ITMD 523 *Advanced Topics in Data Management* HW 2

Student Name **Henry Post** Section **422**

Instructor ***Luke Papademas*** Due Date **09/21**

Part	1	2	3	4	**TOTAL**	**Score**
Maximum Points	25 points	25 points	25 points	25 points	**100** points	

Textbook Reading Assignment Thoroughly read Week 1 - 4 course lecture notes and course textbook Chapter(s) 1 - 4 .

Part 1 Concept Check - Advanced Topics in Data Management

(1) (Data Models and Business Rules)

Consider the following business facts that a retail company requires to represent in its data model:

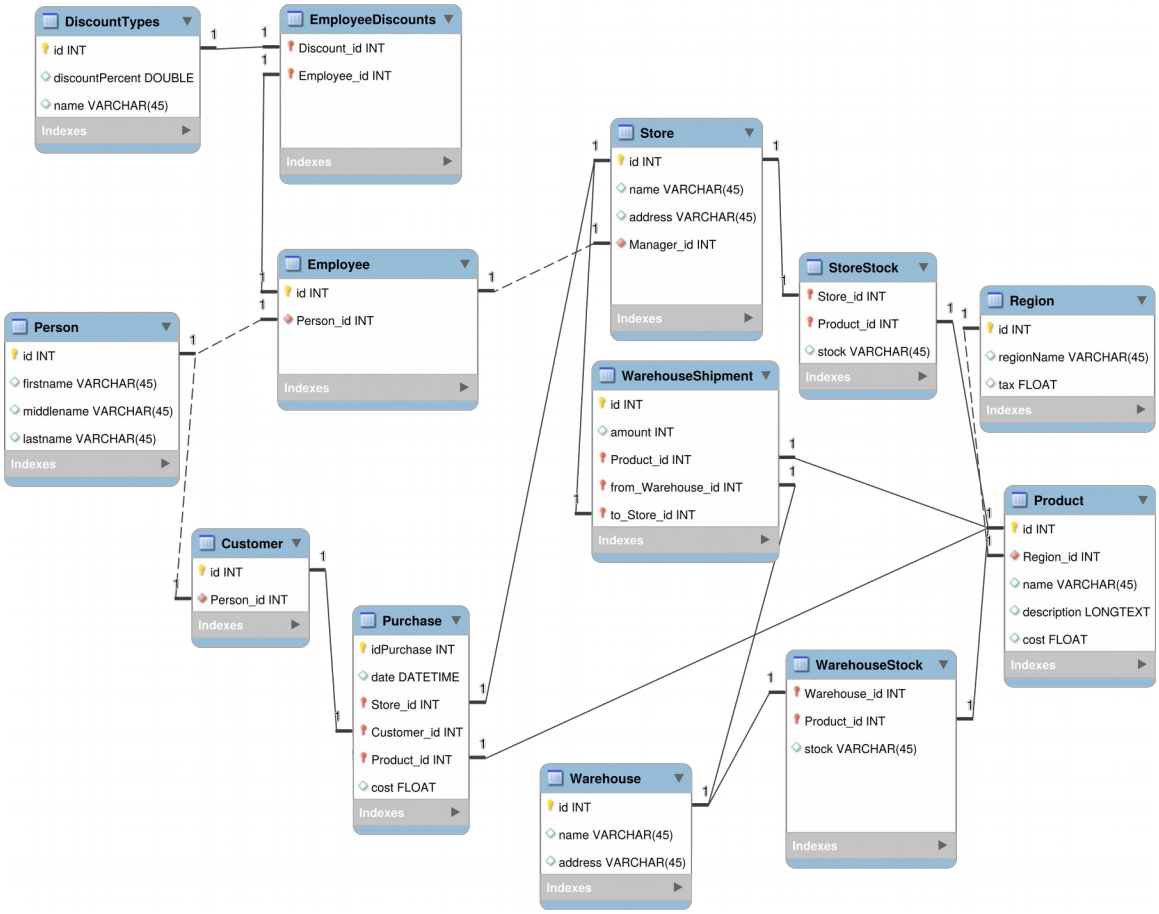
- customers purchase products and goods
- products are subject to various types of sales / county taxes
- retail store members receive special discounts
- products are stored in both the retail stores and in the warehouse
- products can be shipped from the retail warehouse directly to the customer

Based on the above business rules, construct a data model in the form of a database table that can be used to provide information as to the transactional operation of the company.

Customer <purchase> Product
Product <taxed by> RegionTax
Employee <receives> Discount
Products <stored in> Store, Warehouse
Products <shipped to> Customer <from> Store

See below for diagram.

Student Name **Henry Post** Section **422**



Student Name **Henry Post** Section **422**

(2) (Sources of Business Rules)

Business rules are essential to create an accurate data model.

Within an enterprise, the sources of business rules include these categories:

- Company Managers
- Department Managers
- Direct interviews with End Users
- Written documentation Policy Makers

Reflect on the business rules of a typical employee staffing agency. Choose one of the above business rule sources and, from the point of view of your source, discuss five important business rules that would be critical for the agency's data model.

When you conduct interviews with end users, you need to know what person was the POC (point of contact) with that end user. This is very important, as well as recording everything that was said and general impressions.

When PMs (Policy Makers) make policies, they should be able to easily determine who will be affected by their policies. This is critical to ensuring that all end users and business groups are properly informed about requirements that their policies have. Without this, policies that PMs make can be created, deployed, and enforced without the knowledge of the people that they affect.

When an interview is conducted, if there is a problem that a user identifies that affects the quality of their experience with the product significantly, then this must be resolved within a specific timeframe. If you offer support for this product (i.e. a blender), then the person responsible for resolving that issue must track and resolve that problem.

1. End user interviews must include a POC and impression details, as these details are very important to get useful stories from users.

2. PMs must be able to easily specify the business groups and end user groups that their policies affect.

3. PMs must be able to compare the differences between policies and who they affect.

4. Issues that users identify must be tracked and resolved within a timely manner. If an escalation is required (i.e. ship a new product, lawsuit threats, etc), then that must be tracked as well.

5. Company Managers, CEOs, Department Managers, and all other

managers must be able to see who they manage and who manages them.

Part 2 DBMS Concepts - Advanced Topics in Data Management

(1) (Business Rules)

Business rules define one or more of the following modeling components:

- entities
- relationships
- attributes
- connectivity
- cardinalities

Some examples of business rules are:

- An invoice contains one or more invoice lines.
- Each invoice line is associated with a single invoice.
- A store employs many employees.

List FIVE other examples of business rules from various forms of businesses.

A store contains many products.

A manager manages multiple employees.

A customer can perform a purchase of one or more products.

A customer can submit a complaint about one or more products to a store.

A customer can return a purchase involving a product to a store.

Student Name **Henry Post** Section **422**

(2) (Types of Relationships)

Give an example of each of the three types of relationships: 1:1 , 1:M , M:N

Why is an M:N relationships not appropriate in a relational model?

A 1-1 relationship could be one PURCHASE performed by one CUSTOMER.

(I buy one tire,
I buy 3 m&ms,
etc)

A 1:M relationship could be one PERSON owning many PETS.

(I own my cat,
I own my dog,
My dad owns his goldfish)

A M:N relationship could be many people living at many houses.

Example:
(I live at my condo,
I live at my parent's house,
I live at my boyfriend's condo
my boyfriend lives at his condo,
my boyfriend lives at my condo,
my boyfriend lives at his parent's house)

Student Name **Henry Post** Section **422**

Part 3 Data Models / Analytics - Advanced Topics in Data Management

(1) (Data Modeling)

Given the following group of tables, list FIVE written business rules that are reflected in the tables and their contents.

Artists

ArtistNum	LName	FName	MI
100	Jennings	Joyce	D
101	Matthews	Shruti	F
102	Sims	Silas	
103	Dean	Daisy	H

Galleries

GalleryNum	GalleryName	ZipCode
22	Lawrence ArtWorks	60625
25	Cragin Corners	60651
27	Bronzeville Galleries	60616
28	Harlem HangUps	60634

Paintings

PaintNum	Title	ArtistNum	GalleryNum
100	Morning	100	22
101	Outer Space	100	25
102	The Poet	101	28
103	Jane	102	28

One artist can paint one or more paintings.
(PaintNum, ArtistNum)

One painting can be stored in one gallery.
(PaintNum, GalleryNum)

One artist can have one painting in one gallery.
(PaintNum, GalleryNum, ArtistNum)

Student Name **Henry Post** Section **422**

An artist can have a first, middle, and last name.
(**Artists**)

A gallery can have a name and a zip code.
(**Galleries**)

(2) (Data Analytics / Predictive Analytics: The Standard Deviation)

In Oracle and SQL the standard deviation is used as an aggregate group function as well as an analytical function. The following data illustrates company sales for the first half of the year. For the data below, compute both the population standard deviation and the sample standard deviation.

tblValues Table

index	month	values
1	January	$20
2	February	$26
3	March	$28
4	April	$31
5	May	$35
6	June	$30

Student Name **Henry Post** Section **422**

Hints: examine the MS Access SQL statements given below or use equivalent Oracle statements.

Sample Standard Deviation

```
SELECT STDEV(values) FROM tblValues;
```

Population Standard Deviation

```
SELECT STDEVP(values) FROM tblValues;
```

Population stdev = 4.64279609239471

Sample stdev = 5.08592829940284

deviation	squared dev	sum of squared devs	variance	Population stdev
				4.64279609
-8.3333	69.4444	129.3333	21.5556	239471
-2.3333	5.4444			
			sample variance	sample stdev
-0.3333	0.1111			5.08592829
2.6667	7.1111		25.8667	940284
6.6667	44.4444			
1.6667	2.7778			

Student Name **Henry Post** Section **422**

Part 4 Data Design Concepts - Advanced Topics in Data Management

(1) (Database Models)

Compare and contrast the hierarchical and network database models. Which of these database models was historically implemented first? List some advantages and disadvantages of each of these models. Provide some examples of types of enterprises that would favor hierarchical database models as well as some types of enterprises that are best suited for network database models.

Hierarchical Models:
Advantages:
- Rigidly structured
- Can model with directories and flat files
- Promotes data sharing
- Is efficient with 1:M relationships

Disadvantages:
- Queries can be very complicated
- You cannot easily get a parent's information without an expensive query (information is not usually stored in the child)

Hierarchical Models were implemented before Network Models. They are suited for very rigidly structured data that can be stored in an n-ary tree format such as the military chain of command or a manager/employee hierarchy.

Network Models:
Advantages:
- Simpler conceptually than Hierarchical models
- Data access is more flexible
- There is still conformance to standards despite being more flexible

Disadvantages:
- The system can become inefficient to modify because all objects must cascade the data getting updated upwards/downwards
- Not structurally independent
- Has some drawbacks of hierarchical models, but not all.

Network models are suitable for somewhat rigidly structured data, perhaps a centralized operation that still has 1-1 or 1-n relationships.

Student Name **Henry Post** Section **422**

(2) (Data Models and eCommerce)

Design a Data Model for a database application on behalf of an Office Administrative Services enterprise. This type of enterprise forms the backbone of business operations across a variety of industries, such as physicians' offices, rental equipment firms, food services by providing them the day - to - day administrative services including record keeping, financial planning and billing.