Normalization Questions

Each of the below questions deals with normal forms and data modeling, and should serve as examples for what sorts of questions might appear on the exam.

1. Is the following table in 1st normal form? If not, what is wrong with it?

🗝 student_id	student_name	subject	grade
1	Alice	Math, Science, English	A, B, A
2	Bob	English, History	C, B
3	Charlie	Math, Art	B, A
4	Diana	Science, Math	A, B
5	Eve	History, Science	B, A
6	Frank	English, Science, History	B, A, C

Solution: No, this is not in first normal form. While it does have a unique primary key, a requirement of 1st normal form is that each cell contains only a single value.

2. Is the following table in 2nd normal form? If not, what is wrong with it?

🗝 EmployeeID	🗝 DepartmentID	DepartmentName	EmployeeName	HourlyWage
1	101	Sales	John	15
2	102	Marketing	Sarah	18
3	101	Sales	Tom	15
4	103	IT	Kate	20
5	104	HR	Linda	17
6	101	Sales	Mike	15

Solution: No, there are two partial dependencies here. The DepartmentName depends only on the department id, and the employee name depends only on the employee id.

3. What normal form is this table in? Explain your reasoning for full points.

🔑OrderID	🔑ProductID	ProductName	SupplierID	SupplierName
101	1	Widget	5001	Widgets Co.
101	2	Gadget	5002	Gadgets Inc.
102	1	Widget	5001	Widgets Co.
102	3	Doohickey	5003	Doohickey Ltd
102	4	Thingamajig	5004	Things Corp.
103	2	Gadget	5002	Gadgets Inc.

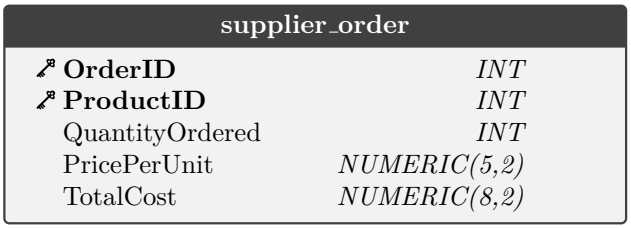
Solution: This table is in 1NF, but it is not in 2NF as it has a partial dependency between product name and product id. It also has a transitive dependency between supplier name and supplier id.

4. What normal form is this table in? Explain your reasoning for full points.

🔑RegistrationID	StudentID	StudentName	CourseID	CourseName	InstructorID	InstructorName
1	1	Alice	100	Math	10	Dr. Smith
2	2	Bob	101	Science	11	Dr. Jones
3	1	Alice	101	Science	11	Dr. Jones
4	3	Charlie	102	Art	12	Dr. Brown
5	4	Diana	103	History	13	Dr. Green
6	5	Eve	101	Science	11	Dr. Jones

Solution: This table is in 2NF. It has a singular primary key, so it is automatically in 2NF if it is in 1NF, since partial dependencies are impossible. And it has one piece of information in every cell, thus securing 1NF. It is definitely not in 3NF, as there are many transitive dependencies between various names and ids.

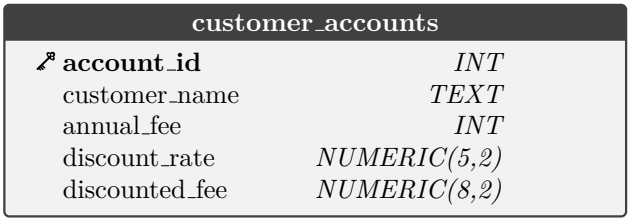
5. Below is the ERD of a table and the accompanying unnormalized JSON data that should be placed into said table. What normal form is this table currently in? Explain yourself for full points.




```
[
  {
    "OrderID": 1,
    "ProductID": "A123",
    "QuantityOrdered": 10,
    "PricePerUnit": 15,
    "TotalCost": 150
  },
  {
    "OrderID": 2,
    "ProductID": "B456",
    "QuantityOrdered": 5,
    "PricePerUnit": 25,
    "TotalCost": 125
  },
  {
    "OrderID": 3,
    "ProductID": "A123",
    "QuantityOrdered": 12,
    "PricePerUnit": 15,
    "TotalCost": 180
  }
]
```

Solution: This table is only in 1NF atm. There is a partial dependency on the price-per-unit to the product id, which can be seen in the JSON. And not that it matters, but there would also be a transitive dependency with the total cost.

6. Below is the ERD of a table and a sample of the accompanying unnormalized JSON data that should be placed into said table. What normal form is this table currently in? Explain yourself for full points.




```
[
  {
    "AccountID": 101,
    "CustomerName": "Jane␣Doe",
    "AnnualFee": 200,
    "DiscountRate": 0.10,
    "DiscountedFee": 180
  },
  {
    "AccountID": 102,
    "CustomerName": "John␣Smith",
    "AnnualFee": 300,
    "DiscountRate": 0.15,
    "DiscountedFee": 255
  },
  {
    "AccountID": 103,
    "CustomerName": "Sarah␣Brown",
    "AnnualFee": 200,
    "DiscountRate": 0.20,
    "DiscountedFee": 160
  }
]
```

Solution: This table looks to be in 2NF. It has a singular unique primary key and 1 value in each cell, so it is easily in 1NF and thus in 2NF. But there is a transitive dependency on the discounted fee depending on the annual fee and discount rate, not directly on the account id.