tapply(dataf.1$AER.soma, dataf.1$local, min) #coluna
criada com os valores mínimos de AER.soma por local
    dataf.2$aer.max <- tapply(dataf.1$AER.soma, dataf.1$local, max) #coluna
criada com os valores máximos de AER.soma por local
    dataf.2$aer.media <- round(tapply(dataf.1$AER.soma, dataf.1$local,
mean), digits=2) #média da AER.soma por local arredondada
    dataf.2$aer.mediana <- round(tapply(dataf.1$AER.soma, dataf.1$local,
median), digits=2) #mediana de AER.soma por local arredondada
    dataf.2$aer.desvio <- round(tapply(dataf.1$AER.soma, dataf.1$local, sd),
digits=2) #desvio padrão de AER.soma por local arredondado
    dataf.2$aer.var <- round(tapply(dataf.1$AER.soma, dataf.1$local, var),
digits=2) #variancia de AER.soma por local arredondada
    dataf.2$iif.medio <- rep(NA, times=length(levels(dataf.1$local)))
```

```
#criando coluna vazia no data frame com o mesmo comprimento de "local" para
iif.medio
    #iserindo classe de iif.medio para locais
    dataf.2[dataf.2$aer.media > 42 & dataf.2$aer.media <= 50,8] <-
"excelente"
    dataf.2[dataf.2$aer.media > 34 & dataf.2$aer.media <= 42,8] <- "bom"
    dataf.2[dataf.2$aer.media > 26 & dataf.2$aer.media <= 34,8] <- "regular"
    dataf.2[dataf.2$aer.media > 18 & dataf.2$aer.media <= 26,8] <- "ruim"
    dataf.2[dataf.2$aer.media <= 18,8] <- "muito ruim"
    plot(dataf.2$aer.media, ylim= c(0,60), pch= 16, main= "Avaliação
Comparativa da Integridade Florestal", ylab="AER média", xlab= "Local")
    media.geral.aer <- mean(dataf.2$aer.media)
    aer.media.por.local <- dataf.2$aer.media
    abline(h=media.geral.aer, col="red")
    #grafico 1: compara valores médios de AER com a média de AER para o
conjunto de locais (abline)
    #este gráfico também é plotado quando o usuário utiliza um data frame
com somente com 1 amostra por local (pessupõe-se que sejam os valores médios
das observações)
    #calculando médias para cada variável por local para plotar (quando
leveldata==1 e,portanto, existe mais de uma amostra por local)
    medias.v1 <- tapply(aer[,3], aer$local, mean)
    medias.v2 <- tapply(aer[,4], aer$local, mean)
    medias.v3 <- tapply(aer[,5], aer$local, mean)
    medias.v4 <- tapply(aer[,6], aer$local, mean)
    medias.v5 <- tapply(aer[,7], aer$local, mean)
    medias.v6 <- tapply(aer[,8], aer$local, mean)
    medias.v7 <- tapply(aer[,9], aer$local, mean)
    medias.v8 <- tapply(aer[,10], aer$local, mean)
    medias.v9 <- tapply(aer[,11], aer$local, mean)
    medias.v10 <- tapply(aer[,12], aer$local, mean)
    #criando data frame com as médias das variáveis por local
    medias.var.local <- round(data.frame(medias.v1, medias.v2, medias.v3,
medias.v4, medias.v5, medias.v6, medias.v7, medias.v8, medias.v9,
medias.v10), digits = 2)
    #inserindo coluna "local" no data frame
    medias.var.local$local <- seq(from=1, to=length(levels(dataf.1$local)))
    #calculando média geral para cada variável para plotar
    media.g.v1 <- apply(aer[3], MARGIN=2 ,FUN=mean)
    media.g.v2 <- apply(aer[4], MARGIN=2 ,FUN=mean)
    media.g.v3 <- apply(aer[5], MARGIN=2 ,FUN=mean)
    media.g.v4 <- apply(aer[6], MARGIN=2 ,FUN=mean)
    media.g.v5 <- apply(aer[7], MARGIN=2 ,FUN=mean)
    media.g.v6 <- apply(aer[8], MARGIN=2 ,FUN=mean)
    media.g.v7 <- apply(aer[9], MARGIN=2 ,FUN=mean)
    media.g.v8 <- apply(aer[10], MARGIN=2 ,FUN=mean)
    media.g.v9 <- apply(aer[11], MARGIN=2 ,FUN=mean)
    media.g.v10 <- apply(aer[12], MARGIN=2 ,FUN=mean)
    x11()
    par(mfrow=c(2,5)) #cria quadro de 10 gráficos, 2 linhas com 5 gráficos
cada
```

```
    plot(medias.v1, ylim=c(0,5), pch=16, ylab= "Médias variável 1", xlab=
"Local")
    abline(h=media.g.v1, col="red")
    plot(medias.v2, ylim=c(0,5), pch=16, ylab= "Médias variável 2", xlab=
"Local")
    abline(h=media.g.v2, col="red")
    plot(medias.v3, ylim=c(0,5), pch=16, ylab= "Médias variável 3", xlab=
"Local")
    abline(h=media.g.v3, col="red")
    plot(medias.v4, ylim=c(0,5), pch=16, ylab= "Médias variável 4", xlab=
"Local")
    abline(h=media.g.v4, col="red")
    plot(medias.v5, ylim=c(0,5), pch=16, ylab= "Médias variável 5", xlab=
"Local")
    abline(h=media.g.v5, col="red")
    plot(medias.v6, ylim=c(0,5), pch=16, ylab= "Médias variável 6", xlab=
"Local")
    abline(h=media.g.v6, col="red")
    plot(medias.v7, ylim=c(0,5), pch=16, ylab= "Médias variável 7", xlab=
"Local")
    abline(h=media.g.v7, col="red")
    plot(medias.v8, ylim=c(0,5), pch=16, ylab= "Médias variável 8", xlab=
"Local")
    abline(h=media.g.v8, col="red")
    plot(medias.v9, ylim=c(0,5), pch=16, ylab= "Médias variável 9", xlab=
"Local")
    abline(h=media.g.v9, col="red")
    plot(medias.v10, ylim=c(0,5), pch=16, ylab= "Médias variável 10", xlab=
"Local")
    abline(h=media.g.v10, col="red")
    #gráficos mostram as médias das variáveis por local para comparação com
a média da variável referente ao conjunto de locais (abline)
    lista1 <- list(dataf.1, dataf.2) #retorno quando leveldata==1
    return(lista1)
  }
  if(leveldata==2) #quando leveldata==2 o data frame de entrada só possui
uma amostra por local
  {
    #data frame 2 (igual no passo anterior)
    dataf.1$local <- as.factor(dataf.1$local)
    dataf.2 <- data.frame(levels(dataf.1$local))
    colnames(dataf.2) <- "local"
    dataf.2$aer.min <- tapply(dataf.1$AER.soma, dataf.1$local, min)
    dataf.2$aer.max <- tapply(dataf.1$AER.soma, dataf.1$local, max)
    dataf.2$aer.media <- round(tapply(dataf.1$AER.soma, dataf.1$local,
mean), digits=2)
    dataf.2$aer.mediana <- round(tapply(dataf.1$AER.soma, dataf.1$local,
median), digits=2)
    dataf.2$aer.desvio <- round(tapply(dataf.1$AER.soma, dataf.1$local, sd),
digits=2)
```

```
    dataf.2$aer.var <- round(tapply(dataf.1$AER.soma, dataf.1$local, var),
digits=2)
    dataf.2$iif.medio <- rep(NA, times=length(levels(dataf.1$local)))
    #iserindo classe de iif.medio para locais
    dataf.2[dataf.2$aer.media > 42 & dataf.2$aer.media <= 50,8] <-
"excelente"
    dataf.2[dataf.2$aer.media > 34 & dataf.2$aer.media <= 42,8] <- "bom"
    dataf.2[dataf.2$aer.media > 26 & dataf.2$aer.media <= 34,8] <- "regular"
    dataf.2[dataf.2$aer.media > 18 & dataf.2$aer.media <= 26,8] <- "ruim"
    dataf.2[dataf.2$aer.media <= 18,8] <- "muito ruim"
    if(scatter.plot==TRUE) #se levaldata==2 e scatter.plot=TRUE a função
retorna um scatter.plot
    {
      #médias para cada variável por local (se houver somente uma amostra,
este é repetido)
      medias.v1 <- tapply(aer[,3], aer$local, mean)
      medias.v2 <- tapply(aer[,4], aer$local, mean)
      medias.v3 <- tapply(aer[,5], aer$local, mean)
      medias.v4 <- tapply(aer[,6], aer$local, mean)
      medias.v5 <- tapply(aer[,7], aer$local, mean)
      medias.v6 <- tapply(aer[,8], aer$local, mean)
      medias.v7 <- tapply(aer[,9], aer$local, mean)
      medias.v8 <- tapply(aer[,10], aer$local, mean)
      medias.v9 <- tapply(aer[,11], aer$local, mean)
      medias.v10 <- tapply(aer[,12], aer$local, mean)
      #criando data frame para as médias das variáveis por local
      medias.var.local <- round(data.frame(medias.v1, medias.v2, medias.v3,
medias.v4, medias.v5, medias.v6, medias.v7, medias.v8, medias.v9,
medias.v10), digits = 2)
      #criando uma coluna para local
      medias.var.local$local <- seq(from=1,
to=length(levels(dataf.1$local)))
      #padronizando o eixo y (dividi os valores das médias das variáveis
pelo valor máximo possível = 5)
      medias.var.padron <- medias.var.local[,1:10]/5
      #criei vetor com as médias para as variáveis
      scatter.vetor <- c(medias.var.padron$medias.v1,
medias.var.padron$medias.v2, medias.var.padron$medias.v3,
medias.var.padron$medias.v4, medias.var.padron$medias.v5,
medias.var.padron$medias.v6, medias.var.padron$medias.v7,
medias.var.padron$medias.v8, medias.var.padron$medias.v9,
medias.var.padron$medias.v10)
      plot(scatter.vetor~rep(medias.var.local$local,10), pch=1, ylim=
c(0,1), main= "Avaliação das Variáveis", xlab="Local", ylab="Escala
padronizada AER")
      #scatter plot com valores médios das variáves por local. número de
variáveis é fixo na função (sempre 10).
    }
    else #quando o usuário opta por não ter o scatter.plot, a função retorna
o gráfico com valores de AER.média para os locais e mostra abline que é a
média geral do conjunto de locais
```

```
    {
      plot(dataf.2$aer.media, ylim= c(0,60), pch= 16, main= "Avaliação
Comparativa da Integridade Florestal", ylab="AER média", xlab= "Local")
      media.geral.aer <- mean(dataf.2$aer.media)
      aer.media.por.local <- dataf.2$aer.media
      abline(h=media.geral.aer, col="red")
    }
    return(dataf.2) #e também retorna o data frame 2 (quadro de resumo por
local)
  }
  else{}
}
```

```
##############################################################
##
```

arquivo teste para a função

aer_teste1.txt

arquivo teste2 para a função

aer_teste2.txt

From:
http://ecor.ib.usp.br/ - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2016:alunos:trabalho_final:apliboni:start

Last update: **2020/08/12 06:04**