COS20015 – Fundamentals of Data Management

Distinction Report

KATHY WONG HUI YING (101212259)

Table of Contents

Introduction

The purpose of this Distinction Task is to design a database based a real-life application regarding a certain business. Thus, Walmart Inc. was chosen, as it is a corporation with vast enough information to sufficiently allow for a satisfying database design.

Walmart stores were first started by brothers Sam and James “Bud” Walton in 1962, where the first Wal-Mart Discount City is established in Rogers, Arkansas. By 1970, Wal-Mart Stores have gone public with 18 stores and sales worth of $44million.

Currently, Walmart is the world's number one retailer, as well as the world's largest company by revenue and largest employer with over 2.3 million associates. Walmart conducts retail and/or e-commerce operations for selling groceries and general merchandise in 24 countries. The products and merchandise sold are sourced from more than 100 countries.

Overview of the System and Business Rules

Walmart operates through three segments: Walmart US, Walmart International, and Sam's Club, with Walmart US being the largest segment. Products sold are generally in the categories of groceries, health and wellness, and general merchandise.

More than 75% of Walmart US store operations management members started as hourly employees, by end of 2020, more than 200,000 staff were promoted to higher responsibility jobs and pay. The average hourly wage for full- and part-time staff in Walmart US, at the end of 2020 was more than $14.00 per hour.

The Walmart Leadership is composed of two groups: Executive Management and Board of Directors. Each governor has their own role and position in the corporate (i.e President, CEO, Vice President, co-founder and so on). There is also a committee of investors involved with the corporate governance. Each member, like for the governors, also have their own role and position.

This database management system formed is for Walmart Inc. itself as a company (as in the headquarters), and not for the stores of the respective business branches. It is to help the headquarters in analyzing the business performance of each branch for decision making. For example, the company can analyze how well its products are selling by looking at the reorder point— if the product has a high reorder point, then it sells very well, meaning it can generate more income to the store.

Information stored regarding the staff and suppliers should also be able to help determine the diversity of its staff and suppliers.

Therefore, information stored in the database will be regarding: the segments and their branches, current corporate governing leader, staff, suppliers, and products for each branch.

The database is created, with each table being populated with at least ten rows of data. Following this is more detailed information regarding the tables and their attributes.

Entities, Attributes and Description

Walmart Inc has is a company whose operations carry out as three segments: Walmart US, Walmart International, and Sam's Club. Each segment is given a code for easy reference.

WalmartIncSegment		
sgID	CHAR(3)	Segment code
sgName	VARCHAR(20)	Segment name

Currently, the leadership of Walmart Inc consists of three group types: executive management, board of directors and investors. Each group is given a code for easy reference.

GovernanceGroups		
gpID	CHAR(3)	Group code
gpName	VARCHAR(20)	Leadership group name, which are abbreviated: Executive management 'EXM', board of directors 'BOD' and investors 'INV'.

Each segment has its own governance. Those involved are known as the corporate governors, and will change after a few years. Each governor has a position and role.

So, one segment will have one-to-many governors.

CorporateGovernance		
sgID	CHAR(3)	Segment code
gpID	CHAR(3)	Group code
ldPos	VARCHAR(3)	Position names, which are abbreviated: President 'PD', Vice President 'VP', Director 'DIR', Chairman 'CM', Vice Chairman 'VCM', and Shareholder 'SH'
ldRole	VARCHAR(20)	Role (can be NULL)
ldIC	CHAR(7)	Identity card number

Each governor has one record of their details, which are: their identity card number, name, gender and country of origin.

GovernorDetails		
ldIC	CHAR(7)	Identity card number

ldName	VARCHAR(30)	Name
ldGend	CHAR(1)	Gender, which is abbreviated: male ‘M’, female ‘F’, unknown or other ‘U’
ldCountry	CHAR(3)	ISO country code

Each Walmart Inc. segment has one-to-many branches.

As the stores are named after their segments, each branch is numbered for easy identification. The branches can be located in different states in different countries, and each has its own local address.

SegmentBranches		
sgID	CHAR(3)	Segment code
brID	CHAR(4)	Branch number
brCountry	CHAR(3)	ISO country code of branch
brZIP	VARCHAR(10)	ZIP code
brAdd	VARCHAR(40)	Branch local address

Each segment has one-to-many staff workers.

One staff worker can only work at one branch at a time/per contractual period. In addition to that, each worker also has their own position in the store, and are either part-time or full-time workers. They are also entitled to a salary.

StaffWorker		
wIC	CHAR(7)	Identity card number
sgID	CHAR(3)	Segment code
brID	CHAR(4)	Branch number
wPos	VARCHAR(3)	Staff worker positions, which are abbreviated: supervisor ‘SPV’, manager ‘MAN’, sales clerk ‘CLK, cashier ‘CSH’ and janitor ‘JAN’
wType	CHAR(2)	Part-time ‘PT’ or full-time ‘FT’ working type
wSalPerHour	DOUBLE(4,2)	Staff worker salary (per-hour)

Each staff worker has one record of details. Just like for the leaders, the details are: their identity card number, name, gender and country of origin. Each staff worker is also to be insured, as according to Walmart policies.

StaffWorkerDetails		
wIC	CHAR(7)	Identity card number
wName	VARCHAR(30)	Name
wGend	CHAR(1)	Gender, which is abbreviated: male ‘M’, female ‘F’, unknown or other ‘U’
wCountry	CHAR(3)	ISO country code
wIns	VARCHAR(20)	Insurance plan

Each segment branch has a record of the products it sells. The record contains: the product price, and the reorder point of the products. A high reorder point indicates a product is selling well.

So, one segment branch will have one-to-many product details.

BranchProductDetails		
sgID	CHAR(3)	Segment code
brID	CHAR(4)	Branch number
pEAN	CHAR(13)	EAN barcode number
brPrice	DOUBLE(6,2)	Product pricing
brReorder	INT(3)	Reorder point

One supplier can supply more than one type of product, at more than segment branch.

The products found at each branch are known by the following information: its category as decided by Walmart, and who supplies it.

Products		
pEAN	CHAR(13)	EAN barcode number
pCat	CHAR(3)	Product category names, which are abbreviated: groceries ‘GRO’, health and wellness ‘HNW, general merchandise ‘GEN’
sID	CHAR(8)	Supplier ID

Each product has one product detail.

The following details are known regarding the product: their unique EAN code (barcode number), the product name, and if it is edible, whether it is Halal or not.

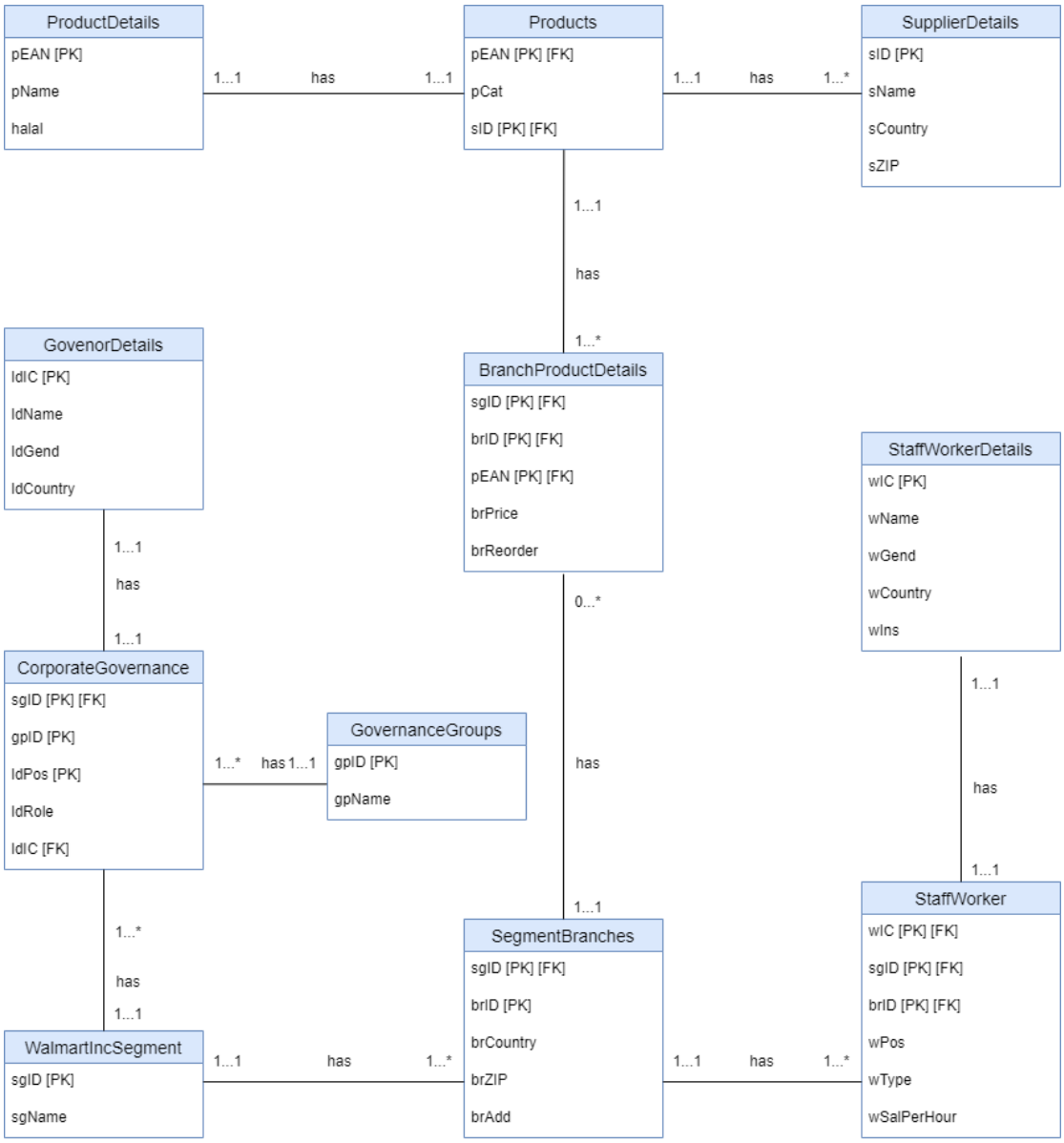
ProductDetails		
pEAN	CHAR(13)	EAN barcode number
pName	VARCHAR(20)	Name
halal	CHAR(1)	Halal status, which are abbreviated: 'Y' yes, certified; 'N' no, not Halal certified, 'X' not applicable

Regarding the suppliers, because Walmart Inc is operating on a world-wide scale, where many countries are taken into consideration, it is possible there could be suppliers coincidently having the same name. Each supplier is thus is recorded with a unique supplier ID, along with their name, located country and state ZIP code.

Each supplier has one supplier detail.

SupplierDetails		
sID	CHAR(8)	Supplier code
sName	VARCHAR(30)	Name
sCountry	CHAR(3)	ISO country code
sZIP	VARCHAR(10)	ZIP code

Entity Relationship Diagram

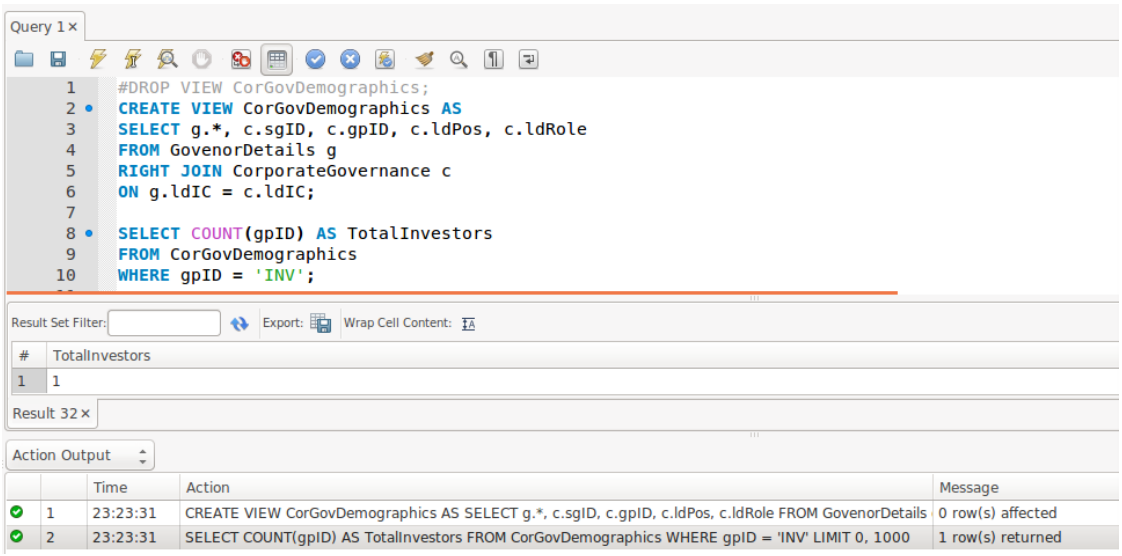


Use Cases

The ten use cases below shows the kind of information that can be retrieved from the database formed for Walmart Inc.

Use Case 01

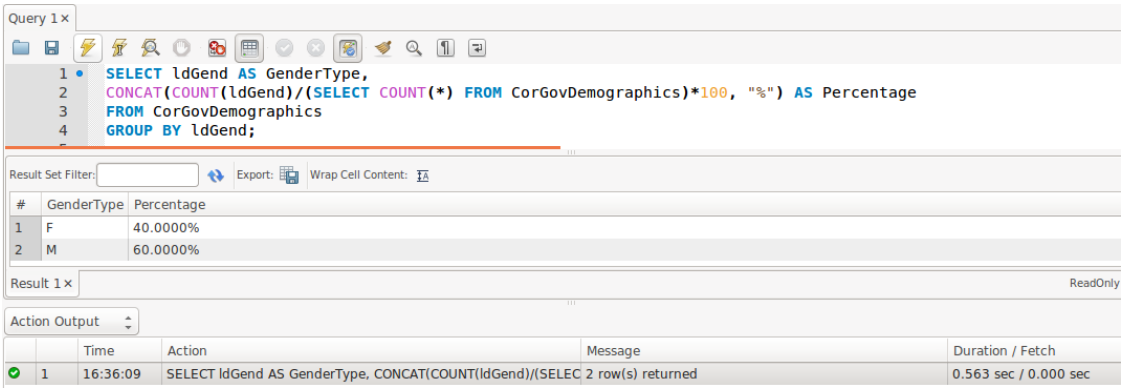
To find how many leaders of Walmart Inc are from a certain group. This helps to determine the number of corporate governances currently involved with the decision-makings in Walmart Inc. This shows how many investors are currently investing in Walmart:



Creating CorGovDemographics View	Query Statement
#DROP VIEW CorGovDemographics; CREATE VIEW CorGovDemographics AS SELECT g.*, c.sgID, c.gpID, c.ldPos, c.ldRole FROM GovenorDetails g RIGHT JOIN CorporateGovernance c ON g.ldIC = c.ldIC;	SELECT COUNT(gpID) AS TotalInvestors FROM CorGovDemographics WHERE gpID = 'INV';

Use Case 02

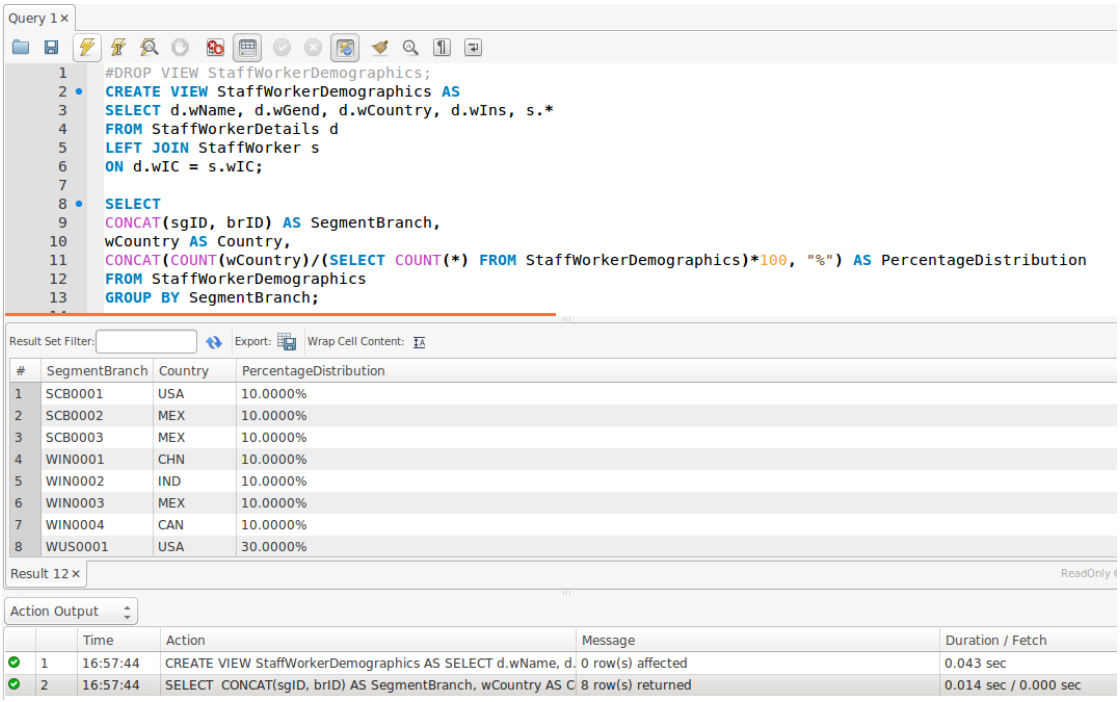
To find out the percentage distribution of corporate governance genders. This can help the headquarters keep track of the gender diversity in the leadership of Walmart.



Query Statement
SELECT ldGend AS GenderType, CONCAT(COUNT(ldGend)/(SELECT COUNT(*) FROM CorGovDemographics)*100, "%") AS Percentage FROM CorGovDemographics GROUP BY ldGend;

Use Case 03

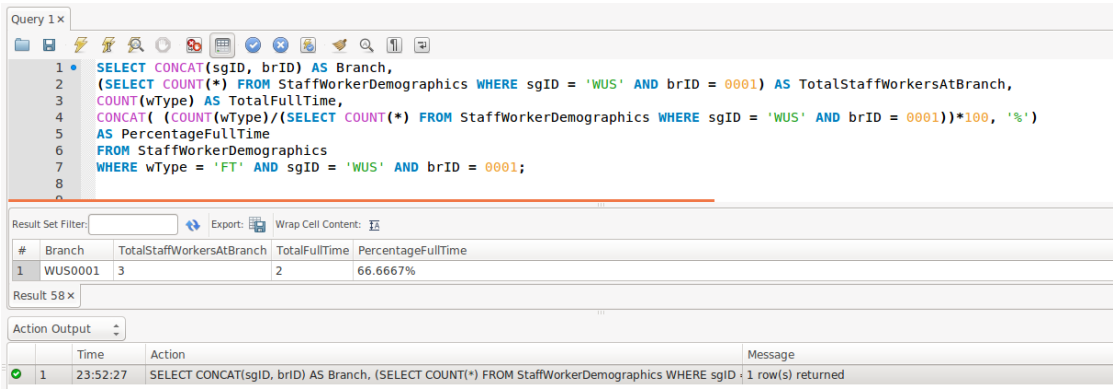
To find out the percentage distribution total staff workers across the segment branches. This can help the headquarters keep track of the diversity of staff workers across different branches of Walmart. This will show the percentage nationality of workers from each branch:



Creating StaffWorkerDemographics View	Query Statement
#DROP VIEW StaffWorkerDemographics; CREATE VIEW StaffWorkerDemographics AS SELECT d.wName, d.wGend, d.wCountry, d.wIns, s.* FROM StaffWorkerDetails d LEFT JOIN StaffWorker s ON d.wIC = s.wIC;	SELECT CONCAT(sgID, brID) AS SegmentBranch, wCountry AS Country, CONCAT(COUNT(wCountry)/(SELECT COUNT(*) FROM StaffWorkerDemographics)*100, "%") AS PercentageDistribution FROM StaffWorkerDemographics GROUP BY SegmentBranch;

Use Case 04

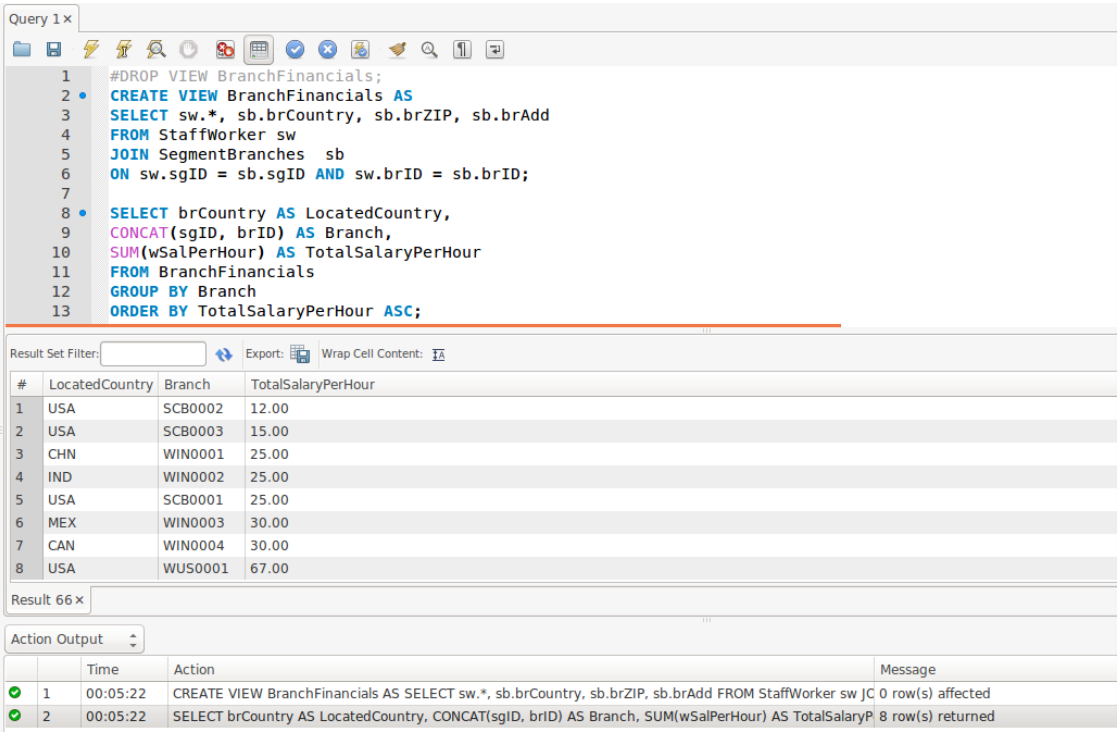
To find out the percentage of staff workers type at a segment branch. If a branch is not doing particularly well, the headquarters can consider decreasing the number of full-time workers, as full-time workers cost more than part-time workers. Or, if the branch is doing well, they can increase in hiring more full-time workers. This shows what percentage of staff workers from WUS0001 are full-timers:



Query Statement
SELECT CONCAT(sgID, brID) AS Branch, (SELECT COUNT(*) FROM StaffWorkerDemographics WHERE sgID = 'WUS' AND brID = 0001) AS TotalStaffWorkersAtBranch, COUNT(wType) AS TotalFullTime, CONCAT((COUNT(wType)/(SELECT COUNT(*) FROM StaffWorkerDemographics WHERE sgID = 'WUS' AND brID = 0001))*100, '%') AS PercentageFullTime FROM StaffWorkerDemographics WHERE wType = 'FT' AND sgID = 'WUS' AND brID = 0001;

Use Case 05

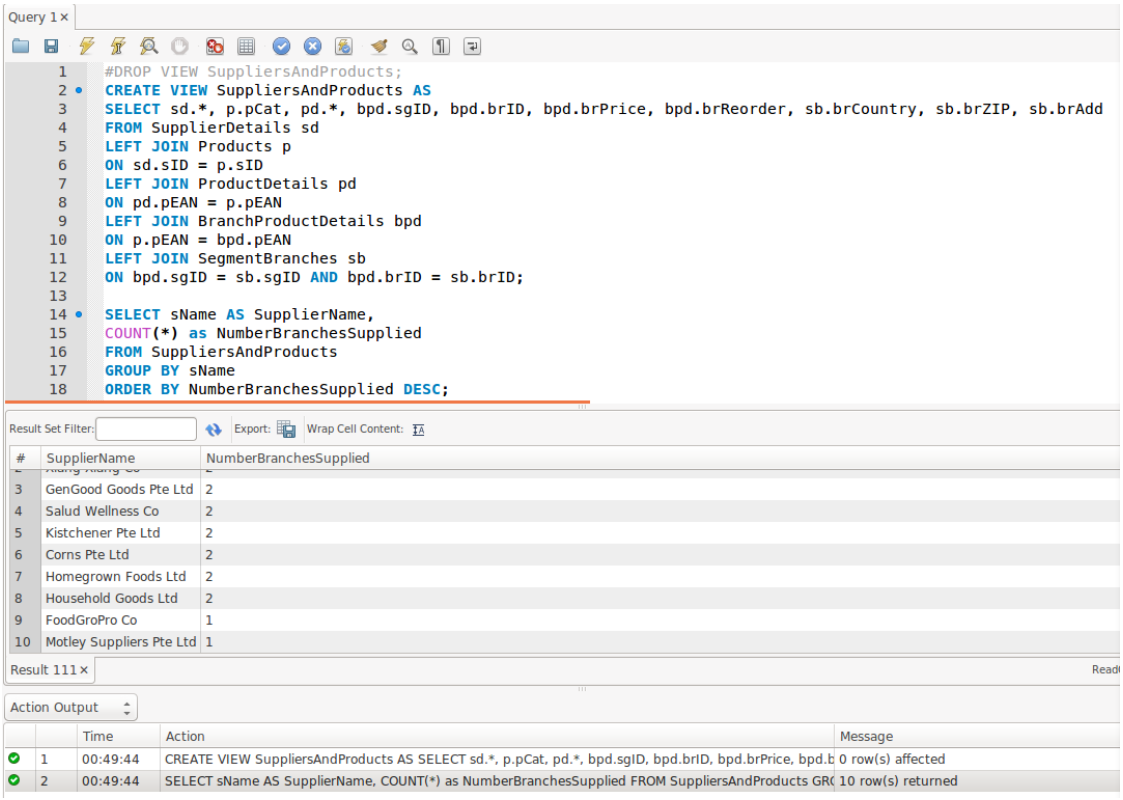
To count how much salary each branch would pay their staff workers (as by a per hour basis). This would help the headquarters analyze the outflow of financials of each branch. If the financial performance of a certain branch is low, but the total salary paid per hour is high, they can consider decreasing the salary value.



Creating BranchFinancials View	Query Statement
#DROP VIEW BranchFinancials; CREATE VIEW BranchFinancials AS SELECT sw.*, sb.brCountry, sb.brZIP, sb.brAdd FROM StaffWorker sw JOIN SegmentBranches sb ON sw.sgID = sb.sgID AND sw.brID = sb.brID;	SELECT brCountry AS LocatedCountry, CONCAT(sgID, brID) AS Branch, SUM(wSalPerHour) AS TotalSalaryPerHour FROM BranchFinancials GROUP BY Branch ORDER BY TotalSalaryPerHour ASC;

Use Case 06

To find out how many branches each supplier supplies to. As Walmart has many suppliers, this will determine what supplier is 'common' to most branches. The headquarters then can consider establishing good relations with that supplier, which can maybe bring in more benefits to Walmart (i.e special discount price for certain supplied goods):

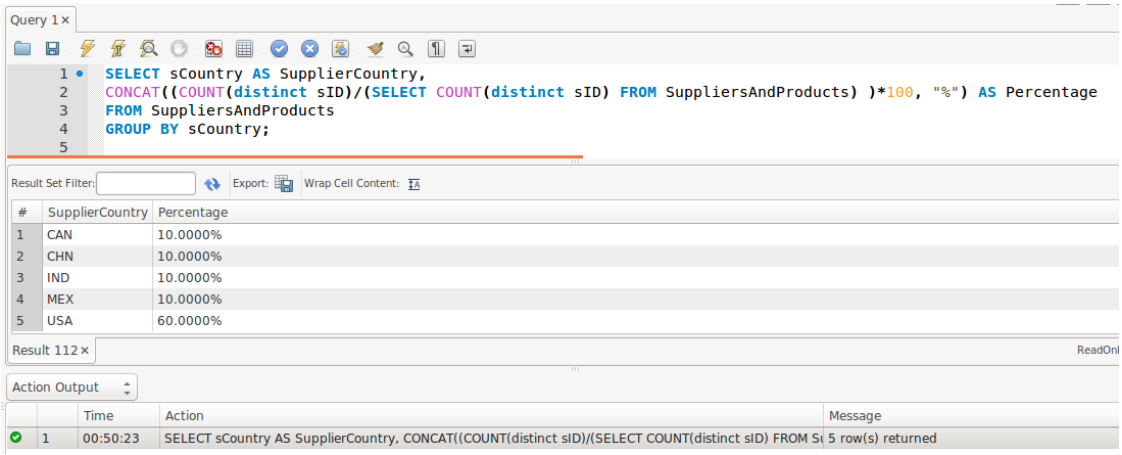


Creating SuppliersAndProducts View	Query Statement
#DROP VIEW SuppliersAndProducts; CREATE VIEW SuppliersAndProducts AS SELECT sd.*, p.pCat, pd.*, bpd.sgID, bpd.brID, bpd.brPrice, bpd.brReorder, sb.brCountry, sb.brZIP, sb.brAdd FROM SupplierDetails sd LEFT JOIN Products p ON sd.sID = p.sID LEFT JOIN ProductDetails pd ON pd.pEAN = p.pEAN LEFT JOIN BranchProductDetails bpd	SELECT sName AS SupplierName, COUNT(*) as NumberBranchesSupplied FROM SuppliersAndProducts GROUP BY sName ORDER BY NumberBranchesSupplied DESC;

ON p.pEAN = bpd.pEAN LEFT JOIN SegmentBranches sb ON bpd.sgID = sb.sgID AND bpd.brID = sb.brID;	

Use Case 07

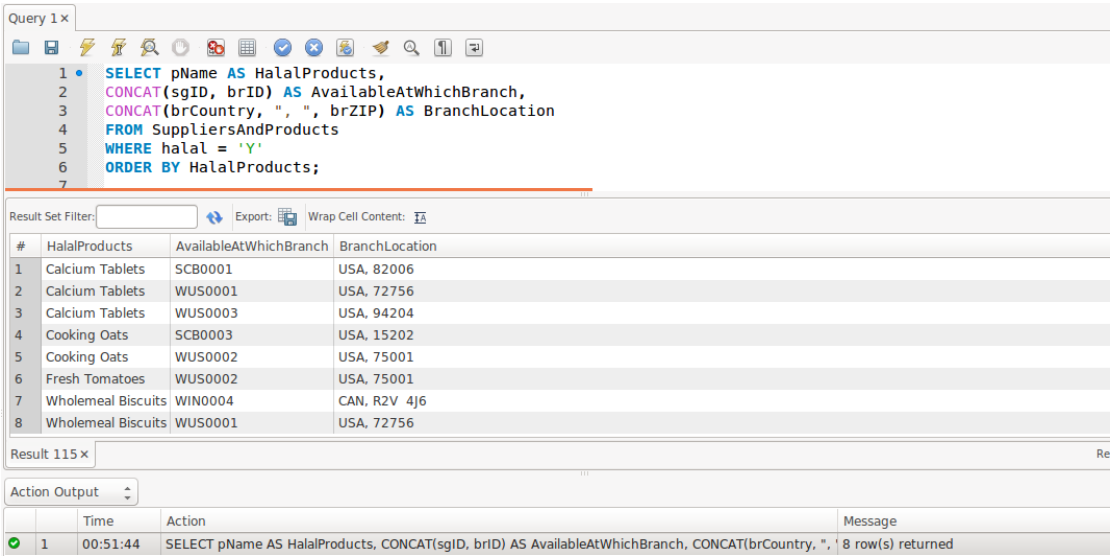
To find out the percentage diversity of Walmart's suppliers. This can help the headquarters determine which country are most of Walmart's suppliers from.



Query Statement
SELECT sCountry AS SupplierCountry, CONCAT((COUNT(distinct sID)/(SELECT COUNT(distinct sID) FROM SuppliersAndProducts))*100, "%") AS Percentage FROM SuppliersAndProducts GROUP BY sCountry;

Use Case 08

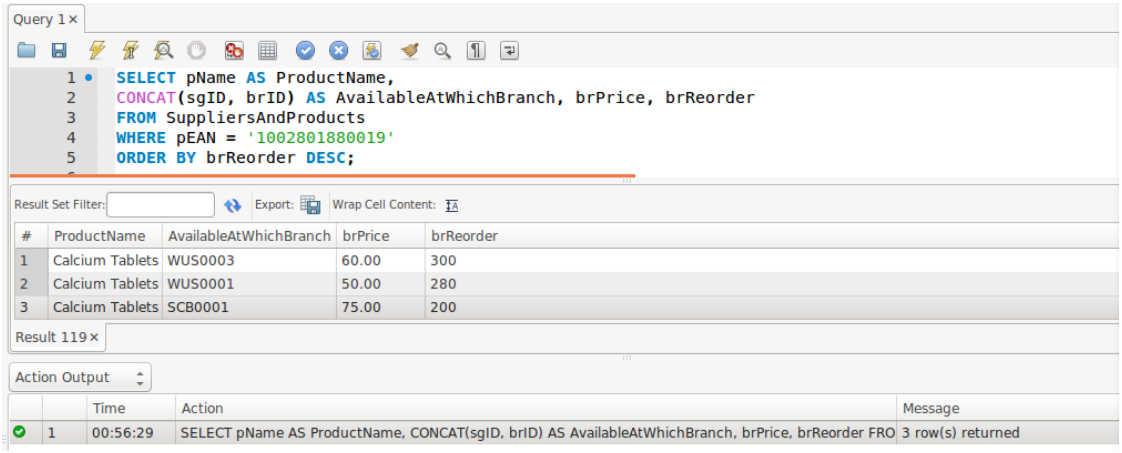
To find out which products supplied are Halal and what branches they are available at. This gives the headquarters a record of products that are suitable for consumption for certain religious groups, in order to ensure a bigger customer base.



Query Statement
SELECT pName AS HalalProducts, CONCAT(sgID, brID) AS AvailableAtWhichBranch, CONCAT(brCountry, '', '', brZIP) AS BranchLocation FROM SuppliersAndProducts WHERE halal = 'Y' ORDER BY HalalProducts;

Use Case 09

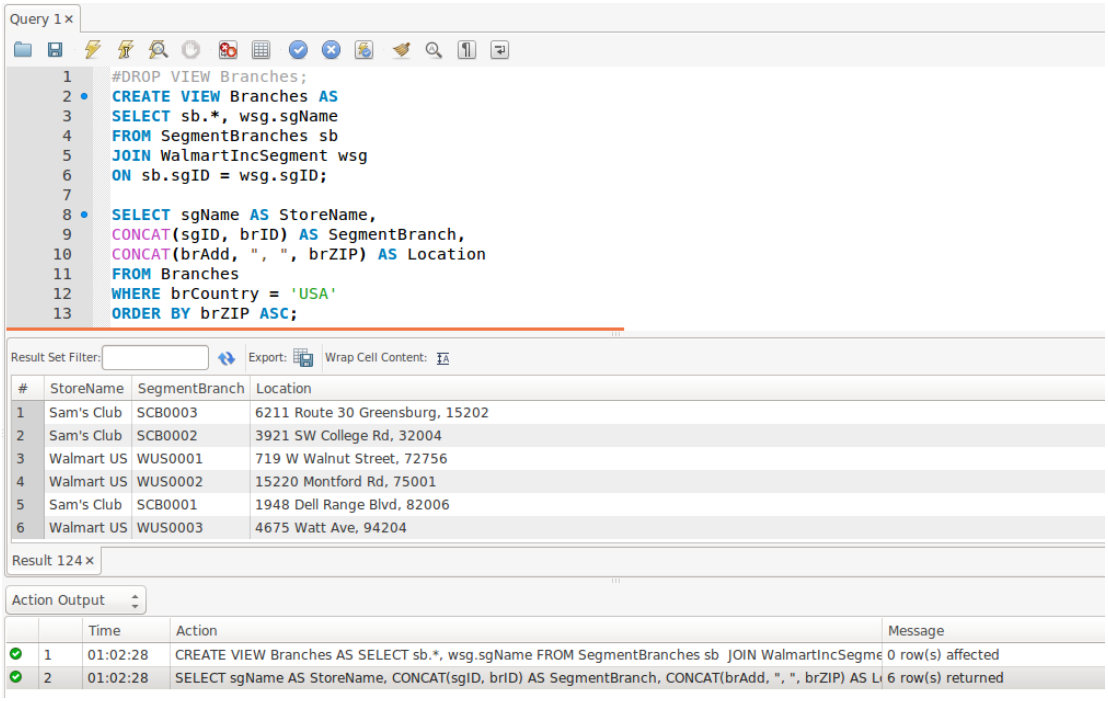
To compare the reorder points and their prices of a certain product sold across branches. This would help headquarters determine how well each product is selling in each branch (if the reorder point is high, this means the product would sell out quite fast). If the reorder point of a product is low, but its selling price is high, then the headquarters can consider reducing the price in order to attract more buyers. For example, the product 'Calcium Tablets' of EAN code 1002801880019 is searched to compare how well it is selling across the branches:



Query Statement
SELECT pName AS ProductName, CONCAT(sgID, brID) AS AvailableAtWhichBranch, brPrice, brReorder FROM SuppliersAndProducts WHERE pEAN = '1002801880019' ORDER BY brReorder DESC;

Use Case 10

To find out the branches in a country and their details. This helps the headquarters keep track of where Walmart has expanded to in a certain country. For example, to find the branch locations of the stores in USA:



Creating Branches View	Query Statement
#DROP VIEW Branches; CREATE VIEW Branches AS SELECT sb.*, wsg.sgName FROM SegmentBranches sb JOIN WalmartIncSegment wsg ON sb.sgID = wsg.sgID;	SELECT sgName AS StoreName, CONCAT(sgID, brID) AS SegmentBranch, CONCAT(brAdd, '', '', brZIP) AS Location FROM Branches WHERE brCountry = 'USA' ORDER BY brZIP ASC;

Conclusion

In conclusion, a database for Walmart Inc has been designed, and in it contains eleven tables to store the relevant information required by the headquarters. These tables are then populated with at least ten rows of data, in order to be used in demonstration of how information can be queried and retrieved from these tables. All in all, the queries made have been shown to successively return expected results.

One thing I have discovered during the course of working on this task was that the datatype INT(n) has a range of values that it can accommodate. If this range is exceeded, it will return an ‘out-of-range’ error. Thus, for declaring an attribute such as ID, which will require a random sequence of numbers, instead of declaring the datatype as INT(n), it would be safer to use CHAR(n) to ensure any random sequence of numbers input will not be met with these kinds of error.

Through this Distinction Task, I was able to put to use what was learnt during lectures to develop a database based on a real-life scenario, and speculate the type of reports that will be expected by the users regarding the given data stored. Through this, whatever knowledge learnt was aptly applied in this task, and it has been quite an enjoyable experience.