**Clustering af skud**

Vi har her lavet clustering på alle skud fra superligaen. Vi har her valgt at lave clustering med k-means. K-means kan kun lave clustering på numeriske data. Vi har udvalgt variablerne lokation x, lokation y, skud var mål og skud på mål. Skud var mål og skud på mål er true/false. Vi har derfor lavet dem binære så true = 1 og false = 0.

For at finde det optimale antal clustre har vi brugt albue metoden:

Chart, line chart

Description automatically generated

Ud fra dette har vi valgt at lave 3 clustrer.

|  |  |  |  |
| --- | --- | --- | --- |
| **Cluster** | 1 | 2 | 3 |
| **Antal** | 722 | 787 | 1157 |

I denne tabel kan vi se den gennemsnitlige værdier vores variabler i de forskellige clustre.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cluster** | **Gennemsnit Lokation X** | **Gennemsnit lokation Y** | **Gennemsnit mål/ikke mål** | **Gennemsnit skud på mål/ikke mål** |
| **1** | 82,7 | 65,5 | 0,05 | 0,35 |
| **2** | 83,8 | 33,8 | 0,06 | 0,34 |
| **3** | 87,1 | 49,5 | 0,18 | 0,4 |

Chart

Description automatically generated

På dette plot kan vi se hvordan de tre clustre er fordelt. Vi kan se at der er to store grupper, en mindre til venstre på x aksen og en større til højre på x aksen. Derudover er der en del outliers nederst og øverst på y aksen, som ligger i cluster 1 & 2. Cluster 3, som ligger i midten, er det cluster med den højeste målscore.

**Clustering af spillere**

Vi laver clustering på spillerne i Superligaen for at finde spillere som ligner hinanden. Dette kan bruges hvis Brøndby f.eks vil finde erstatninger for spillere de skal sælge.

Vi har udvalgt følgende variabler at lave clustering på:

Gennemsnitlige x lokation (skud)

Gennemsnitlige y lokation (skud)

Gennemsnitlige afleverings længde

Gennemsnitlige x lokation for afleveringer

Gennemsnitlige y lokation for afleveringer

Om et skud var blev til et mål

Om et skud var på målet

Om en afleveringer lykkedes

Disse variabler kan dog udelukke målmænd, da de typisk ikke laver hverken skud på mål eller afleveringer.

Vi har lavet 4 clustrer af spillere ud fra ovenstående variabler. Antallet af spillere per cluster er fordelt således:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cluster** | 1 | 2 | 3 | 4 |
| **Antal spillere** | 43 | 99 | 52 | 59 |

Denne tabel viser de gennemsnitlige værdier for de forskellige clustre:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cluster | Lokation X | Lokation Y | Afleverings Længde | Afleverings lokation X | Afleverings lokation Y | Succesfuldt mål | Skud på mål | Succesfuld aflevering |
| 1 | 47,94 | 77,27 | 20,58 | 52,5 | 50,34 | 0,08 | 0,4 | 0,83 |
| 2 | 63,28 | 52,32 | 21,12 | 53,66 | 51,11 | 0,1 | 0,41 | 0,82 |
| 3 | 57,27 | 23,89 | 20,60 | 52,50 | 50,37 | 0,1 | 0,36 | 0,82 |
| 4 | 42,63 | 45,58 | 20,66 | 52,41 | 50,16 | 0,13 | 0,36 | 0,82 |

Ud fra denne tabel kan vi se at spillerne i cluster 4 i gennemsnit har den højeste målscore, til trods for at deres forsøg på mål i gennemsnit er lavere end i de andre clustrer.

Chart

Description automatically generated

Plotter vi clustrerne kan vi se at de fleste spillere ligger tæt sammen, og især cluster 1 og 4 ligger tæt op af hinanden, hvorimod cluster 2 og 3 har nogle outliers.

Vi har undersøgt hvilke spillere Brøndby skal eller eventuelt vil af med til sommerens transfer, for at finde lignende spillere.

**Spillere Brøndby kan erstatte til sommers transfer[[1]](#footnote-1):**   
Mads Hermansen

Anis Slimane

Henrik Heggheim

**Kommende kontraktudløb[[2]](#footnote-2):**

Thomas Mikkelsen (sommer 2022)

Ante Erceg (sommer 2022)

Marko Divkovic (sommer 2022 - lejeaftale udløber)

**Anis** **Ben Slimane:**

Slimane er midtbanespiller, og ligger i cluster 2.

|  |  |
| --- | --- |
| **Variabel** | **Gennemsnitlige værdi** |
| Lokation X | 56,38 |
| Lokation Y | 51,1 |
| Afleveringslængde | 19,67 |
| Aflevering lokation X | 53,2 |
| Aflevering lokation Y | 51,52 |
| Målscore | 0,19 |
| Skud på mål | 0,38 |
| Afleveringssuccess | 0,85 |

En spiller som ligger i cluster 2 og minder om Anis Slimane er Abu Francis, som er midtbanespiller i FC Nordsjælland.

**Marko Divkovic:**

Divkovic er angriber, og ligger i cluster 2.

|  |  |
| --- | --- |
| **Variabel** | **Gennemsnitlige værdi** |
| Lokation X | 66,31 |
| Lokation Y | 58,75 |
| Afleveringslængde | 20,84 |
| Aflevering lokation X | 52,76 |
| Aflevering lokation Y | 46,21 |
| Målscore | 0,21 |
| Skud på mål | 0,57 |
| Afleveringssuccess | 0,85 |

En spiller som ligger i cluster 2 og minder om Marko Divkovic er Stephen Odey. Odey er angriber i Randers FC.

1. <https://footy.dk/footy-fodbold/2022/03/17/broendbys-hotteste-transferemner-kan-indbringe-100-mio-eller-mere/> [↑](#footnote-ref-1)
2. <https://www.tipsbladet.dk/nyhed/superliga/transferplanen-bliver-stjernerne-skudt-afsted> [↑](#footnote-ref-2)