

# STEAMS Methodology of Playing Strategic Empire Video Game

Mason Chen\*

Stanford OHS, San Jose, CA, USA.

\*Corresponding author. Tel.: +1 699 246 2962; email: mason05@ohs.stanford.edu

Manuscript submitted January 3, 2020; accepted March 24, 2020.

doi: 10.17706/ijapm.2020.10.3.153-159

---

**Abstract:** This paper would introduce a STEM+ Methodology STEAMS by adding Artificial Intelligence and Statistics. In modern Big Data World, Artificial Intelligence algorithm is powerful for engineering problem solving to discover the complicated science or/and mechanics. Separating Statistics from Math can draw practical decision and conduct risk assessment. To demonstrate the powerful STEAMS methodology, a case study of playing a strategic Empire Four Kingdoms video game is deployed. STEAMS is powerful for modern educator, teacher, and student to develop their critical thinking and time management. Integrating 6 STEAMS elements are powerful to understand more insight information efficiently and effectively.

**Key words:** Artificial intelligence, STEAMS, strategy, video game.

---

## 1. Introduction

This paper will demonstrate how to use Statistical Modelling to help student study the Environmental Science when playing the video games. There are three sessions to be included in the project introduction: (1) Introduce Empire Four Kingdoms Video Game, (2) Environmental Science, and (3) Build a Powerful Troop to Win the Battle.

### 1.1. Introduce Empire Four Kingdoms Video Game

Playing video game is becoming a critical portion of social activities for most middle school and high school students. However, parents are worrying that kids may play video games too much and most video games may not help develop their critical thinking and teamwork concept. The objective of this project is to convert playing video games to become conducting Projects. Students can learn Environmental Science and Statistics while playing video games. Authors have searched several video games and picked the Empire Four Kingdoms video game not based on the commercial rating. But based on the potential of applying statistical data-driven and engineering problem-solving approach due to its embedded database which can record 7 characteristics of each military unit. The main challenges of playing this video game are: (1) make your dream a reality and become a king and castle lord with the Empire, (2) Produce new resources and build your small castle into a mighty fortress, (3) recruit a powerful army to conquer more land for your kingdom and defend it against enemy attacks, and (4) form alliances with friends to defeat your enemies or fight epic battles against countless other players. Your strategy determines whether you become a legendary king or remain a simple peasant.

## 1.2. Environmental Science

Environmental resource is very critical in the early stage to choose the location of your main castle and the first few outposts (supporting resources to main castle). There are three types of natural resources: Stone, Wood and Food. Resources are used in to construct buildings. There are 3 main types of resources (wood, stone, and food). Resources are produced over time through production buildings Each castle and outpost have its own combination of three natural resources. The functions of each resource in game are described below: Stone can be obtained through establishing a Stone Quarry. Stone is used in construction, creating tools for castle siege or defence. The abundance of the resource can be seen in the form of rock formations. To produce stone, a stone quarry is required - along with a stone quarry in the outpost. Wood is used in construction, creating tools for castle siege or defence through constructing a Woodcutter building. A high wood production area has lush forests around it. Wood is produced by woodcutters only if there is a forest available in the outpost. Production of wood depends on the number of woodcutters, along with what upgrade level they are at. Food is for feeding your citizens, soldiers. This seems to be, among the 3 basic resources, the most important and hard to maintain resource due to just by having citizens and soldiers around, because they need to eat, this resource continuously diminishes at a rate equal to how many your population + the soldiers are. Food is required by your troops to survive. Farmhouses and Windmills increase food production. Food is provided at outposts, though can be 2, 6, or even 8 if you are lucky to find one. The castle type of your first few castles may impact the Kingdom's personality and growing strategies to survive in the neighbor surrounding by weaker or mighty kingdoms.

## 1.3. Build a Powerful Troop to Win the Battle

To expand your Kingdom territory, building a powerful troop is the highest priority. What kind of Troops needed at each Phase? More than 40 types of Troop Units are available: some are for defence and some are for attack. There are seven characteristics for each troop unit: (Melee/Range Défense, Melee/Range Attack, Travel Speed, Looting Capacity, Food Consumption). For example, Deathly Horror & Demon Horror units are some of the most powerful units you can get in the game. They can be obtained when you buy a lot of rubies or attain a rather high position in a tournament or event. These units are mostly hotly contested since having them in an attacking force greatly increases your chances of success. A summary of 7 characteristics across most popular troops is shown in Table 1.

Table 1. Summary of 7 Troop Characteristics

	Melee Attack	Melee Defense	Ranged Attack	Ranged Defense	Looting Capacity	Travel Speed	Food Consumption
Archer	0	53	10	55	9	25	2
Bowman	0	8	24	24	13	75	2
Composite Bowman	0	74	16	159	2	25	3
Crossbowman	0	6	36	16	22	73	2
Crossbowman of the Elite Guard	0	14	129	23	29	28	4
Crossbowman of the Kingsguard	0	14	121	23	29	28	4
Deathly Horror	0	15	162	24	40	28	5
Demon Horror	185	19	0	5	35	28	5
Elite Crossbowman	0	14	129	23	29	28	4
Elite Knight	140	18	0	5	33	28	4
Flame Bearer	14	170	0	68	3	25	3

## 2. STEAMS Approach

Video games have enormous mass appeal, reaching audiences in the hundreds of thousands to millions. They also embed many pedagogical practices known to be effective in other environments. This article [1]

reviews the sparse but encouraging data on learning outcomes for video games in science, technology, engineering, and math (STEM) disciplines, then reviews the infrastructural obstacles to wider adoption of this new medium. Authors have further expanded STEM to deploy STEAMS project management by adopting six Elements: (1) Science, (2) Technology, (3) Engineering, (4) Artificial Intelligence, (5) Mathematics, and (6) Statistics.

## **2.1. Science**

There are several Environmental Science in this Empire Four Kingdom video game including: Stone, Wood, and Food. The growth of population, cities and small-scale industry placed greater demands on the environment for resources- wood, metals, and stone. Highlighted explanation accompanied by poor land management resulted in widespread destruction of the natural environment. Many fertile areas were destroyed by over-grazing, excessive timber cutting, and poor agricultural practices. Through playing this Empire Video Game, students can build their environmental science on protecting the natural resources such as Stone and Wood.

## **2.2. Technology**

The players needed to upgrade their castles through key Technologies through two Categories of Buildings: Civil and Military. The most popular civil buildings are: Stone Quarry, Wood Cutter and Dwelling and Bakery. The Woodcutter produces wood in your castles and outposts. Wood is essential for construction of buildings and the production of tools. The Stone quarry is an important building because it produces stone, which is needed for tools and buildings. A dwelling is a building that determines your population. The more dwellings you have, and the higher the level they are at, will affect how high your population is. The bakery is a civil building that decreases food consumption of soldiers. It has 8 levels giving a final food reduction of 40%. The most popular military buildings are: Barracks, Defense Workshop, Siege Workshop, and Taverns. Barracks enable the player to recruit soldiers. When the barracks are upgraded, a variety of soldiers are unlocked for recruitment. The higher the level of the barracks, the higher the recruitment speed bonus and unit capacity per slot. The Defense Workshop is a building where defense tools are produced. Different defense tools can be made to protect your castle from incoming attacks. Siege workshop is a building where siege tools can be produced. More siege tools are unlocked with each upgrade. The Tavern is a building which sets the maximum possible number of spies you can have. Taverns are required for an espionage. Learning these Technologies are essential to building a great Kingdom.

## **2.3. Engineering**

Engineering problem solving can be applied by setting the game strategy, conducting systematic root cause analysis and apply project management trilogy (schedule, cost, quality) constraints. Each kingdom has its own challenge, strength and weakness. Without adopting engineering thinking, video game players may just play hard on the try and error mode. Our project will define the Kingdom as the beginner and concentrate on Internal Affairs, Enhance Defense Power, and Control Food Consumption. Without well defining the project scope through the systematic engineering problem solving techniques, most players may just blindly develop their kingdoms into different kind of crisis.

## **2.4. Artificial Intelligence**

Artificial Intelligence clustering analysis can help discover the affinity grouping patterns among 40+ troops units to help build the best troops based on kingdom strategy. Hierarchical Clustering Analysis (HCA) [2] was used to further analyze the troop patterns. In data mining and statistics, hierarchical clustering (also called hierarchical cluster analysis or HCA) is a method of cluster analysis which seeks to build a hierarchy of

clusters. Strategies for hierarchical clustering generally fall into two types [3]: (1) Agglomerative: This is a "bottom up" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy and (2) Divisive: This is a "top down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy. In the general case, the computing time of the Agglomerative approach is faster than the Divisive approach. Optimal efficient agglomerative methods have been developed to significantly improve the computing algorithm for large data sets [4], [5]. The main objective of this analysis was to search for the degree of similarity among troops based on 7 characteristics. The Agglomerative approach can identify a clustering pattern faster and more accurately. The Divisive approach may not split the 40+ troop units into clusters which are more concentrated on the bottom level efficiently. Therefore, the authors chose the Agglomerative approach. This approach builds the hierarchy from the individual elements by progressively merging clusters based on a defined distance metric (Euclidean distance). The distance is calculated by the discrepancy of scores among the seven characteristics. This HCA approach can pair the troop unit with similar score patterns and use clustering to group units. JMP statistical software was used to calculate the closest distance (the affinity) among all potential pairs, and grouped the first pair, at the strongest affinity (based on their similar score pattern). The linkage criterion determines the distance between sets of observations as a function of the pairwise distances between observations [6]-[8].

Authors used Data Mining Cluster and Dendrogram to group the similar troops units into eight clusters as shown in Fig. 1. Then, Kingdom can select the best troop unit from each Cluster to build a more diverse troop which can have enough power across all 7 attributes. In this paper, the top 3 attributes are: Higher Melee/Range defence Power, and Lower Food Consumption.

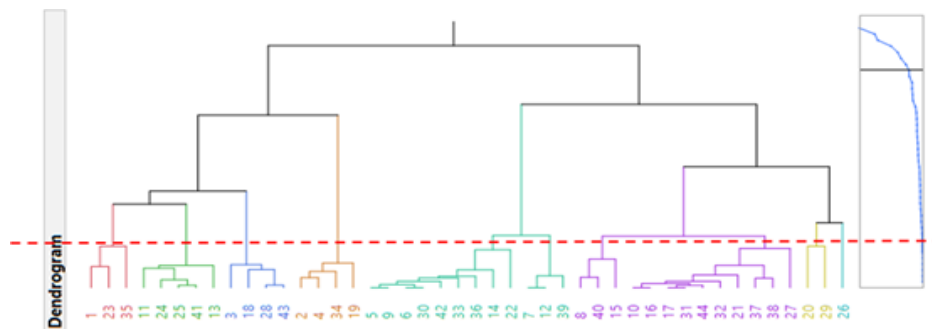


Fig. 1. Cluster analysis of troop types.

Each eight clusters identified has its own identity across 7 troop characteristics as shown in Table 2. For example, 1st cluster is about average across all 7 characteristics. 2nd and 3rd cluster are stronger on the Melee/Ranged Defense. The 4th cluster has faster speed and less food consumption. The remaining four clusters are good at attack and looting. Based on the kingdom strategy set previously, the 2nd-4th clusters may fit to the kingdom strategy better than the other clusters.

Table 2. Cluster Summary of Troop Characteristic

Cluster	Count	Melee Attack	Melee Defense	Ranged Attack	Ranged Defense	Looting Capacity	Travel Speed	Food Consumption	Characteristic
1	3	20.3	43.0	8.7	37.0	12.0	27.0	2.7	About Average
2	5	14.6	153.0	0.0	61.4	13.6	28.0	3.0	Melee Defense + Ranged Defense
3	4	0.0	66.3	16.8	139.3	13.0	27.5	3.0	Ranged Defense + Melee Defense
4	4	16.0	19.5	15.0	13.5	20.3	74.0	2.0	Travel Speed + Less Food
5	12	0.0	14.9	134.8	23.5	32.7	29.3	4.3	Ranged Attack
6	13	145.7	20.8	0.0	8.8	32.2	28.5	4.2	Melee Attack
7	2	139.5	14.5	0.0	3.5	80.0	29.0	5.0	Melee Attack + Looting
8	1	0.0	8.0	144.0	14.0	90.0	28.0	6.0	Ranged Attack +Looting

## 2.5. Mathematics

Previous 2.4 clustering method has identified three car stage groups separated. The clustering patterns were identified based on clustering distance algorithm of calculating the dissimilarity of 7 troop characteristics among 40+ troop types. This section will study the mathematics of various clustering distance algorithms. There are several cluster algorithms: (1) Average, (2) Centroid, (3) Ward, (4) Single, and (5) Complete (Citation). Will these 5 different clustering algorithms have the same results? If different, how to select which algorithm to explore the clustering patterns best? In Fig. 2, three existing clusters (Green, Yellow, Red) are going to join next. Which two clusters should bond first? The joining sequence is determined by the clustering distance algorithms. Centroid, Single, and Complete algorithms are compared show in Fig. 2. The Centroid algorithm connects Green cluster and Yellow cluster through the purple line connecting the two cluster means (purple triangles). The Single algorithm groups Green cluster and Red cluster by the closest points between these two clusters. The Complete algorithm groups Yellow cluster and Green cluster by the farthest points between these two clusters. Depending on which distance algorithm chosen, the clustering sequence and pattern may be different. We must dive into the mathematical calculations for each clustering distance algorithm and understand the benefits and limitations of each algorithm to choose the best algorithm to draw reliable clustering patterns and results. We will compare five major clustering distance algorithms [9], [10]. The calculations of the five different clustering algorithms are shown in Fig. 3. The first algorithm is Average which is the distance pair divided by the number of distances. Since the Average algorithm compares the average distances, it typically joins smaller and similar variances. The 2nd algorithm is centroid which calculates the distance between the cluster means [11].

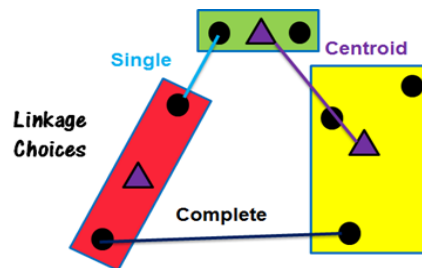


Fig. 2. Diagram of the centroid, single, and complete clustering methods.

**Average Linkage** Distance for the average linkage cluster method is:

$$D_{KL} = \frac{\sum_{i \in C_K} \sum_{j \in C_L} d(x_i, x_j)}{N_K N_L} \quad \text{Average}$$

**Centroid Method** Distance for the centroid method of clustering is:

$$D_{KL} = \|\bar{x}_K - \bar{x}_L\|^2$$

**Ward's** Distance for Ward's method is:

$$D_{KL} = \frac{\|\bar{x}_K - \bar{x}_L\|^2}{\frac{1}{N_K} + \frac{1}{N_L}} \quad \text{ANOVA}$$

**Single Linkage** Distance for the single linkage cluster method is:

$$D_{KL} = \min_{i \in C_K} \min_{j \in C_L} d(x_i, x_j) \quad \text{Minimum}$$

**Complete Linkage** Distance for the Complete linkage cluster method is:

$$D_{KL} = \max_{i \in C_K} \max_{j \in C_L} d(x_i, x_j)$$

Fig. 3. Clustering distance algorithms.

Among five algorithms mentioned, Centroid is the most robust algorithm to outliers. The 3rd algorithm

Ward uses the ANOVA sum/mean of squares (between divided by within). The Ward algorithm is Centroid divided by the degree of freedom [12]. Ward joins smaller numbers of observations and which is the most sensitive to outliers. The 4th Single uses the minimum distance, and therefore typically, joining larger variances/larger number [13]. Clusters (favor in Single algorithm) are large, elongated or irregular. Those clusters may have shorter distances with other similar clusters than with small-sized clusters. The last algorithm Complete joins clusters based on the farthest distance. It is more sensitive to moderate outliers and, very different from single algorithm. Complete algorithm normally joins smaller variances/smaller numbers of clusters. How will these algorithms impact the clustering patterns?

## 2.6. Statistics

This section will use Statistical Transfer Function to optimize the troops power on Melee/Ranged defence Power, and Control Food Consumption. Based on Kingdom Strategy, the JMP Desirability Functions are set to emphasize more on Defense and Food Consumption by setting LSL/USL and Importance as shown in Table 3.

Table 3. JMP Design Optimization Specifications and Desirability (Importance) Setting

Phase I	Direction	LSL	USL	Importance
Attack	Maximum	500		1
Defense	Maximum	1000		2
Loot	Maximum	500		1
Speed	Maximum	500		1
Food	Minimum		250	2

As shown in Table 4, the optimal troop quantity setting is chosen within the total unit constraint. 10 units each chosen from 2nd - 4th cluster as expected. 15 units from 7th cluster and 5 units from the 8th cluster were also chosen to maintain minimum Attack Power, and Looting Capacity. In general, Defense Units are chosen over Offense Units.

Table 4. Optimal Troop Quantity Setting

Unit	Qty	Melee Attack	Melee Defense	Ranged Attack	Ranged Defense	Looting Capacity	Travel Speed	Food Consumption
Archer	0	0	53	10	55	9	25	2
Renegade Bowmaster	0	0	71	16	53	15	32	3
Renegade Katana Warrior	5	14	165	0	66	16	32	3
Renegade Lancer	5	13	150	0	63	16	32	3
Renegade Spear Thrower	5	0	68	14	139	15	32	3
Veteran Longbowman	5	0	67	17	134	14	32	3
Maceman	5	38	38	0	6	32	73	2
Spearman	5	26	26	0	8	14	75	2
Renegade Arrow Thrower	0	0	18	148	27	25	50	4
Veteran Deathly Horror	0	0	17	175	26	50	30	5
Khan Guards	5	175	54	0	50	32	28	5
Renegade Sai Warrior	5	160	20	0	7	29	50	4
Marauder	0	113	18	0	4	70	30	4
Renegade Swashbuckler	5	166	11	0	3	90	28	6
Renegade Sai Ripper	5	0	8	144	14	90	28	6

## 3. Conclusion

The innovative STEAMS methodology is demonstrated through an Empire Video Game. Environmental science and natural resources (Stone, Wood, Food) are discussed. Civil and Military Buildings' Technologies are studied associate with the natural resources. Project scope is set through engineering thinking. Artificial Intelligence Clustering algorithm is adopted to discover the troop characteristic affinity pattern among 40+ melee and ranged troops. Mathematics of clustering algorithm and Statistics are applied to find the optimal Troop design to meet the project scope and objectives. This STEAMS method can be commonly extended to



various educational systems from K-G12, Higher Education to Industrial Education across most Scientific and Engineering fields.

### Conflict of Interest

No conflict of interest.

### Acknowledgment

The author acknowledges Dr. Charles Chen for his support.

### References

- [1] Merrilea, J. M. (2009). Video games: A route to large-scale STEM education? *Science*, 323(5910), 79-82.
- [2] Michael, R. A. (1973). *Cluster Analysis for Applications*. New York: Academic Press.
- [3] Rokach, L., & Oded, M. (2005). *Clustering Methods. Data Mining and Knowledge Discovery Handbook*. Springer US.
- [4] Sibson, R. (1973). Slink: An optimally efficient algorithm for the single-link cluster method. *The Computer Journal, British Computer Society*, 16(1), 30-34.
- [5] Defays, D. (1973). An efficient algorithm for a complete-link method. *The Computer Journal, British Computer Society*, 20(4), 364-366.
- [6] Szekely, G. J., & Rizzo, M. L. (2005). Hierarchical clustering vis joint between-within distances: Extending Ward's minimum variance method. *J. of Classification*, 22, 151-183.
- [7] Ward, J. H. (2005). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236-244.
- [8] Ma, et al. (2007). Segmentation of multivariate mixed data via lossy data coding and compression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9), 1546-1562.
- [9] Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, 45, 325-342.
- [10] Hartigan, J. A. (1981). Consistence of single linkage for high-density clusters. *Journal of the American Statistical Association*, 76, 388-394.
- [11] Florek, K., Lukaszewicz, J., Perkal, J., & Zubrzycki, S. (1951). Sur la liaison et la division des points d'un ensemble fini. *Colloquium Mathematica*, 2, 282-285.
- [12] Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38, 1409-1438.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



**Mason Chen** is currently a student at Stanford OHS, USA and serves as the student ambassador and webmaster for STEAMS. Having started STEAMS since its inception in 2014, he has held various roles such as the president of the student chapter from 2017 to 2019. Through STEAMS, he has published more than 20 conference proceeding papers. As first author, he has won numerous awards including the best conference proceeding paper award in the 2018 JMP Discovery Summit as well as finishing 1st place three times for the STEM presentation competition at IEOM conferences. He has also certified the IBM SPSS Statistics Level I, II, Modeler Level I, and IASSC Yellow Belt, Green Belt, and Black Belt.