

ERNADES SOBREIRA OLIVEIRA JUNIOR  
BRUNO ORMONDE LIMA DE OLIVEIRA  
WILKINSON LOPES LÁZARO

# MODELAGEM E EXPERIMENTAÇÃO EM BIOLOGIA

um guia prático

COMO OBTENHO, CALCULO  
E APRESENTO MEUS DADOS?



**UNEMAT**  
Universidade do Estado de Mato Grosso  
Carlos Alberto Reyes Maitonado

**EDITORA**  
UNEMAT

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

O48m

Modelagem e experimentação em Biologia: um guia prático / Ernandes Sobreira Oliveira Junior; Bruno Ormonde Lima de Oliveira; Wilkinson Lopes Lázaro. – Cáceres: Editora UNEMAT, 2025. 95 p. II.

ISBN: 978-85-7911-288-1

DOI: 10.30681/978-85-7911-288-1

1. Biologia. 2. Modelagem. 3. Experimentação. I. Modelagem e experimentação em biologia. II. Ernandes Sobreira Oliveira Junior.

CDD: 57.08(036)

Bibliotecário WALTER CLAYTON DE OLIVEIRA CRB 1/2049

ERNANDES SOBREIRA OLIVEIRA JUNIOR  
BRUNO ORMONDE LIMA DE OLIVEIRA  
WILKINSON LOPES LÁZARO

# MODELAGEM E EXPERIMENTAÇÃO EM BIOLOGIA

um guia prático

**UNEMAT**  
Universidade do Estado de Mato Grosso  
Carlos Alberto Reyes Maldonado

  
EDITORA  
UNEMAT

Cáceres - MT

2025

## **CONSELHO EDITORIAL**

Portaria nº 1629/2023

## **PRESIDENTE**

Maristela Cury Sarian

## **TITULARES**

**Josemir Almeida Barros**

*Universidade Federal de Rondônia - Unir*

**Lais Braga Caneppele**

*Universidade do Estado de Mato Grosso - Unemat*

**Fabrcio Schwanz da Silva**

*Universidade Federal do Paraná - UFPR*

**Gustavo Rodrigues Canale**

*Universidade Federal de Mato Grosso - UFMT*

**Greciely Cristina da Costa**

*Universidade Estadual de Campinas - Unicamp*

**Edson Pereira Barbosa**

*Universidade Federal de Mato Grosso - UFMT*

**Rodolfo Benedito Zattar da Silva**

*Universidade Federal de Mato Grosso - UFMT*

**Cácia Régia de Paula**

*Universidade Federal de Jataí - UFJ*

**Nilce Vieira Campos Ferreira**

*Universidade Federal de Mato Grosso - UFMT*

**Marcos Antonio de Menezes**

*Universidade Federal de Jataí - UFJ*

**Flávio Bezerra Barros**

*Universidade Federal do Pará - UFPA*

**Luanna Tomaz de Souza**

*Universidade Federal do Pará - UFPA*

## **SUPLENTES**

**Judite de Azevedo do Carmo**

*Universidade do Estado de Mato Grosso - Unemat*

**Rose Kelly dos Santos Martinez Fernandes**

*Universidade do Estado de Mato Grosso - Unemat*

**Maria Aparecida Pereira Pierangeli**

*Universidade do Estado de Mato Grosso - Unemat*

**Célia Regina Araújo Soares**

*Universidade do Estado de Mato Grosso - Unemat*

**Nilce Maria da Silva**

*Universidade do Estado de Mato Grosso - Unemat*

**Rebeca Caitano Moreira**

*Universidade do Estado de Mato Grosso - Unemat*

**Jussara de Araújo Gonçalves**

*Universidade do Estado de Mato Grosso - Unemat*

**Patrícia Santos de Oliveira**

*Universidade Federal de Viçosa - UFV*

**PRODUÇÃO EDITORIAL**  
**EDITORA UNEMAT 2025**

Copyright © dos autores 2025.

A reprodução não autorizada desta publicação,  
por qualquer meio, seja total ou parcial,  
constitui violação da Lei nº 9.610/98.

Esta obra foi submetida à avaliação  
e revisada por pares.

---

**Reitora:** Vera Lucia da Rocha Maquêa

**Vice-reitor:** Alexandre Gonçalves Porto

**Assessora de Gestão da Editora e das Bibliotecas:** Maristela Cury Sarian

**Imagens da capa:** Freepik AI

**Capa:** Potira Manoela de Moraes

**Diagramação:** Potira Manoela de Moraes

**Revisão:** Mara Lúcia Martins Rodrigues

---

# SUMÁRIO

APRESENTAÇÃO..... 7

## SEÇÃO 1

AFINAL, O QUE SÃO DADOS? ..... 9

## SEÇÃO 2

QUE PROBLEMA EU QUERO RESOLVER? ..... 13

## SEÇÃO 3

MONTANDO MINHA PESQUISA COM RÉPLICAS..... 35

## SEÇÃO 4

INTERPRETAÇÕES ESTATÍSTICAS DO TESTE..... 52

## SEÇÃO 5

E A MODELAGEM, O QUE É? ..... 64

REFERÊNCIAS..... 92

SOBRE OS AUTORES ..... 93

# APRESENTAÇÃO

Este trabalho consiste na apresentação de conteúdo importante para ser trabalhado no ensino de biologia, na área de estatística, principalmente como um guia inicial e simples para as aulas e pesquisas. A construção do texto foi baseada nos desafios que os alunos de graduação apresentam quando fazem a disciplina de Modelagem e Desenho Experimental. E, assim, tentar minimizar os efeitos negativos que a estatística, às vezes, apresenta. Tentamos abordar em uma linguagem simples e com passo a passo para que possam executar seus exercícios.

Ao longo do texto há vários exemplos que podem ser utilizados para a construção do TCC, projetos de bolsa, além de questionamentos individuais. Também abordamos questões que precisam ser respondidas e que há poucas pesquisas na área. É importantíssimo salientar que apresentamos a estatística básica para o desenho experimental, principalmente ressaltando os modelos lineares com distribuição normal. Há outros modelos mais adequados para outras distribuição (Binomial e Poisson - GLM). Além disso, apresentamos aqui o uso do Excel como uma ferramenta para a maioria das análises, mas há outras ferramentas disponíveis.

Esperamos que usem o livro como uma base para analisar dados no Excel, uma das ferramentas mais simples e robustas. Não é necessário seguir página a página se o leitor possui algum embasamento na área. Como o objetivo deste trabalho é apresentar o passo a passo (um guia) para a análise e interpretação de dados estatísticos, o leitor pode ir direto onde precisa de informação, desde que os dados apresentem normalidade.

O modo como um cientista desenha seu experimento ou interpreta o mundo molda diretamente a forma como os dados serão coletados, organizados e apresentados. **O mapa apresentado na capa do livro ilustra simbolicamente a subjetividade na construção de uma narrativa visual ou estatística.** Assim como o mapa desafia a visão tradicional do mundo, a estatística exige que questionemos as perspectivas convencionais e adotemos novas formas de análise que reflitam a complexidade dos fenômenos.

A filosofia estatística nos ensina que os dados não são uma realidade absoluta, mas uma interpretação mediada por escolhas metodológicas, culturais e até visuais. A representação de um mapa como este, por exemplo, desafia normas estabelecidas e nos obriga a reconsiderar o papel da posição e da perspectiva ao comunicar informações.

## AFINAL, O QUE SÃO DADOS?

Dados são informações, quantitativas ou qualitativas, que podem ser coletadas, registradas e analisadas para se caracterizar um fenômeno, obter pistas sobre este fenômeno, tirar conclusões sobre sua natureza e tomar decisões informadas.

Em Ciências Biológicas, os dados podem incluir observações, medidas e informações relevantes sobre organismos, processos biológicos, populações e ecossistemas. Eles desempenham um papel fundamental na pesquisa e na obtenção de evidências para apoiar teorias e hipóteses.

Para outras informações sobre estatística, você também pode utilizar obras como: ALVES, AZEVEDO e COSTA.

ALVES, Fábio José da Costa. **Estatística**. Belém: Universidade do Estado do Pará – UEPA, 2008. Disponível em: <http://educapes.capes.gov.br/handle/capes/747881>. Acesso em: 20 mar.2024.

AZEVEDO, Paulo Roberto Medeiros de. **Introdução à estatística** [recurso eletrônico]. 3. ed. Natal: EDUFN, 2016. Disponível em: <https://repositorio.ufrn.br/items/277937e1-5de9-4e1e-81b6-a06bcadc7d90>. Acesso em: 20 mar.2024.

COSTA, Fabrício Martins da. **Estatística**. Belém: UEPA.



## OS DADOS PODEM SER CLASSIFICADOS EM DUAS CATEGORIAS PRINCIPAIS

1. **DADOS QUALITATIVOS:** Também conhecidos como dados categóricos, esses tipos de dados descrevem qualidades ou características e são, frequentemente, expressos em palavras ou categorias. Exemplos de dados qualitativos na biologia incluem a cor do pelo de um animal, o tipo de habitat em que uma espécie vive ou a classificação taxonômica de um organismo. Há, ainda, os dados ordinais, que indicam uma ordem de organização.
2. **DADOS QUANTITATIVOS:** Esses dados representam quantidades e processos, sendo, frequentemente, expressos em números. Os dados quantitativos podem ser discretos (valores inteiros) ou contínuos (valores em uma faixa contínua). Exemplos de dados quantitativos na biologia incluem a contagem de indivíduos em uma população, as medidas de temperatura em um ambiente, ou a concentração de uma substância química em uma amostra biológica.

Veja estas descrições também em outras obras como:

BUSSAB, Wilton de Oliveira; MORETTIN, Pedro Alberto. **Estatística básica**. 6. ed. São Paulo: Saraiva, 2017.

SANTIAGO, Genário Sobreira; PAIVA, Rui Eduardo Brasileiro. **Bioestatística** [recurso eletrônico]. 2. ed. Fortaleza: EdUECE, 2015. (Ciências Biológicas). Disponível em: [https://educapes.capes.gov.br/bitstream/capes/431710/2/Livro\\_Bioestatistica.PDF](https://educapes.capes.gov.br/bitstream/capes/431710/2/Livro_Bioestatistica.PDF). Acesso em: 6 dez. 2024.



## E COMO EU FAÇO PARA OBTER DADOS?

Antes de qualquer coisa, há algumas importantes questões a serem abordadas. A partir de agora, você deverá refletir muito. Um dado é extremamente importante, mas há muitos passos de chegar a obter os dados.

Toda pesquisa, para gerar dados, é advinda de um problema.

Um problema que alguém precisa resolver. Por exemplo, querer saber se mulheres vivem mais do que homens é um problema, pois pode gerar políticas públicas para o atendimento mais eficaz a quem viver menos.

Ou então, o fato de observar que não há tantas abelhas, hoje em dia, quanto havia há dez anos, é também uma preocupação. Pode haver algo que esteja reduzindo a população de abelhas?



Veja no MapBiomias as possibilidades de obter dados  
<https://brasil.mapbiomas.org/>

## QUE PROBLEMA EU QUERO RESOLVER?

A preocupação com estes problemas pode começar em uma disciplina, na universidade, em uma conversa com colegas ou, até mesmo, em uma conversa de bar.

O *start* te deixará inquieto, e esta inquietação é a alma de cientista, aflorando em sua pele. Cientistas querem saber das coisas. Querem saber para poder auxiliar a resolver o problema.

### AQUI SEGUEM ALGUNS PROBLEMAS QUE PODEM SER RESOLVIDOS ATRAVÉS DA OBTENÇÃO DE DADOS

LIBERAÇÃO DE GASES  
DE EFEITO ESTUFA

FALTA DE COMITÊS  
DE BACIAS HIDROGRÁFICAS

POLUIÇÃO DAS ÁGUAS

FALTA DE INTEGRAÇÃO  
ENTRE ATORES SOCIAIS

REDUÇÃO DA BIODIVERSIDADE

MUDANÇAS CLIMÁTICAS

TECNOLOGIAS  
ALTERNATIVAS EM ESCOLAS

AUMENTO DE QUEIMADAS

INTRODUÇÃO DE ESPÉCIES EXÓTICAS

FALTA DE ESTÍMULO À BIOECONOMIA

AUMENTO DO DESMATAMENTO

AUMENTO DE AGROTÓXICOS

DOENÇAS EMERGENTES

Outros exemplos de problemas que podem ser resolvidos através da análise de dados

**Conservação da Biodiversidade:**

a perda de biodiversidade devido à degradação do *habitat*, poluição, mudanças climáticas e outras atividades humanas é um problema crítico.

O estudo de doenças infecciosas, câncer, doenças genéticas e outras condições de saúde é uma área central da biologia. Encontrar tratamentos eficazes e estratégias de prevenção é de importância primordial.

**Mudanças Climáticas:** a biologia está diretamente ligada às mudanças climáticas, pois estuda os impactos dessas mudanças nos ecossistemas e nas espécies. A compreensão das adaptações das plantas e animais às mudanças climáticas é crucial.

**Segurança Alimentar:** com o aumento da população global, a biologia desempenha um papel importante na pesquisa agrícola, no desenvolvimento de culturas resistentes a pragas e condições climáticas adversas, bem como na produção de alimentos de forma sustentável.

**Terapia Genética e Edição de Genes:** a biologia molecular tem avançado na edição de genes e na terapia genética para tratar doenças genéticas. No entanto, isso levanta questões éticas e de segurança que precisam ser abordadas.

**Desenvolvimento de Medicamentos:** a descoberta e o desenvolvimento de novos medicamentos para tratar uma variedade de condições médicas é uma área crucial da pesquisa biológica.

**Ecologia e Sustentabilidade:** o entendimento dos ecossistemas e das interações entre os seres vivos e o ambiente é essencial para a conservação e a gestão sustentável dos recursos naturais.

**Biologia da Evolução:** a biologia evolutiva lida com questões fundamentais sobre a origem e a diversificação da vida. A pesquisa nessa área busca entender como a evolução moldou a vida na Terra.

**Biologia Celular e Molecular:** compreender os processos celulares e moleculares é fundamental para inúmeras aplicações, desde a medicina até a biotecnologia.

**Compreensão do Cérebro e do Comportamento:** a neurobiologia é uma área em crescimento, visando entender o funcionamento do cérebro e o comportamento humano.

**Tecnologia Ambiental:** a biologia também desempenha um papel importante no desenvolvimento de tecnologias ambientais, como biorremediação, para lidar com a poluição e outros problemas ambientais.

FALTA DE COMITÊS  
DE BACIAS  
HIDROGRÁFICAS

DOENÇAS  
EMERGENTES

FALTA DE  
INTEGRAÇÃO ENTRE  
ATORES SOCIAIS  
REDUÇÃO  
DA BIODIVERSIDADE

INTRODUÇÃO DE  
ESPÉCIES EXÓTICAS

É CLARO QUE AQUI FORAM APRESENTADOS APENAS  
ALGUNS DOS MILHARES DE PROBLEMAS EXISTENTES!

AGORA PENSE



QUE PROBLEMA EU QUERO RESOLVER?  
LEMBRE-SE, O PROBLEMA DEVE SER DIRETO,  
CLARO E POSSÍVEL DE SER RESOLVIDO!

CONTINUE PENSANDO...



ESSA PERGUNTA NÃO É FÁCIL, MAS É SEU PONTAPÉ INICIAL!

CONTINUE PENSANDO...

## APÓS A DEFINIÇÃO DO PROBLEMA, SURGE A ELABORAÇÃO DA PERGUNTA QUE IRÁ RESPONDER

A elaboração da pergunta é muito importante. A partir da sua definição é que você irá saber se sua resposta estará correta ou não. Imagine chegar a uma conclusão pulando esta etapa – nunca saberá se a resposta será o suficiente para o que se está pesquisando.

A pergunta é muito importante. Ela é uma das peças fundamentais de sua pesquisa. Não deixe de elaborar com calma.

A pergunta mudará muitas vezes e, isso é importante, pois através disso você verá o quanto aprendeu.

Chegará em um ponto que ela não deverá mudar mais, e se estabilizará. Não se preocupe, isso vai acontecer!



## QUAL É A PERGUNTA QUE DIRECIONA A SUA PESQUISA?

Há vários tipos de perguntas, aqui seguem algumas que podem te auxiliar a desenvolver a sua:

A perda da biodiversidade é um problema grave, mas quais são as espécies com maior risco de desaparecer em ambientes que sofrem com o avanço da agropecuária?

A introdução de espécies exóticas é um dos maiores problemas ambientais, então, qual é o impacto da introdução do tucunaré no rio Paraguai (um rio onde a espécie nunca esteve presente)?

As mudanças climáticas estão afetando as espécies de plantas no planeta como um todo, mas quais são as espécies utilizadas por comunidades tradicionais que tendem a desaparecer?

As escolas públicas possuem um problema de uso de equipamentos robustos para a análise ambiental que possam servir como instrumentos de cálculos em sala de aula. Ou seja, quais são os equipamentos de baixo custo que podem ser produzidos pelos próprios alunos para análises climáticas?



## VEJA QUE, JUNTAMENTE COM A PERGUNTA, SEMPRE ESTÁ APRESENTADO O PROBLEMA!

Isto facilita a construção da pergunta.

Lembre-se que as perguntas devem sempre ser POSSÍVEIS de serem respondidas através da obtenção de dados.

Veja também que as perguntas devem vir associadas aos locais de pesquisa, ou ao grupo a ser investigado.

**UMA PERGUNTA NÃO É FÁCIL DE SER ELABORADA. PENSE, REFLITA  
E MUDE SE FOR NECESSÁRIO, QUANTAS VEZES FOR PRECISO!**



SUA PERGUNTA DEVE SER CLARA,  
OBJETIVA E POSSÍVEL DE SER RESPONDIDA!



## O PROJETO DE PESQUISA SE COMPLETA EM PROBLEMA – PERGUNTA – MÉTODO...

O método é de extrema importância em qualquer pesquisa científica por várias razões:

1. **RIGOR CIENTÍFICO:** O método científico é a estrutura que garante a realização de pesquisas de forma sistemática e rigorosa. Ele ajuda a evitar viés, erros e interpretações subjetivas, tornando os resultados confiáveis.
2. **OBJETIVIDADE:** O método científico promove a objetividade na coleta e interpretação de dados. Ele baseia em evidências empíricas e fatos observáveis, em oposição a suposições ou opiniões pessoais.
3. **REPRODUTIBILIDADE:** Uma pesquisa científica bem conduzida deve ser replicável, ou seja, outros pesquisadores devem poder repetir o estudo e obter resultados semelhantes. Isso ajuda a verificar a validade das descobertas.
4. **VALIDAÇÃO DE HIPÓTESES:** O método científico permite que os pesquisadores testem hipóteses de forma sistemática e controlada. Ele fornece estrutura para projetar experimentos que podem confirmar ou refutar uma hipótese.
5. **PROGRESSO CIENTÍFICO:** O método científico é a base do progresso científico. Ele permite que a pesquisa se baseie em descobertas anteriores, expandindo o conhecimento em um campo específico.

6. **TOMADA DE DECISÕES BASEADA EM EVIDÊNCIAS:** A pesquisa científica fornece evidências sólidas que podem ser usadas para embasar decisões políticas, médicas, ambientais e em muitos outros campos. A confiabilidade dos métodos usados é crucial para tomar decisões informadas.
7. **CONTROLE DE VARIÁVEIS:** O método científico permite que os pesquisadores controlem variáveis para isolar o efeito de uma variável específica, o que é fundamental para estabelecer relações de causa e efeito.
8. **EFICIÊNCIA DE RECURSOS:** Um método bem definido ajuda a otimizar o uso de recursos, como tempo e dinheiro, na condução da pesquisa, garantindo que os recursos sejam direcionados de maneira eficaz.
9. **APRIMORAMENTO DE TEORIAS E MODELOS:** O método científico permite a avaliação crítica de teorias e modelos existentes e o desenvolvimento de novos conceitos e explicações para fenômenos naturais.
10. **MINIMIZAÇÃO DE VIÉS:** O método científico ajuda a minimizar preconceitos e tendências, promovendo a imparcialidade e a integridade na pesquisa.

SEU MÉTODO SERÁ BASEADO, PROVAVELMENTE,  
EM PESQUISAS CIENTÍFICAS JÁ REALIZADAS.

POR ISSO, É FUNDAMENTAL LER ARTIGOS  
CIENTÍFICOS SOBRE O TEMA ABORDADO.



O método varia bastante conforme o problema a ser analisado. Cada método é particular, mas ele deve ser bem elaborado, claro e possível de ser replicado por outros pesquisadores.

Para a boa elaboração do método imagine um passo a passo. O que vou fazer após sair de casa para coletar dados para minha pesquisa?

- Eu vou de quê? Carro? Moto? A pé?
- Quem vai comigo?
- Preciso levar que tipo de equipamento?
- Todos os equipamentos estão calibrados?
- Todos os equipamentos possuem pilhas carregadas?
- Quantas horas vou ficar coletando?
- Qual é o grau de dificuldade de se coletar neste local?
- As pessoas com quem vou falar, estão me esperando?  
Estou levando gravador?



## SE A COLETA FOR VIA INTERNET, PENSE EM ALGO COMO

- Quais são os *websites* em que irei baixar meus dados?
- Estes sites são de quais agências? São confiáveis? Que tipo de dados vou baixar?
- Qual é a extensão dos arquivos?
- Assim que baixar os dados, como eles devem estar organizados?
- Os dados possuem um “leia-me” para facilitar a interpretação?





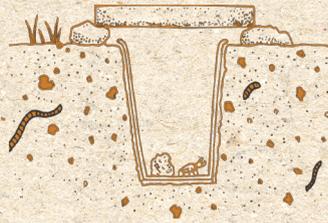
Quais são os métodos no laboratório que devo utilizar?

Existem reagentes suficientes?

Qual é a quantidade mínima de material necessário para as análises?

Quais são os métodos utilizados para a coleta em campo? Quantas armadilhas devo ter no mínimo?

Quanto tempo levo para dispor cada uma? Qual é a distância ideal entre elas? Qual é a profundidade/sombra/exposição/tempo/hora do dia que devo dispor as armadilhas?



Qual é o material necessário para fazer um tipo de observação? Quantas observações são necessárias? Qual é a distância mínima para se manter?

Dentre tantos questionamentos leve em consideração o **local** a ser estudado – se há **sazonalidade**, se há necessidade de pesquisar de dia e de noite, ou só em um período. **O tempo de coleta** de dados é importantíssimo para ser considerado, bem como a necessidade de **réplicas**.

## POR QUE O LOCAL QUE ESCOLHI PARA A PESQUISA É IMPORTANTE?

### POR QUE ESCOLHI ESTA VARIÁVEL?

Pense o seguinte: o que o seu local de estudo tem de tão importante que merece ser estudado. Por exemplo: Você quer saber a quantidade de carbono no solo de praças urbanas, pois há uma indicação de que são importantes fontes de estoque de carbono urbano.

Qual é a praça que devo investigar? Todas? Se eu for naquela mais arborizada estarei influenciando, tendenciando meus dados, e se eu for em uma praça com poucas árvores, também poderei tendenciar?

Escolher uma praça – meio termo – não resolverá meu problema. Tenho que ter a certeza de que aquilo que estou coletando é representativo da realidade.

O mesmo sentido deve ser considerado em relação ao tempo! Muitos pesquisadores investigam um período de tempo, mas por quê? Qual é o melhor período para se analisar um problema? Por exemplo, se você quiser investigar as queimadas na Amazônia, poderá pegar dados via *websites*.

Se o problema está na necessidade de se compreender a tendência das queimadas, então, se você coletar os dados dos últimos 5 anos serão suficientes? Se forem dos últimos 10 anos serão suficientes? Se for dos últimos 15 anos? Mas, qualquer um poderá te questionar, por que não 11 ao invés de 10? Ou então, por que não 9 ao invés de 10? O período de tempo analisado deve ser justificado.

É sempre importante levar em consideração fatores ocorrentes neste período, como a mudança política de um local, o surgimento de novas leis ou, então, a implementação de algum programa político, etc. Anos são somente números, que podem ser substituídos por qualquer nome.

Caso você queira determinar a presença de um agrotóxico presente na água superficial, deve ser levado em consideração que tipo de plantação está sendo utilizado no seu local de estudo. Além disso, deve ser considerada a análise prévia dos tipos de agrotóxicos mais comercializados em sua região de estudo.

Pois, você poderá chegar a uma conclusão de que não há agrotóxico na água, se não analisar o tipo correto, ou seja, você analisará algo que não está presente na região.



## QUANTAS AMOSTRAS SÃO NECESSÁRIAS?

- Esta pergunta é bastante difícil de ser respondida. O número de amostras deve estar de acordo com o seu objeto de estudo. Cada objeto de estudo deve ter um mínimo de amostras para se gerar resultados robustos.
- Entretanto, todas as amostras devem possuir réplicas. As réplicas são parte importante da pesquisa, pois é através delas que podemos confiar nos dados coletados, principalmente, quando o desvio padrão das réplicas ficam pequenos (próximos à média).

## LEMBRANDO ALGUNS PONTOS IMPORTANTES

### POPULAÇÃO DE DADOS

A população de dados refere-se ao conjunto completo de todos os elementos ou observações que são de interesse em um estudo estatístico ou pesquisa.

Esta população pode ser muito grande e, em alguns casos, impossível de ser completamente observada ou medida.

Por exemplo, a população de todos os seres humanos no mundo, a população de árvores em uma floresta ou a população de bactérias em um ambiente.

Em muitos casos, é impraticável ou muito caro coletar dados de toda a população, e é aí que as amostras desempenham um papel importante.

## AMOSTRA

Uma amostra é um subconjunto representativo da população de dados.

A amostragem envolve selecionar um grupo menor de observações ou elementos a partir da população com o objetivo de estimar características ou parâmetros da população maior.

A escolha da amostra deve ser aleatória e representativa, de modo a evitar qualquer viés nos resultados. As amostras fornecem uma maneira prática e econômica de fazer inferências sobre a população maior com base em um estudo de tamanho gerenciável.

## RÉPLICA

Réplicas são múltiplas observações ou medidas obtidas a partir da mesma unidade experimental ou elemento da população.

Réplicas são usadas para medir a variabilidade intrínseca dentro do mesmo elemento da população.

Por exemplo, em um experimento com plantas, você pode medir a altura de vários indivíduos da mesma espécie, sendo cada medida uma réplica. Isso ajuda a avaliar a consistência das observações e a determinar a variabilidade natural dentro desse grupo de elementos.

## UM POUCO MAIS SOBRE RÉPLICA

Uma réplica, no contexto de experimentos científicos e pesquisas, é uma repetição do experimento ou estudo original, realizada sob condições semelhantes, a fim de confirmar a validade e a confiabilidade dos resultados obtidos.

## PARA QUE SERVEM AS RÉPLICAS?

- Realizar réplicas ajuda a confirmar se os resultados obtidos no experimento original eram consistentes, e não eram meramente resultados aleatórios ou anomalias.
- Em experimentos científicos, erros aleatórios podem ocorrer devido a flutuações naturais ou variações nos procedimentos de medição. Realizar réplicas ajuda a reduzir a influência desses erros, permitindo que padrões consistentes se destaquem.
- Réplicas fornecem uma estimativa da variabilidade dos dados. Isso é importante para calcular a precisão e a incerteza associadas aos resultados.

Para referências sobre experimentos, acesse também:

PORTELLA, A. C. F.; NASCIMENTO, I. R. do.; ALVES, A. F.; SCHEIDT, G. N. **Estatística básica para os cursos de Ciências Exatas e Tecnológicas**. Palmas: EDUFT, 2015. Disponível em: <https://historia.uft.edu.br/index.php/eduft/catalogo/obras-academicas/15619-estatistica-basica-para-os-cursos-de-ciencias-exatas-e-tecnologicas>. Acesso em: 20 mar.2024.

SAMPAIO, Nilo Antônio de Souza; ASSUMPTÃO, Alzira Ramalho Pinheiro de; FONSECA, Bernardo Bastos da. **Estatística descritiva**. Belo Horizonte: Poisson, 2018. Disponível em: [https://www.poisson.com.br/livros/estatistica/volume1/Estatistica\\_Descritiva.pdf](https://www.poisson.com.br/livros/estatistica/volume1/Estatistica_Descritiva.pdf). Acesso em: 20 mar. 2024.

SUDARSANAM, N; KANNU, B; FREY, D. (2019). Optimal replicates for designed experiments under the online framework. **Research in Engineering Design**, 30, p. 363 - 379. <https://doi.org/10.1007/s00163-019-00311-x>. Acesso em: 20 mar. 2024.

- Réplicas aumentam a validade externa dos resultados, permitindo que os pesquisadores generalizem suas conclusões para uma população maior ou para uma variedade de condições.



- Réplicas podem ajudar a identificar valores atípicos ou “outliers” que podem distorcer os resultados. Quando esses valores são consistentes em várias réplicas, isso indica uma descoberta válida. No entanto, se forem inconsistentes, podem ser considerados anomalias.
- Réplicas contribuem para a validação de teorias e modelos, pois os resultados consistentes em várias repetições fornecem evidências sólidas de que a teoria é aplicável em várias situações.
- Réplicas podem ajudar a controlar o viés de pesquisa, uma vez que várias repetições podem revelar se um único experimento foi afetado por condições não controladas.
- A reprodutibilidade é uma pedra angular da pesquisa científica. Réplicas são essenciais para garantir que outros pesquisadores possam repetir o estudo e obter resultados semelhantes, validando assim as descobertas originais.



## O QUE É A MÉDIA?

- A média é uma medida de tendência central que representa o valor médio de um conjunto de dados. Ou seja, a representação da média é dada pelo cálculo de todos os seus valores em um único, por isso tende ao centro.
- Para calculá-la, você soma todos os valores dos dados e divide pelo número de observações. A fórmula da média é: Média = Soma de todos os valores / número de observações.
- Em biologia, a média pode ser usada para calcular coisas como a média de altura de uma população de árvores, a média de tamanho de uma amostra de organismos, a média de concentração de uma substância em um conjunto de amostras, etc.
- A média é útil para resumir dados e obter uma ideia geral do valor típico em um conjunto de observações.



## O QUE É O DESVIO PADRÃO?

O desvio padrão mede a dispersão ou variabilidade dos dados em relação a média. Ele indica quão afastados os valores estão da média.

Lembra da média? Então, ela reduz os seus dados em um só, né? Ela centraliza seus valores. Já o desvio padrão mostra a variabilidade dos seus dados, ou seja, o quanto seus dados (réplicas ou outros) variam em relação à média.

Quanto maior o desvio padrão, maior a dispersão dos dados; quanto menor o desvio padrão, menor a dispersão.

Para calculá-lo, você segue uma série de etapas.

$$\text{A FÓRMULA DO DESVIO PADRÃO POPULACIONAL É: DESVIO PADRÃO} = \sqrt{\frac{\text{SOMA DOS QUADRADOS DAS DIFERENÇAS ENTRE CADA VALOR E A MÉDIA}}{\text{NÚMERO DE OBSERVAÇÕES}}}$$

O desvio padrão é importante na biologia quando se deseja entender a variação em dados biológicos, como a variação no tamanho de organismos em uma população, a variação nas concentrações de substâncias em diferentes amostras, a variação genética em uma população, entre outros.



Pode ser usado para avaliar a precisão e a consistência dos dados, identificar anomalias ou pontos fora do comum e ajudar a tomar decisões estatísticas.

**SEM RÉPLICA NÃO SE SABE SE SEU DADO É CONFIÁVEL OU NÃO**

## PENSE NA SEGUINTE SITUAÇÃO

1. O Pantanal é um ambiente sensível.
2. A introdução de espécies exóticas é um grande problema pois pode afetar toda a cadeia alimentar.
3. Houve a introdução de uma espécie de tracajá, a qual parece estar se proliferando em alguns locais.
4. Então, onde os tracajás estão mais frequentes? Existe alguma variável ambiental que possa explicar a sua frequência?



## MONTANDO MINHA PESQUISA COM RÉPLICAS

### NESTAS CONDIÇÕES, PODEMOS IR PARA CAMPO E ANALISAR ALGUMAS QUESTÕES

1. Comparar ambientes onde os indivíduos estão presentes vs. onde estão ausentes

**NÃO ADIANTA OBSERVARMOS SOMENTE ONDE ESTÃO PRESENTES, LÁ É FAVORÁVEL.  
É PRECISO OBSERVAR: ONDE OS INDIVÍDUOS ESTÃO AUSENTES E PARECE  
SER PROPÍCIO A SUA PRESENÇA? POR QUE NÃO ESTÃO AQUI?**

2. Utilizar de equipamentos que possam medir variáveis ambientais que influenciam a presença/ausência da espécie.
3. Teremos que investigar no mínimo (03) locais, pois assim conseguiremos fazer uma análise de média confiável, que possa retornar um desvio padrão.

## DESTA FORMA, PRODUZIMOS OS SEGUINTE DADOS

QUADRO 1 – LOCAIS DE AMOSTRAGEM E NÚMEROS DE INDIVÍDUOS OBSERVADOS CONTENDO A MÉDIA E O DESVIO PADRÃO

LOCAIS DE AMOSTRAGEM	NÚMERO DE INDIVÍDUOS OBSERVADOS
PRAIA 01	12
PRAIA 02	5
PRAIA 03	14
MÉDIA	10,33333333
DESVIO PADRÃO	4,725815626

FONTE: ELABORADO PELOS AUTORES (2024).

Foram observados 31 tracajás. Uma média de 10,33 por praia. Logo, isso não significa que se eu for em uma praia, irei encontrar 10,33, mas algo entre 5 e 14 (amplitude).

A média só reduz a variação dos dados, por isso é importante ter outras medidas de dispersão dos dados, como o desvio padrão. O desvio padrão apresentado se refere ao total de locais analisados.

Veja que na praia 03 foram observados o maior número de indivíduos.

Entretanto, te pergunto: Será que foi um acaso?

Se você voltar amanhã, será que haverá o mesmo número de indivíduos?

Os indivíduos foram observados em um mesmo horário do dia? Houve variação de temperatura?



## POR ISSO, DEVE HAVER RÉPLICAS!

### VEJA A IMPORTÂNCIA DAS RÉPLICAS

QUADRO 2 – LOCAIS DE AMOSTRAGEM, DATAS E NÚMERO DE INDIVÍDUOS OBSERVADOS

LOCAIS DE AMOSTRAGEM	DATA	NÚMERO DE INDIVÍDUOS OBSERVADOS
PRAIA 1	DIA 01	12
PRAIA 2	DIA 01	5
PRAIA 3	DIA 01	14
PRAIA 1	DIA 02	13
PRAIA 2	DIA 02	7
PRAIA 3	DIA 02	12
PRAIA 1	DIA 03	22
PRAIA 2	DIA 03	2
PRAIA 3	DIA 03	6
MÉDIA		10,33333333
DESVIO PADRÃO		6,020797289

FONTE: ELABORADO PELOS AUTORES (2024).

Neste caso, o número de dias foi ampliado, sendo, as informações, coletadas em 03 dias nas mesmas 03 praias.

Veja que a média continua a mesma, são 10,33 indivíduos por praia.

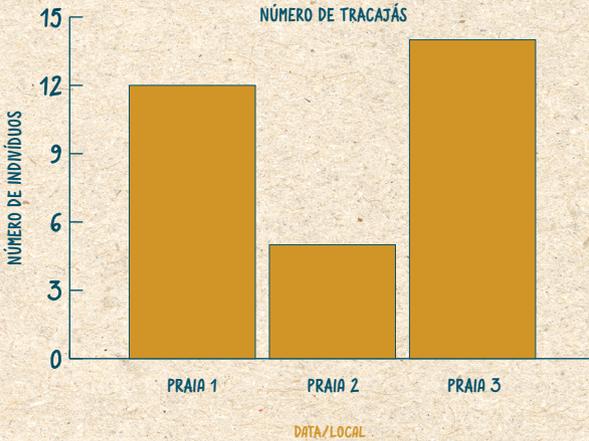
Mas perceba que o desvio padrão mudou.

Aumentou. Este desvio padrão aumentado, provavelmente, se deve a alguma variabilidade encontrada nos dados.

Neste caso a praia número 03 foi aquela com o maior número de indivíduos, e a praia 01 a segunda maior. A praia número 02 foi aquela que apresentou menor variação entre os dados, é só testar via desvio padrão no excel.

### OS GRÁFICOS PODEM SAIR DESTE FORMATO...

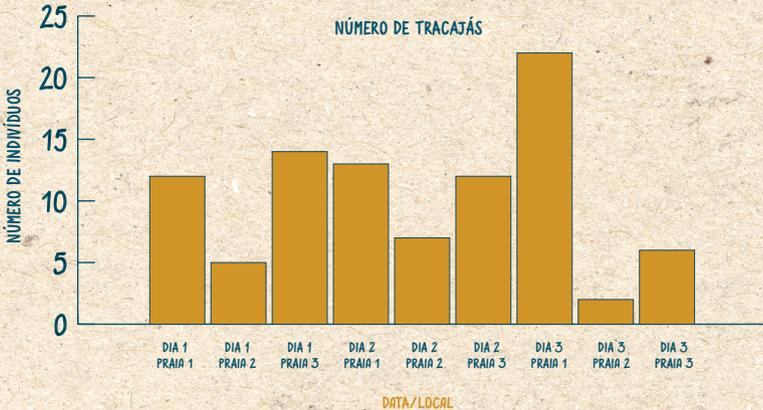
GRÁFICO 1 – NÚMERO DE INDIVÍDUOS ENCONTRADOS EM CADA AMBIENTE AMOSTRADO



FONTE: ELABORADO PELOS AUTORES (2024).

### ...PARA ESTE. PERCEBA AS DINÂMICAS DE ACORDO COM AS DATAS

GRÁFICO 2 – NÚMERO DE INDIVÍDUOS OBSERVADOS EM CADA AMBIENTE AMOSTRADO DURANTE O PERÍODO DE PESQUISA



FONTE: ELABORADO PELOS AUTORES (2024). DADOS FICTÍCIOS.

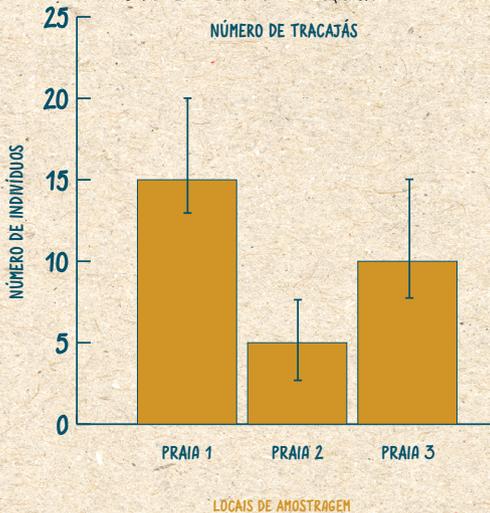
Veja que estão divididos por data. Assim, como há réplicas, é possível fazer um desvio padrão que demonstre a variabilidade dessas médias.

Neste caso, observe que a maior média foi encontrada na praia 01. Ok!

Mas, também observe que na praia 01 há uma maior barra de erros, demonstrando que houve uma maior variação nos dados encontrados. Ou seja, na praia 01, em alguns dias você pode encontrar mais ou menos indivíduos.

Já a praia 02 apresenta um desvio padrão menor. Isso denota uma certa constância no número de indivíduos. Isso tudo implica na discussão de seus resultados posteriormente.

GRÁFICO 3 – NÚMERO DE INDIVÍDUOS OBSERVADOS EM CADA AMBIENTE DURANTE O PERÍODO DE PESQUISA



FONTE: ELABORADO PELOS AUTORES (2024). DADOS FICTÍCIOS.

## MAS... HÁ DIFERENÇA SIGNIFICATIVA ENTRE OS LOCAIS AMOSTRADOS?

Para saber se uma praia, ou outra, é mais visitada pelos tracajás, é importante que seja feito um teste de médias para analisar a sua significância.

Não basta observar que uma média é maior do que outra, ou se o desvio padrão é amplo ou não.

Deve-se fazer um teste de médias!

O TESTE DE MÉDIA É UMA FERRAMENTA ESTATÍSTICA USADA PARA DETERMINAR SE HÁ EVIDÊNCIAS ESTATÍSTICAS SUFICIENTES PARA INFERIR QUE A MÉDIA DE UMA AMOSTRA (OU GRUPO) DE DADOS É ESTATISTICAMENTE DIFERENTE DA MÉDIA DE UMA POPULAÇÃO MAIOR (OU DE OUTRO GRUPO DE DADOS)



## O PASSO A PASSO DO TESTE DE MÉDIA

### PASSO 1 – TUDO INICIA COM A FORMULAÇÃO DAS HIPÓTESES

**HIPÓTESE NULA ( $H_0$ ):** a hipótese nula afirma que não há diferença estatisticamente significativa entre as médias da amostra e da população (ou entre as duas amostras).

**HIPÓTESE ALTERNATIVA ( $H_1$ ):** a hipótese alternativa afirma que há uma diferença estatisticamente significativa entre as médias.

#### POR EXEMPLO

**$H_0$ :** o número de tracajás não varia nas praias estudadas, pois há uma semelhança na disponibilidade de recursos ambientais.

GRÁFICO 4 – MODELO HIPOTÉTICO DO NÚMERO DE INDIVÍDUOS OBSERVADOS EM CADA AMBIENTE AMOSTRADO DURANTE O PERÍODO DE PESQUISA



FONTE: ELABORADO PELOS AUTORES (2024).  
DADOS FICTÍCIOS.

GRÁFICO 5 – MODELO HIPOTÉTICO DO NÚMERO DE INDIVÍDUOS OBSERVADOS EM CADA AMBIENTE AMOSTRADO DURANTE O PERÍODO DE PESQUISA



ELABORADO PELOS AUTORES (2024).  
DADOS FICTÍCIOS.

**$H_1$ :** embora a disponibilidade de recursos ambientais seja semelhante entre as praias, o número de tracajás é maior nas praias onde há menor possibilidade de predação (maior cobertura vegetal).

**DESTA FORMA, CONSTRUIMOS AS HIPÓTESES BASEADAS EM ALGUNS FATOS OBSERVADOS EM CAMPO OU EM FATOS DA LITERATURA**

## PASSO 2 – É AGORA QUE O BICHO PEGA!

**COLETA DE DADOS:** É necessário coletar uma amostra de dados relevante para o teste. A amostra deve ser representativa da população ou do grupo que você está estudando.

Neste passo deve-se considerar as réplicas e os fatores que você irá analisar. Por exemplo, se você for em praias diferentes em horários diferentes, poderá encontrar números de indivíduos diferentes somente pelo fato de horário do dia. Então, esta variável deve ser contabilizada.

Não se esqueça das réplicas. Elas podem ser os diferentes horários do dia. Dias diferentes. Períodos hidrológicos distintos (cheia, seca). Dentre outros fatores.



## PASSO 3 – A ESCOLHA DO TESTE ESTATÍSTICO

A escolha do teste depende das características dos dados (por exemplo, tamanho da amostra e distribuição dos dados) e da natureza da pesquisa. Alguns dos testes de média comuns incluem o Teste t de Student, Teste Z, Teste ANOVA, entre outros.

Neste caso, se você for comparar duas estações hidrológicas, por exemplo (cheia vs seca), deverá haver duas colunas na sua planilha de dados (Excel), desconsiderando o número de dias coletados nas duas estações (os quais devem ser em mesmo número para “parear”).

Assim, pode-se utilizar um Teste t de Student para amostras pareadas.



O TESTE **T DE STUDENT** É USADO PARA RESPONDER A PERGUNTAS COMO

“A AMOSTRA DE TRATAMENTO A É ESTATISTICAMENTE DIFERENTE DA AMOSTRA DE TRATAMENTO B?”

OU

“HÁ UMA DIFERENÇA SIGNIFICATIVA ENTRE AS MÉDIAS DE DOIS GRUPOS?”

Ele compara as médias amostrais e avalia se a diferença observada é maior do que o que seria esperado devido ao acaso.

Escolha o teste **t de Student** apropriado com base nas características dos dados:

Teste t de Student para amostras independentes (t de Student não pareado): Use este teste quando as duas amostras são independentes umas das outras.

Teste t pareado (t de Student pareado): Use este teste quando as observações nas duas amostras estão emparelhadas (por exemplo, antes e depois de um tratamento).

Calcule a estatística de teste t, que leva em consideração a média, o tamanho da amostra e o desvio padrão de cada amostra.



## ESCOLHA O TESTE T DE STUDENT APROPRIADO COM BASE NAS CARACTERÍSTICAS DOS DADOS

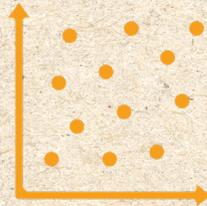
### TESTE T DE STUDENT PARA AMOSTRAS INDEPENDENTES (T DE STUDENT NÃO PAREADO)



Use este teste quando as duas amostras são independentes umas das outras.

### TESTE T PAREADO (T DE STUDENT PAREADO)

Use este teste quando as observações nas duas amostras estão emparelhadas (por exemplo, antes e depois de um tratamento).



Calcule a estatística de **teste t**, que leva em consideração a média, o tamanho da amostra e o desvio padrão de cada amostra.

## SÓ LEMBRANDO

Em estatística, as amostras podem ser classificadas como “**independentes**” ou “**dependentes**” com base na relação entre os dados nas amostras e na maneira como os dados foram coletados.

## AMOSTRAS INDEPENDENTES

Amostras independentes são conjuntos de dados em que as observações de uma amostra não estão relacionadas às observações da outra amostra. Isso significa que as observações em uma amostra são independentes das observações na outra amostra. Em outras palavras, os dados em uma amostra não são afetados ou influenciados pelos dados na outra amostra.

### EXEMPLOS DE AMOSTRAS INDEPENDENTES

TESTE DE EFICÁCIA DE  
DOIS MEDICAMENTOS  
DIFERENTES EM GRUPOS  
DE PACIENTES DIFERENTES

COMPARAÇÃO DAS ALTURAS  
DE DUAS AMOSTRAS DE  
ESTUDANTES DE ESCOLAS  
DIFERENTES



## AMOSTRAS DEPENDENTES (OU EMPARELHADAS)

Amostras dependentes, também conhecidas como amostras emparelhadas, são conjuntos de dados em que as observações de uma amostra estão relacionadas às observações da outra amostra de alguma forma. Isso pode ocorrer quando as observações em uma amostra estão emparelhadas ou relacionadas às observações na outra amostra. Em amostras dependentes, as observações em uma amostra são afetadas ou influenciadas pelas observações na outra amostra.

### EXEMPLOS DE AMOSTRAS DEPENDENTES

COMPARAÇÃO DAS  
PONTUAÇÕES DE  
UM GRUPO DE  
ESTUDANTES ANTES  
E DEPOIS DE UM  
PROGRAMA DE  
TUTORIA

TESTE DE EFICÁCIA DE  
UM NOVO TRATAMENTO  
MÉDICO, ONDE CADA  
PACIENTE SERVE  
COMO SEU PRÓPRIO  
CONTROLE (MEDINDO  
AS CONDIÇÕES DO  
PACIENTE ANTES  
E DEPOIS DO  
TRATAMENTO)



## AGORA É SÓ FAZER O TESTE ESTATÍSTICO

### VEJA COMO OS DADOS DEVEM ESTAR DISPOSTOS E O PASSO A PASSO DO TESTE NO EXCEL

QUADRO 3 – ORGANIZAÇÃO DOS DADOS E PROCEDIMENTO PARA O TESTE NO EXCEL

LOCAIS DE COLETA	CHEIA	SECA
PRAIA 01	12	16
PRAIA 01	13	18
PRAIA 01	22	20
PRAIA 01	10	22
PRAIA 01	15	13
PRAIA 02	5	1
PRAIA 02	7	3
PRAIA 02	2	6
PRAIA 02	5	5
PRAIA 02	4	7
PRAIA 03	14	10
PRAIA 03	12	20
PRAIA 03	6	29
PRAIA 03	5	25
PRAIA 03	12	26

FONTE: ELABORADO PELOS AUTORES (2024).  
DADOS FICTÍCIOS

## TUTORIAL

### MÃO NA MASSA, OU MELHOR, NO TECLADO!

ABRA O EXCEL, VÁ NA ABA – DADOS – DEPOIS ESCOLHA  
A OPÇÃO ANÁLISE DE DADOS

Neste guia introdutório, não abordamos a verificação da normalidade dos dados — isso ficará para o Guia 2; aqui, oferecemos apenas um gostinho das análises.



	Locais de Coleta	Cheia	Seca
7	Praia 01	12	16
8	Praia 01	13	18
9	Praia 01	22	20
10	Praia 01	10	22

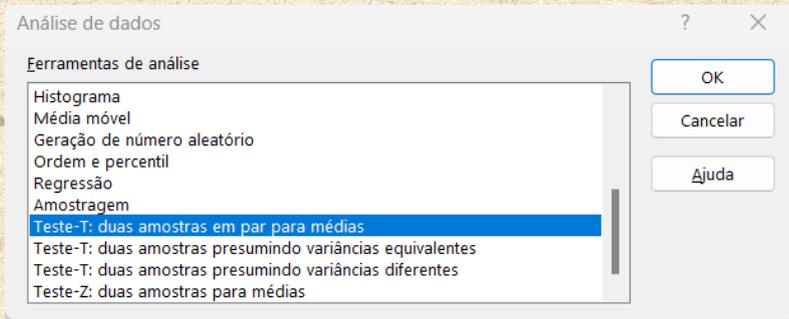


Caso esta opção não esteja habilitada em seu Excel, habilite em:

ARQUIVO > OPÇÕES > SUPLEMENTOS > IR > FERRAMENTAS DE ANÁLISE

## LOGO EM SEGUIDA UMA JANELA IRÁ SE ABRIR

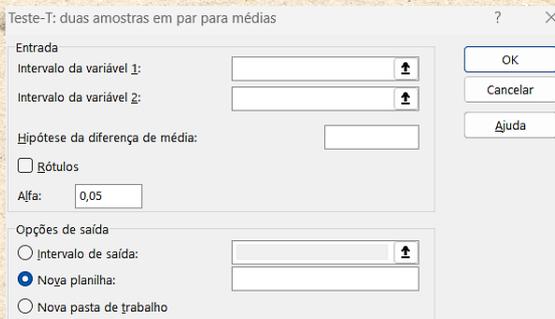
### ESCOLHA A OPÇÃO TESTE-T



### OUTRA JANELA SE ABRIRÁ PARA VOCÊ INSERIR AS VARIÁVEIS

A variável **1** pode ser os dados contidos em **"CHEIA"** e os da variável **2** em **"SECA"**.

Selecione os dados desde os rótulos de linhas para que possam ser melhor apresentados nos resultados. Clique na caixa **"RÓTULOS"**. Deixe seu Alfa como **0,05**, que será a probabilidade de haver diferença significativa.



ESCOLHA ONDE QUER  
QUE SEUS DADOS SEJAM  
PLOTADOS

COMO SUGESTÃO, FAÇA  
PARA "NOVA PLANILHA"

AO CLICAR EM OK, SEUS RESULTADOS APARECERÃO EM OUTRA PLANILHA, COMO MOSTRA A FIGURA ABAIXO. AGORA É SÓ INTERPRETAR!

	A	B	C
1	Teste-t: duas amostras em par para médias		
2			
3		<i>Cheia</i>	<i>Seca</i>
4	Média	9,6	14,73333333
5	Variância	28,82857143	81,35238095
6	Observações	15	15
7	Correlação de Pearson	0,356051267	
8	Hipótese da diferença de média	0	
9	gl	14	
10	Stat t	-2,285129495	
11	P(T<=t) uni-caudal	0,019208764	
12	t crítico uni-caudal	1,761310136	
13	P(T<=t) bi-caudal	0,038417529	
14	t crítico bi-caudal	2,144786688	

Veja que nos resultados ele já mostra a média das estações coletadas “CHEIA” e “SECA”. Veja que houve, em média, mais tracajás avistados no período de Seca do que na Cheia.

Nos dados também estão dispostas a Variância, o número de Observações.

## INTERPRETAÇÕES ESTATÍSTICAS DO TESTE

A correlação de Pearson demonstra a relação linear entre duas variáveis, ou seja, quando aumenta uma também aumenta a outra? Ou vice versa?

Os valores de Pearson variam de 0 a 1, ou -1. Neste caso, quanto mais próximo a 0, menor é a relação entre as duas variáveis, e quanto mais próximo a 1 ou -1 maior é a relação de linearidade.

O “gl” ali presente é o grau de liberdade. Esta análise é baseada no número de observações feitas. Lembrando que, quanto mais observações feitas, mais robusta é sua análise.

Logo em seguida apresenta o “Stat t” este é o valor de “t” de Student. Neste caso é de -2,28.

O valor de t de Student de -2,68 indica que a média amostral está, consideravelmente, abaixo da média populacional (ou de referência), e a magnitude do desvio é 2,68 vezes o erro padrão da amostra.

## CONTINUANDO COM AS INTERPRETAÇÕES...

Agora vamos direto ao ponto que nos interessa. O valor de “P.”, pois é o valor de “P” que vai nos dizer se as estações hidrológicas apresentam dados diferentes significativamente em relação ao número de traçajás.

Veja que o valor de P uni-caudal foi de 0,01 o que significa que o resultado é menor do que 0,05, indicando diferença significativa entre as estações do ano.

O “gl” ali presente é o grau de liberdade. Esta análise é baseada no número de observações feitas. Lembrando que, quanto mais observações feitas, mais robusta é sua análise.

Logo, em seguida, apresenta o “Stat t” este é o valor de “t” de Student. Neste caso é de -2,28.

O valor de t de Student de -2,68 indica que a média amostral está, consideravelmente, abaixo da média populacional (ou de referência), e a magnitude do desvio é 2,68 vezes o erro padrão da amostra.

## RELEMBRANDO A PROBABILIDADE

O valor-p, abreviação de *probability value* (valor de probabilidade), é uma medida estatística fundamental usada para avaliar a significância estatística de resultados em testes de hipóteses. É uma ferramenta essencial em estatística e desempenha um papel crucial na interpretação de resultados de experimentos e estudos.

O valor-p representa a probabilidade de obter resultados tão extremos ou mais extremos do que os observados, supondo que a hipótese nula seja verdadeira. A hipótese nula é uma afirmação inicial que não há efeito, não há diferença ou não há relação entre as variáveis testadas.

Se o valor-p for pequeno (geralmente, menor que um nível de significância escolhido, como 0,05), isso sugere que os resultados são improváveis de ocorrer sob a hipótese nula. Portanto, você pode considerar rejeitar a hipótese nula, indicando que há evidência estatística a favor da hipótese alternativa.

Se o valor-p for grande, isso sugere que os resultados são consistentes com o que seria esperado sob a hipótese nula. Nesse caso, você não rejeita a hipótese nula, indicando que não há evidência estatística para suportar a hipótese alternativa.

Interessante, não é?! Agora veja o seguinte. O valor de “P” é a probabilidade de sua hipótese nula ocorrer, ou não. Quanto menor for o “P”, menor é a probabilidade de sua hipótese nula acontecer.

O valor-p unicaudal é usado quando a hipótese alternativa é direcionada em uma única direção. Em outras palavras, você está interessado em saber se os resultados são significativamente maiores ou menores do que o esperado, mas não em ambas as direções. Ou seja, você afirma que A é maior que B devido certa circunstância, por exemplo.

O valor-p bicaudal é usado quando a hipótese alternativa não especifica uma direção particular e sugere que há uma diferença significativa, mas não especifica se essa diferença será maior ou menor do que o esperado.

## RETORNANDO ÀS HIPÓTESES

### FAZEMOS ENTÃO AS SEGUINTE HIPÓTESES

**H0:** O número de tracajás não varia nas estações hidrológicas estudadas, pois há uma semelhança na disponibilidade de recursos ambientais.

**H1:** Embora a disponibilidade de recursos ambientais seja semelhante entre as estações hidrológicas, o número de tracajás é maior no período de seca onde há maior área disponível. Neste caso, estou direcionando minha análise. Seca tem mais tracajá do que Cheia.

E como vimos, o valor de P foi 0,01, ou seja, menor do que 0,05. Isso significa que o período de Seca tem mais tracajás do que o período de Cheia.



## PASSO 4 – TOMADA DE DECISÃO

COM BASE NO VALOR DE P E NO NÍVEL DE SIGNIFICÂNCIA ESCOLHIDO (GERALMENTE 0,05),  
VOCÊ DECIDE SE REJEITA OU NÃO A HIPÓTESE NULA

SE O VALOR-P FOR MENOR QUE O NÍVEL DE SIGNIFICÂNCIA,  
VOCÊ REJEITARÁ A HIPÓTESE NULA



## PASSO 5 – INTERPRETAÇÃO DOS RESULTADOS

A interpretação dos resultados depende da decisão tomada no passo anterior. Se a hipótese nula for rejeitada, você pode concluir que há evidências de uma diferença estatisticamente significativa nas médias. Caso contrário, não há evidências suficientes para afirmar tal diferença.

DEVE-SE LEVAR EM CONSIDERAÇÃO AINDA OS FATORES QUE FAZEM  
OU NÃO A DIFERENÇA SIGNIFICATIVA NA ESTATÍSTICA.  
PARA ISSO, É NECESSÁRIA MUITA LEITURA!



## MAS CASO VOCÊ QUEIRA COMPARAR AS PRAIAS DIFERENTES, DAÍ O TESTE DEVERÁ SER OUTRO!

Para o teste com este tipo de comparação, normalmente, utilizamos o teste **ANOVA**.

O Teste de Análise de Variância (ANOVA) é uma técnica estatística usada para determinar se há diferenças significativas entre as médias de três ou mais grupos ou amostras independentes. O teste ANOVA avalia se as variações entre os grupos (ou tratamentos) são, estatisticamente, significativas em comparação com a variação dentro dos grupos.

Em outras palavras, ele ajuda a responder se pelo menos um grupo é estatisticamente diferente dos outros.



## ANOVA

O ANOVA é uma ferramenta valiosa em pesquisas biológicas, pois permite a comparação de múltiplos grupos de dados e a avaliação das diferenças entre eles. Ele é amplamente utilizado em estudos de ecologia, genética, botânica e em muitos outros campos da biologia para investigar o efeito de diferentes tratamentos, condições ou variáveis independentes em um conjunto de dados.

Agora sim. Se você for comparar as praias, então terá que ter várias réplicas, seja de dias, ou de estações. Mas, agora a pergunta é se existe uma diferença entre as praias.

Para isso, em sua planilha de dados deverá haver pelo menos 03 colunas (ex. Praia 01, Praia 02, Praia 03).

Essas colunas representarão suas réplicas, uma embaixo da outra.

Agora sim. Se você for comparar as praias, então terá que ter várias réplicas, seja de dias ou de estações. Mas agora, a pergunta é se existe uma diferença entre as praias. Para isso, em sua planilha de dados deverá haver pelo menos três colunas (exemplo: Praia 01, Praia 02, Praia 03). Essas colunas representarão as suas réplicas, dispostas uma embaixo da outra. Desta forma, por exemplo: Dia 01, Dia 02, Dia 03, e assim por diante.

A seguir, veja como os dados devem estar organizados:

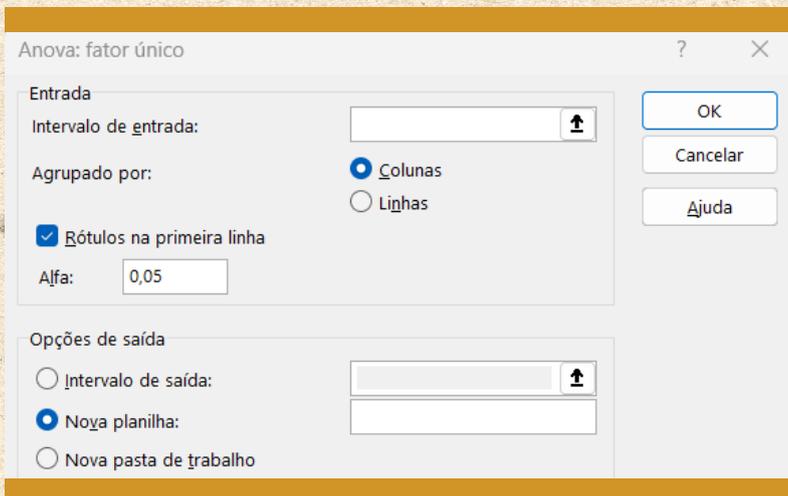
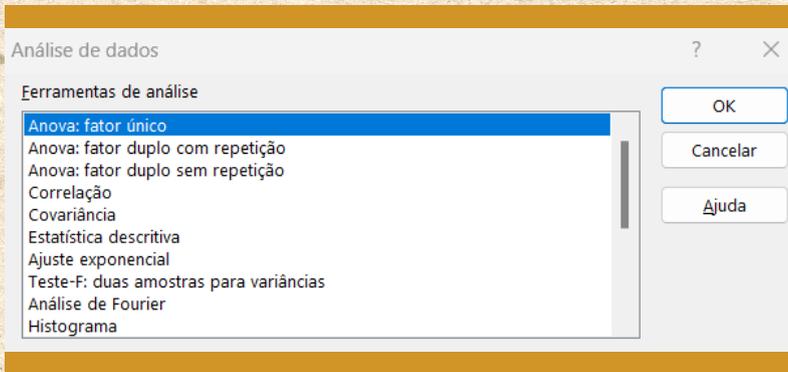
QUADRO 4 – ORGANIZAÇÃO DOS DADOS COLETADOS NAS TRÊS PRAIAS

DIA DE COLETA	PRAIA 01	PRAIA 02	PRAIA 03
1	12	5	14
2	13	7	12
3	22	2	6
4	10	5	5
5	15	4	12
6	16	1	10
7	18	3	20
8	20	6	29
9	22	5	25
10	13	7	26

FONTE: ELABORADO PELOS AUTORES (2024).

## AGORA O TESTE É MUITO PARECIDO. VÁ NA ANÁLISE DE DADOS, ESCOLHA A OPÇÃO ANOVA FATOR ÚNICO

Entretanto, ao invés de selecionar cada variável, separadamente, deverá selecionar todas juntas.



## SUA SELEÇÃO DEVERÁ ESTAR DA SEGUINTE FORMA

Anova: fator único

Entrada  
Intervalo de entrada:

Agrupado por:  Colunas  Linhas

Rótulos na primeira linha

Alfa:

Opções de saída  
 Intervalo de saída:

Nova planilha:

Nova pasta de trabalho

	PRAIA 02	PRAIA 03
	5	14
	7	12
	2	6
	5	5
	4	12
	1	10
	3	20
	6	29
	5	25
	7	26

E SEUS RESULTADOS DEVERÃO ESTAR PLOTADOS NA PLANILHA NOVA

1	Anova: fator único						
2							
3	RESUMO						
4	<i>Grupo</i>	<i>Contagem</i>	<i>Soma</i>	<i>Média</i>	<i>Variância</i>		
5	Dia de coleta	10	55	5,5	9,166667		
6	PRAIA 01	10	161	16,1	18,1		
7	PRAIA 02	10	45	4,5	4,055556		
8	PRAIA 03	10	159	15,9	73,21111		
9							
10							
11	ANOVA						
12	<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>valor-P</i>	<i>F crítico</i>
13	Entre grupos	1215,2	3	405,0666667	15,5	1,23E-06	2,866266
14	Dentro dos grupos	940,8	36	26,13333333			
15							
16	Total	2156	39				

## INTERPRETANDO

A interpretação dos dados é bem semelhante àquela do teste t. Veja as médias, variância e o gl. Agora, o que temos que prestar atenção é no valor de F. O valor de F é calculado ao dividir a variabilidade entre os grupos (variância entre grupos) pela variabilidade dentro dos grupos (variância dentro dos grupos).

A estatística F segue uma distribuição F, cuja forma depende dos graus de liberdade associados à variabilidade entre grupos e dentro dos grupos. Quanto maior a estatística F, maior a evidência de diferenças significativas entre os grupos.

E também o “valor-P”, neste caso, está em notação científica  $1,23E-06$ , isso significa que o valor de P é 0,0000012. Ou seja, menor do que 0,05. Assim, rejeita-se a hipótese nula e aceita a hipótese 1.



## CHEGANDO À CONCLUSÃO DE QUE HÁ UMA DIFERENÇA ESTATÍSTICA SIGNIFICATIVA ENTRE AS PRAIAS AMOSTRADAS

SÓ NÃO SABEMOS QUAIS PRAIAS SÃO DIFERENTES ENTRE SI.

MAS ISSO É PAPO PARA OUTRO MOMENTO

### POR QUE O VALOR DE P TEM QUE SER MENOR DO QUE 0,05?

O valor-p é um componente crítico de testes estatísticos, como testes t, ANOVA, regressão, entre outros. A escolha do nível de significância, que, frequentemente, é estabelecido em 0,05 (ou 5%), é uma convenção amplamente adotada, mas ela não é uma regra rígida e inflexível.

A decisão de considerar um valor-p menor do que 0,05 como estatisticamente significativo é baseada em várias considerações estatísticas, práticas e científicas, mas essa escolha não é arbitrária e deve ser justificada.

Para as tomadas de decisão do valor-p, é importante levar em consideração livros como Legendre e Legendre (2012). Neste livro, os autores alertam que o valor de p não mede a magnitude do efeito, nem a sua importância ecológica, apenas a probabilidade de se obter um resultado igual ou mais extremo que o observado, assumindo que a hipótese nula é verdadeira. Ou seja, é uma medida condicional.



A ESCOLHA DE 0,05 COMO NÍVEL DE SIGNIFICÂNCIA É UMA PRÁTICA COMUM E TRADICIONAL EM ESTATÍSTICA E PESQUISA CIENTÍFICA.

ISSO FACILITA A COMPARAÇÃO DE RESULTADOS ENTRE DIFERENTES ESTUDOS E ÁREAS DE PESQUISA

UM VALOR-P MENOR QUE 0,05 É CONSIDERADO UMA EVIDÊNCIA FORTE DE QUE OS RESULTADOS SÃO, ESTATISTICAMENTE, SIGNIFICATIVOS. ISSO TORNA A INTERPRETAÇÃO DOS RESULTADOS MAIS DIRETA E COMPREENSÍVEL

NESSES TESTES REALIZADOS NÃO FOI LEVADA EM CONSIDERAÇÃO A NORMALIDADE DOS DADOS. E, ENTÃO, CONTINUAREMOS EM OUTRAS OCASIÕES!

## E A MODELAGEM, O QUE É?

Modelos estatísticos, também chamados de modelagem estatística, é um conjunto de técnicas probabilísticas para representar a realidade de forma mais simplificada. A ideia por trás dessa metodologia é criar um modelo que consiga descrever os elementos mais importantes para uma análise.



## REGRESSÃO LINEAR

A regressão linear é, frequentemente, usada para modelar a relação entre duas variáveis contínuas.

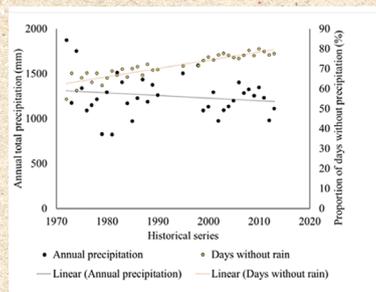
Por exemplo, você pode usar regressão linear para estudar a relação entre a altura das plantas e a quantidade de luz que recebem.

Modelos de crescimento populacional, como o modelo logístico, são usados para entender como as populações de organismos crescem ao longo do tempo, levando em consideração fatores como a disponibilidade de recursos.

Modelo de nicho ecológico pode ser usado para entender a distribuição geográfica de espécies em relação a fatores ambientais.

A análise de séries temporais é usada para estudar dados biológicos coletados ao longo do tempo, como dados de monitoramento de populações, mudanças climáticas e variações sazonais.

Veja no exemplo ao lado. Uma série temporal demonstrando duas linhas de regressão. Uma que identifica os dias sem chuva, e outra demonstrando a quantidade de precipitação. Veja a tendência das duas. Neste caso, a regressão pode dar uma indicação da tendência de aumentar ou diminuir algo em relação ao tempo, por exemplo.



Fonte: SciELO Brasil - Climate change reflected in one of the largest wetlands in the world: an overview of the Northern Pantanal water regime. Climate change reflected in one of the largest wetlands in the world: an overview of the Northern Pantanal water regime. Disponível em: <https://www.scielo.br/j/alb/a/6WY69WMrd8CDsycNgWSwF5j/>. Acesso em: 19 abr. 2024.

## CONTINUANDO COM A REGRESSÃO LINEAR

Imagine que você está com uma pergunta relacionada ao problema de microclima em fragmentos florestais.

Ou seja, qual é o efeito de borda dos fragmentos florestais em relação à temperatura? Quanto mais para dentro de um fragmento florestal, mais fresco será? A seguir, apresenta-se o quadro com os dados coletados, mostrando as temperaturas registradas no centro e na borda de diferentes fragmentos florestais de variados tamanhos.

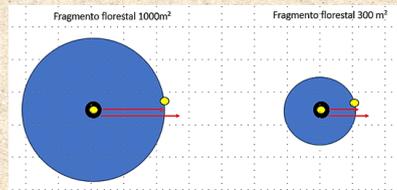
QUADRO 5 – TEMPERATURAS NO CENTRO E NA BORDA DE FRAGMENTOS FLORESTAIS DE DIFERENTES TAMANHOS

LOCAL	TAMANHO (M <sup>2</sup> )	TEMPERATURA NO CENTRO (°C)	TEMPERATURA NA BORDA (°C)
FRAGMENTO	530	30	28
FRAGMENTO	360	27	29
FRAGMENTO	120	29	30
FRAGMENTO	594	31	29
FRAGMENTO	680	31	28
FRAGMENTO	720	32	28
FRAGMENTO	1200	33	29
FRAGMENTO	5000	32	28
FRAGMENTO	5600	31	29
FRAGMENTO	6000	32	28

FONTE: ELABORADO PELOS AUTORES (2024) DADOS FICTÍCIOS..

Considere os dados fictícios acima. Veja que foram realizadas coletas em 10 fragmentos florestais de diferentes tamanhos. Em cada fragmento foram

tomadas a temperatura na borda (entre a área aberta e a fechada) e no centro (bem no centro) veja o esquema ao lado.



Neste caso, eu faço a regressão linear. Seleciono os dados de temperatura no centro e na borda de todos os fragmentos. Vou na aba inserir e daí gráfico de dispersão. Veja, na imagem a seguir, que só estão selecionados os dados de tamanho e de temperatura no centro.

Arquivo Página Inicial **Inserir** Layout da Página Fórmulas Dados Revisão Ex

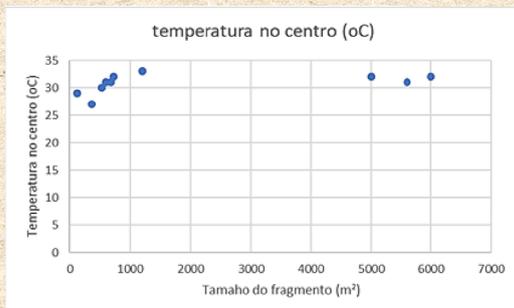
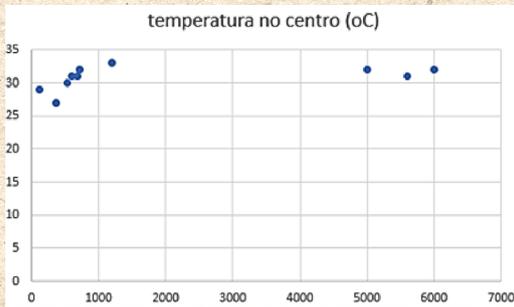
Tabela Dinâmica Tabelas Dinâmicas Recomendadas Tabela Ilustrações Gráficos Recomendados Gráficos

Tabelas Gráficos

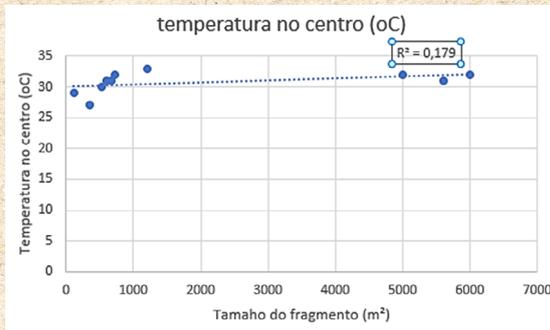
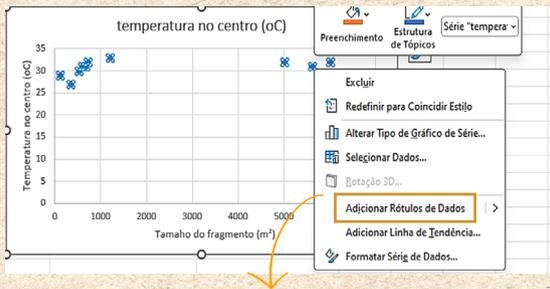
B1 tamanho (m<sup>2</sup>)

	A	B	C	D
1	local	tamanho (m <sup>2</sup> )	temperatura no centro (oC)	temperatura na borda (oC)
2	Fragmento	530	30	28
3	Fragmento	360	27	29
4	Fragmento	120	29	30
5	Fragmento	594	31	29
6	Fragmento	680	31	28
7	Fragmento	720	32	28
8	Fragmento	1200	33	29
9	Fragmento	5000	32	28
10	Fragmento	5600	31	29
11	Fragmento	6000	32	28
12				
13				
14				

Assim, vai sair um gráfico com pontos. Logo após, deverá formatar o gráfico, colocar o nome do eixo y e de do eixo x. Se quiser, tire o título do gráfico.

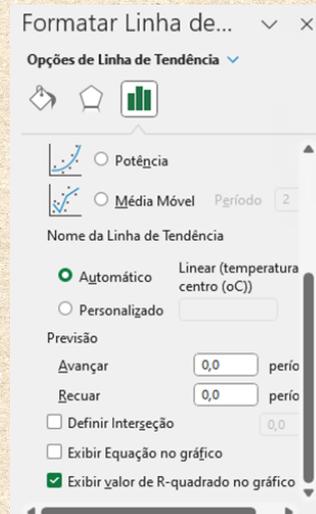


Em seguida, devemos plotar a linha de tendência com o valor de  $R^2$  no gráfico.



Para plotar o valor do  $R^2$  no gráfico, devemos clicar com o botão direito em cima da linha de tendência e, então, abrirá uma janela do lado direito.

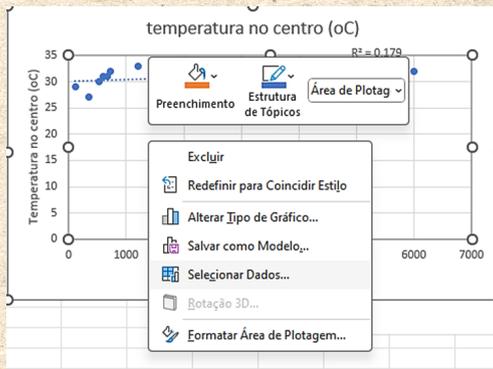
Nesta janela devo colocar Regressão linear (normalmente já estará selecionado) e depois eu vou na última opção – inserir  $R^2$  no gráfico.



## VOCÊ VIU QUE SÓ TEM UMA REGRESSÃO? TEMOS QUE COLOCAR DUAS

O PROCEDIMENTO É BEM SEMELHANTE E SEGUE ABAIXO

CLIQUE COM O BOTÃO DIREITO EM CIMA DO GRÁFICO. ABRIRÁ ESTA JANELA AÍ



VÁ EM SELECIONAR DADOS. ABRIRÁ OUTRA JANELA

Selecionar Fonte de Dados

Intervalo de dados do gráfico:

Entradas de Legenda (Série)	Rótulos do Eixo Horizontal (Categorias)
<input type="button" value="Adicionar"/> <input type="button" value="Editar"/> <input type="button" value="Remover"/> <input type="button" value="↑"/> <input type="button" value="↓"/>	<input type="button" value="Editar"/>
<input checked="" type="checkbox"/> temperatura no centro (oC)	530 360 120 594 680

Células Ocultas e Vazias

NESTA JANELA VOCÊ  
DEVERÁ ADICIONAR  
UM CONJUNTO DE  
VARIÁVEIS

ENTÃO, CLIQUE  
EM ADICIONAR

? X

**Nome da série:**  
 Seleccione um Intervalo

**Valores de X da série:**  
 Seleccione um Intervalo

**Valores de Y da série:**  
 = 1

OK Cancelar

Nesta janela você deverá colocar o nome da série como sendo o nome da sua nova variável Y, que é a Temperatura na borda ( °C). Logo após seleccionar os valores de x (tamanho da área m<sup>2</sup>). E, depois, (apague o valor que está para a variável y nesta janela, e insira os dados da variável y.

local	tamanho (m <sup>2</sup> )	temperatura no centro (oC)	temperatura na borda (oC)
'fragmento	530	30	28
'fragmento	360	27	29
'fragmento	120	29	30
'fragmento	594	31	29
'fragmento	680	31	28
'fragmento	720	32	28
'fragmento	1200	33	29
'fragmento	5000	32	28
'fragmento	5600	31	29
'fragmento	6000	32	28

local	tamanho (m <sup>2</sup> )	temperatura no centro (oC)	temperatura na borda (oC)
'fragmento	530	30	28
'fragmento	360	27	29
'fragmento	120	29	30
'fragmento	594	31	29
'fragmento	680	31	28
'fragmento	720	32	28
'fragmento	1200	33	29
'fragmento	5000	32	28
'fragmento	5600	31	29
'fragmento	6000	32	28

? X

**Nome da série:**  
 =Planilha4!\$D\$1 Seleccione um Intervalo

**Valores de X da série:**  
 Seleccione um Intervalo

**Valores de Y da série:**  
 =1

OK Cancelar

? X

**Nome da série:**  
 =Planilha4!\$D\$1 = temperatura na...

**Valores de X da série:**  
 =Planilha4!\$E\$2:\$E\$11 = 530; 360; 120...

**Valores de Y da série:**  
 =1

OK Cancelar

local	tamanho (m <sup>2</sup> )	temperatura no centro (oC)	temperatura na borda (oC)
'fragmento	530	30	28
'fragmento	360	27	29
'fragmento	120	29	30
'fragmento	594	31	29
'fragmento	680	31	28
'fragmento	720	32	28
'fragmento	1200	33	29
'fragmento	5000	32	28
'fragmento	5600	31	29
'fragmento	6000	32	28

? X

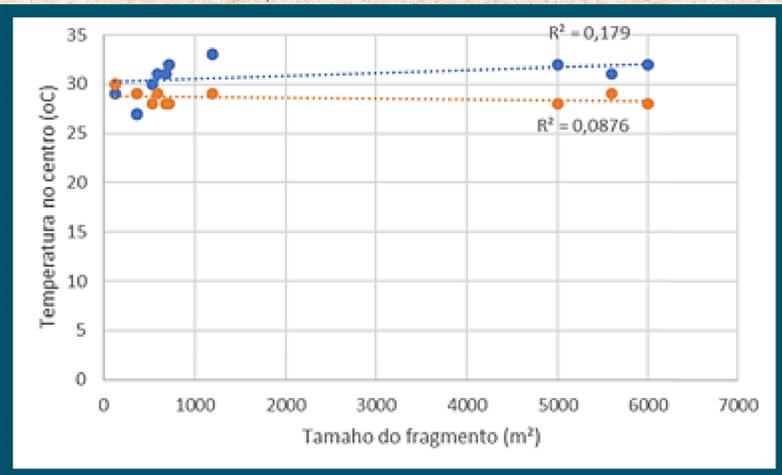
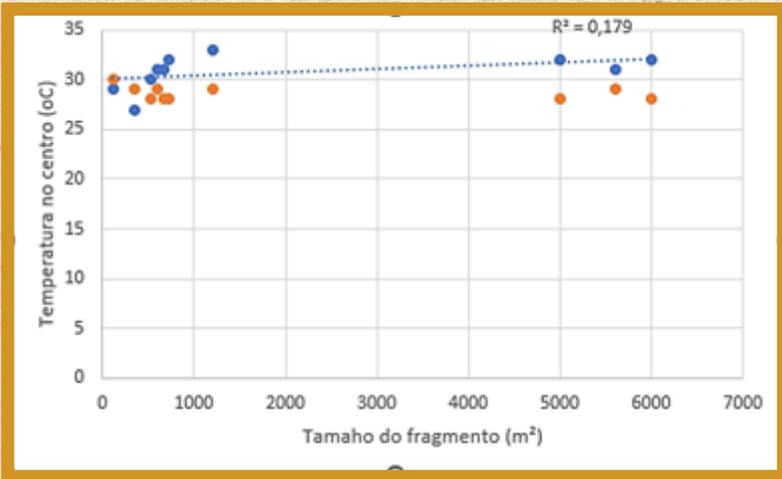
**Nome da série:**  
 =Planilha4!\$D\$1 = temperatura na...

**Valores de X da série:**  
 =Planilha4!\$E\$2:\$E\$11 = 530; 360; 120...

**Valores de Y da série:**  
 =Planilha4!\$D\$2:\$D\$11 = 28; 29; 30; 29...

OK Cancelar

Pronto. Agora você fez o gráfico de dispersão de pontos. Em seguida, faça o procedimento para adicionar a linha de tendência. Você lembra? Só voltar na página anterior.



**AGORA É QUE SÃO ELAS! VEJA QUE A INTERPRETAÇÃO DOS DADOS É CONTIGO. A PERGUNTA FOI RELACIONADA À DIFERENÇA DE TEMPERATURA DA BORDA E DO CENTRO. VEJA QUE A TEMPERATURA TENDE A AUMENTAR NO CENTRO DE FRAGMENTOS MAIORES, ENQUANTO TENDE A DIMINUIR NA BORDA DE FRAGMENTOS MENORES. PARA RESPONDER UMA QUESTÃO COMO ESTA, É NECESSÁRIA A ABORDAGEM E O AMPLO USO DE LITERATURA CIENTÍFICA!**

**MAS SERÁ QUE ISTO FAZ SENTIDO?  
AÍ É CONTIGO!**

## RESUMINDO UM POUCO...

A modelagem estatística é uma ferramenta essencial na biologia, pois ajuda a analisar dados, entender relações ambientais entre variáveis, além de auxiliar a fazer inferências a partir de observações.

## A MODELAGEM ENVOLVE ALGUNS PASSOS, QUAIS SEJAM

★ ANTES DE QUALQUER COISA ★

QUAL É O PROBLEMA E QUAL  
É A PERGUNTA??????

DEPOIS DISSO PODE SEGUIR

1. **COLETA DE DADOS**, como por exemplo: observação de organismos, medindo características, realizando experimentos ou coletando informações de campo.

A escolha adequada de uma amostra é fundamental para a representatividade dos resultados. A biologia muitas vezes lida com populações grandes e variadas, portanto, a amostragem cuidadosa é essencial. **É ESSENCIAL POSSUIR RÉPLICAS.**

2. **IDENTIFICAÇÃO DAS VARIÁVEIS** que mede ou observa. Lembrando que elas podem ser classificadas em variáveis dependentes (aquelas que você está tentando entender ou prever) e variáveis independentes (aquelas que você acha que podem influenciar a variável dependente).

## MODELOS DE TESTE DAS VARIÁVEIS

**REGRESSÃO** – A regressão é uma ferramenta importante para modelar relações entre variáveis. Por exemplo, você pode usar a regressão para entender como uma variável independente afeta uma variável dependente.

**TESTES DE HIPÓTESES** - Os testes de hipóteses ajudam a determinar se as diferenças observadas nos dados são estatisticamente significativas ou se podem ser atribuídas ao acaso.

### DAÍ É CONTIGO!

A interpretação dos resultados estatísticos é crucial. Os resultados não são apenas números, mas informações que ajudam a fazer afirmações sobre o mundo real. É importante entender o contexto biológico e a relevância prática dos resultados e **SE ELAS FAZEM SENTIDO**.



## AGORA NÓS VAMOS FAZER UM EXERCÍCIO DE CONFECÇÃO DE EXPERIMENTO LABORATORIAL CONTROLANDO VARIÁVEIS



## VAMOS PENSAR NA SEGUINTE SITUAÇÃO

AS QUEIMADAS NA AMAZÔNIA E NO PANTANAL TÊM SE INTENSIFICADO A CADA ANO. QUANDO A BIOMASSA É QUEIMADA, MODIFICA A QUALIDADE DO AR. A COMPOSIÇÃO DO MATERIAL PARTICULADO NA ATMOSFERA PODE ALTERAR A QUALIDADE DA ÁGUA DA CHUVA

SERÁ, ENTÃO, QUE A ÁGUA DA CHUVA PODE AFETAR

O CRESCIMENTO DE HORTALIÇAS?

### PARA FAZER ESTA ANÁLISE, DEVEMOS FAZER UM *DESIGN* EXPERIMENTAL QUE POSSA AUXILIAR NA RESPOSTA!

1. Já tenho meu problema.
2. Já tenho minha pergunta.
3. Quais são as variáveis que podem influenciar minha resposta?

A amostragem que vou fazer é de água. Desse modo, já tenho o meu objeto amostral – ÁGUA. Mas, quais são as variáveis que vou analisar em relação a água?

SABE-SE QUE UMA DAS VARIÁVEIS QUE PODE SER AFETADA É O PH, DEVIDO À ACIDEZ, POR EXEMPLO!

Outro exemplo é a presença de mercúrio devido a queima da biomassa e fácil volatilização do elemento.

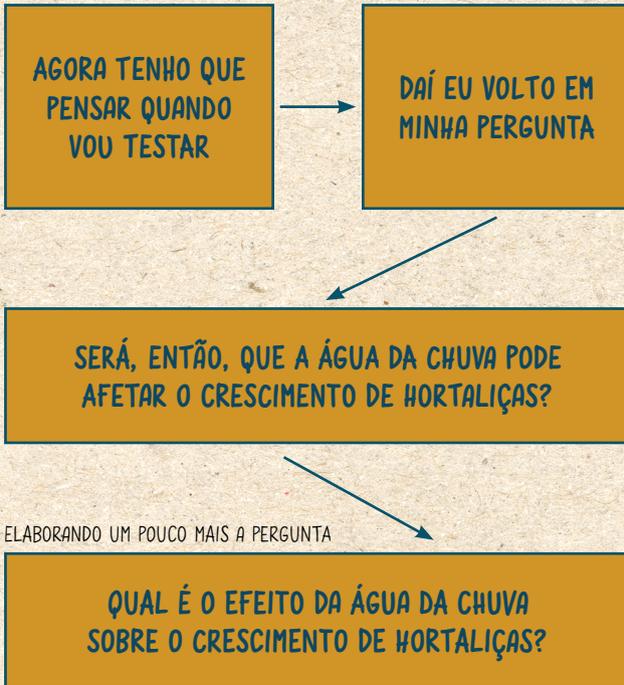


Agora é partir para a coleta? Bom, veja adiante que tem que esperar a planta brotar. E se chover antes? É por isso que o experimento não é fácil. Lembre-se que o pH modifica após 24h. Além disso, lembre-se que a coleta de água de chuva deve ser feita em um recipiente limpo, esterilizado, neste caso, a utilização de caixa d'água seria viável para garantir a irrigação durante todo o período amostral. Garanta a coleta da água da primeira chuva.

**DIFÍCIL, NÃO É? POIS É!**

**SUGESTÃO: OLHE NA METEOROLOGIA PARA  
PREVER O INÍCIO DO EXPERIMENTO!**

Agora eu tenho também que pensar no tipo de hortaliça que vou analisar. Tem que ser uma hortaliça de crescimento rápido (rúcula). A rúcula brota em 5 dias em média.



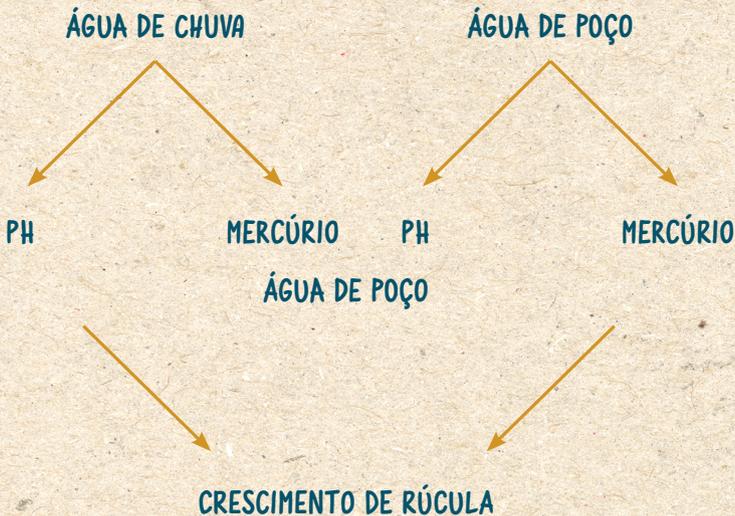
Vejo, aqui, que na minha pergunta eu esclareci que é o crescimento que irei analisar. Então, meu experimento será realizado a partir do brotamento da planta.

Meu design experimental deverá ter também uma fonte de comparação. Pois, imagine só eu dizendo que a água da chuva é boa (ou ruim) para o crescimento de plantas.

**A PESSOA QUE ME OUVIR DIZENDO ISSO  
VAI PERGUNTAR EM RELAÇÃO A QUÊ????**

Pois é. Para isso, deve-se haver o ponto controle, ou seja, algo que eu possa comparar se minha variável possui, realmente, uma relação que possa garantir a robustez do meu resultado.

Neste caso, eu tenho que saber que tipo de água os horticultores utilizam para cultivar as plantas. Sabe-se que a água de poço é muito utilizada. Então, também terei que utilizar a água de poço em meu ponto controle.



## BELEZA PURA!

### AGORA EU VOU PARTIR PARA MONTAR MEU EXPERIMENTO

Sempre tenho que partir do princípio que tenho instrumentos suficientes para a coleta de dados, bem como tenho que contar com um local apropriado para a instalação deste. Lembre-se de algumas questões que podem influenciar seus resultados.

#### 1. TIPO DE SUBSTRATO

Onde vou plantar as rúculas? O tipo de substrato tem que ser o mesmo para todas as plantas, a única variável diferente tem que ser a água recebida.

#### 2. LOCAL DE PLANTAÇÃO (ILUMINAÇÃO, TEMPERATURA, CORRENTE DE VENTO, TELHADO PARA NÃO TER INFLUÊNCIA DE CHUVAS)

Pense no seguinte – se eu deixar as rúculas todas enfileiradas e decidir que as do lado direito receberão água de chuva, e as do esquerdo receberão água de poço, eu entro em um viés que pode ser influenciado pela quantidade de sol que bate em um local, maior ou menor, do que em outro. Por isso é importante ALEATORIZAR meu experimento. Para isso, coloque números em potinhos e sorteie. Sortear quem? Claro, boa resposta! As réplicas.

#### 3. NÚMERO DE RÉPLICAS

Quantas réplicas seriam necessárias para fazer o experimento? Imagine que um horticultor tenha sua renda toda voltada para a venda de rúculas. Se ele plantar uma rúcula, há a possibilidade dela morrer, certo? Se ele plantar 100, é claro que algumas vão morrer, mas não todas. Então, sei que 100 são muitas plantas e, talvez, não haja reagentes suficientes no laboratório para tal análise, mas para 10 há. Somente como sugestão. Entretanto, podem ser mais réplicas.

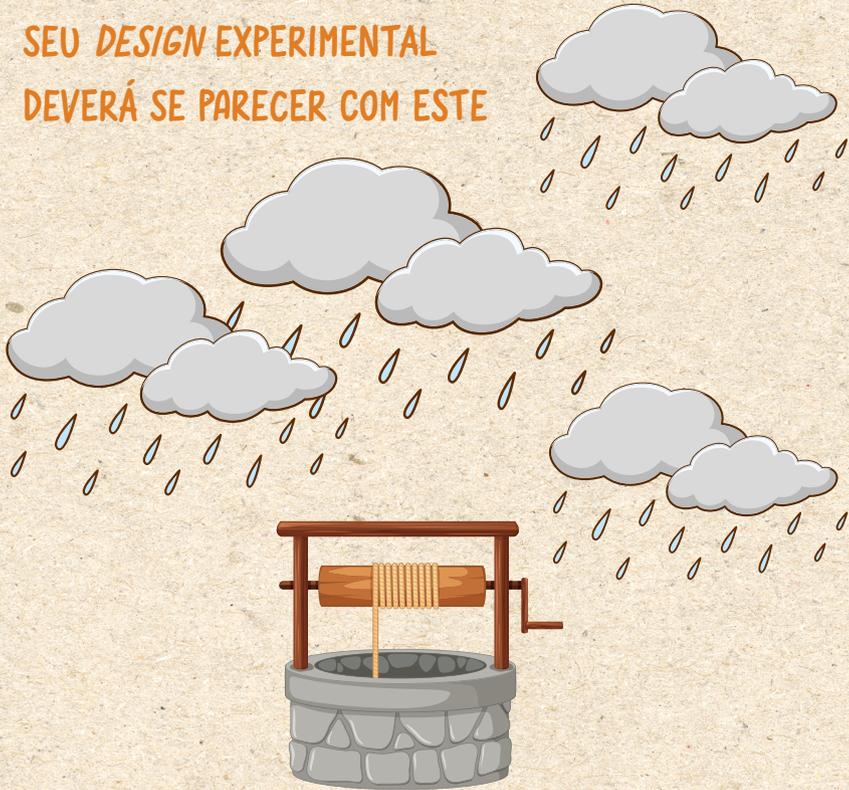
#### 4. FREQUÊNCIA DE COLETAS

Quantas vezes você irá medir as plantas? Todos os dias? Lembre-se que a planta tem um crescimento acelerado e, que está quase pronta para ser consumida após 20 dias. Então, para aumentar a robustez de seu experimento, sugiro que mensure todos os dias. Mensurar o que? O crescimento. Para isso, é necessário somente uma fita métrica, para medir a distância do chão até a folha mais alta. Também pode levar em consideração medir outras variáveis, como número de folhas, coloração, entre outras.

**DEVO MEDIR SÓ A PLANTA? NÃO! MESMO QUE SEJA TUDO CONTROLADO, ELAS ESTEJAM ALEATORIZADAS EM UM MESMO AMBIENTE, DEVE-SE MEDIR AS VARIÁVEIS AMBIENTAIS (EXEMPLO – TEMPERATURA, UMIDADE, LUMINOSIDADE, ETC.)**



SEU DESIGN EXPERIMENTAL  
DEVERÁ SE PARECER COM ESTE



20 plantas, plantadas no mesmo dia e já com brotamento visível. Sendo que 10 dessas serão utilizadas para mensurar o efeito da água da chuva, e outras 10 para a água de poço. Lembre-se de coletar água suficiente para analisar o pH e o mercúrio. Para todas as plantas faça a coleta antes de irrigar. Para algumas análises, 50ml é o suficiente.

## AGORA ALEATORIZE E COMECE O EXPERIMENTO



Irrigue sempre na mesma hora. Colete os dados de crescimento da planta em mesma hora, bem como os ambientais. Analise a água utilizada para a irrigação todos os dias.

Insira os dados na planilha do Excel, e depois faça os testes estatísticos como dito nos exemplos anteriores.



## VEJA O SEGUINTE!

Este exemplo fornecido no exercício anterior (crescimento de rúculas) pode ser modificado para qualquer outro modelo experimental.

Utilização de cinzas das queimadas para a análise do crescimento de hortaliças; análise do crescimento vegetal em virtude de níveis de inundação; análise do efeito da temperatura no crescimento de plantas; análise do efeito da temperatura no comportamento alimentar de lagartos em aquário, ou peixes, ou insetos.

Tem bastante. Dentre as ideias, as próximas páginas serão apresentadas as dúvidas que diversos alunos de biologia tiveram para a configuração de suas perguntas para o *design* experimental.



## EXEMPLO 1

A couve-manteiga (*Brassica oleracea*) pode ser usada como um bioindicador para determinar a saúde e qualidade do solo, levando em consideração sua germinação e, se sim, os sintomas observados nas plantas podem ser atribuídos a contaminações específicas ou deficiências no solo?



Neste exemplo deve-se pensar em quais são as variáveis indicativas de boa qualidade do solo. Uma delas é a diversidade microbiana, por exemplo. Outra variável é a redução da acidez, ou redução da alcalinidade do solo. Para o desenvolvimento deste experimento deve-se ter muitas réplicas (pelo menos 10).

## EXEMPLO 2

Como as mudanças ambientais, o desmatamento e a alteração no ecossistema influenciaram a ocorrência de acidentes com animais peçonhentos no estado de Mato Grosso, no período de 2010 a 2021?

Para esta pergunta deve-se levar em consideração o seguinte:

1. Por que este período? O que houve de tão interessante neste período que você quer analisar nele e não entre 2009 a 2022? E não entre 2011 a 2021, ou 2012 a 2020?
2. Como você irá tirar a influência de outras variáveis como o aumento do número de pessoas indo para o ambiente rural?

Esta análise deve ser feita utilizando dados secundários do sistema de saúde. Assim, poderá comparar acidentes em ambientes rurais e urbanos, por exemplo.



### EXEMPLO 3

Qual o percentual de adolescentes entre a faixa etária de 10 a 14 anos residentes no estado de Mato Grosso que se tornaram mãe durante os anos de 2011 a 2021?

PARA A ANÁLISE DESTA PERGUNTA, DEVE-SE LEVAR EM CONSIDERAÇÃO O PERÍODO. LEMBRE-SE QUE ANOS SÃO SOMENTE NÚMEROS. POR QUE NÃO ANALISAR DE 2010 A 2021? POR QUE NÃO ENTRE 2012 A 2022? ESTES MOTIVOS SÃO IMPORTANTES PARA A ROBUSTEZ DO TRABALHO. AINDA, VEJA QUE A PERGUNTA SE LIMITA AO PERCENTUAL. ESTE PERCENTUAL PODERÁ SER VERIFICADO COM DADOS SECUNDÁRIOS DO SISTEMA DE SAÚDE, MAS AINDA NÃO REFLETE A UMA PERGUNTA QUE GERE UMA PESQUISA.

Para isso, ela deve ser melhor elaborada, como por exemplo – Qual foi o efeito dos programas de governo no índice de gravidez na adolescência na cidade de Cáceres/MT? Para isso terá que analisar quais (e quando) foram os programas de governo de prevenção a gravidez realizados na última década, bem como correlacionar isso com o número de adolescentes grávidas.

### EXEMPLO 4

Qual foi a variação anual da precipitação na cidade de Cáceres, entre os anos de 2000 a 2013?

VOLTAMOS À QUESTÃO DO PERÍODO, POR QUE ESTE PERÍODO?

A pergunta é relevante, mas deve-se atentar ao motivo de escolha dos anos.

## EXEMPLO 5

Quais predadores atuam como controlador biológico de ninhôs de *Podocnemis unifilis* no Rio Paraguai?

**PARA ESTA ANÁLISE, UMA OBSERVAÇÃO DEVERÁ SER FEITA EM DIVERSOS AMBIENTES ONDE A ESPÉCIE OCORRA!**

A análise dos dados poderá ser feita de forma a simplesmente descrever quem são os predadores, ou, então, uma análise estatística de que tipo de local há mais ou menos predadores. Para uma análise estatística, lembre-se da necessidade de haver réplicas.



## EXEMPLO 6

Qual a relação entre a prevalência de geo-helmintos em áreas recreativas e as variações da temperatura e umidade do solo?

Nesta pergunta pode-se realizar uma regressão para analisar o número de indivíduos coletados com a temperatura e umidade, para isso são necessárias várias réplicas, ou de horários, ou de locais.

## EXEMPLO 7

Qual a influência de diferentes práticas de manejo do solo no fluxo de  $\text{CO}_2$ , considerando as variáveis ambientais?

Neste caso, não ficou claro quais são as variáveis ambientais, mas podem ser o tipo de vegetação, ou a presença de queimadas, ou diferente tipo de solo. Para a análise, uma ANOVA pode ser utilizada para comparar diferentes práticas de manejo.



## EXEMPLO 8

Como a sequência de DNA de um gene influencia a estrutura tridimensional da proteína que ele codifica?



**PARA ESTA ANÁLISE, DEVERÁ SER UTILIZADO O MODELO COMPUTACIONAL PARA BIOINFORMÁTICA**

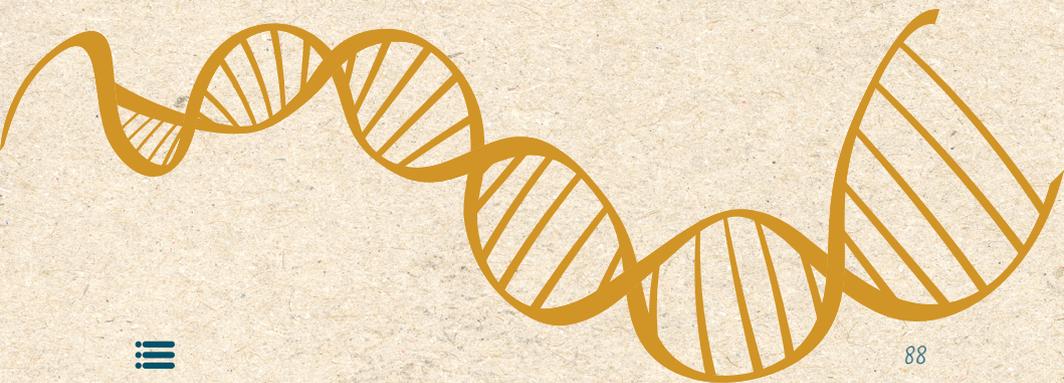
Para a análise deve-se utilizar a linguagem R ou Python, por exemplo. Esta análise deverá levar em consideração variáveis de entrada, como a sequência de DNA apresentada para que possa correlacionar com o tipo de proteína codificada.

## EXEMPLO 9

Como as características fisiológicas das briófitas se relacionam com sua eficácia como bioindicadores de poluentes atmosféricos em ambientes urbanos?

**PARA ESTA ANÁLISE, PODE SER OBSERVADA, POR EXEMPLO, A FORMAÇÃO DE MICRONÚCLEOS NESTES ORGANISMOS**

Ou seja, quanto maior é a incidência de poluente ambiental, maior será a formação de micronúcleos. Para isso, uma regressão linear poderá auxiliar na interpretação dos resultados.



## EXEMPLO 10

*Qualea grandiflora* é árvore ou arbusto?

Para a resposta a esta pergunta deve-se levar em consideração o que a teoria diz sobre a caracterização de um organismo quanto a ser árvore ou arbusto.

Munido desta qualificação, uma coleta em um grande número de indivíduos poderá auxiliar na resposta. Veja que as características deverão ser plotadas em Excel, em uma lista.

Nesta lista, o pesquisador deverá ir a campo para conferir o enquadramento dos indivíduos.



## ALGUMAS QUESTÕES A SEREM LEVADAS EM CONSIDERAÇÃO PARA A REALIZAÇÃO DE TESTES ESTATÍSTICOS, FORMULAÇÃO DE PERGUNTAS E ETC.

1. **TUDO INFLUENCIA TUDO!** Ou seja, quando perguntamos se há influência de uma variável sobre outra, a resposta quase imediata é que há. Entretanto, saber o motivo desta influência é que é o papel do pesquisador.
2. **PROCURAR SABER SOMENTE POR SABER NÃO É PESQUISAR, É ESPECULAR.** Então tenha sempre um motivo para a realização da sua pesquisa;
3. **VEJA O MÉTODO A SER EMPREGADO.** Se no método também consta que são necessárias amplas infraestruturas das quais você não possui – ou seja, inviabilizando sua pesquisa.
4. **TENHA CUIDADO COM A RESPOSTA.** Toda resposta precisa de muita literatura lida anteriormente. Não caia nos erros de dizer que algo é, enquanto ele, na verdade, não é.

## PRA CABAR

À medida que encerramos nossa jornada pelo fascinante mundo da “Modelagem e Experimentação em Biologia”, somos compelidos a refletir sobre a importância monumental desses elementos no avanço da ciência biológica. Através deste livro, exploramos como a coleta, análise e interpretação de dados, tanto qualitativos quanto quantitativos, são pilares fundamentais para o entendimento dos complexos sistemas biológicos e fenômenos naturais.

O coração da biologia moderna, como demonstrado, reside na habilidade de modelar o mundo vivo e experimentar dentro desses modelos. Esta abordagem não apenas amplia nosso conhecimento, mas também, abre portas para inovações e descobertas. A integração de dados qualitativos e quantitativos, discutida extensivamente, ressalta a necessidade de uma visão holística em pesquisa, onde diferentes tipos de dados se entrelaçam para formar uma compreensão mais completa e precisa da vida.

Este guia não é apenas um testemunho ao estado atual da modelagem e experimentação em biologia, mas também, um convite para olhar para o futuro. Os desafios que enfrentamos hoje - mudanças climáticas, doenças emergentes, perda de biodiversidade - exigem uma compreensão mais profunda e integrada da biologia. A modelagem e experimentação, como demonstrado, são fundamentais nesta busca, oferecendo ferramentas e metodologias para abordar questões complexas e urgentes.

Este guia é um chamado à ação para estudantes, educadores e profissionais da área. A biologia é um campo em constante evolução e, permanecer atualizado com as últimas técnicas e descobertas é crucial. Encorajamos a todos que embarcaram nesta jornada conosco a continuar explorando, aprendendo e contribuindo para este campo dinâmico e empolgante. Que este guia sirva não apenas como um recurso educacional, mas também, como uma inspiração para futuras descobertas e inovações no vasto e intrigante campo da biologia.

## REFERÊNCIAS

AZEVEDO, Paulo Roberto. Medeiros de. **Introdução à estatística** [recurso eletrônico]. 3. ed. Natal: EDUFRN, 2016. Disponível em: <https://repositorio.ufrn.br/bitstream/123456789/21298/2/Introduc%CC%A7a%CC%83o%20a%CC%80%20Estati%CC%81stica%20%28digital%29.pdf>. Acesso em: 20 mar. 2024.

BARROS NETO, B.; SARMINIO, I. S.; BRUNS, R. E. **Planejamento e otimização de experimentos**. Campinas: Editora Unicamp, 1995.

COSTA, Fabrício Martins da. **Estatística**. Belém: UEPA/Centro de Ciências Sociais e Educação, 2011. Disponível em: [https://ccse.uepa.br/downloads/material\\_2010/LIVRO\\_ESTADISTICA.pdf](https://ccse.uepa.br/downloads/material_2010/LIVRO_ESTADISTICA.pdf). Acesso em: 20 mar. 2024.

LEGENDTRE, P.; LEGENDRE, L. **Numerical ecology**. 3. ed. Amsterdam: Elsevier, 2012. v. 24.

MORETTIN, Pedro Alberto; BUSSAB, Wilton de Oliveira. **Estatística básica**. 9. ed. São Paulo: Saraiva, 2017. Disponível em: <https://archive.org/details/estatistica-basica-9ed-bussab-e-morettin/page/n19/mode/2up>. Acesso em: 20 mar. 2024.

PORTELLA, Augustus Caesar Franke; NASCIMENTO, Ildon Rodrigues do; ALVES, Anatórcia Ferreira; SCHEIDT, Gessiel Newton. **Estatística básica para os cursos de ciências exatas e tecnológicas**. Palmas: EDUFT, 2015. Disponível em: <https://repositorio.uft.edu.br/bitstream/11612/1434/1/Estat%C3%ADstica%20B%C3%A1sica.pdf>. Acesso em: 20 mar. 2024.

SAMPAIO, Nilo Antônio de Souza; ASSUMPÇÃO, Alzira Ramalho Pinheiro de; FONSECA, Bernardo Bastos da. **Estatística descritiva**. Belo Horizonte: Poisson, 2018. Disponível em: [https://www.poisson.com.br/livros/estatistica/volume1/Estatistica\\_Descritiva.pdf](https://www.poisson.com.br/livros/estatistica/volume1/Estatistica_Descritiva.pdf). Acesso em: 20 mar. 2024.

SANTIAGO, Genário Sobreira; PAIVA, Rui Eduardo Brasileiro. **Bioestatística** [recurso eletrônico]. 2. ed. Fortaleza: EdUECE, 2015. (Ciências Biológicas). Disponível em: [https://educapes.capes.gov.br/bitstream/capes/431710/2/Livro\\_Bioestatistica.PDF](https://educapes.capes.gov.br/bitstream/capes/431710/2/Livro_Bioestatistica.PDF). Acesso em: 20 mar. 2024.

## SOBRE OS AUTORES

### ERNADES SOBREIRA OLIVEIRA JUNIOR

Possui graduação em Licenciatura Plena em Ciências Biológicas pela Universidade do Estado de Mato Grosso (2007), mestrado em Ecologia e Conservação da Biodiversidade pela Universidade Federal de Mato Grosso (2010) e doutorado em Ciência pela Radboud University / Nijmegen, Holanda (2018), onde investigou o impacto de modificadores ecossistêmicos nos fluxos de gases de efeito estufa. Atualmente, desenvolve pesquisas nas regiões do Pantanal, Cerrado e Amazônia, focando na dinâmica das águas, gases de efeito estufa e mudanças climáticas. Também conduz projetos interdisciplinares que exploram as interações entre ar e água e seus impactos na saúde humana. Tem atuado fortemente em Extensão Universitária e Divulgação Científica, conectando as áreas de meio ambiente, tecnologias, educação, ensino e indígenas. Tem experiência em Ecologia, Limnologia e Transformação Ecológica, promovendo a inserção social por meio da pós-graduação. Pai do Davi e da Sara.

*Lattes:* <http://lattes.cnpq.br/7461346615427709>

*Orcid:* <https://orcid.org/0000-0002-6953-6917>

*E-mail:* [ernandes@unemat.br](mailto:ernandes@unemat.br)

### BRUNO ORMONDE LIMA DE OLIVEIRA

Tem formação técnica em Meio Ambiente pela Politécnica Antonio Luiz Pedrosa (2015) e em Petroquímica pela Universidade Estácio de Sá (2016). É tecnólogo em Gestão Ambiental pela Universidade Estácio de Sá (2018) e graduado em Ciência da Computação pela Universidade do Estado de Mato Grosso (2024). Atualmente é professor interino no Curso de Ciência da Computação da Universidade do Estado de Mato Grosso e possui vínculo como pesquisador do Núcleo de Saúde Digital do Estado de Mato Grosso, atuando com o desenvolvimento de Interfaces e softwares para a saúde. É sócio fundador da Abacate Digital, uma software house com uma linha de desenvolvimento focada em projetos de pesquisa que buscam no software

93uma ferramenta transformadora para seu contexto de aplicação. Destacou-se ao conquistar o primeiro lugar na primeira edição do HackaPantanal e ao ser campeão Global na Huawei ICT Competition, em Shenzhen na China, na trilha: Computação. Tem ampla experiência na área de desenvolvimento de software, com especial foco em interfaces amigáveis (UI/UX), Saúde Digital, gestão de times e gerenciamento de projetos.

*Lattes:* <http://lattes.cnpq.br/2124204689044925>

*Orcid:* <https://orcid.org/0009-0007-5576-5544>

*E-mail:* [brunoormonde@outlook.com](mailto:brunoormonde@outlook.com)

### WILKINSON (KIWI) LOPES LÁZARO

Graduado em Ciências Biológicas e Mestre em Ciências Ambientais pela Universidade do Estado de Mato Grosso. Doutor em Ecologia pela Universidade Federal do Rio de Janeiro.

Atualmente é pesquisador do Centro de Estudos em Limnologia, Biodiversidade e Etnobiologia do Pantanal da Universidade do Estado de Mato grosso. Tem experiência na área de Ecologia, com ênfase em Ecologia Aplicada, atuando principalmente nos seguintes temas: biogeoquímica de poluentes metálicos e organo-metálicos em ambientes aquáticos, uso de traçadores em estudos de processos de produção de MeHg em áreas alagadas, limnologia, ecologia de algas perifíticas, ecologia de macrófitas aquáticas, ecologia de comunidades icticas, estatística bayesiana/popperiana, desenho amostral, programação em R e Matlab. Membro do Grupo Conceitos Ecológicos e Etnológicos Aplicados a Conservação da Água e da Biodiversidade do Pantanal. Pai do Miguel, professor de biologia, e, nos bons dias, pesquisador. Me digladio com estatística desde 2010. Fotógrafo e cozinheiro nas horas vagas.

*Lattes:* <http://lattes.cnpq.br/0263718697915954>

*Orcid:* <https://orcid.org/0000-0002-6499-6631>

*E-mail:* [wilkinson.lobes@unemat.br](mailto:wilkinson.lobes@unemat.br)

A OBRA *MODELAGEM E EXPERIMENTAÇÃO EM BIOLOGIA* APRESENTA-SE COMO UM GUIA PRÁTICO, ACESSÍVEL E DIDÁTICO VOLTADO AO ENSINO E À APLICAÇÃO DA ESTATÍSTICA NO CONTEXTO DAS CIÊNCIAS BIOLÓGICAS. COM LINGUAGEM CLARA E EXEMPLOS CONTEXTUALIZADOS, O LIVRO BUSCA ATENDER ÀS PRINCIPAIS DIFICULDADES ENFRENTADAS POR ESTUDANTES DE GRADUAÇÃO AO LIDAR COM O DESENHO EXPERIMENTAL, ANÁLISE ESTATÍSTICA E MODELAGEM DE DADOS. ORGANIZADO EM OITO SEÇÕES, O GUIA ABORDA DESDE A CONCEITUAÇÃO DE DADOS (QUALITATIVOS E QUANTITATIVOS) ATÉ A CONSTRUÇÃO DE PERGUNTAS CIENTÍFICAS E FORMULAÇÃO DE HIPÓTESES, PASSANDO PELA COLETA DE DADOS, USO DE RÉPLICAS, TESTES ESTATÍSTICOS (COMO TESTE T E ANOVA), ATÉ TÉCNICAS BÁSICAS DE REGRESSÃO LINEAR. TAMBÉM PROPÕE EXERCÍCIOS PRÁTICOS E EXPERIMENTOS LABORATORIAIS, COM FOCO EM VARIÁVEIS AMBIENTAIS E BIOLÓGICAS, INCENTIVANDO O PENSAMENTO CIENTÍFICO E A TOMADA DE DECISÃO COM BASE EM EVIDÊNCIAS. A OBRA UTILIZA O EXCEL COMO FERRAMENTA DE APOIO PARA ANÁLISES, VALORIZANDO SUA APLICABILIDADE E SIMPLICIDADE NO ENSINO DE ESTATÍSTICA. AO FINAL, APRESENTA EXEMPLOS REAIS DE PERGUNTAS DE PESQUISA E DESENHOS EXPERIMENTAIS QUE DIALOGAM COM TEMAS AMBIENTAIS E DE SAÚDE PÚBLICA. VOLTADO A ESTUDANTES, PROFESSORES E PESQUISADORES INICIANTE, O LIVRO CONTRIBUI PARA O FORTALECIMENTO DO PENSAMENTO CRÍTICO-CIENTÍFICO, PROMOVEDO A AUTONOMIA NA INVESTIGAÇÃO BIOLÓGICA COM BASE EM FUNDAMENTOS ESTATÍSTICOS SÓLIDOS.

