|  |
| --- |
| Department of Software Engineering  Mehran University of Engineering and Technology, Jamshoro |

|  |  |  |  |
| --- | --- | --- | --- |
| Course: SWE324 - Data Warehousing and Data Mining | | | |
| Instructor | Rabeea Jaffari | **Practical/Lab No.** | 01 |
| Date | 02 April 2019 | **CLOs** | CLO-4: P3 & P4 |
| Signature |  | **Assessment Score** | 1 Marks |

|  |  |
| --- | --- |
| Topic | To become Familiar with OLTP System Design |
| Objectives | * To learn conceptual and logical Transactional Database Design |

|  |
| --- |
| Lab Discussion: Theoretical concepts and Procedural steps |

**OLTP:**  OLTP (Online transaction processing) systems are information systems that perform transaction (can be defined as any day to day or daily activity) oriented tasks and are typically used for data entry and data retrieval operations. The most common OLTP system operations are inserting, updating and deleting data in the OLTP systems. In the light of above statement, it can be said that relational databases are types of OLTP systems which are also termed as operational or transactional databases. These contain related data from daily operations from organizational activities organized in the form of tables/relations.

**OLTP SYSTEM DESIGN:**OLTP systems are designed at the following levels:

1. **Conceptual Design:** Conceptual design of an OLTP system involves the transformation of requirements into a conceptual model which involves identifying the entities and objects required for the organization. It is done with the aid of ER diagrams (to be discussed ahead).
2. **Logical Design:** Logical design maps the conceptual model created with an ER diagram in the above step into a logical model represented via tables with normalization taking place if necessary.
3. **Physical Design:** Physical design is the actual implementation of the tables created in logical design into the system memory with the help of any database product such as ORACLE, MYSQL, SQLSERVER etc. Thus, this design is DBMS dependent.

This lab would be covering the first two OLTP design levels.

**Conceptual Design Constructs: ER Diagram:**

**ER Diagram:** It is a graphical representation of how the data produced by an organization or business area is related. There are five main components of an ERD:

* **Entities**: are represented by rectangles. An entity is an object or concept about which you want to store information.



A weak entity is an entity that is dependent on some other entity and therefore must defined by a foreign key relationship with another entity in the logical design as it cannot be uniquely identified by its own attributes alone.



* **Actions/Relationships**: are represented by diamond shapes and show how two entities share information in the database.



In some cases, entities can be self-linked. For example, employees can supervise other employees.



* **Attributes**: are represented by ovals. A key attribute is the unique, distinguishing characteristic of the entity. For example, an employee's social security number might be the employee's key attribute.



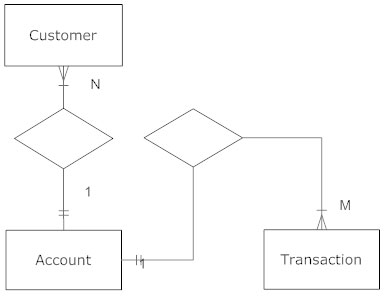
A multivalued attribute can have more than one value. For example, an employee entity can have multiple skill values or multiple contact numbers.



A derived attribute is based on another attribute. For example, an employee's monthly salary is based on the employee's annual salary.



* **Connecting lines** are solid lines that connect attributes to show the relationships of entities in the diagram.
* **Cardinality:** specifies how many instances of an entity relate to one instance of another entity. Ordinality is also closely linked to cardinality. While cardinality specifies the occurrences of a relationship, ordinality describes the relationship as either mandatory or optional. In other words, cardinality specifies the maximum number of relationships and ordinality specifies the absolute minimum number of relationships.



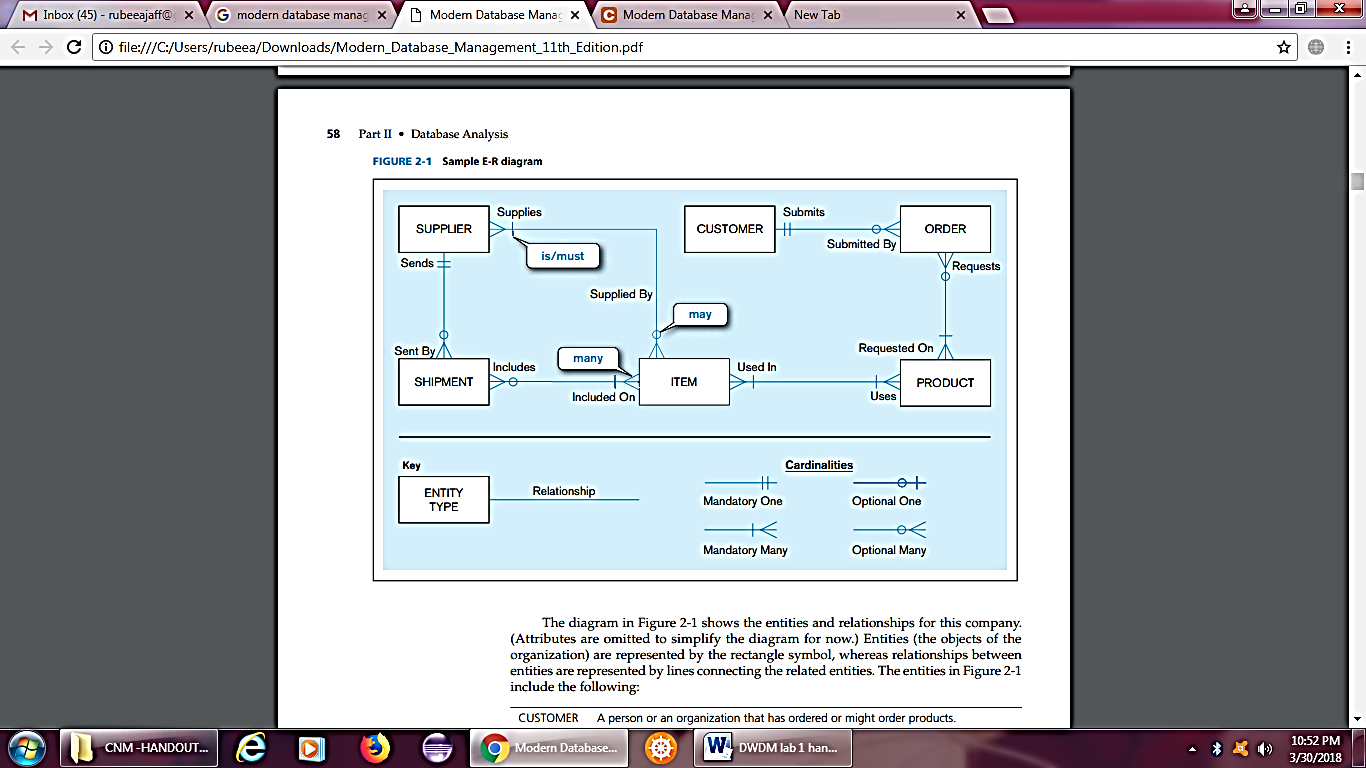
The different notations for cardinality and ordinality are as follows:



**STEPS:** The steps for creating an ER diagram are as follows:

* **Identify the entities:** The first step in making an ERD is to identify all of the entities you will use represented by a rectangle with a description of something that your system stores information about. This could be a customer, a manager, an invoice, a schedule, etc.
* **Identify relationships:** Examine the entities and if they are related, draw a solid line connecting the entities.
* **Describe the relationship:**Draw an action diamond between the two entities on the line you just added. In the diamond write a brief description of how they are related.
* **Add attributes:** Any attributes of the relationship such as cardinality or ordinality can be added at the two ends of relationship lines connecting the entities.
* **Complete the diagram.** Continue to connect the entities with lines, and adding diamonds to describe each relationship until all relationships have been described. Each of your entities may not have any relationships, some may have multiple relationships.

**SAMPLE ERD:**

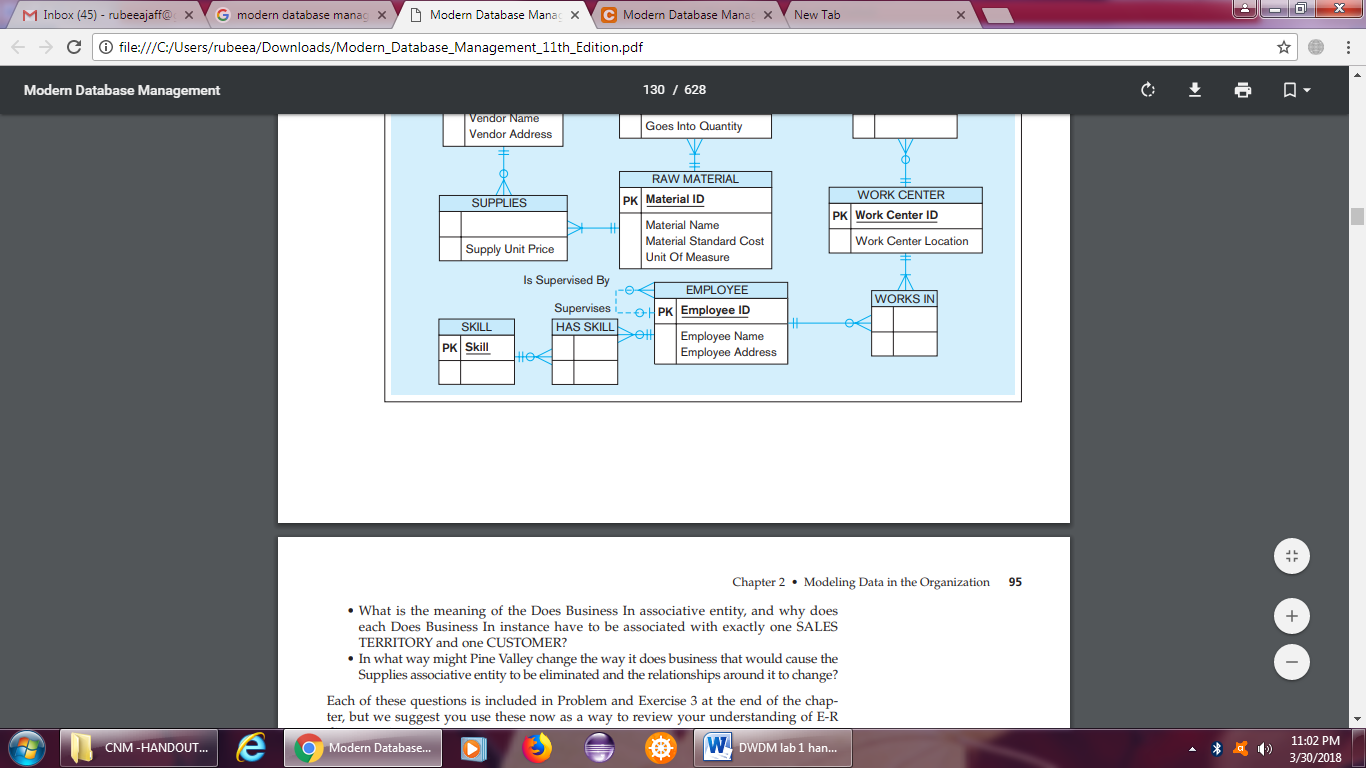
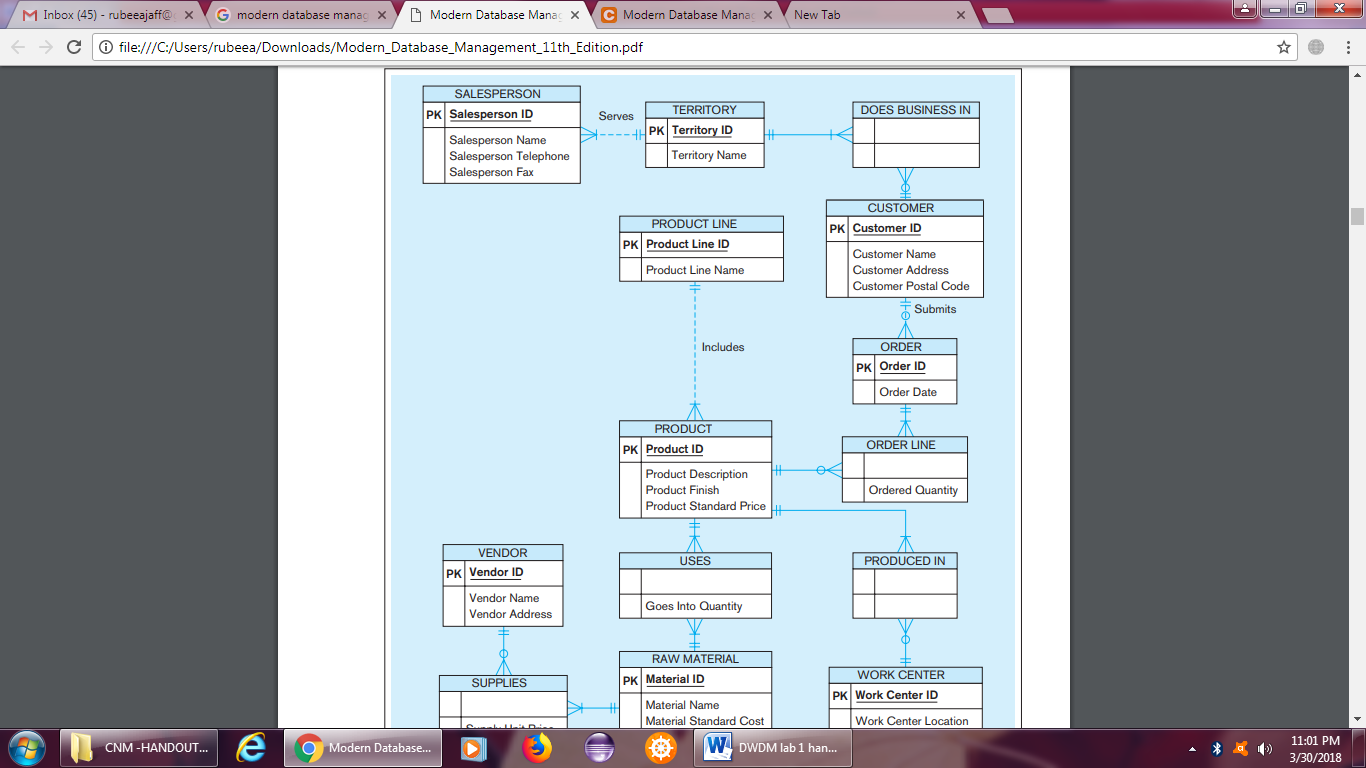


This is an ERD for a small furniture manufacturing company, Pine Valley Furniture Company. A number of suppliers supply and ship different items to Pine Valley Furniture. The items are assembled into products that are sold to customers who order the products. Each customer order may include one or more lines corresponding to the products appearing on that order.

The entities in this figure are CUSTOMER, ORDER, SHIPMENT, ITEM, PRODUCT and SUPPLIER. On examining this ERD, we can determine that the cardinality and the ordinality symbols explain the following rules:

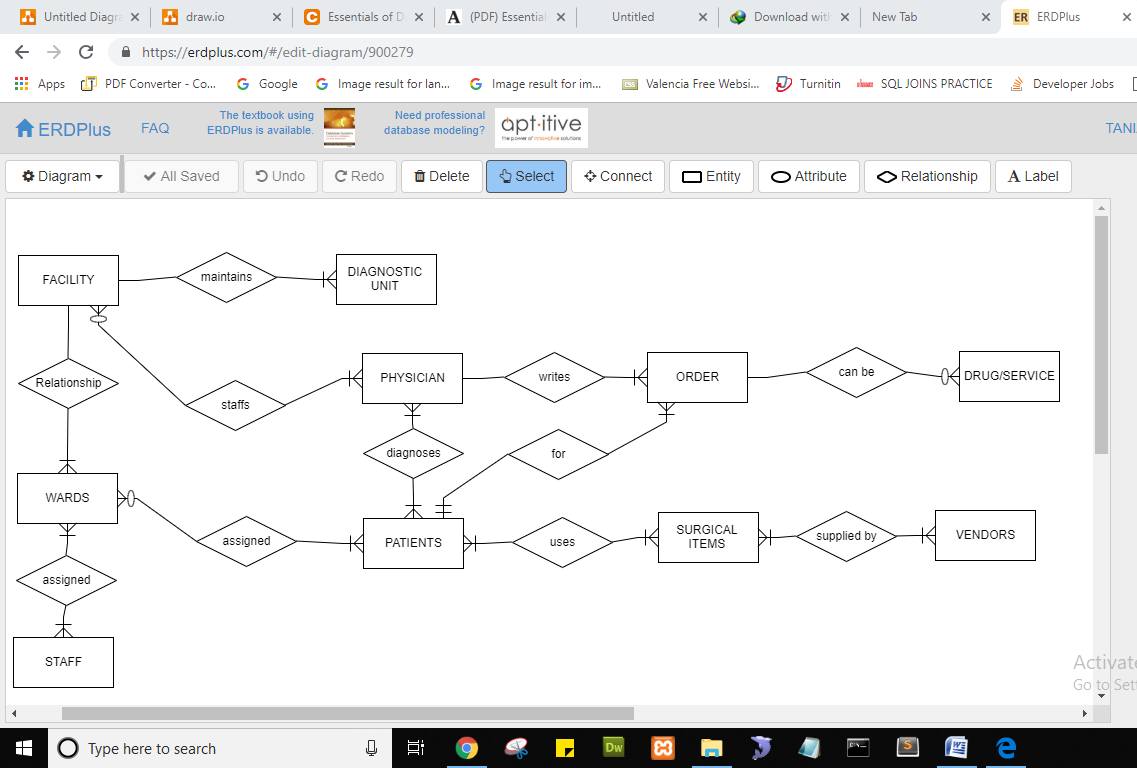
1. A SUPPLIER may supply many ITEMs (by “may supply,” we mean the supplier may not supply any items). Each ITEM is supplied by any number of SUPPLIERs (by “is supplied,” we mean that the item must be supplied by at least one supplier).
2. Each ITEM must be used in the assembly of at least one PRODUCT and may be used in many products. Conversely, each PRODUCT must use one or more ITEMs.
3. A SUPPLIER may send many SHIPMENTs. However, each shipment must be sent by exactly one SUPPLIER. Notice that sends and supplies are separate concepts. A SUPPLIER may be able to supply an item but may not yet have sent any shipments of that item.
4. A SHIPMENT must include one (or more) ITEMs. An ITEM may be included on several SHIPMENTs.
5. A CUSTOMER may submit any number of ORDERs. However, each ORDER must be submitted by exactly one CUSTOMER.
6. An ORDER must request one (or more) PRODUCTs. A given PRODUCT may not be requested on any ORDER or may be requested on one or more orders.

**Logical Design:**Logical design is similar to an ERD but with additional information incorporated into it such as attributes (with data types and maximum lengths) and keys describing the entities. Sample logical design is shown below:



|  |
| --- |
| Lab Tasks |
| Submission Date: 09-04-19 |

1. Draw ERDs for the following case studies:
   1. The study team identified a preliminary set of 11 entity types that describe the data required by the hospital in support of the various business functions: FACILITY, PHYSICIAN, PATIENT, DIAGNOSTIC UNIT, WARD, STAFF, ORDER, SERVICE/ DRUG, MEDICAL/SURGICAL ITEM, SUPPLY ITEM, and VENDOR. From discussions with hospital staff, reviewing hospital documents, and studying existing information systems, the study team developed a list of business rules describing the policies of the hospital and nature of the hospital’s operation that govern the relationships among these entities. Some of these rules follow:
2. A FACILITY maintains one or more DIAGNOSTIC UNITs (radiology, clinical laboratory, cardiac diagnostic unit, etc.).
3. A FACILITY contains a number of WARDs (obstetrics, oncology, geriatrics, etc.).
4. Each WARD is assigned a certain number of STAFF members (nurses, secretaries, etc.); a STAFF member may be assigned to multiple WARDs.
5. A FACILITY staffs its medical team with a number of PHYSICIANs. A PHYSICIAN may be on the staff of more than one FACILITY.
6. A PHYSICIAN treats PATIENTs, and a PATIENT is treated by any number of PHYSICIANs.
7. A PHYSICIAN diagnoses PATIENTs, and a PATIENT is diagnosed by any number of PHYSICIANs.
8. A PATIENT may be assigned to a WARD (outpatients are not assigned to a WARD). The hospital cares only about the current WARD a patient is assigned to (if assigned at all).
9. A PATIENT uses MEDICAL/SURGICAL ITEMs, which are supplied by VENDORs. A VENDOR also provides SUPPLY ITEMs that are used for housekeeping and maintenance purposes.
10. A PHYSICIAN writes one or more ORDERs for a PATIENT. Each ORDER is for a given PATIENT, and a PATIENT may have many ORDERs.
11. An ORDER can be for a SERVICE such as diagnostic test (lab tests such as lipid profile, CBC, liver function tests; diagnostic imaging such as MRIs and X-rays) or a DRUG.



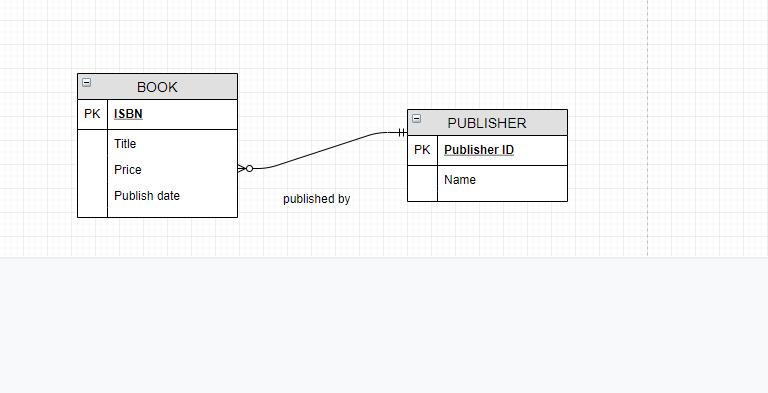
For each of the descriptions below, perform the following tasks:

i. Identify the degree and cardinalities of the relationship.

ii. Express the relationships in each description graphically with a logical design ER diagram.

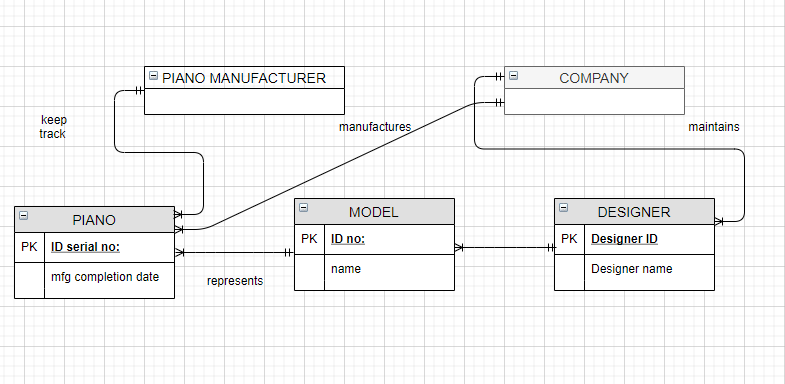
1. A book is identified by its ISBN number, and it has a title, a price, and a date of publication. It is published by a publisher, which has its own ID number and a name. Each book has exactly one publisher, but one publisher typically publishes multiple books over time.

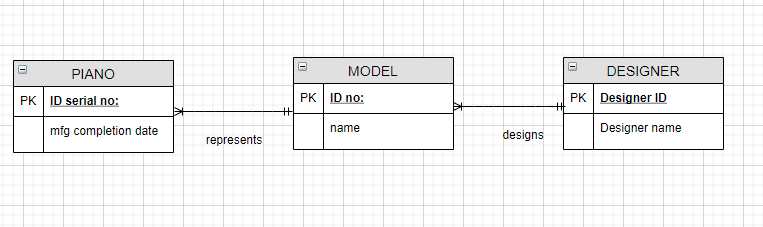
* The relationship is of degree 2(binary).
* The relationship is one-to many from publisher to book.



1. A piano manufacturer wants to keep track of all the pianos it makes individually. Each piano has an identifying serial number and a manufacturing completion date. Each instrument represents exactly one piano model, all of which have an identification number and a name. In addition, the company wants to maintain information about the designer of the model. Over time, the company often manufactures thousands of pianos of a certain model, and the model design is specified before any single piano exists.

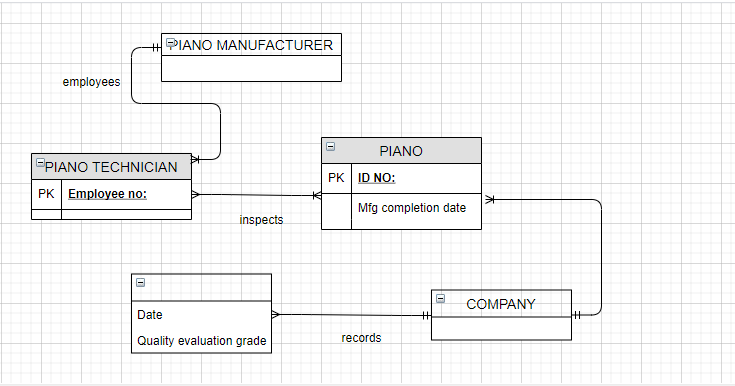
* The relationship is of degree 2.
* The relationship is one to many.

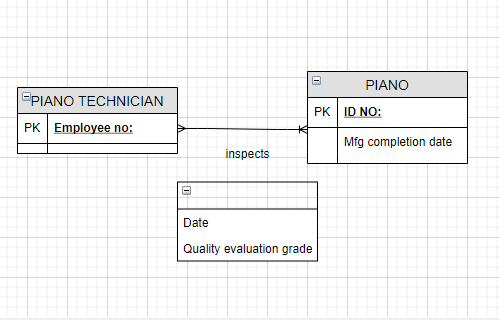




1. A piano manufacturer (see above) employs piano technicians who are responsible for inspecting the instruments before they are shipped to the customers. Each piano is inspected by at least two technicians (identified by their employee number). For each separate inspection, the company needs to record its date and a quality evaluation grade.

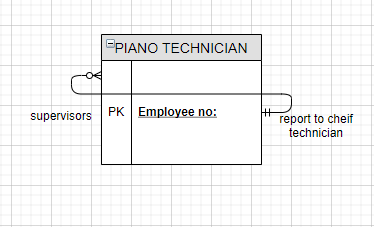
* The relationship is of degree 2
* The relationship is many to many.





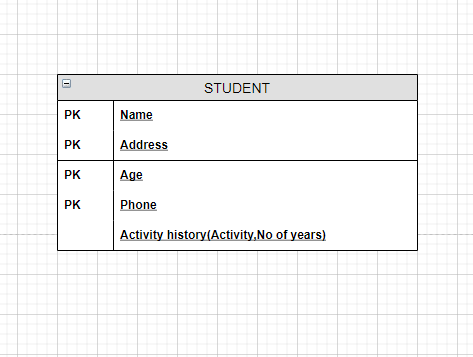
1. The piano technicians (see above) have a hierarchy of reporting relationships: Some of them have supervisory responsibilities in addition to their inspection role and have multiple other technicians report to them. The supervisors themselves report to the chief technician of the company.

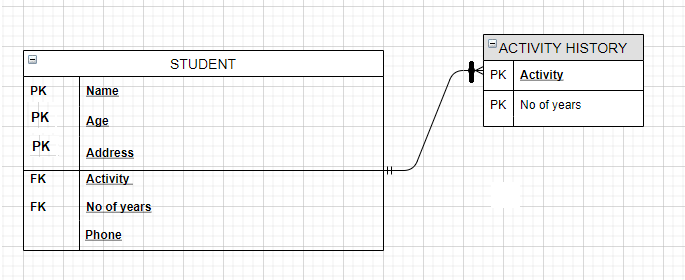
* The relationship is of degree one(unary).
* The relationship is one to many.



1. The entity type STUDENT has the following attributes: Student Name, Address, Phone, Age, Activity, and No of Years. Activity represents some campus-based student activity, and No of Years represents the number of years the student has engaged in this activity. A given student may engage in more than one activity. Draw a logical ERD for this situation. What attribute or attributes did you designate as the identifier for the STUDENT entity? Why?

* Student name only can’t be use as an identifier. Hence student name, address, phone, age will be use as identifiers.





1. A college course may have one or more scheduled sections or may not have a scheduled section. Attributes of COURSE include Course ID, Course Name, and Units. Attributes of SECTION include Section Number and Semester ID. Semester ID is composed of two parts: Semester and Year. Section Number is an integer (such as 1 or 2) that distinguishes one section from another for the same course but does not uniquely identify a section. How did you model SECTION? Why did you choose this way versus alternative ways to model SECTION?

* Section here represents a weak entity
* Both section number and section id are used as an identifier.
* The relationship is one to many.
* The degree is 2 binary.

