1. Create clusters using hierarchical clustering and k-means clustering.  Include screen shots of your cluster results for both techniques.  You must try at least one hierarchical and one k-means technique, but you'll probably want to try a few different settings to determine your best result.

Sol:- Chart

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

A picture containing graphical user interface

Description automatically generated

Table

Description automatically generated

Calendar

Description automatically generatedTable

Description automatically generated

Chart

Description automatically generated

Table

Description automatically generated with medium confidence

k-Means clustering:

Table

Description automatically generated

Text

Description automatically generated with medium confidence

Chart, histogram

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated

Chart, bubble chart

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

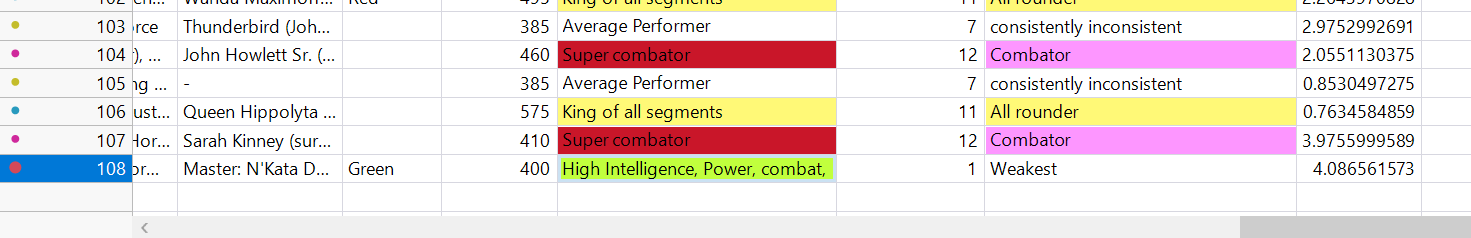
Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application, table, Excel

Description automatically generated



2. Select a "best" set of clusters.  Explain how you chose these clusters as the best.  Explain how they will be useful to Captain America.  This should be a more technical discussion.

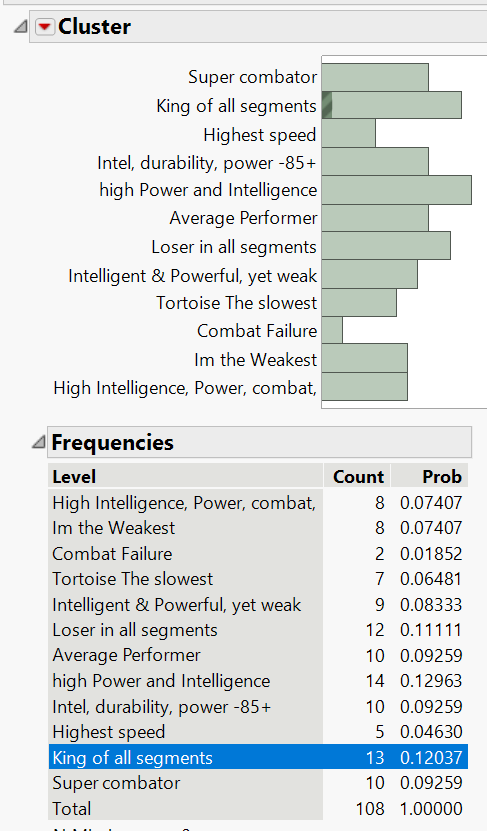
Sol :- Hierarchical clustering: In hierarchical clustering analyzing the outputs, I would say cluster 11 is the best among all, since it has all the aspects greater than 80+ and it is nothing less in any segment, when there is a requirement for all segments, I would not recommend someone who is bits and pieces in one segment and extraordinary in some aspect, the character has rely on its own attributes, hence I recommend him to choose from cluster 11, which is a king of all segments. This could help him in any situation. For example, if he chose someone who is good at combat but low at speed, that would create some trouble in the war. If he has someone who is extraordinary in all segments could fight the war irrespective of situations. If it’s a hierarchical I would recommend him cluster 11.

K-means clustering: Considering the analysis by K-means, I have analyzed 11th cluster and observed that cluster 5 had the all rounder capability among all other clusters, as explained above I would recommend the all-rounder among all segments than one character being extraordinary I one segment. Through parallel coordinate plot I have concluded that 5th cluster is the best and is allrounder among all.

3. Save your final clusters to the data table. Use graphical and descriptive tools to characterize the nature of the clusters.  Assign meaningful names to the clusters in the dataset.  Include a screen shot of the names and row counts.

**K-means clustering Hierarchical - clustering.**

Chart

Description automatically generated

4. Make a recommendation to Captain America based on your clusters.  Be specific in how he should use the results of your work.  Comment on which cluster(s) he should focus on and what he should do with them.  You do not have to recommend specific people.  This answer should focus more on the business interpretation and usage.  Think of this as a people analytics HR hiring problem with a focus on strengthening a team through diversity.

Sol:- I would recommend the Captain America to select people from cluster 11 in hierarchical and cluster 5 from k-means clustering, considering values among speed, intelligence, strength, durability, power and durability in both the clusters, they are allrounders and king in all segments who would be handy, and if with no discrimination based on Gender, color, and physical appearance we have selected the clusters only considering the skill set, yes if he need some specific skill set and extraordinary there are different clusters that have been sub categorized and can be picked upon from categories like super combater, super powerful, lightning speed, super talented.

5. Calculate the support, confidence, expected confidence, and lift for each of the Eye color 2 categories as an antecedent to an Alignment of Bad.  For example, solve for the rule “Eye color 2 = Blue ⇒ Alignment = Bad”.  Repeat the calculations for each eye color.

Table

Description automatically generated

Sol:- Red color:

|  |  |  |
| --- | --- | --- |
| support | 21/472 | 0.0445 |
| confidence | 21/38 | 0.553 |
| Expected confidence | 137/472 | 0.290 |
| Lift | 0.553/0.290 | 1.907 |

Blue:

|  |  |  |
| --- | --- | --- |
| support | 42/472 | 0.089 |
| confidence | 42/203 | 0.207 |
| Expected conf | 137/472 | 0.290 |
| Lift | 0.207/0.290 | 0.713 |

Purple, yellow:

|  |  |  |
| --- | --- | --- |
| support | 11/472 | 0.0233 |
| confidence | 11/23 | 0.478 |
| Expected conf | 137/472 | 0.290 |
| Lift | 0.478/0.290 | 1.648 |

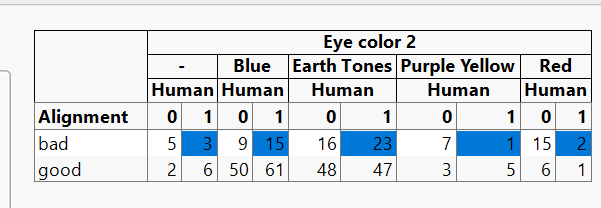
Earth tones:

|  |  |  |
| --- | --- | --- |
| support | 63/472 | 0.133 |
| confidence | 63/208 | 0.303 |
| Expected conf | 137/472 | 0.290 |
| Lift | 0.303/0.290 | 1.045 |

6. Which eye color provides the highest lift in terms of identifying bad characters correctly?

Sol:- Red - color provides the highest in terms of lift identifying the bad characters.

7. If you consider just humans rather than all characters, would you trust your finding in the previous question?  Why or why not?

Sol:- 

No I would not trust the process, since lift is highest in earth tones when its only for humans, considering below values I would say earth tones having lift of 1.807 have a greater bad alignment, and red contributes a lower lift for bad alignment in humans.

Lift for blue: 1.204

Lift for earth tones: 1.807

Lift for Purple yellow: 0.803

Lift for red: 0.16

Vinay Reddy Vangala

NetId: vrv22001