Airline Passenger Satisfaction - SQL Queries

What percentage of airline customers are satisfied?

Query

```sql
SELECT satisfaction, COUNT(id) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
GROUP BY satisfaction;
```

Result

Satisfaction	Number of passengers	Percentage of passengers
Neutral or Dissatisfied	73452	56.55
Satisfied	56428	43.45

Does it vary by customer type?

1) How many passengers belong to each different customer type? How about their percentage?

Query

```sql
SELECT [customer type],
COUNT([customer type]) AS 'Number of passengers',
ROUND(COUNT(*) * 100 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
GROUP BY [customer type];
```

Result

Customer Type	Number of passengers	Percentage of passengers
First-time	23780	18
Returning	106100	81

2) How many satisfied and dissatisfied customers belong to each category? What about their percentage?

Query

```
SELECT satisfaction, [customer type],
COUNT(id) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
GROUP BY [customer type], satisfaction
ORDER BY satisfaction DESC;
```

Result

Satisfaction	Customer Type	Number of passengers	Percentage of passengers
Satisfied	First-time	5700	4.39
Satisfied	Returning	50728	39.06
Neutral or Dissatisfied	First-time	18080	13.92
Neutral or Dissatisfied	Returning	55372	42.63

What about the type of travel?

Query

```
SELECT satisfaction, [type OF travel],
COUNT(id) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
GROUP BY [type OF travel], satisfaction
ORDER BY satisfaction DESC;
```

Result

Satisfaction	Type of Travel	Number of passengers	Percentage of passengers
Satisfied	Business	52356	40.31
Satisfied	Personal	4072	3.14
Neutral or Dissatisfied	Business	37337	28.75
Neutral or Dissatisfied	Personal	36115	27.81

What is the customer profile for a repeating airline passenger?

1) How many passengers were returning passengers?

Query

```
SELECT COUNT(id) AS 'Number of returning passengers'
FROM passengers
WHERE [Customer Type] = 'Returning';
```

Result

Number of returning passengers
106100

2) Among the returning passengers, how many were men and how many were women? What about their percentage?

Query

```
CREATE TABLE returning_by_gender AS
SELECT gender, COUNT(id) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
WHERE [Customer Type] = 'Returning'
GROUP BY gender;

SELECT * FROM returning_by_gender;

SELECT sum([number OF passengers]) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM returning_by_gender;

INSERT INTO returning_by_gender ('Gender','Number of passengers', 'Percentage of passengers')
VALUES ('All', '106100', '100');

SELECT * FROM returning_by_gender;
```

Result

Gender	Number of passengers	Percentage of passengers
Female	53056	50.01
Male	53044	49.99
All	106100	100

3) How old was the youngest returning passenger? How old was the oldest one? What is the average age of returning passengers?

Query

```
SELECT min(age) AS 'Younger passenger', max(age) AS 'Older passenger',
ROUND(avg(age),2) AS 'Average age of passengers'
FROM passengers
WHERE [Customer Type] = 'Returning';
```

Result

Younger passenger	Older passenger	Average age of passengers
7	85	41.46

4) What are the most common age groups among returning customers?

Query

```
SELECT age, COUNT(*) AS 'Passengers per age group'
FROM passengers
WHERE [Customer Type] = 'Returning'
GROUP BY age
ORDER BY COUNT(*) DESC;
```

Result (Displays only the first 11 rows)

Age	Passengers per age group
39	2946
41	2859
40	2767
44	2748
47	2705
42	2688
43	2683
48	2665
45	2660
46	2629
49	2586

5) Which class did returning customers choose the most?

Query

```
SELECT class, COUNT(ID) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
WHERE [customer type] = 'Returning'
GROUP BY class;
```

Result

Class	Number of passengers	Percentage of passengers
Business	52929	49.89
Economy	44675	42.11
Economy Plus	8496	8.01

6) What was the reason for their travels?

Query

```
SELECT [type OF travel], COUNT(id) AS 'Number of passengers',
ROUND(COUNT(*) * 100.0 / sum(COUNT(*)) Over(),2) AS 'Percentage of passengers'
FROM passengers
WHERE [customer type] = 'Returning'
GROUP BY [type OF travel];
```

Result

Type of Travel	Number of passengers	Percentage of passengers
Business	66114	62.31
Personal	39986	37.69

7) How far have returning passengers traveled?

Query

```
SELECT min([flight distance]) AS 'Shortest flight distance in miles',
max([flight distance]) AS 'Longest flight distance in miles',
ROUND(avg([flight distance]),2) AS 'Average flight distance in miles'
FROM passengers
WHERE [customer type] = 'Returning'
```

Result

Shortest flight distance in miles	Longest flight distance in miles	Average flight distance in miles
31	4983	1297.05

8) How many returning passengers traveled that far

Query

```
SELECT min([flight distance]) AS 'Shortest distance (miles)',
COUNT(*) AS 'Number of passengers'
FROM passengers
WHERE [customer type] = 'Returning'
AND [flight distance] <= (SELECT min([flight distance])
                          FROM passengers WHERE [customer type] = 'Returning');
```

Result

Shortest distance (miles)	Number of passengers
31	9

Query

```
SELECT max([flight distance]) AS 'Longest distance (miles)',
COUNT(*) AS 'Number of passengers'
FROM passengers
WHERE [customer type] = 'Returning'
AND [flight distance] >= (SELECT max([flight distance])
                          FROM passengers WHERE [customer type] = 'Returning');
```

Result

Longest distance (miles)	Number of passengers
4983	16

9) What is the most common distance traveled by returning passengers?

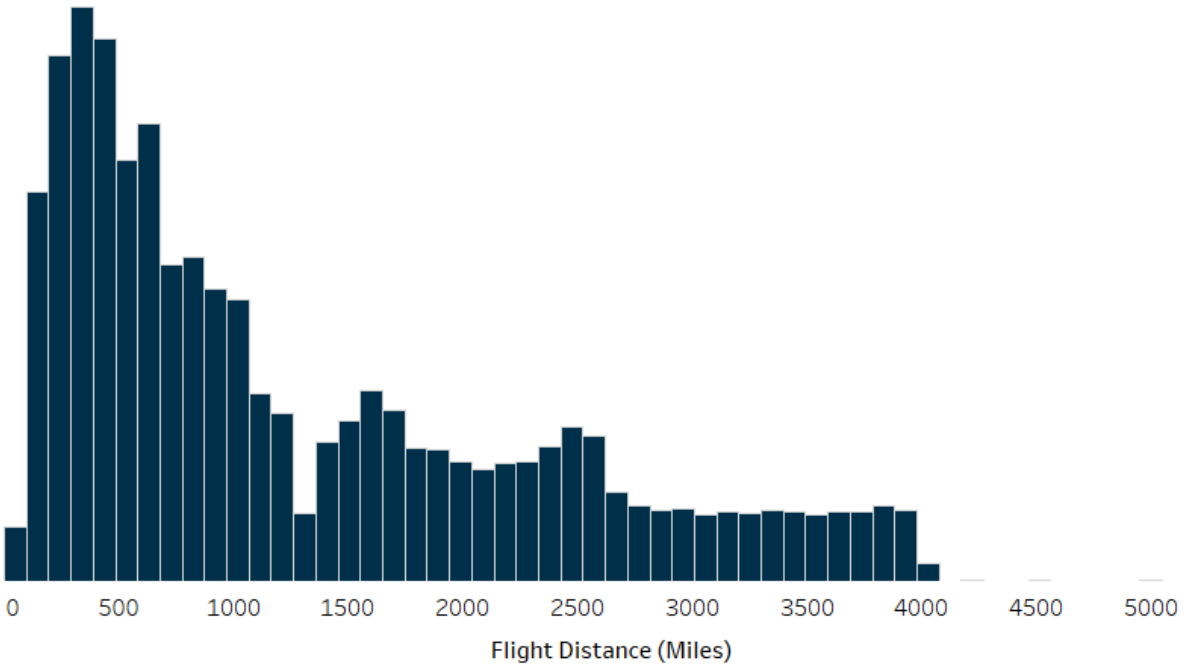
Query

```
SELECT [flight distance], COUNT(*) AS 'Number of passengers'
FROM passengers
WHERE [customer type] = 'Returning'
GROUP BY [flight distance]
ORDER BY COUNT(*) DESC
LIMIT 1;
```

Result

Flight Distance	Number of passengers
337	651

10) How long on average did returning passengers travel the most? (I calculated the median since the distribution on the histogram is skewed to the right).



Query

```
SELECT [flight distance] AS 'Average distance (miles) - Median'
FROM passengers
WHERE [customer type] = 'Returning'
ORDER BY [flight distance]
LIMIT 1
OFFSET (SELECT COUNT(*)
        FROM passengers) / 2
```

Result

Average distance (miles) - Median
1303

Query

```
SELECT COUNT(id) AS 'Number of passengers',
[flight distance] AS 'Average distance (mils) - Median'
FROM passengers
WHERE [customer type] = 'Returning'
AND [flight distance] LIKE
(SELECT [flight distance] AS 'Average distance - Median'
FROM passengers
WHERE [customer type] = 'Returning'
ORDER BY [flight distance]
LIMIT 1
OFFSET (SELECT COUNT(*)
        FROM passengers) / 2)
```

Result

Number of passengers	Average distance (mils) - Median
125	1303

Does flight distance affect customer preferences or flight patterns?

1) Does flight distance affect customer satisfaction?

Query

```
SELECT satisfaction,
COUNT(id) AS 'Number of passengers',
round(COUNT(*) * 100.0 / sum(COUNT(*)) over(),2) AS 'Percentage of passengers',
round(avg([flight distance]),2) AS 'Average distance (miles)'
FROM passengers
GROUP BY satisfaction;
```

Result

Satisfaction	Number of passengers	Percentage of passen...	Average distance (miles)
Neutral or Dissatisfied	73452	56.55	929.72
Satisfied	56428	43.45	1529.54

See the histogram on Tableau for a more detailed view of the distribution.

2) Does flight distance affect class choice?

Query

```
SELECT class,
COUNT(id) AS 'Number of passengers',
round(COUNT(*) * 100.0 / sum(COUNT(*)) over(),2) AS 'Percentage of passengers',
round(avg([flight distance]),2) AS 'Average distance (miles)'
FROM passengers
GROUP BY class;
```

Result

Class	Number of passengers	Shortest distance (miles)	Longest distance (miles)	Average distance (miles)
Business	62160	56	4983	1676.09
Economy	58309	31	4983	744
Economy Plus	9411	31	4983	747.09

See the histogram on Tableau for a more detailed view of the distribution.

3) Does flight distance affect the type of travel?

Query

```
SELECT [type OF travel],
COUNT(id) AS 'Number of passengers',
round(COUNT(*) * 100.0 / sum(COUNT(*)) over(),2) AS 'Percentage of passengers',
round(avg([flight distance]),2) AS 'Average distance (miles)'
FROM passengers
GROUP BY [type OF travel];
```

Result

Type of Travel	Number of passengers	Percentage of passen...	Average distance (miles)
Business	89693	69.06	1368.44
Personal	40187	30.94	792.76

4) Which type of passenger chose longer journeys? Returning customers or first-time customers?

Query

```
SELECT [customer type],
COUNT(id) AS 'Number of passengers',
round(COUNT(*) * 100.0 / sum(COUNT(*)) over(),2) AS 'Percentage of passengers',
round(avg([flight distance]),2) AS 'Average distance (miles)'
FROM passengers
GROUP BY [customer type];
```

Result

Customer Type	Number of passengers	Percentage of passen...	Average distance (miles)
First-time	23780	18.31	714.12
Returning	106100	81.69	1297.05

5) Are departure and arrival delays affected by flight distance? See the scatter plots on Tableau.