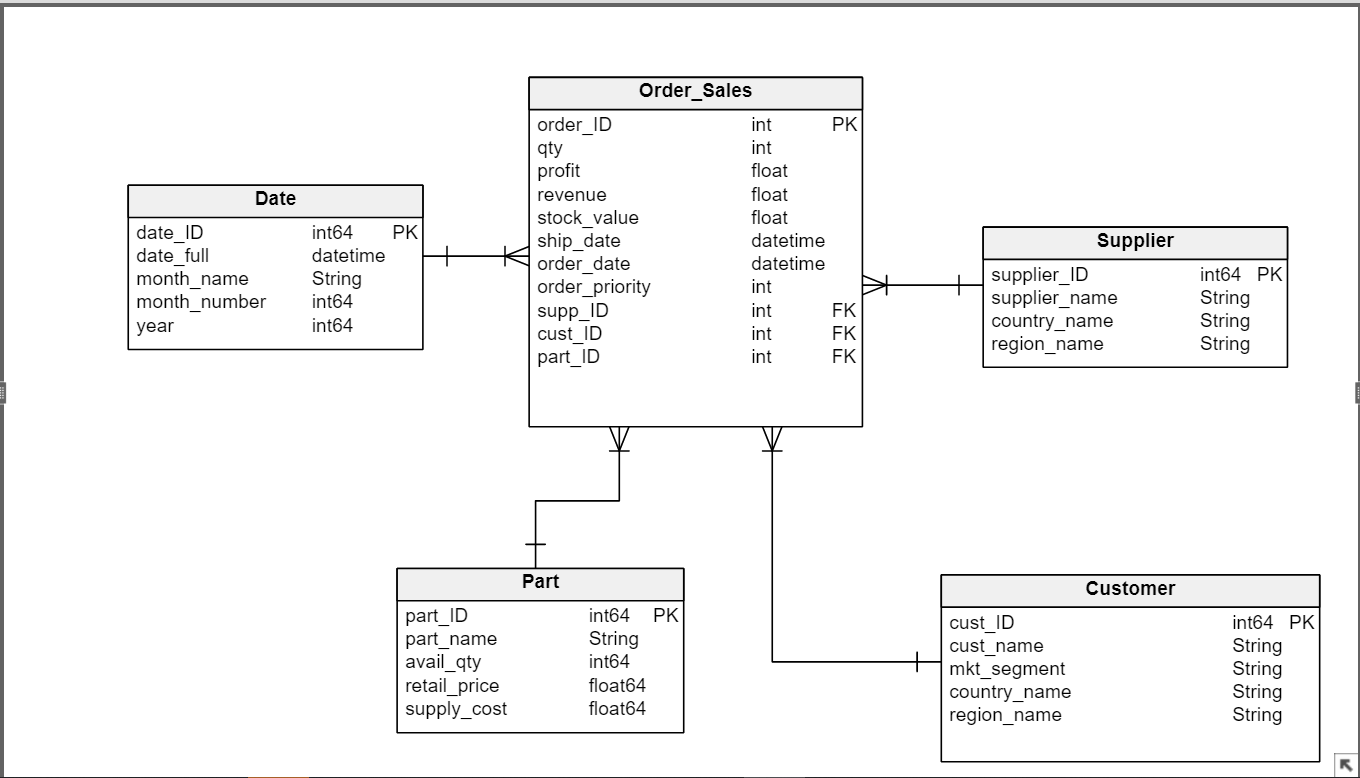
Scott Harris + Noah Dunn

Mini-Project #2

2/29/2020

**Schema**

Please Note: Vertabello cancelled my trial, so all the ID parameters should be classified as strings and not ints

****

**Schema DDL**

-- Created by Vertabelo (http://vertabelo.com)

-- Last modification date: 2020-02-20 16:20:40.989

-- tables

-- Table: Customer

CREATE TABLE IF NOT EXISTS `kcd\_dw.Customer` (

cust\_ID String NOT NULL,

cust\_name String NOT NULL,

mkt\_segment String NOT NULL,

country\_name String NOT NULL,

region\_name String NOT NULL,

);

-- Table: Date

CREATE TABLE IF NOT EXISTS `kcd\_dw.Date` (

date\_ID String NOT NULL,

date\_full datetime NOT NULL,

month\_name String NOT NULL,

month\_number int64 NOT NULL,

year int64 NOT NULL,

);

-- Table: Order\_Sales

CREATE TABLE IF NOT EXISTS `kcd\_dw.Order\_Sales` (

order\_ID String NOT NULL,

qty int64 NOT NULL,

profit float64 NOT NULL,

revenue float64 NOT NULL,

stock\_value float64 NOT NULL,

ship\_date datetime NOT NULL,

order\_date datetime NOT NULL,

order\_priority String NOT NULL,

supp\_ID String NOT NULL,

cust\_ID String NOT NULL,

part\_ID String NOT NULL

);

-- Table: Part

CREATE TABLE IF NOT EXISTS `kcd\_dw.Part` (

part\_ID String NOT NULL,

part\_name String NOT NULL,

avail\_qty int64 NOT NULL,

retail\_price float64 NOT NULL,

supply\_cost float64 NOT NULL,

);

-- Table: Supplier

CREATE TABLE IF NOT EXISTS `kcd\_dw.Supplier` (

supplier\_ID String NOT NULL,

supplier\_name String NOT NULL,

country\_name String NOT NULL,

region\_name String NOT NULL,

);

**Insert Statements**

Note: We left the queries in to ensure that these could be tested thoroughly

**Populating Part Table**

INSERT INTO `mini-project-2-269717.kcd\_dw.Part`

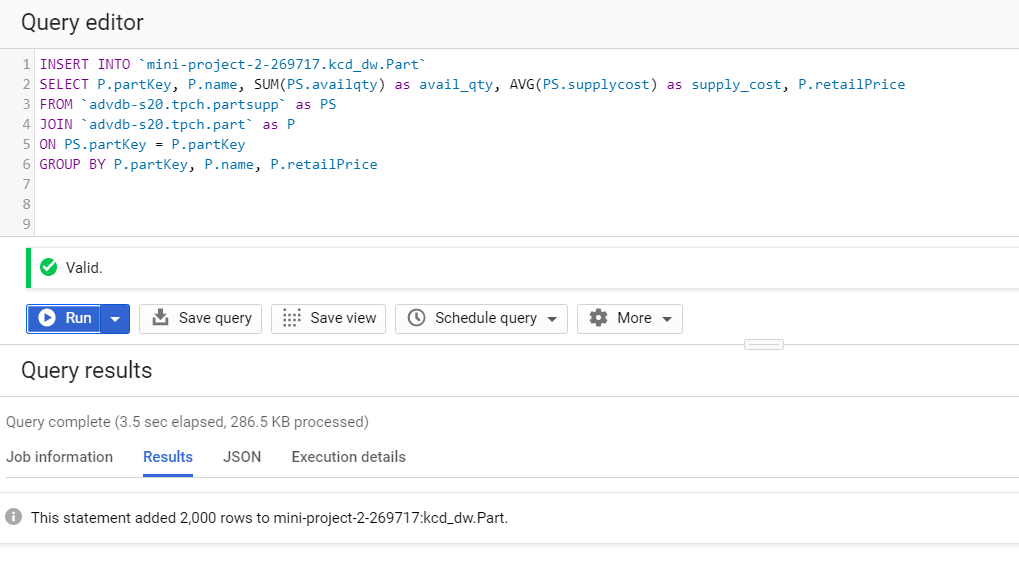
SELECT P.partKey, P.name, SUM(PS.availqty) as avail\_qty, AVG(PS.supplycost) as supply\_cost, P.retailPrice

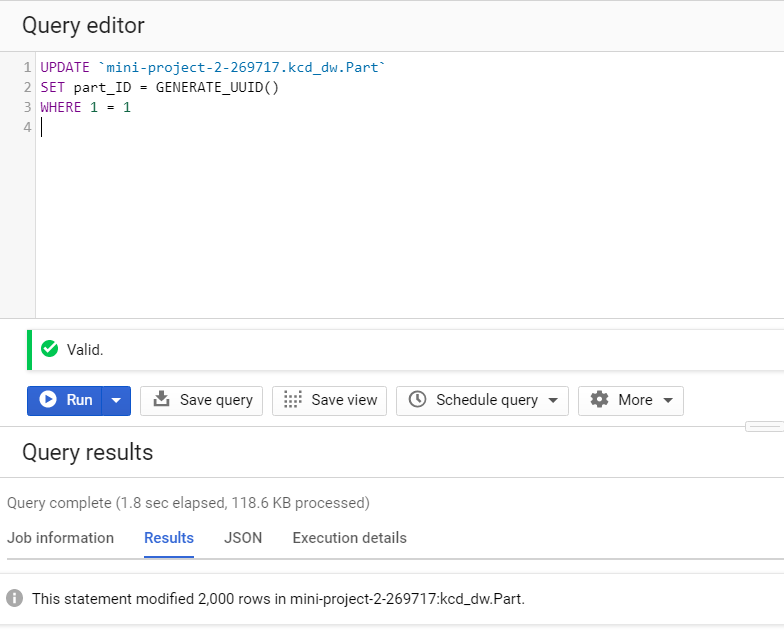
FROM `advdb-s20.tpch.partsupp` as PS

JOIN `advdb-s20.tpch.part` as P

ON PS.partKey = P.partKey

GROUP BY P.partKey, P.name, P.retailPrice



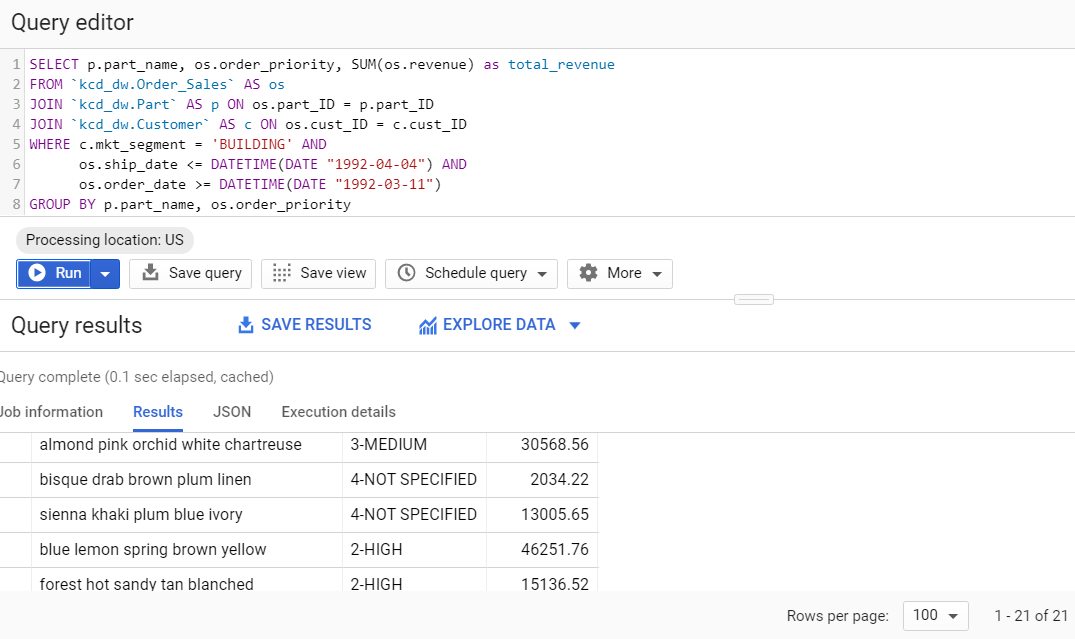


UPDATE `mini-project-2-269717.kcd\_dw.Part`

SET part\_ID = GENERATE\_UUID()

WHERE 1 = 1

**Populating Supplier Table**

****

INSERT INTO `kcd\_dw.Supplier` (supplier\_ID, supplier\_name, country\_name, region\_name)

SELECT DISTINCT “1” as supplier\_ID,

s.name AS supplier\_name,

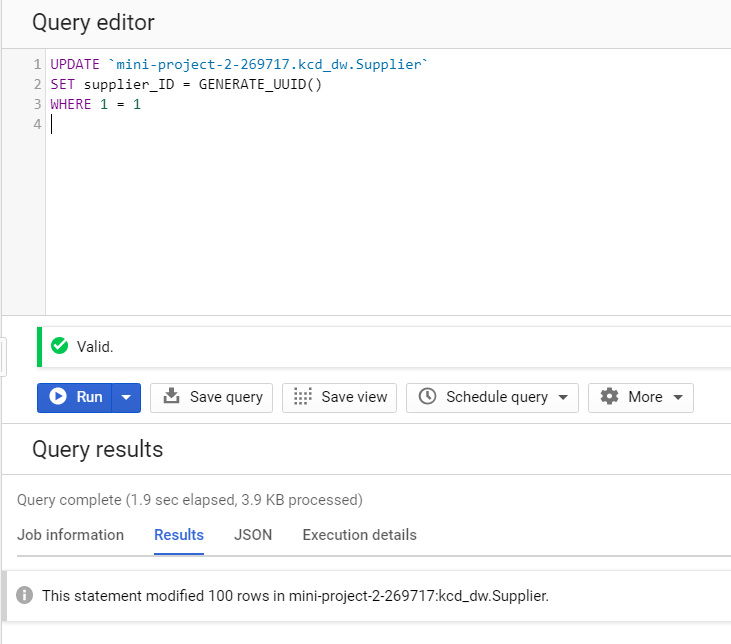
n.name AS country\_name,

r.name AS region\_name

FROM `advdb-s20.tpch.supplier` AS s

JOIN `advdb-s20.tpch.nation` AS n ON s.nationkey = n.nationkey

JOIN `advdb-s20.tpch.region` AS r ON n.regionkey = r.regionkey;

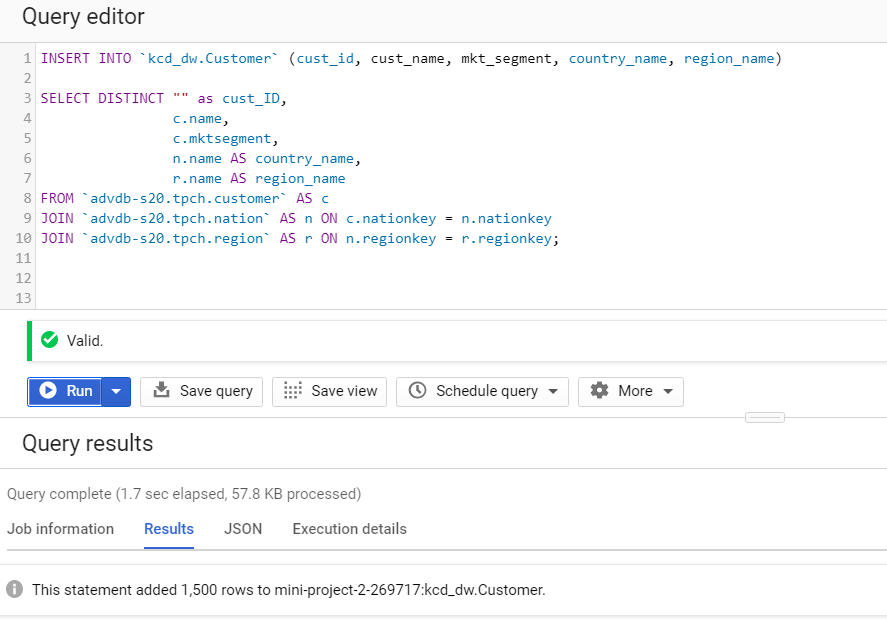


UPDATE `mini-project-2-269717.kcd\_dw.Supplier`

SET supplier\_ID = GENERATE\_UUID()

WHERE 1 = 1

**Populating Customer Table**



INSERT INTO `kcd\_dw.Customer` (cust\_id, cust\_name, mkt\_segment, country\_name, region\_name)

SELECT DISTINCT "" as cust\_ID,

c.name,

c.mktsegment,

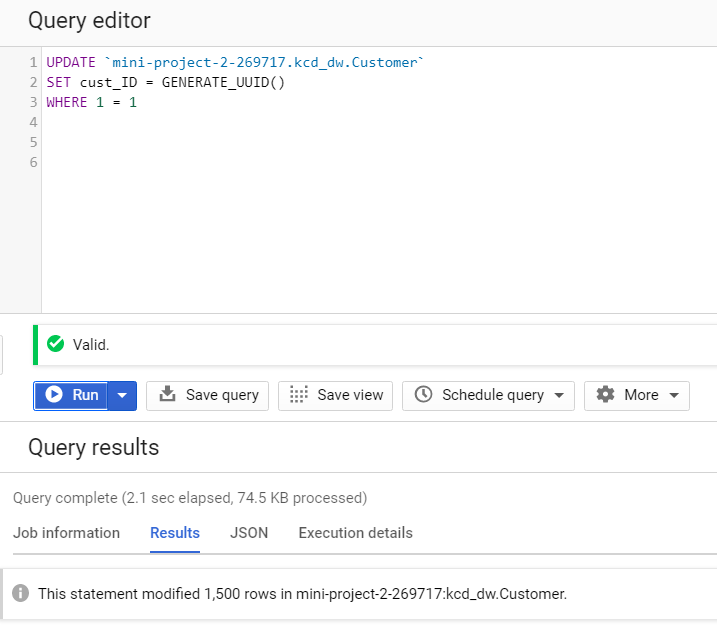
n.name AS country\_name,

r.name AS region\_name

FROM `advdb-s20.tpch.customer` AS c

JOIN `advdb-s20.tpch.nation` AS n ON c.nationkey = n.nationkey

JOIN `advdb-s20.tpch.region` AS r ON n.regionkey = r.regionkey;

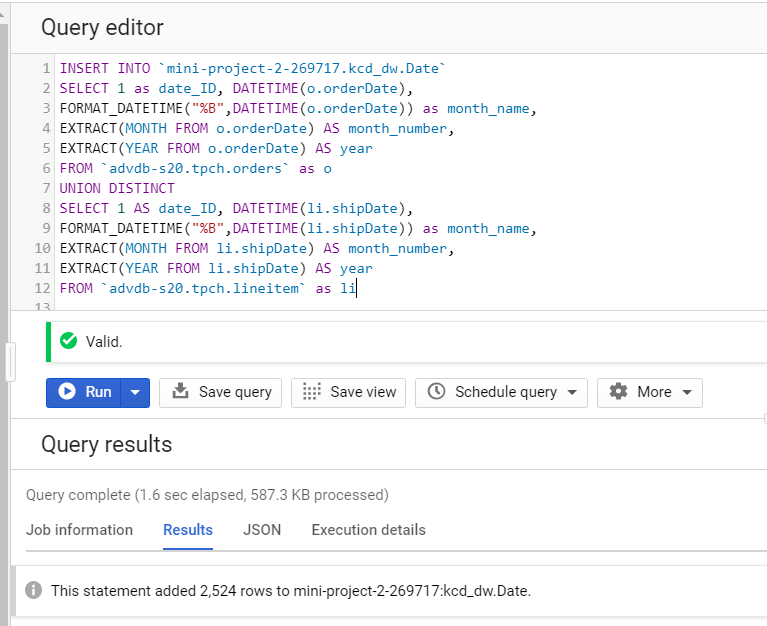


UPDATE `mini-project-2-269717.kcd\_dw.Customer`

SET cust\_ID = GENERATE\_UUID()

WHERE 1 = 1

**Populating Date Table**



INSERT INTO `mini-project-2-269717.kcd\_dw.Date`

SELECT "1" as date\_ID, DATETIME(o.orderDate),

FORMAT\_DATETIME("%B",DATETIME(o.orderDate)) as month\_name,

EXTRACT(MONTH FROM o.orderDate) AS month\_number,

EXTRACT(YEAR FROM o.orderDate) AS year

FROM `advdb-s20.tpch.orders` as o

UNION DISTINCT

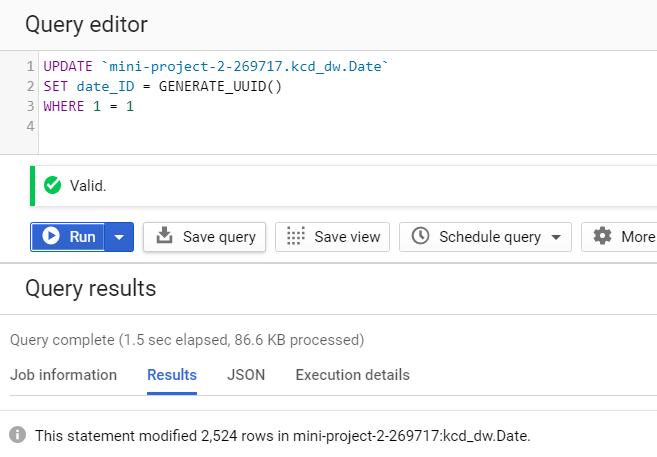
SELECT "1" AS date\_ID, DATETIME(li.shipDate),

FORMAT\_DATETIME("%B",DATETIME(li.shipDate)) as month\_name,

EXTRACT(MONTH FROM li.shipDate) AS month\_number,

EXTRACT(YEAR FROM li.shipDate) AS year

FROM `advdb-s20.tpch.lineitem` as li



UPDATE `mini-project-2-269717.kcd\_dw.Date`

SET date\_ID = GENERATE\_UUID()

WHERE 1 = 1

**Order Sales Table**

INSERT INTO `mini-project-2-269717.kcd\_dw.Order\_Sales`

SELECT

GENERATE\_UUID() as order\_id,

li.quantity,

li.extendedPrice - sp.retailprice AS profit,

li.extendedPrice AS revenue,

sps.availqty \* sps.supplycost AS stock\_value,

DATETIME(li.shipDate) AS ship\_date,

DATETIME(so.orderDate) AS order\_date,

so.orderpriority,

s.supplier\_ID,

c.cust\_ID,

p.part\_ID

FROM `kcd\_dw.Customer` AS c,

`kcd\_dw.Part` AS p,

`kcd\_dw.Supplier` AS s,

`advdb-s20.tpch.orders` AS so,

`advdb-s20.tpch.customer` AS sc,

`advdb-s20.tpch.lineitem` AS li,

`advdb-s20.tpch.partsupp` AS sps,

`advdb-s20.tpch.part` AS sp,

`advdb-s20.tpch.supplier` AS ss ,

`advdb-s20.tpch.nation` AS sn1,

`advdb-s20.tpch.region` AS sr1,

`advdb-s20.tpch.nation` AS sn2,

`advdb-s20.tpch.region` AS sr2

WHERE

--vvvvvvvv Joining the Source Tables vvvvvvvvv

so.custKey = sc.custKey AND

li.suppKey = ss.suppKey AND

li.orderKey = so.orderKey AND

li.partKey = sp.partKey AND

sps.partKey = sp.partkey AND

sps.suppKey = ss.suppKey AND

ss.nationKey = sn1.nationKey AND

sn1.regionKey = sr1.regionKey AND

sc.nationKey = sn2.nationKey AND

sr2.regionKey = sn2.regionKey AND

-- ^^^^^^^ Joining the Source Tables ^^^^^^^

p.part\_name = sp.name AND

s.supplier\_name = ss.name AND

s.country\_name = sn1.name AND

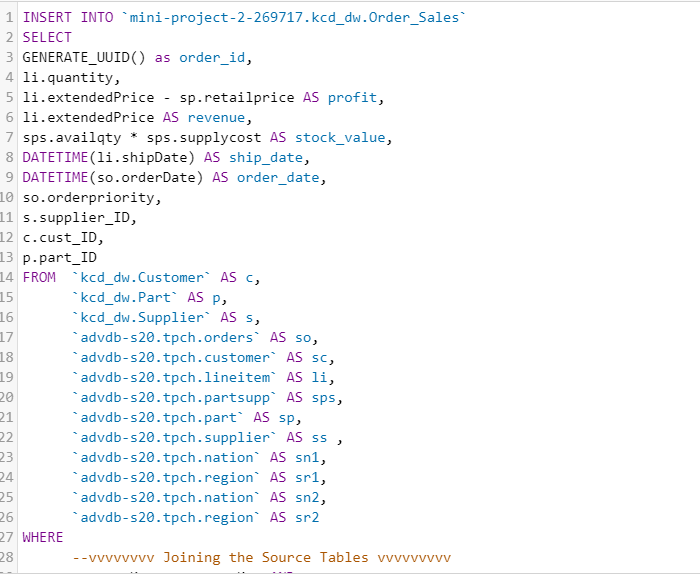
s.region\_name = sr1.name AND

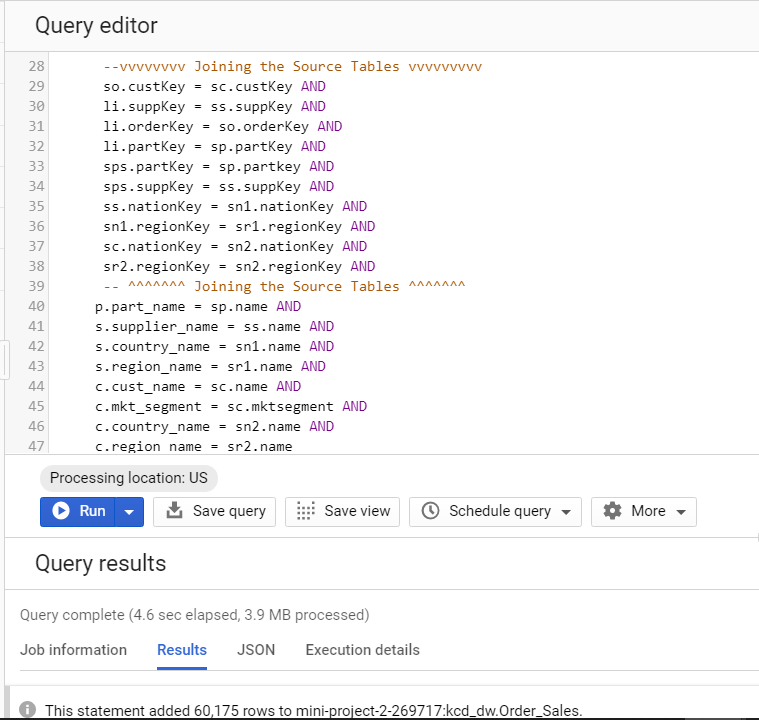
c.cust\_name = sc.name AND

c.mkt\_segment = sc.mktsegment AND

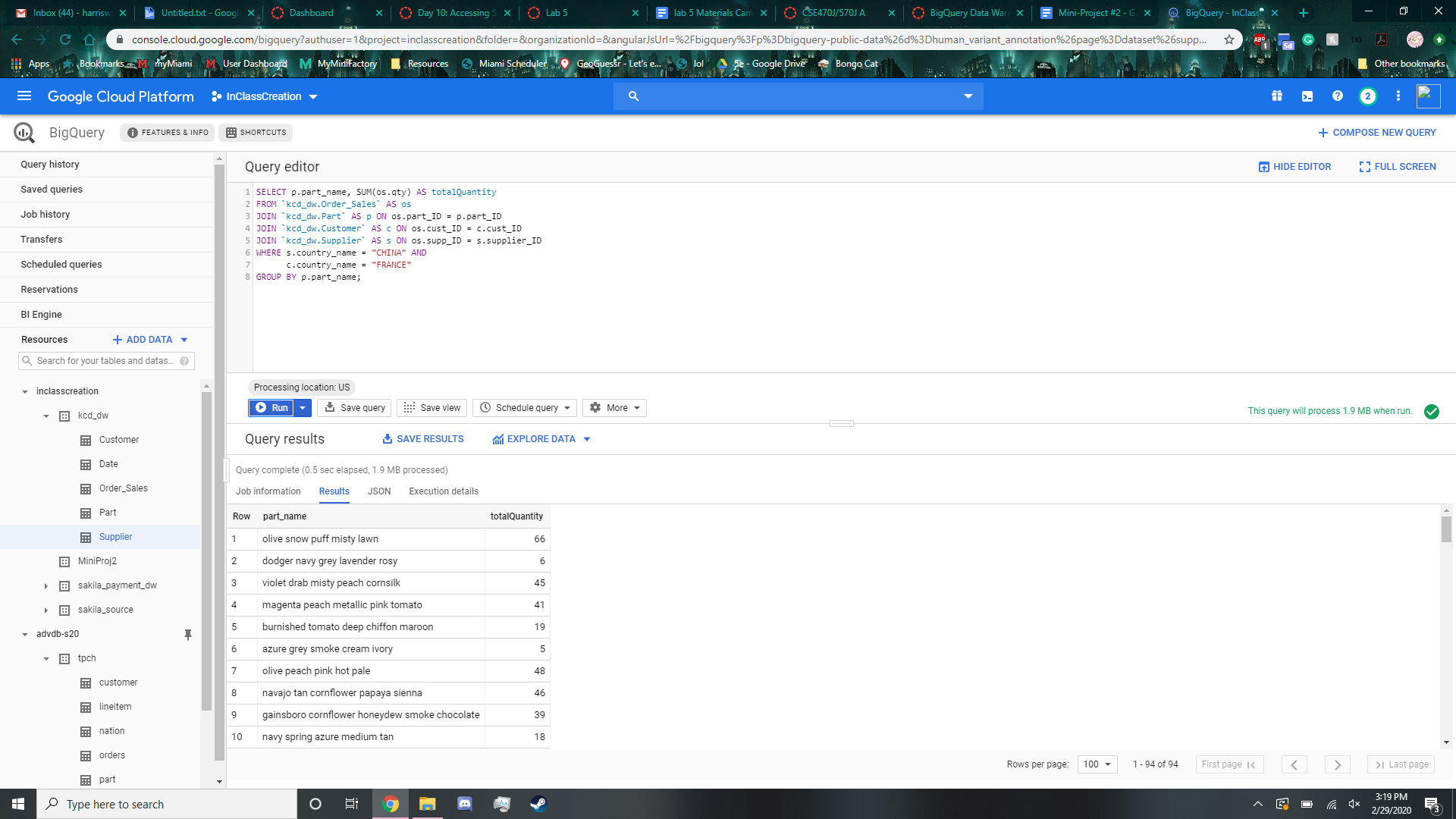
c.country\_name = sn2.name AND

c.region\_name = sr2.name





**Data Demands**

**Q1:** 

SELECT p.part\_name, SUM(os.qty) AS totalQuantity

FROM `kcd\_dw.Order\_Sales` AS os

JOIN `kcd\_dw.Part` AS p ON os.part\_ID = p.part\_ID

JOIN `kcd\_dw.Customer` AS c ON os.cust\_ID = c.cust\_ID

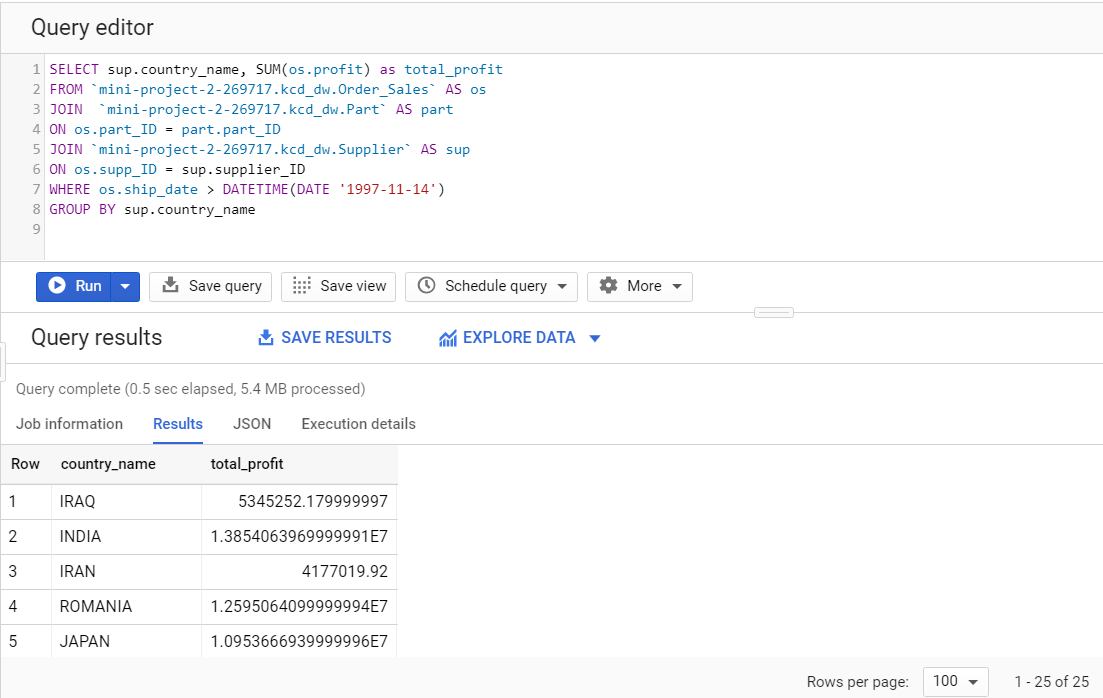
JOIN `kcd\_dw.Supplier` AS s ON os.supp\_ID = s.supplier\_ID

WHERE s.country\_name = "CHINA" AND

c.country\_name = "FRANCE"

GROUP BY p.part\_name;

**Q2:**



SELECT sup.country\_name, SUM(os.profit) as total\_profit

FROM `mini-project-2-269717.kcd\_dw.Order\_Sales` AS os

JOIN `mini-project-2-269717.kcd\_dw.Part` AS part

ON os.part\_ID = part.part\_ID

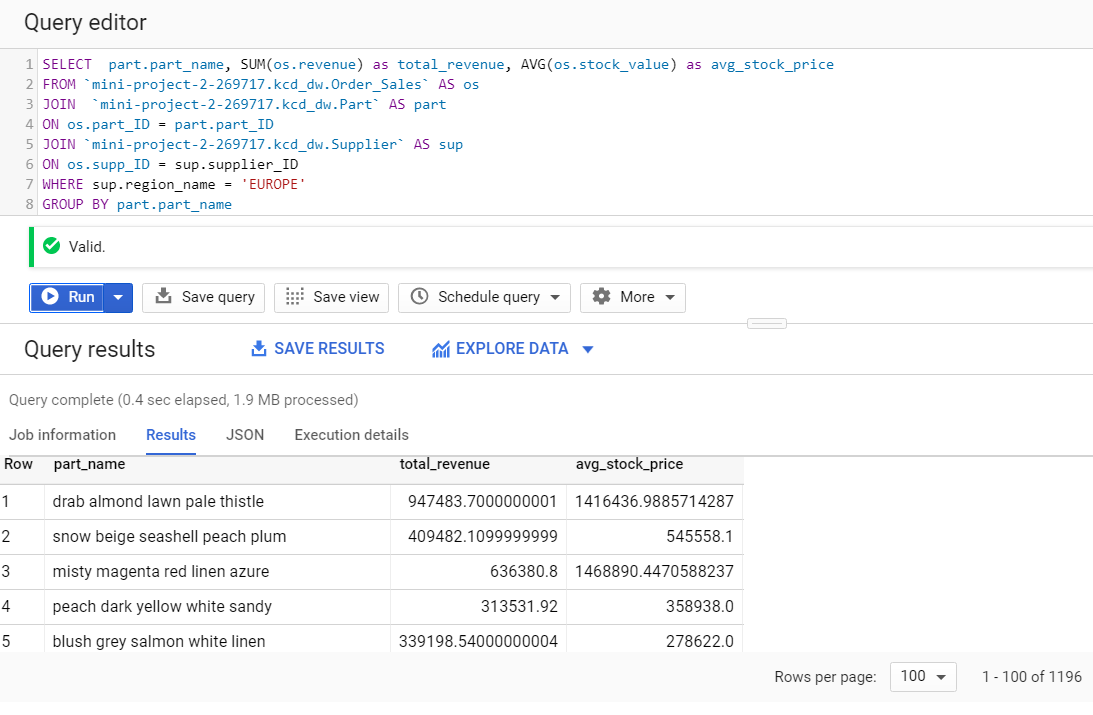
JOIN `mini-project-2-269717.kcd\_dw.Supplier` AS sup

ON os.supp\_ID = sup.supplier\_ID

WHERE os.ship\_date > DATETIME(DATE '1997-11-14')

GROUP BY sup.country\_name

**Q3:**



SELECT part.part\_name, SUM(os.revenue) as total\_revenue, AVG(os.stock\_value) as avg\_stock\_price

FROM `mini-project-2-269717.kcd\_dw.Order\_Sales` AS os

JOIN `mini-project-2-269717.kcd\_dw.Part` AS part

ON os.part\_ID = part.part\_ID

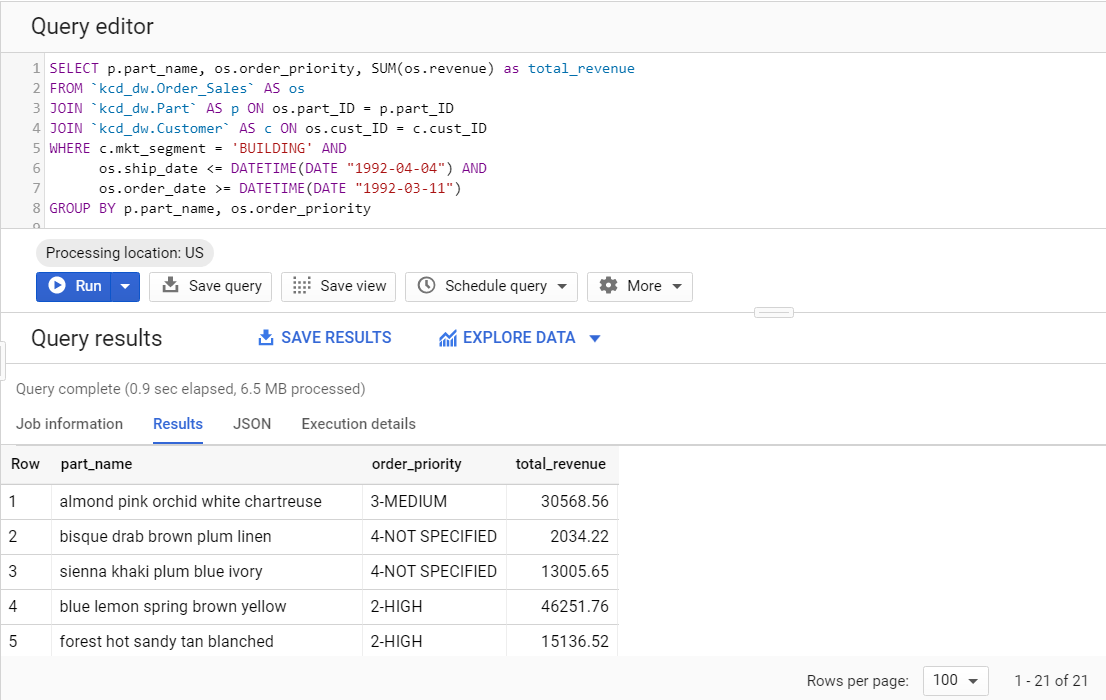
JOIN `mini-project-2-269717.kcd\_dw.Supplier` AS sup

ON os.supp\_ID = sup.supplier\_ID

WHERE sup.region\_name = 'EUROPE'

GROUP BY part.part\_name

**Q4:**



SELECT p.part\_name, os.order\_priority, SUM(os.revenue) as total\_revenue

FROM `kcd\_dw.Order\_Sales` AS os

JOIN `kcd\_dw.Part` AS p ON os.part\_ID = p.part\_ID

JOIN `kcd\_dw.Customer` AS c ON os.cust\_ID = c.cust\_ID

WHERE c.mkt\_segment = 'BUILDING' AND

os.ship\_date <= DATETIME(DATE "1992-04-04") AND

os.order\_date >= DATETIME(DATE "1992-03-11")

GROUP BY p.part\_name, os.order\_priority