

cluster_summary

2024-12-07

```
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

coords_df <- read.csv("avg_team_data_kmeans.csv")
#Show how the countries were categorized
cluster_summary <- coords_df %>%
  group_by(cluster5) %>%
  summarize(team = paste(team, collapse = ", "))
print(cluster_summary)

## # A tibble: 5 x 2
##   cluster5 team
##   <int> <chr>
## 1      1 1 ARGENTINA, BRAZIL, ENGLAND, FRANCE, GERMANY, PORTUGAL
## 2      2 2 BELGIUM, CANADA, CROATIA, DENMARK, KOREA REPUBLIC, MEXICO, TUNISIA, ~
## 3      3 3 SPAIN
## 4      4 4 CAMEROON, ECUADOR, MOROCCO, NETHERLANDS, QATAR, SENEGAL, SWITZERLAND~
## 5      5 5 AUSTRALIA, COSTA RICA, GHANA, IRAN, JAPAN, POLAND, SAUDI ARABIA, SER~

write.csv(cluster_summary, "cluster_summary.csv")
```

Spain is the odd one out. Spain is known for having a distinct, possession-based play style.