

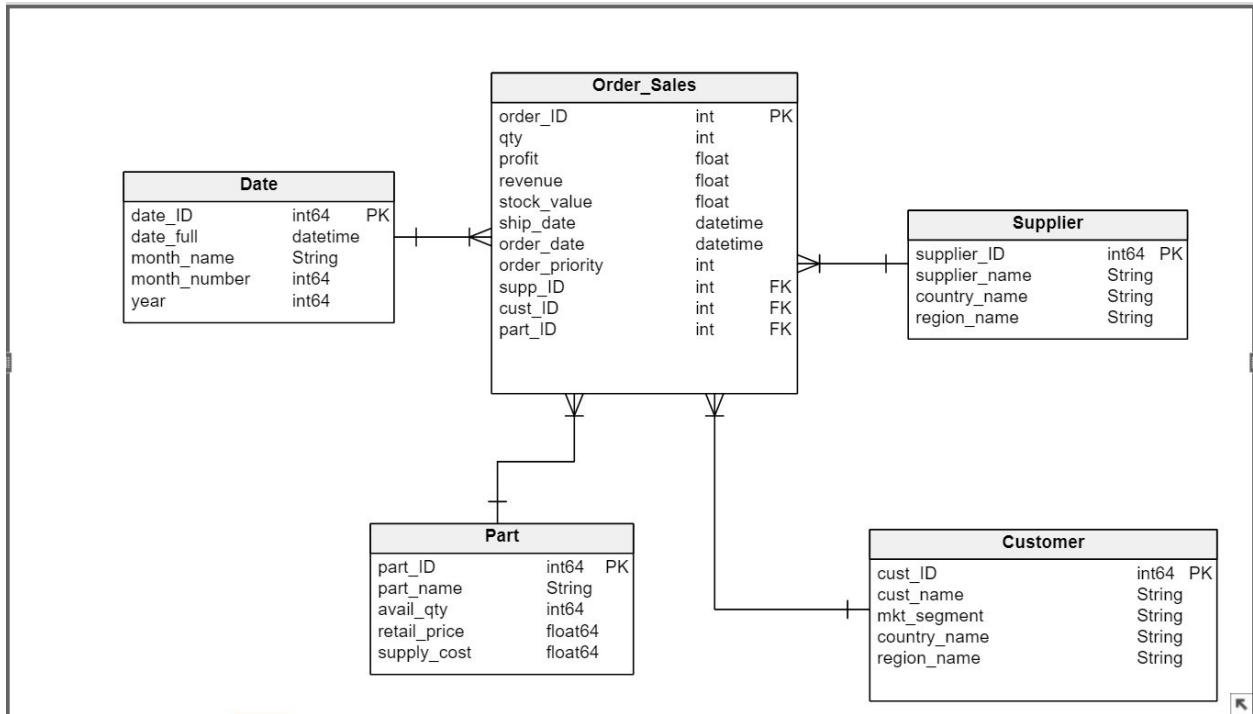
Scott Harris + Noah Dunn

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
```


#### Query editor


```
1 INSERT INTO `mini-project-2-269717.kcd_dw.Part`  
2 SELECT P.partKey, P.name, SUM(PS.availqty) as avail_qty, AVG(PS.supplycost) as supply_cost, P.retailPrice  
3 FROM `advdb-s20.tpch.partsupp` as PS  
4 JOIN `advdb-s20.tpch.part` as P  
5 ON PS.partKey = P.partKey  
6 GROUP BY P.partKey, P.name, P.retailPrice  
7  
8  
9
```


 Valid.

 Run

 Save query

 Save view


 Schedule query

 More

#### Query results

Query complete (3.5 sec elapsed, 286.5 KB processed)

[Job information](#) [Results](#) [JSON](#) [Execution details](#)

 This statement added 2,000 rows to mini-project-2-269717:kcd\_dw.Part.

Query editor

```
1 UPDATE `mini-project-2-269717.kcd_dw.Part`
2 SET part_ID = GENERATE_UUID()
3 WHERE 1 = 1
4 |
```

Valid.

Run

Save query

Save view

Schedule query

More

Query results

Query complete (1.8 sec elapsed, 118.6 KB processed)

Job information

Results

JSON

Execution details

This statement modified 2,000 rows in mini-project-2-269717:kcd\_dw.Part.

```
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 s.suppKey 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;
```

### Query editor

```
1 UPDATE `mini-project-2-269717.kcd_dw.Supplier`  
2 SET supplier_ID = GENERATE_UUID()  
3 WHERE 1 = 1  
4 |
```

 Valid.

Run

Save query

Save view

Schedule query

More

### Query results


Query complete (1.9 sec elapsed, 3.9 KB processed)

Job information

Results

JSON

Execution details

 This statement modified 100 rows in mini-project-2-269717:kcd\_dw.Supplier.

```
UPDATE `mini-project-2-269717.kcd_dw.Supplier`  
SET supplier_ID = GENERATE_UUID()  
WHERE 1 = 1
```

## **Populating Customer Table**

## Query editor

```
1 INSERT INTO `kcd_dw.Customer` (cust_id, cust_name, mkt_segment, country_name, region_name)
2
3 SELECT DISTINCT "" as cust_ID,
4                 c.name,
5                 c.mktsegment,
6                 n.name AS country_name,
7                 r.name AS region_name
8 FROM `advdb-s20.tpch.customer` AS c
9 JOIN `advdb-s20.tpch.nation` AS n ON c.nationkey = n.nationkey
10 JOIN `advdb-s20.tpch.region` AS r ON n.regionkey = r.regionkey;
11
12
13
```

✓ Valid.



Run



Save query



Save view



Schedule query



More

## Query results

Query complete (1.7 sec elapsed, 57.8 KB processed)

Job information

**Results**

JSON

Execution details

**i** This statement added 1,500 rows to mini-project-2-269717:kcd\_dw.Customer.

```
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;
```

## Query editor

```
1 UPDATE `mini-project-2-269717.kcd_dw.Customer`  
2 SET cust_ID = GENERATE_UUID()  
3 WHERE 1 = 1  
4  
5  
6
```

 Valid.



Save query



Save view



Schedule query



More

## Query results


Query complete (2.1 sec elapsed, 74.5 KB processed)

Job information

**Results**

JSON

Execution details


 This statement modified 1,500 rows in mini-project-2-269717:kcd\_dw.Customer.


```
UPDATE `mini-project-2-269717.kcd_dw.Customer`  
SET cust_ID = GENERATE_UUID()  
WHERE 1 = 1
```


## Populating Date Table


### Query editor


```
1 INSERT INTO `mini-project-2-269717.kcd_dw.Date`
2 SELECT 1 as date_ID, DATETIME(o.orderDate),
3 FORMAT_DATETIME("%B",DATETIME(o.orderDate)) as month_name,
4 EXTRACT(MONTH FROM o.orderDate) AS month_number,
5 EXTRACT(YEAR FROM o.orderDate) AS year
6 FROM `advdb-s20.tpch.orders` as o
7 UNION DISTINCT
8 SELECT 1 AS date_ID, DATETIME(li.shipDate),
9 FORMAT_DATETIME("%B",DATETIME(li.shipDate)) as month_name,
10 EXTRACT(MONTH FROM li.shipDate) AS month_number,
11 EXTRACT(YEAR FROM li.shipDate) AS year
12 FROM `advdb-s20.tpch.lineitem` as li
13
```


 Valid.

 Run

 Save query

 Save view


 Schedule query

 More

### Query results

Query complete (1.6 sec elapsed, 587.3 KB processed)

[Job information](#) [Results](#) [JSON](#) [Execution details](#)


 This statement added 2,524 rows to mini-project-2-269717:kcd\_dw.Date.

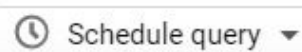
```
INSERT INTO `mini-project-2-269717.kcd_dw.Date`
SELECT "1" as date_ID, DATETIME(o.orderDate),
FORMAT_DATETIME("%B",DATETIME(o.orderDate)) as month_name,
EXTRACT(MONTH FROM o.orderDate) AS month_number,
EXTRACT(YEAR FROM o.orderDate) AS year
FROM `advdb-s20.tpch.orders` as o
UNION DISTINCT
SELECT "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
```



## Query editor

```
1 UPDATE `mini-project-2-269717.kcd_dw.Date`  
2 SET date_ID = GENERATE_UUID()  
3 WHERE 1 = 1  
4
```

 Valid.



## Query results

Query complete (1.5 sec elapsed, 86.6 KB processed)

Job information

**Results**

JSON

Execution details

 This statement modified 2,524 rows in mini-project-2-269717:kcd\_dw.Date.

```
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
    --vvvvvvvvv Joining the Source Tables vvvvvvvvvv
    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

```

```

1 INSERT INTO `mini-project-2-269717.kcd_dw.Order_Sales`
2 SELECT
3 GENERATE_UUID() AS order_id,
4 li.quantity,
5 li.extendedPrice - sp.retailprice AS profit,
6 li.extendedPrice AS revenue,
7 sps.availqty * sps.supplycost AS stock_value,
8 DATETIME(li.shipDate) AS ship_date,
9 DATETIME(so.orderDate) AS order_date,
10 so.orderpriority,
11 s.supplier_ID,
12 c.cust_ID,
13 p.part_ID
14 FROM `kcd_dw.Customer` AS c,
15      `kcd_dw.Part` AS p,
16      `kcd_dw.Supplier` AS s,
17      `advdb-s20.tpch.orders` AS so,
18      `advdb-s20.tpch.customer` AS sc,
19      `advdb-s20.tpch.lineitem` AS li,
20      `advdb-s20.tpch.partsupp` AS sps,
21      `advdb-s20.tpch.part` AS sp,
22      `advdb-s20.tpch.supplier` AS ss ,
23      `advdb-s20.tpch.nation` AS sn1,
24      `advdb-s20.tpch.region` AS sr1,
25      `advdb-s20.tpch.nation` AS sn2,
26      `advdb-s20.tpch.region` AS sr2
27 WHERE
28      --vvvvvvvvv Joining the Source Tables vvvvvvvvvv

```




## Query editor

```

28      --vvvvvvvvv Joining the Source Tables vvvvvvvvvv
29      so.custKey = sc.custKey AND
30      li.suppKey = ss.suppKey AND
31      li.orderKey = so.orderKey AND
32      li.partKey = sp.partKey AND
33      sps.partKey = sp.partKey AND
34      sps.suppKey = ss.suppKey AND
35      ss.nationKey = sn1.nationKey AND
36      sn1.regionKey = sr1.regionKey AND
37      sc.nationKey = sn2.nationKey AND
38      sr2.regionKey = sn2.regionKey AND
39      -- ^^^^^^^ Joining the Source Tables ^^^^^^^
40      p.part_name = sp.name AND
41      s.supplier_name = ss.name AND
42      s.country_name = sn1.name AND
43      s.region_name = sr1.name AND
44      c.cust_name = sc.name AND
45      c.mkt_segment = sc.mktsegment AND
46      c.country_name = sn2.name AND
47      c.region_name = sr2.name

```


Processing location: US

 Run
  Save query
  Save view
  Schedule query
  More

## Query results

Query complete (4.6 sec elapsed, 3.9 MB processed)

[Job information](#)
[Results](#)
[JSON](#)
[Execution details](#)

 This statement added 60,175 rows to mini-project-2-269717:kcd\_dw.Order\_Sales.

# Data Demands

## Q1:

Query editor

```
1 SELECT p.part_name, SUM(os.qty) AS totalQuantity
2 FROM `kcd_dw.Order_Sales` AS os
3 JOIN `kcd_dw.Part` AS p ON os.part_ID = p.part_ID
4 JOIN `kcd_dw.Customer` AS c ON os.cust_ID = c.cust_ID
5 JOIN `kcd_dw.Supplier` AS s ON os.supp_ID = s.supplier_ID
6 WHERE s.country_name = "CHINA" AND
7       c.country_name = "FRANCE"
8 GROUP BY p.part_name;
```

Processing location: US

Run

Save query

Save view

Schedule query

More

Query results

SAVE RESULTS

EXPLORE DATA

Query complete (0.5 sec elapsed, 1.9 MB processed)

Job information

Results

JSON

Execution details

Row	part_name	totalQuantity
1	olive snow puff misty lawn	66
2	dodger navy grey lavender rosy	6
3	violet drab misty peach consilk	45
4	magenta peach metallic pink tomato	41
5	burnished tomato deep chiffon maroon	19
6	azure grey smoke cream ivory	5
7	olive peach pink hot pale	48
8	navajo tan cornflower papaya sienna	46
9	gainsboro cornflower honeydew smoke chocolate	39
10	navy spring azure medium tan	18

Rows per page: 100 1 - 94 of 94

```
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:

### Query editor

```
1 SELECT sup.country_name, part.part_name, SUM(os.profit) as total_profit
2 FROM `mini-project-2-269717.kcd_dw.Order_Sales` AS os
3 JOIN `mini-project-2-269717.kcd_dw.Part` AS part
4 ON os.part_ID = part.part_ID
5 JOIN `mini-project-2-269717.kcd_dw.Supplier` AS sup
6 ON os.supp_ID = sup.supplier_ID
7 WHERE os.ship_date > DATETIME(DATE '1997-11-14')
8 GROUP BY sup.country_name, part.part_name
```

✓ Valid.

[Run](#) [Save query](#) [Save view](#) [Schedule query](#) [More](#)

### Query results

[SAVE RESULTS](#) [EXPLORE DATA](#)

Query complete (0.5 sec elapsed, 1.9 MB processed)

Job information [Results](#) [JSON](#) [Execution details](#)

Row	country_name	part_name	total_profit
1	PERU	olive honeydew misty wheat spring	21893.579999999998
2	PERU	seashell lemon lime thistle purple	12375.019999999999
3	PERU	papaya magenta drab ivory chiffon	39862.68
4	PERU	salmon royal forest white peru	51859.44
5	PERU	orange indian snow sky smoke	40715.71

Rows per page: 100 1 - 100 of 4823

```
SELECT sup.country_name, part.part_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, part.part_name
```

### Q3:

Query editor

```
1 SELECT part.part_name, SUM(os.revenue) as total_revenue, AVG(os.stock_value) as avg_stock_price
2 FROM `mini-project-2-269717.kcd_dw.Order_Sales` AS os
3 JOIN `mini-project-2-269717.kcd_dw.Part` AS part
4 ON os.part_ID = part.part_ID
5 JOIN `mini-project-2-269717.kcd_dw.Supplier` AS sup
6 ON os.supp_ID = sup.supplier_ID
7 WHERE sup.region_name = 'EUROPE'
8 GROUP BY part.part_name
```

Valid.

Run

Save query

Save view

Schedule query

More

Query results

SAVE RESULTS

EXPLORE DATA

Query complete (0.4 sec elapsed, 1.9 MB processed)

Job information

Results

JSON

Execution details

Row	part_name	total_revenue	avg_stock_price
1	drab almond lawn pale thistle	947483.7000000001	1416436.9885714287
2	snow beige seashell peach plum	409482.10999999999	545558.1
3	misty magenta red linen azure	636380.8	1468890.4470588237
4	peach dark yellow white sandy	313531.92	358938.0
5	blush grey salmon white linen	339198.54000000004	278622.0

Rows per page:

100

1 - 100 of 1196

```
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:

Query editor

```
1 SELECT p.part_name, os.order_priority, SUM(os.revenue) as total_revenue
2 FROM `kcd_dw.Order_Sales` AS os
3 JOIN `kcd_dw.Part` AS p ON os.part_ID = p.part_ID
4 JOIN `kcd_dw.Customer` AS c ON os.cust_ID = c.cust_ID
5 WHERE c.mkt_segment = 'BUILDING' AND
6       os.ship_date <= DATETIME(DATE "1992-04-04") AND
7       os.order_date >= DATETIME(DATE "1992-03-11")
8 GROUP BY p.part_name, os.order_priority
```

Processing location: US

Run

Save query

Save view

Schedule query

More

Query results

SAVE RESULTS

EXPLORE DATA

Query complete (0.9 sec elapsed, 6.5 MB processed)

Job information

Results

JSON

Execution details

Row	part_name	order_priority	total_revenue
1	almond pink orchid white chartreuse	3-MEDIUM	30568.56
2	bisque drab brown plum linen	4-NOT SPECIFIED	2034.22
3	sienna khaki plum blue ivory	4-NOT SPECIFIED	13005.65
4	blue lemon spring brown yellow	2-HIGH	46251.76
5	forest hot sandy tan blanchd	2-HIGH	15136.52

Rows per page: 100 1 - 21 of 21

```
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
```