

BOLETIM INFORMATIVO DO SER

V SER

R %>% REMOTES()

A QUINTA edição do Seminário Internacional de Estatística com R marcou seu recorde de inscrições e devido a pandemia da Covid-19, inovou na sua forma de transmissão das palestras com chat ao vivo, dessa vez totalmente online e gratuita.

A REVISTA DO SER foi reformulada e agora conta com a recepção contínua de trabalhos realizados com o uso da linguagem R.

O número de inscritos no V SER chegou a quase 1.500 participantes de todos os Estados do Brasil e Distrito Federal, provenientes de 15 países.

Contou com uma grande equipe de organizadores voluntários formado em sua maioria por professores da AFA, UFF, UNIRIO, UERJ, Colégio Estadual João Bazet, Univ. de Santiago de Compostela e Univ. de Aveiro.

Volume 5/ Edição 1 Ano 2021 – Luciane F. Alcoforado, Ariel Levy e Orlando C. Longo

SER maior evento de R na América Latina totalmente online

O V SER – Seminário Internacional de Estatística com R ocorreu nos dias 09 e 10 de junho de 2021 de forma remota sendo transmitido no canal do youtube (<https://www.youtube.com/seruff>) e contou com a inscrição de 1426 pessoas do Brasil e exterior.



Figura 1: Inscritos no V SER

Foram DOIS DIAS de transmissão de palestras ao vivo no canal do youtube, que registrou 2640 e 1714 visualizações no dia 09 e 10 de junho de 2021, respectivamente.

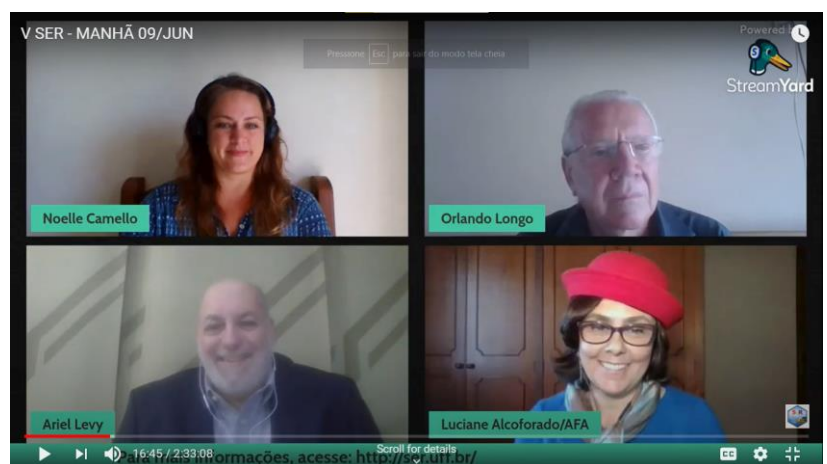


Figura 2: Abertura do V SER no dia 09 de junho de 2021 no canal seruff

A abertura do V SER foi realizada pelos coordenadores desta edição, o prof. **Dr. Orlando Celso Longo** coordenador do Programa de Pós-Graduação em Engenharia Civil da Universidade Federal Fluminense, a profa. **Dra. Luciane Ferreira Alcoforado**, professora da Divisão

de Ensino da Academia da Força Aérea e colaboradora no Programa de Pós Graduação em Engenharia Civil da Universidade Federal Fluminense e o prof. **Dr. Ariel Levy**, professor do Programa de Pós-Graduação em Administração da Universidade Federal Fluminense, assessorados pela mestre de cerimônias do V SER a doutoranda **Noelle Camello** da Universidade de Aveiro que deram as boas-vindas aos participantes, destacando a intensa participação com **quase 1500 inscritos** de todos os Estados do Brasil e Distrito Federal, são pessoas vinculadas à **15 países** como Alemanha, Angola, Brasil, Cabo Verde, Canadá, Chile, Equador, Espanha, França, Irlanda, Itália, Paraguai, Peru, Reino Unido e São Tomé e Príncipe e apoiada por grandes instituições como a Universidade Federal Fluminense (UFF), a Academia da Força Aérea (AFA), a Escola Nacional de Ciências Estatísticas (ENCE), a Universidade Federal do Estado do Rio de Janeiro (UNIRIO), a Universidade Estadual do Rio de Janeiro (UERJ), Universidade de Santiago de Compostela e Universidade de Aveiro.



Figura 3: Revista do Seminário com submissão contínua de trabalhos

Os participantes foram convidados a submeter seus artigos para a **REVISTA DO SER** (<https://periodicos.uff.br/anaisdoser>) que foi reformulada e passa a receber continuamente a submissão de trabalhos. A missão agora é elevar a revista para o mais alto nível de referência da CAPES.

Nesta edição as palestras foram pré-gravadas e exibidas no canal do youtube, a sala de atividades passou a ser o chat ao vivo com o palestrante que foi recepcionado por dois professores moderadores que intermediaram as perguntas dos participantes. A transmissão das palestras foi dividida em 6 blocos, distribuídos pelo turno da manhã, tarde e noite.



Figura 4: Os participantes se comunicavam na sala de bate papo promovendo o encontro remoto do SER

No primeiro bloco tivemos a exibição de quatro palestras:

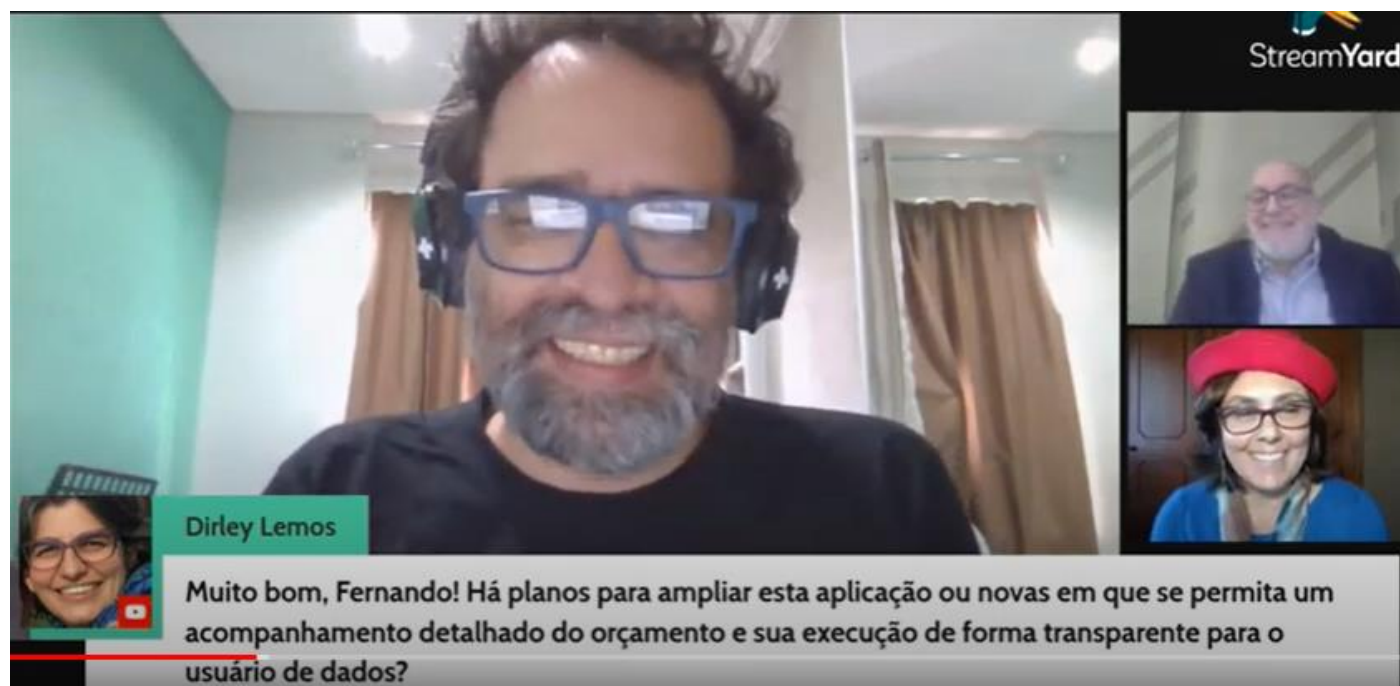


Figura 5: Palestra de Fernando Barbalho, com moderação dos professores Ariel Levy(UFF) e Luciane Alcoforado(AFA)

Abrindo o evento foi dado início com a palestra intitulada “Meu SICONFI: transparência pública, shiny, SQL, máquina de estado, dataviz e machine learning tudo ao mesmo tempo agora” ministrada por Fernando Barbalho, doutor em Administração pela Universidade de Brasília. Atualmente é analista de finanças e controle do Ministério da Fazenda. Tem experiência na área de Administração, com ênfase em Organizações Públicas, atuando academicamente principalmente nos seguintes temas: dados abertos, ciência de dados, governo aberto, transparência, inovação e governo eletrônico. A palestra aborda o interesse em se consumir dados de finanças públicas dos entes federativos brasileiros. As opções disponíveis para esse consumo trazem vários constrangimentos para quem não tem conhecimento aprofundado da Lei de Responsabilidade Fiscal e da arquitetura do Sistema que centraliza essas informações na Secretaria do Tesouro Nacional. A aplicação Meu SICONFI foi desenvolvida para tornar mais fácil esse consumo a partir de um diálogo entre aplicação e usuário. Esse produto foi desenvolvido usando Shiny e flexdashboard com recursos de máquina de estado, consumo de dados usando SQL, visualização de dados em gráficos e mapas e clusterização. A palestra mostra rapidamente como é o uso da aplicação e em seguida navega pelos principais códigos.



Figura 6: Palestra de Larissa Alves com moderação das professoras Paulina Montejano(AFA) e Daniele Magierski(AFA)

A segunda palestra “Inferência variacional bayesiana como alternativa ao MCMC”, ministrada por Larissa de Carvalho Alves, que possui graduação em Matemática, Mestrado e Doutorado em Estatística pela Universidade Federal do Rio de Janeiro. Atualmente é professora e pesquisadora da Escola Nacional de Ciências Estatísticas do IBGE. Tem experiência na área de Probabilidade e Estatística atuando principalmente nos seguintes temas: Inferência Bayesiana, Modelos Dinâmicos e Previsões Bayesianas. A palestra aborda a inferência variacional (ou Bayes variacional - VB) que é um método de machine learning que aproxima densidades de probabilidade por meio de otimização usando a divergência de Kullback-Leibler. Sabe-se que a inferência Bayesiana tem como principal meta obter a distribuição a posteriori que na maioria dos problemas é desconhecida. Portanto, o VB pode ser utilizado como um método alternativo ao tradicional MCMC para obter uma aproximação para a densidade a posteriori. É apresentado resultados simulados para modelos de regressão linear e seleção de variáveis, comparando o MCMC e o VB. É indicado os principais pacotes já desenvolvidos no software R que lança mão do VB para fazer inferência. Propõe ainda uma aplicação aos dados de Covid-19 com o objetivo de capturar a tendência do número de casos do Brasil por meio do ajuste de um modelo de regressão spline utilizando o VB.

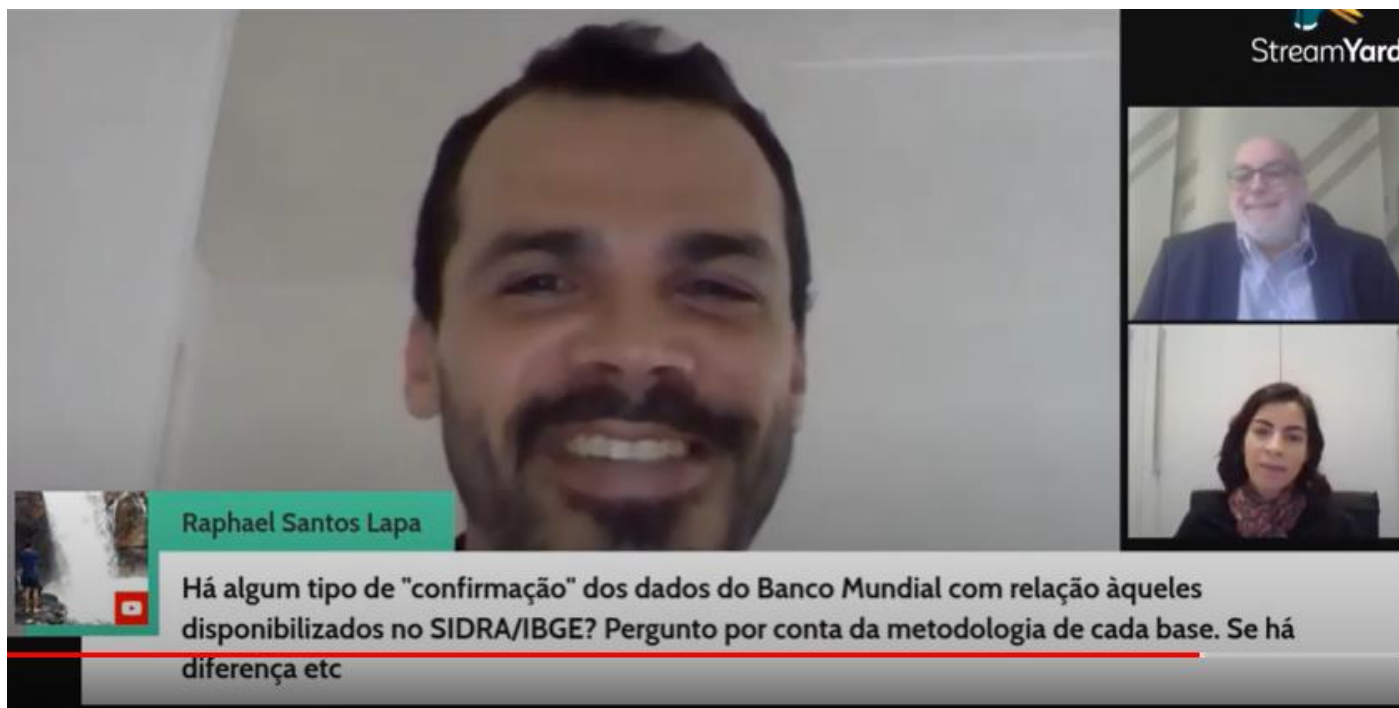


Figura 7: Palestra de Bruno Tômio com moderação dos professores Ariel Levy(UFF) e Iliane Foresti(AFA)

A terceira palestra intitulada “Collect and use open access World Bank data to know your country” de Bruno Thiago Tômio, professor de economia na Universidade de Blumenau – FURB e doutorando na Université Grenoble Alpes na França. A palestra aborda o pacote {WDI} permite explorar 40 conjuntos de dados diferentes hospedados pelo Banco Mundial. A base de dados World Development Indicators (‘WDI’) contém 1.600 indicadores de séries temporais para 217 economias e mais de 40 grupos de países. Para muitas séries, os dados datam de mais de 50 anos. O pacote é muito útil para explorar o desenvolvimento dos países. Faremos o download e plotaremos os dados juntos durante esta sessão.

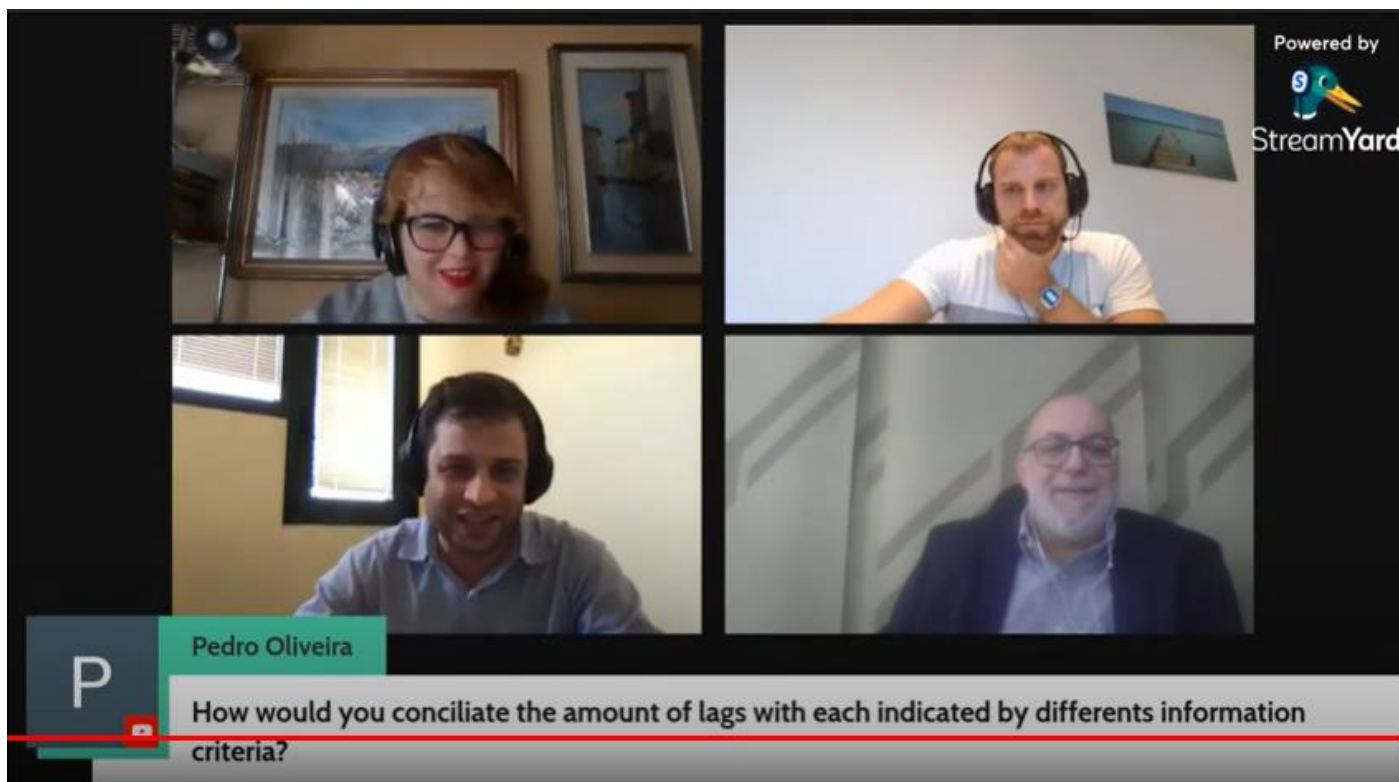


Figura 8: Palestra de Alessia Paccanhini e Fabio Parla, com moderação dos professores Andrea Ugolini(UERJ) e Ariel Levy(UFF)

A quarta palestra intitulada “Vector Autoregressive Models in R.” Apresentada por Alessia Paccanhini, professora assistente na University College Dublin, onde é Diretora Acadêmica do Mestrado em Finanças Quantitativas e Co-Diretora do Women in STEM Committee e pesquisadora. A palestra replica o exemplo em Kilian e Lutkepohl (2017), Structural Vector Autoregressive Analysis, Cambridge University Press, Capítulo 9 (página 239-240). O exercício centra-se na identificação de um choque nos preços do petróleo e seu impacto na inflação e no PIB real. O choque do preço do petróleo é identificado usando um esquema recursivo (isto é, através da decomposição de Cholesky da matriz de covariância de resíduos de forma reduzida do VAR). Os códigos R propostos permitem calcular funções de resposta ao impulso (IRFs), Decomposição de Variância de Erro de Previsão (FEVD) e Decomposição Histórica (HD).

No segundo bloco tivemos a exibição de cinco palestras:

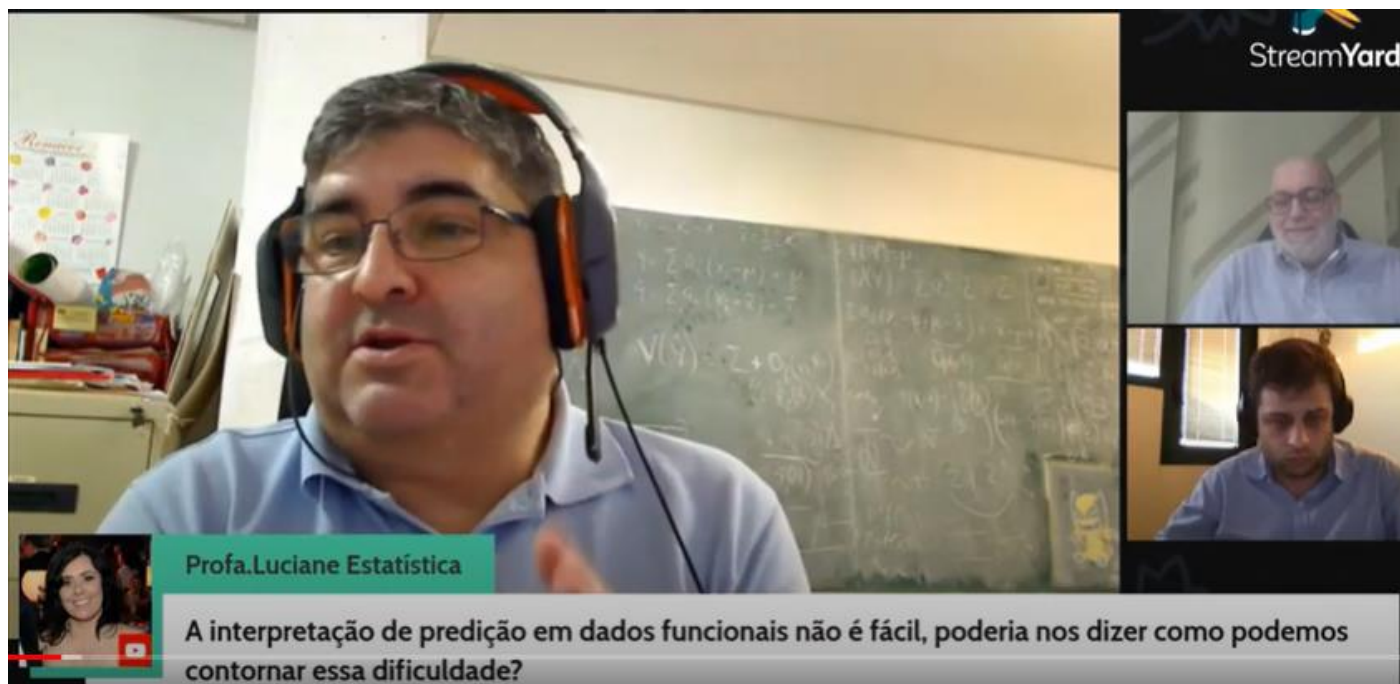


Figura 9: Palestra de Manuel Febrero, com moderação dos professores Ariel Levy(UFF) e Andrea Ugolini(UERJ)

A primeira deste bloco foi a palestra de título “Building a ShinyApp for CoVid-19 from the very begining” ministrada por Manuel Febrero Bande, professor catedrático de Estatística e Pesquisa Operacional na Universidad Santiago de Compostela. Seus interesses de pesquisa incluem: séries temporais, Bootstrap, estimativa não paramétrica, estatística espacial, Redes neurais dados funcionais. Nesta apresentação, são mostradas as etapas necessárias para publicar uma previsão do Shiny Apps for COVID19. Desde a primeira ideia estatística que envolve técnicas de Análise de Dados Funcionais estendendo os modelos epidemiológicos clássicos baseados em equações diferenciais até os detalhes práticos de implantação do Shiny Server e seu custo.



Figura 10: Palestra de Eduardo e Morgany com moderação das professoras Iliane Foresti(AFA) e Renata Mori(AFA)

A segunda palestra “Bibliometria com R: Aplicação e uso do pacote Bibliometrix na função Biblioshiny”, ministrada por Eduardo Soares Jangutta e Morgany Leite dos Santos, Eduardo é Mestrando em Administração (PPGAd/UFF). Possui especialização em Finanças (UNICSUL) e formado em gestão empreendedora (UFF) com experiência profissional no setor público e privado. Morgany é Mestranda em Administração (PPGAd/UFF) Administradora (UFRRJ). Possui Magistério em educação com experiência profissional em ensino, pesquisa e consultoria. Palestra dividida em Parte 1 com Morgany Leite: Contextualização de proposta de pesquisa bibliométrica sobre feminismo e gestão em duas bases internacionais: SCOPUS e Web of Science. Tutorial de como buscar e baixar as buscas. Parte 2 com Eduardo Jangutta: Como reunir as buscas realizadas e analisá-las em uma única base de dados sem duplicados no formato csv ajuda do RStudio e sua interface gráfica biblioshiny.

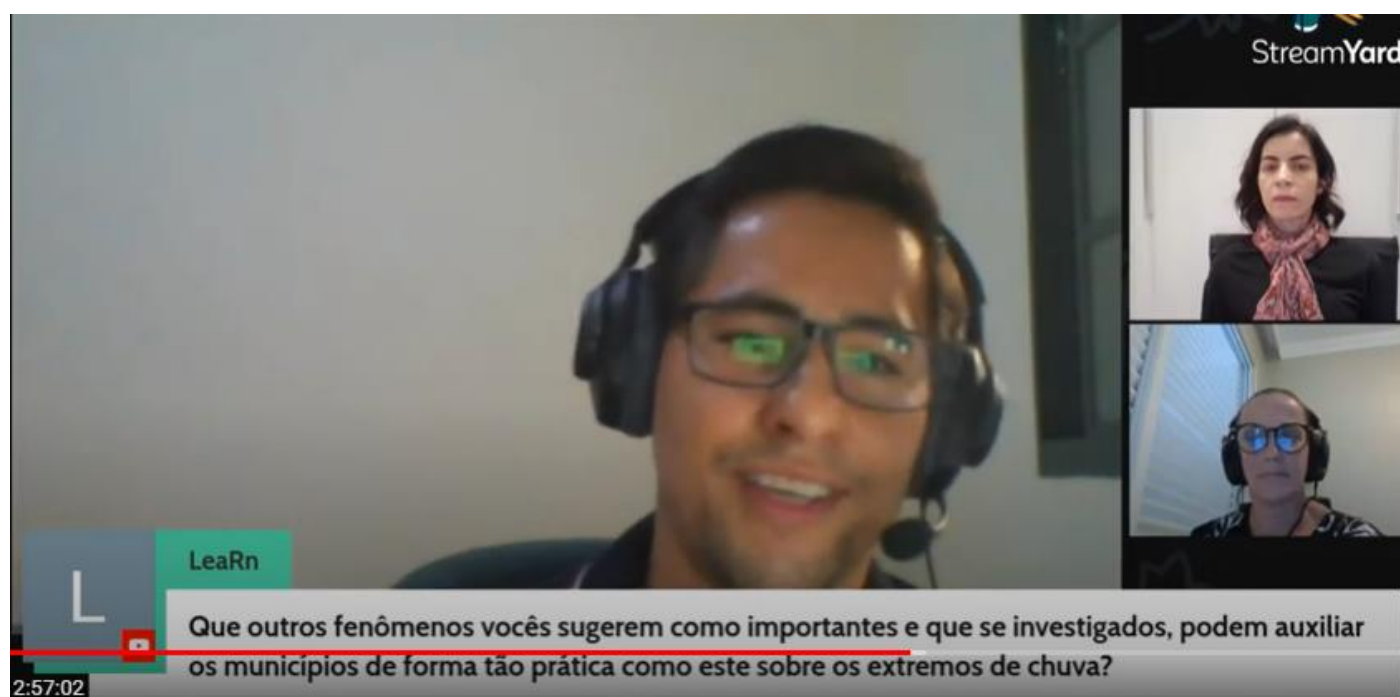


Figura 11: Palestra de Gilberto Liska com moderação das professoras Iliane Foresti (AFA) e Renata Mori(AFA)

A terceira palestra intitulada “Distribuição Generalizada de Pareto aplicada à análise de eventos extremos de chuva em Uruguaiana-RS” apresentada por Gilberto Rodrigues Liska, que possui graduação em Matemática Licenciatura pela Universidade Federal de Alfenas (2010), mestrado (2012) e doutorado (2016) em Estatística e Experimentação Agropecuária pela Universidade Federal de Lavras. É professor de estatística, matemática e fundamentos de programação na Universidade Federal de São Carlos. A palestra aborda sobre como o monitoramento das chuvas permite compreender o ciclo hidrológico que não só influencia a dinâmica ecológica e ambiental, mas também afeta as atividades econômicas e sociais. Esses setores são muito afetados quando as chuvas ocorrem em valores superiores à média, chamados de eventos extremos; além disso, as metodologias estatísticas baseadas na ocorrência média desses eventos são inadequadas para analisar esses eventos extremos. A Teoria dos Valores Extremos fornece modelos teóricos adequados para este tipo de evento; portanto, a Distribuição Generalizada de Pareto (Doravante GPD) é usada para analisar os eventos extremos que excedem um determinado limite. O presente trabalho aplicou o GPD e sua versão aninhada, a Distribuição Exponencial, em dados pluviométricos mensais da cidade de Uruguaiana, no estado do Rio Grande do Sul, Brasil. Para apoiar os

resultados, foi utilizado alguns critérios de qualidade de ajuste e proposto um procedimento de simulação de Monte Carlo para detectar a verdadeira distribuição de probabilidade em cada mês analisado.

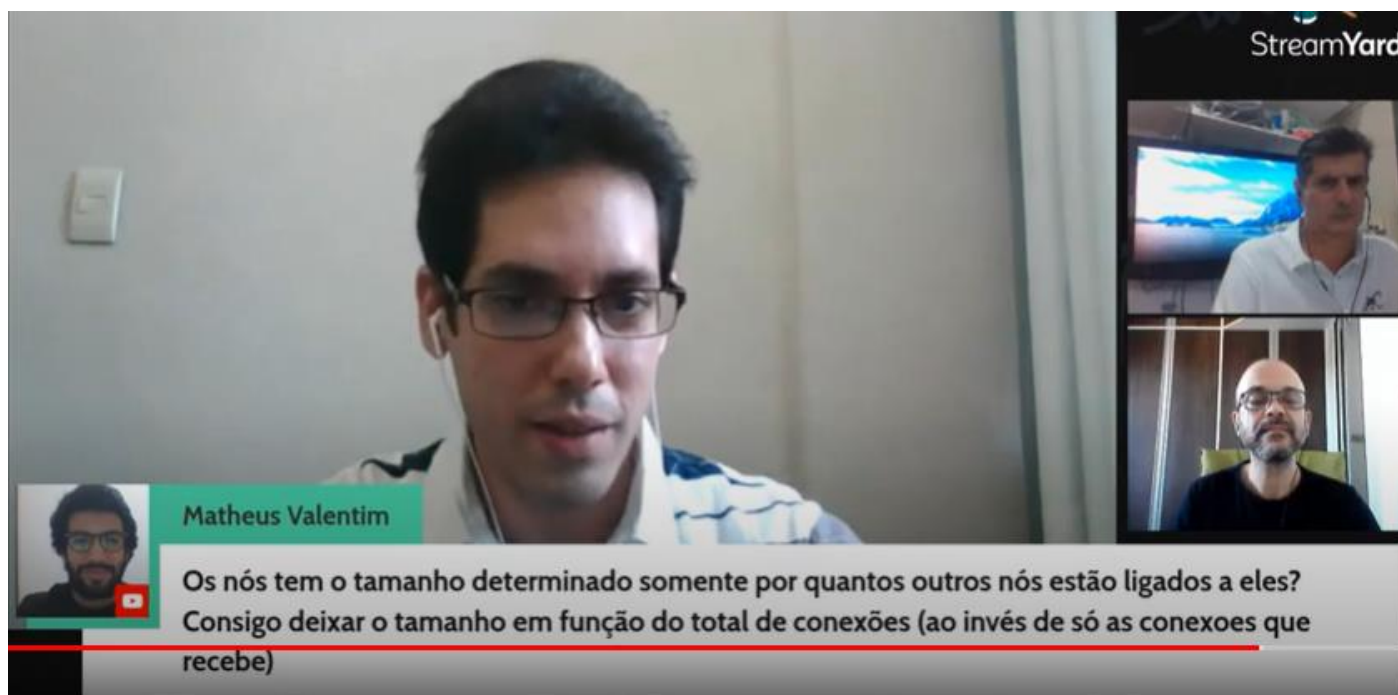


Figura 12: Palestra de Orlando Guilarte, com moderação do professor Paulo Rioli(AFA) e do mestrando Cleuson Alves(UFF)

A quarta palestra intitulada “Manipulação de grafos no R”, apresentada por Orlando Fonseca Guilarte, pesquisador da PUC Rio. Doutor em Ciências - Matemática pela Pontifícia Universidade Católica do Rio de Janeiro, Brasil (2019). Possui graduação em Ciência da Computação (2012) e mestrado em Ciências Matemáticas (2014) pela Universidad de La Habana. Tem experiência na área de Matemática Aplicada, principalmente nos seguintes temas: Visualização de informações, Data Science, Otimização e Computação gráfica. A palestra trata sobre a manipulação de grafos no R e o uso de banco de dados orientado a grafos são temáticas que ultimamente vêm merecendo grande destaque, em particular na área de Ciência de Dados. Nesta palestra são tratados os pacotes `igraph`, `threejs`, `networkD3` e `visNetwork` (análise de grafos e visualização) e o pacote `neo4r` (driver do banco de dados Neo4j).



Figura 13: Palestra de Matheus Barros, com moderação do mestrando Cleuson Alves(UFF) e do professor Paulo Rioli(AFA)

A quinta e última palestra desta rodada intitulada “Como usar a plataforma Base dos Dados com R”, apresentada por Matheus Valentim de Barros, formado em Economia pela FEAUSP, Mestrando em Forest and Management na Universidade de Copenhague, Fundador da página @jfeddados nas redes sociais e Assistente de Dados da Base dos Dados. A palestra apresenta a utilização "pura" da biblioteca da Base dos Dados para R, apresentando a sua estrutura e suas principais funções e apresenta ainda como combinar o pacote com outros pacotes e funcionalidades do R.

No **terceiro bloco**, fechando o primeiro dia do evento tivemos a exibição de mais duas palestras:



Figura 14: Palestra de Paulo Guimarães, com moderação dos professores Carlos Alberto Barbosa(AFA) e Steven Dutt-Ross(UNIRIO)

Iniciou com a palestra de título “Best R Libraries for Machine Learning” ministrada por Paulo Henrique Sales Guimarães, professor no Instituto de Ciências Exatas e Tecnológicas da Universidade Federal de Lavras (UFLA). Doutor em Estatística e Experimentação Agropecuária. Trabalha na área de Séries Temporais, Econometria e Aprendizado de Máquinas. Esta palestra fala sobre os principais pacotes usados no aprendizado de máquina supervisionado.

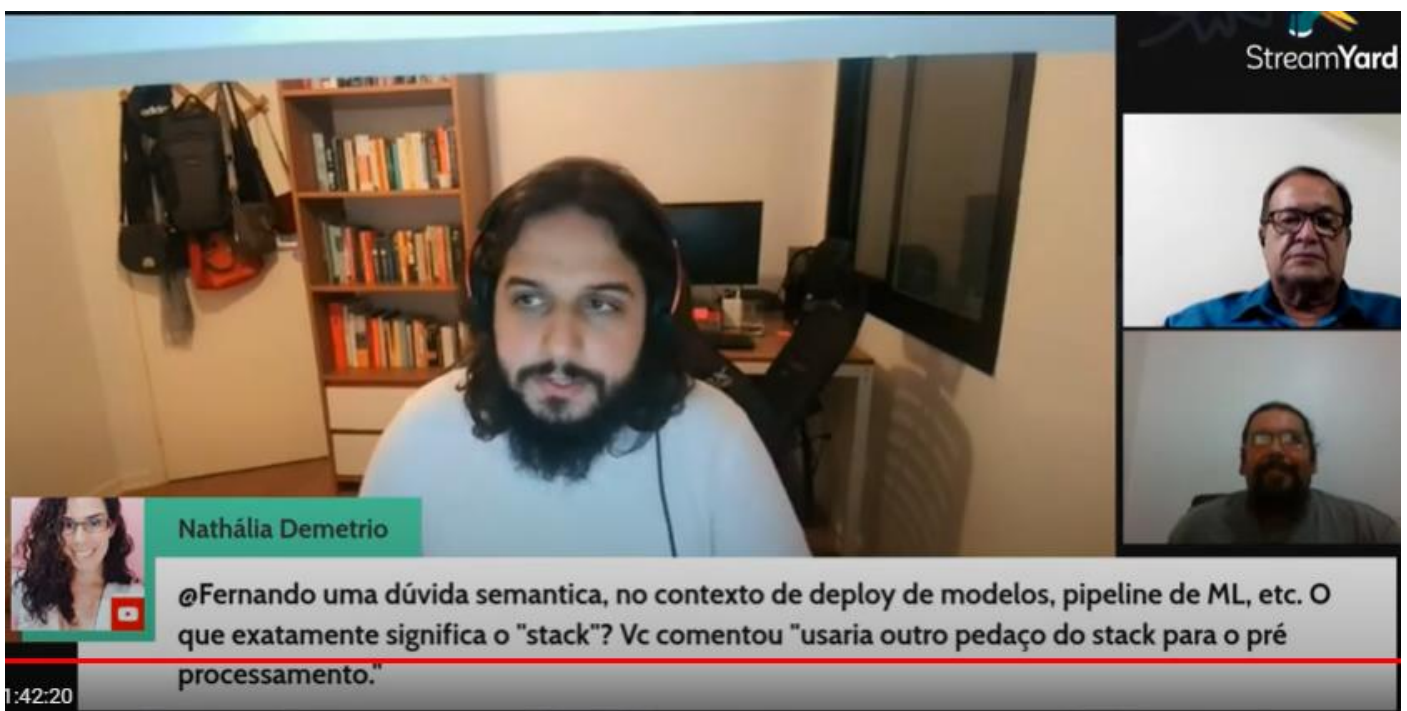


Figura 15: Palestra de Fernando Corrêa, com moderação dos professores Carlos Alberto Barbosa(AFA) e Steven Dutt-Ross(UNIRIO)

A segunda palestra “Deploy de modelos de Machine Learning em R”, ministrada por Fernando Corrêa, Bacharel e mestrando em Estatística pelo IME-USP. Diretor-técnico na Associação Brasileira de Jurimetria. Usa R para tudo, mas tem interesse especial em web scraping, visualização de dados e modelagem bayesiana. Analisar dados e construir modelos preditivos ou de decisão são as principais atividades de Cientistas de Dados, mas o processo de entrega desses modelos vem logo depois na ordem de prioridades. Nesta apresentação é mostrado como usar um dos mais flexíveis frameworks de modelagem para disponibilizar modelos de Machine Learning em APIs REST, abarcando desde o pré-processamento dos dados até a devolução do resultado dos modelos de previsão em formato JSON. A solução utilizará docker containers para garantir a robustez e reprodutibilidade do resultados, além os pacotes tidymodels e plumber da linguagem R.

No **segundo dia** as atividades começaram com o **quarto bloco** de palestras, composto de quatro palestras:



Figura 16: Palestra de Bruna Wundervald, com moderação das professoras Dirley Lemos (AFA) e Luciane Alcoforado (AFA)

A primeira deste bloco foi a palestra de título “Feature Selection via Gain Penalization in Random Forests” ministrada por Bruna Wundervald, doutoranda em Bayesian Machine Learning na Maynooth University, membro de alguns grupos R-Ladies (São Paulo, Curitiba e Dublin), usuária do R ativa e líder de comunidade R. Especialmente interessada em inferência Bayesiana, análise multivariada, extração de características, modelos gráficos probabilísticos e todos os tipos de coisas relacionadas a R. Florestas aleatórias são um algoritmo de aprendizado de máquina baseado em conjunto de árvores muito popular, que são conhecidos por não fazer a seleção de recursos muito bem. Eles normalmente usam todos ou a maioria dos recursos que lhes são fornecidos e se esforçam muito para detectar recursos altamente correlacionados. Nesta palestra, é dada uma visão geral sobre um novo método de seleção de recurso de penalização de ganho para florestas aleatórias, com foco em sua aplicação a conjuntos de dados de alta dimensão (p grande). É apresentada uma nova ideia de penalização de ganho que exibe uma regularização local-global geral para modelos baseados em árvore. O novo método permite flexibilidade total na escolha de pesos de importância específicos de recursos, ao mesmo tempo que aplica uma penalização global. Valida-se o método em dados simulados e reais, explorando como os hiperparâmetros interagem e fornece a implementação como uma extensão do popular pacote Ranger.

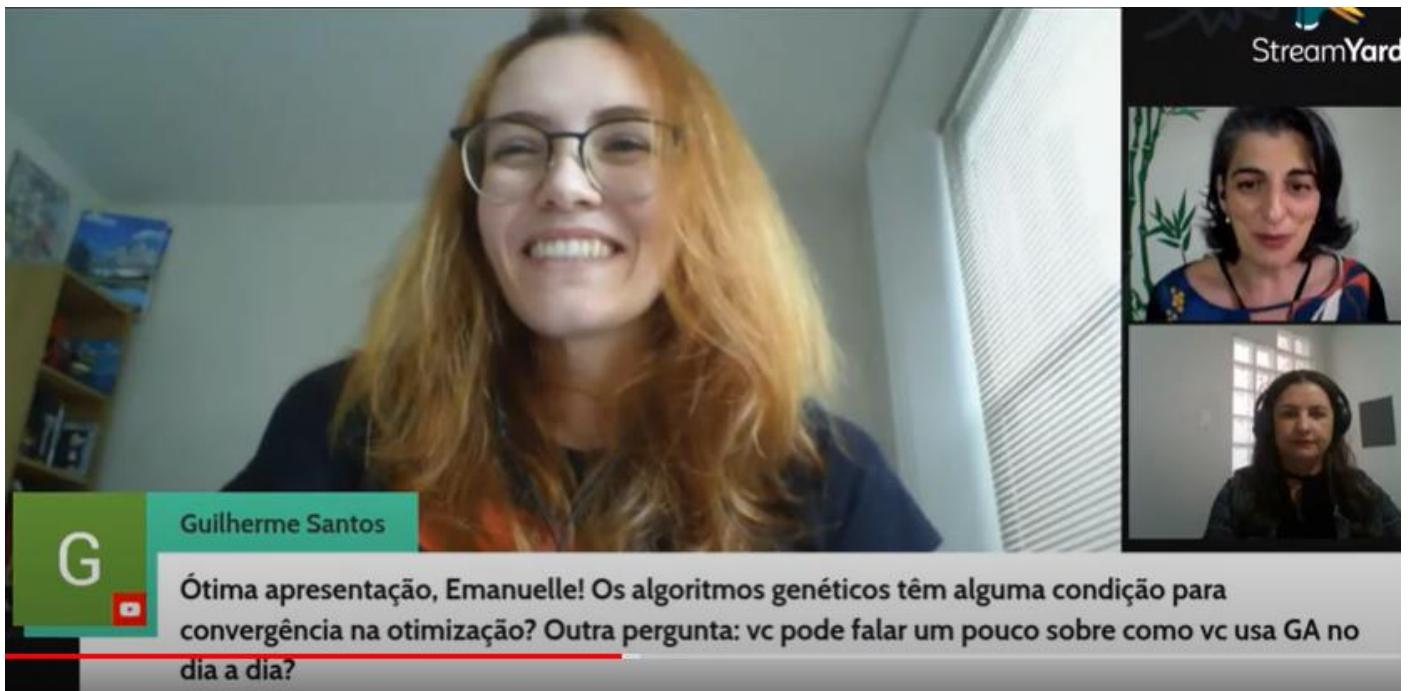


Figura 17: Palestra de Emmanuelle Nunes, com moderação das professoras Dirley Lemos (AFA) e Paulina Montejano (AFA)

A segunda palestra “Genetic Algorithms in R”, ministrada por Emmanuelle Rodrigues Nunes, que é bacharel em Estatística pela USP com mestrado em epidemiologia pela USP e em Ciência da Computação Avançada pela Universidade de Reading. Atualmente atua como cientista de dados senior na Universidade de Reading na Inglaterra. Esta palestra tem como objetivo apresentar o conceito de Algoritmo Genético e demonstrar como aplicar esta técnica no R com as bibliotecas GA. Técnicas de otimização são usadas para encontrar os valores de um conjunto de parâmetros que maximizam ou minimizam alguma função objetivo de interesse. Esses métodos tornaram-se de grande importância nas estatísticas para estimativa, ajuste de modelo, etc. As soluções potenciais estão em uma população e são codificadas como "genes". Podemos produzir novas soluções 'mutando' membros da população atual e 'cruzando' duas soluções para formar uma nova solução. As melhores soluções são selecionadas para se reproduzir e sofrer mutação, e as piores são descartadas. Por se tratar de um método probabilístico, as soluções que exploram não são determinadas apenas pelas propriedades do problema de interesse. Algoritmos genéticos são usados na inteligência artificial como uma alternativa aos métodos mais tradicionais quando nos deparamos, por exemplo, com problemas que atuam em espaços discretos, onde métodos baseados em gradientes não podem ser usados.



Figura 18: Palestra de Fernando Ribeiro, com moderação dos professores Paulo Rioli (AFA) e Daniele Magierski (AFA)

A terceira palestra intitulada “League of Legends - Como ser um bom jogador?” de Fernando Ribeiro, graduando de Administração Pública pela UNIRIO. A palestra aborda sobre a Aplicação de recursos do R em variáveis do jogo mais jogado do mundo.

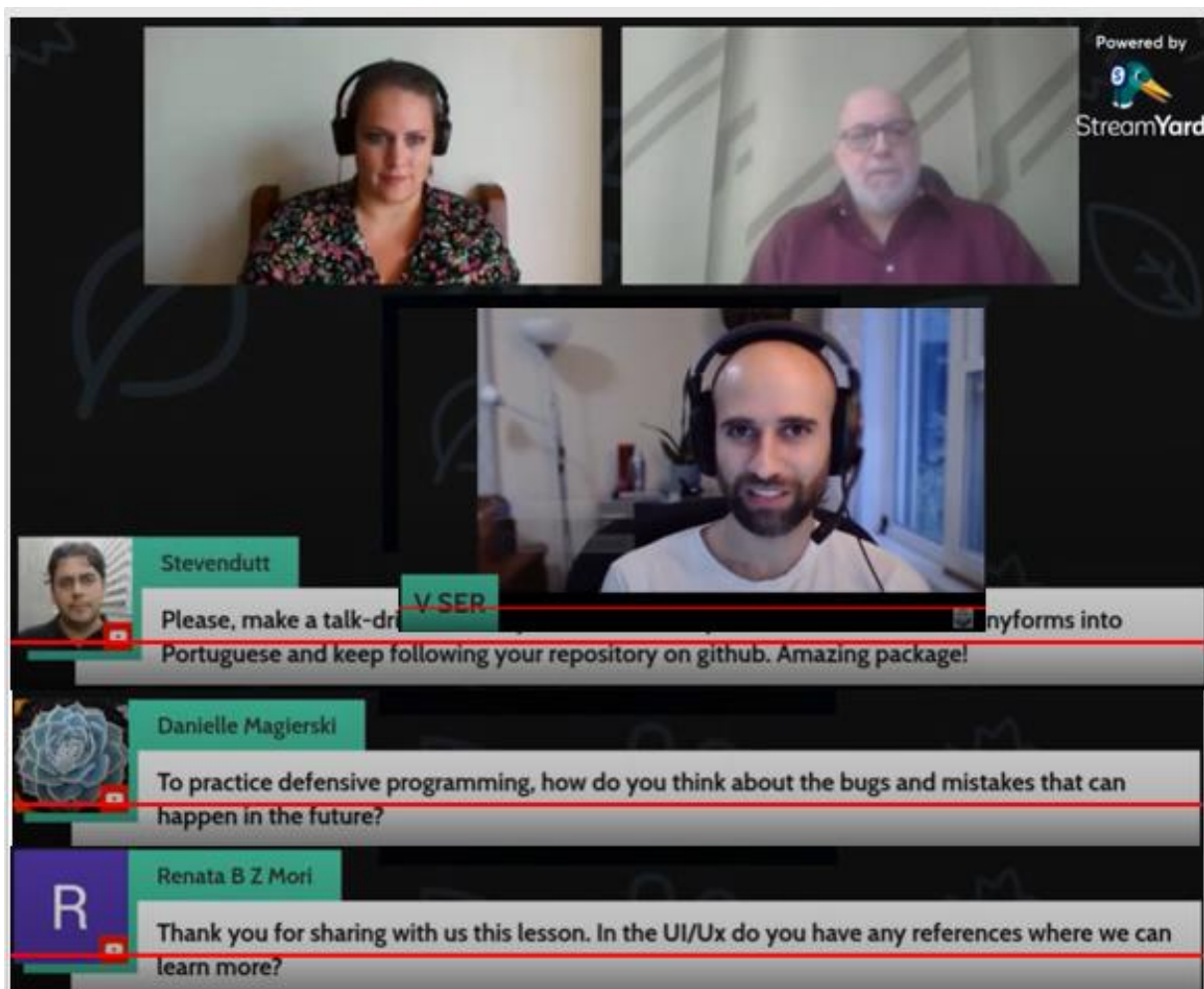


Figura 19: Palestra de Dean Attali, com moderação da doutoranda Noelle Camello (Univ.Aveiro) e do professor Ariel Levy(UFF)

A quarta e última palestra desta rodada intitulada “Taking your shiny apps to the next level: Tips for building great user-facing tools”, apresentada por Dean Attali. Dean é o cara do Shiny. Ele trabalha como consultor Shiny, e se você já fez algum trabalho com o Shiny, provavelmente o encontrou. Você pode ter lido um tutorial em seu blog, ou usado suas respostas no StackOverflow, ou usado um de seus pacotes. Se você não usa o Shiny, poderá reconhecer o site dele, já que muitas pessoas na comunidade R constroem seus sites usando seu modelo chamado Beautiful Jekyll. Em sua palestra nos conta que constrói aplicativos Shiny com pacotes R. Sempre que se cria uma ferramenta para uso de outras pessoas, é claro que se concentra em fazê-la funcionar. Mas há muitas outras considerações a levar em conta se você deseja que seu software sobreviva e se torne popular. Nesta palestra, Dean compartilha algumas lições aprendidas com o desenvolvimento do CRANalerts e sobre como criar ótimas ferramentas voltadas para o usuário.

No quinto bloco tivemos a exibição de cinco palestras:



Figura 20: Palestra de Vanessa e Alexandre, com moderação dos professores Renata Mori (AFA) e Paulo Rioli (AFA)

Iniciamos com a palestra de título “O R na Educação Básica” ministrada por Vanessa Manhães e Alexandre Sousa da Silva, Alexandre Sousa da Silva é doutor em Estatística pela Universidade Federal do Rio de Janeiro (2012), mestre em Estatística e Experimentação Agronômica pela Escola Superior de Agricultura "Luiz de Queiroz" (2007) e graduado em Estatística pela Universidade Estadual Paulista Júlio de Mesquita Filho (2004). Atualmente é professor adjunto da Universidade Federal do Estado do Rio de Janeiro. Tem experiência na área de Probabilidade e Estatística, com ênfase em Análise Espacial, Modelos Espaço-Temporais Bayesianos e Educação Estatística. Vanessa Manhães é licenciada em Matemática e atua na rede de ensino do Estado do Rio de Janeiro desde 2015. Iniciou um curso de Extensão Estatística e Probabilidade no Ensino Médio em 2019 no projeto Livro Aberto, participou do I Educa-SER UFF/UNIRIO, a partir do qual percebeu a aplicabilidade da linguagem R para o Ensino Médio. Integrou-se ao projeto de Extensão Estatística é com R! da UFF e seu objetivo passou a ser o desenvolvimento de atividades utilizando a linguagem R para estudantes do ensino médio. Nessa palestra, os professores Vanessa e Alexandre fazem uma reflexão sobre a importância do estudo de Estatística para o engajamento cívico e apresentam sites e projetos que utilizam o R como ferramenta facilitadora do ensino da mesma, aplicada à realidades da Educação Básica.

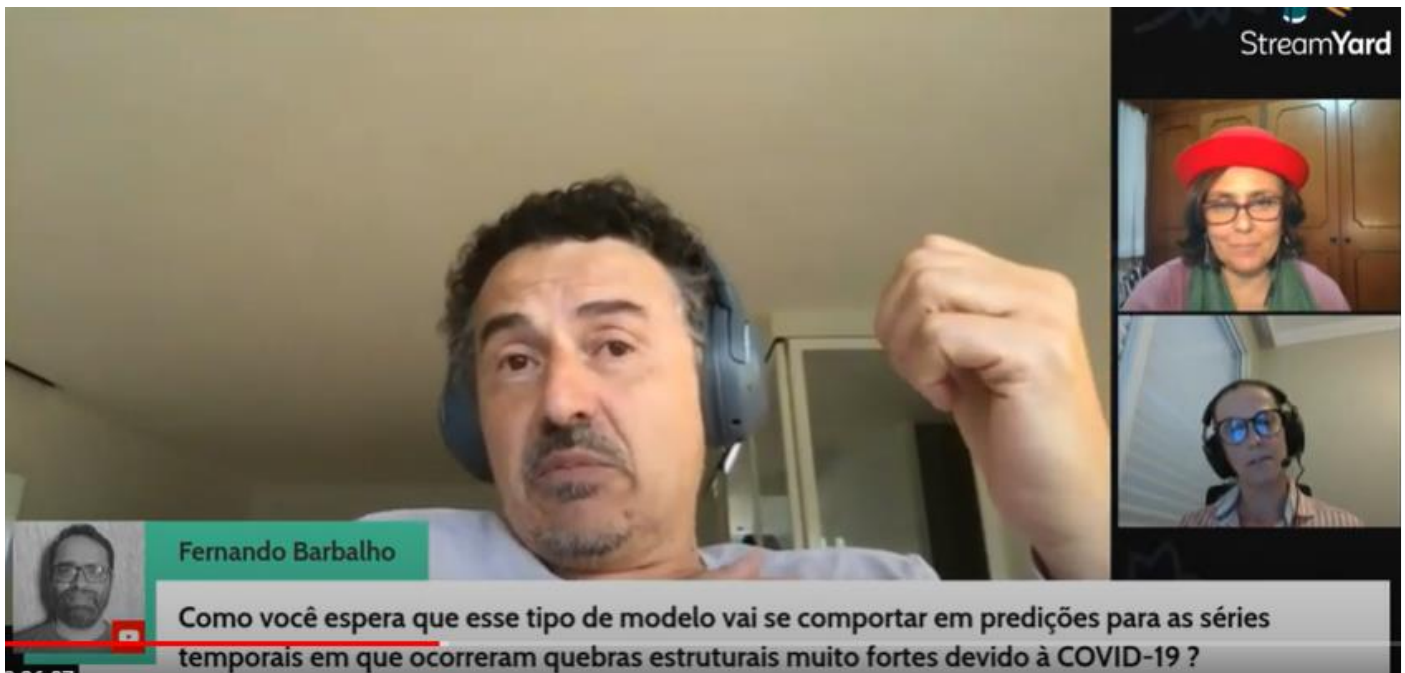


Figura 21: Palestra de Luis Torgo, com moderação de Luciane Alcoforado (AFA) e Renata Mori (AFA)

A segunda palestra “Ensembles for Time Series Forecasting”, ministrada por Luis Torgo, é líder de Pesquisa no Canada e Professor da Faculty of Computer Science da Dalhousie University, é ainda Professor Associado da Faculdade de Ciências da Universidade do Porto, Portugal, e Professor Convidado da Stern Business School da New York University, USA. Luis Torgo é o vice diretor do Institute for Big Data da Dalhousie University e líder do Ocean Data Analytics Lab desse instituto. Nesta palestra descreve e ilustra duas abordagens de previsão de séries temporais, usando ensembles. Ambas as abordagens são baseadas na ideia de gerar diversidade entre os modelos usando propriedades dinâmicas das séries temporais. Resultados experimentais demonstram a eficácia das abordagens, em particular o método ADE disponível na package `tsensemble` do R.

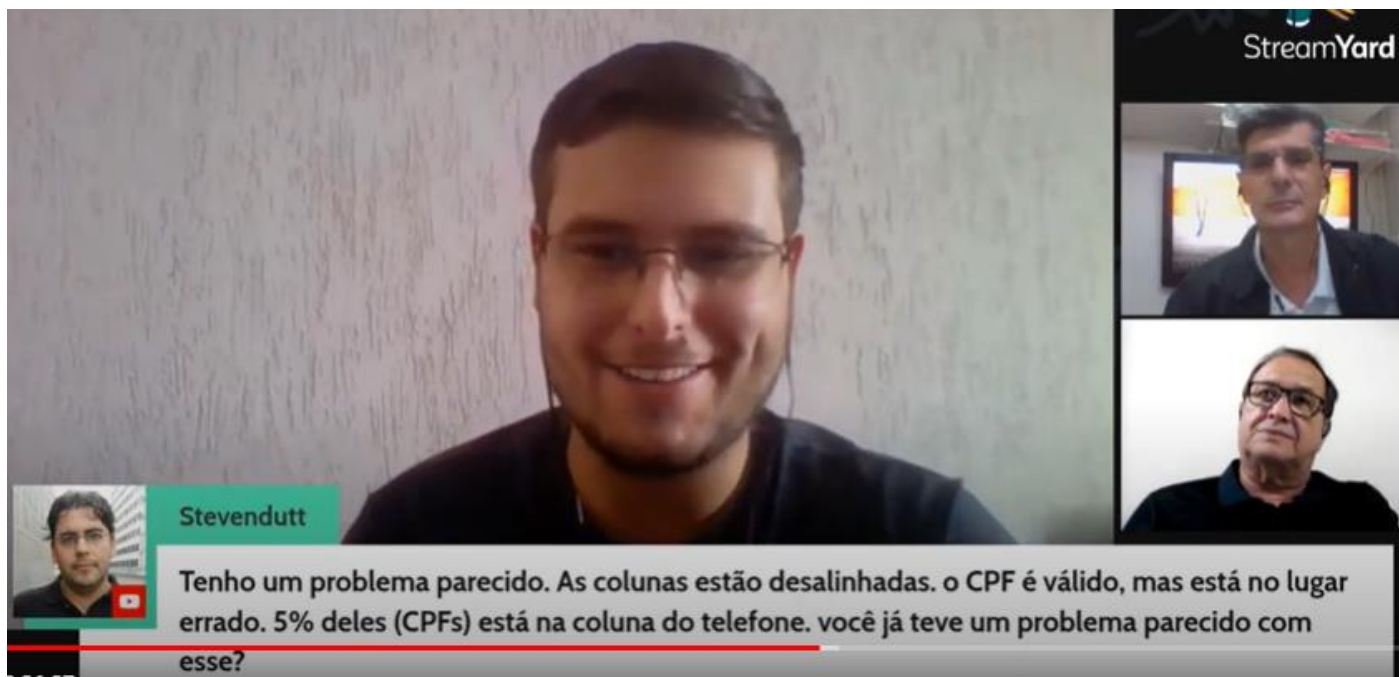


Figura 22: Palestra de Leonardo Filgueira, com moderação dos professores Paulo Rioli (AFA) e Carlos Alberto Barbosa(AFA)

A terceira palestra intitulada “Qualithron: a qualidade faz a diferença apresentada por Leonardo Filgueira, graduado em Estatística pela UFF e cientista de dados na DTM. Nesta palestra ele compartilha um pouco do seu dia a dia como Cientista de Dados na DTM. Dentro do time de Data Science é responsável pela automatização de processos, criação de pacotes e, mais recentemente, pelo desenvolvimento de um produto. Ele conta detalhes sobre o pacote de qualificação de dados feito em R e Python (qualithron) que é usado por diversos clientes e foi desenvolvido por ele, com o apoio da equipe de Data Science e que obteve um resultado surpreendentemente positivo logo que teve sua primeira versão analisada.



Figura 23: Palestra de Beatriz Milz, com moderação dos professores Steven Ross (UNIRIO) e Ariel Levy (UFF).

A quarta palestra intitulada “Coleta de dados automatizada e integração contínua utilizando GitHub Actions: Exemplos com o Pacote Mananciais”, apresentada por Beatriz Milz, que é doutoranda em Ciência Ambiental na Universidade de São Paulo, e começou a programar R em 2018 para usar na pesquisa. Está envolvida em várias comunidades, eventos e atividades sobre R, como: R-Ladies, Curso-R, satRday São Paulo, Latin-R, useR 2021, pacote dados, pacote mananciais, e o livro Ciência de Dados em R. O GitHub Actions (GHA) (<https://docs.github.com/pt/actions>) permite automatizar fluxos de trabalho de desenvolvimento de pacotes em R. Nessa palestra, é mostrado alguns exemplos de uso de GitHub Actions no pacote Mananciais, que disponibiliza dados sobre o volume operacional em reservatórios utilizados para abastecimento público na Região Metropolitana de São Paulo. Os exemplos que utilizam GHA são: atualização da base de dados, checagem do pacote, execução de testes e avaliação da cobertura de testes. O pacote foi desenvolvido por Beatriz Milz, e pode ser acessado em: <https://beatrizmilz.github.io/mananciais/index.html>. Este pacote foi pensado para atender demandas de pessoas pesquisadoras que não somente utilizam R, disponibilizando também a base de dados atualizada diariamente em csv e em xlsx.



Figura 24: Palestra de Julio Trecenti, com moderação dos professores Steven Ross (UNIRIO) e Ariel Levy (UFF).

A quinta palestra “GitHub Actions”, é apresentada por Julio Trecenti, que costuma se dizer faxineiro de Dados na Curso-R, é bacharel, mestre e doutorando em Estatística na Universidade de São Paulo. Secretário-Geral da Associação Brasileira de Jurimetria. Conselheiro no Conselho Federal de Estatística, com interesse em Estatística computacional, raspagem de dados, aprendizado estatístico e Jurimetria. Ele fala sobre GitHub Actions (GHA). Trata-se de uma ferramenta do Github que permite fazer integração contínua de códigos desenvolvidos em quaisquer linguagens de programação, incluindo o R. É possível utilizar o GHA para diversas atividades, como testar um pacote em várias distribuições atualizar dados periodicamente e até mesmo automatizar relatórios com envio de e-mails, tudo em R. Sua palestra complementa a de Beatriz Milz que já faz uma introdução bem completa, portanto ele complementa com um exemplo em live coding!

No sexto e último bloco tivemos a exibição de duas palestras:



Figura 25: Palestra de Eric Ferreira com moderação dos professores Ariel Levy(UFF) e Luciane Alcoforado(AFA)

A palestra de título “ExpDes e ExpDes.pt: pacotes R para análise de experimentos” ministrada por Eric Batista Ferreira, ou professor Eric Barba, da Universidade Federal de Alfenas. A palestra trata sobre a pesquisa científica que está constantemente utilizando-se de experimentos para provar suas hipóteses. Os pacotes Experimental Designs (ExpDes), e sua versão em Português, ExpDes.pt são bibliotecas de funções do software R que propiciam análise de experimentos balanceados segundo modelos de efeitos fixos. Eles permitem fazer a análise de variância, checando a normalidade e homogeneidade de variâncias dos resíduos, comparações múltiplas de tratamentos qualitativos e ajuste de modelos de regressão para fatores quantitativos. Além desse pacotes, em 2019 a família ExpDes aumento com a chegada de uma interface gráfica que facilita o processo de ensino-aprendizagem: o GExpDes. Com esses três pacotes, ganham o estudante, o professor e o analista de dados.



Figura 26: : Palestra de Daniel Falbel com moderação dos professores Ariel Levy(UFF) e Luciane Alcoforado(AFA)

A segunda palestra “Implementation of torch”, ministrada por Daniel Falbel, engenheiro de software na RStudio e sócio da Curso-R. A palestra fala sobre como o torch é implementado e suas motivações e em seguida apresenta o pacote luz que faz com que seja muito fácil ajustar modelos e fazer previsões usando o torch.

V SER 15H DE TRANSMISSÃO REMOTA!

O evento contou com o envolvimento de quase 1500 pessoas, entre voluntários, organizadores, palestrantes e inscritos.

Foram 25 palestrantes do Brasil e exterior que se apresentaram durante os dois dias de evento.

Entre os participantes inscritos havia alunos de graduação e pós-graduação, professores e analistas de diversas instituições como UFF, ENCE, USP, UERJ, UFRJ, UFES, UFPR, UFRN, UNESP, UFV, UNIRIO, AFA, UFMG, UFPE, UFLA entre outras.

O evento contou com o envolvimento de 24 pessoas voluntárias formada por professores e estudantes de oito instituições como Universidade Federal Fluminense, Academia da Força Aérea, Universidade Federal do Estado do Rio de Janeiro, Universidade Estadual do Rio de Janeiro, Escola Nacional de Ciências Estatística, Universidade de São Paulo, Universidade de Santiago de Compostela, Universidade de Aveiro, Colégio Estadual Dr. João Bazet. Foram oferecidos dois minicursos online e gratuito **de Introdução ao R e de Introdução ao Tidyverse**, organizado pela professora Luciane Alcoforado que contou com 353 participantes inscritos.

O evento envolveu 25 palestrantes de diferentes regiões e países como Canadá, Portugal, Espanha, Irlanda, Itália, Reino Unido e do Brasil das seguintes cidades: Araras-SP, Alfenas-MG, Blumenau-SC, Brasília-DF, Duque de Caxias-RJ, Nova Friburgo-RJ, Rio de Janeiro-RJ e São Paulo-SP, totalizando 22 palestras, 11 em cada dia do evento.

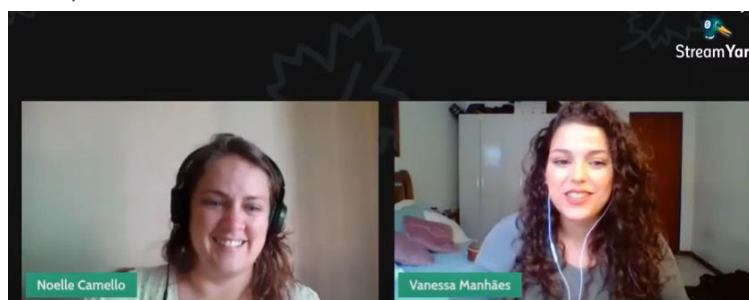
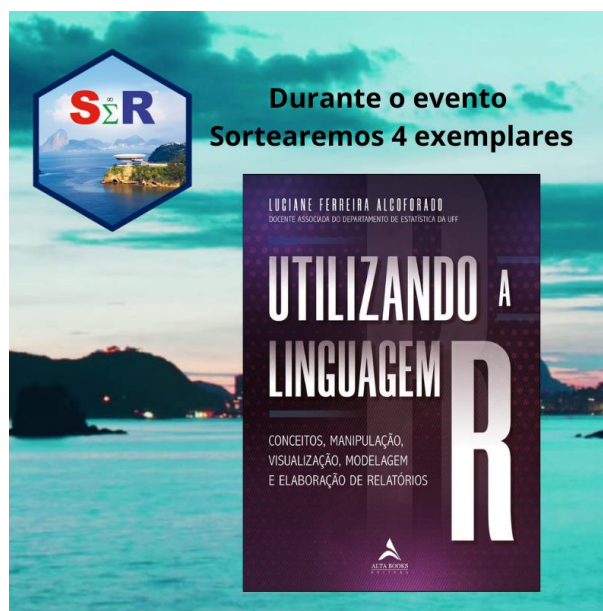


Figura 27: Noelle Camello coordenadora do cerimonial e Vanessa Manhães coordenadora do audio-visual do V SER

Foram exibidos 15h e 17 min de transmissão ao vivo no canal youtube coordenado pela professora de Matemática Vanessa Manhães do Colégio Estadual Dr. João Bazet de Nova Friburgo-RJ contando com a apresentação da mestre de cerimônias e doutoranda em Turismo Noelle Camello da Universidade de Aveiro-PT.



A coordenadora do SER, profa. Luciane Alcoforado recentemente lançou o livro de intitulado **Utilizando a Linguagem R: conceitos, manipulação, visualização, modelagem e elaboração de relatórios** da editora Alta Books que disponibilizou quatro exemplares para sorteio e um **desconto de 35%** em compras do site durante todo o mês de junho (cupom **UFF35**). Os

contemplados com o livro neste sorteio foram Haydée Svab de São Paulo-SP, Guilherme dos Santos de Duque de Caxias-RJ, Thais Maia Cortizo e Vitor Borges Junior do Rio de Janeiro-RJ



Figura 28: Encerramento do V SER

No encerramento do evento estavam presentes os coordenadores gerais e toda a equipe envolvida que agradeceu a presença e o apoio de todos que participaram desta edição. O professor **Ariel Levy** convidou a mestranda do curso de Administração da UFF, Luciana Fernandes a dar seu depoimento sobre o evento. Ela relatou que aprendeu muito com os palestrantes, parabenizou a organização do evento pela escolha da programação e já vai iniciar seu artigo para a revista do SER. O professor **Orlando Longo** conclui dizendo que em 2020 o evento iria acontecer em maio, já contava com o apoio da CAPES mas não houve tempo hábil para vencer as limitações de isolamento social imposta pelo início da pandemia em março de 2020. No início de 2021 a equipe se reuniu e decidiu que iria fazer o V SER, ele confessa que ficou com muito receio pois já havia participado da organização de 14 outros e nunca havia vivido algo semelhante, mas acreditou na força da equipe rumo ao desconhecido pois ninguém possuía a experiência com o formato remoto de evento. Considera que o evento atingiu e superou sua missão de disseminar a linguagem R no Brasil e América Latina, conseguindo manter a qualidade e o sucesso de uma forma remota e segura diante desta pandemia da Covid-19, sem que ninguém precisasse sair de casa. Parabenizou toda a equipe SER, **de A a Z**, que vestiu a camisa e ao final da partida entrou no vestiário com ela molhada mas com o semblante feliz pelo realizado.

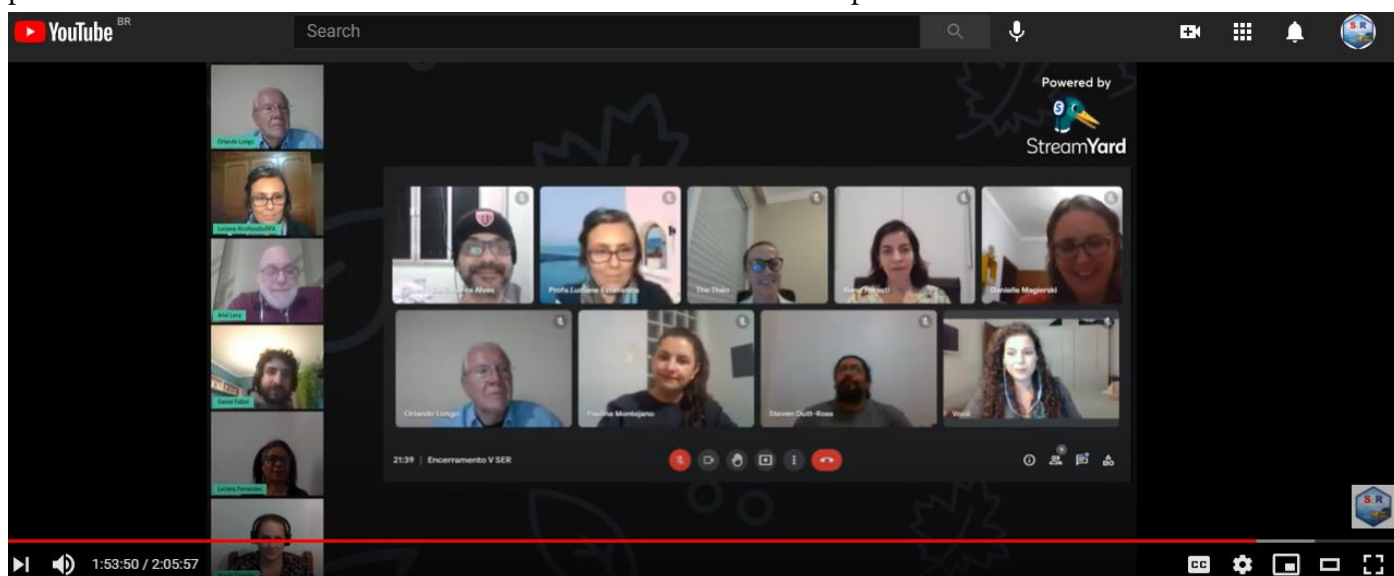


Figura 29: Equipe SER no confraternização do encerramento

Toda a equipe trabalhou de forma sincronizada, destacando especial elogio à mais nova integrante do SER, a professora Vanessa Manhães que tornou possível a produção e transmissão dos vídeos pelo canal youtube. A professora **Luciane Alcoforado** destacou que o evento fortalece muito a comunidade R e acredita que o evento dá oportunidade a muitos pesquisadores e praticantes de análise de dados e deseja que o conhecimento em R saia mais fortalecido, especialmente para quem ainda se inicia no R, que possa se motivar a aprender mais. O professor **Ariel Levy** destaca que temos uma tradição de apoiar o grupo de R-ladies e lembrou que em 2018 a mestre de cerimônia **Noelle Camello** apoiada pela profa. Luciane inaugurou o capítulo de R-ladies Niterói e agora vai ser iniciado um novo capítulo R-ladies de Brasília.

Agradeceu aos novos amigos da AFA trazidos pela mudança institucional da profa. Luciane. Agradeceu a profa Vanessa que trouxe inovações no formato do SER. Terminamos o evento parabenizando a todos que participaram do SER e constatamos que fizemos o MAIOR EVENTO DE R NA AMÉRICA LATINA EM PLENA PANDEMIA, fizemos história e vamos fazer muito mais. MUITO OBRIGADO A TODOS.

“O V SER foi o MAIOR evento de R da América Latina em plena pandemia, fizemos história e vamos fazer muito mais!” Ariel Levy



Nossos sinceros agradecimentos a todos os participantes, palestrantes, apoiadores, voluntários, professores das comissões organizadoras e científicas.

Coordenadores do V SER – Prof. Orlando Celso Longo, Profa. Luciane Ferreira Alcoforado e Prof. Ariel Levy



Redação: Profa. Luciane F. Alcoforado**Comissão Organizadora**

Docentes

Orlando Celso Longo – PPGEC/UFF

Luciane Ferreira Alcoforado – AFA

Ariel Levy – PPGAD/UFF

Manuel Febrero Bande – UN. SANTIAGO COMPOSTELA

Steven Dutt Ross – UNIRIO

Andrea Ugolini – UERJ

Renata Belluzzo Zironi Mori – AFA

Vanessa Manhães Almeida – Colégio Estadual Dr. João Bazet

Discentes

Noelle Santos Camello – doutoranda na Univ. de Aveiro

Cleuson de Oliveira Alves – mestrando na UFF

Marcio – mestrando na UFF

Comissão Científica

Orlando Celso Longo – PPGEC/UFF

Luciane Ferreira Alcoforado – AFA

Ariel Levy – PPGAD/UFF

José Rodrigo de Moraes – IME/UFF

Manuel Febrero Bande – UN. SANTIAGO COMPOSTELA

Steven Dutt Ross – UNIRIO

Andrea Ugolini – UERJ

Renata Belluzzo Zironi Mori – AFA

Maysa Sacramento de Magalhães – ENCE/IBGE

Assed Naked Haddad – UFRJ

Carlos Alberto Pereira Soares – PPGEC/UFF

Equipe SER

Vanessa Manhães Almeida – coordenação áudio-visual

Noelle Santos Camello – coordenação do cerimonial

Orlando Celso Longo, Luciane Alcoforado e Ariel Levy – coordenação geral

Apoio ao evento

Andrea Ugolini - UERJ

Ariel Levy – UFF

Assed Naked Haddad – UFRJ

Beatriz Milz - USP

Carlos Alberto Pereira Soares – PPGEC/UFF

Carlos Alberto Ribeiro Barbosa - AFA

Cleuson de Oliveira Alves - UFF

Danielle Magierski - AFA

Dirley Lemos Vilela - AFA

Emilia Takahashi - AFA

Iliane Jesuína Silva Foresti - AFA

João Paulo Martins dos Santos – AFA

José Rodrigo de Moraes - UFF

Luciane Ferreira Alcoforado - AFA

Manuel Febrero Bande – Un. Santiago de Compostela

Marcio - UFF

Maysa Magalhães - ENCE

Noelle Santos Camello – Un. Aveiro

Orlando Celso Longo - UFF

Paulina Montejano - AFA

Paulo César Rioli Duarte de Souza - AFA

Renata Belluzzo Zironi Mori - AFA

Steven Dutt Ross - UNIRIO

Vanessa Manhães Almeida – Col Dr. João Bazet

Palestrantes no evento

Alessia Paccagnini – UNIVERSITY COLLEGE DUBLIN

Alexandre Sousa da Silva - UNIRIO

Beatriz Milz - USP

Bruna Wundervald – MAYNOOTH UNIVERSITY

Bruno Thiago Tomio - FURB

Daniel Falbel - RSTUDIO

Dean Attali – ATTALITECH

Eduardo Soares Jangutta - UFF

Emmanuelle Rodrigues Nunes – UNIVERSITY OF READING

Eric Batista Ferreira – UNIFAL

Fabio Parla – CENTRAL BANK OF IRELAND

Fernando Almeida Barbalho – TESOURO NACIONAL

Fernando Corrêa – CURSO-R

Fernando Ribeiro - UNIRIO

Gilberto Rodrigues Liska – UFSCAR-ARARAS

Julio Trecenti – CURSO-R

Larissa de Carvalho Alves – ENCE/IBGE

Leonardo Filgueira - DTM

Luis Torgo – DALHOUSE UNIVERSITY

Manuel Febrero Bande – UNIV. SANTIAGO DE COMPOSTELA

Matheus Valentim de Barros – BASE DOS DADOS

Morgany Leite - UFF

Orlando Fonseca Guilarte – PUC-RJ

Paulo Henrique Sales Guimarães - UFLA

Vanessa Manhães Almeida – COLÉGIO DR. JOÃO BAZET

Apoio IME

Professor José Rodrigo de Moraes do departamento de Estatística/UFF

Apoio Escola de Engenharia

Divulgação

Apoio PPGEC

Professores e alunos do Programa de Pós-Graduação

Apoio PPGAD

Professores e alunos do Programa de Pós-Graduação

Apoio Faculdade de Administração

Coordenação do curso de Administração da UFF

Apoio STI

Site do evento

Apoio Datauff

Coordenação do projeto SER, Estatística é com R!

Apoio AFA

Professores da Divisão de Ensino

Apoio ENCE

Coordenação Geral da ENCE

SBMAC

Apoio para emissão dos certificados dos minicursos online

ALTA BOOKS

Livros para sorteio, cupom de desconto

CAPES

Aprovação do projeto no edital de eventos 2020

PAÍSES

Alemanha, Angola, Brasil, Cabo Verde, Canadá, Chile, Equador,

Espanha, França, Irlanda, Itália, Paraguai, Peru, Reino Unido e São

Tomé e Príncipe

ESTATÍSTICAS

Inscritos no evento 1426

Inscritos nos Minicursos Online 353

Inscritos no canal youtube 623

Instagram 968

Facebook 957

Palestrantes 25

Equipe de apoio 24

Horas de transmissão ao vivo: 15h e 17 min

Total de Views nos 2 dias de transmissão 4354

Total de países participantes: 15

Total de Estados do Brasil: 27 (TODOS)

