

## Data Science e Tecnologie per le Basi di Dati

### Quaderno 4 - MongoDB

Rispondere alle seguenti domande con interrogazioni MongoDB, riportando sia l'interrogazione che il suo risultato nel dataset fornito.

#### Domanda 1

Quante stazioni hanno rispettivamente status (extra.status) “online” e “offline”?

##### Query

```
db.bike_stations.aggregate(  
  {$match:  
    {  
      "extra.status": {$in: ["online", "offline"]}  
    }  
  },  
  {$group:  
    {  
      _id: "$extra.status",  
      "count": {$sum: 1}  
    }  
  }  
)
```

##### Risultato

```
[  
  { _id: 'offline', count: 28 },  
  { _id: 'online', count: 33 }  
]
```

#### Domanda 2

Quante stazioni hanno uno status diverso da “online” e “offline”?

##### Query

```
db.bike_stations.aggregate(  
  {$match:  
    {  
      "extra.status": {$nin: ["online", "offline"]}  
    }  
  },  
  {$group:  
    {  
      _id: "$extra.status",  
      "count": {$sum: 1}  
    }  
  }  
)
```

```

        {
          _id: "$extra.status",
          "count": { $sum: 1 }
        }
      }
    )

```

#### Risultato

```

[
  { _id: 'maintenance', count: 4 }
]

```

#### Domanda 3

Per le stazioni che hanno uno status diverso da “offline” e “online”, visualizzare solo il valore del campo status.

#### Query

```

db.bike_stations.aggregate(
  { $match:
    {
      "extra.status": { $nin: ["online", "offline"]}
    }
  },
  { $project:
    {
      status: "$extra.status",
      _id: 0
    }
  }
)

```

#### Risultato

```

[
  { status: 'maintenance' },
  { status: 'maintenance' },
  { status: 'maintenance' },
  { status: 'maintenance' }
]

```

#### Domanda 4

Quali sono le stazioni attive (status = online) con una valutazione media (extra.score) maggiore o uguale a 4?

Estrarre l’elenco dei nomi di queste stazioni, ordinato in ordine alfabetico.

#### Query

```

db.bike_stations.aggregate(
  {$match:
    {
      "extra.status": {$eq: "online"}
    }
  },
  {$group:
    {
      _id: "$name",
      "avg_score": {$avg: "$extra.score"}
    }
  },
  {$match:
    {avg_score: {$gte: 4}}
  },
  {$project:
    {
      _id: 0,
      name: "$_id"
    }
  },
  {$sort: {_id: 1}}
)

```

### Risultato

```

[
  { name: '02. Pettiti' },
  { name: 'Borromini' },
  { name: 'Principi d`Acaja 1' },
  { name: 'Porta Palatina' },
  { name: '06. Municipio' },
  { name: 'Politecnico 3' },
  { name: '04. Reggia' },
  { name: 'Principi d`Acaja 2' },
  { name: '10. Gallo Praile' },
  { name: 'Corte d`Appello' },
  { name: 'Belfiore' },
  { name: 'Giolitti 1' },
  { name: 'Politecnico 1' },
  { name: 'Castello 1' },
  { name: 'San Francesco da Paola' },
  { name: 'Sant`Anselmo' },
  { name: '08. San Marchese' },
  { name: 'Tribunale' }
]

```

### Domanda 5

Qual è il nome delle stazioni non attive (`status = offline`) che hanno almeno una postazione libera (`empty_slots > 0`) oppure hanno almeno una bici a disposizione (`free_bikes > 0`)? Quante postazioni libere e quante bici sono a disposizione?

#### Query

```
db.bike_stations.aggregate(  
  {$match:  
    {  
      "extra.status": {$eq: "offline"},  
      $or: [  
        {free_bikes: {$gt: 0}},  
        {empty_slots: {$gt: 0}}  
      ]  
    }  
  },  
  {$project:  
    {  
      _id: 0,  
      name: 1,  
      free_bikes: 1,  
      empty_slots: 1  
    }  
  }  
)
```

#### Risultato

```
[  
  { empty_slots: 1, free_bikes: 0, name: '06. Le Serre' },  
  { empty_slots: 0, free_bikes: 5, name: '05. Corso Garibaldi' }  
]
```

### Domanda 6

Qual è il numero totale di recensioni (`extra.reviews`) per tutte le stazioni?

#### Query

```
db.bike_stations.aggregate(  
  {$group:  
    {  
      _id: null,  
      total_reviews: {$sum: "$extra.reviews"}  
    }  
  },  
)
```

```

    {$project:
      {
        _id:0,
        total_reviews:1
      }
    }
  )

```

#### Risultato

```

[
  { total_reviews: 15311 }
]

```

#### Domanda 7

Per ciascun valore di valutazioni medie (score), quante sono le stazioni a cui è stata assegnata quella valutazione? Ordinare il risultato per valutazione decrescente.

#### Query

```

db.bike_stations.aggregate(
  {$group:
    {
      _id: "$extra.score",
      num_stations: {$sum: 1}
    }
  },
  {$project:
    {
      _id:0,
      num_stations:1,
      score:"$_id"
    }
  },
  {$sort: {score: -1}}
)

```

#### Risultato

```

[
  { num_stations: 1, score: 4.7 },
  { num_stations: 2, score: 4.5 },
  { num_stations: 2, score: 4.4 },
  { num_stations: 2, score: 4.3 },
  { num_stations: 7, score: 4.2 },
  { num_stations: 5, score: 4.1 },
  { num_stations: 9, score: 4 }
]

```

```

{ num_stations: 9, score: 3.9 },
{ num_stations: 1, score: 3.8 },
{ num_stations: 2, score: 3.7 },
{ num_stations: 1, score: 3.6 },
{ num_stations: 4, score: 3.5 },
{ num_stations: 3, score: 3.4 },
{ num_stations: 1, score: 3.2 },
{ num_stations: 4, score: 3 },
{ num_stations: 2, score: 2.8 },
{ num_stations: 1, score: 2.7 },
{ num_stations: 1, score: 2.5 },
{ num_stations: 1, score: 2.4 },
{ num_stations: 1, score: 2.1 },
{ num_stations: 1, score: 1.5 },
{ num_stations: 1, score: 1.4 },
{ num_stations: 1, score: 1.2 },
{ num_stations: 3, score: 1 }
]

```

### Domanda 8

Qual è la valutazione media per le stazioni attive (status = online) e non attive (status = offline)?

#### Query

```

db.bike_stations.aggregate(
  {$match:
    {"extra.status":
      {$in: ["online", "offline"]}}
  },
  {$group:
    {
      _id: "$extra.status",
      avg_score: {$avg: "$extra.score"}
    }
  },
  {$project:
    {
      _id: 0,
      status: "$_id",
      avg_score: 1
    }
  }
)

```

### Risultato

```
[
  { avg_score: 3.0285714285714285, status: 'offline' },
  { avg_score: 3.8454545454545457, status: 'online' }
]
```

### Domanda 9

Quali sono le valutazioni medie per le stazioni senza bici (`free_bikes = 0`) e per quelle con almeno una bici a disposizione (`free_bikes > 0`)?

### Query

```
db.bike_stations.aggregate(
  {$addFields:
    {
      category: {
        $cond: {
          if: {$eq: ["$free_bikes", 0]},
          then: "no_free_bikes",
          else: "has_free_bikes"
        }
      }
    }
  },
  {$group:
    {
      _id: "$category",
      score: {$push: "$extra.score"}
    }
  },
  {$project:
    {
      _id:0,
      category:"$_id",
      score:1
    }
  }
)
```

### Risultato

```
[
  {
    score: [
      3.2, 3, 3.4, 2.8, 3.9, 3, 4.4, 4.2,
      3.9, 2.8, 4, 3.9, 3.5, 4.5, 1, 4.2,
    ]
  }
]
```

```

        4.7, 1.2, 3.9, 2.7, 1.4, 1, 2.5, 2.1,
        4.1, 3.7, 3.4, 3, 3, 3.8, 2.4, 3.5,
        4, 1.5, 4.4, 4.3
    ],
    category: 'no_free_bikes'
},
{
    score: [
        3.9, 4.1, 3.9, 4.3, 3.9, 4.2, 3.4, 4.1,
        4.1, 4, 4, 4, 4.2, 4, 4.5, 3.7,
        4.2, 1, 4.1, 3.9, 4, 3.9, 4, 3.5,
        4, 3.5, 3.6, 4.2, 4.2
    ],
    category: 'has_free_bikes'
}
]

```

### Domanda 10

Rispondere alla domanda 9, facendo riferimento solamente alle stazioni attive (status = online).

### Query

```

db.bike_stations.aggregate(
    {
        $match: {
            "extra.status": { $eq: "online" }
        }
    },
    {
        $addFields: {
            category: {
                $cond: {
                    if: { $eq: ["$free_bikes", 0] },
                    then: "no_free_bikes",
                    else: "has_free_bikes"
                }
            }
        }
    },
    {
        $group: {
            _id: "$category",
            score: { $push: "$extra.score" }
        }
    }
)

```



```

    {$project:
      {
        _id:0,
        category:"$_id",
        score:1
      }
    }
  )

```

### Risultato

```

[
  {
    score: [
      3.9, 4.1, 3.9, 4.3, 3.9, 4.2, 3.4, 4.1,
      4.1, 4, 4, 4, 4.2, 4, 4.5, 3.7,
      1, 4.1, 3.9, 4, 3.9, 4, 3.5, 4,
      3.5, 3.6, 4.2, 4.2
    ],
    category: 'has_free_bikes'
  },
  {
    score: [ 3.9, 4.1, 3.4, 3.8, 3.5 ],
    category: 'no_free_bikes'
  }
]

```

### Domanda 11

Quali sono i nomi delle 3 stazioni con bici disponibili (`free_bikes > 0`) più vicine al punto `[45.07456, 7.69463]`? Quante bici sono disponibili?

### Query

```

db.bike_stations.find(
  {$and:
    [
      {
        free_bikes: {$gt: 0}
      },
      {location:
        {$near:
          {$geometry:
            {
              type: "Point",
              coordinates: [45.07456, 7.69463]
            }
          }
        }
      }
    ]
  }
)

```

```

    }
  }
]
},
{
  _id:0,
  name:1,
  free_bikes:1
}
).limit(3)

```

### Risultato

```

[
  { free_bikes: 5, name: 'Palermo 2' },
  { free_bikes: 5, name: 'Castello 1' },
  { free_bikes: 4, name: 'San Francesco da Paola' }
]

```

### Domanda 12

Quali sono i nomi delle 3 stazioni con bici disponibili (`free_bikes > 0`) più vicine alla stazione “Politecnico 4”? Quante bici sono disponibili?

### Query

```

db.bike_stations.find(
  {$and:
    [
      {
        free_bikes: {$gt: 0}
      },
      {location:
        {$near:
          {$geometry:
            db.bike_stations.find(
              {name: {$eq: "Politecnico 4"}},
              {_id:0,location:1}
            ).next().location
          }
        }
      }
    ]
  },
  {
    _id:0,
    name:1,
    free_bikes:1
  }
)

```

```
    }  
  ).limit(3)
```

#### Risultato

```
[  
  { free_bikes: 9, name: 'Politecnico 1' },  
  { free_bikes: 5, name: 'Politecnico 3' },  
  { free_bikes: 3, name: 'Tribunale' }  
]
```