



Universidade Federal de Uberlândia
Faculdade de Matemática

Bacharelado em Estatística

**ANÁLISE DO COMPORTAMENTO DAS
MÉTRICAS DE SALTO AO LONGO DE
UMA TEMPORADA EM UM TIME DE
VOLEIBOL PROFISSIONAL FEMININO**

Raiane França Rosseti

Uberlândia-MG
2025

Raiane França Rosseti

**ANÁLISE DO COMPORTAMENTO DAS
MÉTRICAS DE SALTO AO LONGO DE
UMA TEMPORADA EM UM TIME DE
VOLEIBOL PROFISSIONAL FEMININO**

Trabalho de conclusão de curso apresentado à Coordenação do Curso de Bacharelado em Estatística como requisito parcial para obtenção do grau de Bacharel em Estatística.

Orientador: Leandro Alves Pereira

Uberlândia-MG
2025



**Universidade Federal de Uberlândia
Faculdade de Matemática**

Coordenação do Curso de Bacharelado em Estatística

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Estatística.

Uberlândia, 26 de setembro de 2025

BANCA EXAMINADORA

Leandro Alves Pereira

Ednaldo Carvalho Guimarães

Cristiano Lino Monteiro de Barros

**Uberlândia-MG
2025**

AGRADECIMENTOS

Gostaria de expressar meu profundo agradecimento à minha mãe, Flávia, e à minha avó, Marileide, por todo o apoio e pela confiança depositada em mim ao longo desta jornada. Agradeço, de forma especial, ao meu filho Miguel, pela compreensão, pelo companheirismo e pela motivação durante todo o processo. Estendo ainda minha gratidão à minha família itapolitana, que, mesmo distante, nunca deixou de estar presente para mim e para o Miguel, bem como aos meus amigos de Uberlândia, por tornarem meu caminho mais leve.

Registro também meu sincero agradecimento ao professor Leandro, pela orientação e dedicação durante o desenvolvimento deste trabalho. Obrigada, inclusive, por acolher o tema por mim proposto, mesmo sem saber se seria viável. Agradeço ainda ao professor Cristiano, por disponibilizar, com o aval da comissão, os dados necessários para a pesquisa.

RESUMO

Este trabalho analisa o comportamento das métricas de salto em jogadoras da equipe feminina de voleibol Dentil Praia Clube ao longo da temporada 2024/2025. A base de dados foi obtida por meio de dispositivos de monitoramento que registraram altura, data e horário de cada salto para cada atleta. A metodologia envolveu análise exploratória, técnicas de agrupamento hierárquico (método de Ward) e modelos de regressão múltipla, possibilitando a identificação de dois padrões de salto distintos: o tipo 0, de menor altura média e comportamento mais estável, e o tipo 1, de maior altura média, mas com tendência de declínio ao longo das partidas. Os resultados evidenciam diferenças individuais entre as atletas, destacando a importância da análise estatística como suporte à tomada de decisões técnicas e à preparação física personalizada.

Palavras-chave: voleibol, salto vertical, análise de desempenho, clusterização, regressão múltipla.

ABSTRACT

This study analyzes the behavior of jump metrics in players from the Dentil Praia Clube women's volleyball team throughout the 2024/2025 season. The database was obtained using monitoring devices that recorded the height, date, and time of each jump for each athlete. The methodology involved exploratory analysis, hierarchical clustering techniques (Ward's method), and multiple regression models, enabling the identification of two distinct jump patterns: type 0, with lower average height and more stable behavior, and type 1, with higher average height but a tendency to decline throughout the matches. The results highlight individual differences between athletes, emphasizing the importance of statistical analysis as a support for technical decision-making and personalized physical preparation.

Keywords: volleyball, vertical jump, performance analysis, clustering, multiple regression.

SUMÁRIO

Lista de Figuras	I
Lista de Tabelas	II
1 Introdução	1
2 Fundamentação Teórica	4
2.1 Salto Vertical (SV)	4
2.2 Análise de Agrupamento	4
2.2.1 Método Não Supervisionado	5
2.2.2 Distância Euclidiana	5
2.2.3 Método Hierárquico	5
2.2.4 Método de Ward	5
2.2.5 Dendrograma	5
2.3 Regressão múltipla	5
2.3.1 Variável regressiva dummy	6
2.4 Propriedade assintótica dos estimadores de máxima verossimilhança (EMV) . . .	6
3 Metodologia	8
3.1 Dados	8
3.2 Equipamento de Medição de Saltos	8
3.3 Tratamento e manipulação de dados	9
3.4 Análise Estatística	9
4 Resultados	11
4.1 Análise Exploratória	11
4.2 Agrupamento das Alturas de Salto	12
4.3 Modelagem do Desempenho	13
4.3.1 Resultados gerais	13
5 Conclusões	17
Referências Bibliográficas	18
Apêndice A Apêndice	19
A.1 Script R	19

LISTA DE FIGURAS

4.1	Alturas dos saltos das atletas (em centímetros) pela ordem de ocorrência em cada partida	11
4.2	Dendrogramas das alturas dos saltos das atletas registrados a cada partida	12
4.3	Curvas de desempenho de cada tipo de salto ao decorrer da partida para cada uma das atletas estudadas	13

LISTA DE TABELAS

4.1 Estatísticas descritivas por jogadora e tipo de salto (quantidade total de saltos, altura média, desvio-padrão e intervalo de confiança)	12
--	----

1. INTRODUÇÃO

Criado em 1895 por William G. Morgan, o voleibol passou por diversas transformações ao longo do tempo e rapidamente se difundiu em escala global. Atualmente, o voleibol está entre os cinco esportes mais praticados no cenário internacional, com mais de 220 federações nacionais filiadas à Federação Internacional de Voleibol (FIVB), conforme pode ser visto no site da FIVB^[3].

Do ponto de vista fisiológico, o voleibol é caracterizado por exigir elevada potência muscular do atleta, pois exige esforços repetitivos de alta intensidade intercalados por períodos de recuperação [6]. Para que o atleta tenha um bom desempenho é necessário desenvolver suas habilidades técnicas e táticas [10]. Ainda de acordo o autor, a parte técnica refere-se à capacidade do jogador de se adaptar às exigências do jogo, abrangendo a percepção das situações, a tomada de decisão e a execução das jogadas. Já o aspecto tático, está associado ao domínio técnico, à condição física e às habilidades cognitivas, motoras e de coordenação, permitindo que o atleta analise e reaja com precisão e eficiência aos desafios impostos durante a partida.

O alto rendimento no voleibol exige disciplina, dedicação e preparação intensa. A temporada envolve meses de treinos e jogos sob elevada carga física, técnica e tática, demandando acompanhamento multidisciplinar para equilibrar desempenho e recuperação, prevenindo sobrecargas e maximizando o rendimento nos momentos decisivos.

Durante os períodos de intensificação, acúmulo e variação aguda da carga de trabalho, a falta de controle adequado pode aumentar significativamente o risco de lesões [4]. Em esportes coletivos, as estratégias de treinamento são desenvolvidas para otimizar a performance competitiva, ao mesmo tempo em que reduzem o impacto do desgaste físico e minimizam o risco de lesões. Nesse contexto, a periodização ao longo da temporada torna-se fundamental para garantir que os atletas alcancem seu pico de desempenho nos momentos decisivos.

Ao longo de cada partida, os atletas são submetidos a esforços repetitivos de alta intensidade, que demandam predominantemente capacidade anaeróbica, além de uma contribuição moderada do sistema aeróbico para sustentação da performance. Os movimentos mais frequentes durante o jogo incluem sprints curtos, mergulhos, saltos e deslocamentos rápidos em múltiplas direções, exigindo dos jogadores elevado condicionamento físico, com ênfase em força, potência e agilidade [9].

Atualmente, o salto vertical (SV) é considerado um dos principais indicadores de desempenho no voleibol. Essa habilidade está presente em diversas ações do jogo, como ataques, bloqueios e levantamentos, exigindo explosão muscular e precisão. Dada a complexidade técnica

e tática da modalidade, as métricas relacionadas aos saltos como altura, frequência e variação ao longo da partida, tornam-se relevantes para avaliar o rendimento e o condicionamento físico das atletas em alto nível competitivo [9].

Nas últimas duas décadas, o voleibol passou por um crescimento significativo em nível técnico e competitivo, acompanhado de maiores investimentos para o esporte. Com isso, tornou-se indispensável a adoção de ferramentas analíticas no acompanhamento das partidas, adotando a análise de jogo como uma estratégia para a otimização do desempenho das equipes [10]. Essa prática envolve a coleta e interpretação de indicadores técnicos, táticos e físicos observados durante as competições, com o objetivo de orientar os treinamentos de forma mais eficiente e personalizada, contribuindo para o aprimoramento contínuo do rendimento dos atletas.

Com a consolidação da análise de jogo, a estatística tornou-se uma ferramenta indispensável para analisar tendências, mapear a distribuição de jogadas e avaliar a eficácia e eficiência dos jogadores. A análise de desempenho permite identificar os principais fatores que influenciam o rendimento esportivo, possibilitando o desenvolvimento de programas de treinamento personalizados, voltados para o aprimoramento das habilidades individuais. Além disso, essa abordagem desempenha um papel crucial na prevenção de lesões e na mitigação de efeitos adversos decorrentes de limitações previamente detectadas, promovendo um crescimento sustentável e duradouro [1].

A análise das métricas de salto no voleibol de alto rendimento representa uma possibilidade para o avanço da ciência do esporte, especialmente quando associada a métodos estatísticos. Com o uso crescente de tecnologias de monitoramento, torna-se possível explorar padrões de desempenho físico ao longo das partidas e da temporada e identificar padrões existentes. Nesse cenário, a estatística exerce papel fundamental ao permitir a organização, modelagem e interpretação precisa desses dados, apoiando decisões técnicas.

Por esse motivo, este estudo tem como objetivo analisar os dados de saltos da temporada 2024/2025 da equipe feminina do Dentil Praia Clube, buscando identificar padrões de comportamento ao longo das partidas. Especificamente, pretendeu-se verificar se havia diferentes padrões de saltos das jogadoras individualmente durante as partidas e também avaliar a possibilidade de modelar o desempenho de salto, de modo a compreender melhor seu comportamento em quadra.

Ao longo do trabalho, a fundamentação teórica foi organizada em quatro tópicos principais. O primeiro abordou o salto vertical, destacando sua relevância como indicador físico na análise de desempenho. O segundo tratou da análise de agrupamento, destacando conceitos, medidas de similaridade e critérios de formação dos clusters para a classificação dos diferentes tipos de salto. O terceiro discutiu a regressão múltipla utilizada para modelar o desempenho das atletas. Por fim, apresentou-se a estimativa por máxima verossimilhança, ressaltando suas propriedades de consistência e eficiência assintótica, que embasaram a aplicação do modelo estatístico proposto.

A metodologia contemplou quatro etapas principais. Inicialmente, descreveram-se os dados das atletas da equipe feminina do Dentil Praia Clube disponibilizados pela comissão técnica e

que foram utilizados no trabalho. Na sequência, apresentou-se o funcionamento do equipamento de medição, responsável por registrar a altura, data e horário de cada salto para cada uma das atletas monitoradas. Em seguida, discutiu-se o tratamento e manipulação da base, destacando os critérios de seleção de jogadoras, a organização da ordem dos saldos e a exclusão de valores nulos. Por fim, detalhou-se a análise estatística, que incluiu a etapa exploratória, a aplicação da análise de cluster, a construção da variável tipo de salto e o ajuste do modelo de regressão múltipla, sob a hipótese de normalidade assintótica dos estimadores.

No capítulo seguinte, foram apresentados os resultados da análise dos dados, com destaque para os modelos regressivos ajustados individualmente para cada atleta do estudo. Esses modelos permitem identificar possíveis padrões distintos de salto, bem como tendências de evolução ou queda no desempenho.

Por fim, concluiu-se que foi possível identificar e descrever matematicamente os padrões de salto, permitindo reconhecer quais atletas apresentaram maior ou menor perda de desempenho ao longo do tempo e, assim, auxiliaram nas decisões técnicas e, sobretudo, nas estratégias de preparação física individualizadas. Cabe destacar que este estudo não considerou as posições em que as atletas atuam, de modo que os resultados não deveriam ser utilizados para comparações diretas entre as jogadoras, apenas como referência individual. Ressaltou-se, ainda, que se tratou de uma abordagem inicial, passível de aprofundamento em pesquisas futuras.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 SALTO VERTICAL (SV)

Saltar é considerado um dos fenômenos motores fundamentais. Dentre os tipos existentes, destaca-se o salto vertical pelo seu amplo uso em avaliação de aptidão física e planejamento de treinamentos, conforme descrito pelo autor Michael Inácio dos Santos[8]. Segundo a obra, desde antes do século XX, os testes motores já eram utilizados para mensurar o condicionamento físico, tendo Dudley Allen Sargent como pioneiro, ao desenvolver os testes de salto vertical. Ainda de acordo com o livro, o salto vertical se destaca pela facilidade de aplicação do protocolo de avaliação e pela diversidade de desfechos possíveis. Esse indicador é amplamente empregado na avaliação do desempenho de atletas de alto rendimento, uma vez que reflete a potência muscular, componente essencial da aptidão física relacionada ao desempenho esportivo.

2.2 ANÁLISE DE AGRUPAMENTO

A análise de agrupamento, ou *clusterização*, é uma técnica multivariada utilizada para classificar objetos de acordo com seu grau de similaridade, definido a partir das características de interesse. O objetivo é formar grupos internamente homogêneos e externamente heterogêneos. Neste trabalho, empregou-se um método de aprendizado não supervisionado, adequado quando não há rótulos ou categorias previamente definidas.

Entre as motivações para a utilização de técnicas de agrupamento estão a possibilidade de sintetizar conjuntos complexos de variáveis em perfis representativos, facilitar interpretações e até mesmo subsidiar análises posteriores, como testes de hipóteses. No contexto dos saltos no voleibol, essa abordagem permite investigar padrões nas alturas registradas ao longo das partidas, especialmente considerando a ordem em que os saltos ocorrem.

Para conduzir a análise, alguns elementos foram definidos: a medida de similaridade entre os objetos, o procedimento de formação dos grupos e a escolha do número final de clusters. Idealmente, busca-se um número reduzido de grupos que preserve a homogeneidade interna sem perder a capacidade de distinguir padrões relevantes. Neste estudo, a medida de similaridade foi dada pela distância euclidiana, e os agrupamentos foram formados por meio do método hierárquico. A definição do número de clusters baseou-se na interpretação do dendrograma.

2.2.1 MÉTODO NÃO SUPERVISIONADO

O aprendizado não supervisionado busca identificar padrões ou estruturas latentes nos dados sem a necessidade de categorias previamente estabelecidas. Embora não exija rótulos, sua aplicação pressupõe indícios de que diferentes grupos possam existir, uma vez que qualquer conjunto de dados pode ser artificialmente particionado. Assim, sua utilidade depende da interpretação substantiva e do suporte teórico ou empírico que justifique a presença de grupos distintos.

2.2.2 DISTÂNCIA EUCLIDIANA

A distância euclidiana é uma das métricas utilizadas pelo método hierárquico. Ela é obtida através do cálculo das distâncias entre os valores amostrais nas diferentes variáveis de agrupamento e permite quantificar a proximidade ou afastamento dos objetos. Para o atual estudo, essa métrica foi aplicada à variável altura do salto, que é medida em centímetros.

2.2.3 MÉTODO HIERÁRQUICO

O método consiste em construir os agrupamentos iniciando com cada observação representando um grupo. A cada etapa, os agrupamentos mais semelhantes são unidos, formando um novo grupo. Esse procedimento é repetido sucessivamente até que todos os elementos estejam reunidos em um único agrupamento.

2.2.4 MÉTODO DE WARD

Para a definição dos grupos, utilizou-se o método de Ward, que tem como objetivo minimizar a variabilidade interna dos agrupamentos. Esse método baseia-se na soma dos quadrados entre dois grupos, considerando todas as variáveis.

2.2.5 DENDROGRAMA

Para selecionar o número mais adequado de clusters, utilizou-se a representação gráfica fornecida pelo dendrograma. Esse gráfico mostra, de forma hierárquica, como os grupos vão sendo formados a cada etapa do processo, até que todos os dados sejam reunidos em um único agrupamento.

2.3 REGRESSÃO MÚLTIPLA

A regressão múltipla é uma extensão da regressão linear simples em que duas ou mais variáveis explicativas (X_1, X_2, \dots, X_n) são consideradas simultaneamente para explicar ou prever a variável resposta (Y).

No presente trabalho, adotou-se um modelo que incorpora as variáveis de interesse e sua interação, com o objetivo de avaliar o desempenho das atletas selecionadas para o estudo. A forma geral do modelo para as duas variáveis é dada por:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \epsilon \quad (2.1)$$

em que Y é a variável resposta, X_1 e X_2 são as covariáveis do modelo, β_0 é o intercepto, β_1 e β_2 representam o efeito para cada uma das covariáveis respectivamente e β_3 indica o efeito da interação entre X_1 e X_2 no modelo. O termo ϵ representa o erro aleatório. No modelo utilizado, a variável X_2 é uma variável dummy.

Inicialmente, o modelo foi testado também com termos quadráticos (X_1^2 , X_2^2), a fim de captar possíveis relações de natureza curvilínea. No entanto, os resultados mostraram valores de coeficiente de determinação (R^2) muito próximos aos da regressão linear. Dessa forma, optou-se pelo modelo linear, pelo princípio da parcimônia.

2.3.1 VARIÁVEL REGRESSIVA DUMMY

A variável dummy, também chamada de variável indicadora, é utilizada para representar características qualitativas em modelos de regressão. Sua principal função é distinguir entre grupos ou indicar a presença ou ausência de determinada condição. Em geral, assume valores binários, sendo 0 para ausência e 1 para presença. Esse recurso permite incluir variáveis categóricas em modelos estatísticos originalmente voltados para variáveis numéricas, ampliando a capacidade de análise e interpretação[2].

2.4 PROPRIEDADE ASSINTÓTICA DOS ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA (EMV)

Sob condições de regularidade, os **Estimadores de Máxima Verossimilhança** (EMVs) apresentam consistência e eficiência assintótica. Formalmente, se X_1, X_2, \dots, X_n são *iid* e $\hat{\theta}$ é o EMV de θ , e $\tau(\theta)$ é uma função contínua de θ , então, para a função de densidade $f(X|\theta)$ e a função de verossimilhança $L(\theta|X)$, vale que:

$$\sqrt{n} (\tau(\hat{\theta}) - \tau(\theta)) \xrightarrow{d} N(0, v(\theta)), \quad (2.2)$$

em que $v(\theta)$ é o limite inferior de Cramér–Rao.

Em termos práticos, isso significa que o EMV converge para o verdadeiro valor do parâmetro (**consistência**) e, assintoticamente, atinge a menor variância possível entre os estimadores não-viesados (**eficiência**). Em consequência, a **normalidade assintótica** também garante a consistência do EMV.

De fato, suponha que

$$\frac{\sqrt{n}(\bar{W}_n - \mu)}{\sigma} \xrightarrow{d} Z, \quad Z \sim N(0, 1) \quad (2.3)$$

Aplicando o Teorema de Slutsky , concluímos:

$$W_n - \mu = \left(\frac{\sigma}{\sqrt{n}} \right) \left(\frac{\sqrt{n}(\bar{W}_n - \mu)}{\sigma} \right) \rightarrow \lim_{n \rightarrow \infty} \left(\frac{\sigma}{\sqrt{n}} \right) Z = 0 \quad (2.4)$$

Isso mostra que, se a amostra é suficientemente grande, a convergência em distribuição para uma normal implica convergência em probabilidade para o parâmetro verdadeiro, garantindo a consistência do estimador.

3. METODOLOGIA

3.1 DADOS

A análise foi realizada a partir do conjunto de dados fornecidos pela comissão técnica (CT) da equipe feminina do Dentil Praia Clube ao professor Cristiano de Barros da Faculdade de Educação Física e Fisioterapia (FAEFI/UFU), contendo as informações coletadas durante os jogos e treinos da temporada 2024/2025. A base contempla 13 das 14 atletas do elenco, uma vez que não há, até o momento, interesse da comissão em aplicar este tipo de monitoramento às jogadoras que atuam na posição de líbero. Para garantir a privacidade individual de cada jogadora, a base de dados foi recebida apenas com a codificação de cada atleta.

Os dados são coletados através de um equipamento de medição e geram como produto as informações referentes a cada salto executado durante as partidas por cada uma das 13 profissionais da equipe que são monitoradas. Para o atual trabalho, foram selecionadas 6 atletas com padrões de saltos semelhantes e as variáveis abaixo:

- Nome da Jogadora: campo codificado para identificação da atleta
- Altura do salto: estimada com base no tempo de deslocamento vertical, em cm;
- Horário: momento exato em que o salto ocorreu, contendo horas, minutos e segundos;
- Data: dia da ocorrência do salto;
- Evento: identificação do tipo de evento (jogo ou treino).

3.2 EQUIPAMENTO DE MEDIÇÃO DE SALTOS

A coleta dos dados foi realizada por meio do dispositivo tecnológico Vertec Jump Tester, amplamente reconhecida na área esportiva. O equipamento é fixado na região lombar ou próxima ao umbigo das atletas, tanto nos treinos quanto nos jogos. O sensor registra automaticamente cada salto, estimando a altura a partir do tempo de deslocamento vertical, além de registrar a data e o horário de cada ocorrência.

3.3 TRATAMENTO E MANIPULAÇÃO DE DADOS

Inicialmente, foram selecionadas seis jogadoras para a análise, cujo objetivo é avaliar individualmente o desempenho ao longo dos jogos disputados. A escolha foi feita com base na semelhança de seus padrões de salto, identificados por meio da análise gráfica. Assim, foram consideradas apenas as atletas de números 105, 193, 216, 350, 864 e 947, restringindo-se os dados ao contexto de jogos, uma vez que treinos podem apresentar diferenças no esforço das atletas. Após esse filtro, a base final resultou em 15.998 observações.

Ainda com relação ao objetivo de avaliar o desempenho ao longo da partida, foi usado como variável explicativa a ordem em que os saltos aconteceram. Assim, para cada uma das atletas citadas anteriormente, foi criada uma contagem indicando a ordem em que cada salto ocorreu dentro da partida considerando a data do jogo e a ordenação do horário em que eles foram registrados pelo equipamento de medição.

A variável *tipo de salto* foi obtida por meio de análise multivariada, que permitiu agrupar os saltos de acordo com suas características. A divisão resultou em dois tipos de saltos para as seis atletas estudadas. Nesse contexto, a variável indicadora X_2 , que representa o tipo de salto, foi criada para diferenciá-los.

Embora a base de dados seja bem organizada, para acrescentar a coluna do número do salto foi necessário criar uma solução para os dias em que o jogo ultrapassa o horário da meia-noite (00:00) seja contabilizado com a ordenação correta. Além disso, os valores nulos para a altura do salto foram desconsiderados em função da extensa base de dados trabalhada.

3.4 ANÁLISE ESTATÍSTICA

Nessa etapa, foi realizada a análise exploratória dos dados para cada uma das jogadoras com a intenção de identificar os possíveis padrões de saltos ao longo das partidas. Foram produzidos gráficos de dispersão através do pacote *ggplot*[11] considerando as variáveis *Ordem do Salto* versus *Altura do Salto* e também histogramas considerando a frequência em que os dados de *Altura do Salto* ocorrem. A partir deles, foi possível notar que as atletas selecionadas possuem dois padrões de salto. Vale ressaltar que a diferença entre estes padrões podem estar relacionados com o tipo de ação desempenhada no jogo (ataque, bloqueio e/ou simulação de ataque) e não devem ser utilizados como meio de comparação entre as atletas. Para estudos futuros, pode-se separar estes tipos de saltos por posição de atuação a fim de obter uma análise mais apurada.

A partir da identificação de mais de um padrão de salto no passo anterior, foram aplicadas as técnicas de análise de cluster para as jogadoras que obtiveram esse perfil utilizando o pacote *Stats*[5]. Nesse contexto, a metodologia utilizada tem a intenção de agrupar os saltos de cada atleta individualmente conforme o modelo escolhido sugere. Para essa classificação, foram consideradas as distâncias euclidianas entre os saltos registrados e em seguida utilizada uma função de agrupamento hierárquico com o método *ward.D2*. Após isso, a árvore de decisão foi

gerada considerando o número de clusters sugerido pelo dendrograma que foi produzido pelo método hierárquico e resultando na classificação de cada um dos saltos. Por fim, o tipo de salto foi integrado na base para cada atleta selecionada para o estudo.

A partir disso, os saltos foram classificados como do tipo 0, para representar o padrão de salto com menor média, e como do tipo 1, a fim de englobar os saltos com média maior.

Em seguida, foi validado um modelo de regressão múltipla para descrever o desempenho de cada atleta durante a partida conforme equação descrita anteriormente 2.1, em que X_1 representa o número do salto na partida, X_2 indica o tipo de salto para cada atleta selecionada, β_0 é o intercepto, correspondente à altura média inicial dos saltos do tipo 0, β_1 representa o efeito da ordem dos saltos nesse mesmo grupo, β_2 corresponde à diferença média entre os saltos do tipo 1 e os do tipo 0 e β_3 descreve a diferença no efeito da ordem entre os dois tipos de salto.

Ainda, considerando os agrupamentos construídos no passo anterior com a criação da variável *Tipo de Salto*, que assume valor 0 ou 1, tem-se que os coeficientes β_2 e β_3 , só terão influencia para o salto do tipo 1. Além disso, quando os coeficientes são significativos à nível de 5% de significância, as interpretações a respeito de cada um deles é:

- β_0 : quando significativo, representa a altura média inicial dos saltos do tipo 0;
- β_1 : descreve a inclinação da reta para os saltos do tipo 0. Se for significativo, indica a taxa de variação da altura a cada salto. Caso contrário, sugere que a inclinação é próxima de zero, de modo que a reta se mantém praticamente constante em torno de β_0 ;
- β_2 : representa a diferença entre as médias dos saltos do tipo 0 e do tipo 1, de modo que, sua significância indica que a diferença é estatisticamente significativa;
- β_3 : informa a diferença média aproximada entre as inclinações das retas dos saltos tipo 0 e tipo 1. Se não for significativo, entende-se que não há distinção estatística entre os coeficientes angulares dos dois grupos.

Devido ao grande tamanho da amostra, o modelo foi ajustado sob a hipótese de que os coeficientes estimados seguem uma distribuição aproximadamente normal, suposição essa, fundamentada pela propriedade assintótica dos estimadores de máxima verossimilhança. Com o resultado do procedimento aplicado, obteve-se a significância dos coeficientes das variáveis explicativas do modelo ajustado que são representativos para a modelagem em cada um dos casos e o que elas representam na prática.

Além disso, para a validação dos modelos, foram utilizados diferentes recursos, incluindo a análise da normalidade dos resíduos e a avaliação de gráficos de ajuste versus os valores observados, os quais indicaram que os modelos estavam de acordo com os critérios de validação adotados.

Todos os procedimentos foram realizados utilizando o software R[5] e RStudio[7], consideraram o nível de significância padrão ($\alpha = 0,05$) para os testes estatísticos realizados.

4. RESULTADOS

Neste capítulo, são apresentados os principais resultados obtidos a partir das análises realizadas, que são organizados de forma a evidenciar os padrões de salto observados entre as atletas, as estruturas de agrupamento identificadas e o desempenho modelado por meio de regressões lineares múltiplas.

4.1 ANÁLISE EXPLORATÓRIA

Os gráficos exploratórios permitiram visualizar a evolução das alturas de salto ao longo das partidas para as atletas codificadas como 105, 193, 216, 350, 864 e 947. Cada ponto representa a altura registrada em um salto sucessivo. A Figura 4.1 ilustra esses padrões.

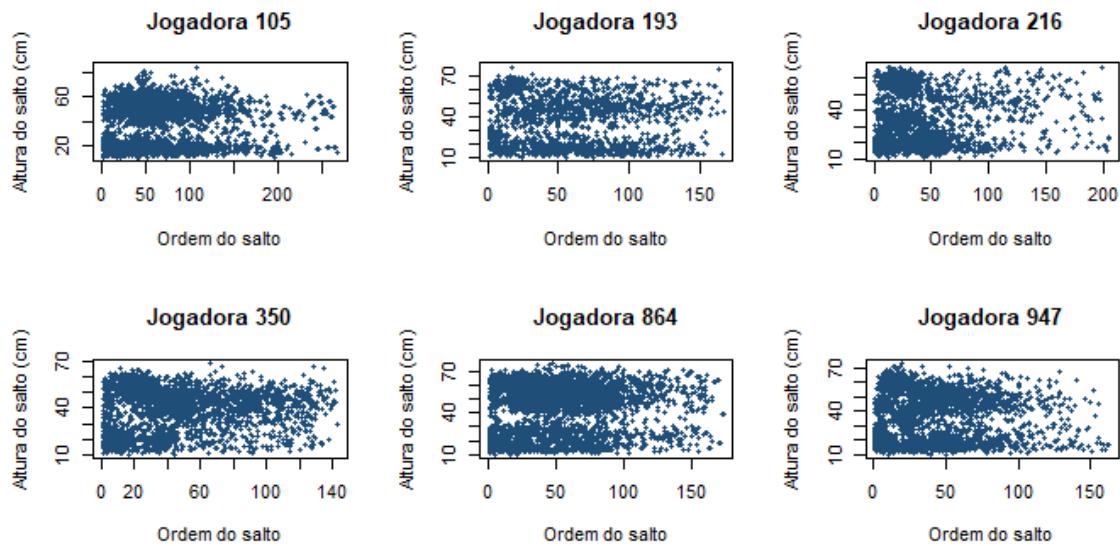


Figura 4.1: Alturas dos saltos das atletas (em centímetros) pela ordem de ocorrência em cada partida

A inspeção visual sugere a existência de mais de um padrão de salto para as jogadoras analisadas, motivando a etapa de agrupamento.

4.2 AGRUPAMENTO DAS ALTURAS DE SALTO

A análise de agrupamento aplicada individualmente a cada atleta indicou, de forma consistente, a existência de dois grupos principais de saltos. Os dendrogramas obtidos, apresentados na Figura 4.2, reforçam essa divisão.

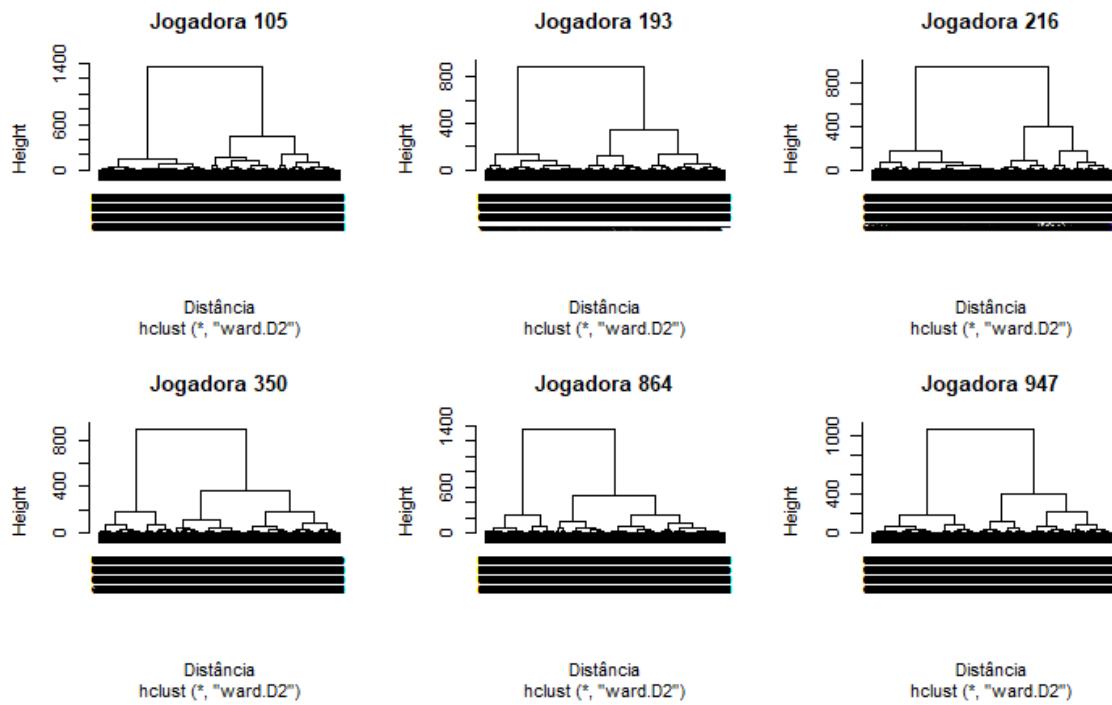


Figura 4.2: Dendrogramas das alturas dos saltos das atletas registrados a cada partida

Com base nessa estrutura, os saltos foram classificados conforme o grupo de menor média (*tipo 0*) e o de maior média (*tipo 1*). As estatísticas descritivas de cada uma das jogadoras encontram-se resumidas na Tabela 4.1.

Tabela 4.1: Estatísticas descritivas por jogadora e tipo de salto (quantidade total de saltos, altura média, desvio-padrão e intervalo de confiança)

Jogadora	Saltos	Salto tipo 0		Salto tipo 1	
		Média \pm DP	IC(95%)	Média \pm DP	IC(95%)
105	3431	$18,51 \pm 3,16$	18,40–18,62	$52,22 \pm 9,38$	51,91–52,54
193	1586	$19,77 \pm 4,79$	19,53–20,00	$51,87 \pm 9,49$	51,40–52,33
216	2174	$20,13 \pm 4,48$	19,94–20,31	$49,69 \pm 10,15$	49,26–50,12
350	2573	$19,25 \pm 5,40$	19,04–19,46	$50,43 \pm 7,55$	50,13–50,72
864	3677	$19,02 \pm 5,46$	18,85–19,20	$55,82 \pm 8,20$	55,56–56,09
947	2557	$18,05 \pm 4,63$	17,87–18,23	$49,86 \pm 9,30$	49,50–50,22

Os grupos apresentam diferenças nas alturas médias, reforçando a distinção entre movimentos de menor deslocamento vertical (tipo 0) e saltos mais vigorosos (tipo 1).

4.3 MODELAGEM DO DESEMPENHO

Em seguida, deu-se início a análise de regressão, a fim de ajustar um modelo que represente cada um dos padrões de saltos observados para cada uma das atletas citadas anteriormente. Como resultado, os modelos que descrevem o comportamento do desempenho para cada uma das jogadoras podem ser observados nos gráficos abaixo.

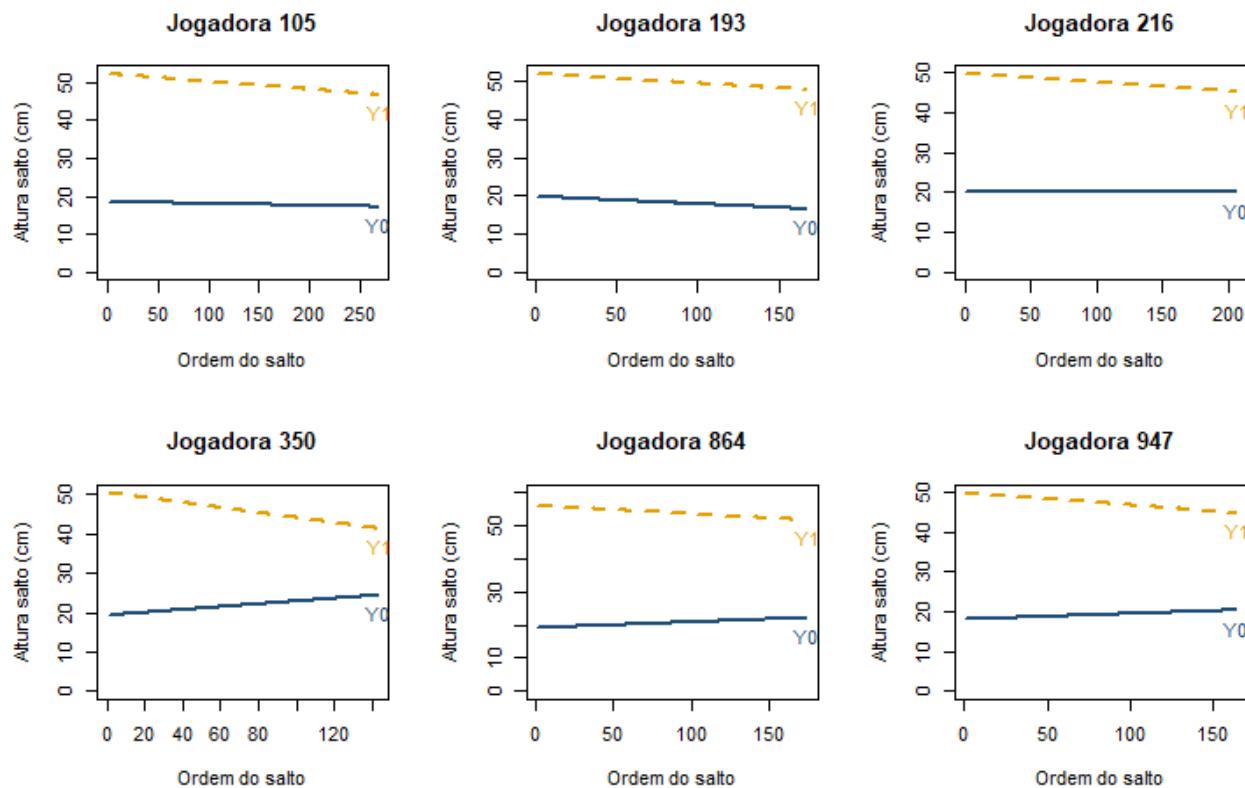


Figura 4.3: Curvas de desempenho de cada tipo de salto ao decorrer da partida para cada uma das atletas estudadas

4.3.1 RESULTADOS GERAIS

Para a atleta 105, o modelo foi ajustado utilizando as variáveis *Tipo de Salto* e *Número do Salto*, além da interação entre as duas, o que resultou nas equações abaixo para descrever cada um dos grupos de salto:

- Equação geral: $Y = 18,51 - 0,004X_1 + 33,71X_2 - 0,016X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 18,51 - 0,004X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 52,22 - 0,020X_1$

O modelo obteve $R^2 = 0,8386$, ou seja, o ajuste consegue explicar cerca de 83,86% da variação das alturas dos saltos da atleta 105 para cada um dos grupos segmentados. Em

relação aos coeficientes, β_0 foi de 18,51 ($p < 2^{-16}$), $\beta_2 = 33,71$ ($p < 2^{-16}$), $\beta_3 = -0,016$ ($p < 0,001$) e $\beta_1 = -0,004$, que no entanto, não apresentou significância estatística.

A partir da contextualização do tema com os modelos projetados acima, observa-se que para os saltos do tipo 0, a altura média inicial é de 18,51 cm com decrescimento de aproximadamente 0,004 cm a cada movimento. No entanto, como o coeficiente linear não é significativo, acredita-se que a reta ajustada mantém-se constante ao longo da partida. Isso significa que a ordem em que o salto ocorre não afeta o desempenho da atleta ao longo do jogo para esse tipo específico de salto. Em contrapartida, nos saltos do tipo 1, a média inicial da altura é de 55,22 cm e possui declínio mais acentuado, em torno de 0,020 cm a cada salto, uma vez que para esse tipo de salto, o coeficiente obtido através da interação entre as variáveis também é significativo e, portanto, contabilizado.

Já para a atleta codificada como 193, os ajustes encontrados foram:

- Equação geral: $Y = 19,77 + -0,0194X_1 + 32,1X_2 + -0,0055X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 19,77 - 0,0194X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 51,85 - 0,0249X_1$

O ajuste produziu um coeficiente de determinação $R^2 = 0,7991$, indicando que o modelo consegue explicar cerca de 79,90% da variabilidade da variável resposta. Os coeficientes obtidos foram $\beta_0 = 19,77$ ($p < 2^{-10}$), $\beta_1 = -0,0194$ ($p = 0,0124$), $\beta_2 = 32,1$ ($p < 2^{-10}$) e $\beta_3 = -0,0055$, que não se mostrou significativo ao nível de significância de 5%.

Com isso, deduz-se que tanto a regressão para o tipo de salto 0 quanto para o salto 1, o decrescimento ocorre de maneira bem similar, mas que diferem fortemente nas médias em que os saltos ocorrem. Ou seja, para o salto do tipo 0, a média do descolamento inicial é de 19,76 cm com um decaimento estatisticamente considerável de aproximadamente 0,0194 cm a cada salto realizado. Já quando analisamos o salto denominado como tipo 1, a média inicial é de 51,85 cm , com decaimento de aproximadamente 0,0249 cm a cada salto. No entanto, o decréscimo para esse grupo é calculado a partir da soma entre o β_1 e o β_3 , em que o último apresentou-se não estatisticamente significativo. Por isso, acredita-se que o decaimento em ambas retas seguem taxas de perdas de desempenho parecidas.

Em seguida, as métricas da atleta 216 foi analisada, resultando nos seguintes modelos:

- Equação geral: $Y = 20,13 - 0,0003X_1 + 29,56X_2 - 0,0215X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 20,13 - 0,0003X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 49,69 - 0,0219X_1$

O coeficiente de determinação gerado pelo modelo indica que aproximadamente 77,77% da variabilidade dos dados é explicada pelas variáveis dependentes ($R^2 = 0,7777$). Já em relação aos coeficientes, a regressão obteve: $\beta_0 = 20,13$ ($p < 2^{-10}$), $\beta_1 = -0,0003$ (não significativo), $\beta_2 = 29,56$ ($p < 2^{-10}$) e $\beta_3 = -0,0215$ ($p = 0,0131$).

Assim, para o salto do tipo 0, apesar da reta possuir um coeficiente linear descrecente (diminuição de cerca de $0,0003\text{ cm}$ a cada pulo), ele não é estatisticamente significativo à nível de 5% de significância, e indica que o desempenho da atleta 216 se mantém praticamente constante para esse grupo de salto. Já em relação ao segundo tipo ($X_2 = 1$), a média inicial do salto é de cerca de $49,68\text{ cm}$, com decréscimo de $0,0219\text{ cm}$ a cada salto. A significância do coeficiente β_3 define que há influência dessa taxa de diminuição para o caso em que a atleta 216 realiza o salto 1.

O modelo obtido para a atleta 350 pode ser visto abaixo:

- Equação geral: $Y = 19,25 + 0,0358X_1 + 31,18X_2 - 0,0994X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 19,25 + 0,0358X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 50,43 - 0,0635X_1$

O ajuste acima conseguiu captar cerca de 78,07% da variabilidade da variável altura do salto ($R^2 = 0,7807$). Além disso, todos os coeficientes mostraram-se estatisticamente significativos a nível de 5% de significância, de maneira que: $\beta_0 = 19,25(p < 2^{-10})$, $\beta_1 = 0,0358(p < 1,26^{-10})$, $\beta_2 = 31,17(p < 2^{-10})$ e $\beta_3 = -0,0994(p < 2^{-10})$.

A partir das informações acima, define-se que para o grupo do salto 0, a atleta 350 possui uma média inicial em torno de $19,25\text{ cm}$ que vai aumentando em aproximadamente $0,0358\text{ cm}$ a altura de cada deslocamento vertical realizado pela jogadora. Por outro lado, ao analisar o salto do tipo 1, a média inicia-se em cerca de $50,42\text{ cm}$ e há tendência de decaimento no desempenho no decorrer da partida (cerca de $0,0635\text{ cm}$ a cada movimento). Com isso, pode-se perceber que a atleta apresenta tendência contrária para cada um dos saltos.

Para a atleta 864, adquiriu-se o modelo abaixo:

- Equação geral: $Y = 19,02 + 0,0174X_1 + 36,8X_2 - 0,04X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 19,02 + 0,0174X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 55,82 - 0,0249X_1$

O modelo ajustado explica cerca de 81,70% da variabilidade da variável resposta ($R^2 = 0,8170$). Além disso, os coeficientes obtidos foram: $\beta_0 = 19,02(p < 2^{-10})$, $\beta_1 = 0,0174(p = 0,0030)$, $\beta_2 = 36,79(p < 2^{-10})$ e $\beta_3 = -0,04(p < 4,16^{-9})$.

Dessa forma, consideramos que para o salto do tipo 0, a atleta tem um crescimento estatisticamente significativo, ou seja, ela inicia o salto com uma média de $19,02\text{ cm}$ e após isso aumenta em torno de $0,0174\text{ cm}$ a cada pulo realizado. Já quando o salto é do tipo 1, ocorre o movimento contrário ao anterior, com tendência de queda a cada deslocamento. Para essa situação, a atleta possui uma média inicial próxima de $55,81\text{ cm}$, mas com o decorrer do jogo, vai perdendo cerca de $0,0249\text{ cm}$ da altura de salto sequencialmente.

Por fim, os modelos ajustados para a jogadora 947 foram:

- Equação geral: $Y = 18,05 + 0,0142X_1 + 31,81X_2 - 0,0452X_1X_2 + \epsilon$
- Salto do tipo 0 ($X_2 = 0$): $Y_0 = 18,05 + 0,0142X_1$
- Salto do tipo 1 ($X_2 = 1$): $Y_1 = 49,86 - 0,0310X_1$

O ajuste realizado possui um $R^2 = 0,7945$, ou seja, consegue explicar 79,45% da variabilidade dos dados. já os coeficientes obtidos foram: $\beta_0 = 18,04(p < 2^{-10})$, $\beta_1 = 0,0142(p = 0,0324)$, $\beta_2 = 31,81(p < 2^{-10})$ e $\beta_3 = -0,0452(p < 1,11^{-6})$.

Com isso, identifica-se que para o salto 0, a atleta tende a aumentar a altura dos saltos ao longo da partida. Inicialmente, a média destes é de 18,04 cm, que é acrescido em cerca de 0,0142 cm a cada deslocamento vertical. O crescimento é estatisticamente significativo. Já quando se trata do outro tipo de salto, a atleta perde altura com o decorrer do jogo. A média inicial é de 49,86 cm, mas a cada salto projeta-se que ela diminua a altura destes em aproximadamente 0,0310 cm.

De modo geral, os resultados indicam que as atletas analisadas apresentam dois padrões distintos de salto. O primeiro, denominado tipo 0, é caracterizado por médias iniciais menores e variações sutis ou estatisticamente pouco significativas. Por outro lado, o salto do tipo 1, que é marcado por médias iniciais mais altas e maior variabilidade ao longo da partida.

Para o salto do tipo 0, as atletas 105 e 216 apresentaram desempenho praticamente constante ao longo da partida, com discreto decréscimo que não altera de forma relevante a altura média final em relação à inicial. Já as jogadoras 193, 350, 864 e 947 apresentaram inclinações mais notáveis nas retas ajustadas: enquanto a atleta 193 demonstrou tendência de queda nas alturas de salto, as demais exibiram tendência de crescimento ao longo da partida para esse tipo de salto. Já para o tipo de salto 1, todas as jogadoras analisadas apresentaram tendência de queda no desempenho.

Ao analisar os coeficientes que comparam as tendências dos saltos dos tipos 0 e 1, observa-se que, quando não são significativos, o padrão de desempenho se mantém semelhante entre os tipos de salto, variando apenas a altura inicial. A inclinação das retas, ou seja, a taxa de variação da altura ao longo dos saltos, é praticamente a mesma para ambos os tipos. Nesse contexto, nota-se que as atletas 105, 216, 350, 864 e 947 apresentam padrões de desempenho diferentes para cada tipo de salto, enquanto a atleta 193 apresenta um declínio proporcional, semelhante entre os dois tipos.

Por fim, para o salto do tipo 0, a atleta que mais apresenta queda de desempenho é a 193, com redução média de cerca de 0,0194 cm a cada salto. Em contrapartida, a que mais evolui nesse padrão é a 350, aumentando em torno de 0,0358 cm por movimento. Já no salto do tipo 1, a atleta 350 se destaca novamente, mas dessa vez por ser a que mais perde altura ao longo da partida, apresentando também a maior diferença entre os dois tipos de salto, em aproximadamente 0,0994 cm.

5. CONCLUSÕES

Através deste estudo, foi possível identificar diferentes padrões de salto durante as partidas e, a partir disso, desenvolver uma modelagem matemática capaz de descrever o comportamento de desempenho de cada atleta ao longo do jogo. Observou-se que, de modo geral, as atletas apresentam médias e desvios de salto próximas, tanto para o salto do tipo 0 quanto para o do tipo 1. O salto do tipo 0, considerado menor e mais estável, manteve-se praticamente constante para algumas atletas, enquanto em outras apresentou variação significativa, mesmo que de pequena magnitude. Já para o salto do tipo 1, verificou-se um decréscimo da altura ao longo do jogo para todas as jogadoras, com diferenças individuais pequenas, mas perceptíveis.

Os resultados podem servir de apoio à tomada de decisões técnicas e à preparação física, permitindo ajustes conforme as necessidades do time. Vale reforçar que este estudo não considerou as posições em que as atletas atuam, o que pode influenciar as métricas obtidas. Por isso, os dados não devem ser utilizados para comparações diretas entre jogadoras.

Por fim, a abordagem utilizada abre espaço para análises adicionais que podem complementar a avaliação do desempenho individual, podendo ser exploradas em trabalhos futuros. Uma possibilidade seria incluir a variável *Evento* considerando também os treinos. Outra sugestão seria investigar diferenças de desempenho entre jogos ou entre períodos específicos de treino, como do primeiro treino após uma folga até o treino anterior ao jogo. Além disso, outras variáveis ainda não exploradas, como o resultado da partida, o time adversário e o local do jogo, também podem ser incorporadas à análise.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Cartaxo, F. B.: *Análise de desempenho para atletas: alcançando resultados em todas as modalidades esportivas!*, 2024. <https://www.drfredcartaxo.com.br/analise-de-desempenho-para-atletas-alcancando-resultados-em-todas-as-modalidades-esportivas/>, acessado em 27/02/2025.
- [2] Fabrício Missio, L. F. J.: *Variáveis dummy: especificações de modelos com parâmetros variáveis*. UFSM, 2007.
- [3] FIVB: *História*, 2025. <https://www.fivb.com/volleyball/the-game/history/>, Acesso em: setembro de 2025.
- [4] Jones, C. M., Griffiths, P. C. e Mellalieu, S. D.: *Training Load and Fatigue Marker Associations with Injury and Illness: A Systematic Review of Longitudinal Studies*. Sports medicine (Auckland, N.Z.), 47(5):943–974, 2017. <https://doi.org/10.1007/s40279-016-0619-5>.
- [5] R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2024. <https://www.R-project.org/>.
- [6] Reeser, J. C. e Bahr, R.: *Handbook of Sports Medicine and Science Volleyball*. Blackwell Science, 2017.
- [7] RStudio Team: *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA, 2025. <https://posit.co/>.
- [8] Santos, Matheus Ribeiro Quaglio, L. C. d. O. R. G. d. O. Michael Inácio dos: *Guia de avaliação do salto vertical*. CRV Ltda, 2023.
- [9] Sheppard, J., Gabbett, T. e Riggs, M.: *Volleyball. In: Physiological testing of elite athletes, 2nd edition*. Australian Institute of Sport, 2013.
- [10] Sucupira, G. e Cardoso, F.: *O Jogo de Voleibol: Análise da eficiência/eficácia das ações e dos complexos estratégicos no alto rendimento*. Editora e Livraria Appris Ltda, 1^a ed., 2023.
- [11] Wickham, H.: *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

A. APÊNDICE

A.1 SCRIPT R

```
library(dplyr)
library(stringr)
library(hms)
library(MASS)
library(ggplot2)
library(class)

df = readxl::read_xlsx("tabela_saltos.xlsx")

df_jogo =
  df |>
  filter(`Jogo/treino` == 'Jogo')

## TRANSFORMANDO AS VARIÁVEIS

str(df_jogo)

{
  df_jogo$Time = str_extract_all(df_jogo$Time,
  pattern = '\\\\d{2}:\\\\d{2}:\\\\d{2}')
  df_jogo$Name = as.factor(df_jogo$Name)
  df_jogo$`Jogo/treino` = as.factor(df_jogo$`Jogo/treino`)
  df_jogo$Date = as.Date(df_jogo$Date)
  df_jogo$Time = as_hms(unlist(df_jogo$Time))

}

#AJUSTE PARA JOGOS QUE PASSARAM DE MEIA NOITE
{
```

```
## CORREÇÃO NA ORDENAÇÃO PARA AS DATAS "2024-12-03"

df_jogo = df_jogo |>
  mutate(
    Time = if_else(
      Date == "2024-12-03" & as.numeric(Time) >1000,
      as_hms(as.numeric(Time) - 1800),
      Time
    )
  )

df_jogo = df_jogo |>
  mutate(
    Time = if_else(
      Date == "2024-12-03" & as.numeric(Time) <= 1000,
      as_hms(as.numeric(Time) + 84575),
      Time
    )
  )

df_jogo |> filter(Date == "2024-11-19") |> arrange(Time)

df_jogo = df_jogo |>
  mutate(
    Time = if_else(
      Date == "2024-11-19" & as.numeric(Time) > 3600,
      as_hms(as.numeric(Time) - 3600),
      Time
    )
  )

df_jogo = df_jogo |>
  mutate(
    Time = if_else(
      Date == "2024-11-19" & as.numeric(Time) <= 3600,
      as_hms(as.numeric(Time) + 82520),
      Time
    )
  )
```

```
)  
  
df_jogo = df_jogo |>  
  mutate(  
    Time = if_else(  
      Date == "2025-02-27" & as.numeric(Time) > 3600,  
      as_hms(as.numeric(Time) - 3600),  
      Time  
    )  
  )  
  
df_jogo = df_jogo |>  
  mutate(  
    Time = if_else(  
      Date == "2025-02-27" & as.numeric(Time) <= 3600,  
      as_hms(as.numeric(Time) + 82520),  
      Time  
    )  
  )  
  
}  
  
## ACRESCENTANDO A CONTAGEM DE SALTOS POR ATLETA (1)  
# df_jogo  
{  
  df_jogo = df_jogo |>  
    group_by(Name, Date) |>  
    arrange(Time, .by_group = TRUE) |>  
    mutate(nsalto = row_number()) |>  
    ungroup()  
  
}  
  
## DADOS DE CADA JOGADORA SELECCIONADA PARA O MODELO:  
  
{  
  df_jogo = df_jogo[is.finite(df_jogo$Jump_Height),]
```

```

Name_105 = df_jogo |> dplyr::filter(Name == 105) |>
  ~ dplyr::select(Jump_Height, nsalto)
Name_193 = df_jogo |> dplyr::filter(Name == 193) |>
  ~ dplyr::select(Jump_Height, nsalto)
Name_216 = df_jogo |> dplyr::filter(Name == 216) |>
  ~ dplyr::select(Jump_Height, nsalto)
Name_350 = df_jogo |> dplyr::filter(Name == 350) |>
  ~ dplyr::select(Jump_Height, nsalto)
Name_864 = df_jogo |> dplyr::filter(Name == 864) |>
  ~ dplyr::select(Jump_Height, nsalto)
Name_947 = df_jogo |> dplyr::filter(Name == 947) |>
  ~ dplyr::select(Jump_Height, nsalto)

}

## GRÁFICOS DE DISPERSÃO

```

```

{
  par(mfrow = c(2, 3))

  plot(x = Name_105$nsalto,
        y = Name_105$Jump_Height,
        main = "Jogadora 105",
        xlab = "Ordem do salto",
        ylab = "Altura do salto (cm)",
        col = "#1F4E79",
        pch = 16,
        cex = 0.7)

  plot(x = Name_193$nsalto,
        y = Name_193$Jump_Height,
        main = "Jogadora 193",
        xlab = "Ordem do salto",
        ylab = "Altura do salto (cm)",
        col = "#1F4E79",
        pch = 16,
        cex = 0.7)

  plot(x = Name_216$nsalto,

```

```
y = Name_216$Jump_Height,  
main = "Jogadora 216",  
xlab = "Ordem do salto",  
ylab = "Altura do salto (cm)",  
col = "#1F4E79",  
pch = 16,  
cex = 0.7)  
  
plot(x = Name_350$nsalto,  
y = Name_350$Jump_Height,  
main = "Jogadora 350",  
xlab = "Ordem do salto",  
ylab = "Altura do salto (cm)",  
col = "#1F4E79",  
pch = 16,  
cex = 0.7)  
  
plot(x = Name_864$nsalto,  
y = Name_864$Jump_Height,  
main = "Jogadora 864",  
xlab = "Ordem do salto",  
ylab = "Altura do salto (cm)",  
col = "#1F4E79",  
pch = 16,  
cex = 0.7)  
  
plot(x = Name_947$nsalto,  
y = Name_947$Jump_Height,  
main = "Jogadora 947",  
xlab = "Ordem do salto",  
ylab = "Altura do salto (cm)",  
col = "#1F4E79",  
pch = 16,  
cex = 0.7)  
}  
  
## CLUSTERIZAÇÃO  
  
{  
#105
```

```
{  
  distancia_105 = dist(Name_105$Jump_Height, method = "euclidean")  
  agrup_hierarquico_105 = hclust(distancia_105, method = "ward.D2")  
  
  
  clusters_105 = cutree(agrup_hierarquico_105, k = 2)  
  clusters_105 = ifelse(clusters_105 == 2, 0, clusters_105)  
  
  Name_105 = cbind(Name_105, clusters_105)  
}  
  
  
# 193  
{  
  distancia_193 = dist(Name_193$Jump_Height, method = "euclidean")  
  agrup_hierarquico_193 = hclust(distancia_193, method = "ward.D2")  
  
  clusters_193 = cutree(agrup_hierarquico_193, k = 2)  
  clusters_193 = clusters_193 - 1  
  
  Name_193 = cbind(Name_193, clusters_193)  
}  
  
  
# 216  
{  
  distancia_216 = dist(Name_216$Jump_Height, method = "euclidean")  
  agrup_hierarquico_216 = hclust(distancia_216, method = "ward.D2")  
  
  clusters_216 = cutree(agrup_hierarquico_216, k = 2)  
  clusters_216 = clusters_216 -1  
  
  Name_216 = cbind(Name_216, clusters_216)  
}  
  
  
# 350  
{  
  distancia_350 = dist(Name_350$Jump_Height, method = "euclidean")  
  agrup_hierarquico_350 = hclust(distancia_350, method = "ward.D2")
```

```
clusters_350 = cutree(agrup_hierarquico_350, k = 2)
clusters_350 = clusters_350-1

Name_350 = cbind(Name_350, clusters_350)
}

# 864
{
  distancia_864 = dist(Name_864$Jump_Height, method = "euclidean")
  agrup_hierarquico_864 = hclust(distancia_864, method = "ward.D2")

  clusters_864 = cutree(agrup_hierarquico_864, k = 2)
  clusters_864 = clusters_864 -1

  Name_864 = cbind(Name_864, clusters_864)
}

# 947
{
  distancia_947 = dist(Name_947$Jump_Height, method = "euclidean")
  agrup_hierarquico_947 = hclust(distancia_947, method = "ward.D2")

  clusters_947 = cutree(agrup_hierarquico_947, k = 2)
  clusters_947 = clusters_947 - 1

  Name_947 = cbind(Name_947, clusters_947)
}

## DENDROGRAMA
{
  ### PLOTS
  par(mfrow = c(2,3))
```

```
plot(agrup_hierarquico_105,
  main = "Jogadora 105",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

plot(agrup_hierarquico_193,
  main = "Jogadora 193",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

plot(agrup_hierarquico_216,
  main = "Jogadora 216",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

plot(agrup_hierarquico_350,
  main = "Jogadora 350",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

plot(agrup_hierarquico_864,
  main = "Jogadora 864",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

plot(agrup_hierarquico_947,
  main = "Jogadora 947",
  xlab = "Distância",
  ylab = "Altura do salto (cm)")

}

## MODELAGEM REGRESSÃO

{

# 105
{
  mod105 = lm(Jump_Height ~ nsaldo + clusters_105 + nsaldo:clusters_105,
  ↵  data = Name_105)
```

```
summary(mod105)

b0 = coalesce(mod105$coefficients[1], 0)
b1 = coalesce(mod105$coefficients[2], 0)
b2 = coalesce(mod105$coefficients[3], 0)
b3 = coalesce(mod105$coefficients[4], 0)

{
  y0 = function(n.salto){
    b0 + b1*n.salto
  }

  y1 = function(n.salto){
    (b0 + b2) + (b1 + b3)*n.salto
  }

  ylim_min = 0
}

valores_x_105 = seq(1, max(Name_105$nsalto), by = 1)

valores_y0_105 = y0(valores_x_105)
valores_y1_105 = y1(valores_x_105)

ylim_max_105 = max(valores_y0_105, valores_y1_105, na.rm = TRUE)

### tabela:

Numero_saltos_105 = count(Name_105)
Name_105_media_0 = b0
Name_105_sd_0 = Name_105 |> dplyr::filter(clusters_105 == 0) |>
  summarise(sd(Jump_Height))
Name_105_media_1 = b0 + b2
Name_105_sd_1 = Name_105 |> dplyr::filter(clusters_105 == 1) |>
  summarise(sd(Jump_Height))
}
```

```

# 193
{
  mod193 = lm(Jump_Height ~ nsalto + clusters_193 + nsalto:clusters_193,
  ↵  data = Name_193)

  summary(mod193)

  valores_x_193 = seq(1, max(Name_193$nsalto), by = 1)

  b0 = coalesce(mod193$coefficients[1], 0)
  b1 = coalesce(mod193$coefficients[2], 0)
  b2 = coalesce(mod193$coefficients[3], 0)
  b3 = coalesce(mod193$coefficients[4], 0)

  B0_Y1 = b0 + b2
  B1_Y1 = b1 + b3
  B0_Y1
  B1_Y1

  {
    y0 = function(n.salto){
      b0 + b1*n.salto
    }

    y1 = function(n.salto){
      (b0 + b2) + (b1 + b3)*n.salto
    }

    ylim_min = 0
  }

  valores_y0_193 = y0(valores_x_193)
  valores_y1_193 = y1(valores_x_193)

  ylim_max_193 = max(valores_y0_193, valores_y1_193, na.rm = TRUE)

  ### tabela:

```

```
Numero_saltos_193 = count(Name_193)
Name_193_media_0 = b0
Name_193_sd_0 = Name_193 |> dplyr::filter(clusters_193 == 0) |>
  summarise(sd(Jump_Height))
Name_193_media_1 = b0 + b2
Name_193_sd_1 = Name_193 |> dplyr::filter(clusters_193 == 1) |>
  summarise(sd(Jump_Height))
}

# 216
{
  mod216 = lm(Jump_Height ~ nsalto + clusters_216 + nsalto:clusters_216,
  ~ data = Name_216)

  summary(mod216)

  valores_x_216 = seq(1, max(Name_216$nsalto), by = 1)

  b0 = coalesce(mod216$coefficients[1], 0)
  b1 = coalesce(mod216$coefficients[2], 0)
  b2 = coalesce(mod216$coefficients[3], 0)
  b3 = coalesce(mod216$coefficients[4], 0)

  B0_Y1 = b0 + b2
  B1_Y1 = b1 + b3
  B0_Y1
  B1_Y1

  {
    y0 = function(n.salto){
      b0 + b1*n.salto
    }

    y1 = function(n.salto){
      (b0 + b2) + (b1 + b3)*n.salto
    }

    ylim_min = 0
  }
}
```

```

valores_y0_216 = y0(valores_x_216)
valores_y1_216 = y1(valores_x_216)

ylim_max_216 = max(valores_y0_216, valores_y1_216, na.rm = TRUE)

### tabela:
Numero_saltos_216 = count(Name_216)
Name_216_media_0 = b0
Name_216_sd_0 = Name_216 |> dplyr::filter(clusters_216 == 0) |>
  summarise(sd(Jump_Height))
Name_216_media_1 = b0 + b2
Name_216_sd_1 = Name_216 |> dplyr::filter(clusters_216 == 1) |>
  summarise(sd(Jump_Height))
}

# 350
{
  mod350 = lm(Jump_Height ~ nsalto + clusters_350 + nsalto:clusters_350,
  ~ data = Name_350)
  summary(mod350)

  valores_x_350 = seq(1, max(Name_350$nsalto), by = 1)

  b0 = coalesce(mod350$coefficients[1], 0)
  b1 = coalesce(mod350$coefficients[2], 0)
  b2 = coalesce(mod350$coefficients[3], 0)
  b3 = coalesce(mod350$coefficients[4], 0)

  B0_Y1 = b0 + b2
  B1_Y1 = b1 + b3
  B0_Y1
  B1_Y1

  {
    y0 = function(n.salto){
      b0 + b1*n.salto
    }
  }
}

```

```
y1 = function(n.salto){  
  (b0 + b2) + (b1 + b3)*n.salto  
}  
  
ylim_min = 0  
}  
  
valores_y0_350 = y0(valores_x_350)  
valores_y1_350 = y1(valores_x_350)  
  
ylim_max_350 = max(valores_y0_350, valores_y1_350, na.rm = TRUE)  
  
### tabela:  
  
Numero_saltos_350 = count(Name_350)  
Name_350_media_0 = b0  
Name_350_sd_0 = Name_350 |> dplyr::filter(clusters_350 == 0) |>  
  summarise(sd(Jump_Height))  
Name_350_media_1 = b0 + b2  
Name_350_sd_1 = Name_350 |> dplyr::filter(clusters_350 == 1) |>  
  summarise(sd(Jump_Height))  
}  
  
# 864  
{  
  
  mod864 = lm(Jump_Height ~ nsaldo + clusters_864 + nsaldo:clusters_864,  
  ~ data = Name_864)  
  
  summary(mod864)  
  valores_x_864 = seq(1, max(Name_864$nsaldo), by = 1)  
  
  b0 = coalesce(mod864$coefficients[1], 0)  
  b1 = coalesce(mod864$coefficients[2], 0)  
  b2 = coalesce(mod864$coefficients[3], 0)  
  b3 = coalesce(mod864$coefficients[4], 0)  
  
  B0_Y1 = b0 + b2  
  B1_Y1 = b1 + b3
```

```

B0_Y1
B1_Y1

{

y0 = function(n.salto){
  b0 + b1*n.salto
}

y1 = function(n.salto){
  (b0 + b2) + (b1 + b3)*n.salto
}

ylim_min = 0
}

valores_y0_864 = y0(valores_x_864)
valores_y1_864 = y1(valores_x_864)

ylim_max_864 = max(valores_y0_864, valores_y1_864, na.rm = TRUE)

## tabela:

Numero_saltos_864 = count(Name_864)
Name_864_media_0 = b0
Name_864_sd_0 = Name_864 |> dplyr::filter(clusters_864 == 0) |>
  → summarise(sd(Jump_Height))
Name_864_media_1 = b0 + b2
Name_864_sd_1 = Name_864 |> dplyr::filter(clusters_864 == 1) |>
  → summarise(sd(Jump_Height))

}

# 947
{

mod947 = lm(Jump_Height ~ nsaldo + clusters_947 + nsaldo:clusters_947,
  → data = Name_947)
summary(mod947)

```

```
valores_x_947 = seq(1, max(Name_947$nsalto), by = 1)

b0 = coalesce(mod947$coefficients[1], 0)
b1 = coalesce(mod947$coefficients[2], 0)
b2 = coalesce(mod947$coefficients[3], 0)
b3 = coalesce(mod947$coefficients[4], 0)

B0_Y1 = b0 + b2
B1_Y1 = b1 + b3
B0_Y1
B1_Y1

{
  y0 = function(n.salto){
    b0 + b1*n.salto
  }

  y1 = function(n.salto){
    (b0 + b2) + (b1 + b3)*n.salto
  }

  ylim_min = 0
}

valores_y0_947 = y0(valores_x_947)
valores_y1_947 = y1(valores_x_947)

ylim_max_947 = max(valores_y0_947, valores_y1_947, na.rm = TRUE)

### tabela:

Numero_saltos_947 = count(Name_947)
Name_947_media_0 = b0
Name_947_sd_0 = Name_947 |> dplyr::filter(clusters_947 == 0) |>
  summarise(sd(Jump_Height))
Name_947_media_1 = b0 + b2
Name_947_sd_1 = Name_947 |> dplyr::filter(clusters_947 == 1) |>
  summarise(sd(Jump_Height))
}
```

```
}

## TABELA
{

Nomes = c("105", "193", "216", "350", "864", "947")

NúmeroSaltos = c(Número_saltos_105, Número_saltos_193, Número_saltos_216,
                 Número_saltos_350, Número_saltos_864, Número_saltos_947)
NúmeroSaltos = unlist(NúmeroSaltos)

Medias0 = c(Name_105_media_0, Name_193_media_0, Name_216_media_0,
            Name_350_media_0, Name_864_media_0, Name_947_media_0)

desvio0_inf = c(Name_105_media_0 - Name_105_sd_0,
                 Name_193_media_0 - Name_193_sd_0,
                 Name_216_media_0 - Name_216_sd_0,
                 Name_350_media_0 - Name_350_sd_0,
                 Name_864_media_0 - Name_864_sd_0,
                 Name_947_media_0 - Name_947_sd_0)

desvio0_inf = unlist(desvio0_inf)

desvio0_sup = c(Name_105_media_0 + Name_105_sd_0,
                 Name_193_media_0 + Name_193_sd_0,
                 Name_216_media_0 + Name_216_sd_0,
                 Name_350_media_0 + Name_350_sd_0,
                 Name_864_media_0 + Name_864_sd_0,
                 Name_947_media_0 + Name_947_sd_0)

desvio0_sup = unlist(desvio0_sup)

Medias1 = c(Name_105_media_1, Name_193_media_1, Name_216_media_1,
            Name_350_media_1, Name_864_media_1, Name_947_media_1)

desvio1_inf = c(Name_105_media_1 - Name_105_sd_1,
                 Name_193_media_1 - Name_193_sd_1,
                 Name_216_media_1 - Name_216_sd_1,
                 Name_350_media_1 - Name_350_sd_1,
                 Name_864_media_1 - Name_864_sd_1,
                 Name_947_media_1 - Name_947_sd_1)
```

```
desvio1_inf = unlist(desvio1_inf)

desvio1_sup = c(Name_105_media_1 + Name_105_sd_1,
                 Name_193_media_1 + Name_193_sd_1,
                 Name_216_media_1 + Name_216_sd_1,
                 Name_350_media_1 + Name_350_sd_1,
                 Name_864_media_1 + Name_864_sd_1,
                 Name_947_media_1 + Name_947_sd_1)

desvio1_sup = unlist(desvio1_sup)

## MONTAGEM DO DATAFRAME

df_medias <- data.frame(
  Jogadora = Nomes,
  NumeroSaltos = NumeroSaltos,
  Media0 = Medias0,
  Limite0_inf = desvio0_inf,
  Limite0_sup = desvio0_sup,
  Media1 = Medias1,
  Limite1_inf = desvio1_inf,
  Limite1_sup = desvio1_sup,
  stringsAsFactors = FALSE
)

print(df_medias)

}

### PLOTS REGRESSAO

par(mfrow = c(2,3))
{
  ## 105

  {
    plot(valores_x_105, valores_y0_105, type = "l", col = "#1F4E79", lwd =
         2,
```

```

  xlab = "Ordem do salto", ylab = "Altura salto (cm)",
  main = "Jogadora 105",
  ylim = c(0, ylim_max_105))

lines(valores_x_105, valores_y1_105, col = "#E69F00", lwd = 2, lty = 2)

text(x = max(valores_x_105), y = tail(valores_y0_105, 1),
  labels = "Y0", pos = 1, col = "#1F4E79")

text(x = max(valores_x_105) +1, y = tail(valores_y1_105, 1),
  labels = "Y1", pos = 1, col = "#E69F00")
}

## 193
{

plot(valores_x_193, valores_y0_193, type = "l", col = "#1F4E79", lwd =
  2,
  xlab = "Ordem do salto", ylab = "Altura salto (cm)",
  main = "Jogadora 193",
  ylim = c(ylim_min, ylim_max_193))

lines(valores_x_193, valores_y1_193, col = "#E69F00", lwd = 2, lty = 2)

text(x = max(valores_x_193), y = tail(valores_y0_193, 1),
  labels = "Y0", pos = 1, col = "#1F4E79")

text(x = max(valores_x_193) +1, y = tail(valores_y1_193, 1),
  labels = "Y1", pos = 1, col = "#E69F00")
}

## 216
{

plot(valores_x_216, valores_y0_216, type = "l", col = "#1F4E79", lwd =
  2,
  xlab = "Ordem do salto", ylab = "Altura salto (cm)",
  main = "Jogadora 216",
  ylim = c(ylim_min, ylim_max_216))
}

```

```
lines(valores_x_216, valores_y1_216, col = "#E69F00", lwd = 2, lty = 2)

text(x = max(valores_x_216), y = tail(valores_y0_216, 1),
      labels = "Y0", pos = 1, col = "#1F4E79")

text(x = max(valores_x_216) +1, y = tail(valores_y1_216, 1),
      labels = "Y1", pos = 1, col = "#E69F00")

}

## 350
{
  plot(valores_x_350, valores_y0_350, type = "l", col = "#1F4E79", lwd =
    2,
    xlab = "Ordem do salto", ylab = "Altura salto (cm)",
    main = "Jogadora 350",
    ylim = c(ylim_min, ylim_max_350))

  lines(valores_x_350, valores_y1_350, col = "#E69F00", lwd = 2, lty = 2)

  text(x = max(valores_x_350), y = tail(valores_y0_350, 1),
        labels = "Y0", pos = 1, col = "#1F4E79")

  text(x = max(valores_x_350) +1, y = tail(valores_y1_350, 1),
        labels = "Y1", pos = 1, col = "#E69F00")
}

#####
864
{
  plot(valores_x_864, valores_y0_864, type = "l", col = "#1F4E79", lwd =
    2,
    xlab = "Ordem do salto", ylab = "Altura salto (cm)",
    main = "Jogadora 864",
    ylim = c(ylim_min, ylim_max_216 + 10))

  lines(valores_x_864, valores_y1_864, col = "#E69F00", lwd = 2, lty = 2)
```

```
text(x = max(valores_x_864), y = tail(valores_y0_864, 1),
      labels = "Y0", pos = 1, col = "#1F4E79")

text(x = max(valores_x_864) +1, y = tail(valores_y1_864, 1),
      labels = "Y1", pos = 1, col = "#E69F00")
}

#####
947

{
  plot(valores_x_947, valores_y0_947, type = "l", col = "#1F4E79", lwd =
    2,
    xlab = "Ordem do salto", ylab = "Altura salto (cm)",
    main = "Jogadora 947",
    ylim = c(ylim_min, ylim_max_947))

  lines(valores_x_947, valores_y1_947, col = "#E69F00", lwd = 2, lty = 2)

  text(x = max(valores_x_947), y = tail(valores_y0_947, 1),
        labels = "Y0", pos = 1, col = "#1F4E79")

  text(x = max(valores_x_947) +1, y = tail(valores_y1_947, 1),
        labels = "Y1", pos = 1, col = "#E69F00")
}

}
```