

# Business Intelligence per i Big Data

## Esercitazione n. 6

### Passi preliminari per lo svolgimento dell'esercitazione

L'obiettivo di questa esercitazione è quello di prendere confidenza con la struttura dei database NoSQL, esplorando e interrogando una base dati grazie al tool MongoDB Compass.

#### Connessione alla base dati remota

1. Scaricare la versione 1.30.1 di MongoDB Compass dal seguente link:  
<https://github.com/mongodb-js/compass/releases/tag/v1.30.1>

Scegliere la versione corrispondente al sistema operativo utilizzato e, al termine del download, lanciare l'eseguibile.

2. Connettersi alla base dati remota tramite le seguenti credenziali:
  - a. **Hostname:** bigdatadb.polito.it
  - b. **Port:** 27017
  - c. **Authentication:** Username/Password
  - d. **Username:** Compass
  - e. **Password:** Compass19!
  - f. **Authentication database:** dbdmg
  - g. **SSL:** Unvalidated (insecure)
3. Click su **Connect**.
4. Accedere su **dbdmg**.
5. Accedere alla base **Parkings**.

## Descrizione del Database *Parkings*

Il database Parkings contiene informazioni relative ai parcheggi di auto associate ad un servizio di *Car Sharing*. Gli attributi più importanti per ogni documento sono elencati nella tabella 1.

Nome	Tipo	Descrizione
<b>_id</b>	objectid	identificativo del documento
<b>address</b>	string	Indirizzo di parcheggio del veicolo
<b>city</b>	string	Città nella quale è avvenuto il parcheggio
<b>engineType</b>	string	Identificativo del tipo di motore del veicolo
<b>exterior</b>	string	Stringa identificativa delle condizioni esterne del veicolo
<b>final_date</b>	date	Data e ora di fine stazionamento del veicolo
<b>fuel</b>	int32	Livello di carburante del veicolo durante il periodo di stazionamento
<b>init_date</b>	date	Data e ora di inizio stazionamento del veicolo
<b>interior</b>	string	Stringa identificativa delle condizioni interne del veicolo
<b>loc</b>	coordinates	Coordinate della posizione di parcheggio del veicolo
<b>plate</b>	int32	Identificativo della targa del veicolo
<b>smartphoneRequired</b>	boolean	Valore binario (Vero/Falso) che indica qualora sia necessario lo smartphone per il noleggio del veicolo
<b>vendor</b>	string	Azienda proprietaria del veicolo da noleggiare
<b>vin</b>	string	Identificativo del telaio del veicolo

Tabella 1: descrizione del database Parkings.

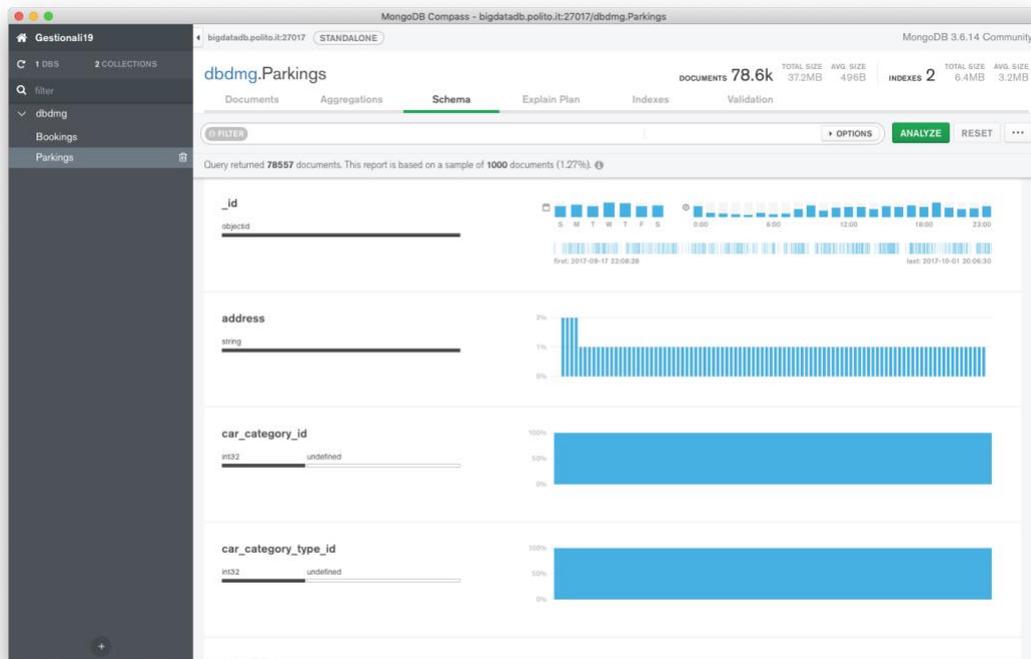
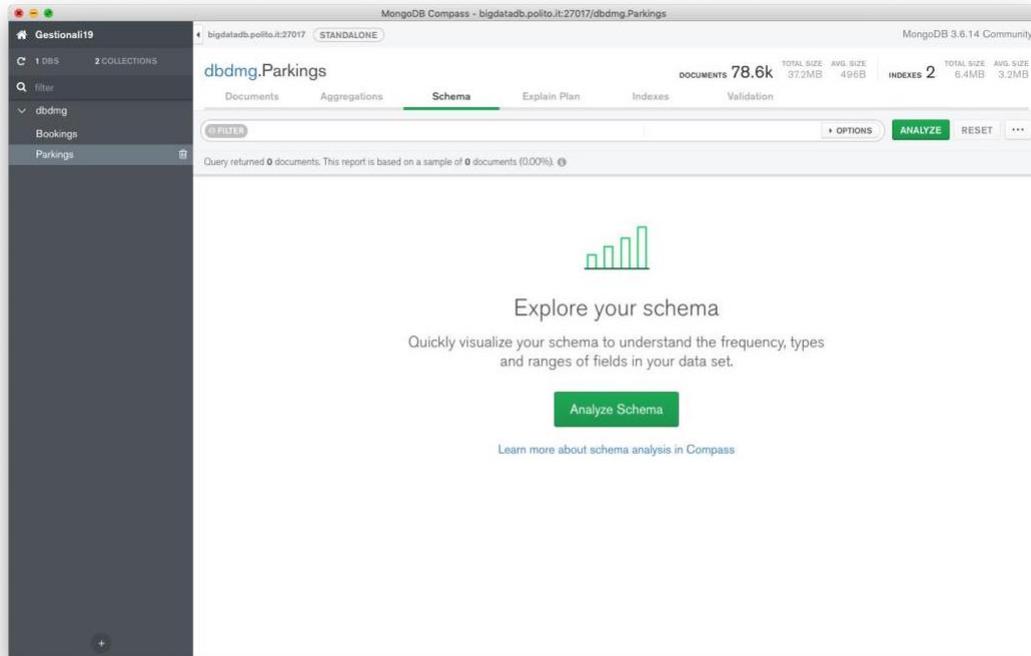
## Descrizione del Database *Bookings*

Il database Bookings contiene informazioni relative al noleggio di veicoli per un servizio di *Car Sharing*. Gli attributi più importanti per ogni documento sono elencati nella tabella 2.

Nome	Tipo	Descrizione		
<b>_id</b>	objectid	identificativo del documento		
<b>car_name</b>	string	Modello del veicolo		
<b>city</b>	string	Città nella quale è avvenuto il noleggio		
<b>distance</b>	int32	Distanza percorsa durante il noleggio (in metri)		
<b>driving</b>	object	<b>distance</b>	int32	Distanza percorsa durante il noleggio (in metri)
		<b>duration</b>	int32	Durata del noleggio (in secondi)
<b>engineType</b>	string	Identificativo del tipo di motore del veicolo		
<b>exterior</b>	string	Stringa identificativa delle condizioni esterne del veicolo		
<b>final_address</b>	string	Indirizzo della posizione di fine noleggio		
<b>final_date</b>	date	Data e ora di fine noleggio		
<b>final_fuel</b>	int32	Livello del carburante a fine noleggio		
<b>init_address</b>	int32	Indirizzo della posizione di inizio noleggio		
<b>init_date</b>	date	Data e ora di inizio noleggio		
<b>init_fuel</b>	int32	Livello del carburante a inizio noleggio		
<b>interior</b>	string	Stringa identificativa delle condizioni interne del veicolo		
<b>plate</b>	int32	Identificativo della targa del veicolo		
<b>smartphoneRequired</b>	boolean	Valore binario (Vero/Falso) che indica qualora sia necessario lo smartphone per il noleggio del veicolo		
<b>vendor</b>	string	Azienda proprietaria del veicolo da noleggiare		
<b>walking</b>	object	<b>distance</b>	int32	Distanza percorsa a piedi per raggiungere il veicolo.
		<b>duration</b>	int32	Durata del percorso a piedi per raggiungere il veicolo.

Tabella 2: descrizione del database Bookings.

# 1. Analizzare la base dati con lo *schema analyzer* (Parkings)

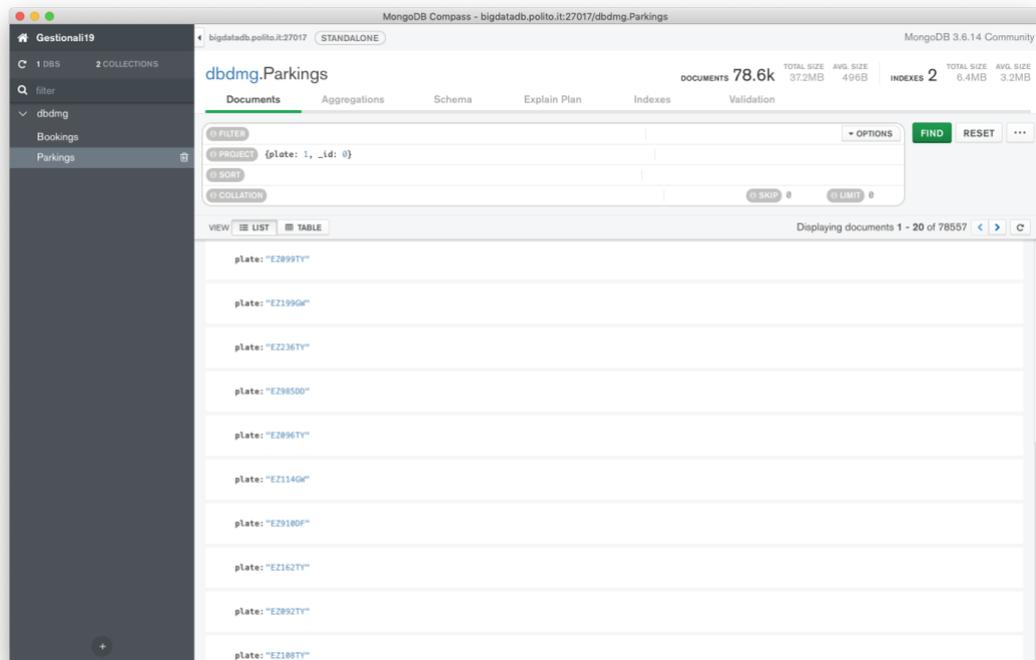


1. Identificare l'intervallo/gli intervalli orari con maggiore **richiesta di parcheggio (inizio stazionamento)** di veicoli.
2. Identificare l'intervallo/gli intervalli orari nei quali i veicoli **vengono noleggiati (fine stazionamento)** più di frequente.
3. Filtrare sulla mappa una zona di interesse e analizzare l'intervallo/gli intervalli orari di **inizio noleggio (fine stazionamento)** più frequenti.
4. Per i veicoli filtrati al passo precedente, visualizzare solo quelli che hanno un livello di carburante residuo maggiore del 90%.
5. Visualizzare su mappa i veicoli che hanno un livello di carburante residuo inferiore al 5%.

Concentrarsi ora sul database *Bookings*.

6. Identificare la/le percentuali più frequenti di livello di carburante a inizio noleggio.
7. Identificare la/le percentuali più frequenti di livello di carburante a fine noleggio.
8. Identificare il range di distanza più frequente se il percorso fosse stato fatto a piedi.

## 2. Interrogare la base dati



1. (Parkings) Trovare le targhe e gli indirizzi di parcheggio dei veicoli che hanno iniziato il noleggio (finito stazionamento) dopo le 6.00 del 30-09-2017. (Hint: usare la funzione `Date("<YYYY-mm-ddTHH:MM:ss>")` )
2. (Parkings) Trovare gli indirizzi e il livello di carburante residuo per le auto che hanno avuto durante lo stazionamento almeno il 70% di carburante residuo e ordinare i risultati in base al loro livello di carburante decrescente.
3. (Parkings) Trovare la targa, tipo di motore, e livello di carburante dei veicoli di 'car2go' che hanno buone condizioni interne ed esterne.

4. (Bookings) Per i noleggi che hanno un percorso a piedi maggiore di 15 km, visualizzare la data e l'orario di inizio noleggio e il livello di carburante a inizio noleggio. Visualizzare i risultati ordinati in base al livello di carburante iniziale decrescente.
5. (Bookings) Raggruppare i documenti in base al loro livello di carburante a **fine** noleggio. Per ogni gruppo visualizzare il livello di carburante medio a **inizio** noleggio.
6. (Bookings) Visualizzare la distanza media percorsa nei noleggi per ciascun fornitore del servizio. In media con quale fornitore del servizio gli utenti percorrono una distanza maggiore?
7. (Bookings) Quali sono i primi tre veicoli che hanno percorso una maggiore distanza considerando tutti i noleggi? Visualizzarne la targa e il conteggio in ordine decrescente per conteggio.
8. (Bookings) Quanti sono i veicoli che, considerando tutti i noleggi, hanno percorso almeno 350000m?
9. (Bookings) Ripetere la query precedente considerando solamente i veicoli di proprietà 'car2go'.

### **Interrogazioni Bonus**

10. (Parkings) Quanti sono i parcheggi effettuati a meno di 1km da Piazza San Carlo (coordinate 7.683016, 45.067764).

**Hint:** usare l'operatore [\\$geoWithin](#) insieme a [\\$centerSphere](#).

11. Ripetere l'interrogazione al passo precedente con un punto di interesse personale nell'area metropolitana di Torino (e.g. indirizzo di casa) usando Open Street Maps per trovare le coordinate esatte (www.openstreetmap.org, invertire l'ordine delle coordinate).