**PARHAN ASSIGNMENT**

**\**

# 

# **DAFTAR ISI**

[DAFTAR ISI ii](#_Toc159710208)

[Instruction Samsung Assignment 1](#_Toc159710209)

[1. Task 1: Raw Data Analytics & Data Modeling 2](#_Toc159710210)

[2. Task: Create SQL Query 3](#_Toc159710211)

# **Instruction Assignment**

**Task 1 : Raw Data Analytics & Data Modeling**

Please download dataset from <https://www.kaggle.com/datasets/ylchang/coffee-shop-sample-data-1113>, and use the file from the given link to work on this task.

Please put the data into database with the most effective and optimal structure

* It is allowed to create any data model and structure that you think suitable (you can normalize or denormalize as needed). Please let us know if the data has some flaws that makes it hard to be processed (if any) during presentation
* The artifact of this tasks is the ERD diagram, please send it when you submit the result

**Task 2: Create SQL Query**

From the data on task 1, we will provide several question that you need to answer, for this task please provide the:

* SQL Query needed to produce the result
* Query result

The questions are the following:

1. Please produce a table with the following columns: customer\_id, customer\_first\_name, customer\_email, and transaction\_date which has the total line\_item\_amount the same with the total line\_item\_amount of user with customer\_email ' Charissa@Integer.us' and transaction\_date at April 20th, 2019. Please sort it with the latest transaction\_date on top
2. Please produce a table with the following columns: product\_type, total quantity and status\_sold, for transaction\_date from April 6th, 2019 to April 14th, 2019. Please sort it from the largest to the smallest. For status\_sold column please apply the following rules:
   1. If total quantity more than or equal to 2000 then 'High Sold'
   2. If total quantity more than or equal to 1000 but less than 2000 then 'Medium Sold'
   3. If total quantity less than 1000 then 'Low Sold'
3. Produce a table with the following columns: product\_type and total quantity sold. Please show all product\_type for transaction\_date at April 2nd, 2019, if total total quantity sold is empty (no transaction) then fill it with 0. Please sort it based on the largest to smallest total quantity sold
4. Produce a table with the following columns: customer\_id, customer\_first\_name, and customer\_email which has the largest total quantity for each transaction date from April 12th, 2019 to April 20th, 2019. Please sort it with the latest date on top
5. Produce a table with the following columns: product\_id, product. Show only 10 product with the largest total quantity for transaction\_date from April 11th, 2019 to April 13th, 2019. Please sort it from the largest to the lowest total quantity

# **Task 1: Raw Data Analytics & Data Modeling**

A diagram of a company

Description automatically generated

# **Task: Create SQL Query**

* Please produce a table with the following columns: customer\_id, customer\_first\_name, customer\_email, and transaction\_date which has the total line\_item\_amount the same with the total line\_item\_amount of user with customer\_email ' Charissa@Integer.us' and transaction\_date at April 20th, 2019. Please sort it with the latest transaction\_date on top

**SELECT** c.customer\_id, c.customer\_first\_name, c.customer\_email, s.transaction\_date, **SUM**(s.line\_item\_amount) **AS** total\_line\_item\_amount

**FROM** Customer c

**JOIN** Sales s **ON** c.customer\_id = s.customer\_id

**WHERE** c.customer\_email = 'Charissa@Integer.us' **AND** s.transaction\_date = '2019-04-20'

**GROUP** **BY** c.customer\_id, c.customer\_first\_name, c.customer\_email, s.transaction\_date

**ORDER** **BY** s.transaction\_date **DESC**;

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **customer\_id** | **customer\_first\_name** | **customer\_email** | **transaction\_date** | **total\_line\_item\_amount** |
| **8144** | Leslie | [Charissa@Integer.us](mailto:Charissa@Integer.us) | 2019-04-20 | 30.5 |

* Please produce a table with the following columns: product\_type, total quantity and status\_sold, for transaction\_date from April 6th, 2019 to April 14th, 2019. Please sort it from the largest to the smallest. For status\_sold column please apply the following rules:
  1. If total quantity more than or equal to 2000 then 'High Sold'
  2. If total quantity more than or equal to 1000 but less than 2000 then 'Medium Sold'
  3. If total quantity less than 1000 then 'Low Sold'

**SELECT**

p.product\_type,

**SUM**(s.quantity) **AS** total\_quantity,

**CASE**

**WHEN** **SUM**(s.quantity) >= 2000 **THEN** 'High Sold'

**WHEN** **SUM**(s.quantity) >= 1000 **THEN** 'Medium Sold'

**ELSE** 'Low Sold'

**END** **AS** status\_sold

**FROM**

Sales s

**JOIN**

Product p **ON** s.product\_id = p.product\_id

**WHERE**

s.transaction\_date **BETWEEN** '2019-04-06' **AND** '2019-04-14'

**GROUP** **BY**

p.product\_type

**ORDER** **BY**

total\_quantity **DESC**;

|  |  |  |
| --- | --- | --- |
| **product\_type** | **total\_quantity** | **status\_sold** |
| **Brewed Chai tea** | 2586 | High Sold |
| **Gourmet brewed coffee** | 2546 | High Sold |
| **Barista Espresso** | 2399 | High Sold |
| **Brewed Black tea** | 1797 | Medium Sold |
| **Hot chocolate** | 1688 | Medium Sold |
| **Brewed herbal tea** | 1645 | Medium Sold |
| **Premium brewed coffee** | 1369 | Medium Sold |
| **Drip coffee** | 1283 | Medium Sold |
| **Organic brewed coffee** | 1199 | Medium Sold |
| **Regular syrup** | 1049 | Medium Sold |
| **Scone** | 1028 | Medium Sold |
| **Pastry** | 856 | Low Sold |
| **Brewed Green tea** | 824 | Low Sold |
| **Biscotti** | 631 | Low Sold |
| **Sugar free syrup** | 451 | Low Sold |
| **Chai tea** | 78 | Low Sold |
| **Housewares** | 70 | Low Sold |
| **Organic Beans** | 66 | Low Sold |
| **Gourmet Beans** | 48 | Low Sold |
| **Premium Beans** | 45 | Low Sold |
| **Drinking Chocolate** | 44 | Low Sold |
| **Herbal tea** | 41 | Low Sold |
| **Espresso Beans** | 40 | Low Sold |
| **Black tea** | 40 | Low Sold |
| **Organic Chocolate** | 32 | Low Sold |
| **House blend Beans** | 23 | Low Sold |
| **Clothing** | 23 | Low Sold |
| **Green tea** | 16 | Low Sold |
| **Green beans** | 14 | Low Sold |

* Produce a table with the following columns: product\_type and total quantity sold. Please show all product\_type for transaction\_date at April 2nd, 2019, if total total quantity sold is empty (no transaction) then fill it with 0. Please sort it based on the largest to smallest total quantity sold

**SELECT**

p.product\_type,

**COALESCE**(**SUM**(s.quantity), 0) **AS** total\_quantity\_sold

**FROM**

Product p

**LEFT** **JOIN**

Sales s **ON** p.product\_id = s.product\_id **AND** s.transaction\_date = '2019-04-02'

**GROUP** **BY**

p.product\_type

**ORDER** **BY**

total\_quantity\_sold **DESC**;

|  |  |
| --- | --- |
| **product\_type** | **total\_quantity\_sold** |
| **Brewed Chai tea** | 323 |
| **Gourmet brewed coffee** | 314 |
| **Barista Espresso** | 272 |
| **Hot chocolate** | 210 |
| **Brewed Black tea** | 205 |
| **Brewed herbal tea** | 196 |
| **Organic brewed coffee** | 173 |
| **Drip coffee** | 171 |
| **Brewed Green tea** | 133 |
| **Premium brewed coffee** | 127 |
| **Scone** | 107 |
| **Biscotti** | 64 |
| **Pastry** | 59 |
| **Seasonal drink** | 0 |
| **Organic Beans** | 0 |
| **House blend Beans** | 0 |
| **Drinking Chocolate** | 0 |
| **Clothing** | 0 |
| **Sugar free syrup** | 0 |
| **Organic Chocolate** | 0 |
| **Herbal tea** | 0 |
| **Gourmet Beans** | 0 |
| **Green tea** | 0 |
| **Premium Beans** | 0 |
| **Espresso Beans** | 0 |
| **Black tea** | 0 |
| **Green beans** | 0 |
| **Regular syrup** | 0 |
| **Housewares** | 0 |
| **Chai tea** | 0 |
| **Specialty coffee** | 0 |

* Produce a table with the following columns: customer\_id, customer\_first\_name, and customer\_email which has the largest total quantity for each transaction date from April 12th, 2019 to April 20th, 2019. Please sort it with the latest date on top

**WITH** TotalQuantityPerDate **AS** (

**SELECT**

s.customer\_id,

c.customer\_first\_name,

c.customer\_email,

s.transaction\_date,

**SUM**(s.quantity) **AS** total\_quantity,

**ROW\_NUMBER**() **OVER**(**PARTITION** **BY** s.transaction\_date **ORDER** **BY** **SUM**(s.quantity) **DESC**) **AS** rn

**FROM**

Sales s

**INNER** **JOIN**

Customer c **ON** s.customer\_id = c.customer\_id

**WHERE**

s.transaction\_date **BETWEEN** '2019-04-12' **AND** '2019-04-20'

**GROUP** **BY**

s.customer\_id,

c.customer\_first\_name,

c.customer\_email,

s.transaction\_date

)

**SELECT**

customer\_id,

customer\_first\_name,

customer\_email,

transaction\_date

**FROM**

TotalQuantityPerDate

**WHERE**

rn = 1

**ORDER** **BY**

transaction\_date **DESC**;

|  |  |  |  |
| --- | --- | --- | --- |
| **customer\_id** | **customer\_first\_name** | **customer\_email** | **transaction\_date** |
| **221** | Emery Hooper | [Xander@vel.edu](mailto:Xander@vel.edu) | 2019-04-20 |
| **8146** | Ciara | [Rigel@luctus.com](mailto:Rigel@luctus.com) | 2019-04-19 |
| **34** | Murphy Mckee | [Hunter@Ut.edu](mailto:Hunter@Ut.edu) | 2019-04-18 |
| **8311** | Hanna | [Rylee@taciti.org](mailto:Rylee@taciti.org) | 2019-04-17 |
| **5055** | Catherine | [Abel@luctus.gov](mailto:Abel@luctus.gov) | 2019-04-16 |
| **8352** | Todd | [Kane@ad.org](mailto:Kane@ad.org) | 2019-04-15 |
| **3** | Elvis Cardenas | [Brianna@tellus.edu](mailto:Brianna@tellus.edu) | 2019-04-14 |
| **8071** | Simone | [Oren@eu.net](mailto:Oren@eu.net) | 2019-04-13 |
| **5874** | Shannon | [Xanthus@nibh.net](mailto:Xanthus@nibh.net) | 2019-04-12 |

* Produce a table with the following columns: product\_id, product. Show only 10 product with the largest total quantity for transaction\_date from April 11th, 2019 to April 13th, 2019. Please sort it from the largest to the lowest total quantity

**SELECT**

s.product\_id,

p.product,

**SUM**(s.quantity) **AS** total\_quantity

**FROM**

Sales s

**JOIN**

Product p **ON** s.product\_id = p.product\_id

**WHERE**

s.transaction\_date **BETWEEN** '2019-04-11' **AND** '2019-04-13'

**GROUP** **BY**

s.product\_id, p.product

**ORDER** **BY**

total\_quantity **DESC**

**limit** 10;

|  |  |  |
| --- | --- | --- |
| **product\_id** | **product** | **total\_quantity** |
| **84** | Chocolate syrup | 188 |
| **65** | Sugar Free Vanilla syrup | 180 |
| **50** | Earl Grey Rg | 179 |
| **48** | English Breakfast Rg | 179 |
| **63** | Carmel syrup | 177 |
| **42** | Lemon Grass Rg | 170 |
| **27** | Brazilian Lg | 169 |
| **37** | Espresso shot | 168 |
| **71** | Chocolate Croissant | 165 |
| **31** | Ethiopia Sm | 164 |