BC2402: BC2402 Design and Develop Databases

NANYANG TECHNOLOGICAL UNIVERSITY

Final Project Report

Professor:

Ben Choi

Seminar 7, Group 4

Team Members:

Jordan Cheng Zong Jun (Zhuang Zongjun) U2310241H
Nicholas Ng Joon Eng U2310415E
Foo Jek Yen, Alden U2310747E
Xia Jinyin U2240684K
Koh Zhi Liang U2310769E

Content Page

BC2402 Group Project – Task Allocation

S/N	Tasks	Member In-charge
1	*[relational database] SQL development (specific to query 1)*	*Jordan*
2	*[relational database] SQL development (specific to query 2)*	*Zhi Liang*
3	*[relational database] SQL development (specific to query 3)*	*Alden*
4	*[relational database] SQL development (specific to query 4)*	*Nicholas*
5	*[relational database] SQL development (specific to query 5)*	*Jinyin*
6	*[relational database] SQL development (specific to query 6)*	*Jinyin*
7	*[relational database] SQL development (specific to query 7)*	*Nicholas*
8	*[relational database] SQL development (specific to query 8)*	*Alden*
9	*[relational database] SQL development (specific to query 9)*	*Zhi Liang*
10	*[relational database] SQL development (specific to query 10)*	*Jordan*
11	*[nonrelational database] noSQL development (specific to query 1)*	*Jinyin*
12	*[nonrelational database] noSQL development (specific to query 2)*	*Nicholas*
13	*[nonrelational database] noSQL development (specific to query 3)*	*Alden*
14	*[nonrelational database] noSQL development (specific to query 4)*	*Zhi Liang*
15	*[nonrelational database] noSQL development (specific to query 5)*	*Jordan*
16	*[nonrelational database] noSQL development (specific to query 6)*	*Jordan*
17	*[nonrelational database] noSQL development (specific to query 7)*	*Zhi Liang*
18	*[nonrelational database] noSQL development (specific to query 8)*	*Alden*
19	*[nonrelational database] noSQL development (specific to query 9)*	*Nicholas*
20	*[nonrelational database] noSQL development (specific to query10)*	*Jinyin*

Q6 – Correlations between flight hours, length of stay, and various preferences

It is important for airline service providers to predict if given certain variables, would they affect the length of stay, wanting preferred seats or wanting extra baggage. That way, it would be easier to foresee a possible conflict for preferred seats, or for more informed pricing on seating plans and baggage.

SQL Development

In MySQL, I utilized the GROUP BY function to aggregate data by each “sale_channel” and “route” from the passenger booking dataset. These two categories were initially selected to display the results for each channel and route. Following this, I created four new engineered variables to analyze the relationship between flight hours and passenger preferences:

```
AVG(length_of_stay) / AVG(flight_hour) AS Stay_flight,
AVG(wants_extra_baggage) / AVG(flight_hour) AS Baggage_flight,
AVG(wants_preferred_seat) / AVG(flight_hour) AS Seat_flight,
AVG(wants_in_flight_meals) / AVG(flight_hour) AS Meal_flight
```

NoSQL Development

I started off with a pipeline on an aggregate function on the passenger booking collection. The first step, as required by the question itself, was to group the data first by sales_channel and second by route. Thus, using $group : { _id: {sales_channel: “$sales_channel” , route: “$route” }. The second step would be to make use of the grouping aggregate to find the averages as demanded by the question.

```
avgLOS: { $avg: "$length_of_stay" },
avgFH: { $avg: "$flight_hour" },
avgEB: { $avg: "$wants_extra_baggage" },
avgPS: { $avg: "$wants_preferred_seat" },
avgIFM: { $avg: "$wants_in_flight_meals" }
```

The third step would be to calculate the ratios by $addFields because these new variables need to be set.

```
LOStoFH:{
$cond:{
    if: { $ne: ["$avgFH", 0] },
    then: { $divide: ["$avgLOS", "$avgFH"] },
    else: null}
},
EBtoFH:{
$cond:{
    if: { $ne: ["$avgFH", 0] },
    then: { $divide: ["$avgEB", "$avgFH"] },
    else: null}
},
PStoFH: {
$cond:{
    if: { $ne: ["$avgFH", 0] },
    then: { $divide: ["$avgPS", "$avgFH"] },
    else: null}
},
IFMtoFH: {
$cond: {
    if: { $ne: ["$avgFH", 0] },
    then: { $divide: ["$avgIFM", "$avgFH"] },
    else: null}
```

The first ratio in the above script can be understood as:

“LOStoFH, which means ratio of length of stay to Flight Hours, can be calculated by taking $avgLOS divided by $avgFH in all groups provided ($cond) that $avgLOS is non zero as we do not want an operation error.”

The rest of the ratios can then be read the same way as the first ratio.

Lastly, we print the results that we want to see by $project.

```
,
 /*Print*/
 {$project: {
   sales_channel: "$_id.sales_channel",
   route: "$_id.route",
   LOStoFH: 1,
   EBtoFH: 1,
   PStoFH: 1,
   IFMtoFH: 1}
```

Key Insights

Since the answers for both the SQL and NoSQL developments are the same, the insights derived will be the same too. We are determined to find out if there are any relationships between the following pairs:

Length of Stay to Flight Hours

Hypothesis: Passengers who are staying longer at their destination or want to make the most of their weekends or holidays are likely to choose early morning or late-night departures. This allows them to maximize their time at the destination.

Results: We will focus on the top three routes with the highest length of stay to flight hour ratio, as well as the bottom three routes where the ratio is not null. For the top three routes, the average length of stay values are 97, 369, and 28 hours, with corresponding flight hours at 3 AM, 13 PM, and 1 AM. This supports the hypothesis that late-night departures are associated with longer stays to certain extends. In contrast, the bottom three routes have average length of stays of 1 day each, with flight hours at 14 PM and 8 AM. These departures are not associated with long stays, nor do they align with early morning or late-night departures.

sales_channel	route	Stay_flight
Mobile	HANOOL	32.33333333
Mobile	ICNKNO	28.38461538
Internet	KCHXIY	28.00000000

sales_channel	route	Stay_flight
Internet	CTUSBW	NULL
Internet	DACMEL	NULL
Internet	ICNTRZ	NULL
Mobile	KCHXIY	NULL
Mobile	KIXKNO	0.07142857
Mobile	SINWUH	0.07142857
Internet	GOIKUL	0.12000000

Analysis: This further analysis examines the correlation between flight departure times and the length of stay to flight hour ratio by categorizing flights into early morning departures (0-6 AM) and those departing on weekends (Saturday or Sunday).

```
CASE
    WHEN flight_hour BETWEEN 0 AND 6 THEN 'Early Morning'
    ELSE 'Daytime'
END AS flight_time_category,
CASE
    WHEN flight_day IN ('Sat', 'Sun') THEN 'Weekend'
    ELSE 'Weekday'
END AS is_weekend
```

The findings support our hypothesis that the highest length of stay to flight hour ratio occurs for weekend early morning departures.

flight_time_categ...	is_weekend	Stay_flight
Daytime	Weekday	1.62811556
Early Morning	Weekday	5.88872853
Daytime	Weekend	1.37791199
Early Morning	Weekend	6.09215501

Wanting Extra Baggage to Flight Hours

Hypothesis: Passengers traveling in the early morning or late evening are more likely to opt for extra baggage. This is likely because these time slots often suggest connecting flights that involve multiple destinations or longer journeys.

Results: For the top three routes, all flight departures occur at 1 AM, supporting the hypothesis that late-night departures are associated with extra baggage. In contrast, the flight times for the bottom three routes are at 5 PM, 11 AM, and 7:30 AM, which are not considered early morning or late-night departures.

sales_channel	route	Stay_flight	Baggage_flig...
Internet	KCHXIY	28.00000000	1.00000000
Mobile	MELXIY	18.00000000	1.00000000
Mobile	SYDTPE	24.00000000	1.00000000

sales_channel	route	Stay_flight	Baggage_flig...
Internet	CTUSBW	NULL	NULL
Internet	DACMEL	NULL	NULL
Internet	ICNTRZ	NULL	NULL
Mobile	KCHXIY	NULL	NULL
Mobile	AKLPEK	1.82352941	0.00000000
Mobile	CGKTPE	1.54545455	0.00000000
Internet	CTUMEL	4.20000000	0.00000000

Analysis: When we categorize flight hours into early morning (0 - 6 AM) and length of stay into short, medium, and long stays, the correlation further supports the hypothesis.

```
CASE
    WHEN flight_hour BETWEEN 0 AND 6 THEN 'Early Morning'
    ELSE 'Daytime'
END AS flight_time_category,
CASE
    WHEN length_of_stay between 0 and 30 then "Short Stay"
    WHEN length_of_stay between 31 and 90 then "Medium Stay"
    WHEN length_of_stay > 90 then "Long Stay"
END AS Stay_Length,
AVG(wants_extra_baggage) AS baggage
```

flight_time_category	Stay_Length	baggage
Early Morning	Long Stay	0.9385
Daytime	Long Stay	0.8862
Early Morning	Medium Stay	0.8808
Daytime	Medium Stay	0.8707
Daytime	Short Stay	0.7206
Early Morning	Short Stay	0.7140

Wanting Preferred Seat to Flight Hours

Hypothesis: we propose that passengers are more likely to choose their preferred seats for flights departing late at night or early in the morning to ensure a more comfortable experience during potential sleeping hours.

Results: For the top three routes, all flight departures occur between 1 AM and 3 AM, supporting the hypothesis that late-night departures are associated with a higher likelihood of opting for preferred seating. In contrast, the flight times for the bottom three routes are at 7 AM, 8 AM, and 4 PM, indicating a lower preference for preferred seating.

sales_channel	route	Stay_flight	Baggage_flig...	Seat_flight
Mobile	SYDTPE	24.00000000	1.00000000	1.00000000
Mobile	MFMPER	4.00000000	1.00000000	1.00000000
Internet	PERSDK	5.25000000	0.50000000	0.50000000

sales_channel	route	Stay_flight	Baggage_flig...	Seat_flight
Internet	CTUSBW	NULL	NULL	NULL
Internet	DACMEL	NULL	NULL	NULL
Internet	ICNTRZ	NULL	NULL	NULL
Mobile	KCHXIY	NULL	NULL	NULL
Internet	AKLPVG	4.28571429	0.14285714	0.00000000
Internet	AORKTM	8.06666667	0.13333333	0.00000000
Internet	BDOTPE	2.00000000	0.06250000	0.00000000

Analysis: To test our hypothesis that passengers departing late at night or early in the morning are more likely to opt for preferred seats, flight hours were categorized accordingly. The results show that passengers departing between 9 PM and 12 AM have the highest rate of choosing preferred seats and experience the longest flight durations. However, the overall rate is still lower than the preference for extra baggage observed in the previous analysis.

This suggests that, in addition to flight hours, other factors may play a more significant role in influencing their choices.

```
CASE
    WHEN flight_hour BETWEEN 0 AND 6 THEN 'Early Morning'
    WHEN flight_hour BETWEEN 21 AND 24 THEN "Late Night"
    ELSE 'Daytime'
END AS flight_time_category,
AVG(wants_preferred_seat) AS seats,
AVG(flight_duration) AS duration
```

flight_time_category	seats	duration
Late Night	0.3812	7.337376237623755
Daytime	0.3553	6.8876421547669295
Early Morning	0.3426	6.883229407236364

Wanting in-flight meals to Flight Hours

Hypothesis: it is expected that the flight scheduled during meal times (7 - 9, 12 - 14, 18 - 20), will show a higher preference for in-flight meals, as passengers are more likely to opt for in-flight meals if the flight overlaps with breakfast, lunch, or dinner times.

Results: For the top three routes, all flight departures occur between 1 AM and 2:45 AM, which doesn't fall within the hypothesized time slots. Similarly, the flight times for the bottom three routes—7:30 AM, 4 PM, and 1 AM—also contradict our initial hypothesis. It was also found that, although the top two routes show a 100% preference for in-flight meals, the overall rate is low. From the third route onward, it drops to below 40%, indicating a generally low preference for eating on the plane.

sales_channel	route	Stay_flight	Baggage_flig...	Seat_flight	Meal_flight
Mobile	MELXIY	18.00000000	1.00000000	0.00000000	1.00000000
Mobile	MFMPER	4.00000000	1.00000000	1.00000000	1.00000000
Internet	DADOOL	12.18181818	0.36363636	0.09090909	0.36363636

sales_channel	route	Stay_flight	Baggage_flig...	Seat_flight	Meal_flight
Internet	CTUSBW	NULL	NULL	NULL	NULL
Internet	DACMEL	NULL	NULL	NULL	NULL
Internet	ICNTRZ	NULL	NULL	NULL	NULL
Mobile	KCHXIY	NULL	NULL	NULL	NULL
Internet	AORKTM	8.06666667	0.13333333	0.00000000	0.00000000
Internet	BDOTPE	2.00000000	0.06250000	0.00000000	0.00000000
Internet	BWNSYD	21.00000000	0.50000000	0.00000000	0.00000000

Analysis: Guided by the hypothesis, we categorized flight hours into breakfast, lunch, and dinner times. Flights during dinner hours showed the highest preference for in-flight meals, followed by those during daytime meal times. Other factors, such as flight duration, meal price, and personal habits like resting, are also considered influential.

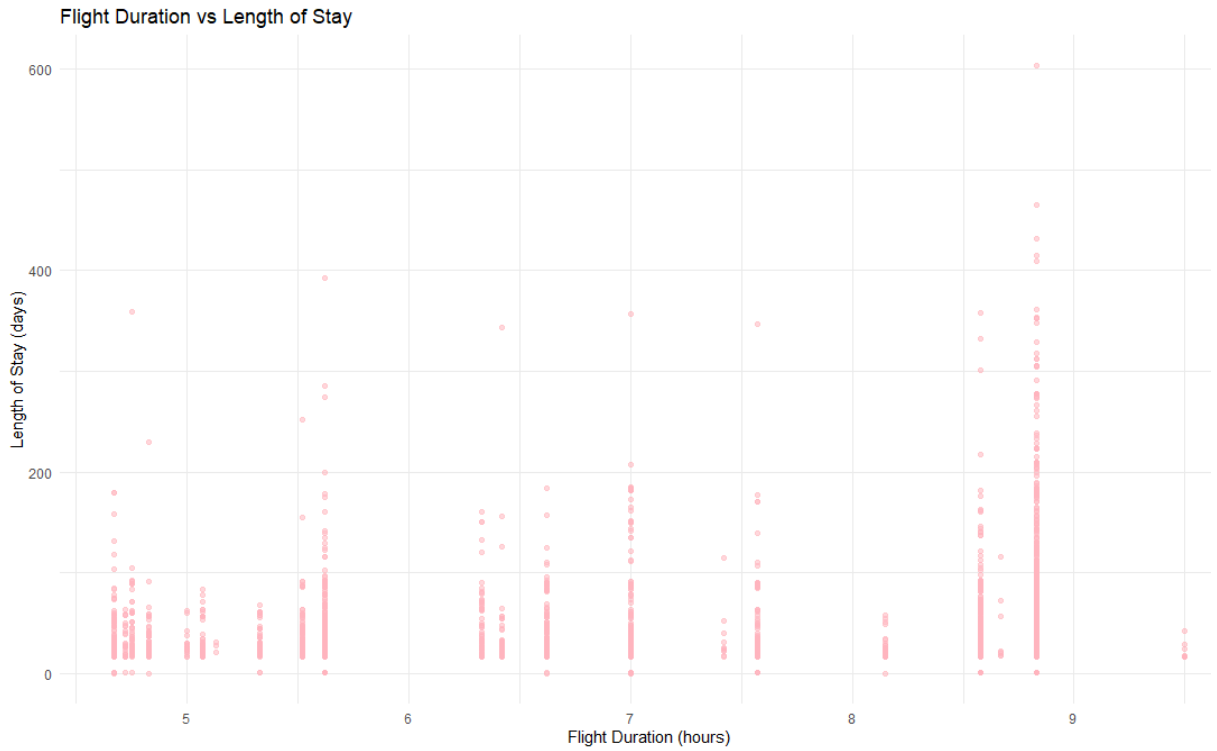
```
CASE
    WHEN flight_hour BETWEEN 7 AND 9 THEN 'Breakfast'
    WHEN flight_hour BETWEEN 12 AND 14 THEN 'Lunch'
    WHEN flight_hour BETWEEN 18 AND 20 THEN 'Dinner'
    ELSE 'Daytime'
```

flight_time_category	meals
Dinner	0.5034
Daytime	0.4619
Lunch	0.4559
Breakfast	0.4452

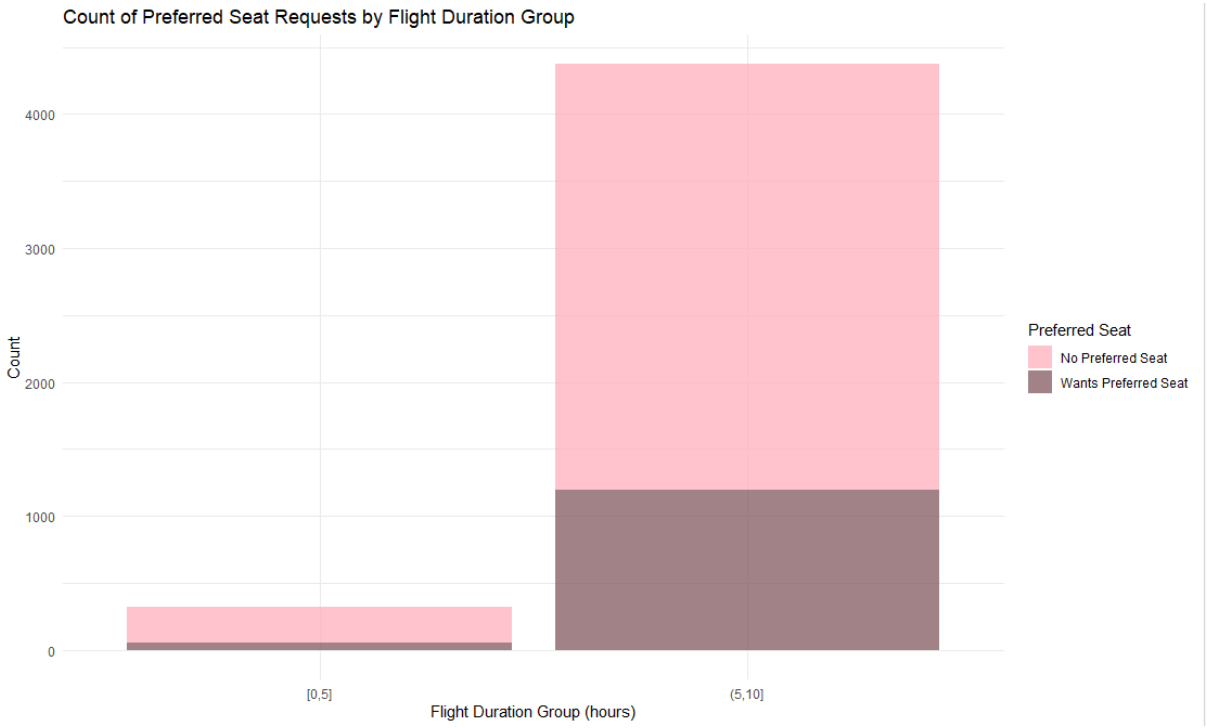
Suggestion

Seeing as to how Flight Hours had minimal relationship with length of stay, wanting extra baggage and wanting preferred seats, we want to propose another variable as a proxy to flight hours which is flight duration.

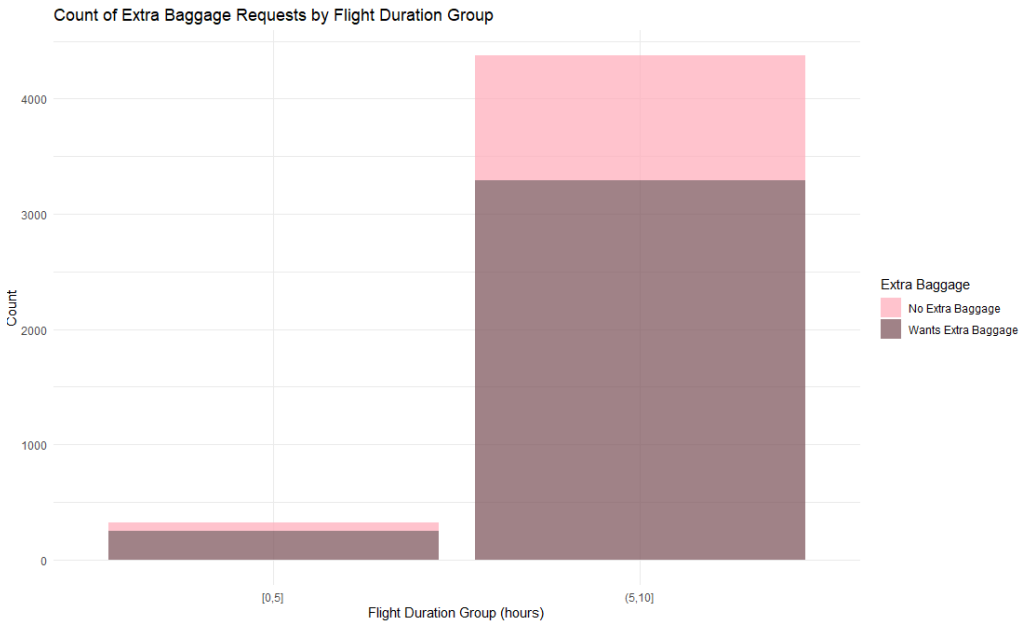
Firstly, flight duration and length of stay had a weak positive relationship where generally the longer the flight duration, the longer the length of stay. (See diagram below) It does make sense that flying over a longer distance meant a further country, and most people would want to stay longer to make the trip worthwhile.



Secondly, flight duration and wanting a preferred seat also have a weak positive relationship where if the flight is more than 5 hours, the likelihood of someone wanting a preferred seat is higher. (See diagram below) The hypothesis makes sense because when the flight is longer, the passenger would want to get as comfortable as possible.



Lastly, flight duration and wanting extra baggage has a stronger positive relationship as compared to the previous two scenarios, when flight duration is more than 5 hours, a significant number of requests would be raised for extra baggage. (See diagram below) This could be due to the general assumption that if flight duration is shorter, the passenger is going to a country that is nearer for a shorter vacation and hence require less baggage.



Implications

By combining our analysis on flight duration's impact on extra baggage, preferred seat and length of stay with the analysis on flight hours' impacts, we will be able to, with a decent level of accuracy, predict the volume of demand for those services. That way, an aircraft at any given time will be able to adequately meet the customers' demand for baggage space, or even before the flight to be able to price the baggage appropriately to maximise pricing strength due to its inelasticity. Furthermore, meals and "good" seats on an aircraft can also be allocated sufficiently to customers who demand them. Ultimately, making use of analytics to enhance both SIA's service and customers' satisfaction.

Q7 – Airline seasonality

Knowing how airline seasonality affects customer perceptions of the overall flight experience is important for airlines to determine key differences in preferences between these two customer segments that airlines face annually to optimise business operations. In the below 2 codes, we explore how average ratings for various aspects of flights differ between these two groups of customers.

SQL Development:

Firstly, I started off with a query to derive the “Month” of the flight that the review was referring to by using substring() on the “MonthFlown” column, which has a uniform string structure of MMM-YY, and other defining columns of each individual review.

```
(SELECT SUBSTRING(MonthFlown, 1, 3) as Month, MonthFlown, Name, ReviewDate, Title
FROM airlines_reviews) as t1
```

Using the above as a nested table, I then joined this back with the original table, using the WHERE clause to join on “MonthFlown”, “Name”, “ReviewDate” and “Title”, and filter the results to the seasonal period using IN to compare it to a list of months. As I require the average of all the ratings grouped by Airline and Class, I SELECT the required information and GROUP BY “Airline” and “Class”.

```
(SELECT Class, Airline, AVG(SeatComfort) as AvgSeatComfort_Seasonal, AVG(FoodnBeverages) as AvgFoodnBeverages_Seasonal,
        AVG(InflightEntertainment) as AvgInflightEntertainment_Seasonal, AVG(ValueForMoney) as AvgValueForMoney_Seasonal
        AVG(OverallRating) as AvgOverallRating_Seasonal
FROM
    (SELECT SUBSTRING(MonthFlown, 1, 3) as Month, MonthFlown, Name, ReviewDate, Title
    FROM airlines_reviews) as t1
JOIN
    airlines_reviews as t2
WHERE t1.MonthFlown = t2.MonthFlown AND t1.Name = t2.Name
      AND t1.ReviewDate = t2.ReviewDate AND t1.Title = t2.Title
      AND t1.Month IN ("Jun", "Jul", "Aug", "Sep")
GROUP BY Class, Airline
) as t5
```

I would repeat this exact same code, but now I would add a NOT to the IN clause in order to obtain the NonSeasonal average ratings. I have also changed the naming conventions to

differentiate between the Seasonal and NonSeasonal average ratings.

```
(SELECT Class, Airline, AVG(SeatComfort) as AvgSeatComfort_NonSeasonal, AVG(FoodnBeverages) as AvgFoodnBeverages_NonSeasonal,
        AVG(InflightEntertainment) as AvgInflightEntertainment_NonSeasonal, AVG(ValueForMoney) as AvgValueForMoney_NonSeasonal,
        AVG(OverallRating) as AvgOverallRating_NonSeasonal
FROM
    (SELECT SUBSTRING(MonthFlown, 1, 3) as Month, MonthFlown, Name, ReviewDate, Title
    FROM airlines_reviews) as t3
    JOIN
    airlines_reviews as t4
WHERE t3.MonthFlown = t4.MonthFlown AND t3.Name = t4.Name AND t3.ReviewDate = t4.ReviewDate AND t3.Title = t4.Title
      AND t3.Month NOT IN ("Jun", "Jul", "Aug", "Sep")
GROUP BY Class, Airline
) as t6
```

Finally, for presentation, as I would like to project all the results in a single table, I would JOIN the two tables on “Airline” and “Class”, and SELECT all relevant information. This would allow us to arrive at our final code.

The results have been exported to Excel for easy viewing and will be discussed later in our Insights section.

noSQL Development:

Firstly, I would create an .aggregate() pipeline to find out the required average ratings for “SeatComfort”, “FoodnBeverages”, “InflightEntertainment”, “ValueforMoney” and “Overall rating”. I started off by creating a new field and extracting the month from “MonthFlown”, which has the format of “MMM-YY” via the $substr function. Moving on, I classified the extracted months by categorising it according to the specified seasonality in the question.

```
db.airline_reviews.aggregate([
    {
        $addFields: {
            month: {
                $substr: ["$MonthFlown", 0, 3]
            }
        }
    },
    {
        $addFields: {
            seasonality: {
                $cond: [
                    { $in: ["$month", ["Jun", "Jul", "Aug", "Sep"]] },
                    "Seasonal",
                    "Non-Seasonal"
                ]
            }
        }
    },
```

Next, I group by Airline, Class and Seasonality to calculate the averages

```
{
    $group: {
        _id: { Airline: "$Airline", Class: "$Class", Seasonality: "$seasonality" },
        AvgSeatComfort: { $avg: "$SeatComfort" },
        AvgFoodnBeverages: { $avg: "$FoodnBeverages" },
        AvgInflightEntertainment: { $avg: "$InflightEntertainment" },
        AvgValueForMoney: { $avg: "$ValueForMoney" },
        AvgOverallRating: { $avg: "$OverallRating" }
    }
},
```

Next, I regrouped the data by Airline and Class, and collected the seasonal ratings into an array. The $push operator creates an array (seasonalRatings) that contains each of the ratings for both "Seasonal" and "Non-Seasonal" periods.

```
{
    $group: {
        _id: { Airline: "$_id.Airline", Class: "$_id.Class" },
        seasonalRatings: {
            $push: {
                seasonality: "$_id.Seasonality",
                AvgSeatComfort: "$AvgSeatComfort",
                AvgFoodnBeverages: "$AvgFoodnBeverages",
                AvgInflightEntertainment: "$AvgInflightEntertainment",
                AvgValueForMoney: "$AvgValueForMoney",
                AvgOverallRating: "$AvgOverallRating"
            }
        }
    }
},
```

Lastly is to reshape the data by creating two separate fields for the seasonal and non-seasonal ratings. Using the $filter function, it filters the data according to their

seasonality and the $arrayElemAt operator ensures the first and only match gets returned. Finally we will display the output using the $project function.

```
{
    $project: {
        _id: 0,
        Airline: "$_id.Airline",
        Class: "$_id.Class",
        Seasonal: {
            $arrayElemAt: [
                { $filter: {
                    input: "$seasonalRatings",
                    as: "rating",
                    cond: { $eq: ["$$rating.seasonality", "Seasonal"] }
                }},
                0
            ]
        },
        NonSeasonal: {
            $arrayElemAt: [
                { $filter: {
                    input: "$seasonalRatings",
                    as: "rating",
                    cond: { $eq: ["$$rating.seasonality", "Non-Seasonal"] }
                }},
                0
            ]
        }
    }
}
```

Insights:

Seasonal

Class	Airline	AvgSeatComfort_Seasonal	AvgFoodnBeverages_Seasonal	AvgInflightEntertainment_Seasonal	AvgValueForMoney_Seasonal	AvgOverallRating_Seasonal
Business Class	Singapore Airlines	3.5714	3.4048	3.75	3.8929	7.5595
Economy Class	Singapore Airlines	3.6242	3.497	3.8	3.497	6.6364
Premium Economy	Singapore Airlines	3.5385	3.6154	4.2308	2.8846	6
First Class	Singapore Airlines	4	4	2.5	5	10
Economy Class	Qatar Airways	3.9153	3.9266	4.1977	3.4802	6.6158
Business Class	Qatar Airways	4.0704	3.8592	4.1268	3.7183	7.5211
First Class	Qatar Airways	4	2.5	4.5	3.25	6.75
Premium Economy	Qatar Airways	4	5	2	5	10

NonSeasonal

Class	Airline	AvgSeatComfort_NonSeasonal	AvgFoodnBeverages_NonSeasonal	AvgInflightEntertainment_NonSeasonal	AvgValueForMoney_NonSeasonal	AvgOverallRating_NonSeasonal
Business Class	Singapore Airlines	3.6531	3.6122	4.0136	3.5986	7.0612
Economy Class	Singapore Airlines	3.7324	3.5523	3.91	3.4234	6.3747
Premium Economy	Singapore Airlines	3.614	3.4386	3.807	2.9298	5.6667
First Class	Singapore Airlines	3.7273	3.5455	4.4545	4.5455	9.2727
Economy Class	Qatar Airways	3.9841	4.2024	4.2659	3.8492	7.2738
Business Class	Qatar Airways	4.1139	4.0696	4.2911	3.7532	7.2911
First Class	Qatar Airways	4	4	4.1111	3.8889	9.3333
Premium Economy	Qatar Airways	3.75	4	3.5	5	10

Percentage Differences [(Seasonal-NonSeasonal)/NonSeasonal]

Class	Airline	AvgSeatComfort_Change	AvgFoodnBeverages_Change	AvgInflightEntertainment_Change	AvgValueForMoney_Change	AvgOverallRating_Change
Business Class	Singapore Airlines	-2.24%	-5.74%	-6.57%	8.18%	7.06%
Economy Class	Singapore Airlines	-2.90%	-1.56%	-2.81%	2.15%	4.11%
Premium Economy	Singapore Airlines	-2.09%	5.14%	11.13%	-1.54%	5.88%
First Class	Singapore Airlines	7.32%	12.82%	-43.88%	10.00%	7.84%
Economy Class	Qatar Airways	-1.73%	-6.56%	-1.60%	-9.59%	-9.05%
Business Class	Qatar Airways	-1.06%	-5.17%	-3.83%	-0.93%	3.15%
First Class	Qatar Airways	0.00%	-37.50%	9.46%	-16.43%	-27.68%
Premium Economy	Qatar Airways	6.67%	25.00%	-42.86%	0.00%	0.00%

As our SQL and noSQL results are the same, we will use the same results to derive our insights, and have placed them into an excel sheet as shown above. We have then

calculated the percentage difference between the Seasonal ratings and the non-Seasonal Ratings, with the formula used being [(Seasonal - NonSeasonal)/Seasonal].

Based on the percentage differences, some interesting findings that could be derived. Firstly, Business and Economy class for Singapore Airlines saw a drop in average ratings for seat comfort, food, beverages and inflight entertainment. However, the rating regarding value for money has improved, which aligns with the increase in overall rating as well. Therefore, this could explain that consumers view value for money as a more important factor for their travel experience since there could be many promotion packages during seasonal periods. Therefore, pricing strategy is something that airline companies could prioritise and look into.

When comparing Singapore Airlines to Qatar Airways's Business and Economy class, although both airline companies seem to be performing negatively, Singapore Airlines fared better in retaining overall rating scores. This could explain that there might be other factors that were more impactful in determining the airline companies' ratings that were not discussed in this report.

Inflight entertainment seems to take the most hit in ratings for both Singapore Airlines First Class (43.86%) and Qatar Airways Premium Economy Class (42.86%). This would be a key factor for airline companies to look out for as well.

Most importantly, there seems to be a lot more downward trend in ratings for both airline companies during the Seasonal period. This could indicate that these companies may have a harder time accommodating the larger increase in passenger volumes during these busy periods. The possibility of these airlines experiencing overstretched resources and staff is highly possible as well, as seen from the major drop in ratings such as food, beverages and inflight entertainment.

Overall, Singapore Airlines tend to perform more consistently as compared to Qatar Airways despite the drop in ratings for other categories. As such, Singapore Airlines and Qatar

Airways should investigate and analyse further in-depth into key customer expectations for seasonal travel and improve on weaker areas while at the same time continue to leverage on their own strengths.

Q8 – The common complaints and top 5 common issues for each Airline and TypeofTraveller

Understanding the common complaints and identifying the top 5 issues for each airline and type of traveller is crucial to improve airline performance and increase customer satisfaction. The analysis allows airlines to gain valuable insight into areas that are underperforming and require improvements. The airlines can also use this data to identify recurring issues and spot trends.

SQL Development:

I started off by analysing the ratings that travellers gave at the end of their flights for Singapore Airlines and Qatar Airways. Specifically I counted the total number of ratings that were less than 3 and split them based on the AIrline and Type of Traveller. This will allow the Airlines to identify weak service areas and implement improvement strategies for each specific type of travellers.

This is done by first counting the total number of reviews for each traveller type of each airline, this would allow the analysis to show the proportion of low ratings. Next, sum(case when ___ < 3 then 1 else 0 end) would count the total number of low ratings for each traveller type.

The results indicate that Qatar Airways has a pricing issue as its Value For Money rating has the highest number for all traveller types out of all other ratings with 40/130 couple leisure travellers rating it less than 3.

Singapore Airlines on the other hand have a large number of low ratings in both Food and Beverages as well as Value For Money, with 74/248 Couple Leisure travellers rating the Food and Beverages less than 3.

This analysis would allow the airlines to better allocate resources towards the areas where low ratings are often found in order to increase overall customer satisfaction.

```
/* Q8 */
/*Count number of low ratings for each rating*/
select Airline, TypeofTraveller,
    count(*) as TotalCount,
    sum(case when SeatComfort < 3 then 1 else 0 end) as SeatComfort_LessThan3,
    sum(case when StaffService < 3 then 1 else 0 end) as StaffService_LessThan3,
    sum(case when FoodnBeverages < 3 then 1 else 0 end) as FoodnBeverages_LessThan3,
    sum(case when InflightEntertainment < 3 then 1 else 0 end) as InflightEntertainment_LessThan3,
    sum(case when ValueForMoney < 3 then 1 else 0 end) as ValueForMoney_LessThan3
from airlines_reviews
group by Airline, TypeofTraveller
order by Airline, TypeofTraveller;
```

Airline	TypeofTraveller	TotalCount	SeatComfort_LessThan3	StaffService_LessThan3	FoodnBeverages_LessThan3	InflightEntertainment_LessThan3	ValueForMoney
Qatar Airways	Business	112	15	18	16	8	28
Qatar Airways	Couple Leisure	130	27	13	17	15	40
Qatar Airways	Family Leisure	112	21	14	25	3	25
Qatar Airways	Solo Leisure	322	33	28	41	24	57
Singapore Airlines	Business	143	26	17	31	17	45
Singapore Airlines	Couple Leisure	248	58	52	74	31	79
Singapore Airlines	Family Leisure	182	31	31	42	27	51
Singapore Airlines	Solo Leisure	330	63	54	82	43	73

MySQL Code and Findings

```
db.airlines_reviews.aggregate([
    {
        $project: {
            Airline: 1,
            TypeofTraveller: 1,
            TotalCount: 1,
            SeatComfort: 1,
            StaffService: 1,
            FoodnBeverages: 1,
            InflightEntertainment: 1,
            ValueForMoney: 1,
        }
    },
    {
        $project: {
            Airline: 1,
            TypeofTraveller: 1,
            TotalCount: 1,
            SeatComfort_LessThan3: { $cond: [{ $lt: ["$SeatComfort", 3] }, 1, 0] },
            StaffService_LessThan3: { $cond: [{ $lt: ["$StaffService", 3] }, 1, 0] },
            FoodnBeverages_LessThan3: { $cond: [{ $lt: ["$FoodnBeverages", 3] }, 1, 0] },
            InflightEntertainment_LessThan3: { $cond: [{ $lt: ["$InflightEntertainment", 3] }, 1, 0] },
            ValueForMoney_LessThan3: { $cond: [{ $lt: ["$ValueForMoney", 3] }, 1, 0] }
        }
    },
    {
        $group: {
            _id: { Airline: "$Airline", TypeofTraveller: "$TypeofTraveller" },
            TotalRatings: { $sum: 1 },
            SeatComfort_LessThan3: { $sum: "$SeatComfort_LessThan3" },
            StaffService_LessThan3: { $sum: "$StaffService_LessThan3" },
            FoodnBeverages_LessThan3: { $sum: "$FoodnBeverages_LessThan3" },
            InflightEntertainment_LessThan3: { $sum: "$InflightEntertainment_LessThan3" },
            ValueForMoney_LessThan3: { $sum: "$ValueForMoney_LessThan3" }
        }
    },
    {
        $sort: { "_id.Airline": 1, "_id.TypeofTraveller": 1 }
    }
```

```
/* 1 */
{
    "_id" : {
        "Airline" : "Qatar Airways",
        "TypeofTraveller" : "Business"
    },
    "TotalRatings" : 112,
    "SeatComfort_LessThan3" : 15,
    "StaffService_LessThan3" : 18,
    "FoodnBeverages_LessThan3" : 16,
    "InflightEntertainment_LessThan3" : 8,
    "ValueForMoney_LessThan3" : 28
},

/* 2 */
{
    "_id" : {
        "Airline" : "Qatar Airways",
        "TypeofTraveller" : "Couple Leisure"
    },
    "TotalRatings" : 130,
    "SeatComfort_LessThan3" : 27,
    "StaffService_LessThan3" : 13,
    "FoodnBeverages_LessThan3" : 17,
    "InflightEntertainment_LessThan3" : 15,
    "ValueForMoney_LessThan3" : 40
},
```

NoSQL code and Findings

Next, to identify the top 5 common issues and complaints for each type of traveller in each airline, I first started by creating a list of common words associated with various complaints. This is done by creating a temporary table and inserting the keywords into said table.

```
'expensive', 'seat', 'legroom', 'quality', 'food', 'drinks', 'beverage',
'staff', 'employee', 'service', 'crew', 'comfort', 'meal', 'entertainment',
'shows', 'price', 'toilet', 'noise', 'dirty', 'clean', 'hygiene'
```

Words Used for MySQL and NoSQL

SQL Development:

In SQL, I used 2 different methods to count the number of words associated with common issues in order to identify the top 5 common issues.

For the 1st method, the keywords and their corresponding SOUNDEX codes were used for matching with words in the reviews. An additional table was also created to extract individual words from Reviews by splitting them.

Individual words from the Reviews where the recommended column had values of ‘no; were extracted and their SOUNDEX words were computed to be matched with the keywords identified.

```
/* Top 5 common complains using wordcount of wordcount of specific words using soundex*/
/*Calculate soundex for each word in the reviews*/
with SplitReviews as (
    select
        r.Airline,
        r.TypeofTraveller,
        regexp_substr(r.Reviews, '[[:alpha:]]+', 1, n.n) as word,
        soundex(regexp_substr(r.Reviews, '[[:alpha:]]+', 1, n.n)) as word_soundex
    from airlines_reviews r
    join numbers n on char_length(r.Reviews) - char_length(replace(r.Reviews, ' ', '')) >= n.n - 1
    /*Only take bad reviews to prevent counting from good reviews*/
    where r.Recommended = 'no'
),
```

For each airline and traveller type, the total mentions of each keyword are calculated regardless of spelling variations.

```
/*Count the instances of each keyword for each airline and traveller type*/
KeywordInstances as (
    select
        r.Airline,
        r.TypeofTraveller,
        k.keyword,
        count(*) as KeywordCount
    from SplitReviews r
    join category_keywords k
        on r.word_soundex = k.soundex_code
        where r.word not in ('the', 'and', 'is', 'of', 'to', 'in', 'are', 'were')
    group by r.Airline, r.TypeofTraveller, k.keyword
),
```

The keywords are then ranked by frequency within each Airline and traveller type in order to identify the top 5 most common issues for each.

```
/*Rank the keywords by frequency for each airline and traveller type*/
RankedKeywords as (
    select
        Airline,
        TypeofTraveller,
        keyword,
        KeywordCount,
        row_number() over (partition by Airline, TypeofTraveller order by KeywordCount desc) as ranked
    from KeywordInstances
)

select Airline, TypeofTraveller, keyword as Complains, KeywordCount
from RankedKeywords
where ranked <= 5
order by Airline, TypeofTraveller, KeywordCount desc;
```

Airline	TypeofTraveller	Complains	KeywordCount
Qatar Airways	Business	seat	19
Qatar Airways	Business	service	11
Qatar Airways	Business	food	8
Qatar Airways	Business	drinks	5
Qatar Airways	Business	crew	5
Qatar Airways	Couple Leisure	seat	16
Qatar Airways	Couple Leisure	service	11
Qatar Airways	Couple Leisure	food	11
Qatar Airways	Couple Leisure	staff	8
Qatar Airways	Couple Leisure	toilet	6
Qatar Airways	Family Leisure	staff	16
Qatar Airways	Family Leisure	food	13
Qatar Airways	Family Leisure	seat	12
Qatar Airways	Family Leisure	service	11
Qatar Airways	Family Leisure	toilet	6
Qatar Airways	Solo Leisure	seat	51
Qatar Airways	Solo Leisure	staff	23
Qatar Airways	Solo Leisure	service	23
Qatar Airways	Solo Leisure	toilet	15
Qatar Airways	Solo Leisure	crew	13
Singapore Airl...	Business	service	20
Singapore Airl...	Business	seat	19
Singapore Airl...	Business	crew	16
Singapore Airl...	Business	meal	12
Singapore Airl...	Business	drinks	12
Singapore Airl...	Couple Leisure	seat	54
Singapore Airl...	Couple Leisure	food	26
Singapore Airl...	Couple Leisure	meal	20
Singapore Airl...	Couple Leisure	crew	18
Singapore Airl...	Couple Leisure	service	14
Singapore Airl...	Family Leisure	seat	29
Singapore Airl...	Family Leisure	service	18
Singapore Airl...	Family Leisure	food	15
Singapore Airl...	Family Leisure	crew	13
Singapore Airl...	Family Leisure	staff	10
Singapore Airl...	Solo Leisure	seat	41
Singapore Airl...	Solo Leisure	service	33
Singapore Airl...	Solo Leisure	food	21
Singapore Airl...	Solo Leisure	crew	16
Singapore Airl...	Solo Leisure	staff	13

MySQL 1st Method Output

The second method that was used in SQL did the same counting of keywords and filtering but without the use of SOUNDEX. Although this prevents any spelling errors from being counted, it also prevents other words that sound similar but have different meanings from being counted, providing a more reliable output.

Airline	TypeofTraveller	Complains	KeywordCount
Qatar Airways	Business	seat	17
Qatar Airways	Business	service	15
Qatar Airways	Business	food	10
Qatar Airways	Business	comfort	7
Qatar Airways	Business	staff	6
Qatar Airways	Couple Leisure	seat	27
Qatar Airways	Couple Leisure	service	21
Qatar Airways	Couple Leisure	food	13
Qatar Airways	Couple Leisure	staff	12
Qatar Airways	Couple Leisure	comfort	7
Qatar Airways	Family Leisure	staff	20
Qatar Airways	Family Leisure	service	20
Qatar Airways	Family Leisure	food	14
Qatar Airways	Family Leisure	seat	13
Qatar Airways	Family Leisure	meal	5
Qatar Airways	Solo Leisure	seat	37
Qatar Airways	Solo Leisure	service	35
Qatar Airways	Solo Leisure	staff	29
Qatar Airways	Solo Leisure	food	25
Qatar Airways	Solo Leisure	meal	17
Singapore Airl...	Business	service	28
Singapore Airl...	Business	seat	24
Singapore Airl...	Business	food	19
Singapore Airl...	Business	crew	16
Singapore Airl...	Business	meal	14
Singapore Airl...	Couple Leisure	seat	56
Singapore Airl...	Couple Leisure	food	45
Singapore Airl...	Couple Leisure	service	38
Singapore Airl...	Couple Leisure	meal	35
Singapore Airl...	Couple Leisure	staff	23
Singapore Airl...	Family Leisure	seat	30
Singapore Airl...	Family Leisure	food	23
Singapore Airl...	Family Leisure	service	22
Singapore Airl...	Family Leisure	meal	18
Singapore Airl...	Family Leisure	crew	17
Singapore Airl...	Solo Leisure	service	53
Singapore Airl...	Solo Leisure	seat	53
Singapore Airl...	Solo Leisure	food	37
Singapore Airl...	Solo Leisure	crew	30
Singapore Airl...	Solo Leisure	meal	27

MySQL 2nd Method Output

NoSQL Development:

In NoSQL, I started by filtering the reviews to only look at the reviews who had ‘no’ in the recommended column.

The reviews were then split into words and unwind to process each word individually. Each word was then trimmed and converted into lower case using $tolower.

```
db.airlines_reviews.aggregate([
    {
    |    $match: { Recommended: "no" }  // Only consider bad reviews
    },
    {
    |    $addFields: {
    |    |    splitWords: {
    |    |    |    $split: [{ $ifNull: ["$Reviews", ""] }, " "]  // Split reviews into words
    |    |    }
    |    }
    },
    { $unwind: "$splitWords" },  // Unwind the split words array to process each word
    {
    |    $addFields: {
    |    |    word: { $toLower: { $trim: { input: "$splitWords" } } }  // Trim and convert to lowercas
    |    }
    },
```

The words were then used to match the list of predefined keywords that was used in SQL as well to identify the most common words used in bad reviews. Each match was then counted and only the top 5 issues were displayed.

```
    {
    |    $match: {
    |    |    word: { $in: [  // Match the predefined keywords
    |    |    |    'expensive', 'seat', 'legroom', 'quality', 'food', 'drinks', 'beverage',
    |    |    |    'staff', 'employee', 'service', 'crew', 'comfort', 'meal', 'entertainment',
    |    |    |    'shows', 'price', 'toilet', 'noise', 'dirty', 'clean', 'hygiene'
    |    |    ]}
    |    }
    },
    {
    |    $group: {
    |    |    _id: { Airline: "$Airline", TypeofTraveller: "$TypeofTraveller", Keyword: "$word" },
    |    |    KeywordCount: { $sum: 1 }
    |    }
    },
    {
    |    $sort: { "_id.Airline": 1, "_id.TypeofTraveller": 1, "KeywordCount": -1 }
    },
```

```
    {
        $group: {
            _id: { Airline: "$_id.Airline", TypeofTraveller: "$_id.TypeofTraveller" },
            top_keywords: {
                $push: {
                    keyword: "$_id.Keyword",
                    count: "$KeywordCount"
                }
            }
        }
    },
    {
        $project: {
            Airline: "$_id.Airline",
            TypeofTraveller: "$_id.TypeofTraveller",
            top5_issues: { $slice: ["$top_keywords", 5] }  // Only take the top 5 issues
        }
    },
    {
        $sort: { Airline: 1, TypeofTraveller: 1 }  // Sort by Airline and TypeofTraveller
    }
]);
```

```
/* 1 */
{
    "Airline" : "Qatar Airways",
    "TypeofTraveller" : "Business",
    "top5_issues" : [
        {
            "keyword" : "seat",
            "count" : 19
        },
        {
            "keyword" : "food",
            "count" : 12
        },
        {
            "keyword" : "service",
            "count" : 12
        },
        {
            "keyword" : "staff",
            "count" : 8
        },
        {
            "keyword" : "crew",
            "count" : 4
        }
    ]
}
```

NoSQL Sample Output

Insights From Results:

By analysing the low ratings of each traveller type from each airline and their proportions, we can identify which are the areas that the airlines need to work on.

‘

For example, Qatar Airways had a pricing issue as it has the highest number of low Value For Money ratings for all types of travellers. In addition, it might also have to work on quality of Food and Beverages as it has the second highest number for Solo Leisure and Family Leisure.

Singapore Airlines also faces the same issue with pricing as Business, Couple Leisure and Family Leisure have the highest number of low ratings for Value For Money. In addition, 82 out of 330 Solo Leisure travellers gave a low rating for Food and Beverages. It can be argued that 24.85% low ratings may indicate an emergency that needs to be solved.

These findings would allow airlines to introduce specific targeting strategies to tackle these issues. Competitive pricing can be introduced and food quality assessments could be conducted to improve their products and services. With limited resources available, these findings would allow for better allocation of resources and achieve higher customer satisfaction with the same amount of resources available.

In analysing the top 5 common issues for each type of traveller in each airline, there were multiple trends identified.

FIrstly, the keyword “seat” has appeared in every single top 5 list, with it being the number 1 in 8 of 12 lists for Qatar Airways (6 from MySQL and 2 from NoSQL) and 6 of 12 lists for Singapore Airlines (5 from MySQL 1 from NoSQL). While it is clear that issues related to seats are the most common issue identified in negative reviews, it does not indicate that the seats of both airlines are of poor quality. It is worth noting that negative reviews makeup 34.2% for Qatar Airways and 29.3% of Singapore Airlines.

Airline	Recommended	Count
Singapore Airlines	yes	594
Singapore Airlines	no	309
Qatar Airways	no	198
Qatar Airways	yes	478

The Findings in the lists will provide a reference for both Airlines to look at when determining the priority of areas to allocate resources to. As each type of travellers will have different priorities in regards to areas such as services, food and beverages and seats, it is extremely important to address these needs based on priority due to limited resources.

For example, Business travellers tend to focus on seat, services and food for both Airlines which indicates that the Airlines should prioritise these areas for Business Travellers. Other types of travellers for Qatar Airways have a high frequency of “staff” mentioned in negative reviews while those of Singapore Airlines display the same priority as that of Business travellers which might indicate that Qatar Airways needs to put more focus on Staff Training.

.

In conclusion, these analyses and findings can help to understand customers’ priority and tailor improvements and responses based on patterns identified across the various types of travellers. Some future plans that can be made are to increase allocation of resources to seats and constant monitoring of negative reviews is essential to improve overall customer satisfaction.

Q9 – Singapore Airlines Performance

It is extremely important for SIA to look at passenger preferences and complaints over time as this allows them to understand and assess how they are performing, as well as provide more details so as to provide direction for SIA to improve their operations.

To look at any systematic changes in passenger preferences and complaints pre- and post-COVID specific to Singapore Airlines, we would like to look specifically at all the ratings and “Recommended” columns in the [airlines_reviews] table.

For simplicity’s sake, considering a lack of consensus on the start and end of the Covid-19 pandemic, we will take dates before 2019-12-31 as pre-Covid, and dates after 2022-12-31 as post-Covid. This is because lockdowns around the world started in the year 2020, and Covid measures eased in early 2023.

Let’s first talk about our thought processes in developing our code to develop these insights in the SQL and the noSQL (MongoDB) environment.

SQL Development:

Number of Yes/No Recommendations during Post and Pre-Covid :

The MonthFlown column is used to categorise data into two distinct periods:

- Pre-COVID (flights before the year 2020): Months ending in '19' or earlier.
- Post-COVID (flights from the year 2023 onward): Months ending in '23' or later.

The CASE statement is used to assign the period labels (Pre-COVID or Post-COVID) based on the last two digits of MonthFlown. I then calculate the total number of reviews and the number of "Recommended" reviews for each period. Next, the data is grouped by their periods and whether a customer recommends it or not. Lastly, the output esults are ordered by period and recommendation status.

```sql
SELECT
    CASE
        WHEN RIGHT(MonthFlown, 2) <= '19' THEN 'Pre-COVID'
        WHEN RIGHT(MonthFlown, 2) >= '23' THEN 'Post-COVID'
        ELSE NULL
    END AS Period, Recommended, COUNT(*) AS totalcount
FROM airlines_reviews
WHERE Airline = 'Singapore Airlines' AND (RIGHT(MonthFlown, 2) <= '19' OR RIGHT(MonthFlown, 2) >= '23')
GROUP BY Period, Recommended
ORDER BY Period, Recommended;
```

Period	Recommended	totalcount
Post-COVID	no	58
Post-COVID	yes	77
Pre-COVID	no	163
Pre-COVID	yes	441

Average rating for each category :

The second query evaluates the average ratings for various service categories:

Overall rating,Seat comfort,Staff service,Food and beverages,In-flight entertainment and Value for money.

Finally, I grouped the data by period to compare the changes over time and sorted the output results by period.

```sql
SELECT
    CASE
        WHEN RIGHT(MonthFlown, 2) <= '19' THEN 'Pre-COVID'
        WHEN RIGHT(MonthFlown, 2) >= '23' THEN 'Post-COVID'
        ELSE NULL
    END AS Period,
    COUNT(*) AS totalcount,
    AVG(OverallRating) AS average_overall_rating,
    AVG(SeatComfort) AS average_seat_comfort,
    AVG(StaffService) AS average_staff_service,
    AVG(FoodnBeverages) AS average_food_and_beverages,
    AVG(InflightEntertainment) AS average_inflight_entertainment,
    AVG(ValueForMoney) AS average_value_for_money
FROM airlines_reviews
WHERE Airline = 'Singapore Airlines' AND (RIGHT(MonthFlown, 2) <= '19' OR RIGHT(MonthFlown, 2) >= '23')
GROUP BY Period
ORDER BY Period;
```

Period	totalcount	average_overall_rating	average_seat_comfort	average_staff_service	average_food_and_beverages	average_inflight_entertainment	average_value_for_money
Post-COVID	135	6.2444	3.5333	4.0444	3.4889	3.9259	3.3333
Pre-COVID	604	7.0646	3.7401	3.9768	3.6854	3.9205	3.6689

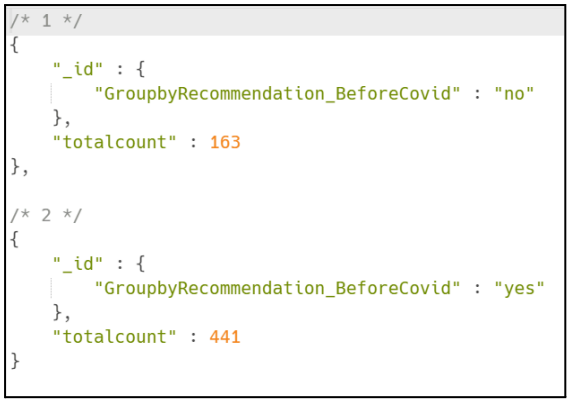
noSQL Development:

Firstly, I would like to obtain the number of “yes” recommendations and “no” recommendations for pre-Covid and post-Covid. I would do this by constructing an .aggregate() pipeline.

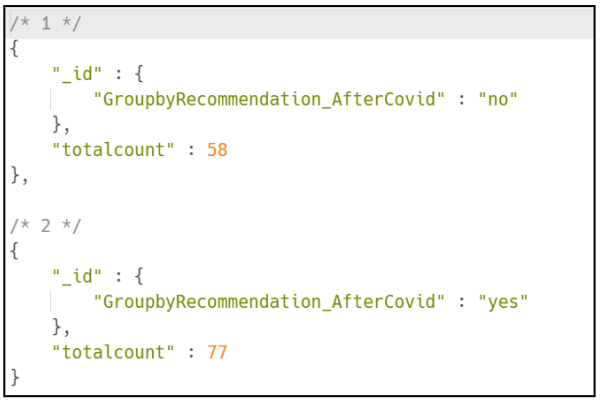
```
db.airlines_reviews.aggregate([
    {$addFields: {Year: {$substr: [ "$MonthFlown", 4, 2 ]}}
    {$match: {"Year": {$lte: "19"} }
    {$match: {"Airline": "Singapore Airlines"}}
    {$group: { _id: {GroupbyRecommendation_BeforeCovid: "$Recommended"}, totalcount: {$sum: 1}}}
]) ;
```

```
db.airlines_reviews.aggregate([
    {$addFields: {Year: {$substr: [ "$MonthFlown", 4, 2 ]}}
    {$match: {"Year": {$gt: "22"}}}
    {$match: {"Airline": "Singapore Airlines"}}
    {$group: { _id: {GroupbyRecommendation_AfterCovid: "$Recommended"}, totalcount: {$sum: 1}}}
]);
```

I started off by first doing $addFields to add “Year”, being derived from obtaining a substring from “MonthFlown”, which has a standard format of “MMM-YY”, via the $substr function. Afterwards, I used $match to filter down the documents based on the year to determine if it is a pre-Covid review or a post-Covid review by using {"Year": {$lte: "19"}} for pre-Covid and {"Year": {$gt: "22"}} for post-Covid. I would then do another $match to filter only the reviews from “Singapore Airlines” using {"Airline": "Singapore Airlines"}. Finally, to count the number of data points of “yes” and “no”, I would use $group to group by “$Recommended” and use a {$sum: 1} to count. Repeating this code 2 times for pre-Covid and post-Covid, allows us to arrive at the following results:

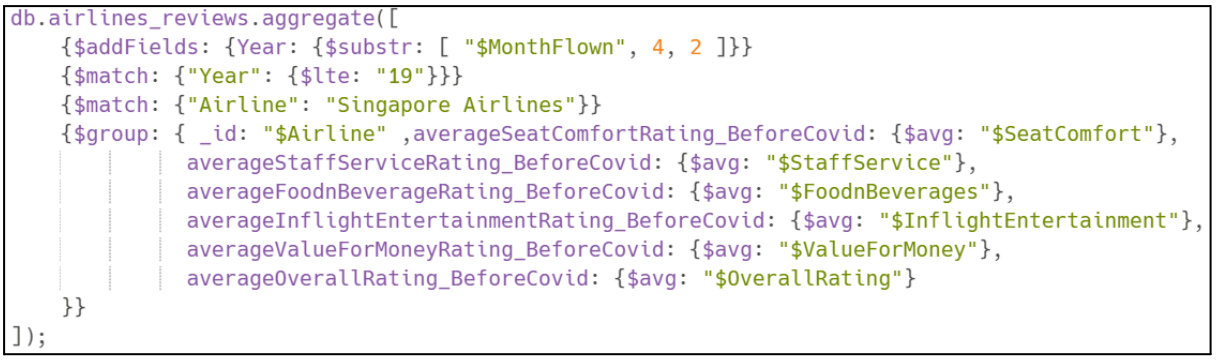


```
/* 1 */
{
    "_id" : {
        "GroupbyRecommendation_BeforeCovid" : "no"
    },
    "totalcount" : 163
},

/* 2 */
{
    "_id" : {
        "GroupbyRecommendation_BeforeCovid" : "yes"
    },
    "totalcount" : 441
}
```

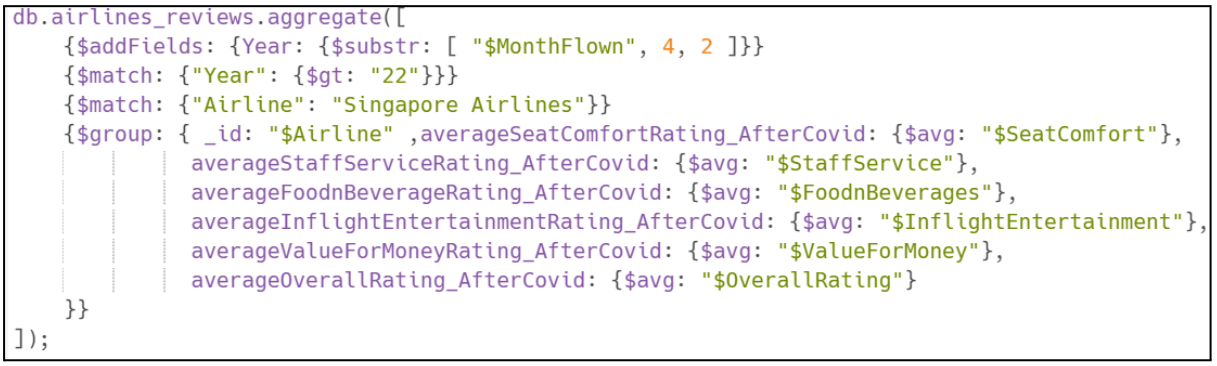


```
/* 1 */
{
    "_id" : {
        "GroupbyRecommendation_AfterCovid" : "no"
    },
    "totalcount" : 58
},

/* 2 */
{
    "_id" : {
        "GroupbyRecommendation_AfterCovid" : "yes"
    },
    "totalcount" : 77
}
```

Next, I would like to obtain the average overall rating for pre-Covid and post-Covid. I would do this by constructing another .aggregate() pipeline.



```
db.airlines_reviews.aggregate([
    {$addFields: {Year: {$substr: [ "$MonthFlown", 4, 2 ]}}
    {$match: {"Year": {$lte: "19"}}}
    {$match: {"Airline": "Singapore Airlines"}}
    {$group: { _id: "$Airline" ,averageSeatComfortRating_BeforeCovid: {$avg: "$SeatComfort"},
               averageStaffServiceRating_BeforeCovid: {$avg: "$StaffService"},
               averageFoodnBeverageRating_BeforeCovid: {$avg: "$FoodnBeverages"},
               averageInflightEntertainmentRating_BeforeCovid: {$avg: "$InflightEntertainment"},
               averageValueForMoneyRating_BeforeCovid: {$avg: "$ValueForMoney"},
               averageOverallRating_BeforeCovid: {$avg: "$OverallRating"}
    }}
]);
```



```
db.airlines_reviews.aggregate([
    {$addFields: {Year: {$substr: [ "$MonthFlown", 4, 2 ]}}
    {$match: {"Year": {$gt: "22"}}}
    {$match: {"Airline": "Singapore Airlines"}}
    {$group: { _id: "$Airline" ,averageSeatComfortRating_AfterCovid: {$avg: "$SeatComfort"},
               averageStaffServiceRating_AfterCovid: {$avg: "$StaffService"},
               averageFoodnBeverageRating_AfterCovid: {$avg: "$FoodnBeverages"},
               averageInflightEntertainmentRating_AfterCovid: {$avg: "$InflightEntertainment"},
               averageValueForMoneyRating_AfterCovid: {$avg: "$ValueForMoney"},
               averageOverallRating_AfterCovid: {$avg: "$OverallRating"}
    }}
]);
```

Similar to counting recommendations previously, I first used $addFields to add “Year”, the same way as explained above. The next parts of the pipeline are the exact same as the one from above. Using $match to trim to pre-Covid or post-Covid, and using another $match to only take the documents from “Singapore Airlines”. Lastly, I used a $group with $avg to obtain all average ratings pre-Covid and post-Covid. Repeating this code 2 times for pre-Covid and post-Covid, allows us to arrive at the following results:

```
{
    "_id" : "Singapore Airlines",
    "averageSeatComfortRating_BeforeCovid" : 3.7400662251655628,
    "averageStaffServiceRating_BeforeCovid" : 3.9768211920529803,
    "averageFoodnBeverageRating_BeforeCovid" : 3.685430463576159,
    "averageInflightEntertainmentRating_BeforeCovid" : 3.9205298013245033,
    "averageValueForMoneyRating_BeforeCovid" : 3.6688741721854305,
    "averageOverallRating_BeforeCovid" : 7.064569536423841
}
```

```
{
    "_id" : "Singapore Airlines",
    "averageSeatComfortRating_AfterCovid" : 3.533333333333333,
    "averageStaffServiceRating_AfterCovid" : 4.044444444444444,
    "averageFoodnBeverageRating_AfterCovid" : 3.488888888888889,
    "averageInflightEntertainmentRating_AfterCovid" : 3.925925925925926,
    "averageValueForMoneyRating_AfterCovid" : 3.3333333333333335,
    "averageOverallRating_AfterCovid" : 6.2444444444444445
}
```


Insights From Results:

Using our results, we are able to calculate the proportions of positive recommendations by dividing the number of “yes” recommendations by the total number of recommendations

Placing all our results into an excel file and we obtain the following results:

Recommendation Results

Covid Status	"Yes"	"No"	Proportion of positive recommendations ["Yes"/("Yes" + "No")]
Pre-Covid	441	163	73.01%
Post-Covid	77	58	57.04%
		Difference:	-15.98%

Ratings Results

Covid Status	Seat Comfort	Staff Service	FoodnBeverage	Inflight Entertainment	Value For Money	Overall
Pre-Covid	3.74	3.98	3.69	3.92	3.67	7.06
Post-Covid	3.53	4.04	3.49	3.93	3.33	6.24
% Change:	-5.61%	1.51%	-5.42%	0.26%	-9.26%	-11.61%

From our results, it actually seems that the passenger satisfaction of Singapore Airlines has actually dropped post-Covid from pre-Covid levels. This is because as per our findings, two

metrics for passenger satisfaction, proportion of positive recommendations and Overall ratings, have fallen by 15.98% and 11.61% when comparing pre-Covid reviews to post-Covid reviews. This means that from an overall standpoint, it can be said that there is a drop in the quality of the customer experience for SIA.

Zooming into the specifics, from our table, we can see that there are some variations in the changes in the various ratings. In fact, contrary to the general trend, ratings for Staff Service and Inflight Entertainment have increased marginally, by 1.51% and 0.26% respectively. However, since these increases are rather marginal, we can attribute this to natural fluctuations and not a true indication of improvements in standards. On the other hand, all other ratings of Seat Comfort, Food and Beverage, and Value for Money have decreased by 5.61%, 5.42% and 9.26% respectively. These changes are rather high and are less likely to be due to natural fluctuations and should be addressed.

That being said, while these metrics have indeed decreased, it is important to contextualise and understand these decreases. Our group believes that following SIA commitments to high standards of elevating the customer experience, coupled with many improvements during and after the covid period, such as improved personalised options at booking (*Singapore Airlines Drives Friction Out of the Customer Experience*, n.d.), increased smartphone integration into inflight entertainment system and their recent “Book the Cook” initiative (*Book the Cook | Singapore Airlines*, n.d.), this decrease should not be attributed to an absolute decrease in the quality of SIA. Rather, it shows an increased customer standard that results in a relative decrease in perceived quality of the various indicators which can be attributed to other factors like an increasing demand for low-cost carriers (Guang-Xi, 2024), which essentially perform a similar function but at a much lower cost.

Overall, however, these two metrics are strong indicators of passenger satisfaction that cannot be denied, even if it is subjective in nature, and shows that Singapore Airlines has

seen a decline in perceived customer satisfaction levels. If they decide to work on this, our group believes they should focus on the 3 key ratings that have fallen, Seat Comfort, Food and Beverage, and Value for Money. This could be done by methods such as investing in things like increased legroom, further customer directed R&D to improve food quality, and improving operational efficiency to cut airfare costs.

This means that interestingly, SIA group's strong performance in the financial year 2023 and 2024 cannot be attributed to the improvement of their passenger satisfaction, which in fact would negatively affect SIA's performance.

As such, to understand what other factors contributed to their strong performance, we looked at the SIA 2023/2024 Annual Report to find out further reasons for the large increase in profits.

Essentially, SIA attributes the increase in profits to increased passenger number, mostly due to the North Asia Markets, and increases in cargo demand.

To quote the annual report on the reasons for its stellar performance: "SIA and Scoot carried a combined total of 36.4 million passengers in FY2023/24, up 37.6% from the previous year. The higher passenger numbers were buoyed by the full reopening of borders in the key North Asia markets such as China, Hong Kong SAR, Japan, and Taiwan." and "Cargo demand saw an uptick towards the end of the financial year, driven by robust e-commerce demand, resilient and growing segments such as perishable goods and concerts, and a shift to air freight by shippers due to concerns over security in the Red Sea region" (SIA, 2023).

For the increase in passenger numbers in North Asia, this indeed tracks with what we can find about the market, where according to the International Air Transport Association (IATA), "Asia-Pacific is expected to be responsible for half of the world's RPK growth in 2024 driven

largely by recovering domestic markets in China, Japan, and Australia." (IATA, 2023), which allows SIA, which is based in Singapore, a global hub for aviation, to definitely capitalise on this to increase passenger numbers, as seen from the annual report, where Singapore Airlines passenger numbers have increased from 18,155 thousand to 23,741 thousand (SIA, 2023).

Singapore Airlines					
Passengers carried (thousand)	23,741	18,155	+	30.8	
Revenue passenger-km (million)	109,942.9	91,025.2	+	20.8	
Available seat-km (million)	126,240.5	106,099.3	+	19.0	
Passenger load factor (%)	87.1	85.8	+	1.3	points
Passenger yield (cents/pkm)	12.1	12.7	-	4.7	
Revenue per available seat-km (cents/ask)	10.6	10.9	-	2.8	
Passenger unit cost (cents/ask)	9.0	9.8	-	8.2	
Passenger unit cost ex-fuel (cents/ask)	5.8	6.0	-	3.3	
Passenger breakeven load factor (%)	74.4	77.2	-	2.8	points

The increase in cargo demand also tracks with what we can find about the market, according to IATA's General Director, Willie Walsh, who said that "The air cargo business continues to benefit from growth in global trade, booming e-commerce and capacity constraints on maritime shipping. With the peak season still to come, it is shaping to be a very strong year for air cargo" (*Strong Air Cargo Demand Growth Extends Into July*, 2024). As such, SIA group, who has business in the air cargo industry, definitely benefited from this growth to increase profits, which is reflected in their annual report, having an increased cargo and mail carried from 923 million kg to 952.4 million kg (SIA, 2023).

Group Airlines (Cargo)					
Cargo and mail carried (million kg)	952.4	923.0	+	3.2	
Cargo load (million tonne-km)	5,347.9	5,260.8	+	1.7	
Gross capacity (million tonne-km)	9,804.8	9,165.4	+	7.0	
Cargo load factor (%)	54.5	57.4	-	2.9	points
Cargo yield (cents/ltk)	39.6	68.5	-	42.2	
Cargo unit cost (cents/ctk)	21.1	24.9	-	15.3	
Cargo breakeven load factor (%)	53.3	36.4	+	16.9	points

Finally, something that is perhaps not mentioned is that SIA is in fact a group of legal entities, which includes Scoot, a low-cost carrier. This means that despite the decrease in customer satisfaction of Singapore Airlines, this does not affect Scoot, and in fact, coupled with the increasing demand for low-cost carriers as mentioned previously, Scoot's revenue would increase as well. This is reflected in the annual report as well, where Scoot has also

increased its passenger count from 8,331 thousand to 12,702 thousand as well, overall increasing the revenue for SIA group (SIA, 2023).

Scoot					
Passengers carried (thousand)	12,702	8,331	+	52.5	
Revenue passenger-km (million)	33,946.7	22,602.9	+	50.2	
Available seat-km (million)	37,227.4	26,932.6	+	38.2	
Passenger load factor (%)	91.2	83.9	+	7.3	points
Passenger yield (cents/pkm)	6.9	8.2	-	15.9	
Revenue per available seat-km (cents/ask)	6.3	6.9	-	8.7	
Passenger unit cost (cents/ask)	6.2	6.5	-	4.6	
Passenger unit cost ex-fuel (cents/ask)	4.2	4.4	-	4.5	
Breakeven load factor (%)	89.9	79.3	+	10.6	points

All in all, although customer satisfaction may perhaps have negatively affected SIA group's performance, the combination of other factors of increased passenger demand especially in North Asia, increased cargo demand, and a diversified portfolio in low-cost carrier Scoot contributed to the strong economic performance of SIA group in recent years.

Q10 – Designing a chatbot to handle exceptional circumstances

Exceptional circumstances refer to situations where SIA faces unexpected troubles with delays and accidents or bad customer service. In the case of the article given, SQ321 faced an accident and SIA had to pay compensation, and that would mean the customer support service would be loaded with work. Having a chatbot on the side as a proxy to deliver apologetic messages to grievances, and simple instructions to compensation orders will be an operational relief to the overloaded customer support service.

SQL Development

From the introduction, our chatbot must be able to:

1. Apologise with context
2. Provide guidance to customer support service

From the given response column in the customer_support data, we have drawn the following inspirations: (See Appendix A)

1. Acknowledgement at the front and Pleasantries in the end
2. Guidelines, contact information and service hours must be provided
3. Context is important in personalising the message with a focus in an apologetic tone

Proposed new table structure:

Title: The title of the feedback input by the user

Name: Full name of the user

ReviewDate: The date of the Review

Reviews: Full comment of the user

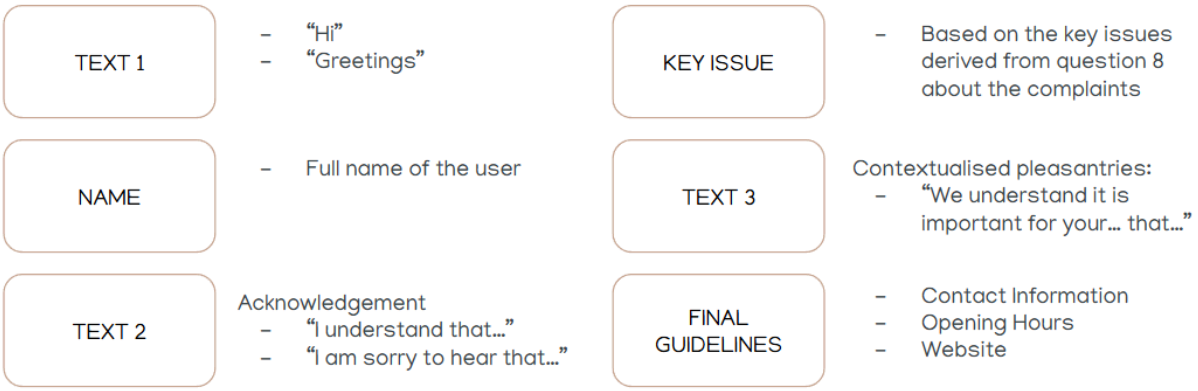
KeyIssues: The category of complaint (i.e. Seat, Service, Comfort, etc)

Response: The message generated by the chatbot using other columns

```
# Make a new table
CREATE TABLE `table` (
    title VARCHAR(255),
    name VARCHAR(255),
    reviewdate DATE,
    Reviews VARCHAR(255),
    key_issues TEXT,
    responses VARCHAR(255)
);
```

```
INSERT INTO `table` (title, name, reviewdate, airline, reviews)
SELECT title, name, STR_TO_DATE(reviewdate, '%d/%m/%Y'), reviews
FROM airlines_reviews;
```

Proposed Function structure:



The function will be generating the response for the chatbot, which is going to mainly contain the CONCAT function of the above items: TEXT 1 + NAME + TEXT 2 + KEY ISSUE + TEXT 3 + FINAL GUIDELINES

```
DELIMITER //

CREATE FUNCTION response_generation(
    customer_name VARCHAR(255),
    key_issue VARCHAR(255),
    customer_support_hours VARCHAR(255),
    customer_support_phone VARCHAR(255),
    website_url VARCHAR(255)
) RETURNS TEXT
    NOT DETERMINISTIC
    READS SQL DATA
```

TEXT 1: a random selection from the list of possible greetings

- Hi
- Dear
- Good day

```
BEGIN
    RETURN CONCAT(
        -- TEXT 1: Random Greeting
        CASE
            WHEN RAND() < 0.33 THEN 'Hi '
            WHEN RAND() BETWEEN 0.33 AND 0.66 THEN 'Dear '
            ELSE 'Good day '
        END,
```

NAME: referenced the NAME column of the table

```
    customer_name, ', ',
```

TEXT 2: a random selection from the list of possible acknowledgement

- I understand that you have experienced…
- I am incredibly sorry to hear that you went through…
- We would like to apologise for the inconvenience caused as you have suffered…

```
-- TEXT 2: Random Acknowledgment
CASE
    WHEN RAND() < 0.33 THEN 'I understand that you have experienced '
    WHEN RAND() BETWEEN 0.33 AND 0.66 THEN 'I am incredibly sorry to hear that you went through '
    ELSE 'We would like to apologise for the inconvenience caused as you have suffered '
END,
```

KEY ISSUE: From Q8, we were able to sieve out the main complaints (category of complaints) by using soundex function and with the results, we can apply the same methodology on potential complaints (from the review column).

Therefore, according to the key issue sieved from the reviews, we would retrieve the corresponding text.*

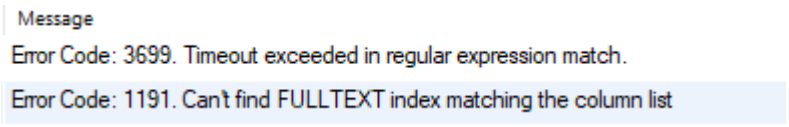
Seat/Comfort: “an uncomfortable journey”

Service/Staff/Crew: “poor quality of customer service”

Food/Meal: “unsatisfactory food and dining experience”

```
-- Conditions for Key_issues based on Q8 findings
CASE
    WHEN key_issue = 'Seat' OR key_issue = 'Comfort' THEN 'an uncomfortable journey'
    WHEN key_issue = 'Service' OR key_issue = 'Staff' OR key_issue = 'Crew' THEN 'poor quality of customer service'
    WHEN key_issue = 'Food' OR key_issue = 'Meal' THEN 'unsatisfactory food and dining experience'
    ELSE 'an unspecified issue.'
END, '. ',
```

*We understand the limitations of our approach, however, if we were to use variables and FULLTEXT capabilities on MySQL, it would be too complex to carry out. We have faced the following errors while using the FULLTEXT approach.



TEXT 3: according to the key issue, we would then follow up with pleasantries that are contextualised with an apologetic tone.

Seat/Comfort: “We understand that it is important to have a comfortable journey, and it is our duty to ensure that all customers have a pleasant experience flying with SIA.”

Service/Staff/Crew: “SIA takes pride in our excellent customer service and we will not condone any dissatisfactory customer service on our part.”

Food/Meal: “We are deeply apologetic to know that the food served was not up to standard and we will give feedback to the relevant parties.”

```
CASE
    WHEN key_issue = 'Seat' OR key_issue = 'Comfort' THEN
        'We understand that it is important to have a comfortable journey, and it is our duty to ensure that all customers have a pleasant experience flying with SIA.'
    WHEN key_issue = 'Service' OR key_issue = 'Staff' OR key_issue = 'Crew' THEN
        'SIA takes pride in our excellent customer service and we will not condone any dissatisfactory customer service on our part.'
    WHEN key_issue = 'Food' OR key_issue = 'Meal' THEN
        'We are deeply apologetic to know that the food served was not up to standard and we will give feedback to the relevant parties.'
    ELSE
        'We are currently reviewing your feedback and will address your concerns promptly.'
END, ' ',
```

FINAL GUIDELINES: to offer the final advice as to where the customer can go to for any claims or compensations.

- Customer Support Hours
- Customer Support Phone Number
- Website URL

```
-- FINAL GUIDELINES
'If you wish to further your claim and seek compensation, our team is available during ',
customer_support_hours,
'. You can reach us at ',
customer_support_phone,
' or via Live Chat on our ',
website_url,
'. We value your satisfaction and are here to assist you.'
```

<u>Example Outputs of chatbot’s response:</u>

‘Dear Alison Soetantyo, I understand that you have experienced unsatisfactory food and dining experience. We are deeply apologetic to know that the food served was not up to standard and we will give feedback to the relevant parties. If you wish to further your claim and seek compensation, our team is available during 9AM - 6PM. You can reach us at +XX XXXX XXXX or via Live Chat on our EXAMPLEWEBSITE.COM. We value your satisfaction and are here to assist you.’

'Dear Robert Watson, I am incredibly sorry to hear that you went through an uncomfortable journey. We understand that it is important to have a comfortable journey, and it is our duty to ensure that all customers have a pleasant experience flying with SIA. If you wish to further your claim and seek compensation, our team is available during 9AM - 6PM. You can reach us at +XX XXXX XXXX or via Live Chat on our EXAMPLEWEBSITE.COM. We value your satisfaction and are here to assist you.'

'Hi S Holger, We would like to apologise for the inconvenience caused as you have suffered poor quality of customer service. SIA takes pride in our excellent customer service and we will not condone any dissatisfactory customer service on our part. If you wish to further your claim and seek compensation, our team is available during 9AM - 6PM. You can reach us at +XX XXXX XXXX or via Live Chat on our EXAMPLEWEBSITE.COM. We value your satisfaction and are here to assist you.'

NoSQL Development

In NoSQL, we will design our approach to address two main concerns:

- To what extent should Singapore Airlines compensate passengers?
- How can a customer service chatbot be designed to respond to exceptional circumstances?

For the first concern, we propose examining four key categories:

- For one or more ratings below the average (3), this could signal an issue worthy of compensation.
- For the overall rating falling below the average (5), this could signal an issue worthy of compensation.

- Certain words or phrases in the review comments (such as “discomfort”, “disappointing”, “poor”, “delayed”, “cancelled”, “damaged”) can be flagged to identify issues that might warrant compensation.
- Passengers who choose not to recommend Singapore Airlines.

```
$match: {
  Airline: "Singapore Airlines",
  $or: [
    { "SeatComfort": { $lt: 3 } },
    { "StaffService": { $lt: 3 } },
    { "FoodnBeverages": { $lt: 3 } },
    { "InflightEntertainment": { $lt: 3 } },
    { "ValueForMoney": { $lt: 3 } },
    { "OverallRating": { $lt: 3 } },
    { "Recommended": "no" }
```

It is believed that if passengers experience significant delays, cancellations or reschedules, damaged luggage or facilities, mishandled services, poor food quality, or food contamination, compensations should be issued, whether financially or non-financially.

```
$addFields: {
  issue_severity: {
    $cond: {
      if: { $regexMatch: { input: "$Reviews", regex: /delayed|cancelled|damaged/i } },
      then: "Severe",
      else: {
        $cond: {
          if: { $regexMatch: { input: "$Reviews", regex: /poor|disappointing|discomfort/i } },
          then: "Minor",
          else: "None"
```

However, it is understood that review comments may overlook many other factors. We cannot fully capture every passenger’s opinion with simplified categories or summarized words.

With this in mind, the NoSQL queries have been developed accordingly:

```
$addFields: {
  compensation_type: {
    $cond: {
      if: { $eq: ["$issue_severity", "Severe"] },
      then: "Financial Compensation: Refund or Travel Credit",
      else: {
        $cond: {
          if: { $eq: ["$issue_severity", "Minor"] },
          then: "Non-Financial Compensation: Voucher or Miles",
          else: "No Compensation"
```

One example of the output is shown below:


```
    "Title" : "seats on this aircraft are dreadful ",
    "Name" : "Robert Watson",
    "Airline" : "Singapore Airlines",
    "Verified" : "TRUE",
    "Reviews" : "Booking an emergency exit seat still meant huge discomfort in a seat
    "SeatComfort" : 5,
    "StaffService" : 3,
    "FoodnBeverages" : 4,
    "InflightEntertainment" : 4,
    "ValueForMoney" : 1,
    "OverallRating" : 3,
    "Recommended" : "no",
    "issue_severity" : "Minor",
    "compensation_type" : "Non-Financial Compensation: Voucher or Miles"
```


For the second concern, NoSQL queries will be designed in a similar manner to the MySQL queries, with the goal of responding to the passenger as follows:

- Greeting & Name: Randomly select a greeting from a list and address the specific passenger by their name.
- Key issue & Contextual Apologetic Response: List all ratings below 3 from the reviews, such as SeatComfort, StaffService, FoodnBeverages, etc and respond to each key issue with a contextual, apologetic tone.
- Final Guidelines: Provide the final advice with customer support details and how to proceed with claims (Service time, contact number, website)

```
    $addFields: {
      response: {
        $concat: [
          // Random Greeting
          { $arrayElemAt: [["Hi", "Dear", "Good day."], { $floor: { $multiply: [{ $rand: {} }, 3] } }] }, " ",
          // Passenger Name
          "$Name", ",\n",

          // Key Issue (seat comfort, staff service, etc.)
          { $cond: {
              if: { $lt: ["$SeatComfort", 3] },
              then: "We understand that it is important to have a comfortable journey, and it is our duty to ensure that all customers have a pleasant experience flying with SIA. ",
              else: ""
            }
          },
          { $cond: {
              if: { $lt: ["$StaffService", 3] },
              then: "SIA takes pride in our excellent customer service and we will not condone any dissatisfactory customer service on our part. ",
              else: ""
            }
          },
          { $cond: {
              if: { $lt: ["$FoodnBeverages", 3] },
              then: "We are deeply apologetic to know that the food served was not up to standard and we will give feedback to the relevant parties. ",
              else: ""
            }
          },
          { $cond: {
              if: { $lt: ["$InflightEntertainment", 3] },
              then: "We strive to offer a variety of entertainment options, and we are sorry that this did not meet your expectations. ",
              else: ""
            }
          },
          { $cond: {
              if: { $lt: ["$ValueForMoney", 3] },
              then: "We understand the importance of value, and we will take your feedback into consideration to improve our offerings. ",
              else: ""
            }
          },

          // Final Guidelines (compensation and support details)
          "\nIf you wish to further your claim and seek compensation, our team is available during 9 AM - 6 PM. You can reach us at +XX XXXX XXXX or via Live Chat on our EXAMPLEWEBSITE.COM."
```

One final output is shown below:


```
        "Title" : "seats on this aircraft are dreadful ",
        "Name" : "Robert Watson",
        "Airline" : "Singapore Airlines",
        "Verified" : "TRUE",
        "Reviews" : "Booking an emergency exit seat still meant huge discomfort in a seat far too narrow and poor padding meaning back ache in 90 minute flight. The seats on this aircraft are dreadful. The headphones and sound on the entertainment system was dreadful.",
        "SeatComfort" : 5,
        "StaffService" : 3,
        "FoodnBeverages" : 4,
        "InflightEntertainment" : 4,
        "ValueForMoney" : 1,
        "OverallRating" : 3,
        "Recommended" : "no",
        "issue_severity" : "Minor",
        "compensation_type" : "Non-Financial Compensation: Voucher or Miles",
        "response" : "Hi Robert Watson,\nWe understand the importance of value, and we will take your feedback into consideration to improve our offerings. \nIf you wish to further your claim and seek compensation, our team is available during 9 AM - 6 PM. You can reach us at +XX XXXX XXXX or via Live Chat on our EXAMPLEWEBSITE.COM."
```

Appendix:

Appendix A

Inspiration from customer support response column:

'I'm sensitive to the fact that you're facing financial difficulties and need to cancel your purchase with the order number {{Order Number}}. We're here to assist you with that, and I appreciate your transparency in sharing your situation with us.

If you encounter any difficulties or have further questions along the way, our team is ready to assist you. Reach out to us during {{Customer Support Hours}} at {{Customer Support Phone Number}} or via Live Chat on our {{Website URL}}. We value your satisfaction and understand that situations can change.'

References:

Book the Cook | Singapore Airlines. (n.d.). https://www.singaporeair.com/en_UK/sg/flying-withus/dining/book-the-cook/

Guang-Xi, J. O. (2024, November 10). 20 years of budget airlines: Once the underdogs of commercial aviation, now a force to reckon with. CNA. https://www.channelnewsasia.com/singapore/budget-airlines-20-years-scoot-jetstar-asia-changi-4733956

IATA. (2023). IATA Annual General Meeting. In iata.org. https://www.iata.org/en/iata-repository/pressroom/presentations/asia-pacific-north-asia-briefing-agm-2024/

SIA. (2023). SIA Annual Report FY2023/24. In https://www.singaporeair.com. https://www.singaporeair.com/content/dam/sia/web-assets/pdfs/about-us/information-for-investors/annual-report/annualreport2324.pdf

Singapore Airlines drives friction out of the customer experience. (n.d.). iTnews Asia. https://www.itnews.asia/news/singapore-airlines-drives-friction-out-of-the-customer-experience-576454

Strong Air Cargo Demand Growth Extends into July. (2024, August 28). https://www.iata.org/en/pressroom/2024-releases/2024-08-28-01/