



Universidade de Brasília - UnB
Instituto de Ciências Biológicas
Departamento de Zoologia

BIOLOGIA QUANTITATIVA Código ZOO-0054

Segundo Semestre de 2020 REMOTO

Local: A disciplina será ministrada em regime remoto segundo o calendário da UnB, usando a plataforma Microsoft Teams da UnB. Os exercícios práticos farão uso das plataformas de estatística e programação Rstudio Cloud, GitHub, Google Colabora, Google Cloud, Binder e outras.

Horários: Terças-Feiras de 14:00 às 15:40 (Teórico-Prática), Quintas-Feiras de 14:00 às 15:40 (Teórico-Prática) e Sábados de 08:00 às 10:00 (Prática de Campo somente quando autorizada pela UnB, ou Tutorial no modo remoto)

Ementa: Apresentação de fundamentos teóricos e de exercícios sobre métodos quantitativos usados em biologia, nas áreas de ciência de dados, biologia de campo, bioestatística com R, geoprocessamento e biologia da paisagem com QGIS, e técnicas de inteligência artificial com redes neurais, algoritmos genéticos e autômatos celulares.

Objetivo: Fornecer aos alunos de graduação um conjunto de ferramentas básicas para análise de dados em biologia e para desenvolvimento de modelos usando métodos computacionais modernos.

Professor: Roberto Cavalcanti (Departamento de Zoologia)

Monitoras e Monitores: Ana Terra Sousa Guedes, Marcela Ponce de Leon Soares, Vinícius Telheiro do Nascimento.

Programa: A disciplina é dirigida a estudantes de graduação em Biologia, Biotecnologia, e Ciências Ambientais a partir do terceiro semestre. O conteúdo neste semestre 2/2020 irá versar sobre abordagens e métodos quantitativos para análise de dados em Biologia, Zoologia e Ecologia, aplicáveis para análise de dados obtidos no campo e para metodologias usuais de interpretação e modelagem de encontradas na literatura atual. Iremos focar em quatro temáticas: 1) conceitos básicos de ciência de dados, combinando dados já existentes na literatura com dados originais coletados no campo; 2) uso dos softwares R, Rstudio, Git, jupyter e python para gestão, análise e ciência de dados, exemplificando com os métodos de análise de variância, regressão, modelos lineares e linearizados; 3) uso do QGIS e outros pacotes de software no geoprocessamento e tratamento de imagens para análises de paisagens, modelagem de mudanças nos ecossistemas e na biota na escala local e regional. 4) Aplicações de inteligência artificial na biologia, usando redes neurais, algoritmos genéticos e autômatos celulares.

Métodos:

O programa será apresentado através de ensino remoto por meio da plataforma Microsoft Teams no site da UnB. Os docentes e monitores irão apresentar e tutoriais nos horários de aula da disciplina via Microsoft Teams ou Zoom. Usaremos também os serviços de chat, email e quadro de avisos do Teams para resolver dúvidas e atender a consultas, no espaço virtual da sala de aula da disciplina. O conteúdo será ministrado por meio da realização de 15 módulos de ensino remoto contendo materiais para leitura, apresentações powerpoint, vídeos, tutoriais em R e python, e exercícios. Cada módulo será considerado concluído pelo estudante após a entrega do exercício do módulo e a realização com sucesso de um teste de conteúdo do mesmo com nota acima de 5. Não há limites para o número de tentativas para o teste final de cada módulo. Os estudantes poderão concluir até dois módulos por semana, sempre seguindo a sequência do curso. Para aprovação será necessário concluir ao menos 08 módulos. A matéria será encerrada segundo o calendário da UnB do 2/2020, e as menções divulgadas na data indicada no calendário da UnB 2/2020. Portanto muita atenção para o cronograma, de forma a não acumular módulos e ficar sem tempo para completá-los, pois são dois por semana no máximo! A menção serão atribuída conforme o número de módulos concluídos até a data de encerramento da disciplina.

É necessário que todos estudantes se inscrevam na plataforma Microsoft da UnB (<https://login.microsoftonline.com/?whr=unb.br>) na sala de aula Biologia Quantitativa para acompanhamento da disciplina. Os materiais específicos relacionados com os temas das aulas serão postados no site. A plataforma será também utilizada para realizar avaliações online e para distribuição e recolhimento de tarefas do curso.

Projeto Final: Os estudantes terão a opção de apresentar um projeto final, valendo até 10% do total de pontos atribuíveis na disciplina, correspondente ao módulo 15 e mais um a escolher. O projeto poderá ser feito individualmente ou em grupo de até 3 pessoas e poderá adotar um dos seguintes formatos: a) um tutorial no formato R notebook para um dos temas do curso; b) um relatório em formato de artigo realizando ajuste de modelo linear usando dados de campo coletados em projetos dos alunos participantes; c) uma função, script, ou pacote no R para resolver um problema prático em biologia.

Aulas Teórico/Práticas, Tutoriais via Microsoft Teams ou Zoom com docente e/ou monitores, e atividades de campo:

Iremos realizá-los no horário listado na lista de oferta 2/2020. Há a Os tutoriais cobrirão o tema dos módulos da semana e/ou dúvidas e temas levantados pelos alunos, ou faremos coletas de dados no campo no CO desde que haja autorização da UnB para este tipo de atividade presencial.

Recursos Online a serem utilizados

Iremos usar primariamente o Microsoft Teams, o Github, e o Rstudio Cloud ou o Google Colaboratory para executar programas e tarefas. É recomendado aos alunos instalar o Rstudio e o QGIS nos seus computadores pessoais. Todos os programas listados são de código livre e/ou têm planos de acesso gratuito para os alunos.

Frequência: Para aprovação os estudantes deverão completar no mínimo 8 módulos e assistir a pelo menos 8 tutoriais.

Módulos da Disciplina:

Módulo 1 : Introdução à disciplina. Tipos de dados, rstudio, github, jupyter, colaboratory, azure. Apresentação do professor, monitora, e alunos.

Módulo 2 : Bases de dados online e gestão de dados: o exemplo do SiBBR. Palestrantes a designar.

Módulo 3 : Testes de Hipóteses e Planejamento de Experimentos. Visualização e distribuição de dados, normal, estatísticas centrais e de dispersão.

Módulo 4: Populações e suas distribuições: normal, t. Comparação de Amostras.

Módulo 5: Conceito de modelos lineares, mínimos quadrados, anova, partição de variância. Uso de armadilhas fotográficas e gravadores.

Módulo 6: modelos lineares de variáveis contínuas, regressão.

Módulo 7: distribuição espacial, binomial, poisson, uniforme, aleatória

Módulo 8: amostragens por transectos e quadrat, distance sampling, rarefação

Módulo 9: modelos lineares e linearizados com dados normais, poisson, binomial, log, AIC

Módulo 10: Análise de agrupamento, cluster

Módulo 11: Distribuições Espaciais e Ecologia da Paisagem Introdução ao SIG, matrizes, objetos espaciais, QGIS

Módulo 12: Modelagem usando SIG

Módulo 13: Análise de componentes principais

Módulo 14: Inteligência Artificial - Redes Neurais e Autômatos Celulares

Módulo 15: Inteligência Artificial - Algoritmos Genéticos

Projeto Final:

Segundo o Calendário da UnB: Divulgação dos Resultados e Revisão das Notas

Segundo o Calendário da UnB : Lançamento das Menções

Avaliação: Os pontos da disciplina serão distribuídos em 14 módulos de conteúdo (sendo um bônus) e um de projeto final. As menções serão dadas da seguinte forma: de 8 a 11 módulos concluídos menção MM; de 12 a 14 módulos concluídos menção MS; 15

módulos concluídos menção SS. O aluno deve cumprir no mínimo 8 módulos e registrar presença em 8 tutoriais para ter aprovação.

Observações gerais: Os *testes, questionários e quizzes* de cada módulo serão apresentados individualmente, exceto quando for explicitamente informado que a atividade pode ser apresentada por um grupo. O *projeto final* poderá ser feito individualmente ou em grupo de até 3 pessoas e poderá adotar um dos seguintes formatos: a) um tutorial no formato R notebook para um dos temas do curso; b) um relatório em formato de artigo realizando ajuste de modelo linear usando dados de campo coletados no CO ou em outros projetos dos alunos participantes; c) uma função, script, ou pacote no R para resolver um problema prático em biologia. d) um modelo de “data paper” publicando dados coletados pelos estudantes durante o semestre, usando protocolo do SIBBr usando padronização no Darwin Core e publicação no IPT/GBIF.

Bibliografia:

Andrade, D. F. e P. J. Ogliari. 2013. Estatística para as ciências agrárias e biológicas - com noções de experimentação. 3a Ed revista e ampliada. Editora UFSC, Florianópolis, 478pp.

Bolker, B. M. 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ, USA, 396 pp.

Bruce, A, e P. Bruce. 2019. Estatística Prática para Cientistas de Dados: 50 conceitos essenciais. Editora Alta Books, Rio de Janeiro, RJ, 392 pp/

Borcard, D, F. Gillet, and P. Legendre. 2011. Numerical Ecology with R. Springer Science+Business Media, New York Dordrecht London Heidelberg, xi + 306 pp.

Gotelli, N. J. e A. M. Ellison. 2011. Princípios de Estatística em Ecologia. Artmed, Porto Alegre, RS, xiii + 527 pp.

Novo, Evlyn M. de Moraes. 2010. Sensoriamento Remoto - Princípios e Aplicações. Editora Edgard Blucher, São Paulo, SP, 388 pp.

Silva, I. N. da. 2016. Redes Neurais Artificiais para Engenharia e Ciências Aplicadas. Artliber Editora, São Paulo, SP, 431 pp.

Vist, H. L, A. M. Gonçalves, R. M. dos Santos. 2013. Apostila do Curso: Quantum GIS Básico - Módulo 1. CVIST, Porto Alegre, RS, 39 pp. Disponível na internet em <http://www.clickgeo.com.br/apostila-em-portugues-de-introducao-ao-qgis/>

Vieira, Sonia. 1980. Introdução à Bioestatística. 3a Ed revista e ampliada. Editora Campus, Rio de Janeiro, RJ, xvii + 196 pp.

Wickham, H, & G. Grolemund. 2019. R para Data Science. Editora Alta Books, Rio de Janeiro, RJ, 528 pp.