|  |  |
| --- | --- |
| **employees** | **Field Description** |
| Ssn (pk) | Unique SSN ID# for employee |
| First\_name | Employee first name |
| Last\_name | Employee last\_name |
| Dept (fk) | Dept ID# |
| Start\_year | Year of employment |

|  |  |  |
| --- | --- | --- |
| **trips** | | **Field Description** |
| Employee (pk, fk) | | SSN of employee travelling |
| Trip\_ID (pk) | | Unique Trip ID# |
| Start\_date | | Start date of trip |
| End\_date | | End date of trip |
| Reason\_code (fk) | | Code for reason for trip |
|  | |  |
| **expenses** | **Field Description** | | |
| Employee (pk, fk)  Trip\_id (pk, fk) | SSN of employee travelling  Unique Trip ID# | | |
| Expense\_seq (pk) | Sequence# for expense report line item | | |
| Account\_no (fk) | Account number for line item | | |
| Gross\_amount | Gross dollar amount of line item | | |
| tax | Sales tax (if applicable) of line item | | |
|  |  | | |

|  |  |
| --- | --- |
| **dept\_codes** | **Field Description** |
| Dept\_ID (pk) | Dept ID# |
| Dept\_name | Name of department |

|  |  |  |
| --- | --- | --- |
| **Reason\_codes** | **Field Description** | |
| Reason\_code (pk) | Reason ID# | |
| Reason\_description | Description of reason for trip | |
|  |  | |
| **account\_codes** | **Field Description** |
| Account\_no (pk) | Account ID# |
| Account\_description | Description of account |
| Account\_type | Category of account |
| **reimbursements** | **Field Description** |
| Employee (pk, fk) | SSN of employee travelling |
| Trip\_id (pk, fk) | Unique Trip ID# |
| Auditor | Auditor last name |
| Reimbursement\_amount | Amount of reimbursement |
| Reimbursement\_date | Date of reimbursement |

Diagram

Description automatically generated

1. (Design and create a data warehouse for the Expense database. The decisions about which fields to include and how to aggregate the data are left to you. You do not need to include every single data point from the 7 tables given. Use your judgement as to what will be interesting/useful for the organization. But please make sure that you pull (combine) data from **at least four tables** and compute relevant aggregate statistics. Please see many examples from class lectures and you may adapt those codes for your purpose (for this dataset).

USE expense;

CREATE OR REPLACE VIEW expense\_dw\_report AS

SELECT ssn,first\_name,last\_name, dept\_name,

COUNT(trip\_id) number\_trips,

ROUND(sum(amount),2) total\_spent

FROM

(SELECT ssn, first\_name, last\_name, dept\_name, t.trip\_id,

reason\_code,

gross\_amount+tax AS amount

FROM trips t

JOIN employees e

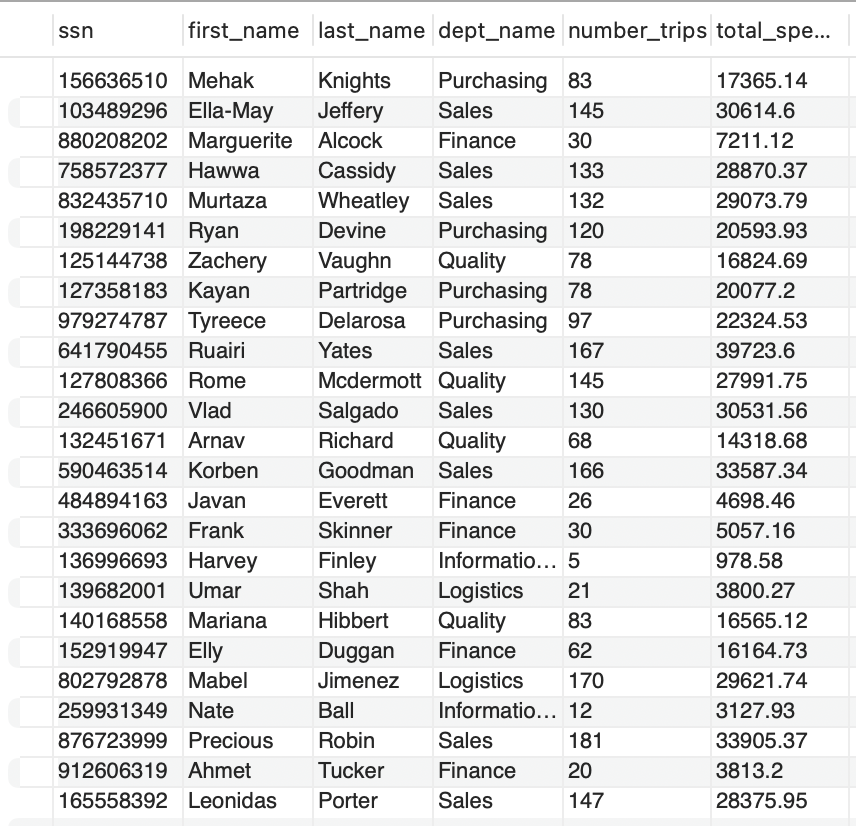
ON e.ssn=t.employee

JOIN expenses x

ON t.trip\_id=x.trip\_id

JOIN dept\_codes d ON e.dept=d.dept\_id) t

GROUP BY ssn;

****

2. Create **four** SQL queries on your data warehouse that answer interesting important questions. At least two queries should be more than simple queries. For example, more complex queries could include Joins, a Group By element or a subquery or use some aggregate functions and summary calculations (see examples in the class lectures’ slides).

**#1 no. of trips for each reason in descending order**

SELECT reason\_description, count(\*) total\_trips

FROM expense\_dw\_report e

JOIN trips t

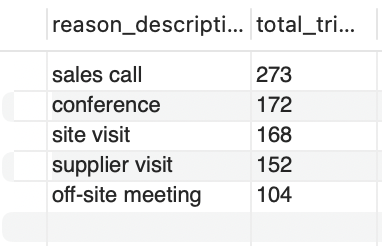
ON t.employee=e.ssn

JOIN reason\_codes r

ON t.reason\_code=r.reason\_code

GROUP BY t.reason\_code

ORDER BY total\_trips DESC;



**#2 total amt spent by dept**

SELECT dept\_name,sum(total\_spent) total

FROM expense\_dw\_report

GROUP BY dept\_name

ORDER BY total DESC;

A screenshot of a computer

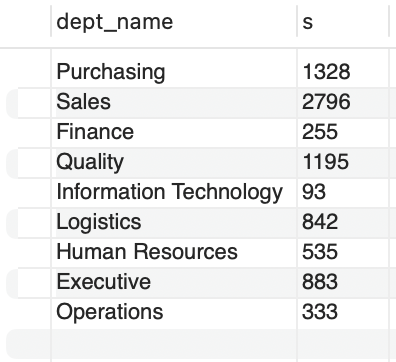
Description automatically generated

**#3 total trips by dept**

SELECT dept\_name,sum(number\_trips) s

FROM expense\_dw\_report

GROUP BY dept\_name;



#**4 reimbursements per dept**

SELECT dept\_name, ROUND(SUM(reimbursement\_amount),2) reimbursement

FROM expense\_dw\_report exp

JOIN reimbursements reim

ON exp.ssn = reim.employee

GROUP BY dept\_name

ORDER BY reimbursement DESC;

