

**ANÁLISE DE DESEMPENHO DE
CORREDORES UTILIZANDO DADOS DO
STRAVA: UMA ABORDAGEM DE
REGRESSÃO LINEAR.**

Uberlândia-MG

2024



Universidade Federal de Uberlândia
Faculdade de Matemática

Bacharelado em Estatística

**ANÁLISE DE DESEMPENHO DE
CORREDORES UTILIZANDO DADOS
DO STRAVA: UMA ABORDAGEM DE
REGRESSÃO LINEAR.**

João Marcos de Souza Matos

Uberlândia-MG

2024

João Marcos de Souza Matos

**ANÁLISE DE DESEMPENHO DE
CORREDORES UTILIZANDO DADOS
DO STRAVA: UMA ABORDAGEM DE
REGRESSÃO LINEAR.**

Relatório com resultados parciais do Trabalho de Conclusão de Curso (TCC1) apresentado na Faculdade de Matemática da Universidade Federal de Uberlândia como requisito básico para a conclusão do Curso de Estatística.

Orientador: Pedro Franklin Cardoso Silva

**Uberlândia-MG
2024**

SUMÁRIO

1	Introdução	2
2	Justificativa	3
3	Objetivos	4
3.1	Objetivo geral:	4
3.2	Objetivos específicos:	4
4	Metodologia	5
5	Cronograma	6
6	Cronograma	7
	Referências	8

1 INTRODUÇÃO

Os cuidados com a saúde e a prática de esportes têm se tornado, nos últimos anos, hábitos essenciais na vida das pessoas. Dentre os esportes mais praticados destaca-se a corrida (STRAVA, 2023)[1]. Buscando maior contato com a natureza e seus benefícios, corredores têm explorado novas modalidades ao ar livre, dentre elas está a prática de montanhismo (FRANÇA et al, 2021)[2], que se insere no mundo esportivo com as seguintes modalidades: Corrida de Montanha(MC), o Trail Running(TR), o Skyrunnig, entre outros.

A obtenção de melhores resultados em corridas depende de alguns fatores, como por exemplo, a capacidade do atleta de controlar a intensidade do esforço ao longo do percurso, adaptando-se aos diferentes tipos de terrenos encontrados. Essa regulação eficiente é determinada pela estratégia de ritmo por quilômetro (pacing) adotada, sendo crucial para equilibrar a utilização dos recursos energéticos e retardar a fadiga muscular.

Dentre outros fatores, as características ambientais do percurso, como inclinação, tipo de superfície e horário de largada, influenciam significativamente nos resultados da prova. O desnível vertical total do percurso, que engloba as alturas acumuladas nas subidas e descidas, é um fator crucial a ser considerado, pois impacta diretamente no esforço requerido e no desempenho do corredor em diferentes trechos da corrida [3].

Tais informações e dados dos corredores são encontradas em serviços de rastreamento de exercícios físicos, dentre eles destaca-se o Strava, uma plataforma esportiva que registra os dados dos dispositivos de monitoramento pessoal de cada atleta. Neste serviço, é possível obter informações pessoais dos atletas e fatores como, pacing, frequência cardíaca, ganho de elevação, distância percorrida, tempo de movimentação, dentre outros, captados durante atividades (STRAVA, 2024) [4].

Previamente, será possível obter conclusões apenas analisando as informações adquiridas dos corredores, entretanto, visando identificar impactos mais profundos, será adotada uma abordagem estatística utilizando a regressão linear múltipla como principal ferramenta no estudo, com propósito de quantificar o impacto de cada variável no desempenho do corredor, bem como, estimar o valor de uma variável dependente modelando a relação com uma ou mais variáveis independentes[5].

Devido a discrepância entre o alto crescimento de adeptos à modalidade de TR e a baixa quantidade de estudos envolvendo esse estilo de corrida, considerando aspectos de provas reais, fica evidente que o tema tem alto potencial para ser abordado. Em síntese, o objetivo deste trabalho é analisar fatores que influenciam no desempenho final atletas em prova de TR.

2 JUSTIFICATIVA

Ao observar a lacuna existente na exploração acadêmica ao que se diz respeito a aplicação de técnicas estatísticas envolvendo temas relacionados ao desempenho de atletas, notou-se a viabilidade de elaborar um projeto de pesquisa com ênfase em corredores, sendo eles amadores ou profissionais.

Entender o desempenho dos corredores é fundamental porque fornece ideias valiosas não apenas para os próprios atletas, mas também para treinadores, profissionais de saúde esportiva e até mesmo pesquisadores da área. Ao compreender as variáveis que afetam o desempenho, como ritmo, desnível do percurso, entre outras, podemos identificar estratégias mais eficazes de treinamento e competição, maximizando o potencial de cada corredor.

Além disso, analisar essas variáveis não apenas fornece uma compreensão mais profunda do desempenho individual, mas também contribui para o desenvolvimento de métodos de treinamento mais personalizados e eficazes. Por exemplo, um corredor pode se beneficiar de ajustes em sua estratégia de pacing com base em suas características físicas e objetivos de desempenho. Da mesma forma, conhecer o impacto do desnível do percurso pode ajudar os atletas a planejar suas estratégias de corrida de forma mais eficiente, adaptando-se melhor às condições específicas da prova.

Quanto ao uso da regressão linear, essa técnica estatística é uma ferramenta poderosa para modelar e compreender as relações entre variáveis independentes e dependentes. No contexto da análise do desempenho dos corredores, a regressão linear nos permite quantificar o impacto de cada variável sobre o desempenho geral do corredor, fornecendo uma base sólida para a tomada de decisões informadas no treinamento e na competição.

Portanto, utilizar a regressão linear neste estudo não apenas nos permite explorar essas relações de forma mais precisa, mas também nos ajuda a identificar padrões e tendências nos dados, fornecendo insights valiosos que podem ser aplicados de maneira prática para melhorar o desempenho e a saúde dos corredores. Assim, este estudo não só contribui para o avanço do conhecimento científico na área, mas também tem o potencial de impactar positivamente a comunidade de corredores, auxiliando na melhoria contínua de suas práticas de treinamento e competição.

3 OBJETIVOS

3.1 OBJETIVO GERAL:

O objetivo do presente estudo será aplicar modelos de regressão linear para analisar o desempenho de corredores de *skyrunning* utilizando dados do Strava.

3.2 OBJETIVOS ESPECÍFICOS:

1. Utilizar técnicas de web scraping através da linguagem de programação python para coletar e preparar uma base de dados, contendo informações como tempo, distância, altitude, ritmo, entre outros.
2. Explorar e analisar os dados coletados do Strava para identificar padrões e tendências no desempenho dos corredores, diferenças entre corredores de diferentes níveis de habilidade, entre outros aspectos.
3. Aplicar técnicas de regressão linear para modelar a relação entre variáveis independentes (como desnível, pacing, distância, frequência cardíaca) e dependente (desempenho do corredor), explorando como elas podem influenciar o desempenho dos corredores.
4. Interpretar os resultados da análise de regressão linear para identificar os principais fatores que influenciam o desempenho dos corredores.
5. Auxiliar assessorias esportivas e/ou corredores amadores a traçar estratégias para melhorarem o desempenho em uma prova de *skyrunning*.

4 METODOLOGIA

Neste trabalho de conclusão de curso será realizado um estudo descritivo. Iniciando-se pela extração dos dados de uma prova específica de *skyrunning* da plataforma de rastreamento de exercícios físicos Strava, através de técnicas de webscraping com a linguagem de programação python.

A população alvo do estudo consiste em corredores maratonistas de provas de skyrunning. Os dados coletados serão armazenados e organizados em um arquivo Comma Separated Values(CSV), em que as variáveis serão agrupadas por cada corredor participante da prova. Os participantes não serão identificados no conjunto.

A análise dos dados coletados se dará por meios de técnicas de estatísticas descritivas buscando compreender características e comportamentos dos corredores da amostra extraída, assim como identificar perfis de corredores através de técnicas de agrupamentos e detecção de perfil.

Ao ajustar o modelo de regressão linear multivariada utilizando como características o pacing, desnível, frequência cardíaca, entre outras, como variáveis independentes, será possível prever o tempo final de prova do atleta. Tendo como objetivo determinar fatores que influenciarão no desempenho do corredor.

Para finalizar, serão aplicadas técnicas de avaliação de modelos de regressão linear por meio da análises de resíduos e testes de significância dos coeficientes estimados buscando, assim, entender a relação entre as variáveis independentes e o desempenho dos corredores.

5 CRONOGRAMA

Atividades	Janeiro	Fevereiro	Março	Abril
Entrega do termo de compromisso	X			
Pesquisa e definição do tema	X			
Revisão bibliográfica	X	X	X	X
Elaboração do pré- projeto	X	X	X	
Entrega do pre-projeto			X	
Estudos sobre o assunto	X	X	X	X
Coleta de Dados (se for o caso)		X	X	
Análise preliminar dos dados (se for o caso)				X
Preparação do relatório com os resultados preliminares				X
Entrega do relatório com os resultados preliminares				X

6 CRONOGRAMA

Atividades	Janeiro	Fevereiro	Março	Abril
Entrega do termo de compromisso	X			
Pesquisa e definição do tema	X			
Revisão bibliográfica	X	X	X	X
Elaboração do pré- projeto	X	X	X	
Entrega do pre-projeto			X	
Estudos sobre o assunto	X	X	X	X
Coleta de Dados (se for o caso)		X	X	
Análise preliminar dos dados (se for o caso)				X
Preparação do relatório com os resultados preliminares				X
Entrega do relatório com os resultados preliminares				X

REFERÊNCIAS

- [1] STRAVA. Strava year in sport trend report: Insights on the world of exercise, 2024. acessado em 14/03/2024.
- [2] G. L. DE FRANÇA, D. L. DE FRANÇA, and A. F. CAREGNATO. Motivos da adesão em atividades de aventura na natureza e as influências na qualidade de vida de seus praticantes. *LICERE - Revista Do Programa De Pós-graduação Interdisciplinar Em Estudos Do Lazer*, 24(3):182–203, 2021. <https://doi.org/10.35699/2447-6218.2021.36311>.
- [3] D. L. BORGES. Dinâmica de tempo e pacing no campeonato brasileiro feminino de sky-running, 2023.
- [4] STRAVA. About, 2024. acessado em 14/03/2024.
- [5] HOFFMANN RODOLFO. *Análise de regressão: uma introdução à econometria*. Universidade de São Paulo. Escola Superior de Agricultura Luiz de Queiroz, 4 edition, 2016.