

## **Technical Paper: Database Administrator for Department Store**

**Introduction:**

This paper will talk about the meaning of the potential sales, and it can be measured by calculating the amount of the consumer sales. The potential sales have many advantages such as the evaluating the information performance. Also, the paper will demonstrate the transactions that are the core of the DBMS; and when the transactions are founded in the retail stores. The database design has many components: Entities, Attributes, and relationships among the entities. Every entity has its attributes, and there are relationships types, such as 1: 1, 1: M, and M: N. Every one of these relations has its particular symbol that you will see its drawing in this paper; with the representing of all database components of the store. There are many types of models to design the database, and one of these models is the Crow's Foot notation that has the relationship name that is written above of the relationship line; the relationship line has two different sides that represent the relationship type. This paper will have the design of the Crow's Foot notation model of the store. Then, it will explain what the big data model is and how it would be used to be the solution of the tremendous grown of the data in these days. Next, I will represent the SQL Stored Procedures that use SQL functions to help sales associates perform explanatory or predictive analytics; with all the screen shots of the SQL programming. The cloud computing is the new software technique that permits a large number of users the corporate usage of the computing resources. It has many trends; therefore it has various pricing considerations with different security phases. This technique of the cloud computing service has options of Software as a Service such as the Platform as a Service and Infrastructure as a Service, which permits the user to use the software components from many far locations. Next, the paper will explain about the distributed DBMS, and it is appropriate for the store database, but with more attention to the

privacy of the database. After that, I will provide the meaning of the lost updates and uncommitted data with examples of how they may occur with the mulch transactions at the same attributes. Then, how we can control the concurrency problem by the many approaches. Lastly, the paper will explain about the database security; what is the disaster mean; how we can manage the database through the disaster; what are the types of database threats; and what is the threats countermeasures.

Firstly, according to the author (**Havens, et al.,1977**), the potential sales can be measured by calculating some consumer sales. The advantage of computing the potential sales is evaluating the performance of the information, to develop the market program. That is, it can be determined the leakage and the surplus of the sales in the market location. If the actual sales more than the potential sales, then there are weak sales. In contrast, if the actual sales are less than the potential sales, then there are the excess sales. These measurements are helping the new market program to be developed. Therefore, the potential sales refer to "the maximum sales opportunities for the product by a particular company". (**Havens, 1977, pg# 574**).

The author (**JARKE et al., 1984**) explained the transactions as the essential core of the business information system; therefore the transactions processing must be efficient and safe. The database management system DBMS supports the transactions processing that is supporting the processing analysis. In the business, there are a high number with a low variety of operations that is governed by the business transactions processing rules. This transactions processing must be parallel. The transactions in the market begin when the customer brings some items to the counter employer to purchase it. These items are entered into the computer program, by scan them by the barcode. Then, the customer pays the money by entering the card into the money

machine that is recorded in the program. All these items names, costs, and date are stored in the transactions table that produced the bill. While the department and other products information are stored in another transactions table. The bill is printed, and the transactions processes are completed. "Essentially, the business program defines (planned) processes together with a script that defines the relationship between these processes. A planned process can be further refined into sub-processes; subatomic processes are called activities. The distinction between processes and activities is left to the discretion of the system designer. Processes of the same type may occur in multiple higher-level processes." (JARKE et al., 1984, pg# 65).

Secondly, according to the opinion of the author (Cornel et al., 2013), the database design composed of the entities (which are the objects, such as the place, location, person, etc.). Every entity has some attributes (which are the characteristics of the entity, example: the Customer (entity) has the attributes customer name, customer number, etc.). Also, there are many relationships among the entities, such as one to one (1: 1), one to many (1, M), many to many (M: N) relationships. Here, in this database there are many entities, attributes, and relationships, as the following:

The entities: **Customer**: the person who buy the product from the store. **Vendor**: who is the employee works in the store to sell the products. **Store**: it is the location (place) has the products that are be purchased. **Product** is the thing that is purchased by the customers. **Sales** are the process of purchasing. **The employee** is the staff deal with the customers. Each entity has many attributes as the following:

ENTITIES	CUSTOMER	VENDOR	EMPLOYEE	SALES	PRODUCT	STORE
	CUS_ID	VEN_NUMBER	EMP_NUMBER	SA_NUMBER	PRO_NUMBER	ST_NUMBER
ATTRIBUTES	CUS_NAME	VEN_NAME	EMP_NAME	SA_NAME	PRO_NAME	ST_NAME
		VEN_SEX	EMP_SEX	SA_DATE	PRO_TYPE	ST_ADDRESS
		VEN_AGE	EMP_AGE		PRO_EXPIREDATE	ST_TYPE

There are many roles that house the sales transactions, such as the following: **Rule 1:** There are five stores that sell the products. All these store will be the fields in the database, and it stored with its information and it must be distinguished by its primary key such as the number of the store. **Rule 2:** each vendor must sell at least one product to the at least one customer. (Cornel et al., 2013).

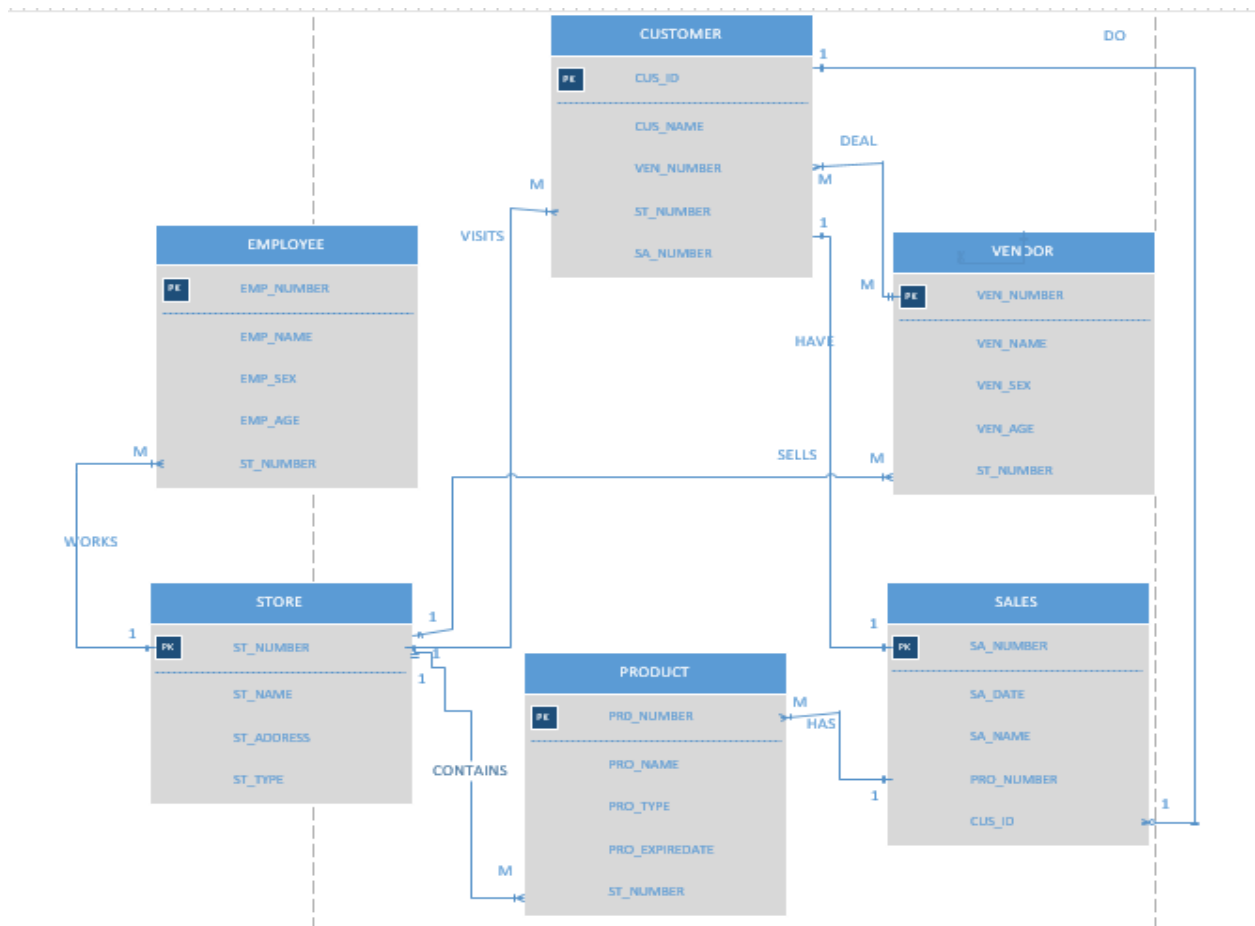
Thirdly, According to the author (Cornel et al., 2013), The Crow's Foot notation is the model that is used to describe the relationships among the entities. The relationship name is written above of the relationship line, and the relationship line has two different sides that represent the relationship type, 1: 1, 1: M, and M: N.

————— ∅< Zero or more

————— < One or more

————— # One and only one

————— ∅+ Zero or one



In this diagram there many relationships types:

1 and only one customer DOES zero or more sales.

1 and only one sales HAVE 1 and only one customer.

1 and only one vendor DEALS with one or more customers.

1 or more customers VISIT 1 and only one store.

1 or more vendors SELL in 1 and only one store.

1 or more employees WORK in 1 and only one store.

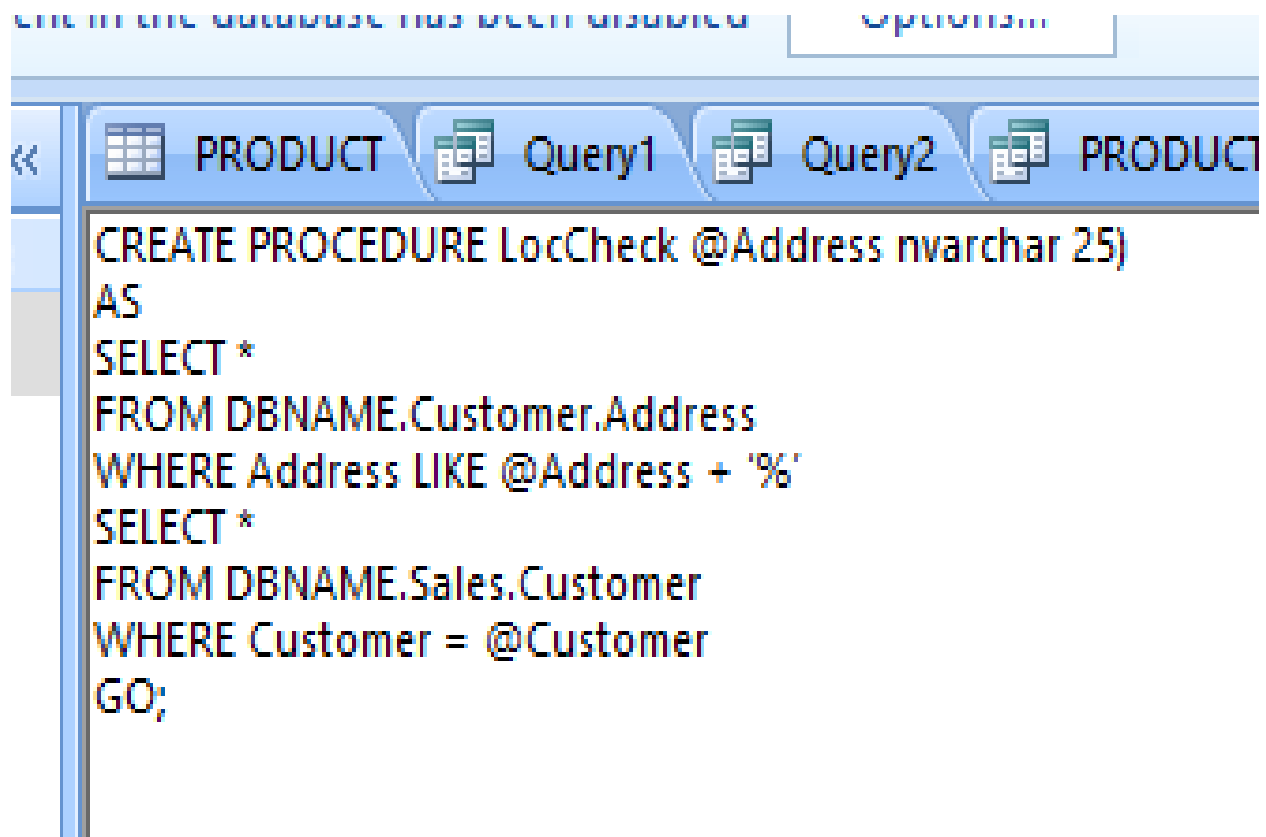
1 and only one store CONTAINS 1 or more products.

1 and only one sales HAS 1 or more products. (**Cornel et al., 2013**).

Fourthly, As the (**Barker, 2015**) opinion, the quick growing of the data is a big problem because all these data need so big storage space, the complexity when retrieving the data, and the difficulty of updating it. This complexity and difficulty of dealing with the data are minimizing the sales with reducing the customers. Therefore, the using of big data tools will be the solution to all these problems. The big data tools provide the removing of complexity, fast time response, and efficient interaction. All these facilities by the using of the visualization and data analytics that used by the user without the programming. (**Barker, 2015**).

As the (**Overby, 2014**) opinion, the big data tools help the companies that have quick growing of their data by the data analysis. The analysis of the data provides the kind imagine about the customers who will purchase from the market. Therefore, the market has the informed decision about the use, store, retrieve, and update of the data fast. This speed increases the sales and helps the database store for maintaining the customers. (**Overby, 2014**).

**Fig (1): The first procedure of the (update the address).**

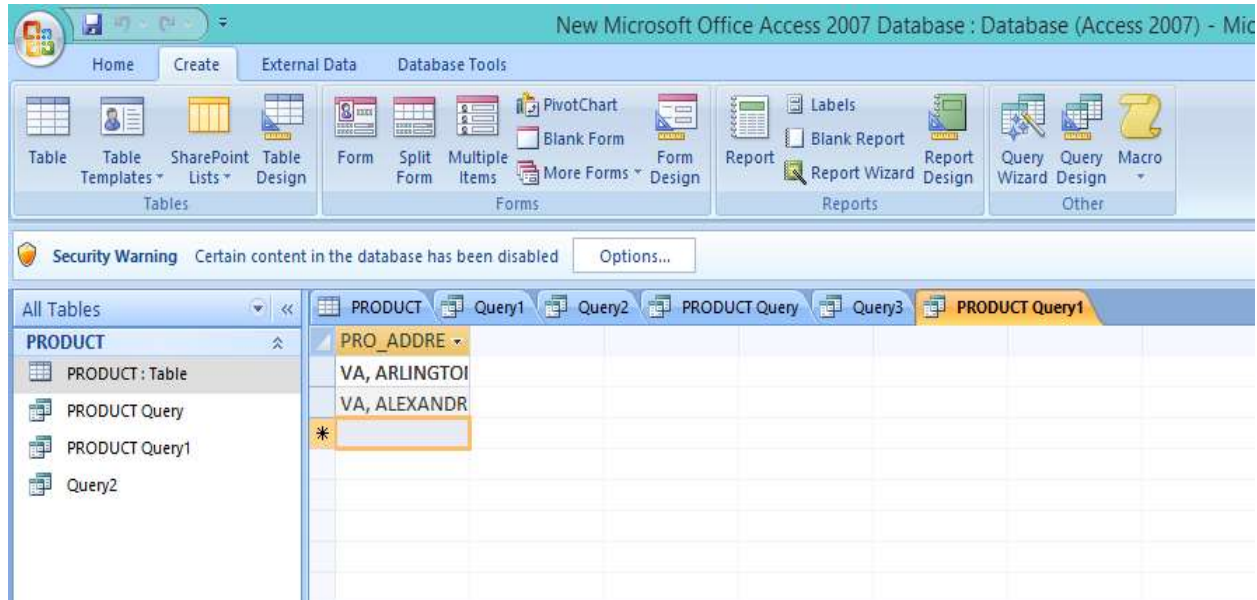


The screenshot shows a SQL Server Enterprise Manager interface. At the top, a status bar indicates "CRL in the database has been disabled" and "Optimization". Below this, a toolbar contains icons for "PRODUCT", "Query1", "Query2", and "PRODUCT". The main area displays the following SQL code:

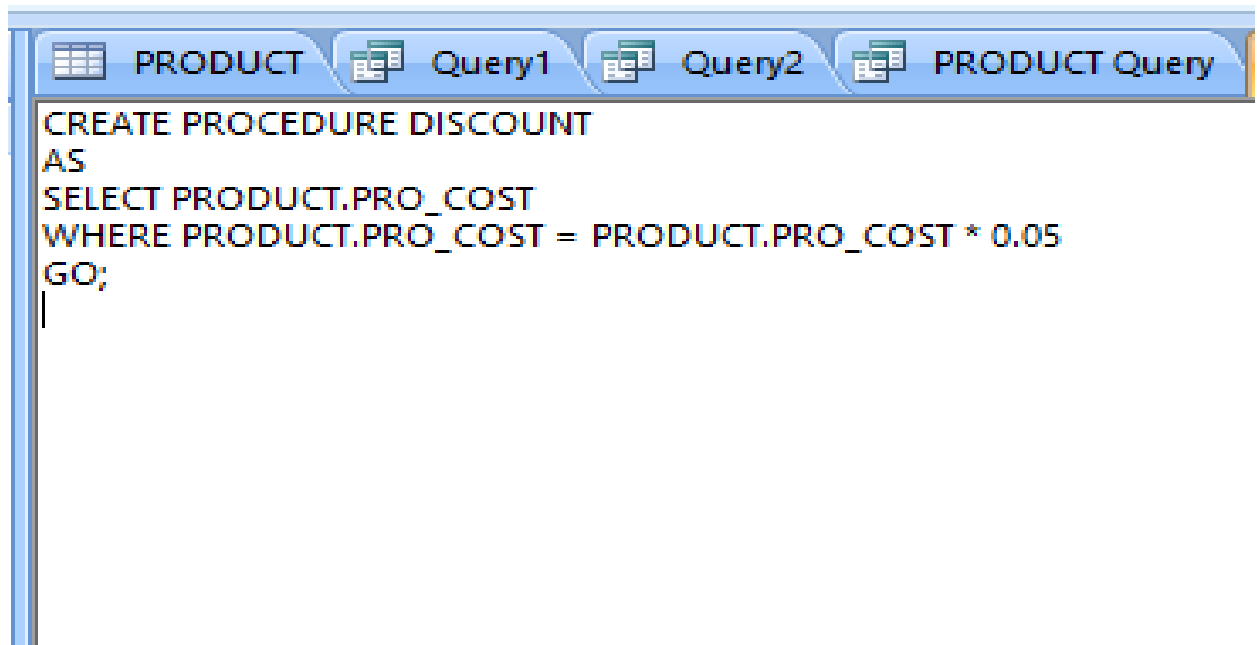
```
CREATE PROCEDURE LocCheck (@Address nvarchar 25)
AS
SELECT *
FROM DBNAME.Customer.Address
WHERE Address LIKE @Address + '%'
SELECT *
FROM DBNAME.Sales.Customer
WHERE Customer = @Customer
GO;
```

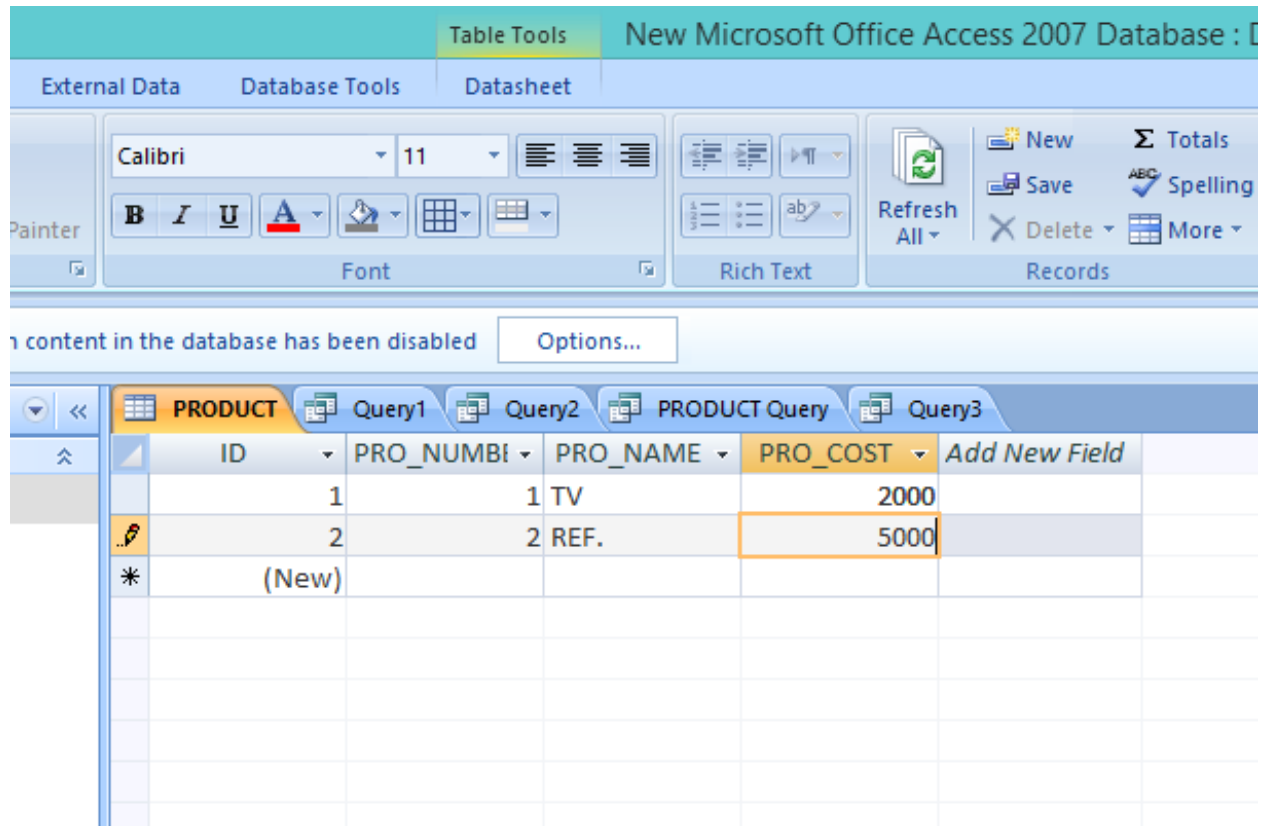


**Fig (2): The screenshot of the SQL program to find the updated address.**

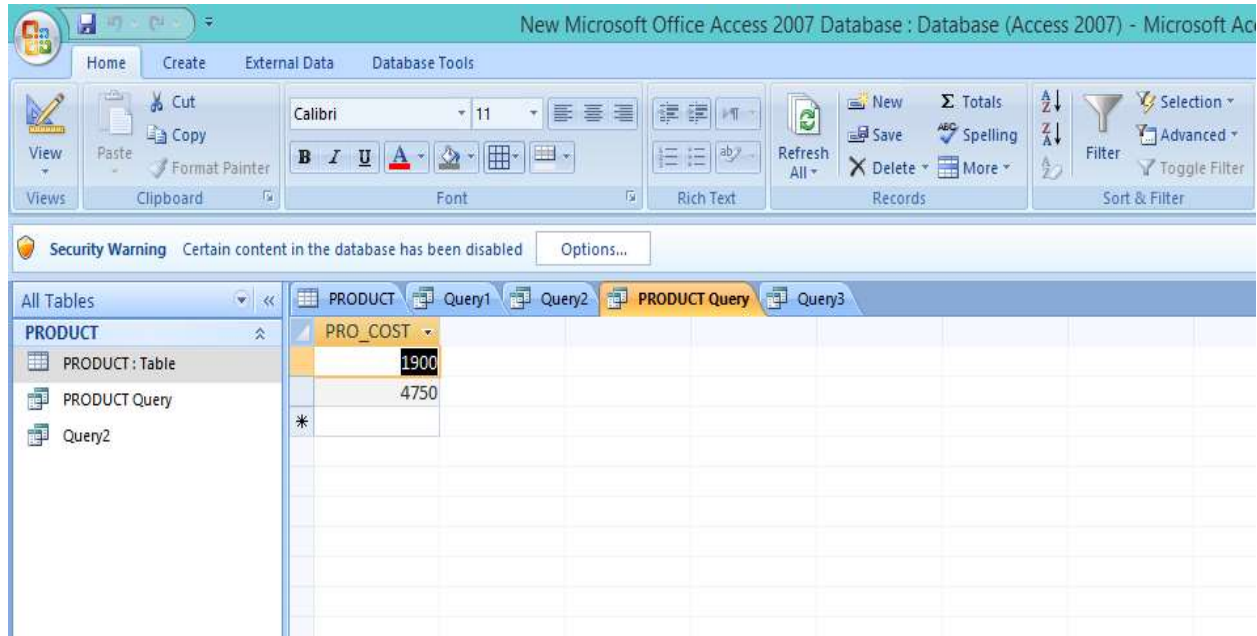


**Fig (3): The second procedure (Discount value).**



**Fig (4): The screen shot of the procedure two.**

**Fig (5): The screen shot of the procedure tow (which has the discount values).**



Fifthly, according to the author (**Tilley et al., 2013**) opinion, the cloud computing is the new software technique that permits a large number of users the corporate usage of the computing resources. There are many trends in the cloud computing such as the pricing, security, and the efficiency consideration. All these proposal will be produced here. Firstly, the author (**Tilley et al., 2013**) explained the pricing structure required when implementing a cloud-hosted solution for a database. The scalability is the ability to grow in the work performance. The cloud computing provides the high scalability; and with an increase of the resources, the scalability also must be increased. The pricing rate depends on the scalability and performance rate. Because of the different performance rate; therefore, there are different pricing rate. "Pricing Problem: Cloud aims to provide economy of scale. However, the promise of unlimited scalability is difficult to achieve. Given various pricing plan and granularity by different providers, such as

infrastructure services and platform services, it leads to different overall cost.". "Cost: The economy of cloud performance is measured as \$/WIPS, where WIPS is for web interactions per second used by the conventional TPC-W benchmark. Also, it also measures the standard deviation of the cost during the scaling,". (Tilley et al., 2013, pg# 356-381). Also, as the author (JR. et al., 2000) opinion, there are many strategies for accessing the database service online: The pricing according to the Internet delivery, The charge must be for every record, The pricing according to the working value, and the connecting time pricing. (JR. et al., 2000).

Also, the author (Kayem, 2014) explained the security consideration of the cloud computing system. The security is one of the most critical problems in the CC system. The risks increase with the multiple providers of the service, and then the costs increase according with the increasing of security risks. The using of the techniques of 'erasure code' to delete the stripping data to enhance the speed of the system performance. With increasing of the system performance, the Cloud-RAID performance also will be increased. The RAID technique increases the data reliability. "Today, storage providers inevitably use RAID like methods to improve the reliability of the entrusted data to their customers. The procedure causes costs that are covered by providers price structure,". (Kayem, 2014, pg# 280-302).

Then the author (Parveen et al., 2013, pg# 54- 67) said about the cloud services options and their ability to provide the performance and integrity of the data. The cloud computing technique provides the Software as a Service (SaaS) to the customers. SaaS permits to the user to use the software tools (such as the storage, server, CPU, etc.) on demand, with flexibility, high performance work, pay per use, and many other advantages. There is a platform as a service that

"Provides a computing service model to its clients including operating system, programming language execution environment, database, and web server. Developers can develop and run their applications on a cloud platform without buying and managing the underlying hardware and software layers.". (**Parveen et al., 2013, pg# 54- 67**); and Infrastructure as a service that is "A platform in which a company outsources the equipment used to support operations, including hardware, servers, storage, and networking components at pay-per-use service. The service provider owns the equipment and handles housing, running, and maintaining it.". (**Parveen et al., 2013, pg# 54- 67**). The both services have risks when continue of the depending on the CC service. Also, the users must not change the CC service provider more times, because that let the user system under the dangerous of threats. Then "the biggest concern for a company is cloud security which includes the abuse of cloud computing, insecure software, malicious insiders, technological vulnerabilities, data loss, traffic hijacking, and unknown risk profiles.". (**Parveen et al., 2013, pg# 54- 67**). Therefore, it must to monitor the tools and test the CC service providers before using their services. The users must put in their minds that the CC service don't provide the high performance of integrity always. (**Parveen et al., 2013, pg# 54- 67**).

Sixth, according to the author (**PETRINI, 2009**) opinion, It is so difficult to application the usage of distributed data in the local locations; but the huge number of data and increase the needs of direct sales marketing led to the designer to design the distributed DBMS. Then, to exceed the critical situation in the distributed data, the stylists put several mechanisms and techniques for enhancing the operations in the database:

- The first method is the Remote Database Access, which is one of the simplest ways to manage the data that are stored in the many locations. This mechanism lets the user to USE, SELECT, UPDATE, and RETRIEVE the data quickly.
- The second method is the Remote Data Transparency which is using the synonyms and aliases in the SQL to exceed the confusing among the columns names and then achieve the remote access transparency. (PETRINI, 2009).

Seventy, the author (Cornel et al., 2013) explained the lost update case is the concurrency control problem that occurs when two transactions update the same data, and one of the transaction is lost through the concurrent execution. For example, if there are two transactions (from two stores to the same table attribute and at the same time) T1 and T2; and the two transactions want to update the value of the same table attribute, such as the table is product (PRODUCT), and the attribute is price (PRO\_PRI). The first transaction T1 has not be committed when the second transaction T2 executes; therefore, the stored value of the T1 is still 50. Then, the T2 reads the false value (50) and executes its operation ( $50 - 10 = 40$ ) to write the end value (40) instead the right value (60). The lost update process happened as the following:

TIME	TRANSACTION	STEP	STORED VALUE
1	T1	READ PRO_PRI	50
2	T2	READ PRO_PRI	50
3	T1	PRO_PRI = 50 + 20	
4	T2	PRO_PRI = 50 - 10	
5	T1	WRITE PRO_PRI (LOST UPDATE)	70
6	T2	WRITE PRO_PRI	40

The right process must be as the following:

TIME	TRANSACTION	STEP	STORED VALUE
1	T1	READ PRO_PRI	50
2	T1	PRO_PRI = 50 + 20	
3	T1	WRITE PRO_PRI	70
4	T2	READ PRO_PRI	70
5	T2	PRO_PRI = 70 - 10	
6	T2	WRITE PRO_PRI	60

Also, the author (**Cornel et al., 2013**) demonstrated the uncommitted data that it happens when the two transactions at the same time at the same table attributes, but the first transaction T1 is rolled back after the transaction two T2 accessed to the uncommitted data. For example, if there are two transactions T1 and T2 occurred at the same time, and the T1 reads the value of the table attributes (PRO\_PRI) = 50, then T1 did its process (purchase 30 units: PRO\_PRI = 50 + 30 = 80); the T2 read this value (80) and after that the T1 is rolled back (because there are something wrong). But, the T2 read the value (80) and did its process (sale 40 units: PRO\_PRI = 80 - 40 = 40) instead of the right process (PRO\_PRI = 50 - 40 = 10). Here the end value is the wrong value 40 instead of the right value 10. The uncommitted data event as the following:

TIME	TRANSACTION	STEP	STORED VALUE
1	T1	READ PRO_PRI	50
2	T1	PRO_PRI = 50 + 30	
3	T1	WRITE PRO_PRI	80
4	T2	READ PRO_PRI (Read uncommitted data)	80
5	T2	PRO_PRI = 80 - 40	
6	T1	*** ROLL BACK ***	50
7	T2	WRITE PRO_PRI	40

While the right execution as the following:

TIME	TRANSACTION	STEP	STORED VALUE
1	T1	READ PRO_PRI	50
2	T1	PRO_PRI = 50 + 30	
3	T1	WRITE PRO_PRI	80
4	T1	*** ROLL BACK ***	50
5	T2	READ PRO_PRI	50
6	T2	PRO_PRI = 50 - 40	
7	T2	WRITE PRO_PRI	10

Eighthly, according to the author (**Cornel et al., 2013**) opinion, there are many concurrency control options: The locking method, Timestamping method, and the optimistic method.

- The locking method which means that any transaction has its exclusive time to complete its process, and any other transaction can't access the same table attributes unless the previous transaction completes its operation with this attributes;
- Time stamping method that is the technique by which every transaction has its specific value, then all the transactions have a scheduled values. The timestamping method has two conditions: the uniqueness (every transaction has a unique value) and monotonicity (transaction value is always increased ). All operation that specified to the transaction (read, write, etc.) must have the same timestamping value of this transaction;
- And the optimistic method that is assuming there are no (concurrency) in the database. According to this approach the transaction moves with three levels: read, validation, and



write. This method is accepted by the DB system that didn't have many update processes. **(Cornel et al., 2013)**.

Ninthly, according to the author **(Gregory, 2009)**, the disaster is an event is happened by the natural or man-made. The reasons for a natural disaster are geographical, health, and other. The reason for the man-made disaster are the stolen, sabotage, denial of service, political/ religion, and competitions reasons (among the companies). There are two options are made to achieve the disaster management: Business continuity (BC), and disaster recovery planning (DRP). BC is the ensuring of continuing the business to work through the disaster. DRP is the plans to recover any defect that occurs by the disaster are many standards are provided as the BC and DRP, such as the NIST 800-34, ISO 17799, NFPA 1620, and others. **(Gregory, 2009)**.

According to the author **(Oppel et al., 2009)** opinion, there are many types of database threats, such as the following:

- The physical threats: This type of threat is threatening the hardware components that are included the database information, such as the hard disk, servers, and other. This component may be stolen by the intruders or, it may be damaged by any one of the natural disasters that damage them.
- The application software: This type of threats is threatening the software components that are including the programs, applications, and all database information. The hackers are always trying to access to the DB information by many ways, such as the Viruses, malicious codes (Malware), phishing, the denial of service DoD, and many ways. Also, the spying may be happened by the social engineering which is occurred by the employees are working in the company.

Then, to have the countermeasures from these threats, it must do the following measures:

- The physical security protection: This standard means protecting the hardware components from any stealing or damaging. This protecting occurs by putting the hardware components in locked room with camera surveillance, some security devices (keys), and others.
- The software application protection: This type of protection protects the database programs and its application from the hackers. To achieve this protection, it must use the anti-viruses, Anti-Malware, and mulch layers firewalls.
- The protection from the social engineering that is occurred by the employees. It must use much application to limit the accessing, such as the authentication, authorization, privilege, accreditation, identification, and others. (Oppel et al., 2009).

### **Conclusion:**

This paper talked about the meaning of the potential sales, and it can be measured by calculating the amount of the consumer sales. The potential sales have many advantages such as the evaluating the information performance. Also, the paper demonstrated the transactions that are the core of the DBMS; and when the transactions are founded in the retail stores. The database design has many components: Entities, Attributes, and relationships among the entities. Every entity has its attributes, and there are relationships types, such as 1: 1, 1: M, and M: N. Every one of these relations has its specific symbol that you saw its drawing in this paper; with the representing of all database components of the store. There are many types of models to design the database, and one of these models is the Crow's Foot notation that has the relationship name that is written above of the relationship line; the relationship line has two

different sides that represent the relationship type. This paper had the design of the Crow's Foot notation model of the store. Then, it explained the big data model, and it would be used to be the solution of the huge grown of the data in these days. Next, I represented the SQL Stored Procedures that use SQL functions to help sales associates perform explanatory or predictive analytics; with all the screen shots of the SQL programming. The cloud computing is the new software technique that permits a large number of users the corporate usage of the computing resources. It has many trends; therefore it has various pricing considerations with different security phases. This technique of the cloud computing service has options of Software as a Service such as the Platform as a Service and Infrastructure as a service that permits the user to use the software components from many far locations. Next, the paper explained about the distributed DBMS and it is appropriate for the store database, but with more attention to the privacy of the database. After that, I provided the meaning of the lost updates and uncommitted data with examples of how they may occur with the mulch transactions at the same attributes. Then, how we can control the concurrency problem by the many approaches. Lastly, the paper explained about the database security; what is the disaster mean; how we can manage the database through the disaster; what are the types of database threats; and what is the threats countermeasures.

**References:**

**Cox Jr., et al., (1977).** The article of (Determination of Sales Potentials and Performance for an Industrial Goods Manufacturer), Journal of Marketing Research (JMR). Nov77, Vol. 14 Issue 4, p574-578. 5p. 5 Charts. , Database: Business Source Complete.

**JARKE et al., (1984).** The article of (A Database Architecture for Supporting Business Transactions). Journal of Management Information Systems. Summer84, Vol. 1 Issue 1, p63-80. 18p.

**Cornel et al., (2013).** Database Systems: Design, Implementation, and Management, 10e - 1 Year Option, 10th Edition. (eBook).

**Barker, (2015).** The article of (Kyvos aims to change the big data analytics landscape). Retrieved from <http://betanews.com/2015/06/30/kyvos-aims-to-change-the-big-data-analytics-landscape/>

**Overby, (2014).** The article of (The CIO and CMO perspective on big data). Retrieved from <http://www.itworld.com/article/2697057/big-data/the-cio-and-cmo-perspective-on-big-data.html?page=2>

**JR. et al., (2000),** The article of (Private Markets for Public Goods: Pricing Strategies of Online Database Vendors). Journal of Management Information Systems. Summer2000, Vol. 17 Issue 1, p59-85. 27p. 6 Graphs

**Tilley et al., (2013).** The article of (Cloud Scalability Measurement and Testing).Cloud Technology: Concepts, Methodologies, Tools, and Applications. (eBook).

**Kayem, (2014).** The article of (Critical issues/ Solving Security and Availability Challenges in Public Clouds). Cloud Technology: Concepts, Methodologies, Tools, and Applications. (eBook).

**Parveen et al., (2013).** The article of (Tools and Technologies/ Cloud-Enabled Software Testing Based on Program Understanding).Cloud Technology: Concepts, Methodologies, Tools, and Applications. (eBook).

**PETRINI, (2009).** The article of (DISTRIBUTED DATABASES MANAGEMENT USING REMOTE ACCESS METHOD). Annals of the University of Petrosani Economics. 2009, Vol. 9 Issue 2, p203-208. 6p. 2 Diagrams.

**(Gregory, 2009).** The article of (Chapter 4: Business Continuity and Disaster Recovery). CISSP Guide to Security Essentials, 1e - 1 Year Option. (eBook)

**(Oppel et al., 2009).** The article of (Chapter 10: Database Security). Databases: A Beginner's Guide (Book).