

Wydarzenia piłkarskie

Michał Kiełb, Łukasz Szlas, Tomasz Smogór

Informacje o komputerach

Michał Kiełb:

Nazwa	Wartość
System	Ubuntu 16.04.2 LTS x64
Procesor	i7-4790K 4.0 GHz, 4.4 GHz Turbo, 4 rdzenie, 8 wątków
Pamięć	16 GB, HyperX Savage 2x8GB 2400MHz DDR3 CL11
Dysk	SSD, Crucial MX200 250GB SATA3

Informacje o komputerach

Łukasz Szlas:

Nazwa	Wartość
System	Windows 8 x64
Procesor	i7-2720QM 2.2 GHz, 3.3 GHz Turbo, 4 rdzenie
Pamięć	8 GB
Dysk	SSD GoodRam Iridium Pro 240GB

Informacje o komputerach

Tomasz Smogór:

Nazwa	Wartość
System	Windows 8.1 x64
Procesor	Intel Core i7 4700MQ
Pamięć	8GB
Dysk	Seagate ST1000LM014 1TB

Kolekcja *events*

Opis: zawiera informacje o wydarzeniach sportowych.

Import:

```
mongoimport -d football-events -c events --type csv --file ever
```

Czas importu: 16 sekund

Ilość zaimportowanych dokumentów: 941 009

Wyjaśnienie pól w kolekcji *events*

Nazwa	Opis
id_odsp	unique identifier of game (odsp stands from oddsportal.com)
id_event	unique identifier of event (id_odsp + sort_order)
sort_order	chronological sequence of events in a game
time	minute of the game
text	text commentary
event_type	primary event. 11 unique events (1-Attempt(shot), 2-Corner, 3-Foul, 4-Yellow Card, 5-Second yellow card, 6-(Straight) red card, 7-Substitution, 8-Free kick won, 9-Offside, 10-Hand Ball, 11-Penalty conceded)

Wyjaśnienie pól w kolekcji *events* c.d.

Nazwa	Opis
event_type2	secondary event. 4 unique events (12 - Key Pass, 13 - Failed through ball, 14-Sending off, 15-Own goal)
side	1-Home, 2-Away
event_team	team that produced the event. In case of Own goals, event team is the team that benefited from the own goal
opponent	Help us describe this column
player	name of the player involved in main event (converted to lowercase and special chars were removed)
player2	name of player involved in secondary event

Wyjaśnienie pól w kolekcji *events* c.d.

Nazwa	Opis
player_in	player that came in (only applies to substitutions)
player_out	player substituted (only applies to substitutions)
shot_place	placement of the shot, 13 possible placement locations (1-Bit too high, 2-Blocked, 3-Bottom left corner, 4-Bottom right corner, 5-Centre of the goal, 6-High and wide, 7-Hits the bar, 8-Misses to the left, 9-Misses to the right, 10-Too high, 11-Top centre of the goal, 12-Top left corne, 13-Top right corner)
shot_outcome	4 possible outcomes (1-On target, 2-Off target, 3-Blocked, 4-Hit the post)

Wyjaśnienie pól w kolekcji *events* c.d.

Nazwa	Opis
is_goal	binary variable if the shot resulted in a goal (own goals included)
location	location on the pitch where the event happened, 19 possible locations(1-Attacking half, 2-Defensive half, 3-Centre of the box, 4-Left wing, 5-Right wing, 6-Difficult angle and long range, 7-Difficult angle on the left, 8-Difficult angle on the right, 9-Left side of the box, 10-Left side of the six yard box, 11-Right side of the box, 12-Right side of the six yard box, 13-Very close range, 14-Penalty spot, 15-Outside the box, 16-Long range, 17-More than 35 yards, 18-More than 40 yards, 19-Not recorded)

Wyjaśnienie pól w kolekcji *events* c.d.

Nazwa	Opis
bodypart	(1- right foot, 2-left foot, 3-head)
assist_method	in case of an assisted shot, 5 possible assist methods (0-None, 1-Pass, 2 Cross, 3-Headed pass, 4-Through ball)
situation	4 types: 1-Open Play, 2-Set piece (excluding Direct Free kicks), 3-Corner, 4-Free kick
fast_break	binary

Kolekcja *ginf*

Opis: zawiera metadane i kursy rynkowe dotyczące każdej gry.

Import:

```
mongoimport -d football-events -c ginf --type csv --file ginf.c
```

Czas importu: 170 ms

Ilość dokumentów: 10 112

Wyjaśnienie pól w kolekcji *ginf*

Nazwa	Opis
id_odsp	unique game identifier
link_odsp	link to oddsportal page
adv_stats	boolean if the game has detailed event data
date	Help us describe this column
league	Help us describe this column
season	Help us describe this column
country	Help us describe this column
ht	home team
at	away team

Wyjaśnienie pól w kolekcji *ginf* c.d.

Nazwa	Opis
fthg	full time home goals
ftag	full time away goals
odd_h	highest home win market odds
odd_d	highest draw market odds
odd_a	highest away market odds
odd_over	highest over 2.5 market odds
odd_under	highest under 2.5 market odds
odd_bts	highest both teams to score market odds
odd_bts_n	highest both teams NOT to score market odds

Agregacja nr 1: Przedziały czasowe, w których trafiano bramki

Etap 1:

Wybranie trafionych bramek

```
var stage1 = {  
  "$match": {  
    "is_goal": {  
      "$eq": 1  
    }  
  }  
};
```

Agregacja nr 1: Przedziały czasowe, w których trafiano bramki

Etap 2:

- pogrupowanie względem czasu trafienia bramki,
- podział przedziały czasu $[0,46)$, $[46,91)$, $[91-120)$,
- zliczenie bramek.

```
var stage2 = {  
  "$bucket": {  
    "groupBy": "$time",  
    "boundaries": [0,46,91,120],  
    "default": "Other",  
    "output": {  
      "count": { "$sum": 1 }  
    }  
  }  
};
```

Agregacja nr 1: Przedziały czasowe, w których trafiano bramki

Cała agregacja:

```
db.events.aggregate({  
  "$facet":{  
    "categorizedByGoal":[  
      stage1,  
      stage2  
    ]  
  }  
});
```


Agregacja nr 1: Przedziały czasowe, w których trafiano bramki

Wynik:

```
{
  "categorizedByGoal": [
    { "_id": 0, "count": 10744 },
    { "_id": 46, "count": 13433 },
    { "_id": 91, "count": 269 }
  ]
}
```

Wniosek

Najwięcej bramek padło w drugiej połowie meczów, a najmniej w dogrywkach.

Agregacja nr 2: Celne strzały względem pozycji z której oddano strzał, a miejsca w które trafiono

Etap 1:

Wybranie celnych strzałów

```
var stage1 = {  
  "$match": {  
    "shot_outcome": { "$eq": 1 }  
  }  
};
```

Agregacja nr 2: Celne strzały względem pozycji z której oddano strzał, a miejsca w które trafiono

Etap 2:

Pogrupowanie względem miejsca trafienia oraz pozycji i zsumowanie

```
var stage2 = {  
  "$group": {  
    "_id": {  
      "shot_place" : "$shot_place",  
      "location" : "$location"  
    },  
    "count": {  
      "$sum" : 1  
    }  
  }  
};
```

Agregacja nr 2: Celne strzały względem pozycji z której oddano strzał, a miejsca w które trafiono

Etap 3:

Posortowanie względem ilości

```
var stage3 = {  
  "$sort": {  
    "count" : -1  
  }  
};
```

Agregacja nr 2: Celne strzały względem pozycji z której oddano strzał, a miejsca w które trafiono

Cała agregacja:

```
db.events.aggregate(stage1, stage2, stage3);
```

Agregacja nr 2: Celne strzały względem pozycji z której oddano strzał, a miejsca w które trafiono

Wynik:

```
{ "_id" : { "shot_place" : 5, "location" : 3 }, "count" : 9452
{ "_id" : { "shot_place" : 5, "location" : 15 }, "count" : 7302
{ "_id" : { "shot_place" : 4, "location" : 3 }, "count" : 6213
{ "_id" : { "shot_place" : 3, "location" : 3 }, "count" : 5837
{ "_id" : { "shot_place" : 3, "location" : 15 }, "count" : 5547
{ "_id" : { "shot_place" : 4, "location" : 15 }, "count" : 5497
...
```

Wniosek

Najwięcej bramek strzelano ze środka boiska i w środek bramki.

Agregacja nr 3: Wpływ własnego boiska na wygraną

Etap 1:

Porównanie goli zdobytych u siebie i na wyjeździe

```
var stage1 = {  
  "$project":{  
    "ab":{  
      "$cmp":[ "$fthg", "$ftag" ]  
    }  
  }  
};
```

Agregacja nr 3: Wpływ własnego boiska na wygraną

Etap 2:

Wybranie odpowiednich wartości dla wygranych i przegranych

```
var stageWins2 = {  
  "$match":{  
    "ab":{ "$gt":0 }  
  }  
};  
  
var stageLost2 = {  
  "$match":{  
    "ab":{ "$eq":-1 }  
  }  
};
```


Agregacja nr 3: Wpływ własnego boiska na wygraną

Etap 3:

Zliczenie wygranych i przegranych

```
var stageWins3 = {  
  "$count": "Wygrane"  
};  
  
var stageLost3 = {  
  "$count": "Porażki"  
};
```

Agregacja nr 3: Wpływ własnego boiska na wygraną

Cała agregacja:

```
db.ginf.aggregate({  
  "$facet":{  
    "Wygrane":[  
      stage1,  
      stageWins2,  
      stageWins3  
    ],  
    "Porazki":[  
      stage1,  
      stageLost2,  
      stageLost3  
    ]  
  }  
});
```

Agregacja nr 3: Wpływ własnego boiska na wygraną

Wynik:

```
{  
  "Wygrane" : 4649,  
  "Porażki" : 2890  
}
```

Wniosek

Granie u siebie jest atutem.

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 1:

Wyszukanie Barcelony

```
var stage1 = {  
  "$match":{  
    "$or":[  
      {  
        "$and":[{ "ht":"Barcelona" },{ "at":"Real Madrid" }]  
      },  
      {  
        "$and":[{ "ht":"Real Madrid" },{ "at":"Barcelona" }]  
      }  
    ]  
  }  
};
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2:

Obliczenie wygranych grając jako gospodarz

```
var stage2 = {  
  "$facet":{  
    "WygraneDom":[  
      {  
        "$project":{ "ht":1, "ab":{ "$cmp":["$fthg","$ftag"] }  
      },  
      {  
        "$match":{  
          "$and":[ { "ab":{ "$gt":0 } }, { "ht":"Barcelona" } ]  
        }  
      },  
      { "$count":"Wygrane" }  
    ],  
    ...  
  }  
}
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2 c.d.:

Obliczenie remisów grając jako gospodarz

```
"RemisDom":[
  {
    "$project":{
      "ht":1,
      "ab":{ "$cmp":["$fthg","$ftag"] }
    }
  },
  {
    "$match":{
      "$and":[ { "ab":{ "$eq":0 } }, { "ht":"Barcelona" } ]
    }
  },
  { "$count":"Remis" }
],
...
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2 c.d.:

Obliczenie przegranych grając jako gospodarz

```
"PorazkiDom":[
  {
    "$project":{ "ht":1, "ab":{ "$cmp":["$fthg","$ftag"] }
  },
  {
    "$match":{
      "$and":[
        { "ab":{ "$eq":-1 } },
        { "ht":"Barcelona" }
      ]
    }
  },
  { "$count":"Porażki" }
],
...
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2 c.d.:

Obliczenie wygranych grając jako gość

```
"WygraneGosc":[
  {
    "$project":{ "at":1, "ab":{ "$cmp":["$ftag", "$fthg"] } },
  },
  {
    "$match":{
      "$and":[ { "ab":{ "$gt":0 } }, { "at":"Barcelona" } ]
    }
  },
  { "$count":"Wygrane" }
],
...
```


Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2 c.d.:

Obliczenie remisów grając jako gość

```
"RemisGosc":[
  {
    "$project":{ "at":1, "ab":{ "$cmp":["$ftag", "$fthg"] } },
  },
  {
    "$match":{
      "$and":[ { "ab":{ "$eq":0 } }, { "at":"Barcelona" } ]
    }
  },
  { "$count":"Remis" }
],
...
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Etap 2 c.d.:

Obliczenie porażek grając jako gość

```
"PorazkiGosc":[
  {
    "$project":{ "at":1, "ab":{ "$cmp":["$ftag", "$fthg"] } },
  },
  {
    "$match":{
      "$and":[ { "ab":{ "$eq":-1 } }, { "at":"Barcelona" } ]
    }
  },
  { "$count":"Porażki" }
]
};
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Cała agregacja

```
db.ginf.aggregate(stage1, stage2);
```

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Wynik:

```
{  
  "WygraneDom": [ { "Wygrane": 2 } ],  
  "RemisDom": [ { "Remis": 2 } ],  
  "PorazkiDom": [ { "Porażki": 2 } ],  
  "WygraneGosc": [ { "Wygrane": 3 } ],  
  "RemisGosc": [ ],  
  "PorazkiGosc": [ { "Porażki": 2 } ]  
}
```

Czyli łącznie

Wygrane: 5 (2 jako gospodarz, 3 jako gość)

Remisy: 2 (2 jako gospodarz, 0 jako gość)

Porażki: 4 (2 jako gospodarz, 3 jako gość)

Agregacja nr 4: Bilans Barcelony w 'El Clasico'

Dla potwierdzenia poprawności wyszukanych danych wyniki meczów:

2011/12

Real Madryt – FC Barcelona 1:3 Benzema – Alexis Sánchez, Xavi, Fàbregas

FC Barcelona – Real Madryt 1:2 Alexis Sánchez – Khedira, Cristiano Ronaldo

2012/13

FC Barcelona – Real Madryt 2:2 Messi (2) – Cristiano Ronaldo (2)

Real Madryt – FC Barcelona 2:1 Benzema, Sergio Ramos – Messi

2013/14

FC Barcelona – Real Madryt 2:1 Neymar, Alexis Sánchez – Jesé
Real Madryt – FC Barcelona 3:4 Benzema (2), Cristiano Ronaldo – Iniesta, Messi (3)

2014/15

Real Madryt – FC Barcelona 3:1 Cristiano Ronaldo, Pepe, Benzema – Neymar
FC Barcelona – Real Madryt 2:1 Mathieu, Suárez – Cristiano Ronaldo

2015/16

Real Madryt – FC Barcelona 0:4 Neymar, Iniesta, Suárez (2)
FC Barcelona – Real Madryt 1:2 Piqué – Benzema, Cristiano Ronaldo

2016/17

FC Barcelona - Real Madryt 1:1 Suárez - Sergio Ramos

Agregacja nr 5: Statystyki goli Lionela Messiego

Etap 1:

Wybranie bramek zdobytych przez Lionela Messiego

```
var stage1 = {  
  $match: {  
    $and: [  
      { is_goal: { $eq:1 } },  
      { player: "lionel messi" }  
    ]  
  }  
};
```

Agregacja nr 5: Statystyki goli Lionela Messiego

Etap 2:

Zliczenie goli względem części ciała, pozycji, miejsca strzału i sytuacji

```
var stage2 = {  
  $facet: {  
    "Z czego": [ { $sortByCount: "$bodypart" } ],  
    "Skad": [ { $sortByCount: "$location" } ],  
    "Gdzie": [ { $sortByCount: "$shot_place" } ],  
    "Jak": [ { $sortByCount: "$situation" } ]  
  }  
};
```

Agregacja nr 5: Statystyki goli Lionela Messiego

Cała agregacja

```
db.events.aggregate(stage1, stage2);
```


Agregacja nr 5: Statystyki goli Lionela Messiego

Wynik:

```
{
  "Z czego":[
    { "_id":2, "count":167 }, // Gole z lewej nogi
    { "_id":1, "count":30 }, // Gole z prawej nogi
    { "_id":3, "count":8 } // Gole z główki
  ],
  "Skad":[
    { "_id":3, "count":95 }, // Gol z środka pola karnego
    { "_id":14, "count":30 }, // Gol z punktu karnego
    { "_id":13, "count":17 }, // Gol z bardzo bliskiej odległoś
    { "_id":15, "count":16 }, // Gol z poza pola karnego
    { "_id":19, "count":14 }, // Nie zarejestrowano
    { "_id":9, "count":9 }, // Gol z lewej strony pola karnego
    { "_id":11, "count":8 }, // Gol z prawej strony pola karne
    { "_id":10, "count":7 }, // Gol z lewej strony pola bramkow
    { "_id":12, "count":5 }, // Gol z prawej strony pola bramko
    ...
  ]
}
```

Agregacja nr 5: Statystyki goli Lionela Messiego

Wynik c.d.:

```
    { "_id":7, "count":2 }, // Gol z ostrego prawego kąta
    { "_id":8, "count":2 } // Gol z ostrego lewego kąta
  ],
  "Gdzie":[
    { "_id":3, "count":70 }, // Gol w lewy dolny róg
    { "_id":4, "count":55 }, // Gol w prawy dolny róg
    { "_id":5, "count":36 }, // Gol w środek bramki
    { "_id":13, "count":25 }, // Gol w prawy górny róg
    { "_id":12, "count":14 }, // Gol w prawy górny róg
    { "_id":"NA", "count":5 } // Nie określono
  ],
  ...
```

Agregacja nr 5: Statystyki goli Lionela Messiego

Wynik c.d.:

```
"Jak":[
  { "_id":1, "count":154 }, // Gol z gry
  { "_id":2, "count":33 }, // Gol ze stałego fragmentu gry
  { "_id":4, "count":14 }, // Gol z rzutu wolnego / po rzucie
  { "_id":3, "count":4 } // Gol po rzucie różnym
]
```

Wniosek

Lionel Messi strzelił najwięcej bramek z lewej nogi, ze środka pola karnego, w lewy dolny róg i były to bramki z gry.

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Etap 1:

Wyszukanie sezonu 2012/2013 (Ferguson)

```
var FergusonStage1 = {  
  "$match":{  
    "$and":[  
      {  
        "$or":[ { "ht":"Manchester Utd" }, { "at":"Manchester Utd" } ],  
      },  
      {  
        "date":{" $gte":"2012-06-01", "$lt":"2013-06-01" }  
      }  
    ]  
  }  
};
```

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Etap 1 c.d.:

Wyszukanie sezonu 2013/2014 (Moyes)

```
var MoyesStage1 = {  
  "$match":{  
    "$and":[  
      {  
        "$or":[ { "ht":"Manchester Utd" }, { "at":"Manchester Utd" } ],  
      },  
      {  
        "date":{ "$gte":"2013-06-01", "$lt":"2014-06-01" }  
      }  
    ]  
  }  
};
```

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Etap 2:

Wybranie odpowiednich pól i dodanie jednego, którego wynikiem jest porównanie wartości

```
var stage2 = {  
  "$project": {  
    "ht": 1,  
    "at": 1,  
    "ab": {  
      "$cmp": [ "$fthg", "$ftag" ]  
    }  
  }  
};
```

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Etap 3:

Wybranie odpowiednich rekordów

```
var stage3 = {  
  "$match": {  
    "$or": [  
      {  
        "$and": [ { "ab": { "$gt": 0 } }, { "ht": "Manchester Utd" } ],  
      },  
      {  
        "$and": [ { "ab": { "$eq": -1 } }, { "at": "Manchester Utd" } ],  
      }  
    ]  
  }  
};
```

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Etap 4:

Zliczenie wygranych

```
var stage4 = {  
  "$count": "Wygrane"  
};
```


Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Cała agregacja

```
db.ginf.aggregate({  
  "$facet":{  
    "Ferguson":[ FergusonStage1, stage2, stage3, stage4 ],  
    "Moyes":[ MoyesStage1, stage2, stage3, stage4 ]  
  }  
});
```

Agregacja nr 6: Zmiana trenera w zespole Manchester United a liczba zwycięstw

Wynik:

```
{
  "Ferguson": [
    { "Wygrane": 56 }
  ],
  "Moyes": [
    { "Wygrane": 38 }
  ]
}
```

Wniosek

Nie łatwo jest zastąpić legendę.

Czasy

Użyta wersja MongoDB: 3.4.4

	Michał Kiełb	Łukasz Szlas	Tomasz Smogór
Agregacja nr 1	1521 ms	2220 ms	1801 ms
Agregacja nr 2	377 ms	1113 ms	801 ms
Agregacja nr 3	662 ms	1232 ms	953 ms
Agregacja nr 4	944 ms	1334 ms	1147 ms
Agregacja nr 5	730 ms	1757 ms	12108 ms
Agregacja nr 6	973 ms	1835 ms	1347 ms