Geometric mean

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

options(OutDec = ",")

load("data/HDA_dc_2_v2.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
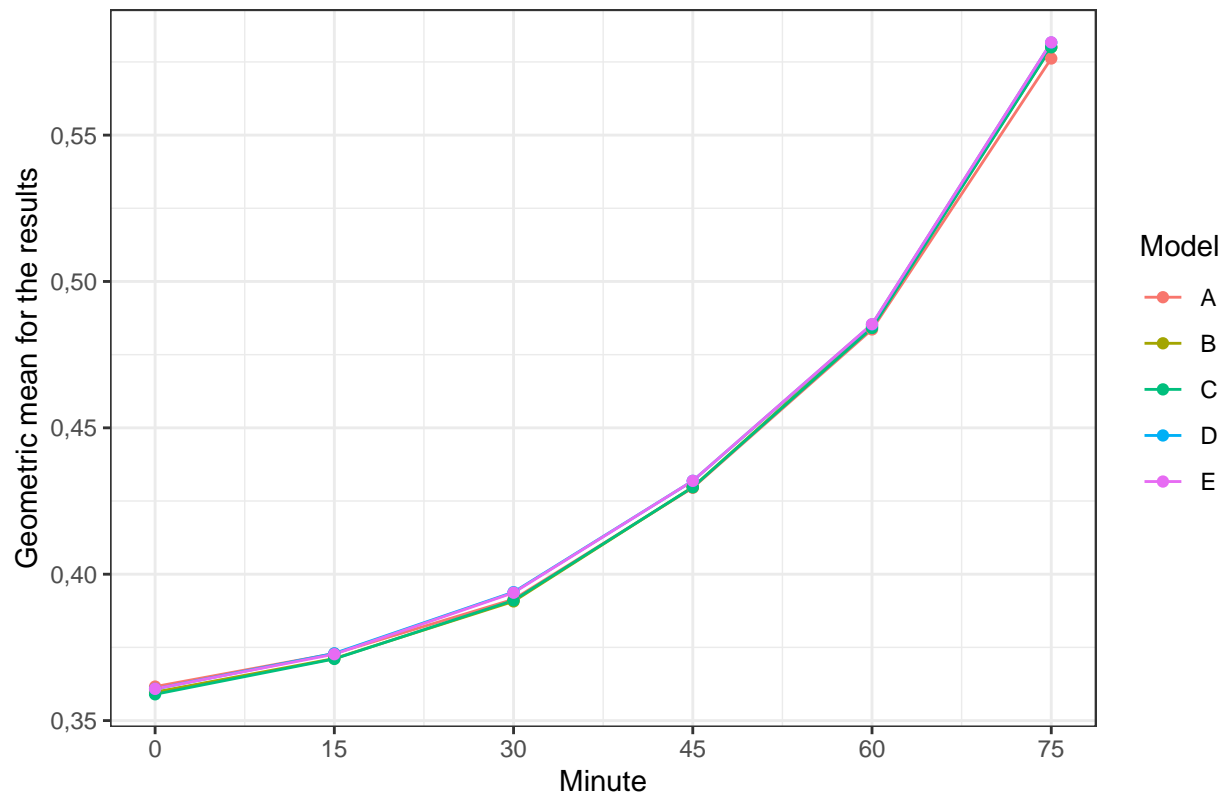

```
df = results %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ")
kable(df)
```

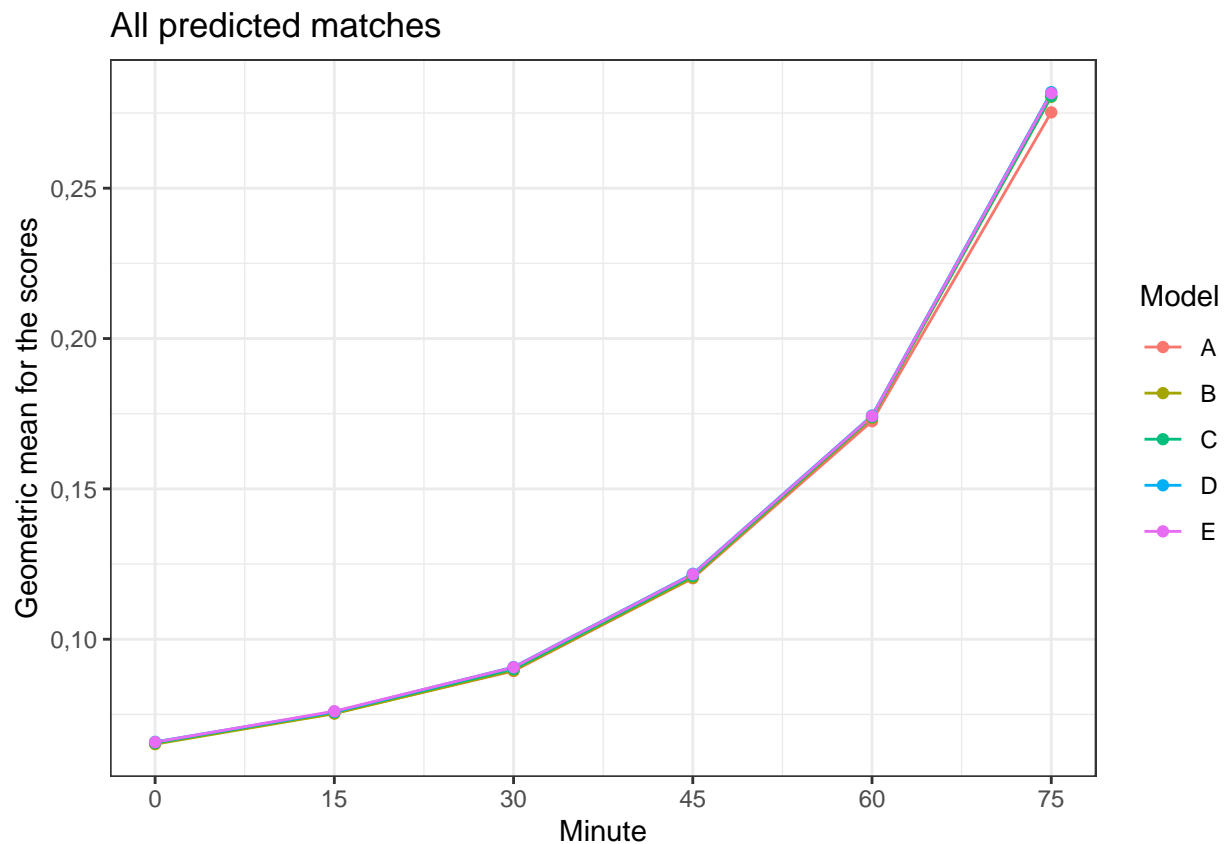| Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,3616020	0,3728834	0,3914252	0,4295289	0,4836092	0,5761762
B	0,3597211	0,3711876	0,3907086	0,4297610	0,4845987	0,5800708
C	0,3589731	0,3710589	0,3910287	0,4297209	0,4841180	0,5800120
D	0,3608178	0,3729549	0,3938743	0,4318778	0,4853132	0,5815989
E	0,3608391	0,3726740	0,3936784	0,4318957	0,4854239	0,5816908

```
xtable(df, digits = 4)
```

```
## % latex table generated in R 4.0.4 by xtable 1.8-4 package
## % Fri Aug 27 03:20:00 2021
## \begin{table}[ht]
## \centering
## \begin{tabular}{rlrrrrrr}
##   \hline
##  & Model & Minute 0 & Minute 15 & Minute 30 & Minute 45 & Minute 60 & Minute 75 \\
##   \hline
## 1 & A & 0,3616 & 0,3729 & 0,3914 & 0,4295 & 0,4836 & 0,5762 \\
```

```
##    2 & B & 0,3597 & 0,3712 & 0,3907 & 0,4298 & 0,4846 & 0,5801 \\
##    3 & C & 0,3590 & 0,3711 & 0,3910 & 0,4297 & 0,4841 & 0,5800 \\
##    4 & D & 0,3608 & 0,3730 & 0,3939 & 0,4319 & 0,4853 & 0,5816 \\
##    5 & E & 0,3608 & 0,3727 & 0,3937 & 0,4319 & 0,4854 & 0,5817 \\
##    \hline
## \end{tabular}
## \end{table}
```

```r
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")
```

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

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,0654750	0,0754421	0,0894467	0,1203027	0,1725042	0,2752126
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

```r
xtable(df, digits = 4)
```

```
## % latex table generated in R 4.0.4 by xtable 1.8-4 package
## % Fri Aug 27 03:20:00 2021
## \begin{table}[ht]
## \centering
## \begin{tabular}{rlrrrrrr}
##   \hline
##  & Model & Minute 0 & Minute 15 & Minute 30 & Minute 45 & Minute 60 & Minute 75 \\
##   \hline
## 1 & A & 0,0655 & 0,0754 & 0,0894 & 0,1203 & 0,1725 & 0,2752 \\
##   2 & B & 0,0651 & 0,0753 & 0,0896 & 0,1205 & 0,1736 & 0,2804 \\
##   3 & C & 0,0655 & 0,0757 & 0,0901 & 0,1211 & 0,1739 & 0,2806 \\
##   4 & D & 0,0659 & 0,0759 & 0,0907 & 0,1218 & 0,1744 & 0,2819 \\
##   5 & E & 0,0658 & 0,0761 & 0,0907 & 0,1216 & 0,1742 & 0,2816 \\
##    \hline
## \end{tabular}
## \end{table}
```

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

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

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

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

```r
HDA_no_reds = HDA %>%
  anti_join(matches)
```

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

```
nrow(HDA_reds)
```

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

```
results_reds = tibble(GeoMean = apply(HDA_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_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 results")
```

All predicted matches with a red card in the first half
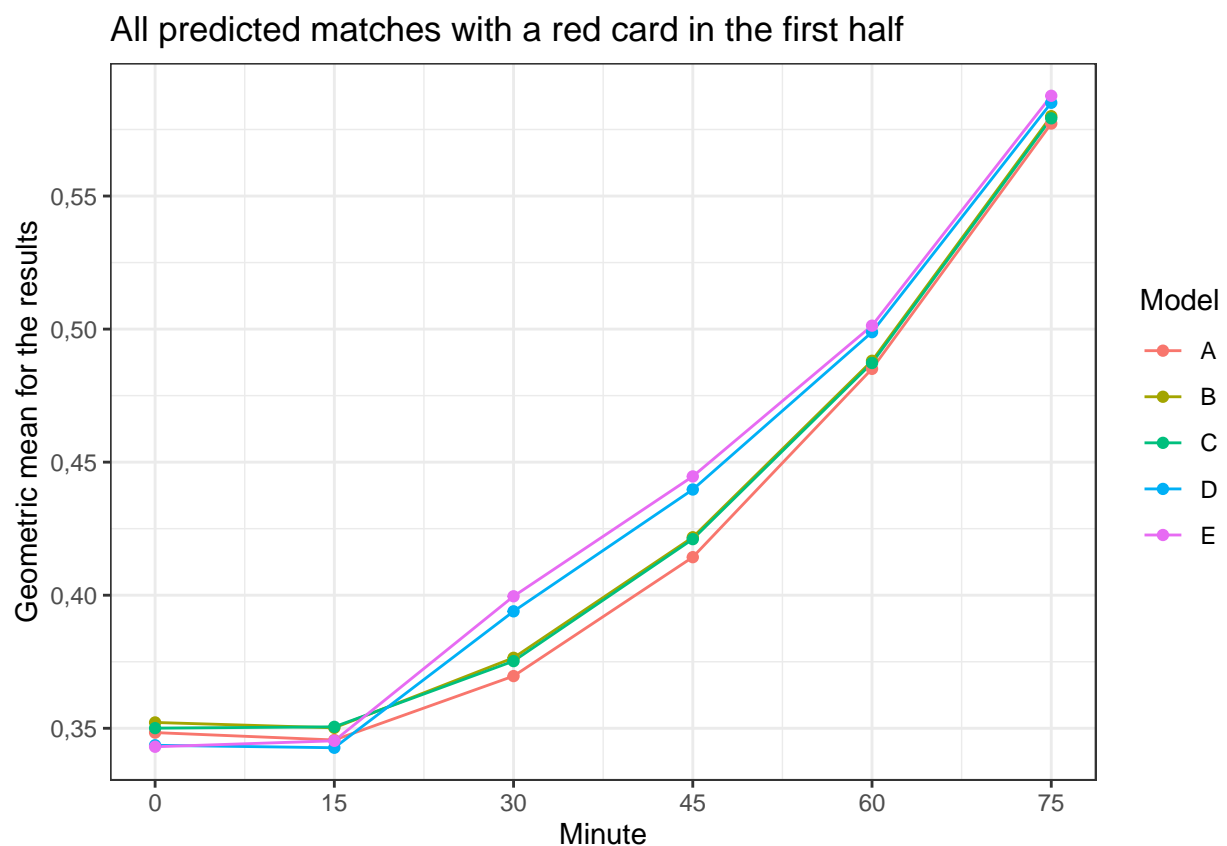

```
df = results_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ")
kable(df)
```

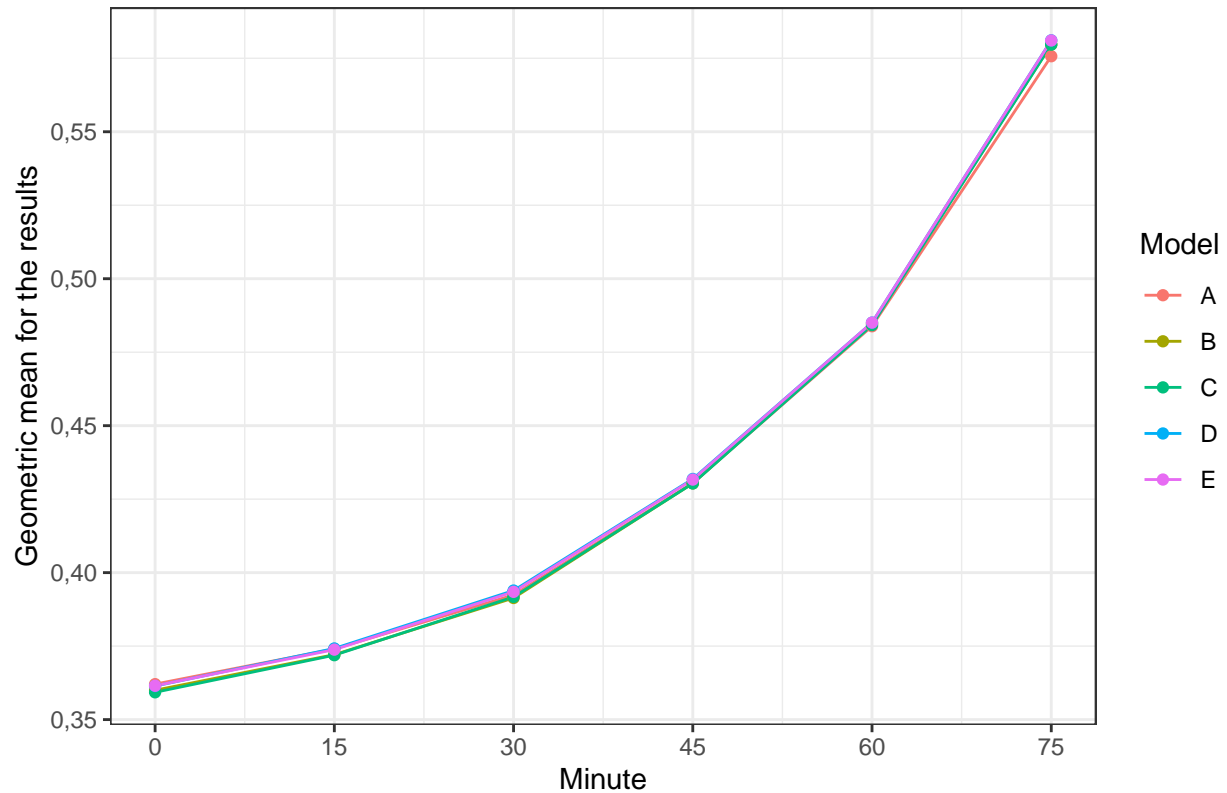| Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,3483818	0,3456032	0,3696152	0,4142996	0,4850619	0,5772395
B	0,3521896	0,3501696	0,3764855	0,4217736	0,4880827	0,5800068
C	0,3500544	0,3505401	0,3752422	0,4210734	0,4872617	0,5792023
D	0,3436144	0,3426827	0,3939457	0,4397161	0,4988566	0,5850398
E	0,3430631	0,3452915	0,3996023	0,4446400	0,5012863	0,5876665

```r
xtable(df, digits = 4)
```

```
## % latex table generated in R 4.0.4 by xtable 1.8-4 package
## % Fri Aug 27 03:20:01 2021
## \begin{table}[ht]
## \centering
## \begin{tabular}{rlrrrrrr}
##   \hline
##  & Model & Minute 0 & Minute 15 & Minute 30 & Minute 45 & Minute 60 & Minute 75 \\
##   \hline
## 1 & A & 0,3484 & 0,3456 & 0,3696 & 0,4143 & 0,4851 & 0,5772 \\
##   2 & B & 0,3522 & 0,3502 & 0,3765 & 0,4218 & 0,4881 & 0,5800 \\
##   3 & C & 0,3501 & 0,3505 & 0,3752 & 0,4211 & 0,4873 & 0,5792 \\
##   4 & D & 0,3436 & 0,3427 & 0,3939 & 0,4397 & 0,4989 & 0,5850 \\
##   5 & E & 0,3431 & 0,3453 & 0,3996 & 0,4446 & 0,5013 & 0,5877 \\
##    \hline
## \end{tabular}
## \end{table}
```

```r
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")
```

All predicted matches with no red cards in the first half
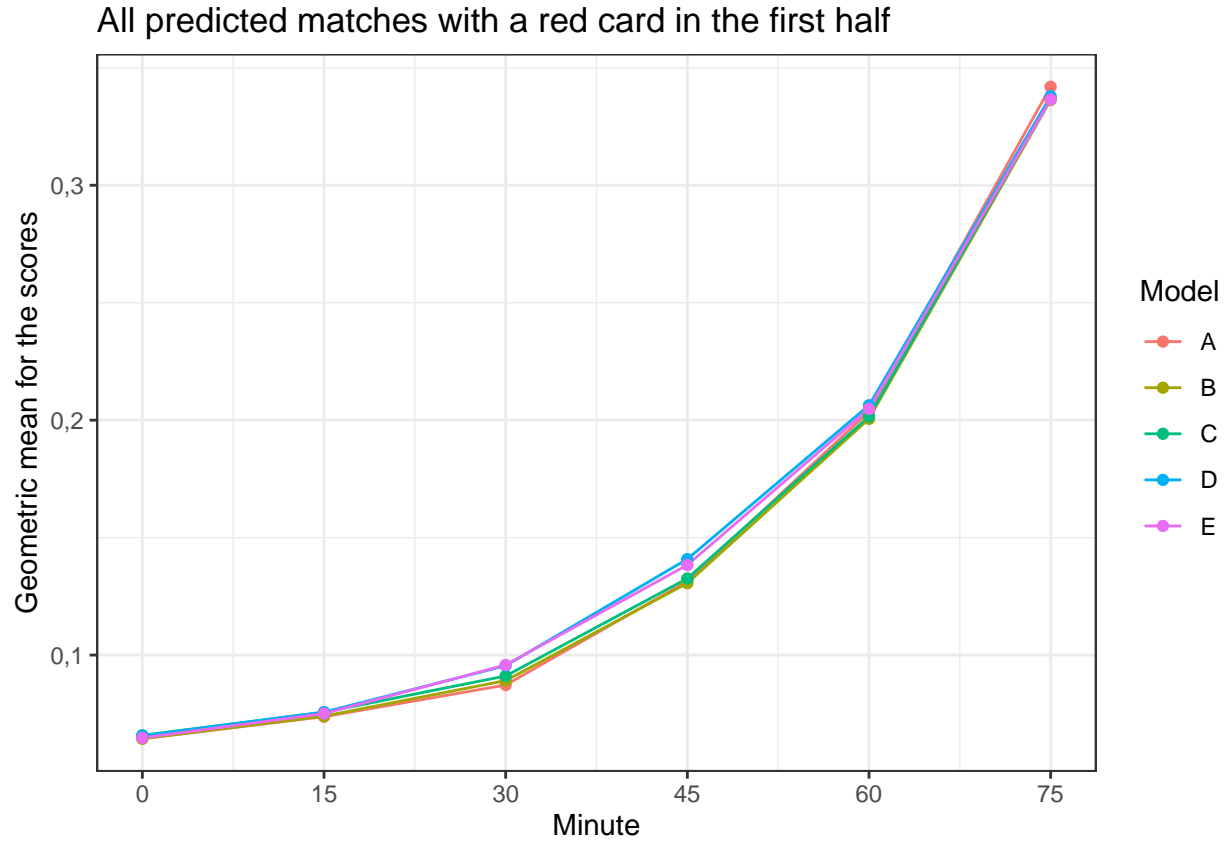


```
results_no_reds %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,3620605	0,3740245	0,3923863	0,4304904	0,4838324	0,5757014
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),
                                      rep("E", 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")
```

All predicted matches with a red card in the first half


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

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,0644110	0,0737552	0,0872426	0,1314436	0,2036446	0,3419332
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

```
xtable(df, digits = 4)
```

```
## % latex table generated in R 4.0.4 by xtable 1.8-4 package
## % Fri Aug 27 03:20:01 2021
## \begin{table}[ht]
```
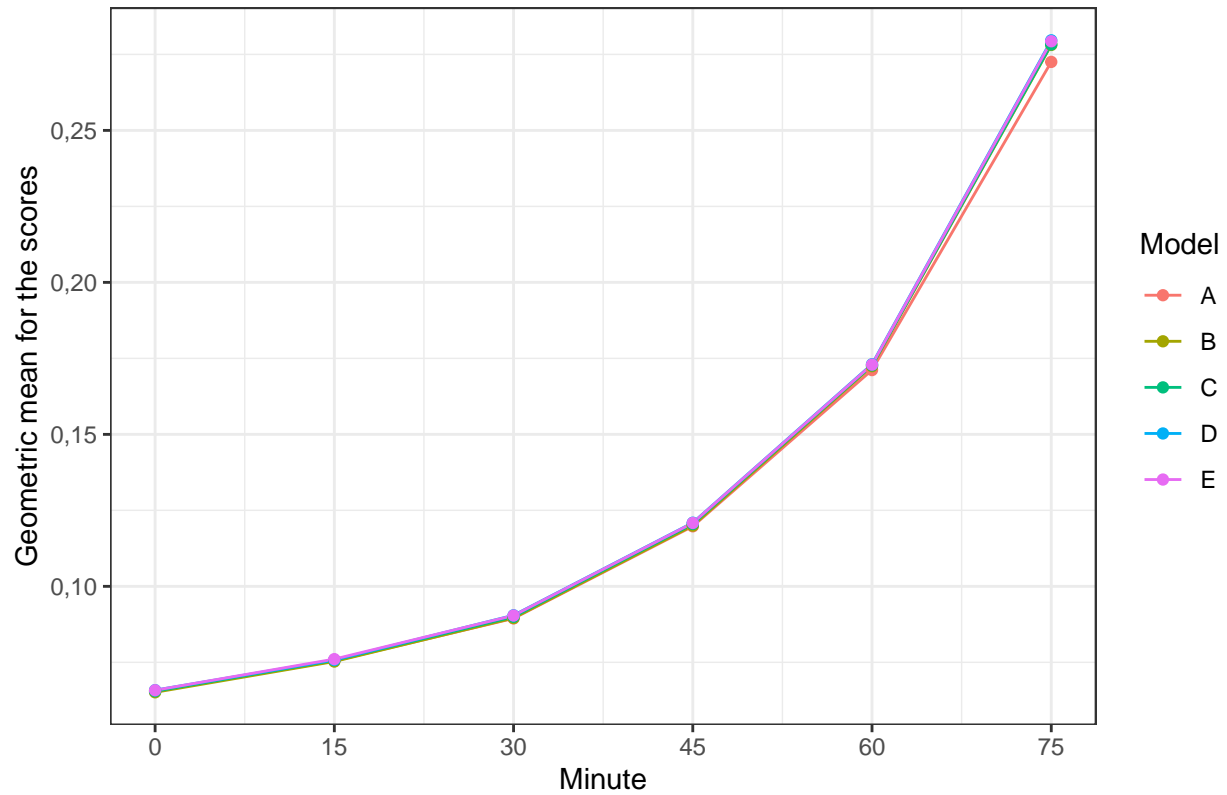
```
## \centering
## \begin{tabular}{rlrrrrrr}
##   \hline
##  & Model & Minute 0 & Minute 15 & Minute 30 & Minute 45 & Minute 60 & Minute 75 \\
##   \hline
## 1 & A & 0,0644 & 0,0738 & 0,0872 & 0,1314 & 0,2036 & 0,3419 \\
##   2 & B & 0,0644 & 0,0740 & 0,0891 & 0,1306 & 0,2006 & 0,3363 \\
##   3 & C & 0,0658 & 0,0757 & 0,0911 & 0,1326 & 0,2016 & 0,3372 \\
##   4 & D & 0,0657 & 0,0757 & 0,0955 & 0,1408 & 0,2063 & 0,3379 \\
##   5 & E & 0,0647 & 0,0751 & 0,0958 & 0,1383 & 0,2048 & 0,3366 \\
##    \hline
## \end{tabular}
## \end{table}
```

```r
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")
```

All predicted matches with no red cards in the first half


```
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	0,0655008	0,0754502	0,0894448	0,1197479	0,1711560	0,2725312
B	0,0651388	0,0752655	0,0895886	0,1199859	0,1724249	0,2781116
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")
```
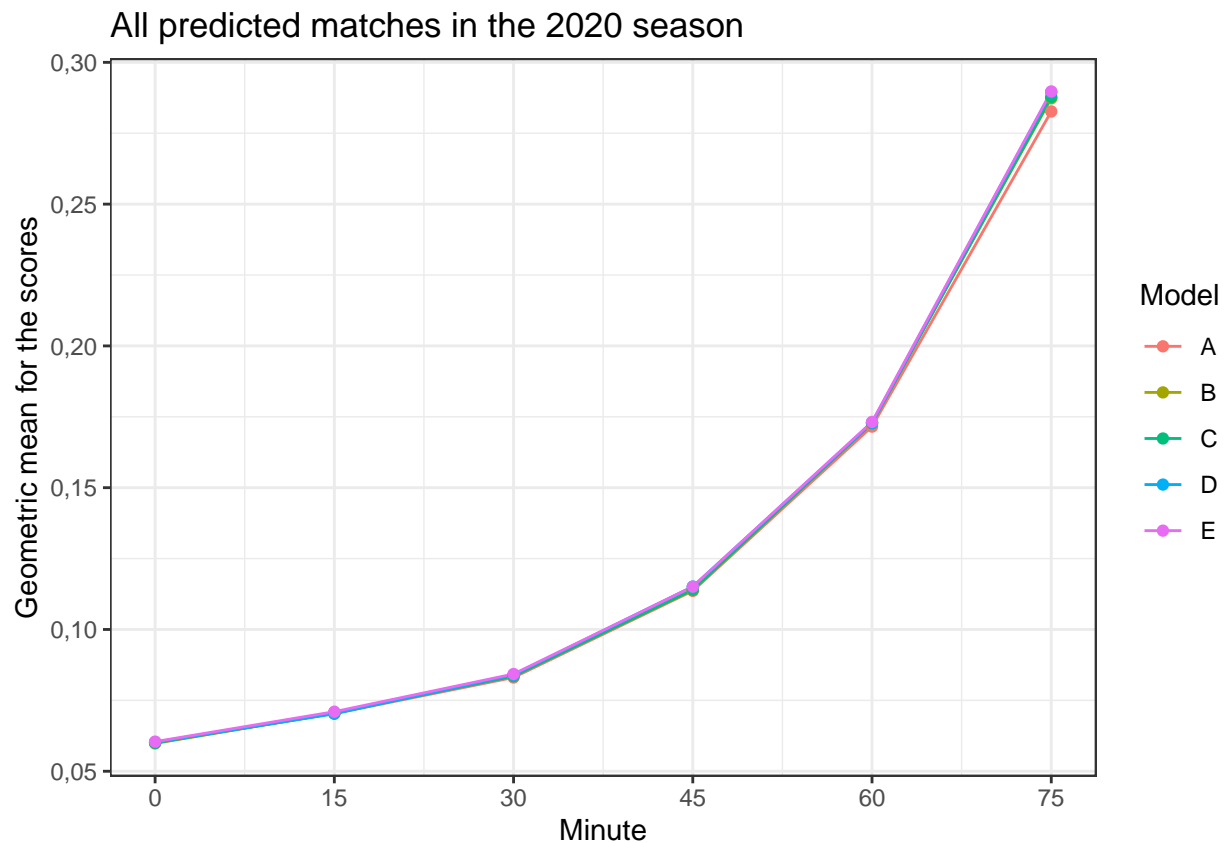

All predicted matches in the 2020 season
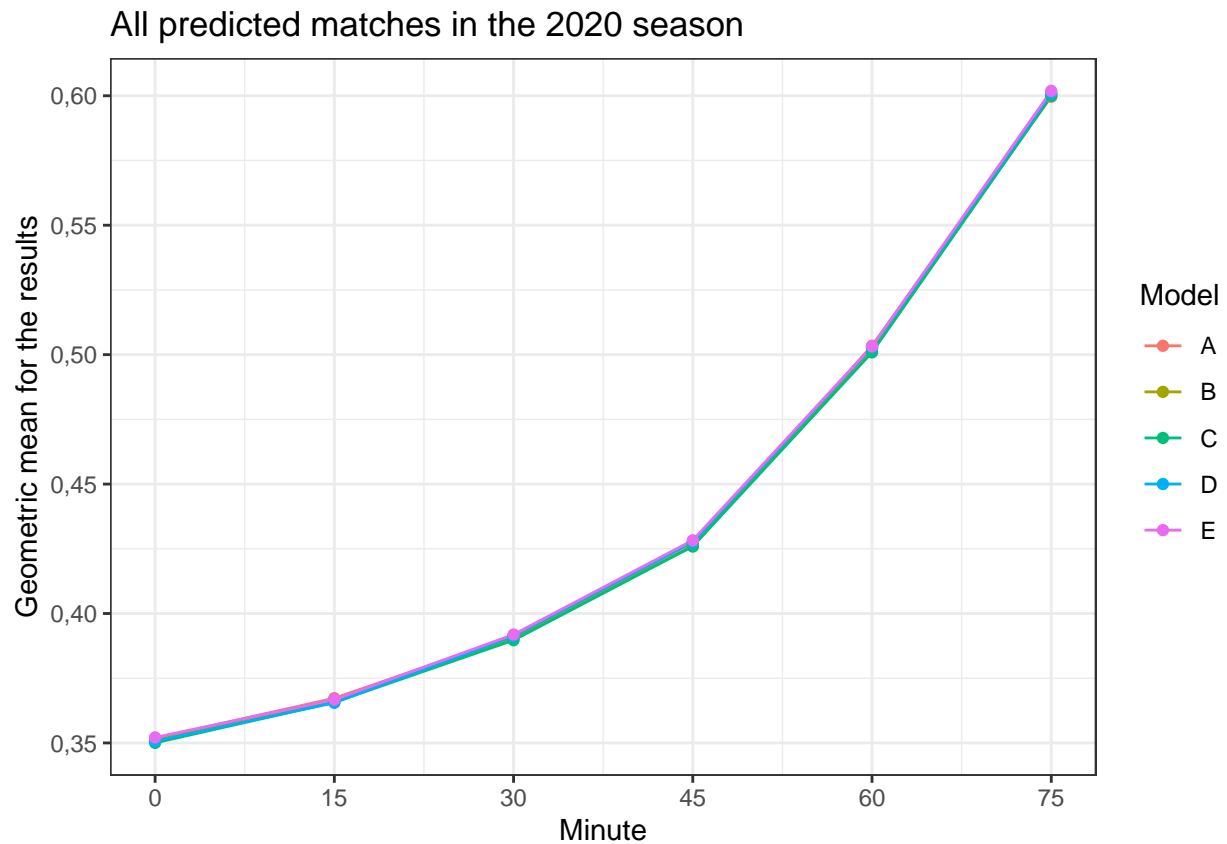
```
results_2020 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,3517607	0,3672820	0,3911344	0,4259202	0,5034214	0,5996358
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	0,0602567	0,0703762	0,0830651	0,1136109	0,1715565	0,2826861
B	0,0598204	0,0704569	0,0837564	0,1139177	0,1728954	0,2873481

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
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

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

```r
at_45 = results %>%
  select(Season, Match) %>%
  filter(Season > 2015)
```

```r
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()
}
```

```r
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))
```

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

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

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

```r
nrow(HDA_00)
```
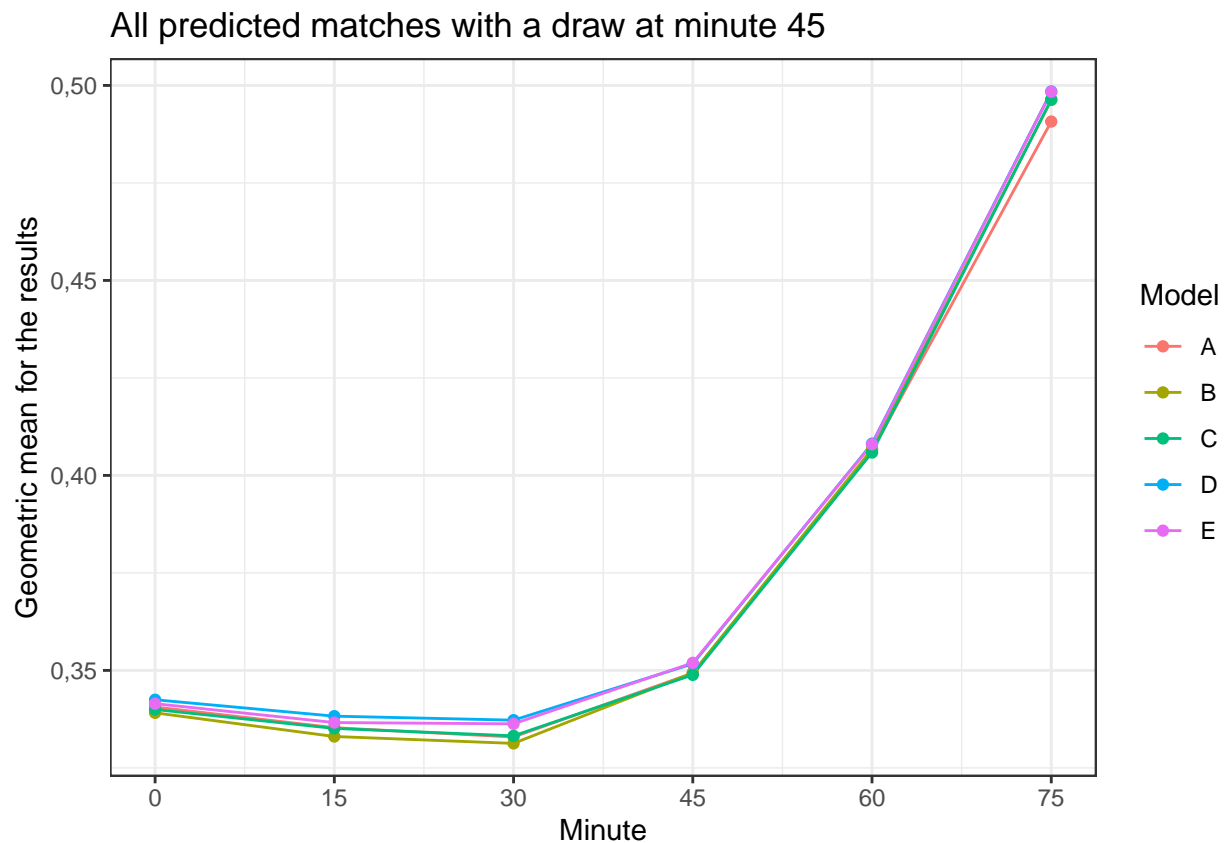
```
## [1] 838
```

```
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()
```
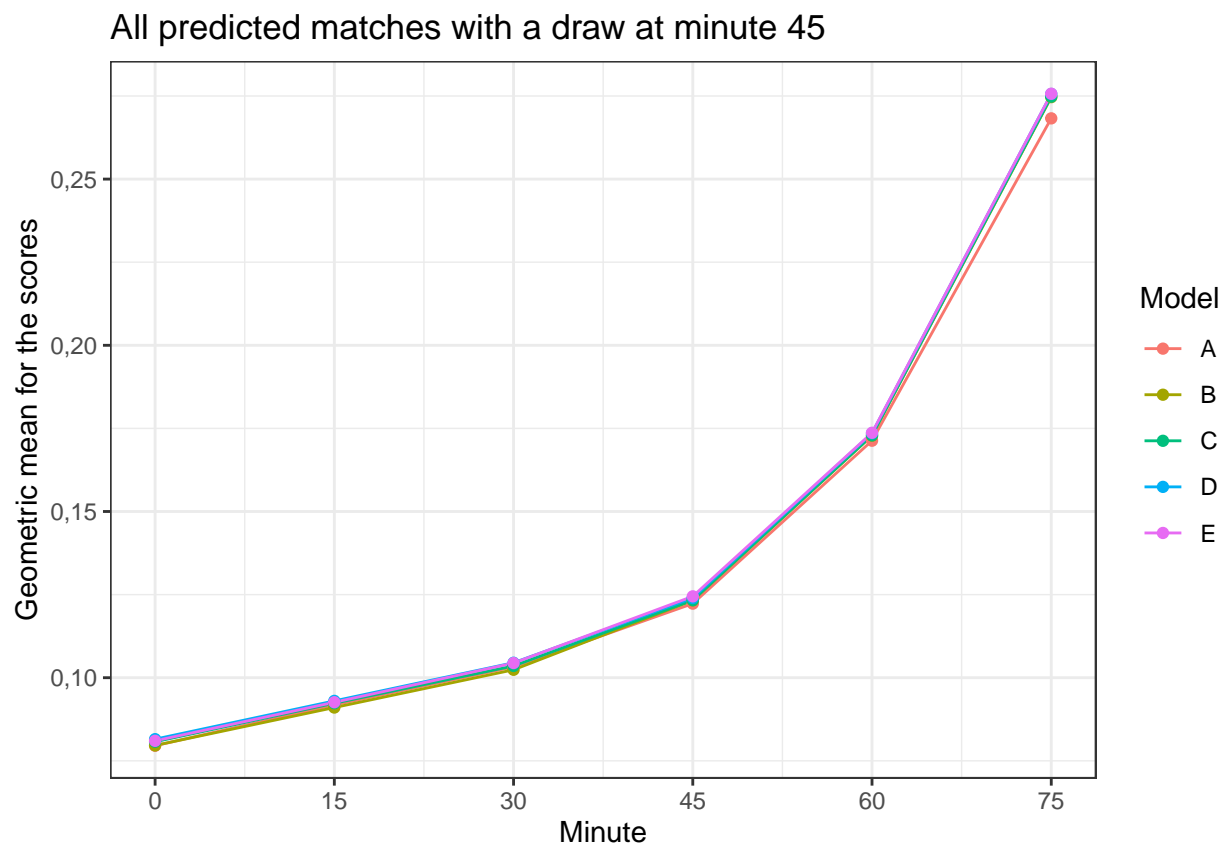
Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,3406170	0,3353077	0,3329185	0,3493102	0,4062844	0,4907415
B	0,3391190	0,3330485	0,3312800	0,3493648	0,4067135	0,4962825

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
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

```r
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")
```



```r
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	0,0795408	0,0918085	0,1032429	0,1222840	0,1713016	0,2682718
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")
```

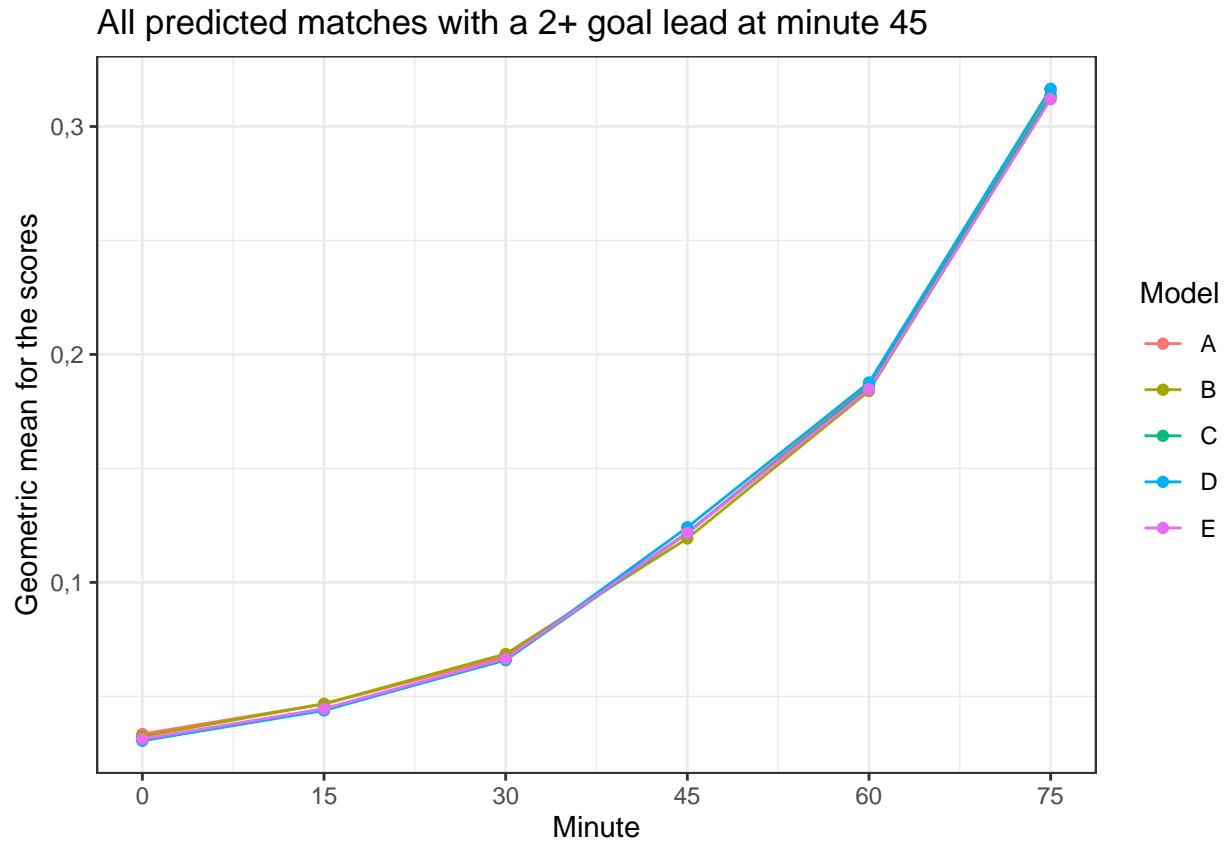## All predicted matches with a 2+ goal lead at minute 45
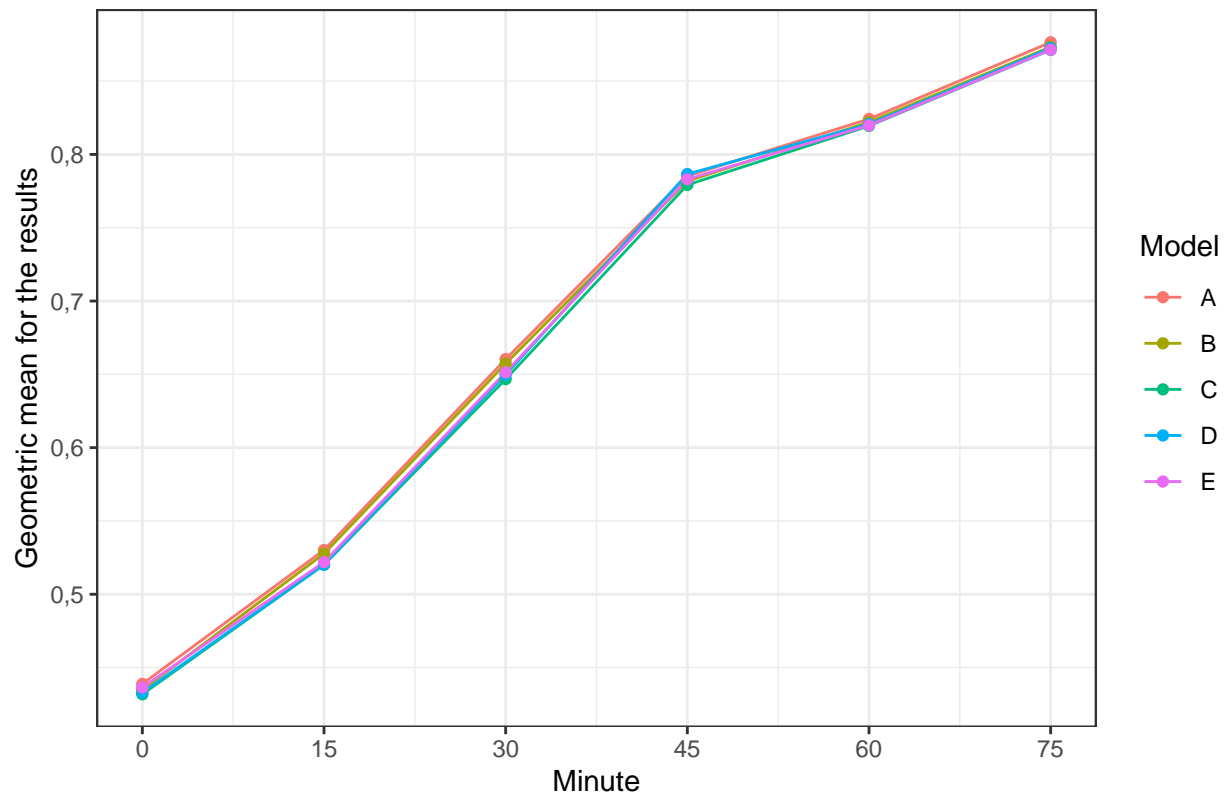

```
results_20 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,4388741	0,5301640	0,6603500	0,7852923	0,8241582	0,8764415
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")
```

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


```
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	0,0334958	0,0466985	0,0674678	0,1218122	0,1867375	0,3159467
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
```

```r
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")
```


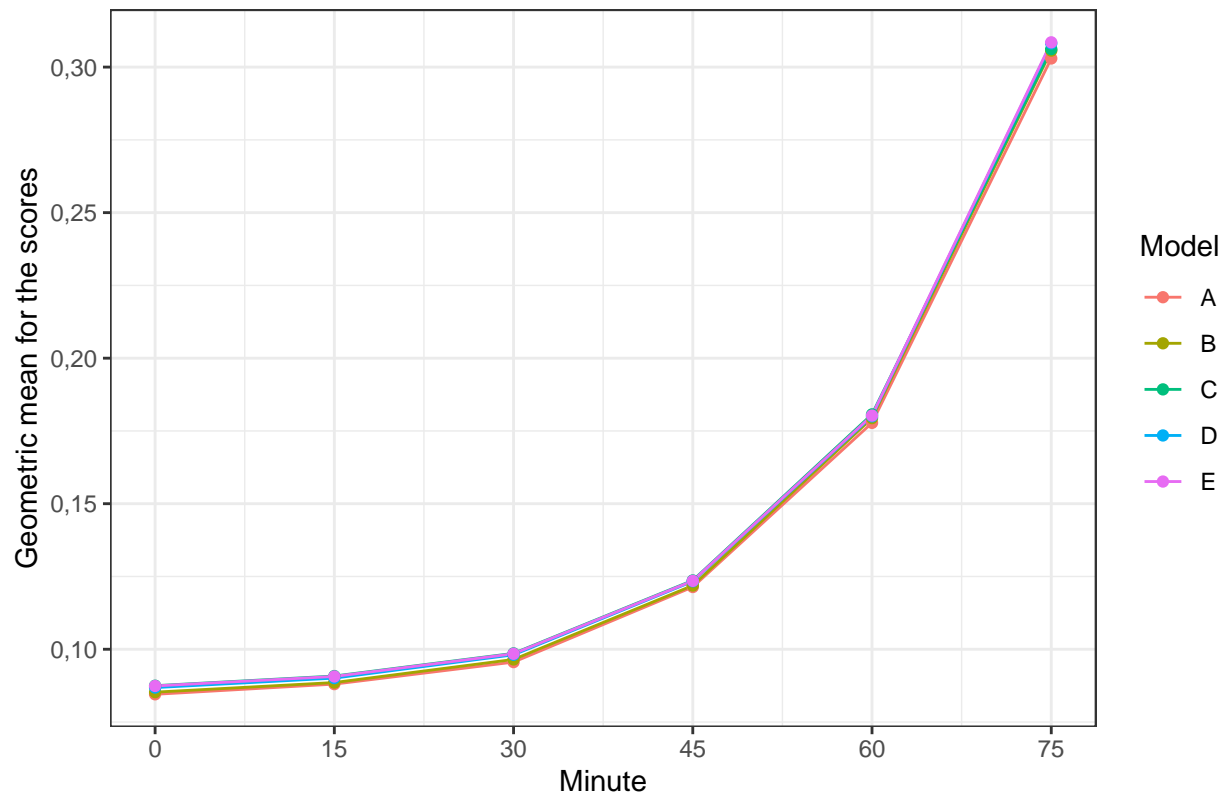All predicted matches with score 1−0 at minute 45
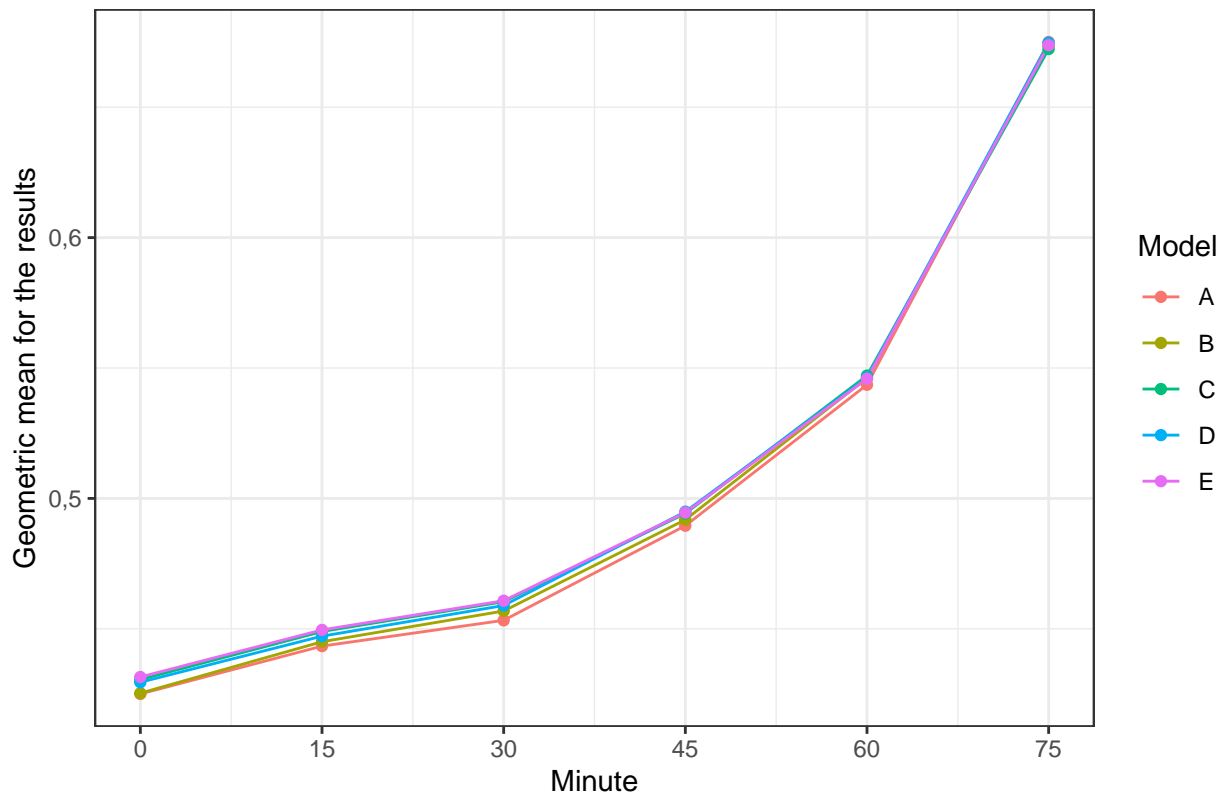
```r
results_10 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,4250278	0,4433712	0,4532632	0,4895368	0,5435804	0,6750085
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

```r
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	0,0845245	0,0880271	0,0955883	0,1213874	0,1777734	0,3029728
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")
```
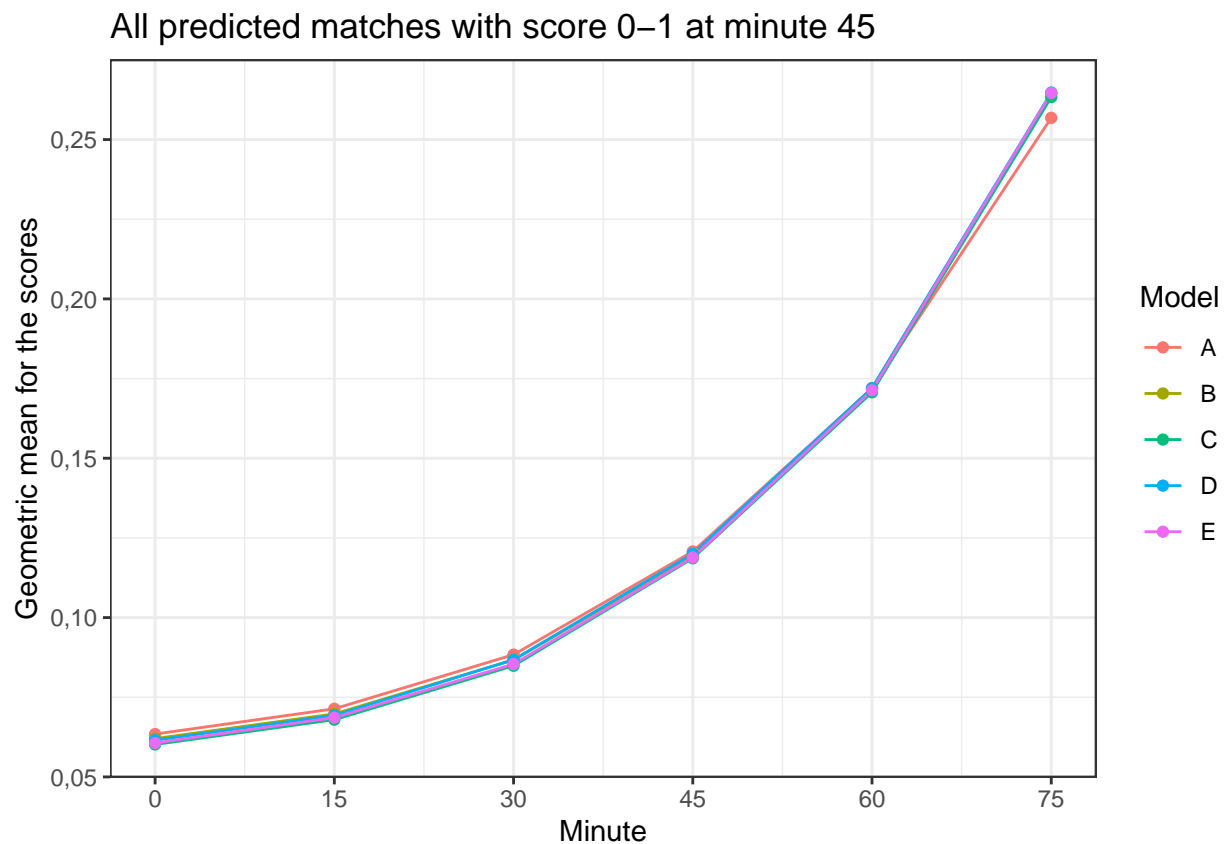
All predicted matches with score 0−1 at minute 45
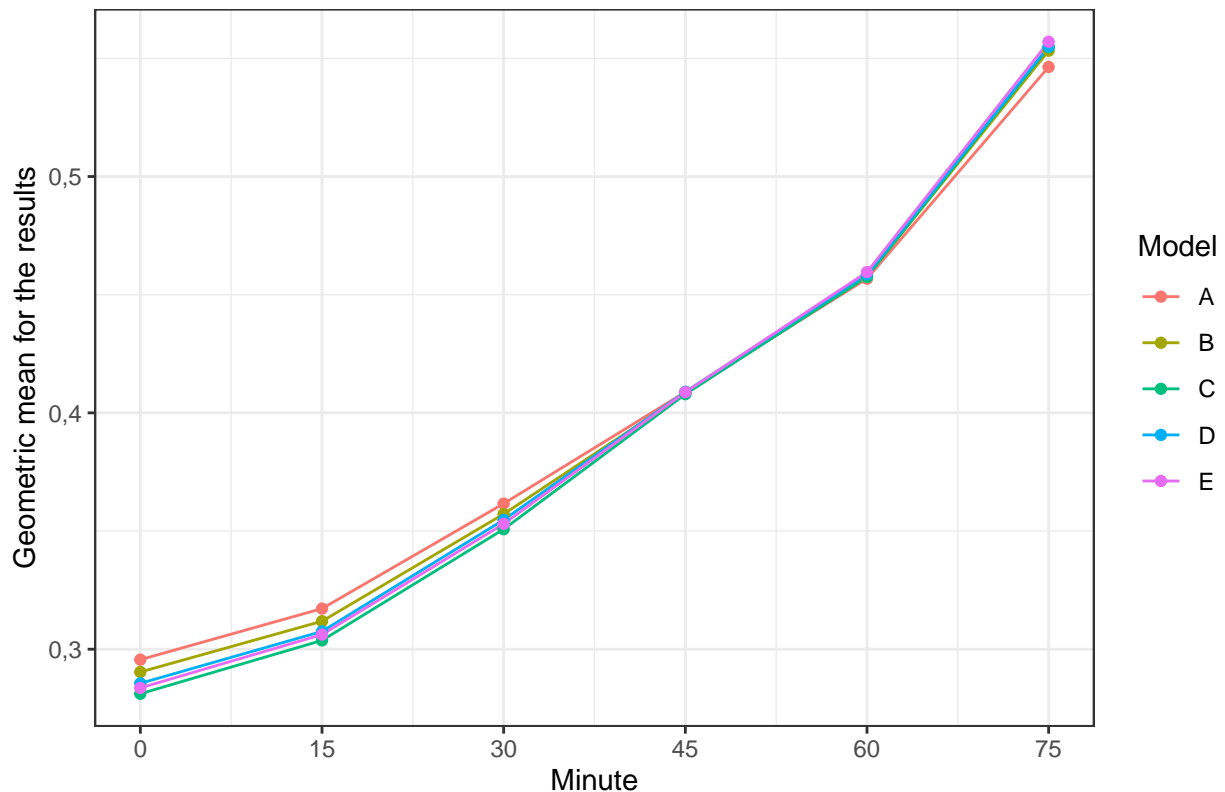

```
results_01 %>%
  pivot_wider(id_cols = "Model", values_from = "GeoMean", names_from = "Minute",
              names_prefix = "Minute ") %>%
  kable()
```

Model	Minute 0	Minute 15	Minute 30	Minute 45	Minute 60	Minute 75
A	0,2955724	0,3171840	0,3616106	0,4088270	0,4567841	0,5463802
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

```r
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")
```


All predicted matches with score 0−1 at minute 45

```
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	0,0634443	0,0713724	0,0883598	0,1207407	0,1718783	0,2568151
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