

Geometric mean

```
library(dplyr)
library(ggplot2)
library(tidyr)
library(knitr)

load("data/HDA_dc_2.RData")
load("data/first_matches.RData")

HDA = HDA_dc %>%
  anti_join(first_matches)

nrow(HDA)

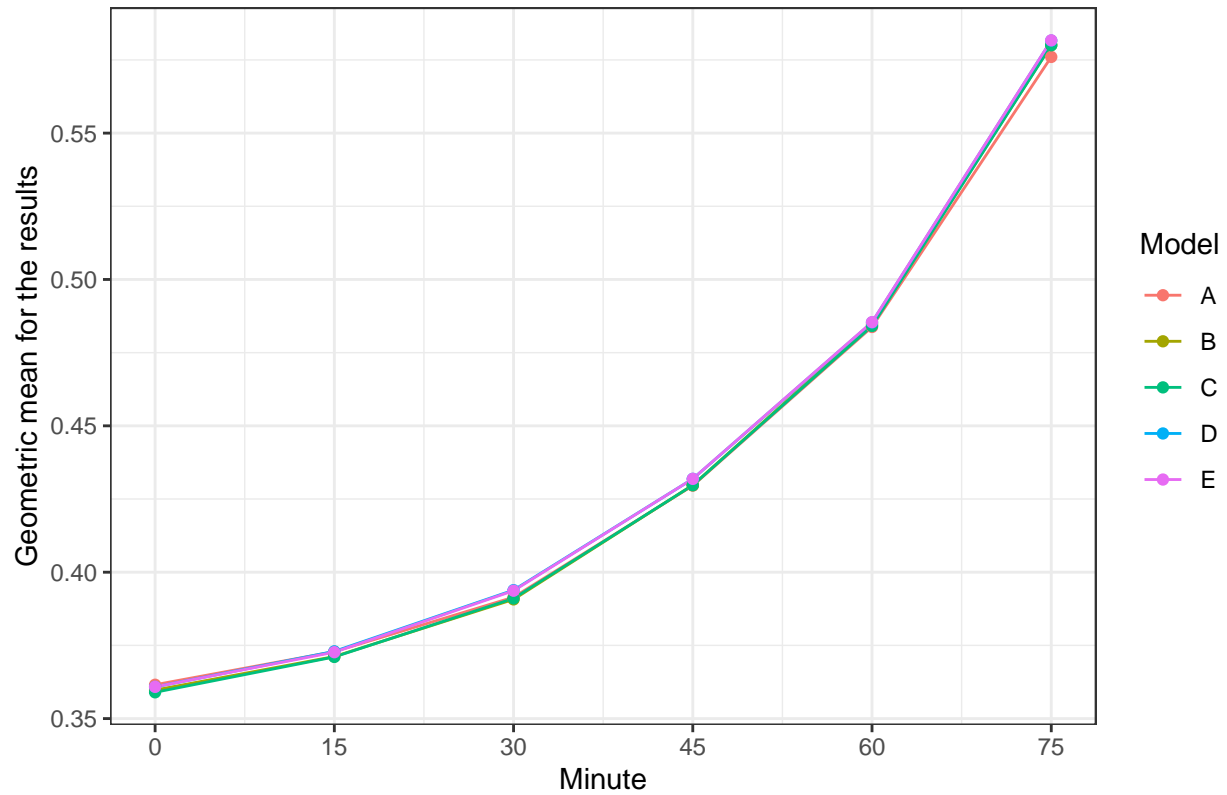
## [1] 1858

HDA[,c(9:158)][which(HDA[,c(9:158)] == 0, arr.ind = TRUE)] = 10^-5

results = tibble(GeoMean = apply(HDA[,c(99:128)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))

results %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches") +
  ylab("Geometric mean for the results")
```

All predicted matches



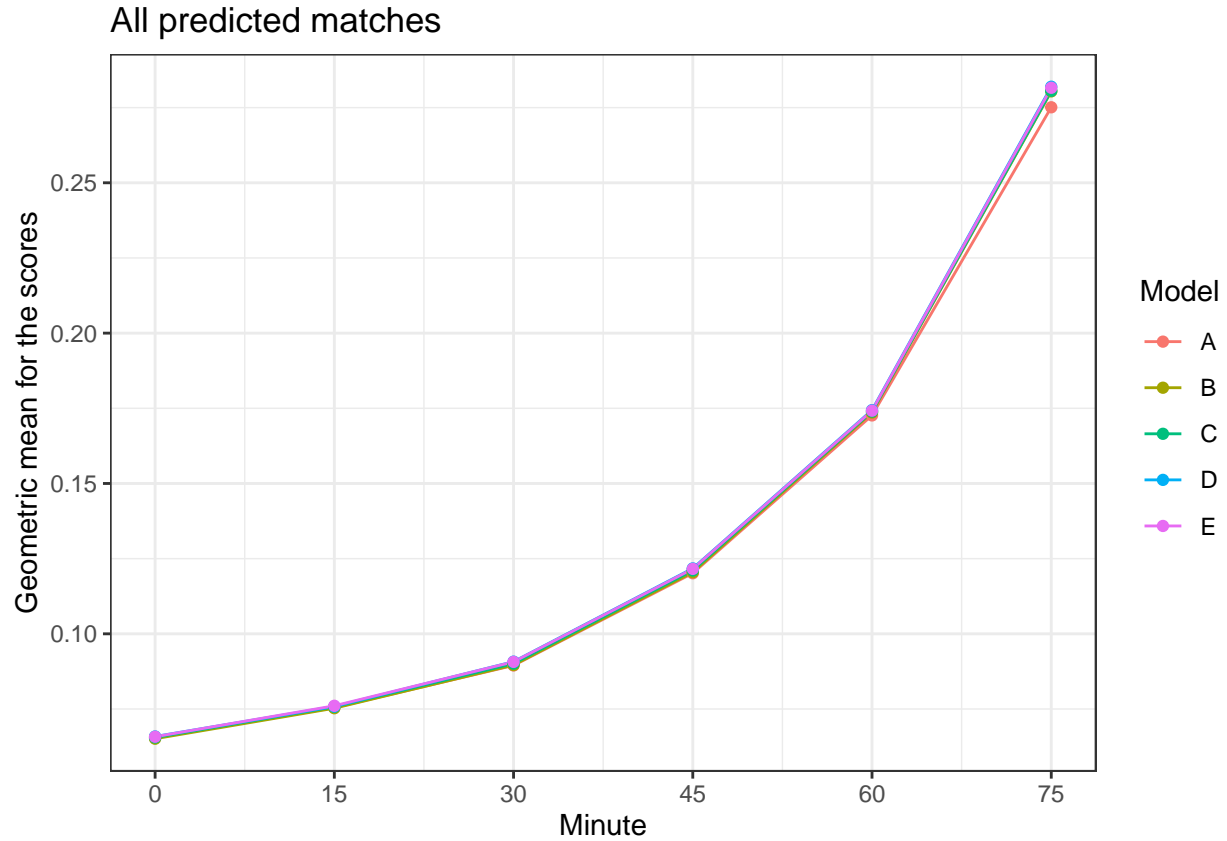
```
results %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.3615542	0.3728595	0.3914571	0.4295694	0.4837697	0.5760223
B	0.3597211	0.3711876	0.3907086	0.4297610	0.4845987	0.5800708
C	0.3589731	0.3710589	0.3910287	0.4297209	0.4841180	0.5800120
D	0.3608178	0.3729549	0.3938743	0.4318778	0.4853132	0.5815989
E	0.3608391	0.3726740	0.3936784	0.4318957	0.4854239	0.5816908

```
scores = tibble(GeoMean = apply(HDA[,c(129:158)], 2, EnvStats::geoMean),
                 Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
                 Model = factor(c(rep("A", 6),
                                   rep("B", 6),
                                   rep("C", 6),
                                   rep("D", 6),
                                   rep("E", 6))))

scores %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
```

```
scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
theme_bw() +
ggtitle("All predicted matches") +
ylab("Geometric mean for the scores")
```



```
scores %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0654583	0.0754529	0.0894830	0.1201492	0.1726362	0.2751038
B	0.0651204	0.0752651	0.0896398	0.1204935	0.1736041	0.2804053
C	0.0655202	0.0757454	0.0900929	0.1210661	0.1739258	0.2805999
D	0.0658623	0.0758537	0.0907471	0.1217557	0.1743915	0.2819345
E	0.0657864	0.0760711	0.0906762	0.1215974	0.1742131	0.2815970

```
load("~/GitHub/soccer-live-predictions/soccer-live-predictions/scrape/data/reds.RData")

matches = reds %>%
  filter(Season > 2015, Half == 1) %>%
  select(Season, Match)
```

```
HDA_recs = HDA %>%  
  inner_join(matches)
```

```
## Joining, by = c("Season", "Match")
```

```
HDA_no_recs = HDA %>%  
  anti_join(matches)
```

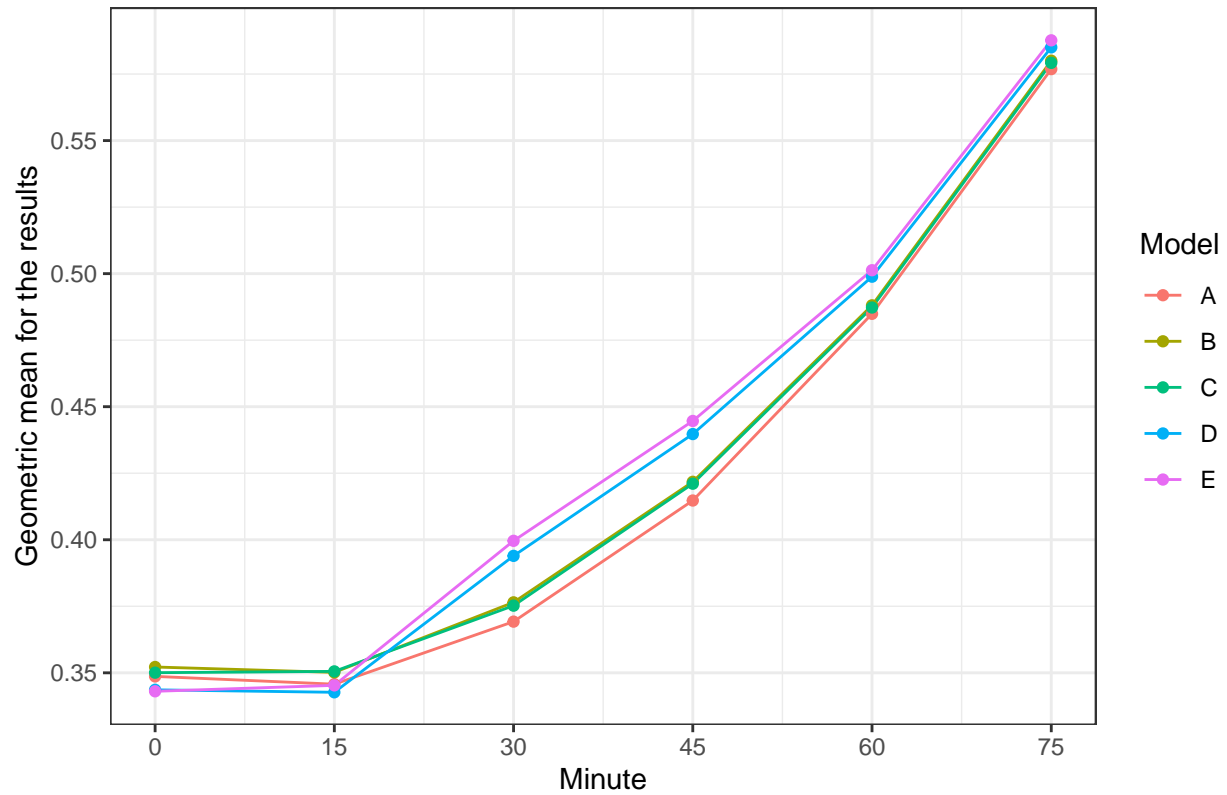
```
## Joining, by = c("Season", "Match")
```

```
nrow(HDA_recs)
```

```
## [1] 82
```

```
results_recs = tibble(GeoMean = apply(HDA_recs[,c(99:128)], 2, EnvStats::geoMean),  
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),  
  Model = factor(c(rep("A", 6),  
    rep("B", 6),  
    rep("C", 6),  
    rep("D", 6),  
    rep("E", 6))))  
results_recs %>%  
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +  
  geom_line() +  
  geom_point() +  
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +  
  theme_bw() +  
  ggtitle("All predicted matches with a red card in the first half") +  
  ylab("Geometric mean for the results")
```

All predicted matches with a red card in the first half



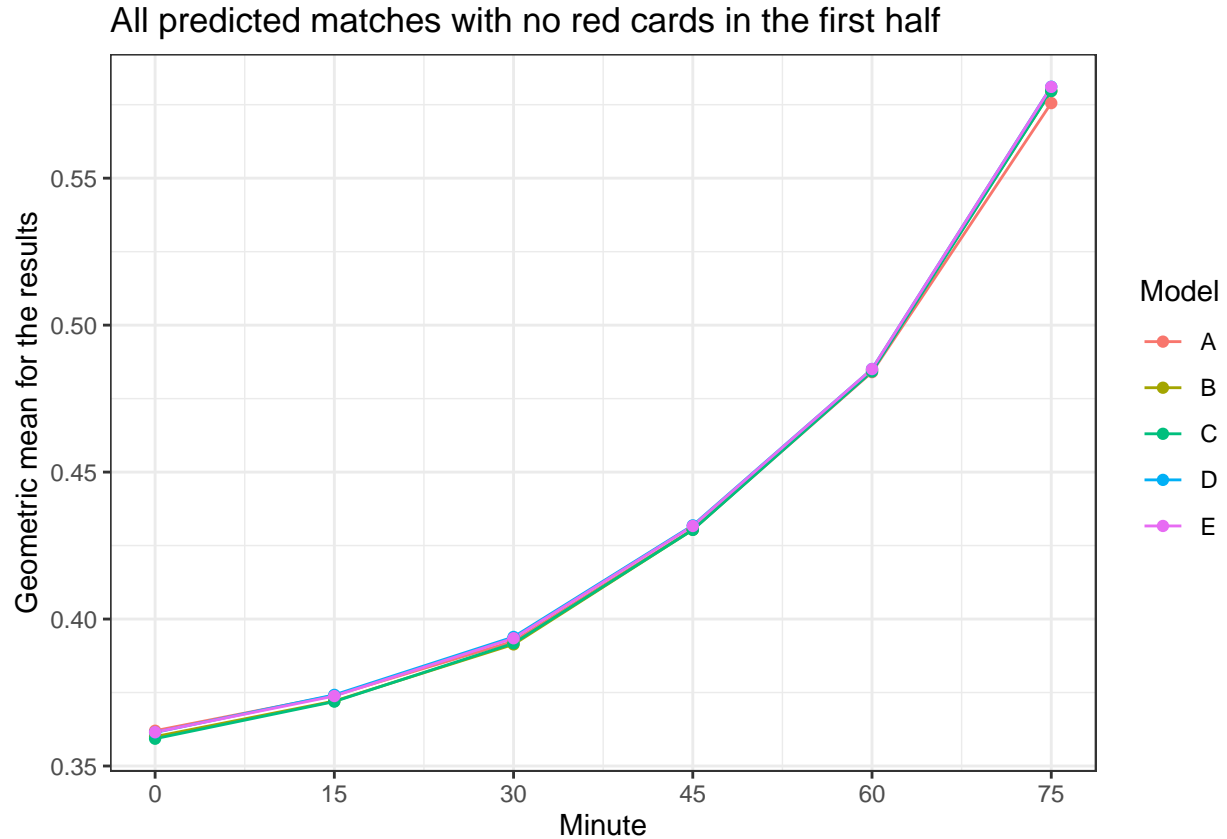
```
results_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.3486645	0.3457368	0.3692382	0.4147450	0.4848982	0.5768597
B	0.3521896	0.3501696	0.3764855	0.4217736	0.4880827	0.5800068
C	0.3500544	0.3505401	0.3752422	0.4210734	0.4872617	0.5792023
D	0.3436144	0.3426827	0.3939457	0.4397161	0.4988566	0.5850398
E	0.3430631	0.3452915	0.3996023	0.4446400	0.5012863	0.5876665

```
results_no_reds = tibble(GeoMean = apply(HDA_no_reds[,c(99:128)], 2, EnvStats::geoMean),
                          Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
                          Model = factor(c(rep("A", 6),
                                             rep("B", 6),
                                             rep("C", 6),
                                             rep("D", 6),
                                             rep("E", 6))))

results_no_reds %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
```

```
scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
theme_bw() +
ggtitle("All predicted matches with no red cards in the first half") +
ylab("Geometric mean for the results")
```



```
results_no_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

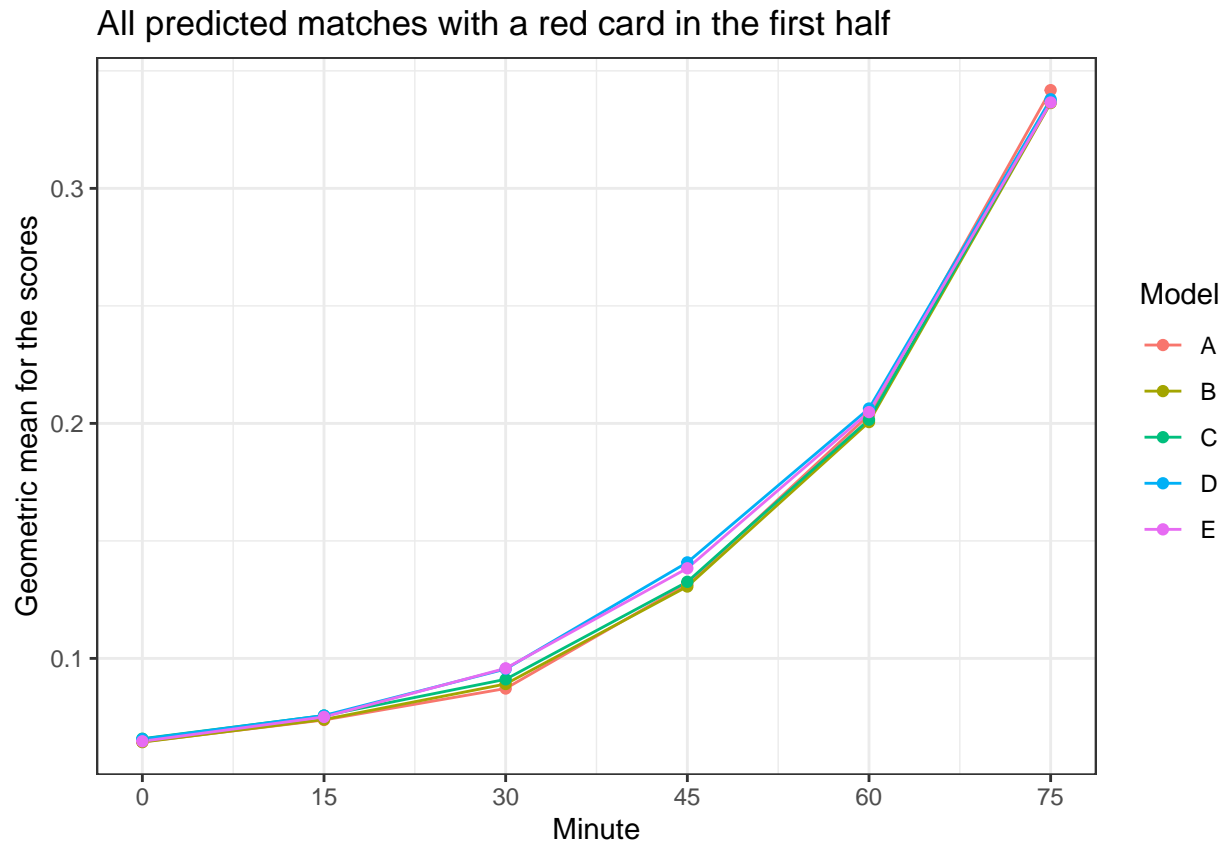
Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.3620003	0.3739905	0.3924370	0.4305123	0.4840060	0.5755491
B	0.3599875	0.3720903	0.3913739	0.4303819	0.4847162	0.5796775
C	0.3592840	0.3719146	0.3917359	0.4303439	0.4842272	0.5796571
D	0.3614173	0.3741974	0.3939049	0.4318678	0.4850609	0.5811073
E	0.3615060	0.3738049	0.3934632	0.4316615	0.4850620	0.5810679

```
scores_reds = tibble(GeoMean = apply(HDA_reds[,c(129:158)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
```

```

scores_reds %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with a red card in the first half") +
  ylab("Geometric mean for the scores")

```



```

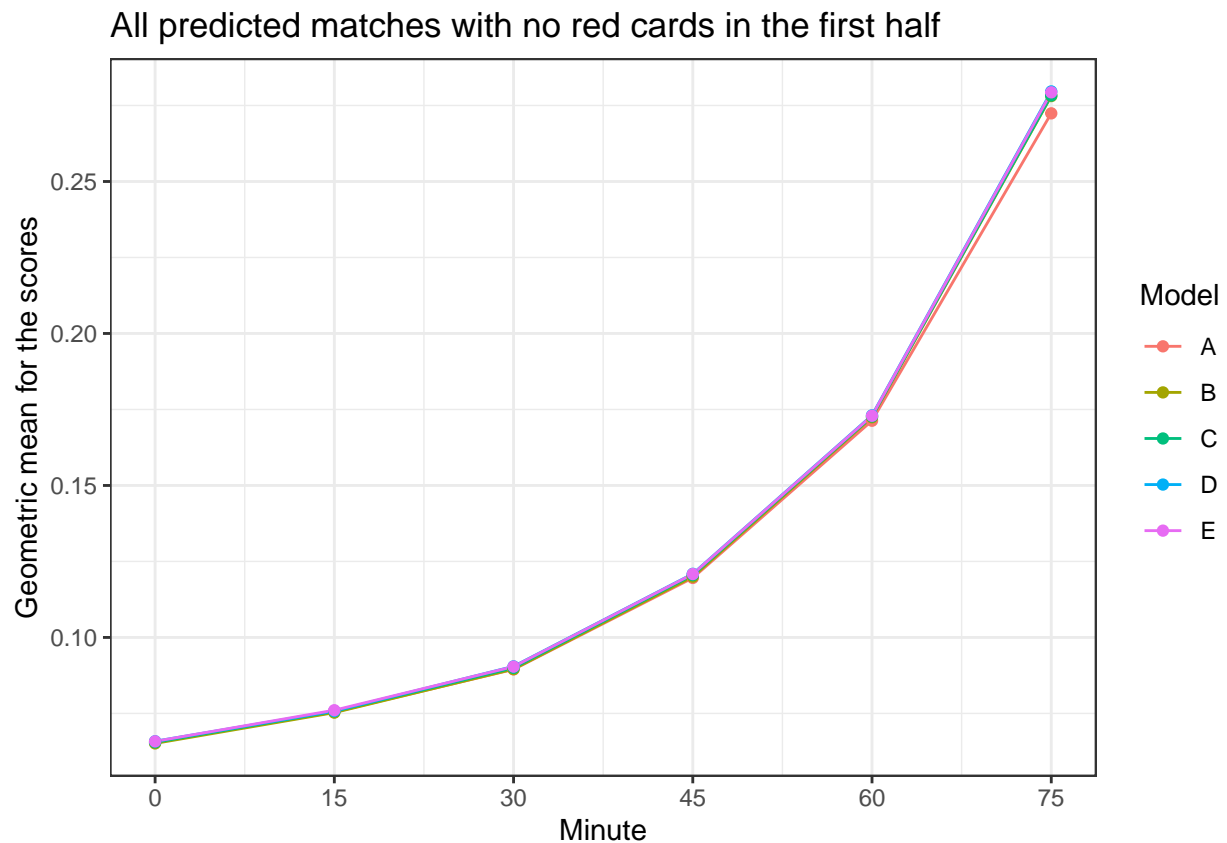
scores_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()

```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0645187	0.0738059	0.0871930	0.1316973	0.2036072	0.3417896
B	0.0643823	0.0740057	0.0891173	0.1305896	0.2005666	0.3363072
C	0.0658344	0.0756710	0.0911056	0.1325957	0.2015781	0.3371531
D	0.0656550	0.0756983	0.0954845	0.1408456	0.2063144	0.3379100
E	0.0647495	0.0750961	0.0957736	0.1383453	0.2048027	0.3365510

```
scores_no_reds = tibble(GeoMean = apply(HDA_no_reds[,c(129:158)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))

scores_no_reds %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with no red cards in the first half") +
  ylab("Geometric mean for the scores")
```



```
scores_no_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
    names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0654788	0.0754585	0.0894830	0.1195775	0.1712945	0.2724196
B	0.0651388	0.0752655	0.0895886	0.1199859	0.1724249	0.2781116

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
C	0.0654939	0.0756954	0.0899762	0.1205013	0.1727187	0.2782880
D	0.0658462	0.0757965	0.0904839	0.1209245	0.1730690	0.2796281
E	0.0658144	0.0760528	0.0903975	0.1208463	0.1729365	0.2793293

```
HDA_2020 = HDA %>%
  filter(Season == 2020)
```

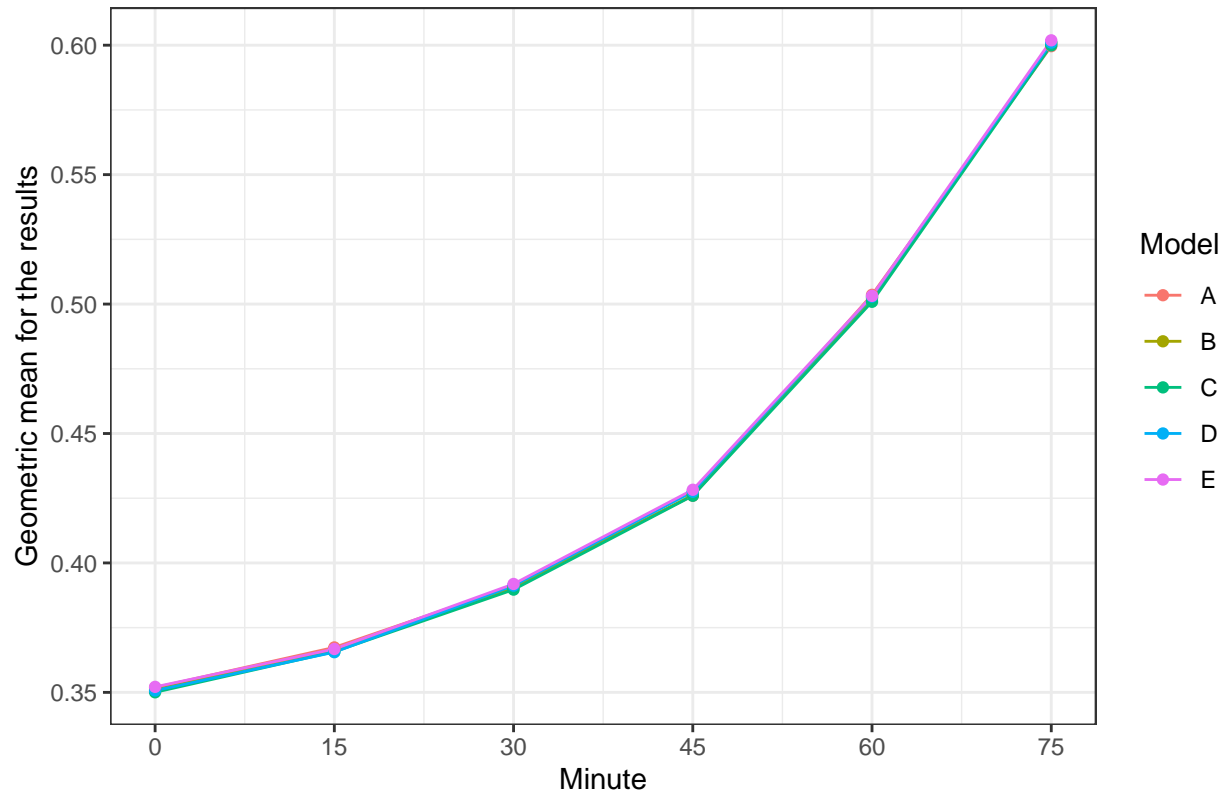
```
nrow(HDA_2020)
```

```
## [1] 376
```

```
results_2020 = tibble(GeoMean = apply(HDA_2020[,c(99:128)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))
```

```
results_2020 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches in the 2020 season") +
  ylab("Geometric mean for the results")
```

All predicted matches in the 2020 season



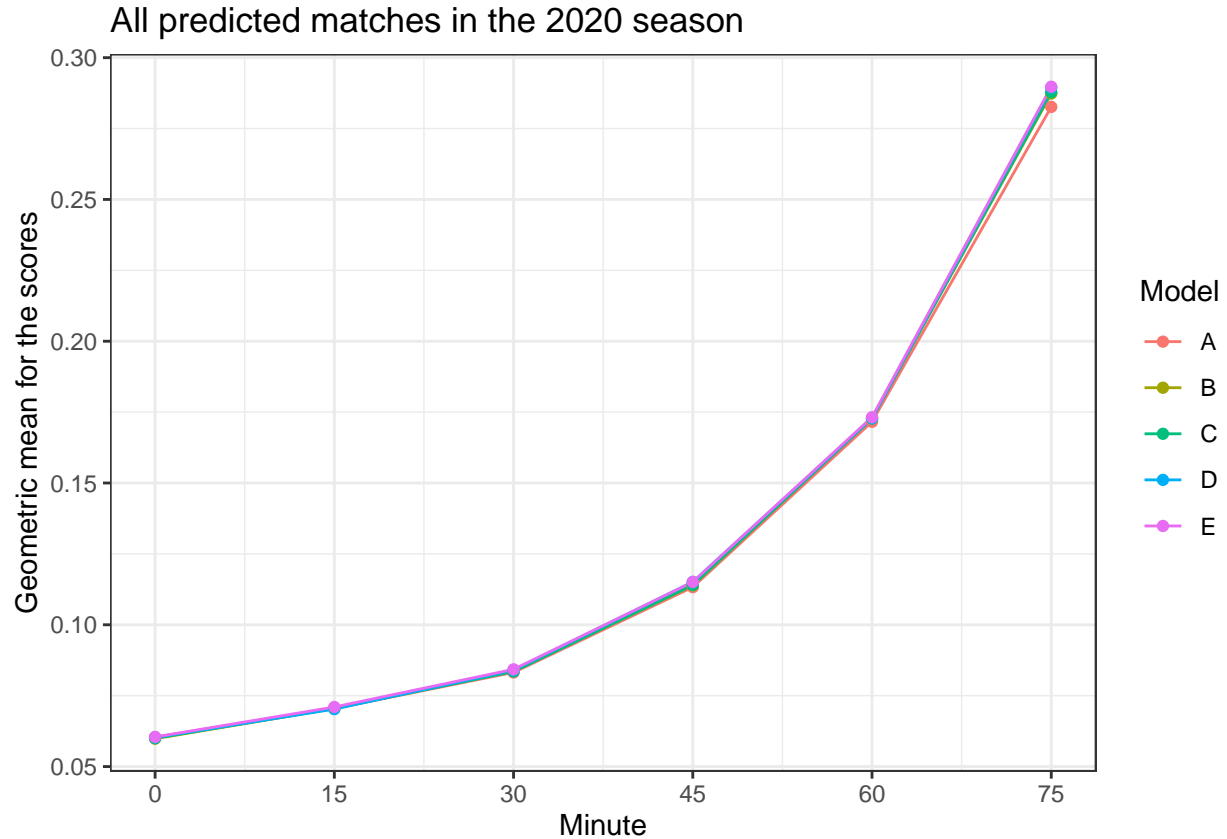
```
results_2020 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.3518546	0.3673683	0.3910833	0.4259717	0.5035654	0.5996244
B	0.3509758	0.3664922	0.3904459	0.4262226	0.5023343	0.6008139
C	0.3500279	0.3657104	0.3896927	0.4259532	0.5008836	0.5999531
D	0.3507179	0.3655351	0.3912212	0.4275058	0.5028590	0.6010051
E	0.3521268	0.3667433	0.3918680	0.4282678	0.5031646	0.6018786

```
scores_2020 = tibble(GeoMean = apply(HDA_2020[,c(129:158)], 2, EnvStats::geoMean),
                      Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
                      Model = factor(c(rep("A", 6),
                                       rep("B", 6),
                                       rep("C", 6),
                                       rep("D", 6),
                                       rep("E", 6))))

scores_2020 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
```

```
scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
theme_bw() +
ggtitle("All predicted matches in the 2020 season") +
ylab("Geometric mean for the scores")
```



```
scores_2020 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0603763	0.0704758	0.0831918	0.1132685	0.1715947	0.2826407
B	0.0598204	0.0704569	0.0837564	0.1139177	0.1728954	0.2873481
C	0.0600008	0.0705530	0.0835151	0.1140324	0.1727221	0.2878218
D	0.0602509	0.0702607	0.0840340	0.1151168	0.1728721	0.2894863
E	0.0604376	0.0710010	0.0842879	0.1150959	0.1732006	0.2897436

```
load("~/GitHub/soccer-live-predictions/soccer-live-predictions/scrape/data/results.RData")
load("~/GitHub/soccer-live-predictions/soccer-live-predictions/scrape/data/goals.RData")
```

```
at_45 = results %>%
  select(Season, Match) %>%
```

```
filter(Season > 2015)
```

```
home_score_at_45 <- function(season, match) {  
  goals %>%  
    filter(Season == season,  
           Match == match,  
           Team == 1,  
           Half == 1) %>%  
    nrow()  
}
```

```
away_score_at_45 <- function(season, match) {  
  goals %>%  
    filter(Season == season,  
           Match == match,  
           Team == 2,  
           Half == 1) %>%  
    nrow()  
}
```

```
at_45 = at_45 %>%  
  rowwise() %>%  
  mutate(Home_Score = home_score_at_45(Season, Match),  
         Away_Score = away_score_at_45(Season, Match),  
         abs_dif = abs(Home_Score - Away_Score))
```

```
tmp_00 = at_45 %>%  
  filter(abs_dif == 0) %>%  
  select(Season, Match)
```

```
HDA_00 = HDA %>%  
  inner_join(tmp_00)
```

```
## Joining, by = c("Season", "Match")
```

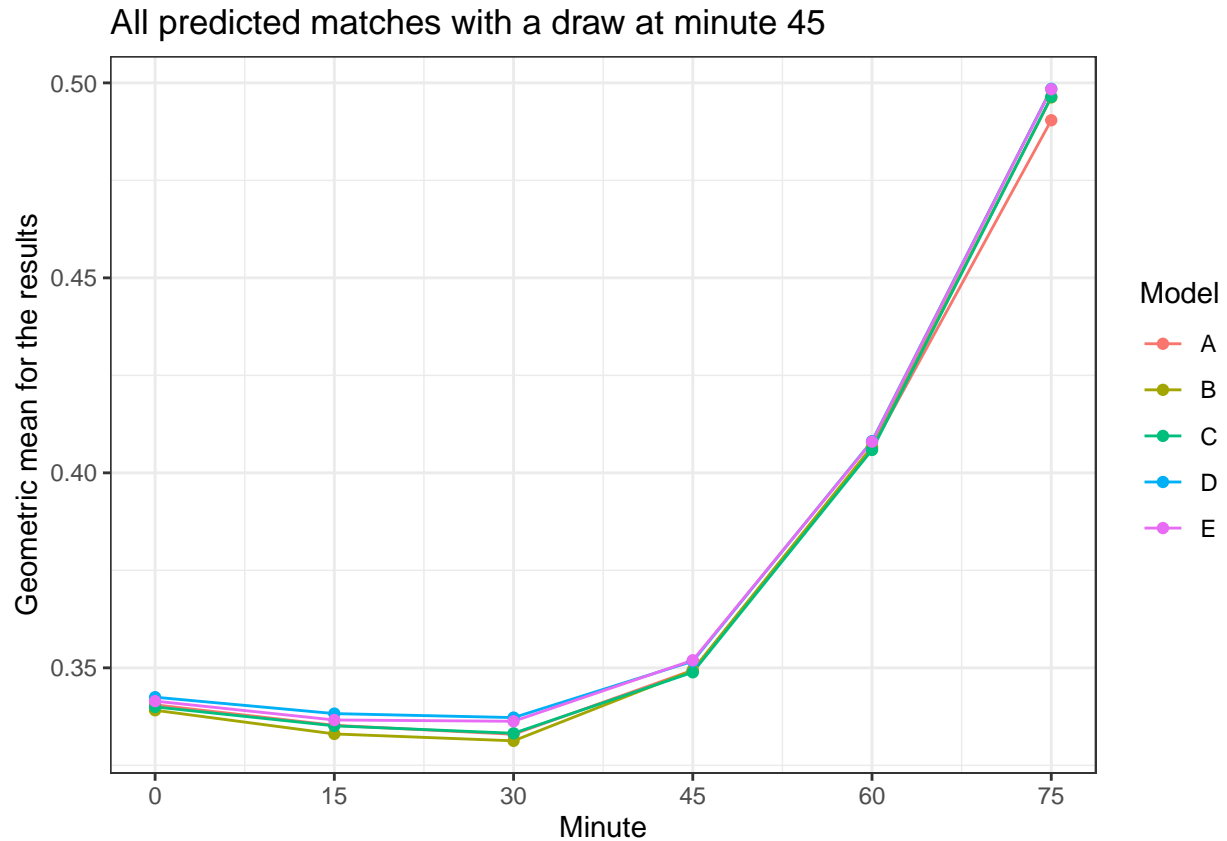
```
nrow(HDA_00)
```

```
## [1] 838
```

```
results_00 = tibble(GeoMean = apply(HDA_00[,c(99:128)], 2, EnvStats::geoMean),  
                    Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),  
                    Model = factor(c(rep("A", 6),  
                                     rep("B", 6),  
                                     rep("C", 6),  
                                     rep("D", 6),  
                                     rep("E", 6))))
```

```
results_00 %>%  
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +  
  geom_line() +  
  geom_point() +  
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
```

```
theme_bw() +
  ggtitle("All predicted matches with a draw at minute 45") +
  ylab("Geometric mean for the results")
```

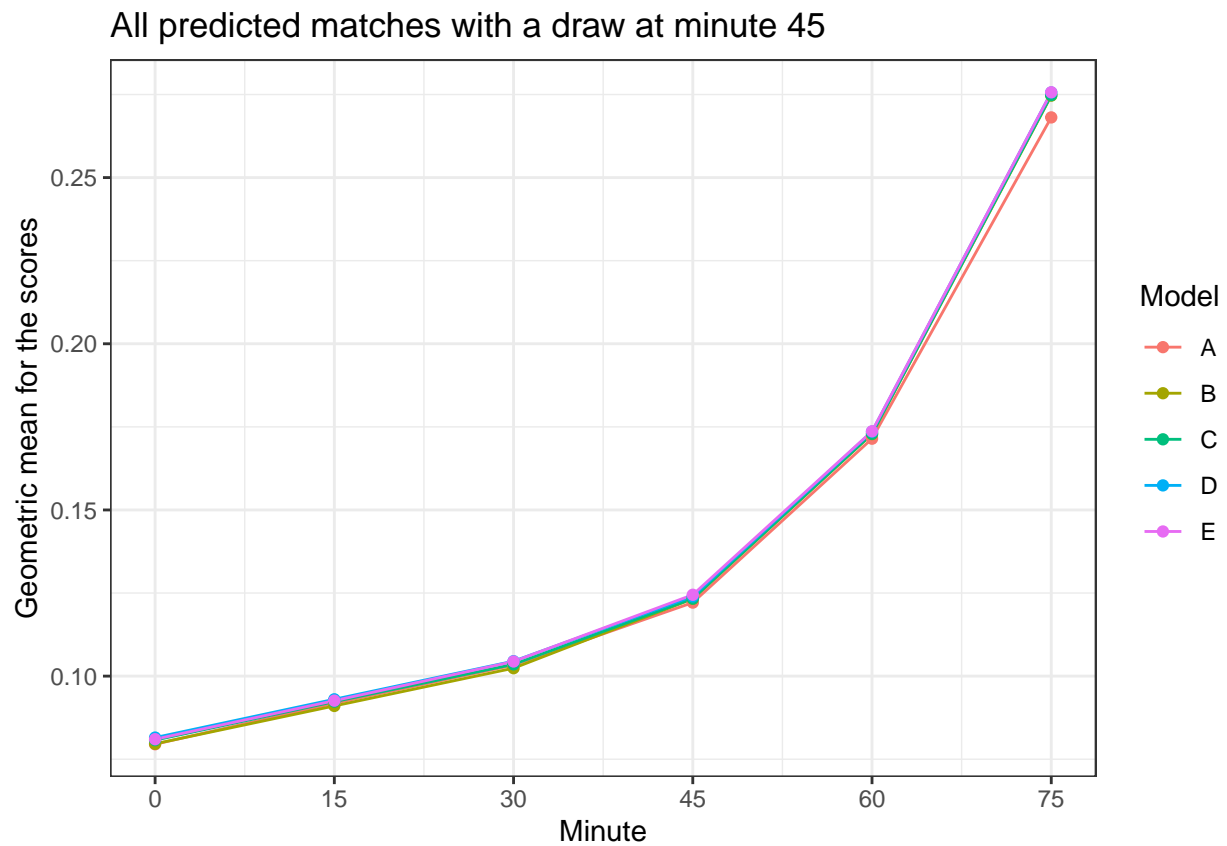


```
results_00 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.3405184	0.3352532	0.3329615	0.3493841	0.4063878	0.4904210
B	0.3391190	0.3330485	0.3312800	0.3493648	0.4067135	0.4962825
C	0.3399719	0.3351255	0.3332167	0.3488567	0.4058726	0.4964008
D	0.3424718	0.3382645	0.3372412	0.3516865	0.4081422	0.4984660
E	0.3414583	0.3366233	0.3362987	0.3519373	0.4079961	0.4983616

```
scores_00 = tibble(GeoMean = apply(HDA_00[,c(129:158)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))
```

```
scores_00 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with a draw at minute 45") +
  ylab("Geometric mean for the scores")
```



```
scores_00 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0795046	0.0918301	0.1033027	0.1221252	0.1714369	0.2681104
B	0.0796256	0.0910280	0.1023848	0.1232615	0.1728151	0.2746641
C	0.0807339	0.0923178	0.1035412	0.1233372	0.1729215	0.2748196
D	0.0815104	0.0930254	0.1044979	0.1239824	0.1735399	0.2756050
E	0.0810113	0.0926142	0.1044074	0.1244750	0.1736701	0.2756736

```
tmp_20 = at_45 %>%
  filter(abs_dif >= 2) %>%
  select(Season, Match)
```

```
HDA_20 = HDA %>%
  inner_join(tmp_20)
```

```
## Joining, by = c("Season", "Match")
```

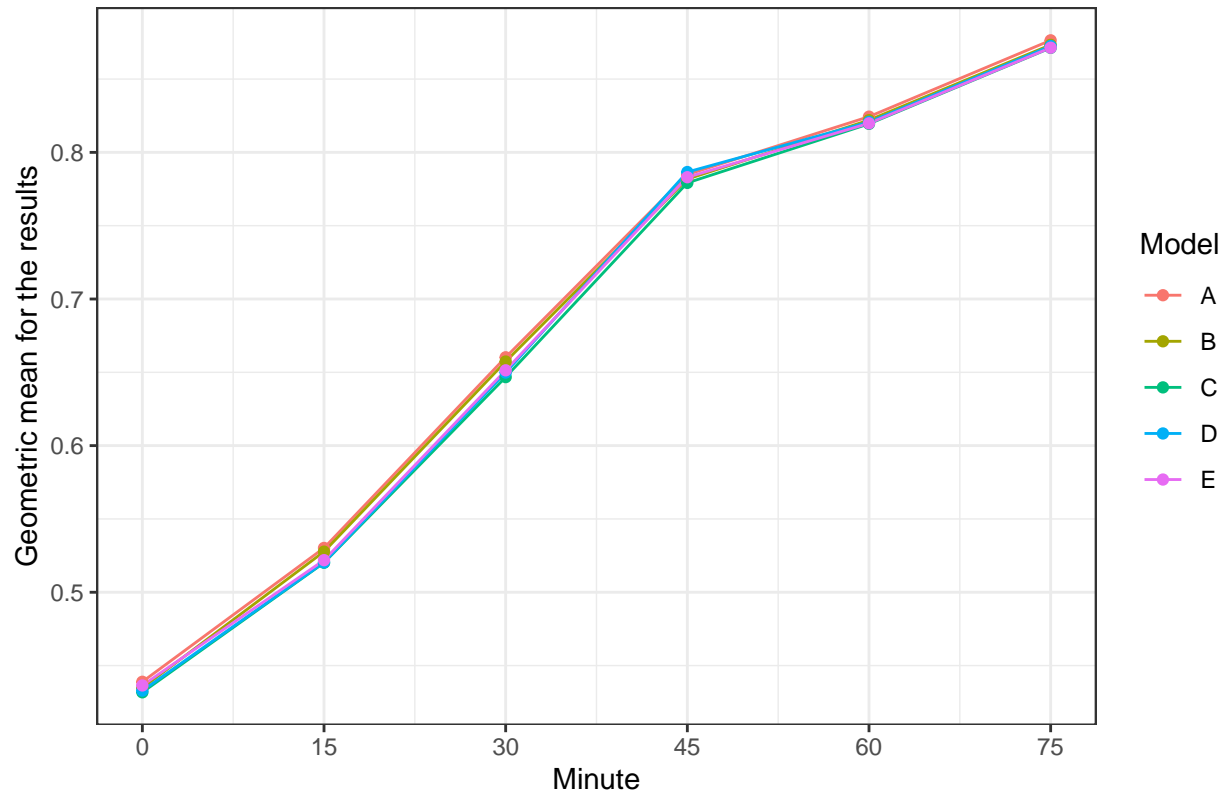
```
nrow(HDA_20)
```

```
## [1] 211
```

```
results_20 = tibble(GeoMean = apply(HDA_20[,c(99:128)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))
```

```
results_20 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with a 2+ goal lead at minute 45") +
  ylab("Geometric mean for the results")
```

All predicted matches with a 2+ goal lead at minute 45



```
results_20 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

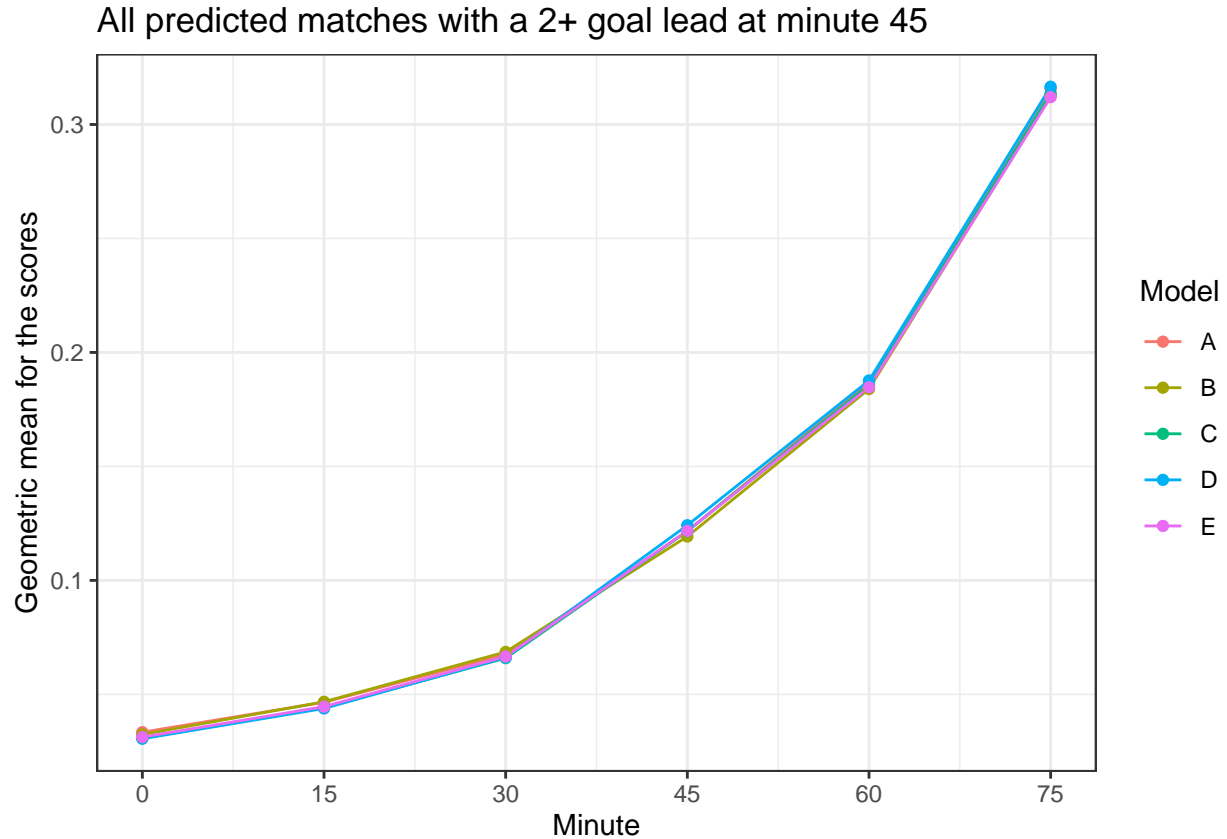
Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.4389412	0.5302146	0.6602535	0.7848906	0.8243508	0.8764061
B	0.4347884	0.5276975	0.6573326	0.7817797	0.8219608	0.8733194
C	0.4317755	0.5201440	0.6467315	0.7791872	0.8194427	0.8712679
D	0.4329183	0.5203485	0.6498265	0.7865919	0.8207394	0.8725380
E	0.4364570	0.5218895	0.6513633	0.7830524	0.8198218	0.8713727

```
scores_20 = tibble(GeoMean = apply(HDA_20[,c(129:158)], 2, EnvStats::geoMean),
                  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
                  Model = factor(c(rep("A", 6),
                                   rep("B", 6),
                                   rep("C", 6),
                                   rep("D", 6),
                                   rep("E", 6))))

scores_20 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
```



```
scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
theme_bw() +
ggtitle("All predicted matches with a 2+ goal lead at minute 45") +
ylab("Geometric mean for the scores")
```



```
scores_20 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0334097	0.0465290	0.0675163	0.1215685	0.1868990	0.3156954
B	0.0324130	0.0467390	0.0686181	0.1192675	0.1839593	0.3125120
C	0.0307227	0.0441779	0.0659004	0.1216406	0.1856787	0.3135589
D	0.0305472	0.0438189	0.0662227	0.1241918	0.1876758	0.3165560
E	0.0312706	0.0446136	0.0666052	0.1217181	0.1847287	0.3119925

```
tmp_10 = at_45 %>%
  filter(Home_Score == 1, Away_Score == 0) %>%
  select(Season, Match)

HDA_10 = HDA %>%
  inner_join(tmp_10)
```

```
## Joining, by = c("Season", "Match")
```

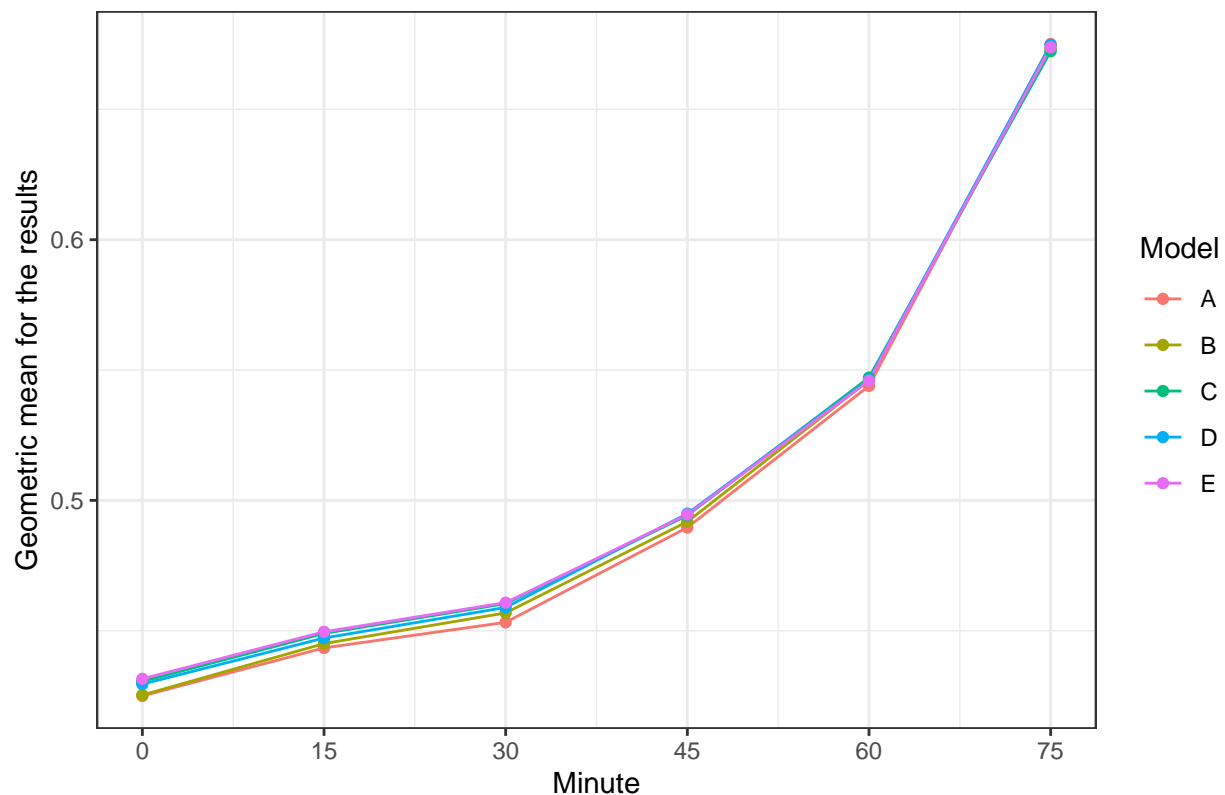
```
nrow(HDA_10)
```

```
## [1] 435
```

```
results_10 = tibble(GeoMean = apply(HDA_10[,c(99:128)], 2, EnvStats::geoMean),  
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),  
  Model = factor(c(rep("A", 6),  
    rep("B", 6),  
    rep("C", 6),  
    rep("D", 6),  
    rep("E", 6))))
```

```
results_10 %>%  
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +  
    geom_line() +  
    geom_point() +  
    scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +  
    theme_bw() +  
    ggtitle("All predicted matches with score 1-0 at minute 45") +  
    ylab("Geometric mean for the results")
```

All predicted matches with score 1-0 at minute 45



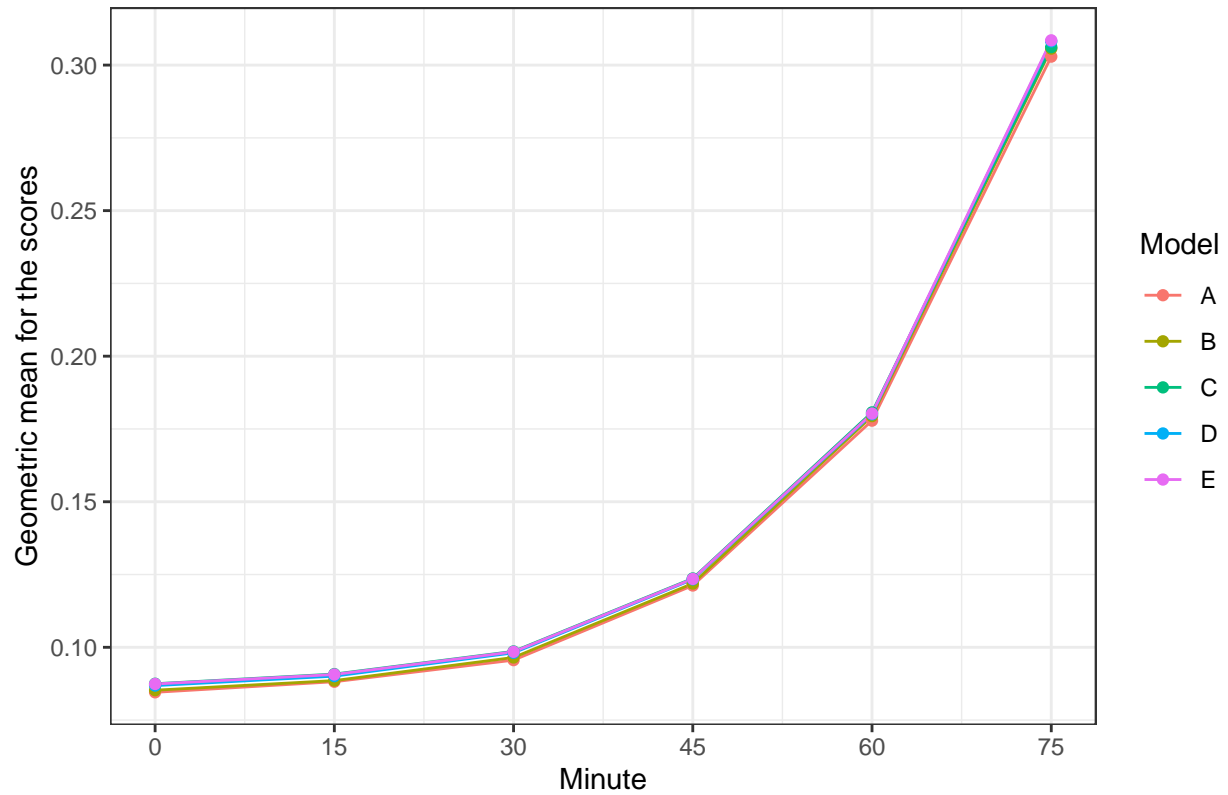
```
results_10 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.4249727	0.4434106	0.4532144	0.4895783	0.5439249	0.6750752
B	0.4253275	0.4450552	0.4568317	0.4918335	0.5461396	0.6732982
C	0.4305525	0.4489664	0.4603948	0.4942267	0.5470987	0.6722989
D	0.4295074	0.4471846	0.4589346	0.4949138	0.5463074	0.6747006
E	0.4315310	0.4495795	0.4607663	0.4945642	0.5458738	0.6737746

```
scores_10 = tibble(GeoMean = apply(HDA_10[,c(129:158)], 2, EnvStats::geoMean),
                    Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
                    Model = factor(c(rep("A", 6),
                                      rep("B", 6),
                                      rep("C", 6),
                                      rep("D", 6),
                                      rep("E", 6))))

scores_10 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with score 1-0 at minute 45") +
  ylab("Geometric mean for the scores")
```

All predicted matches with score 1–0 at minute 45



```
scores_10 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0845383	0.0881690	0.0955940	0.1212249	0.1778640	0.3029202
B	0.0852603	0.0885589	0.0964978	0.1219057	0.1794030	0.3056681
C	0.0874755	0.0907938	0.0985849	0.1236467	0.1806914	0.3061738
D	0.0868414	0.0900916	0.0981007	0.1233328	0.1800243	0.3081358
E	0.0873957	0.0906541	0.0984517	0.1234928	0.1802904	0.3085047

```
tmp_01 = at_45 %>%
  filter(Home_Score == 0, Away_Score == 1) %>%
  select(Season, Match)

HDA_01 = HDA %>%
  inner_join(tmp_01)
```

```
## Joining, by = c("Season", "Match")
```

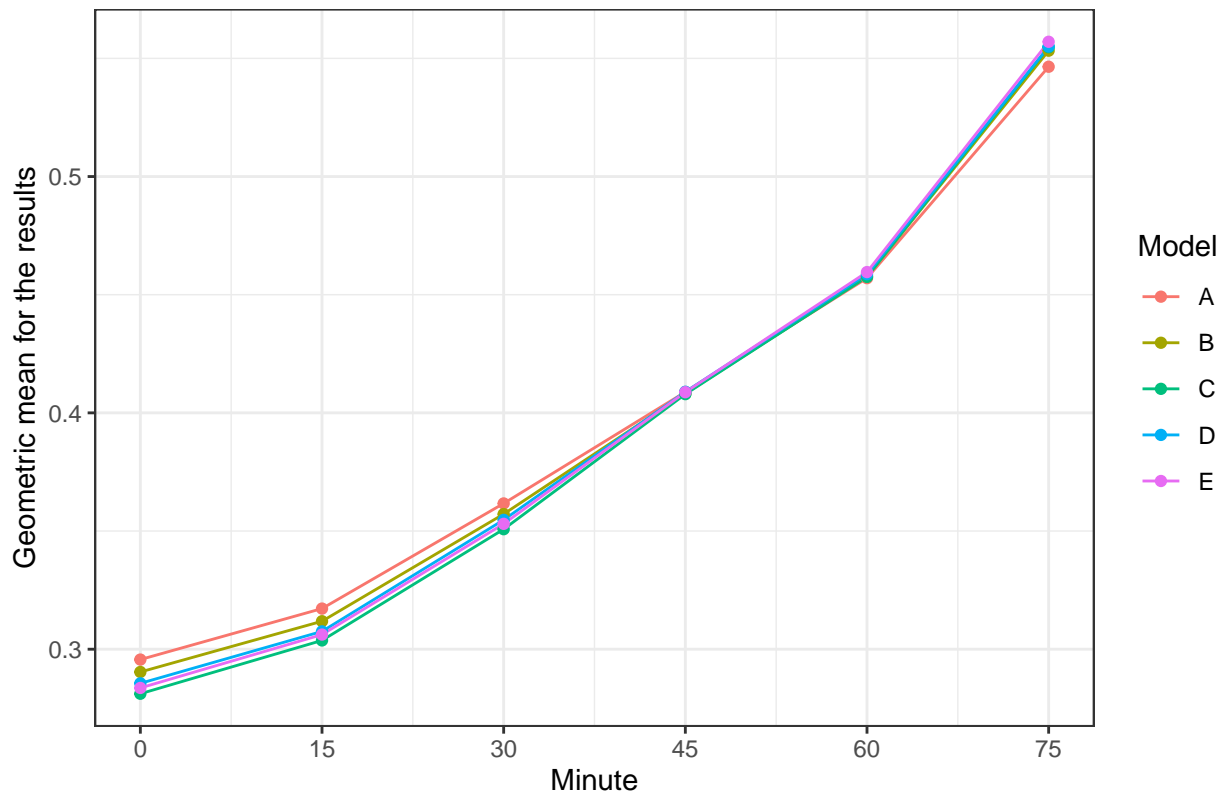
```
nrow(HDA_01)
```

```
## [1] 275
```

```
results_01 = tibble(GeoMean = apply(HDA_01[,c(99:128)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))

results_01 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with score 0-1 at minute 45") +
  ylab("Geometric mean for the results")
```

All predicted matches with score 0–1 at minute 45

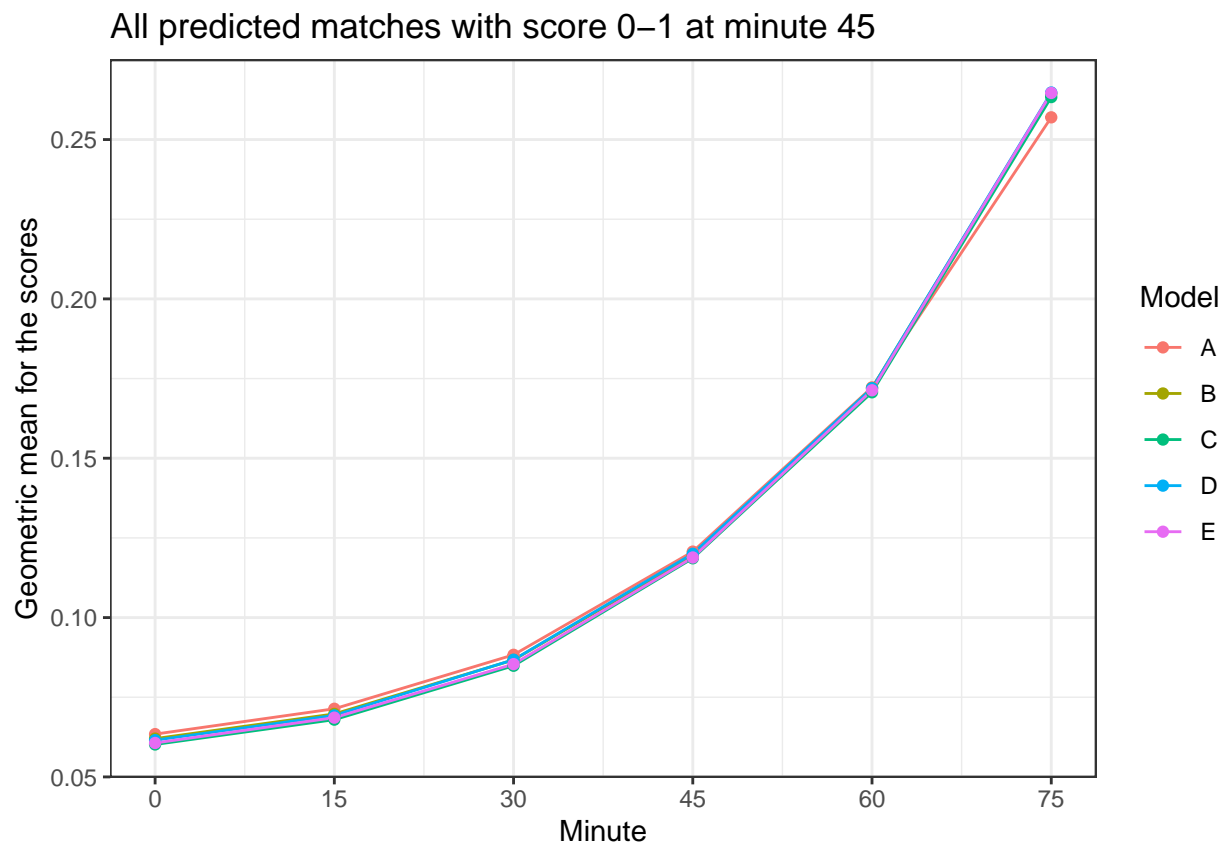


```
results_01 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
    names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.2956123	0.3171910	0.3616782	0.4088358	0.4570105	0.5464660
B	0.2903633	0.3118330	0.3571803	0.4081381	0.4576884	0.5532639
C	0.2811449	0.3037128	0.3507755	0.4079275	0.4575561	0.5547378
D	0.2855482	0.3075360	0.3547167	0.4089197	0.4586845	0.5550264
E	0.2836077	0.3061807	0.3530433	0.4087463	0.4595779	0.5570242

```
scores_01 = tibble(GeoMean = apply(HDA_01[,c(129:158)], 2, EnvStats::geoMean),
  Minute = as.integer(rep(c(0, 15, 30, 45, 60, 75), 5)),
  Model = factor(c(rep("A", 6),
    rep("B", 6),
    rep("C", 6),
    rep("D", 6),
    rep("E", 6))))

scores_01 %>%
  ggplot(aes(x = Minute, y = GeoMean, col = Model)) +
  geom_line() +
  geom_point() +
  scale_x_continuous(breaks = c(0, 15, 30, 45, 60, 75)) +
  theme_bw() +
  ggtitle("All predicted matches with score 0-1 at minute 45") +
  ylab("Geometric mean for the scores")
```



```
scores_01 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0.0634328	0.0713822	0.0883127	0.1206942	0.1722355	0.2569705
B	0.0620208	0.0697676	0.0867108	0.1194028	0.1719953	0.2643466
C	0.0601808	0.0679700	0.0848852	0.1186034	0.1706895	0.2633290
D	0.0615342	0.0693115	0.0867757	0.1200215	0.1719729	0.2647589
E	0.0606467	0.0685952	0.0853952	0.1188613	0.1713194	0.2646236