

Data Science e Tecnologie per le Basi di Dati

Quaderno 4 – MongoDB

Domanda 1

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

Query	Risposta
<code>db["bike stations"].find({ "extra.status": { \$in: ["online", "offline"] } }).count()</code>	61

Domanda 2

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

Query	Risposta
<code>db["bike stations"].find({ "extra.status": { \$nin: ["online", "offline"] } }).count ()</code>	4

Domanda 3

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

Query	Risposta
<code>db.getCollection('bike stations').find({ 'extra.status': { \$nin: ['online', 'offline'] } }, { 'extra.status': 1, _id: 0 });</code>	<code>[{"extra": {"status": "maintenance"}}, {"extra": {"status": "maintenance"}}, {"extra": {"status": "maintenance"}}, {"extra": {"status": "maintenance"}}]</code>

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	Risposta
<code>db["bike stations"].aggregate([{ \$match: { 'extra.status': { \$eq: "online" }, 'extra.score': { \$gte: 4 } } }, { \$sort: { "name": 1 } // Ordina per nome in ordine alfabetico ascendente }, { \$group: { _id: null, names: { \$push: "\$name" } } }]);</code>	<code>{"_id":null,"names":["02. Pettiti","04. Reggia","06. Municipio","08. San Marchese","10. Gallo Praile","Belfiore","Borromini","Castello 1","Corte d'Appello","Giolitti 1","Politecnico 1","Politecnico 3","Porta Palatina","Principi d'Acaja 1","Principi d'Acaja 2","San Francesco da Paola","Sant'Anselmo","Tribunale"]}</code>

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	Risposta
<pre>db["bike stations"].aggregate([{ \$match: { "extra.status": "offline", \$or: [{ "empty_slots": { \$gt: 0 } }, { "free_bikes": { \$gt: 0 } }] } }, { \$group: { _id: null, names: { \$push: "\$name" }, totalEmptySlots: { \$sum: "\$empty_slots" }, totalFreeBikes: { \$sum: "\$free_bikes" } } }])</pre>	<pre>{ _id: null, names: ['06. Le Serre', '05. Corso Garibaldi'], totalEmptySlots: 1, totalFreeBikes: 5 }</pre>

Domanda 6

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

Query	Risposta
<pre>db["bike stations"].aggregate([{ \$group: { _id: null, totalExtraReviews: { \$sum: "\$extra.reviews" } } }])</pre>	<pre>{ _id: null, totalExtraReviews: 15311 }</pre>

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	Risposta
<pre>db["bike stations"].aggregate([{ \$group: { _id: "\$extra.score", stations: { \$sum: 1 } } }, { \$sort: { _id: -1 } }])</pre>	<pre>[{"_id": 4.7, "stations": 1}, {"_id": 4.5, "stations": 2}, {"_id": 4.4, "stations": 2}, {"_id": 4.3, "stations": 2}, {"_id": 4.2, "stations": 7}, {"_id": 4.1, "stations": 5}, {"_id": 4, "stations": 9}, {"_id": 3.9, "stations": 9}, {"_id": 3.8, "stations": 1}, {"_id": 3.7, "stations": 2}, {"_id": 3.6, "stations": 1}, {"_id": 3.5, "stations": 4}, {"_id": 3.4, "stations": 3}, {"_id": 3.2, "stations": 1}, {"_id": 3, "stations": 4}, {"_id": 2.8, "stations": 2}, {"_id": 2.7, "stations": 1}, {"_id": 2.5, "stations": 1}, {"_id": 2.4, "stations": 1}, {"_id": 2.1, "stations": 1}, {"_id": 1.5, "stations": 1}, {"_id": 1.4, "stations": 1}, {"_id": 1.2, "stations": 1}, {"_id": 1, "stations": 3}]</pre>

Domanda 8

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

Query	Risposta
<pre>db["bike stations"].aggregate([{ \$match: { "extra.status": { \$in: ["online", "offline"] } } }, { \$group: { _id: "\$extra.status", averageScore: { \$avg: "\$extra.score" } } }])</pre>	<pre>[{"_id": "offline", "averageScore": 3.0285714285714285}, {"_id": "online", "averageScore": 3.8454545454545452}]</pre>

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	Risposta
<pre>db["bike stations"].aggregate([{ \$group: { _id: { \$cond: { if: { \$eq: ["\$free_bikes", 0] }, then: "No bikes", else: "Has bikes" } }, // Raggruppa per "No bikes" o "Has bikes" avgScore: { \$avg: "\$extra.score" } // Calcola la media dei punteggi } }]);</pre>	<pre>[{ "_id": "Has bikes", "avgScore": 3.8758620689655174 }, { "_id": "No bikes", "avgScore": 3.2305555555555556 }]</pre>

Domanda 10

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

Query	Risposta
<pre>db["bike stations"].aggregate([{ \$match: { "extra.status": { \$eq: "online" } } }, { \$group: { _id: { \$cond: { if: { \$eq: ["\$free_bikes", 0] }, then: "No bikes", else: "Has bikes" } }, avgScore: { \$avg: "\$extra.score" } } }]);</pre>	<pre>[{ "_id": "Has bikes", "avgScore": 3.8642857142857143 }, { "_id": "No bikes", "avgScore": 3.7399999999999998 }]</pre>

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	Risposta
<pre> db.getCollection('bike stations').aggregate([{ \$geoNear: { near: { type: 'Point', coordinates: [7.69463, 45.07456] }, distanceField: 'distance', spherical: true } }, { \$match: { 'extra.status': 'online' } }, { \$limit: 3 }, { \$group: { _id: null, names: { \$push: '\$name' }, bikesNum: { \$sum: '\$free_bikes' } } }]); </pre>	<pre> [{ "_id": null, "names": ["Borromini", "San Francesco da Paola", "Belfiore"], "bikesNum": 15 }] </pre>

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	Risposta
<pre> var station = db['bike stations'].findOne({ name: "Politecnico 4" }); var coordinates = station.location.coordinates; db['bike stations'].aggregate([{ \$geoNear: { near: { type: 'Point', coordinates: coordinates }, distanceField: 'distance', spherical: true } }, { \$match: { 'extra.status': 'online' } }, { \$limit: 3 }, { \$group: { _id: null, names: { \$push: '\$name' }, bikesNum: { \$sum: '\$free_bikes' } } }]); </pre>	<pre> { _id: null, names: ['Politecnico 1', 'Politecnico 3', 'Tribunale'], bikesNum: 17 } </pre>