

# Trabalho Final ME921

Gustavo Leiva (RA 217418)

## Introdução

### O jogo

Terraria é um jogo eletrônico desenvolvido pela Re-Logic e lançado oficialmente no dia 16 de maio de 2011. Originalmente, foi lançado para Microsoft Windows, mas ao longo dos anos foi portado para diversas plataformas, sendo que são agrupadas em: Desktop, Console, Mobile e Nintendo 3DS. As versões são em sua maioria sincronizadas nas atualizações, mas tem suas particularidades. Para a análise do meu trabalho, estudei apenas a versão de desktop. A primeira versão do jogo foi a 1.0. Até a data que esse trabalho foi realizado, a versão mais recente é a 1.4.2.3, lançada em 16 de maio de 2021. Essa é a versão que utilizei para as análises.

O jogo é um sandbox de mundo aberto 2D de ação e aventura com vários elementos de RPG. O jogador começa em um mundo gerado aleatoriamente a partir de alguns parâmetros e pode fazer o que quiser - sendo que o mais comum é começar coletando madeira e alguns recursos. Conforme o jogador avança em seus itens, ele pode derrotar *bosses*, que liberam mais itens, áreas e eventos especiais do jogo. Um ponto de extrema importância do jogo é quando o jogador derrota um *boss* chamado *Wall of Flesh*, que faz com que uma nova parte do jogo comece: o Hardmode. Nessa parte, aparecem dezenas de novos inimigos muito mais fortes, novos *bosses*, novos biomas e itens bem mais poderosos. Ao longo da jogatina, o jogador pode escolher focar em alguma classe de batalha: corpo a corpo, a distância, magia e invocação. Cada classe tem inúmeros itens, armas, armaduras e equipamentos especiais.

Para esse estudo, decidi focar apenas em um tipo de arma da classe corpo a corpo: espadas. O motivo foi que esse tipo de arma é o mais consistente entre as observações. Itens mágicos, por exemplo, pode causar vários efeitos: explosões, dano em área, dano direcionado, dano espalhado e tudo isso com muita aleatoriedade em como o dano é calculado. Portanto, a escolha da espada se dá pela facilidade do uso da arma: apertar o botão de ataque e só, na maioria das observações.

O objetivo do estudo é fazer um agrupamento das espadas a partir de várias variáveis que cada uma das espadas contém, contando também com uma análise descritiva prévia para entender como é a distribuição das

espadas para cada variável.

## Dados

Os dados foram capturados a partir de um webscraping com o pacote `rvest` na página da web <https://terraria.fandom.com/wiki/Swords>, que faz parte da wiki oficial de Terraria. Foram necessários vários tratamentos nos dados, e são eles:

- Mudança de nome das variáveis;
- Capturar somente estatísticas das espadas para a versão de desktop do jogo;
- Retirar etiquetas das variáveis `Use time` e `Knockback` para ficar somente com os valores numéricos;
- Tratamento das variáveis `Autoswing` e `Hardmode` para transformar em fatores (com 1 e 0) ao invés de símbolos;
- Deleção das variáveis `Sell` e `Source`, que não iriam ser úteis para a análise que eu procurava.

Foram no total 97 observações, sendo que as primeiras 10 podem ser vistas na Tabela 1. `Damage` se refere ao dano base que a espada causa no inimigo, `Use time` se refere a quão rápido o jogador pode infligir dano (quanto menor o valor, maior a velocidade), `Knockback` é quão distante o inimigo é lançado ao receber dano, `Crit` é a chance (portanto, entre 0 e 1) de determinado ataque ter um dano crítico, ou seja, infligir o dobro de dano que o comum, `Autoswing` é uma habilidade inerente de algumas espadas que faz com que com apenas um clique segurado a espada inflija repetidos danos, sem precisar dar um clique para cada “swing”, `Hardmode` é uma indicação se aquela espada pode ser obtida apenas no *Hardmode* ou no *pré-hardmode* e, finalmente, `Rarity` é uma sinalização de quão rara é aquela espada de “00\*” até “10\*”. Por exemplo, espadas que são feitas manualmente a partir de minerais ou itens pegos facilmente são menos raras que espadas que vêm a partir da derrota de algum inimigo que só aparece em certa circunstância do jogo, portanto, essa variável indica quão difícil de conseguir é aquele item independente da fase do jogo.

## Metodologia

A clusterização escolhida foi a hierárquica divisiva, através da função `diana` do pacote do R `cluster`, baseado na metodologia de Kaufman and Rousseeuw 2009. Como tenho diferentes tipos de variáveis (discretas e contínuas), calculei a matriz de dissimilaridades a partir da distância de *Gower* com a função `daisy` do mesmo pacote. Eu

Tabela 1: Cabeça da base de dados tratada

Name	Damage	Use time	Knockback	Crit	Autoswing	Hardmode	Rarity
Adamantite Sword	56	26	6.00	0.04	1	1	04*
Beam Sword	52	20	6.50	0.04	0	1	04*
Bee Keeper	26	20	5.30	0.04	1	0	03*
Blade of Grass	28	30	3.00	0.04	0	0	03*
Bladed Glove	12	8	4.00	0.04	1	0	02*
Bladetongue	55	28	5.75	0.04	1	1	05*

escolhi o tipo divisivo de clusterização pois penso que, primeiramente, todas as espadas estão dentro de um único grupo de itens: espadas, claro. Como existem outros grupos de itens (arcos, revólveres, itens mágicos...), o que faz mais sentido pra mim é pensar que espadas são grande grupo único que pode ser subdividido em vários outros subgrupos, necessariamente nessa ordem.

## Resultados

### Análise descritiva

A partir da Tabela 2 podemos ver que a média do dano das espadas pré- *hardmode* é muito mais baixa que a média das espadas *hardmode*, enquanto o *knockback* e o *use time* se mantem bem parecidos. Essas características se mantêm quando olhamos o desvio padrão, que indica que o dano das espadas do *hardmode* são muito mais dispersos que o dano das do pré. Inclusive, o primeiro quartil do dano das espadas do *hardmode* consegue ser mais alto até que a espada do pré-*hardmode* com mais dano - 49 versus 42. A respeito do *use time*, vemos que tanto no pré quanto no *hardmode* tem uma espada com o mesmo valor (que é 8). Ao ponto que o *knockback* não muda muito entre pré e *hardmode*, temos uma mudança significativa no valor máximo, que para o segundo é de 20 pontos.

Tabela 2: Estatísticas descritivas de Damage, Use time e Knockback agrupado por Hardmode

Hardmode	Variável	n	Média	D.P.	Q1	Q2/Med	Q3	Mín.	Máx.
0	Damage	62	19.11	11.29	10.0	15.000	25.0	5.0	42
	Knockback		4.18	1.17	3.0	4.375	5.0	0.5	7
	Use time		20.73	5.31	20.0	21.000	23.0	8.0	35
1	Damage	34	71.18	39.42	49.0	57.500	85.0	11.0	200
	Knockback		5.96	2.74	4.5	6.000	6.5	3.5	20
	Use time		21.06	5.21	18.0	22.000	25.0	8.0	29

Analisando a Figura 1(a), podemos observar que existem espadas na maior parte da extensão de Use time para espadas com dano inferior a 25, e a partir de 25 o Use time se concentra entre 15 e 30. Existem alguns

aparentes outliers, que foram etiquetados. A Meowmere e a Flying Dragon tem um dano similar, mas enquanto a primeira tem velocidade bem alta, a segunda é um bem mais lenta. A Psycho Knife ganha como uma arma extremamente rápida e com dano relativamente alto. Isso faz com que o seu dano por segundo seja bastante rápido também. Já a Purple Clubberfish é a espada mais lenta do jogo e com um dano não tão impressionante.

Pela Figura 1(b) observamos espadas bem menos dispersas na variável Knockback, o que faz com que os outliers sejam bem mais dramáticos. No lado esquerdo, que contém espadas com quase nenhum *knockback* (ou seja, ao bater no inimigo, ele fica praticamente parado) temos a Gladius e a Ruler, que são espadas com dano baixíssimo também. Do lado direito, temos a Slap Hand, que tem um *knockback* absurdo. Provavelmente essa arma foi programada para ser uma “*gimmick weapon*”, ou uma arma de piada, pois basicamente ela arremessa qualquer inimigo voando para longe, ao mesmo tempo que tem um dano bem baixo para uma espada do *hardmode*.

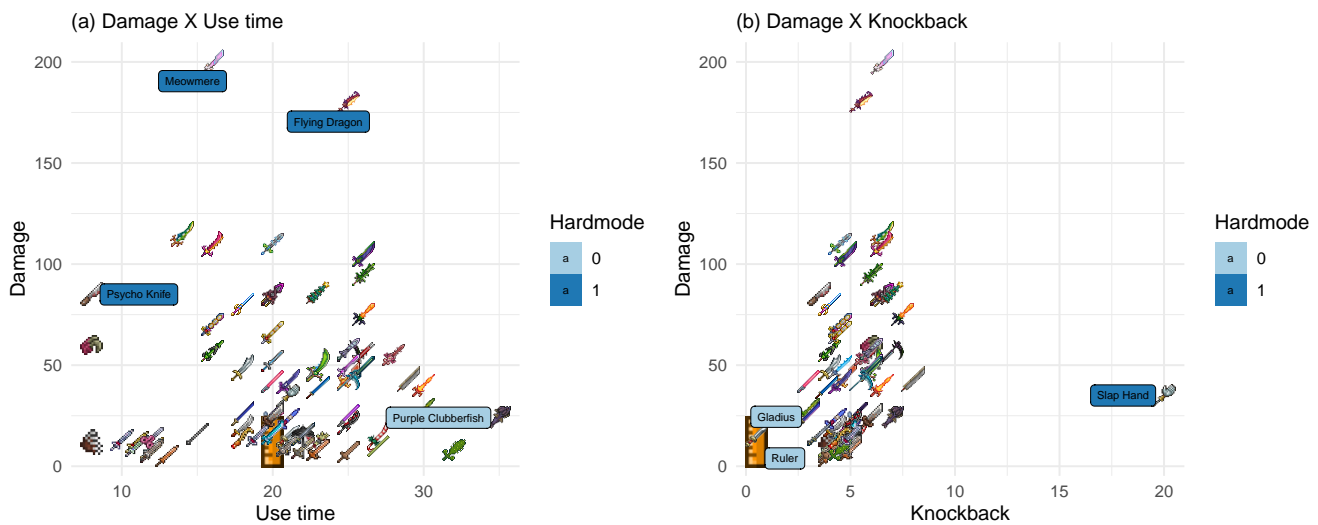


Figura 1: Gráficos de dispersão para as espadas

## Clusterização

Como dito anteriormente, o tipo de clusterização escolhido foi a hierárquica divisiva com dissimilaridades calculadas a partir da distância de *Gower*. Para a escolha do melhor número de clusters, observei os gráficos na Figura 2. Através do gráfico de cotovelo, vemos que o número ideal são 14 clusteres. Ao observar esse valor de *k* no gráfico de silhueta, vemos que é o primeiro valor com uma silhueta média bastante elevada, então acredito que 14 clusteres é realmente o ideal. Além disso, pensando no *gameplay* de *Terraria*, 14 clusteres fazem bastante sentido pensando nas diferentes partes do jogo em que você consegue obter cada item e os diferentes usos de cada espada.

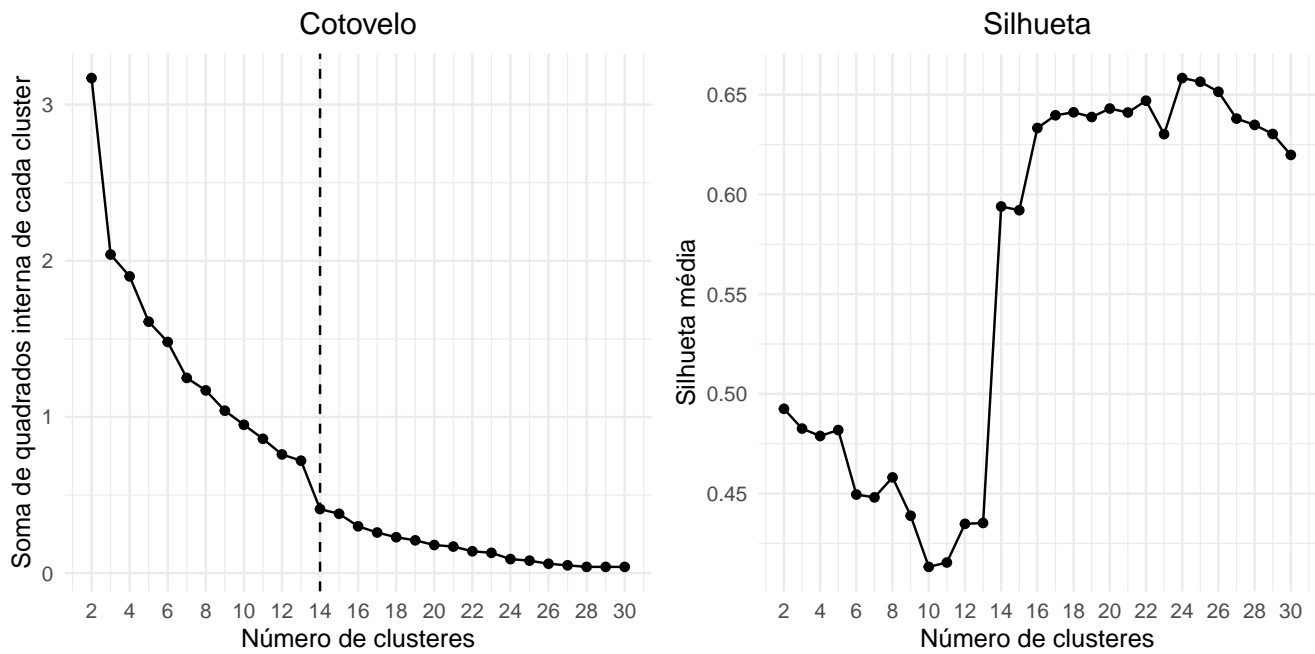


Figura 2: Gráfico de cotovelo e silhueta média ao longo de diferentes tamanhos de cluster

A Figura 3 mostra que o Cluster 8 e 14 tem silhuetas iguais a zero por terem apenas uma espada. O 10 e o 13 são os que tem as piores silhuetas,

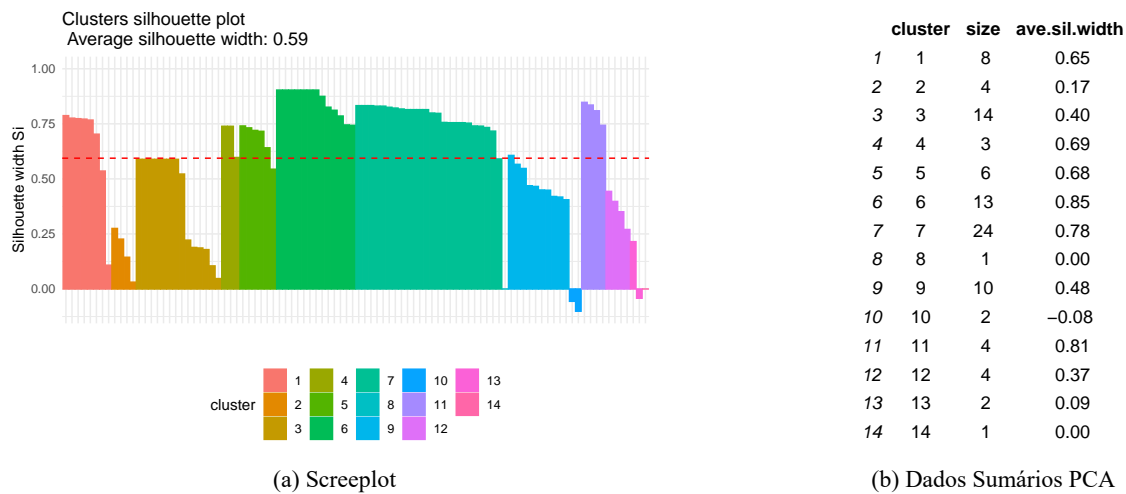


Figura 3: Gráfico de silhueta para k = 14

O dendrograma final ficou da forma que podemos ver na Figura 4. Nele, podemos ver que a Meowmere, a espada com maior dano, ficou sozinha num cluster com a Star Wrath.

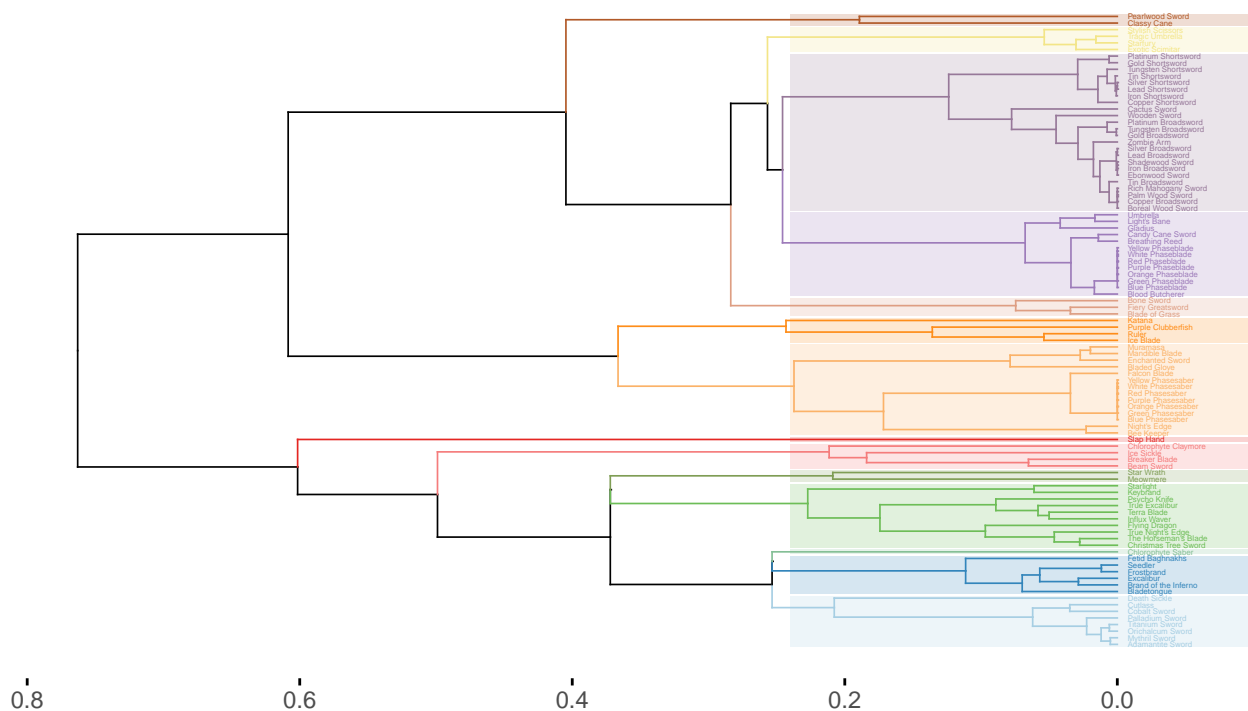


Figura 4: Dendrograma para k = 14

A Figura 5 mostra o comportamento das quatro variáveis numéricas do banco de dados separado por cada cluster. O Cluster 13, que contém Meowmere e Star Wrath são o que tem o maior dano em média. Já o Cluster 14, que contém apenas a Slap Hand se destaca por ter essa arma que, como vimos na análise descritiva, possui um knockback altíssimo. Além de knockback, a Slap Hand também possui a maior taxa de dano crítico, o que faz com que ela ganhe mais um motivo para ter um cluster só para ela. Para a variável Crit, também, podemos ver que o Cluster 9 (que contém 10 espadas) possui uma taxa de dano crítico elevada em relação aos demais clusteres. Além disso, esse mesmo Cluster ganha o segundo lugar quando olhamos as espadas com maior dano. A variável Use time não parece trazer muitas informações.

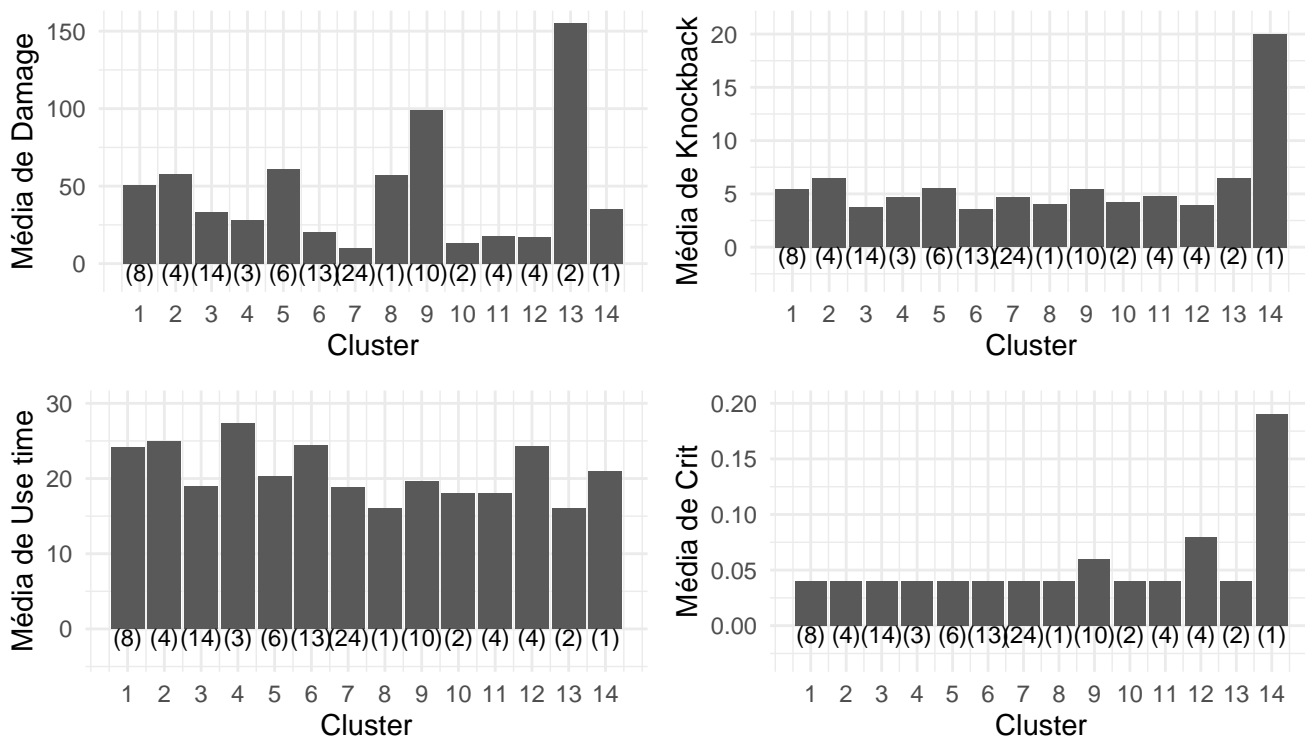


Figura 5: Médias das variáveis numéricas por cluster

A Figura 6 mostra como as variáveis categóricas se comportam para cada Cluster. Vemos que cada Cluster possui somente um valor das duas variáveis binárias (Autoswing e Hardmode). O Cluster 7 é praticamente o único que tem espadas da raridade 00, ou seja, as espadas mais comuns e simples. Por esse mesmo motivo também é o cluster com a maior quantidade de espadas. Nele se encontram praticamente todas as espadas construídas pelo próprio jogador com materiais muito simples de se encontrar no mundo de Terraria. Essas espadas costumam também ter características muito próximas entre si. A única espada que faz parte dessa raridade mas não está no Cluster 7 é a Pearlwood Sword. O motivo disso é que essa é a única espada com raridade 00 que é encontrada somente no Hardmode, portanto teve de ser alocada em um Cluster separado, em conjunto com a Classy Cane no Cluster 10. Essa segunda também tem dano baixo porém é apenas do Hardmode. O fato delas serem espadas bastante simples se confirma quando olhamos na figura anterior que a média desse mesmo cluster para o dano é a menor de todas. O Cluster 9 possui todas as espadas da raridade 08.

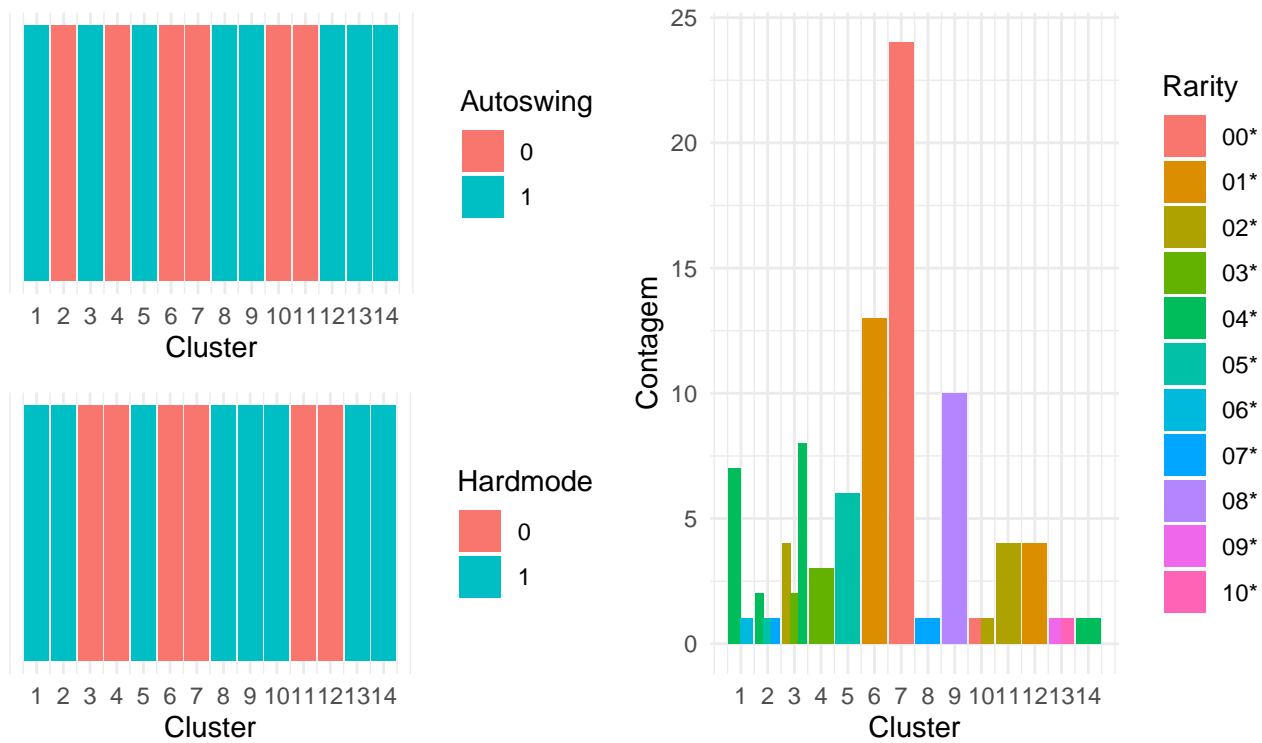


Figura 6: Características das variáveis categóricas por cluster