

Análise do Impacto dos Duelistas no Valorant VCT 2023 usando K-Means

Luiz Felipe De Oliveira Barbosa Nunes (RA 255403)

Resumo

Este trabalho apresenta uma análise do desempenho dos jogadores no Valorant VCT 2023, aplicando o algoritmo K-means com um foco detalhado na classe Duelista. Valorant, o aclamado jogo de tiro em primeira pessoa da Riot Games, coloca equipes de cinco em combates táticos com objetivos claros. O objetivo da análise é avaliar se a classe Duelista, com suas habilidades de combate especializadas, exerce uma influência mais determinante nas partidas do que as outras classes.

1 Introdução

Valorant é organizado em partidas onde a equipe atacante tenta plantar e defender uma bomba, chamada Spike, até sua detonação para ganhar pontos. Enquanto isso, a equipe defensora tenta desarmar a Spike ou eliminar todos os adversários para também ganhar pontos. As partidas são jogadas em um formato de “melhor de 24 rodadas”, onde a equipe que primeiro alcançar 13 vitórias é declarada vencedora.

No Valorant, os jogadores são categorizados em quatro classes: Duelistas, Controladores, Sentinelas e Iniciadores, cada uma com habilidades distintas que definem suas funções estratégicas no jogo. Os Duelistas, especializados em combate direto, são fundamentais para iniciar confrontos, penetrar defesas adversárias e realizar eliminações críticas.

O banco de dados usado neste estudo foi adquirido do site Kaggle, que oferece um vasto conjunto de métricas de desempenho dos jogadores no campeonato Valorant VCT 2023. O conjunto de dados está disponível no seguinte endereço: <https://www.kaggle.com/datasets/vkay616/valorant-vct-2023-player-performance>, incluindo variáveis importantes como Rating, KD (relação Kill/Death), e Kills.Max (número máximo de eliminações em uma partida), entre outras. Além disso, foi adicionada uma coluna de “Funções” ao banco de dados, que categoriza a função de cada jogador no jogo.

Este estudo tem como objetivo principal empregar o método K-means para analisar a performance dos jogadores no Valorant VCT 2023, com atenção especial aos Duelistas.

Procurei determinar se os Duelistas, de acordo com suas habilidades ofensivas, realmente têm o maior impacto nas partidas em comparação com outras classes. Variáveis-chave como Rating, ACS (Average Combat Score) e KD serão exploradas para identificar padrões e perfis de impacto dos jogadores. A análise foi realizada utilizando a linguagem de programação R.

2 Materiais e Métodos

O objetivo principal do algoritmo K-Means é particionar um conjunto de dados $\mathbf{X} = \{x_1, \dots, x_N\}$, onde cada x_n é um vetor em \mathbb{R}^d , em M clusters $C = \{C_1, \dots, C_M\}$. Cada cluster C_k é identificado pelo seu centroide \mathbf{m}_k , que é o ponto médio de todos os pontos atribuídos a C_k . O objetivo é minimizar a variância interna dos clusters, que é dada pela soma das distâncias quadradas dentro de cada cluster:

$$E(\mathbf{m}_1, \dots, \mathbf{m}_M) = \sum_{k=1}^M \sum_{\mathbf{x}_i \in C_k} \|\mathbf{x}_i - \mathbf{m}_k\|^2,$$

onde $\|\mathbf{x}_i - \mathbf{m}_k\|$ denota a distância euclidiana entre o ponto \mathbf{x}_i e o centroide \mathbf{m}_k .

Para implementar o K-Means e analisar os clusters de desempenho dos jogadores, segui estes passos:

1. *Análise de PCA*: Realizei uma análise de componentes principais (PCA) para identificar as variáveis mais significativas, facilitando a redução da dimensionalidade.
2. *Método do Cotovelo*: Utilizei o método do cotovelo para determinar o número ótimo de clusters k .
3. *Inicialização dos centroides*: Selecionei k pontos iniciais como centroides, os k primeiros pontos de dados.
4. *Cálculo das distâncias*: Calculei a distância de cada ponto de dados a todos os k centroides para definir o cluster ao qual cada ponto pertence.
5. *Atribuição dos clusters*: Associei cada ponto \mathbf{x}_i ao cluster cujo centroide \mathbf{m}_k é o mais próximo.
6. *Atualização dos centroides*: Atualizei o centroide de cada cluster \mathbf{m}_k para ser o centro geométrico de todos os pontos atribuídos a C_k , calculado por:

$$\mathbf{m}_k = \frac{1}{|C_k|} \sum_{\mathbf{x}_i \in C_k} \mathbf{x}_i,$$

onde $|C_k|$ é o número de pontos no cluster C_k .

7. *Gráfico de Distribuição*: Gerei gráficos de distribuição dos clusters em duas dimensões após a aplicação do PCA, para visualizar como os clusters são organizados no espaço reduzido.
8. *Gráfico de Silhuetas*: Produzi gráficos de silhuetas para avaliar a qualidade da clusterização, onde cada silhueta mede o quão similar um ponto é aos pontos do seu próprio cluster comparado a pontos de clusters vizinhos.
9. *Construção de Tabelas*: Compilei tabelas para sumarizar as estatísticas chave, incluindo a média das métricas de desempenho e a frequência das funções dos jogadores em cada cluster.

3 Resultados e Discussão

Em primeiro lugar, foi feita a análise de PCA. Esses foram os resultados dos testes:

Tabela 1: Importância dos componentes principais da PCA.

| Componente | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 | PC11 | PC12 | PC13 | PC14 | PC15 | PC16 |
|------------------------|--------|--------|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------|---------|---------|
| Desvio Padrão | 2.8712 | 1.9339 | 1.5219 | 0.91984 | 0.79903 | 0.61312 | 0.55211 | 0.45562 | 0.33985 | 0.24867 | 0.24188 | 0.18969 | 0.17259 | 0.1234 | 0.07722 | 0.05892 |
| Proporção da Variância | 0.4849 | 0.2200 | 0.1363 | 0.04977 | 0.03756 | 0.02211 | 0.01793 | 0.01221 | 0.00679 | 0.00364 | 0.00344 | 0.00212 | 0.00175 | 0.0009 | 0.00035 | 0.0002 |
| Proporção Cumulativa | 0.4849 | 0.7049 | 0.8412 | 0.89095 | 0.9285 | 0.95062 | 0.96855 | 0.98076 | 0.98755 | 0.99119 | 0.99463 | 0.99675 | 0.9985 | 0.9994 | 0.99975 | 0.99995 |

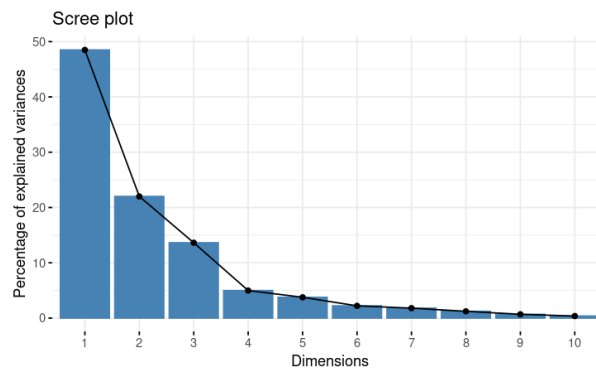


Figura 1: Scree plot da análise de PCA mostrando a porcentagem de variância explicada por cada componente.

A Tabela 1 mostra que a primeira componente principal (PC1) é responsável por 48,49% da variância. As componentes PC2 e PC3 contribuem com 22% e 13,63%, respectivamente, acumulando aproximadamente 84,12% da variância total nas três primeiras componentes. Essa concentração de informações também é evidenciada pelo Gráfico de Scree (Figura 1).

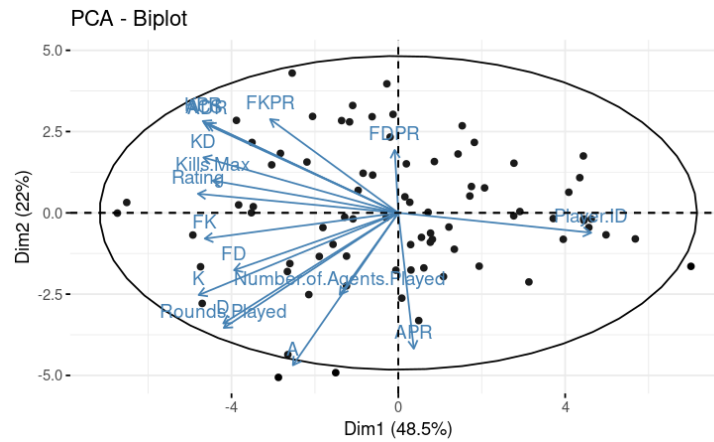


Figura 2: Biplot do PCA indicando a relação entre as variáveis e os componentes principais.

No Biplot do PCA (Figura 2), a direção das setas indica a correlação entre variáveis e componentes principais, e seu comprimento reflete a magnitude da influência.

Selecionei 'Rating', 'KD' (Kill/Death ratio) e 'ACS' (Average Combat Score) para o K-means com base em sua forte correlação com os componentes principais — PC1, PC2 e PC3. Estas métricas capturam o desempenho geral do jogador, a eficiência em combate e a contribuição em ações de jogo.

Em seguida, realizei a análise do Método de Elbow ("Método do Cotovelo"):

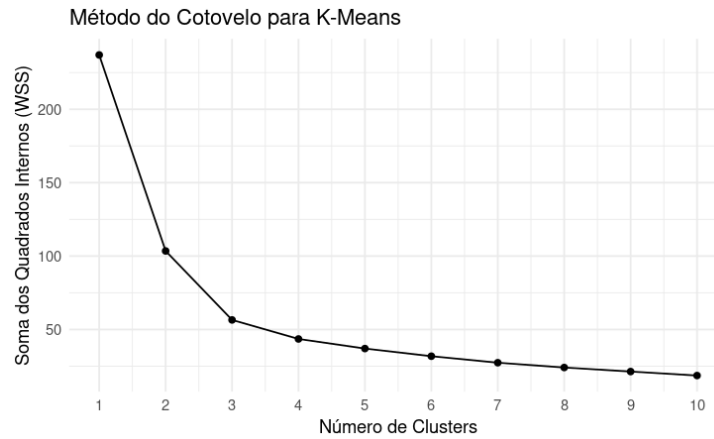


Figura 3: Método do Cotovelo para K-Means, mostrando o WSS por número de clusters.

Com base no gráfico acima, observei que o ponto de inflexão ocorre em três clusters. Este ponto representa o equilíbrio onde há uma diminuição significativa na soma dos quadrados internos (WSS), indicando que a adição de mais clusters não levaria a melhorias significativas na variância explicada, e evitando o sobreajuste.

Por fim, segue a análise dos clusters:

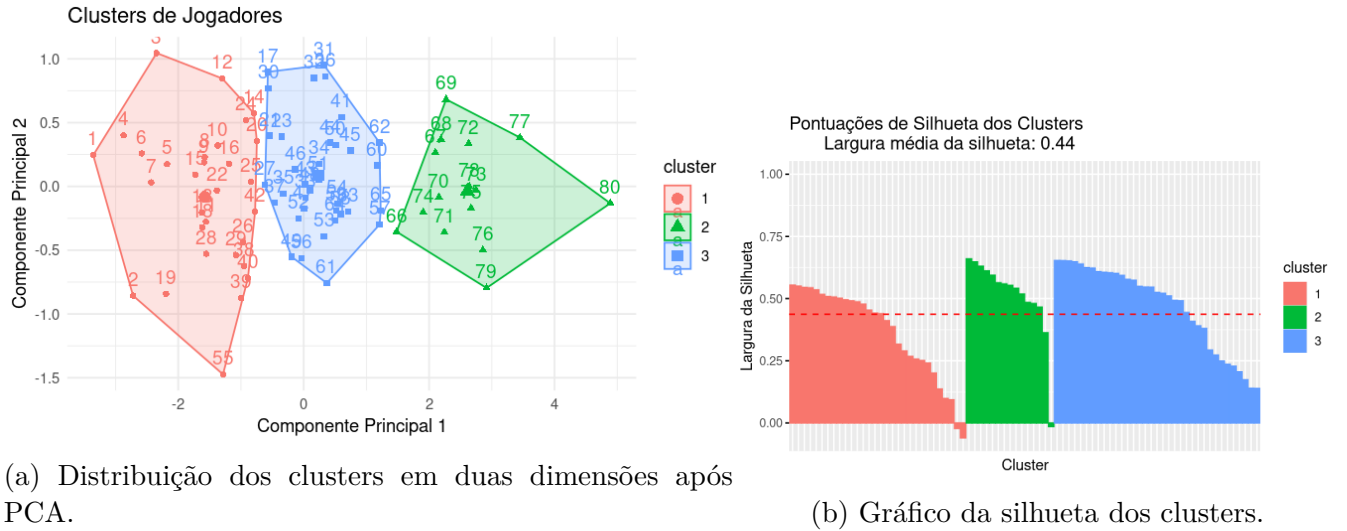


Figura 4: Visualizações complementares dos resultados da PCA e análise de cluster.

Tabela 2: Médias dos clusters para as variáveis de desempenho.

| Cluster | Rating | KD | ACS |
|---------|--------|------|--------|
| 1 | 1.08 | 1.14 | 223.52 |
| 2 | 0.73 | 0.70 | 147.36 |
| 3 | 0.96 | 0.93 | 187.21 |

Tabela 3: Frequência das funções no Cluster 1.

| Função | n |
|-------------|----|
| Controlador | 1 |
| Duelista | 13 |
| Iniciador | 7 |
| Sentinela | 9 |

Tabela 4: Frequência das funções no Cluster 2.

| Função | n |
|-------------|----|
| Controlador | 11 |
| Duelista | 0 |
| Iniciador | 2 |
| Sentinela | 2 |

Conclusão

A partir dos testes realizados, pode constatar que:

- A análise revelou que o **Cluster 1**, predominante de Duelistas (Tabela 3), apresenta as maiores médias em *Rating* (1.08), *KD* (1.14) e *ACS* (223.52), como mostra a

Tabela 2. Este cluster não só supera os outros significativamente em desempenho, mas também valida a importância dos Duelistas.

- Os **Clusters 2 e 3**, que incluíam Controladores, Sentinelas e Iniciadores, mostraram desempenho inferior ou mediano. Especificamente, o Cluster 2, que teve a ausência de Duelistas, registrou as menores médias de desempenho, destacando a influência limitada das outras classes nas métricas chave comparadas aos Duelistas.
- O método do Cotovelo (Figura 3) confirmou que a utilização de três clusters era ideal para este conjunto de dados. A distinção adequada entre os clusters foi evidenciada pelo valor médio da silhueta (0.44) (Gráfico 4b).

Portanto, conclui-se que os Duelistas desempenham um papel vital nas partidas de Valorant, exibindo um impacto significativamente maior do que as outras classes. Esta influência é claramente refletida nos altos índices de desempenho do Cluster 1, o que os caracteriza como os agentes mais decisivos.

Referências

- [1] Universidade Estadual de Campinas. *Notas de Aula de ME921, G. Ludwig*. Disponível em: https://moodle.ggte.unicamp.br/pluginfile.php/3926125/mod_resource/content/1/aula06.pdf.
- [2] Goyal, Samarth. *Agent Role Recommender System - Valorant*. Medium, 2020. Disponível em: <https://medium.com/@goyalsamarth/agent-role-recommender-system-valorant-1bfd66569a2>.
- [3] Likas, A., Vlassis, N., & Verbeek, J. J. *The global k-means clustering algorithm*. Pattern Recognition, 2003. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0031320302000602>.
- [4] Norouzi, Mohammad, et al. *Cartesian K-Means*. CVPR, 2013. Disponível em: https://www.cv-foundation.org/openaccess/content_cvpr_2013/papers/Norouzi_Cartesian_K-Means_2013_CVPR_paper.pdf.