



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS (PUC GOIÁS)
ESCOLA POLITÉCNICA – MESTRADO EM ENGENHARIA DE
PRODUÇÃO E SISTEMAS (MEPROS)

EDITAL: 2º DESAFIO EM CIÊNCIA DE DADOS

O desafio em ciência de dados tem como objetivo oportunizar a alunos, pesquisadores, profissionais e educadores a apresentação e discussão das mais recentes inovações, tendências e preocupações, bem como desafios práticos encontrados em ciência dos dados. Apresentamos à comunidade da PUC Goiás o projeto do “**2º DESAFIO EM CIÊNCIA DE DADOS**” que será realizado como parte integrante da JORNADA CIENTÍFICA DA ESCOLA POLITÉCNICA DA PUC GOIÁS, 2022.

O QUE É O “DESAFIO EM CIÊNCIA DE DADOS”?

O “Desafio em Ciência de Dados” é uma competição que visa reunir alunos da pós-graduação, graduação, iniciação científica, ensino médio e egressos da PUC Goiás (ou de outras IES) para trocar e compartilhar experiências, na solução de questões de pesquisa, a partir de uma base de dados pública.

O propósito do desafio é avaliar conhecimento prático e habilidades dos competidores na solução de problemas de ciências de dados, com trabalho em equipe e foco na aplicação de Estatística e *Machine Learning*.

SOBRE O TEMA

A base de dados será sobre **Transplantes de Rins no Estado de São Paulo, no período de 2000 a 2017, um problema real de modelagem para previsão do tempo de espera na fila, usando ciência dos dados**. Os participantes do desafio receberão a base de dados de um repositório público e um artigo publicado na *Plos One*, em 2021. No artigo um modelo é apresentado e discutido. Os participantes deverão propor outros modelos e as regras serão informadas em documento apropriado e disponibilizado para os participantes

após cadastro e definição das equipes.

Base de dados do desafio: Transplantes no estado de São Paulo, 2000-2017, SP/OAS/Brasil. Os dados utilizados foram utilizados no artigo: Sapiertein Silva JF, Ferreira GF, Perosa M, Nga HS, de Andrade LGM (2021) *A machine learning prediction model for waiting time to kidney transplant*. PLoS ONE 16(5): e0252069. <https://doi.org/10.1371/journal.pone.0252069>. Os dados estão em repositório público: <https://www.kaggle.com/gustavomodelli/waitlist-kidney-brazil>.

O tema foi escolhido por ser um tema relevante na área da saúde. Aproveitando a oportunidade, no sentido de conscientizar a todos sobre a importância da matéria, teremos uma palestra “Processo de Doação de Órgãos e Tecidos para Transplantes” a ser ministrada pela gerente da secretária da saúde de Goiás, Enf. Katiúscia Christiane Freitas. A palestra também será importante para contextualizar a base de dados e as variáveis a serem tratadas no modelo.

SOBRE AS INSCRIÇÕES E O TRABALHO DAS EQUIPES

As inscrições para o “2º DESAFIO EM CIÊNCIA DE DADOS” podem ser feitas por equipes, tendo de 4 a 5 participantes.

- a) Para participar o aluno deverá efetivar inscrição no evento pelo site da PUC Goiás: www.pucgoias.edu.br, no link da Jornada Científica, da Escola Politécnica da PUC Goiás.
- b) No Desafio as equipes poderão usar o Software R ou Linguagem Python para demonstrarem as suas habilidades com a modelagem de um problema real. O Software R poderá ser obtido em <https://vps.fmvz.usp.br/CRAN/> e o Rstudio em <https://www.rstudio.com/products/rstudio/download/#download> e o Python pode ser utilizado no Google Colab, em <https://colab.research.google.com/> ou Python instalado, a critério da equipe.
- c) As equipes deverão se inscrever e submeter os resultados dentro do calendário oficial do desafio. Todas as equipes receberão a mesma base de dados e a(s) questão(ões) de pesquisa.
- d) As inscrições podem ser realizadas no período de 09/05 (início 09:00h) a 10/05 (até 09:00h), usando o formulário:

<https://forms.office.com/Pages/ResponsePage.aspx?id=Qp8xcwiJiUufjVWM9NX>

[XdjAhTDojHD5Fh1u0LRfyVX9UMjhWTjlGTDhZVVBUNEdTSjdUSUhWRE9GUC4u](https://teams.microsoft.com/l/team/19%3aHcrIGbRxMtaHnfgMURvabS_UfdjPkoLxnMLYskgZ3k1%40thread.tacv2/conversations?groupId=9eb3658f-e5d0-4b9a-afcf-dcfbf5a8d97e&tenantId=73319f42-8908-4b89-9f8d-558cf4d5d776)

A inscrição pode ser feita por uma equipe ou por um participante individual. Os participantes individuais serão apresentados a outros na mesma categoria para formar novas equipes, conforme item e).

- e) As equipes receberão orientações em uma sala na Plataforma *Teams*, dia 09/05/2022, das 14h00min-15h00min e das 17h00-18h00min, com a participação de membros da comissão organizadora. Nesta oportunidade as equipes serão confirmadas e as novas equipes formadas. Para as atividades do desafio serão criados canais na sala do desafio na Plataforma *Teams*, com os apelidos das equipes, para que ocorram as reuniões de cada equipe até o dia da apresentação.

Link da sala do 2º DESAFIO EM CIÊNCIAS DE DADOS (canal geral)

https://teams.microsoft.com/l/team/19%3aHcrIGbRxMtaHnfgMURvabS_UfdjPkoLxnMLYskgZ3k1%40thread.tacv2/conversations?groupId=9eb3658f-e5d0-4b9a-afcf-dcfbf5a8d97e&tenantId=73319f42-8908-4b89-9f8d-558cf4d5d776

- f) Os modelos desenvolvidos devem ser apresentados no dia 13/05, das 19h00min às 22h00min, em sala específica para o desafio de dados na Plataforma *Teams* (https://teams.microsoft.com/l/team/19%3aHcrIGbRxMtaHnfgMURvabS_UfdjPkoLxnMLYskgZ3k1%40thread.tacv2/conversations?groupId=9eb3658f-e5d0-4b9a-afcf-dcfbf5a8d97e&tenantId=73319f42-8908-4b89-9f8d-558cf4d5d776), e serão analisados e pontuados objetivamente, segundo:

1. Identificar e visualizar os dados para obter informações úteis
2. Preparar os dados
3. Selecionar e aplicar um modelo
4. Gráficos e análises para aperfeiçoar o modelo selecionado
5. Apresentar a solução comprovando com métricas adequadas

Os projetos serão avaliados por uma banca constituída por professores do Mestrado em Engenharia de Produção e Sistemas. A banca assistirá as apresentações e fará as devidas considerações e após, em reunião deliberativa, as pontuações serão emitidas e as equipes classificadas.

SOBRE A PREMIAÇÃO.

- a) Todos os participantes receberão certificados de participação no “**2º DESAFIO EM CIÊNCIA DOS DADOS**”, com 4 horas atividades.
- b) Todos os participantes das equipes classificadas nos três primeiros lugares receberão medalhas.
- c) Premiações para as equipes patrocinadas pela FLEXIM do BRASIL (<https://www.flexsim.com/>)

1º Lugar, cedido pela FLEXSIM do BRASIL, cursos de Simulação I (32h), no valor de R\$ 4550,00 por participante e Modelagem em Lógica com *Process Flow* (20h), no valor de R\$3500,00 por participante, online e ao vivo, para todos os membros da equipe, com certificado. Plataforma EAD: <https://www.eadsimulacao.com.br/>. Os alunos receberão os *vouchers* para acessarem os cursos.

2º Lugar, cedido pela FLEXSIM do BRASIL curso de Simulação I (32h), no valor de R\$4550 por participante, para todos os membros da equipe, com certificado. Plataforma EAD: <https://www.eadsimulacao.com.br/>. Os alunos receberão os *vouchers* para acessarem os cursos.

3º Lugar, cedido pela FLEXSIM do BRASIL curso de Introdução a Simulação com Flexsim (5h), no valor de R\$ 150,00 por aluno, para todos os membros da equipe, com certificado. Os alunos receberão os *vouchers* para acessarem os cursos. Plataforma EAD:<https://www.eadsimulacao.com.br/>

- d) A premiação será no dia 14/05/2022, 09h00min (local será comunicado no dia das apresentações pelas equipes)
- e) A premiação será realizada pela direção da Escola Politécnica, Coordenador do MEPROS e comissão organizadora.

CONTATOS

Havendo dúvida, sugestão ou reclamação quanto ao Edital de Abertura do “**2º DESAFIO EM CIÊNCIAS DOS DADOS**” entrar em contato com a organização através do endereço eletrônico mjdantas@pucgoias.edu.br ou jelmo@pucgoias.edu.br

OBSERVAÇÕES:

1. Os casos omissos serão discutidos pela comissão organizadora.
2. Este edital poderá sofrer alterações que serão devidamente informadas pelos meios de divulgação.

Comissão Organizadora do “**2º Desafio de Ciência dos Dados**”

Profa. Maria José Pereira Dantas

Prof. José Elmo de Menezes

Coordenação do MEPROS

Prof. Marcos Lajovic Carneiro

Direção da Escola Politécnica

Profa. Mírian Sandra Rosa Gusmão

Goiânia, 24 de abril de 2022.