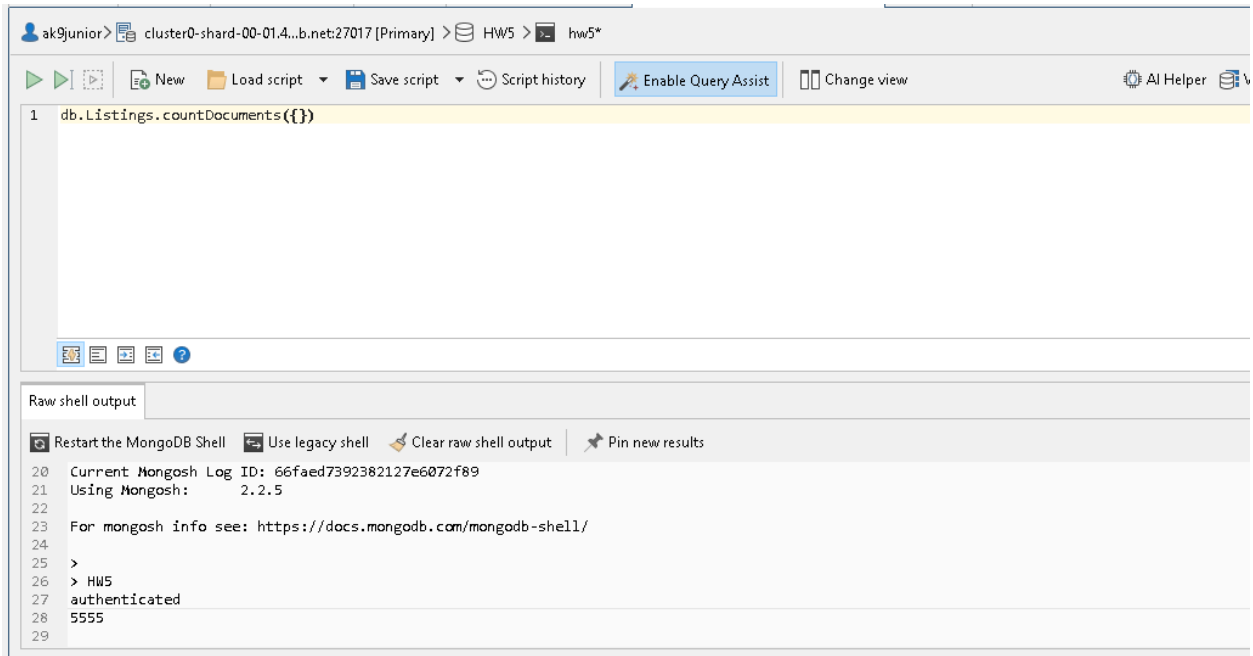


Muhammad Areeb Kazmi – HW 5

Mk07202@st.habib.edu.pk

Q1



The screenshot shows the MongoDB Shell interface. The top bar indicates the user is 'ak9junior' and the current database is 'cluster0-shard-00-01.4...b.net:27017 [Primary]'. The current collection is 'HW5' and the current document is 'hw5*'. The main editor area contains a single query: `db.Listings.countDocuments({})`. Below the editor, the 'Raw shell output' section shows the following text:

```
20 Current Mongosh Log ID: 66faed7392382127e6072f89
21 Using Mongosh: 2.2.5
22
23 For mongosh info see: https://docs.mongodb.com/mongosh-shell/
24
25 >
26 > HW5
27 authenticated
28 5555
29
```

Q2

```
ak9junior> cluster0-shard-00-01.4...b.net:27017 [Primary] > HW5 > hw5
New Load script Save script Script history Enable Query Assist Change view
1 //db.Listings.countDocuments({})
2
3 db.Listings.countDocuments({property_type: "House"})

Raw shell output
Restart the MongoDB Shell Use legacy shell Clear raw shell output Pin new results
20 Current Mongosh Log ID: 66faee078d71d6b533072f89
21 Using Mongosh: 2.2.5
22
23 For mongosh info see: https://docs.mongodb.com/mongodb-shell/
24
25 >
26 > HW5
27 authenticated
28 606
29
```

This shows that around 10.9% of the total property types are that of Houses.

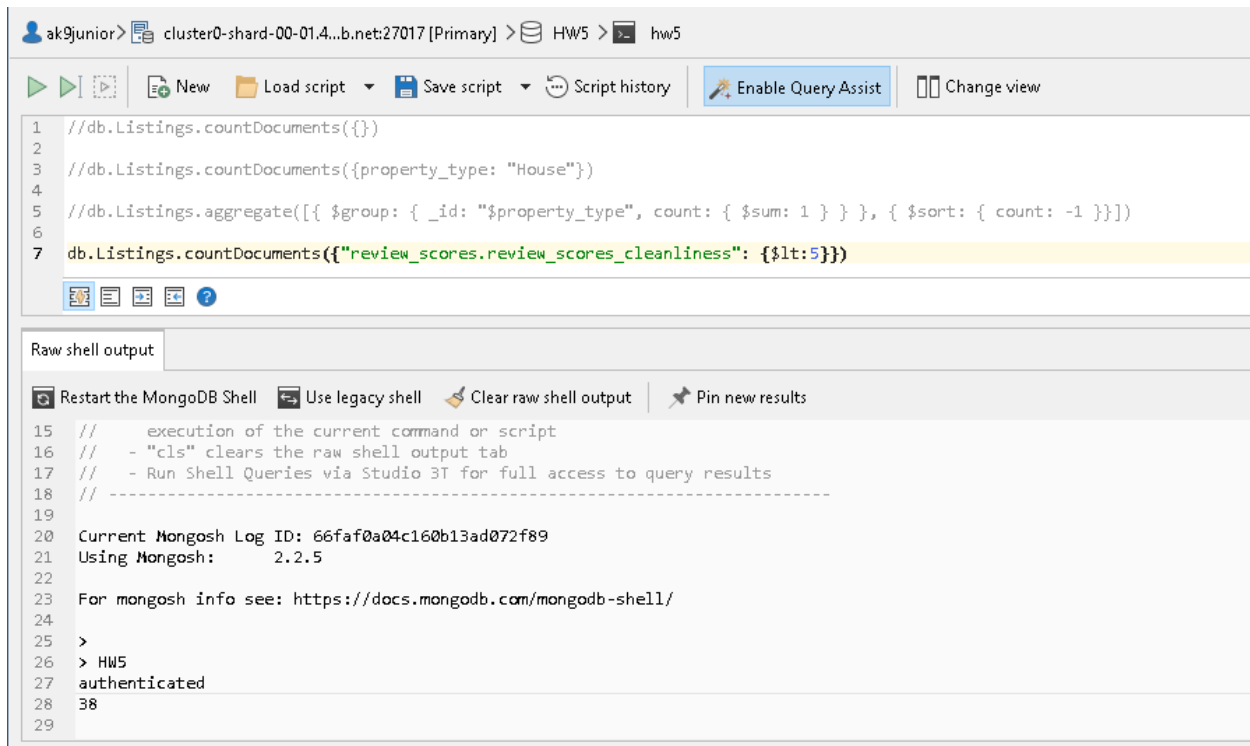
Q3)

```
4 //db.Listings.aggregate([{$group: {'_id': '$property_type', count: {$sum: 1}}}, {$sort: {'count': -1}}])
5 db.Listings.aggregate([{$group: {'_id': '$property_type', count: {$sum: 1}}}, {$sort: {'count': -1}}])

Raw shell output Aggregate Query (line 5)
50 Documents 1 to 36
Query Code Explain Query Table View
Listings > count
_id count
Apartment 3626.0
House 606.0
Condominium 399.0
Serviced apartm... 185.0
Loft 142.0
Townhouse 108.0
Guest suite 81.0
1 document selected
Count Documents 00:00:00.226
```

We observe that majority of the properties are apartments while others are types that we see in suburbs probably and community places like townhouses etc. There are various property types which only exist in single digits or quantity in the listing. While majority are apartments, houses and cordominiums.

Q4)



The screenshot shows the MongoDB Studio interface. The top toolbar includes buttons for 'New', 'Load script', 'Save script', 'Script history', 'Enable Query Assist', and 'Change view'. The script editor contains the following code:

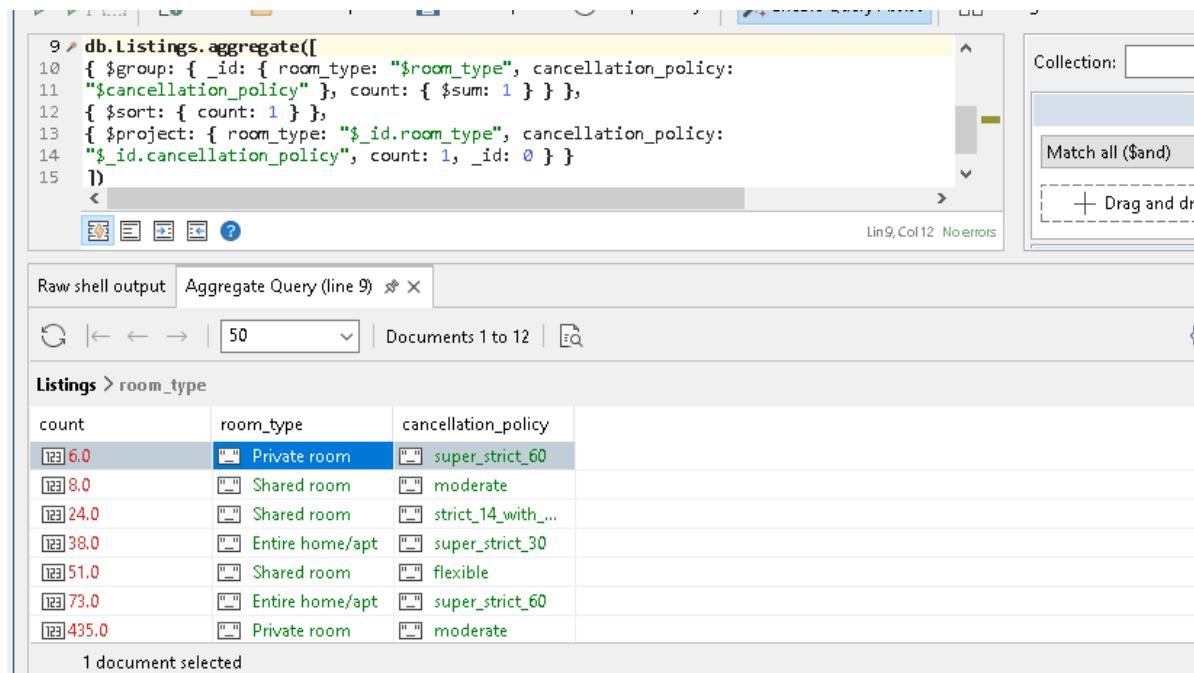
```
1 //db.Listings.countDocuments({})
2
3 //db.Listings.countDocuments({property_type: "House"})
4
5 //db.Listings.aggregate([{$group: { _id: "$property_type", count: { $sum: 1 } } }, {$sort: { count: -1 }}])
6
7 db.Listings.countDocuments({"review_scores.review_scores_cleanliness": {$lt:5}})
```

The 'Raw shell output' tab is active, showing the following text:

```
15 // execution of the current command or script
16 // - "cls" clears the raw shell output tab
17 // - Run Shell Queries via Studio 3T for full access to query results
18 // -----
19
20 Current Mongosh Log ID: 66faf0a04c160b13ad072f89
21 Using Mongosh: 2.2.5
22
23 For mongosh info see: https://docs.mongodb.com/mongodb-shell/
24
25 >
26 > HW5
27 authenticated
28 38
29
```

This shows that the majority of listings have ratings above 5 and that cleanliness scores below 5 are very rare.

Q5)



The screenshot shows the MongoDB Studio interface with an aggregate query in the script editor:

```
9 db.Listings.aggregate([
10 { $group: { _id: { room_type: "$room_type", cancellation_policy:
11 "$cancellation_policy" }, count: { $sum: 1 } } },
12 { $sort: { count: 1 } },
13 { $project: { room_type: "$_id.room_type", cancellation_policy:
14 "$_id.cancellation_policy", count: 1, _id: 0 } }
15 ])
```

The 'Raw shell output' tab is active, showing the results of the query in a table format. The table has columns: count, room_type, and cancellation_policy. The results are as follows:

count	room_type	cancellation_policy
6.0	Private room	super_strict_60
8.0	Shared room	moderate
24.0	Shared room	strict_14_with_...
38.0	Entire home/apt	super_strict_30
51.0	Shared room	flexible
73.0	Entire home/apt	super_strict_60
435.0	Private room	moderate

At the bottom, it indicates '1 document selected'.

The least frequent combination is "Private room" with a "super_strict_60" cancellation policy (6 listings), suggesting that strict policies for private rooms are rare. Shared rooms with "moderate" or "strict_14_with_grace_period" policies are also relatively uncommon, with 8 and 24 listings, respectively.

Entire home/apt listings with stricter cancellation policies (e.g., "super_strict_30" or "super_strict_60") occur more often but are still not very common, with 38 and 73 listings, respectively. The most frequent combination is "Entire home/apt" with "strict_14_with_grace_period" (1698 listings), indicating that stricter policies tend to be associated with entire homes.

Q6)

The screenshot shows a MongoDB query interface. The query is as follows:

```
17 db.Listings.aggregate([
18   { $match: { property_type: "Apartment" } },
19   { $group: { _id: { room_type: "$room_type", cancellation_policy:
20     "$cancellation_policy" }, count: { $sum: 1 } } },
21   { $sort: { "_id.cancellation_policy": 1 } },
22   { $project: { room_type: "$_id.room_type", cancellation_policy:
23     "$_id.cancellation_policy", count: 1, _id: 0 } }
24 ])
```

The results are displayed in a table with columns: count, room_type, and cancellation_policy. The table shows the following data:

count	room_type	cancellation_policy
24.0	Shared room	flexible
550.0	Private room	flexible
567.0	Entire home/apt	flexible
631.0	Entire home/apt	moderate
3.0	Shared room	moderate
287.0	Private room	moderate
12.0	Shared room	strict_14_with...

- "Entire home/apt" with "strict_14_with_grace_period" (1112 listings) and "flexible" policies (567 listings) are the most prevalent, suggesting stricter policies are more common for apartments.
- **Private rooms** frequently have "flexible" (550 listings) and "strict_14_with_grace_period" (416 listings) policies, reflecting a mix of flexibility and strictness in these spaces.
- **Shared rooms** have significantly fewer listings across all policies, with "flexible" (24 listings) being the most common option.
- Strict policies like "super_strict_30" and "super_strict_60" are especially rare for both room types, with counts as low as 1 and 3 for private rooms.

Q7

```
27 db.Listings.find(
28   { "address.market": "New York", monthly_price: { $gt: 5000 } },
29   { _id: 1, monthly_price: 1 }
30 ).sort({ monthly_price: -1 })
```

Raw shell output Find Query (line 27) ✕

50 Documents 1 to 10

Listings > monthly_price

_id	monthly_price
846854	17000.0
256328	12000.0
598612	10000.0
1897001	9495.0
6064471	8820.0
640813	8000.0
1146653	7500.0

We observe here that the monthly price exceeds \$5000-6000 dollars for rent in New York

Q8)

```
32 db.Listings.find(
33   {},
34   { name: 1, property_type: 1, cleaning_fee: 1 }
35 ).sort({ cleaning_fee: -1 }).limit(10)
```

Raw shell output Find Query (line 32) ✕

50 Documents 1 to 10

Listings > name

_id	name	property_type	cleaning_fee
13927230	Casa completa...	House	2000.0
5725151	service apart...	Apartment	1200.0
6147746	Stunning Wate...	House	1000.0
28884716	两房一厅,出租...	Serviced apart...	1000.0
15488401	INSTAGRAM H...	Apartment	942.0
20362690	WORLD CLASS...	Apartment	942.0
1176693	BEST REVIEWS*...	Apartment	942.0

1 document selected

Here, the JSON file would have really worked but we see that the cleaning fee for premium apartments and houses exceeds 900 USD and the lowest cleaning fee for top 10 is \$800.

Q9)

```
36
37 db.Listings.aggregate([
38   { $match: { price: { $lt: 100 }, beds: { $lt: 2 } } },
39   { $group: { _id: { room_type: "$room_type", cancellation_policy:
40     "$cancellation_policy" }, count: { $sum: 1 } } },
41   { $project: { room_type: "$_id.room_type", cancellation_policy:
```

Raw shell output | Aggregate Query (line 37) ✖

50 Documents 1 to 10

Listings > room_type

count	room_type	cancellation_policy
123 5.0	Shared room	strict_14_with_...
123 143.0	Entire home/apt	moderate
123 14.0	Shared room	flexible
123 1.0	Entire home/apt	super_strict_30
123 2.0	Shared room	moderate
123 228.0	Private room	moderate
123 383.0	Private room	flexible

1 document selected

Here we observe that the cancellation policy for private and shared rooms has been flexible and moderate. Note that this has been the case for the price to be lower than \$100 and beds to be less than 2.

Q10)

```
db.Listings.aggregate([
  {
    "$group": {
      "_id": {
        "room_type": "$room_type"
      },
      "averagePrice": { "$avg": "$price" },
      "averageReviews": { "$avg": "$number_of_reviews" }
    },
    {
      "$project": {
        _id: 0,
        "room_type": "$_id.room_type",
        "averagePrice": 1,
        "averageReviews": 1
      }
    },
    {
      "$sort": { "averagePrice": -1 }
    }
  ]
)
```

])

Json output:

```
{
  "count" : 5.0,
  "room_type" : "Shared room",
  "cancellation_policy" : "strict_14_with_grace_period"
}
{
  "count" : 143.0,
  "room_type" : "Entire home/apt",
  "cancellation_policy" : "moderate"
}
{
  "count" : 14.0,
  "room_type" : "Shared room",
  "cancellation_policy" : "flexible"
}
{
  "count" : 1.0,
  "room_type" : "Entire home/apt",
  "cancellation_policy" : "super_strict_30"
}
{
  "count" : 2.0,
  "room_type" : "Shared room",
  "cancellation_policy" : "moderate"
}
{
  "count" : 228.0,
  "room_type" : "Private room",
  "cancellation_policy" : "moderate"
}
{
  "count" : 383.0,
  "room_type" : "Private room",
  "cancellation_policy" : "flexible"
}
{
  "count" : 314.0,
  "room_type" : "Private room",
  "cancellation_policy" : "strict_14_with_grace_period"
}
{
  "count" : 95.0,
  "room_type" : "Entire home/apt",
  "cancellation_policy" : "flexible"
}
{
  "count" : 158.0,
  "room_type" : "Entire home/apt",
  "cancellation_policy" : "strict_14_with_grace_period"
}
```

From the data provided, **Entire home/apt** listings tend to favor stricter policies, with **158 listings** having a **"strict_14_with_grace_period"** policy and **143 listings** opting for a **"moderate"** cancellation policy. However, **95 listings** offer a **"flexible"** policy, indicating that while stricter terms are more common, flexibility is still important for a substantial number of entire homes/apartments. Extremely strict policies like **"super_strict_30"** are quite rare, with just **1 listing** for an entire home/apt.

For **Private rooms**, flexibility seems to be the most attractive option, with **383 listings** having a **"flexible"** policy. However, stricter policies are not entirely uncommon, as **314 listings** use **"strict_14_with_grace_period"**, and **228 listings** have a **"moderate"** policy, indicating a well-distributed preference across different policy types for private rooms.

In contrast, **Shared rooms** have far fewer listings overall, with the most popular cancellation policy being **"flexible"** (14 listings). More restrictive policies, such as **"strict_14_with_grace_period"** (5 listings) and **"moderate"** (2 listings), are much less common.

Overall, the data indicates that flexible policies dominate for smaller or shared accommodations, while stricter or moderate cancellation policies are more prevalent in entire homes and apartments.