## 10 The Sales Rollup

In this exercise we are going to write a ton of queries about an interesting sales dataset.

## The Data

Table D.15: Sales Table, 10,250 Rows

| SID | CID | amount | TID | sales_dt |
| :---: | :---: | :---: | :---: | :---: |
| 1 | 4 | 13.55 | 1 | $1 / 11 / 2011$ |
| 2 | 12 | 18.99 | 2 | $12 / 22 / 2012$ |
| 2 | 12 | 22.01 | 3 | $1 / 12 / 2013$ |

Table D.16: Expenses Table, 425 Rows

| SID | CID | exp_dt | amount | e_type |
| :---: | :---: | :---: | :---: | :---: |
| 12 | 18 | $5 / 4 / 2012$ | 112.24 | Dinner |
| 2 | 44 | $10 / 10 / 2010$ | 112.24 | Sport |

Table D.17: Transactions Table, 25,254 Rows

| TID | IID | num |
| :---: | :---: | :---: |
| 1 | 12 | 1 |
| 1 | 18 | 1 |
| 2 | 45 | 1 |
| 3 | 18 | 1 |

Table D.18: Client Table, 152 Rows

| CID | Name | Address | City | st | Zip |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | John Smith | 12 Blue Bell street | Hayward | CA | 94552 |
| 2 | Julia Xue | 14 Howard Street | Danville | VA | 24543 |

Table D.19: SalesPerson Table, 22 Rows

| SID | Name | Address | City | st | Zip | StartDt | EndDt |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Rachel Adams | 27796 Hanover Hill | Hanover | NY | 14081 | $12-12-2011$ | $1-3-2014$ |
| 2 | Julia Xue | 60 Darwin Court | Mobile | AL | 36602 | $01-11-2012$ | $7-7-2013$ |

Table D.20: Items Table, 440 Rows

| IID | ItemName | Cost |
| :---: | :---: | :---: |
| 1 | 10 lbs. Concrete | 12.95 |
| 2 | 20 lbs. Concrete | 19.95 |
| 3 | 30 lbs. Concrete | 27.95 |

## Questions

## Main Questions

1. How many sales people are there from CA?

Table D.21: Regions Table, 50 Rows

| ST | Region |
| :---: | :---: |
| CA | West |
| PA | East |
| AL | South |
| VT | North |

2. In how many different states does this company have clients?
3. What is the average, min and max cost of an item which is for sale?
4. What is the average, min and max cost of items which have concrete in their name?
5. What is the average, min and max cost of items which have concrete in their name vs. those which do not?
6. Write a query which returns the total sales and number of sales per month.
7. How many sales did each sales person have?
8. How many states are in each region?
9. How many transactions contained item 12 ?
10. What was the average number of items per transaction? (Think about what to do with the number of items column)
11. What are the top- 5 items sold in units-sold?

## A bit harder

1. The total dollars sold and number of items, per salesperson.
2. How many currently employed sales people?
3. Write a query which returns the total number of items that each salesperson sold.
4. What was the average amount expensed per-client and per-salesperson? (Need to do two queries?)
5. How many sales people are in each region?
6. How many sales are from each region? (How would you measure this?)
7. For each client, compute the Revenue - expenses.
8. For each client, compute the profit (revenue - expenses - costs).
9. For each client region, compute the total revenue.
10. For each salesperson region, compute the total revenue.
11. How many clients are in each region?
12. What is our profit per region? (How would you measure this?)
13. How would you calculate profit per transaction?
14. How many different sales people sold each item?
15. How many distinct items has each sales person sold?
16. For each region (based on client), return the $\#$ of salespeople servicing the region, the number of items sold in that region, the number of transactions that occurred in that region and the total cost of the items sold in that region.
17. Calculate, for each client region, the number of sales that occurred in the same region as the salesperson.
18. Because of weird tax rules, we need to collect a tax of $3 \%$ of revenue for those transactions where the client is in the western region. How much tax do we owe per month?

