Jomatech, SQL Introduction to Databases: Assignment 4.

Contents

Jomatech, SQL Introduction to Databases: Assignment 4. 1

A4Q1: Which device type generates most views? 2

A4Q2: Are there any specific days where another device type generates more views than the device from the answer above? 3

A4Q3: Which device type has the highest completion rate? Why do you think that is? 4

A4Q4: Which device type has the highest share rate? Why do you think that is? 5

A4Q5: Which day had the most views? 6

A4VQ1: Plot a chart to visualize the number of quizzes from different types of quizzes 7

A4VQ2: Plot a chart to visualize the total number of views every day. Does the trend look good for us? 8

A4VQ3: Plot a chart to visualize the total number of views every day but for each device type. Does this align with what you found in problem 1, question 1 & 2? 9

A4VQ4: Show the different completion and share rates for each device type in a chart. What is your hypothesis on why there is a difference? 10

A4Q1: Which device type generates most views?

The device_type "desktop" rakes in most views when looking at the sum.

```sql
/* A4Q1: Which device type generates most views? */
SELECT
  device_type, SUM(num_views)
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  device_type
ORDER BY SUM(num_views) DESC
```

✓ 3 rows | 43B returned in 1s

	device_type	sum
1	desktop	68442083
2	mobile	68193244
3	tablet	8167154

A4Q2: Are there any specific days where another device type generates more views than the device from the answer above?

From the returned values we can see that mobile takes over desktop on 2019-03-04, 2019-03-05, 2019-03-07 and numerous other days.

```
SELECT
  "date",
  SUM(num_views) as num_views,
  device_type
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  "date",
  device_type
ORDER BY
  "date",
  num_views DESC;
```

✓ 100 rows | 2KB returned in 572ms

	date	num_views	device_type
1	2019-03-01 00:00:00	504750	desktop
2	2019-03-01 00:00:00	415213	mobile
3	2019-03-01 00:00:00	62486	tablet
4	2019-03-02 00:00:00	665725	desktop
5	2019-03-02 00:00:00	545211	mobile
6	2019-03-02 00:00:00	63859	tablet
7	2019-03-03 00:00:00	693853	desktop
8	2019-03-03 00:00:00	617209	mobile
9	2019-03-03 00:00:00	67863	tablet
10	2019-03-04 00:00:00	1083129	mobile
11	2019-03-04 00:00:00	910988	desktop
12	2019-03-04 00:00:00	112274	tablet
13	2019-03-05 00:00:00	1212496	mobile
14	2019-03-05 00:00:00	1035563	desktop
15	2019-03-05 00:00:00	129322	tablet

A4Q3: Which device type has the highest completion rate? Why do you think that is?

We find that quizzes have the highest completion rate on desktops, based on the data from the table.

Since there is not enough data to go on, the why is just a hypothesis. Perhaps desktop users generally commit to longer sessions when compared to mobile and tablet users.

Desktops are designed for more work to be done on them, while mobile and tablet are not (ergonomically) designed for longer use and long media consumption times.

The odds may therefore be in favor of desktops when it comes to quiz completion rates.

```sql
SELECT
  device_type,
  ROUND(((SUM(num_completes) / SUM(num_views)) * 100)) AS completion_rate
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  device_type
ORDER BY
  completion_rate DESC;
```

✓ 3 rows | 43B returned in 485ms

	device_type	completion_rate
1	desktop	78
2	mobile	73
3	tablet	73

A4Q4: Which device type has the highest share rate? Why do you think that is?

The share rate is highest on tablets according to the given data, with ~1.44% of tablet views leading to a share.

An assumption to make is that the share rate is highest on tablet due to the inherent user experience design on tablets (and mobile phones) being catered towards sharing via social media. This is in stark contrast to desktop browsing of quizzes where the button size of share options is smaller and often gets obscured by multiple activities or tabs on the screen(s).

```sql
SELECT
  device_type,
  ((SUM(num_shares) / SUM(num_views)) * 100) AS share_rate
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  device_type
ORDER BY
  share_rate DESC;
```

✓ 3 rows | 43B returned in 571ms

	device_type	share_rate
1	tablet	1.4363999993143266
2	mobile	1.2221313301945278
3	desktop	0.5675104891240671

A4Q5: Which day had the most views?

April 9th, 2019 is the day with most views according to the table data.

```
75  SELECT
76    DISTINCT "date",
77    SUM(num_views) as total_views
78  FROM
79    public.daily_agg_quiz_metrics
80  GROUP BY
81    "date"
82  ORDER BY
83    total_views DESC;
```

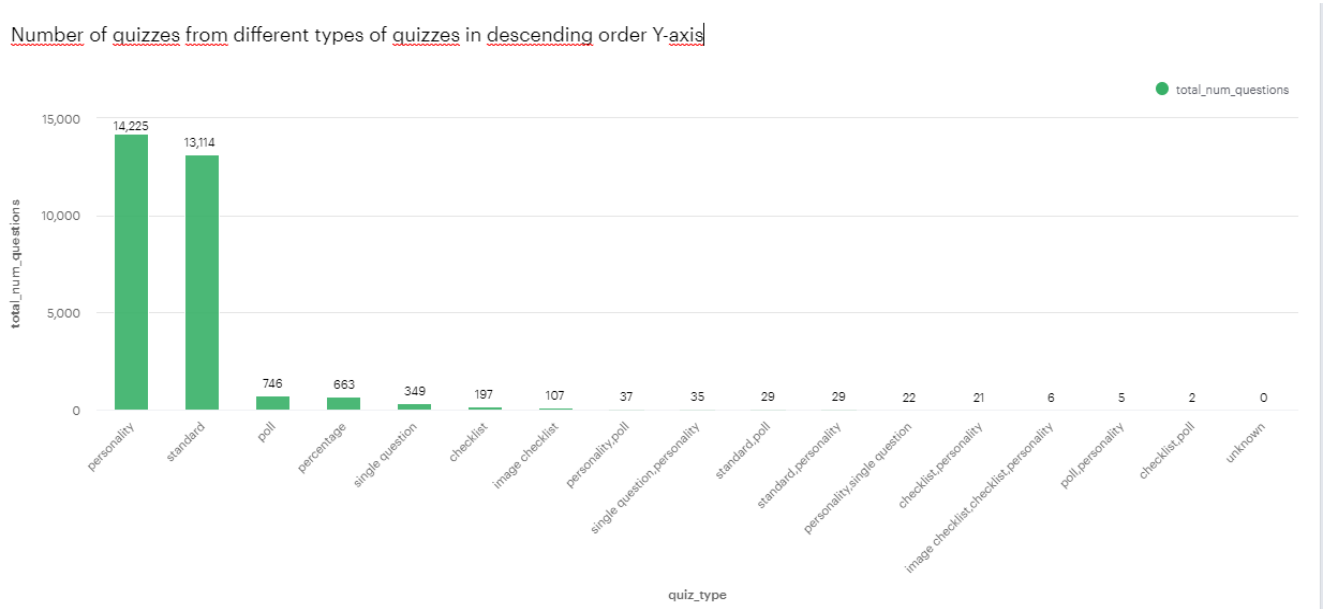
✓ 90 rows | 1KB returned in 574ms

	date	total_views
1	2019-04-09 00:00:00	2820316
2	2019-03-31 00:00:00	2718327
3	2019-04-16 00:00:00	2578897
4	2019-03-30 00:00:00	2410062
5	2019-03-05 00:00:00	2377381
6	2019-04-10 00:00:00	2315357
7	2019-04-02 00:00:00	2315094
8	2019-04-01 00:00:00	2294713
9	2019-04-08 00:00:00	2290698
10	2019-04-03 00:00:00	2133439
11	2019-03-04 00:00:00	2106391
12	2019-04-17 00:00:00	2099340
13	2019-03-23 00:00:00	2097593
14	2019-04-22 00:00:00	2076755
15	2019-04-07 00:00:00	2062515

A4VQ1: Plot a chart to visualize the number of quizzes from different types of quizzes

The personality quiz type scores highest, closely followed by the standard quiz type, leaving the rest far behind.

```
SELECT
  quiz_type,
  SUM(num_questions) as total_num_questions
FROM
  public.quiz_metadata
GROUP BY
  quiz_type
ORDER BY
  total_num_questions DESC;
```



A4VQ2: Plot a chart to visualize the total number of views every day. Does the trend look good for us?

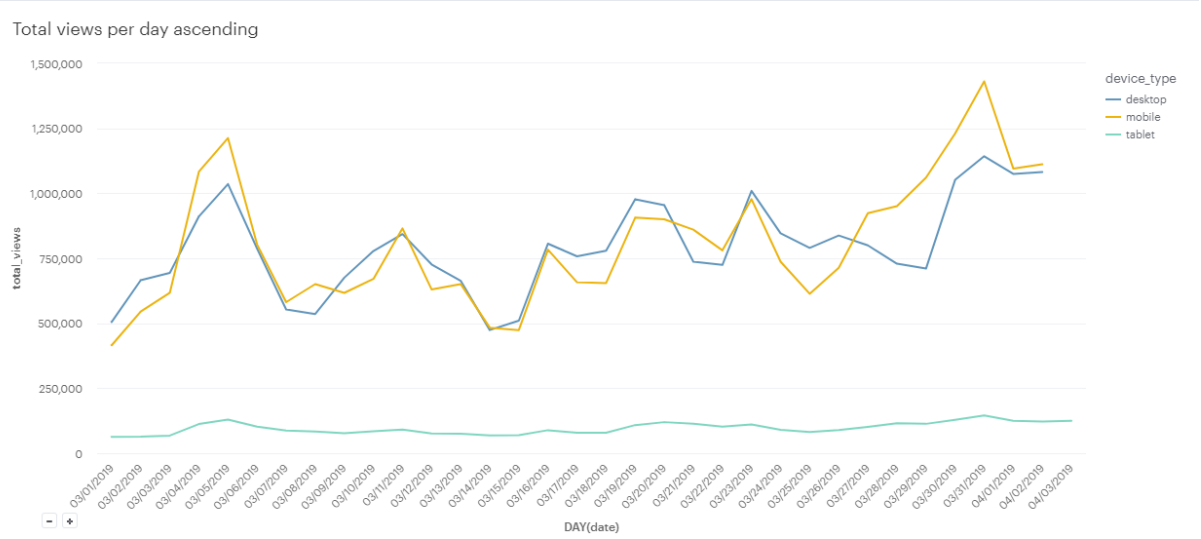
We can observe a steady decline of total views following an upward trend in april. This is not a good development.

```sql
SELECT
  "date",
  SUM(num_views) as total_views
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  "date"
ORDER BY
  "date";
```

A4VQ3: Plot a chart to visualize the total number of views every day but for each device type. Does this align with what you found in problem 1, question 1 & 2?

The total view results vary from the specific quiz view results from assignment 1, problem 1 & 2. It's almost a reversed trend in terms of peaks and throughs when comparing num_views per day from quiz_id 2023144 (mobile) and total views per day via mobile.

```
117 SELECT
118   "date",
119   SUM(num_views) as total_views,
120   device_type
121 FROM
122   public.daily_agg_quiz_metrics
123 GROUP BY
124   "date",
125   device_type
126 ORDER BY
127   "date";
```



	quiz_id	device_type	num_views	num_completes	num_shares	date
1	2023144	mobile	117	63	0	2019-03-01 00:00:00
2	2023144	mobile	63	34	0	2019-03-02 00:00:00
3	2023144	mobile	58	35	0	2019-03-03 00:00:00
4	2023144	mobile	88	49	0	2019-03-04 00:00:00
5	2023144	mobile	98	55	0	2019-03-05 00:00:00
6	2023144	mobile	66	40	0	2019-03-06 00:00:00
7	2023144	mobile	121	79	0	2019-03-07 00:00:00
8	2023144	mobile	104	64	0	2019-03-08 00:00:00
9	2023144	mobile	78	47	0	2019-03-09 00:00:00
10	2023144	mobile	65	46	0	2019-03-10 00:00:00
11	2023144	mobile	82	53	0	2019-03-11 00:00:00
12	2023144	mobile	62	38	0	2019-03-12 00:00:00
13	2023144	mobile	77	51	0	2019-03-13 00:00:00
14	2023144	mobile	108	78	1	2019-03-14 00:00:00
15	2023144	mobile	109	74	1	2019-03-15 00:00:00

A4VQ4: Show the different completion and share rates for each device type in a chart. What is your hypothesis on why there is a difference?

The combined answers from non-visual question 3 and 4 serve as hypothesis here.

```sql
SELECT
  device_type,
  ROUND(((SUM(num_completes) / SUM(num_views)) * 100)) AS completion_rate,
  ((SUM(num_shares) / SUM(num_views)) * 100) AS share_rate
FROM
  public.daily_agg_quiz_metrics
GROUP BY
  device_type
ORDER BY
  completion_rate DESC
```