

Question:

What does the line graph representing the total number of cd's (any type) checked out from 2005-2010 look like?

In Processing, the vertical axis will represent years 2005-2010 and the horizontal will be the total number of checkouts.

Query:

```
SELECT
    itemtype, count(*),
    SUM(CASE WHEN year(ckinDateTime) = '2005' THEN 1 ELSE 0 END) as yr2005,
    SUM(CASE WHEN year(ckinDateTime) = '2006' THEN 1 ELSE 0 END) as yr2006,
    SUM(CASE WHEN year(ckinDateTime) = '2007' THEN 1 ELSE 0 END) as yr2007,
    SUM(CASE WHEN year(ckinDateTime) = '2008' THEN 1 ELSE 0 END) as yr2008,
    SUM(CASE WHEN year(ckinDateTime) = '2009' THEN 1 ELSE 0 END) as yr2009,
    SUM(CASE WHEN year(ckinDateTime) = '2010' THEN 1 ELSE 0 END) as yr2010
FROM
    transactionsallID
WHERE
    itemtype like "%cd%"
```

Explanation:

SELECT itemtype, count(), grab the item type info and show the total count*

SUM show me how many items are being counted in years 2005-2010

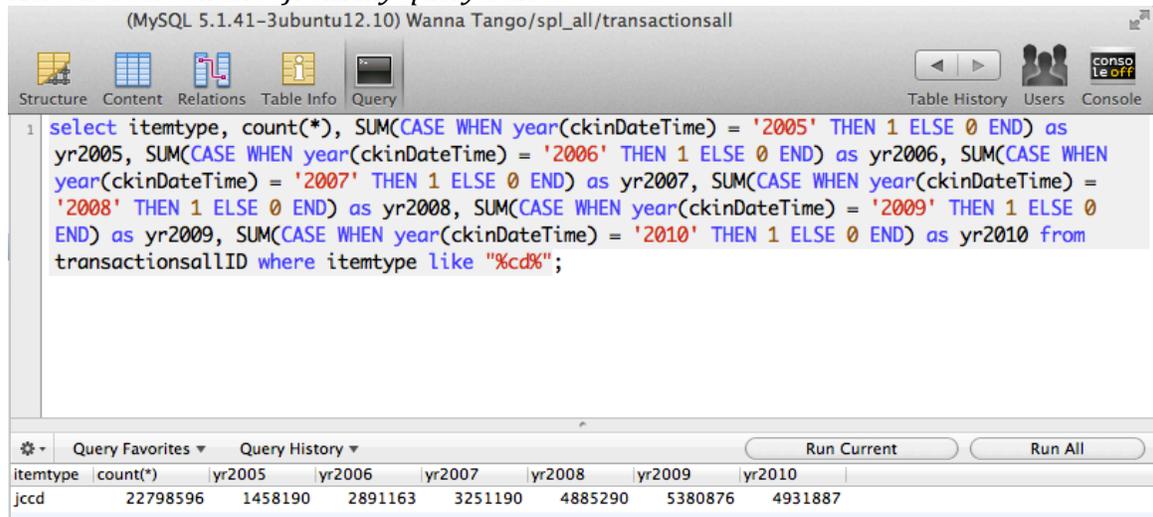
FROM transactionsallID the data set I'm pulling from

WHERE itemtype like %cd% pull only items with "cd" in type

Results, Analysis, & Feedback:

To begin, this query came back with one row of data and took 112ms to complete. I have organized it in this way so I could have 1 row of data containing separate columns of the total counts I was looking for. These numbers are incredibly large too, going up and beyond millions, but are just total counts for the cd-type items.

Here is a screen shot from my query below



The screenshot shows a MySQL query editor window titled "(MySQL 5.1.41-3ubuntu12.10) Wanna Tango/spl_all/transactionsall". The query is as follows:

```
1 select itemtype, count(*), SUM(CASE WHEN year(ckinDateTime) = '2005' THEN 1 ELSE 0 END) as yr2005, SUM(CASE WHEN year(ckinDateTime) = '2006' THEN 1 ELSE 0 END) as yr2006, SUM(CASE WHEN year(ckinDateTime) = '2007' THEN 1 ELSE 0 END) as yr2007, SUM(CASE WHEN year(ckinDateTime) = '2008' THEN 1 ELSE 0 END) as yr2008, SUM(CASE WHEN year(ckinDateTime) = '2009' THEN 1 ELSE 0 END) as yr2009, SUM(CASE WHEN year(ckinDateTime) = '2010' THEN 1 ELSE 0 END) as yr2010 from transactionsallID where itemtype like "%cd%";
```

Below the query, a table displays the results:

itemtype	count(*)	yr2005	yr2006	yr2007	yr2008	yr2009	yr2010
jccd	22798596	1458190	2891163	3251190	4885290	5380876	4931887

I created a visual graph of this data in Processing and was easily able to see the total number of cd's checked out over the selected years



→ This clearly shows that 2009 had the most checkouts with years 2008 and 2010 also relatively close. The same goes for years 2006 and 2007, although 2007 had a higher total value that is easily visible here.

→ The line for 2005 seems rather small in my opinion but I would assume that they don't have the accurate and/or complete data totals. 2005 was the first year they started collecting data, I believe, and if they didn't start exactly on January 1st of that year then it's not as accurate as the following years.

→ I also notice a large and fairly consistent jump in the totals from year 2007 to years 2008 and beyond. I'm trying to think of an explanation for this occurrence but I'm rather stumped. I would have thought there would be a steady decline in cd totals around that time period because of the popularity increase of digital music downloads, Internet streaming sites, iPods, etc.