**TEST**

**1.**

(10 points)

Imagine you have a table claim including 3 columns: customer\_id, claim\_datetime, claim\_amount. Write a query to only return the last claim record for each customer. Hint: One customer may claim multiple times, so only select the latest claim record.

claim

|  |  |  |
| --- | --- | --- |
| customer\_id | claim\_datetime | claim\_amount |
| 1 | 2020-02-01 | 500 |
| 1 | 2020-03-01 | 500 |
| 1 | 2020-04-01 | 500 |
| 2 | 2020-02-01 | 600 |
| 2 | 2020-04-01 | 1200 |
| .... | | |

**2.**

(10 points)

Write a query to output the province where customers have the highest average balance.

customer\_contact\_info

|  |  |  |
| --- | --- | --- |
| account\_number | postal\_code | province |
| 283285 | M2N2A2 | ON |
| 973525 | V5A4A6 | BC |
| 736823 | B6N8M3 | NS |
| 160186 | T1K4X5 | AB |
| 384623 | L6A3C5 | ON |
| 164623 | A1A2B2 | NL |
| 864302 | K1A4V5 | ON |
| 321247 | V1A9J2 | BC |
| 459804 | T6W4B4 | AB |
| .... | | |

customer\_status

|  |  |  |  |
| --- | --- | --- | --- |
| account\_number | status | balance | credit\_limit |
| 283285 | OPEN | 3050 | 5000 |
| 160186 | OPEN | 234 | 10000 |
| 973525 | CLOSED | 0 | 8000 |
| 459804 | OPEN | 2007 | 7500 |
| 321247 | OPEN | 2500 | 20000 |
| 384623 | OPEN | 1805 | 10000 |
| 736823 | OPEN | 342 | 3000 |
| 864302 | OPEN | 900 | 1000 |
| 164623 | CLOSED | 0 | 5000 |
| .... | | | |

**3.**

(15 points)

Write a query based on 4 tables to list the names of supervisors for the crew on the flight to Copenhagen (destination= CPH) on March 4, 2013. Note:

* Supervisors live in the same state as the employees they supervise.
* There is one supervisor for each state and job category. Therefore, to find the supervisor for employee, you need to make sure that supervisor's state = employee's state and supervisor's job category = employee's job category.
* Table flight\_schedule and payroll\_master only contain information for the crew.
* Table staff\_master contains information for the crew and supervisors.
* job\_code with the same first 2 characters belong to the same job category, for example, jobcode FA2 and FA3 are under job category FA. Hint: use substring(jobcode ,1,2)=jobcategory

flight\_schedule

|  |  |  |  |
| --- | --- | --- | --- |
| emp\_id | date | destination | flight\_number |
| 1269 | 2013-03-04 | YYZ | 182 |
| 1556 | 2013-03-04 | CPH | 387 |
| 1830 | 2013-03-04 | CPH | 387 |
| 1739 | 2013-03-04 | CPH | 387 |
| 1437 | 2013-03-04 | CPH | 387 |
| 1545 | 2013-03-04 | FRA | 622 |
| .... | | | |

staff\_master

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| emp\_id | last\_name | first\_name | city | state | phone |
| 1830 | ADAMS | GERALD | STAMFORD | CT | (203)781-1255 |
| 1556 | ALEXANDER | SUSAN | BRIDGEPORT | CT | (203)675-7715 |
| 1739 | BOYCE | JONATHAN | NEW YORK | NY | (212)587-1247 |
| 1545 | BRADLEY | JEREMY | NEW YORK | NY | (212)587-3622 |
| 1126 | BROWN | JASON | STAMFORD | CT | (203)781-0019 |
| 1118 | CARTER | DONALD | NEW YORK | NY | (718)384-2946 |
| 1437 | CARTER | DOROTHY | BRIDGEPORT | CT | (203)675-4117 |
| 1639 | CARTER | KAREN | STAMFORD | CT | (203)781-8839 |
| 1269 | CASTON | FRANKLIN | STAMFORD | CT | (203)781-3335 |
| .... | | | | | |

payroll\_master

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| emp\_id | gender | date\_of\_birth | date\_of\_hire | job\_code | salary |
| 1556 | F | 1973-09-04 | 1998-08-01 | FA3 | 46040 |
| 1545 | M | 1969-04-03 | 1989-02-14 | PT2 | 124048 |
| 1437 | F | 1976-12-26 | 1997-10-09 | TA3 | 54299 |
| 1739 | M | 1972-12-29 | 1999-01-30 | PT1 | 93124 |
| 1830 | M | 1968-05-01 | 2000-12-24 | FA2 | 40001 |
| 1269 | M | 1961-02-28 | 1988-01-04 | PT2 | 127926 |
| .... | | | | | |

supervisors

|  |  |
| --- | --- |
| emp\_id | job\_category |
| 1882 | ME |
| 1118 | PT |
| 1639 | FA |
| 1126 | TA |
| .... | |

**4.**

(25 points)

Using Excel to transform Bucho Burrito point-of-sale (POS) data to provide

1. a monthly summary of sales and traffic for each store (use "Indic Sales" as sales and "# Trans" as traffic).
2. a week over week traffic trend analysis for Alberta and Quebec (use "# Trans" as traffic).
3. a 7 Days (Monday to Sunday) traffic analysis for Ontario stores (use "# Trans" as traffic).

**5.**

(40 points)

**Overview**

You will have received 3 datasets containing account, customer, and transaction level data. We are interested in understanding the process you take in answering the following questions. Please provide the answer and an explanation where applicable. These questions are designed to test your critical thinking, logic, and problem-solving skills. Please document your approach and present your findings as if you were presenting to an (non-technical) business partner executive. You can use any programming language you would like to complete the task.

**Your Task**

Use the data provided to answer the following questions:

1. We are planning to launch a new product focused on a specific merchant category. Which specific merchant category would you like to focus on for this new product?  Please explain your rationale for this category incorporating both the insights derived from the data and other concepts where you see fit.
2. Identify and describe various segments of customers within the data. Consider applying segmenting/clustering techniques to aid in the development of your answer.
3. Of the segments that you created in question 2, which specific segment would you like to target for this new product?  Why would you target them?  What are the potential challenges/risks to consider when targeting this segment vs. others?

**Approach**

Outline your approach to answering the above questions. Include all general assumptions and interpretations of the questions.

**Executive Summary**

Outline your key findings of your analysis. Use visuals to enhance understanding of your analysis.