

/*

Is a small percentage of Hacker News submitters taking the majority of the points?

/*

```
SELECT SUM(score)
```

```
FROM hacker_news;
```

/*

the users who have accumulated a lot of points across their stories.

the individual users who have gotten combined scores of more than 200, and their combined scores.

/*

```
SELECT user, SUM(score)
```

```
FROM hacker_news
```

```
GROUP BY user
```

```
HAVING SUM(score) > 200
```

```
ORDER BY 2 DESC;
```

/*

the percentage

/*

```
SELECT (517 + 309 + 304 + 282) / 6366.0;
```

/*

```
SELECT user,
```

```
    COUNT(*)
```

```
FROM hacker_news
```

```
WHERE url LIKE '%watch?v=dQw4w9WgXcQ%'
```

```
GROUP BY user
```

```
ORDER BY COUNT(*) DESC;
```

/*

Which of these sites feed Hacker News the most:

GitHub, Medium, or New York Times?

categorize each story based on their source

/*

SELECT CASE

WHEN url LIKE '%github.com%' THEN 'GitHub'

-- WHEN statement here

-- WHEN statement here

-- ELSE statement here

END AS 'Source'

FROM hacker_news;

/*

a column for the number of stories from each URL using COUNT().

/*

-- Which sites feed Hacker News?

SELECT CASE WHEN url LIKE '%github.com%' THEN 'GitHub' WHEN url LIKE '%medium.com%' THEN 'Medium' WHEN url LIKE '%nytimes.com%' THEN 'New York Times'

ELSE 'Other' END AS 'Source',

COUNT(*) FROM hacker_news

GROUP BY 1;

/*

What's the best time of the day to post a story on Hacker News?

/*

SELECT timestamp

```
FROM hacker_news
```

```
LIMIT 10;
```

```
SELECT timestamp,
```

```
    strftime('%H', timestamp)
```

```
FROM hacker_news
```

```
GROUP BY 1
```

```
LIMIT 20;
```

```
/*
```

The best hours are in the morning around 7 am and afternoon around 6 - 8 pm! Monster difference!

```
/*
```

```
SELECT strftime('%H', timestamp),
```

```
    AVG(score),
```

```
    COUNT(*)
```

```
FROM hacker_news
```

```
GROUP BY 1
```

```
ORDER BY 2 DESC;
```

```
/*
```

The hours of the timestamp
The *average* score for each hour
The *count* of stories for each hour

```
/*
```

```
SELECT strftime('%H', timestamp),
```

```
    AVG(score),
```

```
    COUNT(*)
```

FROM hacker_news

GROUP BY 1

ORDER BY 2 DESC;

*/**

Best time to post a story

*/**

SELECT strftime('%H', timestamp) AS 'Hour',

ROUND(AVG(score), 1) AS 'Average Score',

COUNT(*) AS 'Number of Stories'

FROM hacker_news

WHERE timestamp IS NOT NULL

GROUP BY 1 ORDER BY 2 DESC;