

Analytics su MongoDB

Antonio Cola, Maristella Simonetti, Rosario Urso

1 Descrizione dataset

Il dataset preso in esame contiene oltre 16000 videogiochi che hanno venduto più di centomila copie. Esso è stato generato mediante web scraping dal sito vgchartz.com. Successivamente il numero di variabili è stato esteso con un'ulteriore operazione di web scraping dal sito metacritic.com. Le variabili presenti nel dataset sono:

1. Name
2. Platform
3. Year (di rilascio)
4. Genre
5. Publisher
6. NASales (vendite in Nord America, espresse in milioni di unità)
7. EUSales (vendite in Unione Europea, espresse in milioni di unità)
8. JPSales (vendite in Giappone, espresse in milioni di unità)
9. OtherSales (vendite nel resto del mondo, espresse in milioni di unità)
10. GlobalSales (vendite totali, espresse in milioni di unità)
11. CriticScore (punteggio totale assegnato dallo staff di Metacritic)
12. CriticCount (numero dei critici di Metacritic che ha assegnato un punteggio)
13. UserScore (punteggio totale assegnato dagli iscritti a Metacritic)
14. UserCount (numero degli utenti di Metacritic che ha assegnato un punteggio)
15. Developer
16. Rating (punteggio di ESRB, esrb.org)

2 Descrizione analytics

Con la prima query effettuata sul dataset abbiamo ottenuto (tra i 4 capitoli della serie di videogiochi denominata Uncharted) il capitolo su console Playstation dal 2007 al 2018 che è stato più venduto a livello globale. Con la seconda query abbiamo ricavato la media dei voti attribuiti dallo staff di Metacritic ed abbiamo ottenuto i 3 generi con la media voti più alta. Con la terza query abbiamo cercato quale fosse la console per la quale sono stati rilasciati il maggior numero di giochi tra il 2013 ed il 2016. Con la quarta query abbiamo generato la top 10 tra i giochi della compagnia Nintendo per le diverse console, ordinata per il punteggio totale assegnato dallo staff di Metacritic. Con la quinta query abbiamo effettuato il conteggio del numero di giochi FIFA e PES, per tutte le console, tra le diverse edizioni. Infine, il risultato dell'ultima query mostra la top 10, per le vendite globali, con percentuale di vendita relativa alle diverse regioni.

3 Implementazione su MongoDB

Capitolo di Uncharted più venduto a livello globale

```
1 [{ $project:
2   { _id: 0, Name: 1, Platform: 1, Year_of_Release: 1, Publisher: 1, Global_Sales: 1 } },
3   { $match:
4     { Name:
5       RegExp('.*Uncharted.*'), Platform: RegExp('PS'), Year_of_Release:
6         { $gt: 2006, $lt: 2019 } } },
7     { $sort:
8       { Global_Sales: -1 } },
9     { $limit:
10      1 } }
```

Output Query 1

```
1 [{ 'Name': "Uncharted 3: Drake's Deception",
2   'Platform': 'PS3',
3   'Year_of_Release': 2011,
4   'Publisher': 'Sony Computer Entertainment',
5   'Global_Sales': 6.74 } ]
```

Top 3 Generi con media voti più alta

```
1 [{ $project: {
2   _id: 0, Name: 1, Genre: 1, Critic_Score: 1, Publisher: 1
3 } },
4   { $group: {
5     _id: '$Genre', Media_voti: { $avg: '$Critic_Score' }
6   } },
7   { $project: {
8     _id: 1, Media_voti: { $round: [ '$Media_voti', 2 ] }
9   } },
10  { $sort: {
11    Media_voti: -1
12  } },
13  { $limit:
14    3 } }
```

Output Query 2

```
1 [{ '_id': 'Role-Playing', 'Media_voti': 72.65 },
2   { '_id': 'Strategy', 'Media_voti': 72.09 },
3   { '_id': 'Sports', 'Media_voti': 71.97 } ]
```

Console che ha rilasciato maggior numero di giochi

```
1 [{ $project: {
2   _id: 0, Platform: 1, Year_of_Release: 1
3 }},
4 { $match: {
5   Year_of_Release: { $ne: null, $gt: 2009, $lt: 2017 }
6 }},
7 { $group: {
8   _id: '$Platform', totale: { $sum: 1 }
9 }},
10 { $sort: {
11   totale: -1
12 }},
13 { $limit:
14   1}]
```

Output Query 3

```
1 [{ '_id': 'PS3', 'totale': 889}]
```

Top 10 giochi Nintendo ordinata per CriticCount

```
1 [{ $match: {
2   Publisher: { $eq: 'Nintendo' }, Name: { $ne: null }, Platform: { $ne: null },
3   Critic_Count: { $ne: null, $gt: 0 }
4 }},
5 { $project: {
6   _id: 0, Critic_Count: 1, Publisher: 1, Name_Platform: { $concat:
7     ['$Name', ' - for ', '$Platform'] }
8 }},
9 { $sort: {
10   Critic_Count: -1
11 }},
12 { $limit:
13   5}]
```

Output Query 4

```
1 [{ 'Publisher': 'Nintendo',
2   'Critic_Count': 88,
3   'Name_Platform': 'Splatoon - for WiiU' },
4 { 'Publisher': 'Nintendo',
5   'Critic_Count': 87,
6   'Name_Platform': 'Xenoblade Chronicles X - for WiiU' },
7 { 'Publisher': 'Nintendo',
8   'Critic_Count': 87,
9   'Name_Platform': 'Super Mario Galaxy 2 - for Wii' },
10 { 'Publisher': 'Nintendo',
11   'Critic_Count': 85,
12   'Name_Platform': 'Super Mario Maker - for WiiU' },
13 { 'Publisher': 'Nintendo',
14   'Critic_Count': 83,
15   'Name_Platform': 'Yoshi's Woolly World - for WiiU' }]
```

Conteggio FIFA e PES tra tutte le edizioni

```
1 [{ $project: {  
2   _id: 0, Name: 1, Platform: 1, Year_of_Release: 1, Genre: 1, Publisher: 1,  
3   Global_Sales: { $multiply: [ '$Global_Sales', 1000000 ] }, Game: { $split: [ '$Name', ' ' ] }  
4 },  
5 { $match: {  
6   Name: { $in: [ RegExp('^PES.*'), RegExp('^FIFA.*') ] }  
7 },  
8 { $group: { _id: { $first: '$Game' }, count:  
9   { $sum: 1 }, Global_Sales: { $sum: '$Global_Sales' }  
10 },  
11 { $sort:  
12   { count: -1  
13 },  
14 { $limit:  
15   2 } ] }
```

Output Query 5

```
1 [ { '_id': 'FIFA', 'count': 130, 'Global_Sales': 169940000.0 },  
2   { '_id': 'PES', 'count': 5, 'Global_Sales': 6860000.0 } ]
```

Top 10 regioni con percentuali di vendita globali

```

1  [{ $project:
2    { _id: 0, Publisher: 1, NA_Sales: 1, EU_Sales: 1, JP_Sales: 1, Other_Sales: 1,
3      Global_Sales: 1
4    } },
5    { $group:
6      { _id: '$Publisher', NA_Sales: { $sum: '$NA_Sales' },
7        EU_Sales: { $sum: '$EU_Sales' }, JP_Sales: { $sum: '$JP_Sales' },
8        Other_Sales: { $sum: '$Other_Sales' }, Global_Sales: { $sum: '$Global_Sales' }
9      } },
10   { $project:
11     { NA_Sales: { $round: [{ $divide: [{ $multiply: ['$NA_Sales', 100] }, '
12       $Global_Sales' ] }, 2] }, EU_Sales: { $round: [{ $divide: [{ $multiply: ['$EU_Sales', 100] },
13         '$Global_Sales' ] }, 2] }, JP_Sales: { $round: [{ $divide: [{ $multiply: ['$JP_Sales', 100] },
14         '$Global_Sales' ] }, 2] }, Other_Sales: { $round: [{ $divide: [{ $multiply:
15         ['$Other_Sales', 100] }, '$Global_Sales' ] }, 2] }, Global_Sales: 1
16     } },
17   { $sort:
18     { Global_Sales: -1
19   } },
20   { $project:
21     { Global_Sales: 0 } },
22   { $limit:
23     5 } ]

```

Output Query 6

```

1  [ { '_id': 'Nintendo',
2    'NA_Sales': 45.67,
3    'EU_Sales': 23.42,
4    'JP_Sales': 25.61,
5    'Other_Sales': 5.29 },
6    { '_id': 'Electronic Arts',
7    'NA_Sales': 53.67,
8    'EU_Sales': 33.48,
9    'JP_Sales': 1.28,
10   'Other_Sales': 11.55 },
11   { '_id': 'Activision',
12   'NA_Sales': 59.16,
13   'EU_Sales': 29.53,
14   'JP_Sales': 0.92,
15   'Other_Sales': 10.37 },
16   { '_id': 'Sony Computer Entertainment',
17   'NA_Sales': 43.89,
18   'EU_Sales': 30.76,
19   'JP_Sales': 12.23,
20   'Other_Sales': 13.14 },
21   { '_id': 'Ubisoft',
22   'NA_Sales': 53.59,
23   'EU_Sales': 34.35,
24   'JP_Sales': 1.59,
25   'Other_Sales': 10.43 }

```