Name Chopra Aman, Ranjan Nidhi Date: 5th November 2022

NYU NetID: ac8511, nr2387

Section: CSCI-GA.2433-001

**Project Part 1**

We Aman Chopra and Nidhi Ranjan, hereby certify by submitting this project Part 1 that all the efforts put into this part are our own. We have referred the project support materials.

**Total in points** (100 points total): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Professor’s Comments:**

**A picture containing arrow

Description automatically generated**

**Affirmation of my Independent Effort:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Description:**

Our project is to create an Auto Insurance system. The customers can buy policies, and claim insurance. Similarly, the agents can sell policies and investigators can investigate the claims. We have tried to build a scalable and simple architecture to support most of the use cases of a functional Auto Insurance Company.

**Aim**:

We want to generate and accumulate data to generate new premium prices for the customers based on their history, the state they reside in, their vehicle type, the amount of premium they have paid, and the money they have claimed.

We also want to generate snapshots to quickly query and get the details of how well the business is doing. We will talk more about these snapshots after we explain the structure.

**ER Diagram:**

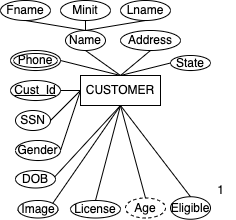
**Diagram

Description automatically generated**

**Entities Involved:**

1. CUSTOMERS
2. HISTORY
3. CLAIMS
4. VEHICLE
5. EMPLOYEE
6. DATA SCIENTIST
7. INVESTIGATOR
8. AGENT
9. PRODUCT (INSURANCE)
10. LOSS DAMAGE WAIVER
11. THIRD-PARTY LIABILITY
12. PROPERTY DAMAGE
13. BODILY INJURY
14. POLICY
15. TRANSACTION
16. FINANCIAL INSTITUTION
17. INVOICE
18. CLAIMED
19. PURCHASED

**CUSTOMER:**

****

**Attributes:**

1. The Cust\_ID is the customer's primary key to uniquely identify a customer.
2. The Customer's name contains first, middle, and last names.
3. The Address of the customer.
4. The state where the Customer resides in as different states has different auto insurance laws and prices.
5. The Phone number is a Multivalued attribute.
6. The Social Security Number of the Customer.
7. The Image of the customer is saved in the Base64 format.
8. The License Number of the Customer.
9. The Date of Birth of the Customer.
10. The Gender of the Customer.
11. The Age of the Customer is derived from the Date of Birth and hence is a derived attribute.
12. The customer’s eligibility indicates if they are eligible to buy an insurance policy.

**Cardinality:**

1. 1 Customer has one History.
2. 1 Customer can have N Claims.
3. 1 Customer can do N Transactions.
4. 1 Customer can buy N Policies.

**HISTORY:**

This entity is essential in determining the premium a customer has to pay. A person with a history of accidents and a lot of claims might have to pay a higher premium price.

Text

Description automatically generated with medium confidence

It is a weak entity type and does not have a primary key. It has total participation with the Customer and history cannot exist without a customer.

**Attributes:**

1. The amount of money the Customer has claimed.
2. The number of accidents the Customer has been a part of.
3. The premium the customer currently pays.

**Cardinality:**

1. 1 History belongs to 1 Customer.
2. 1 History can contain N Claims.

**CLAIMS:**

This entity contains the details of the claims done by Customers.

Graphical user interface, application

Description automatically generated

**Attributes:**

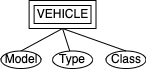
1. Claim\_id is the primary Key.
2. Claim\_note contains the details of the claim.
3. Date indicates when the claim was claimed.
4. Event described which incident led to the accident.
5. Proof image is a Base64 encoded image containing the proof needed to claim the insurance.

**Cardinality:**

1. Many claims constitute the history of a Customer.
2. Many claims can be done by a Customer.
3. Claims have a Transaction associated with them.
4. A Claim contains a policy.
5. A Claim can be investigated by many Investigators.

**VEHICLE:**

The type of Vehicle owned by a Customer can determine the price of the auto insurance. It is a weak entity as in this scenario, it cannot exists on its own. It also has total participation from the Customers.



**Attributes:**

1. Model of the Vehicle
2. Type of the Vehicle like normal, and heavy duty.
3. The Class of the vehicle is based on the vehicle’s price. The Class determines which insurance premium price range the vehicle falls in.

**Cardinality:**

1. Many vehicles can be owned by a Customer.

**EMPLOYEE:**

The Employee is a superclass that has the details of the Employees working at the Insurance Company.

A picture containing text, electronics

Description automatically generated

**Attributes:**

1. The E\_Id is the employee’s primary key to uniquely identify an employee.
2. The Employee’s name contains first, middle, and last names.
3. The Address of the employee.
4. The Phone number is a Multivalued attribute.
5. The Social Security Number of the Employee.
6. The Image of the employee is saved in the Base64 format.
7. The Date of Birth of the Employee.
8. The Gender of the Employee.
9. The Age of the Employee is derived from the Date of Birth and hence is a derived attribute.
10. The customer’s eligibility indicates if they are eligible to buy an insurance policy.
11. The experience of the employee indicates the years of relevant experience.
12. The Salary of the Employee per year.

**Specialization:**

There are 3 specialized classes of Employees. We have used disjoint partial participation as each class has a particular set of responsibilities and are not overlapping. This set of employees is non-exhaustive and hence partial participation.

Text

Description automatically generated with low confidence

1. DATA SCIENTIST: Aggregates and analyze data to determine policy premiums based on Customer’s Vehicle and History. They can have specialization in the fields of Machine Learning and Data Engineering.
2. INVESTIGATOR: Investigates the Claims by analyzing the proofs. Cases Investigated tells the number of cases investigated by a particular Investigator.
3. AGENT: The Agent sells the policies to the Customer. They have a field called Deals which tells the number of deals closed by an agent.

**Cardinality:**

1. One Agent can sell multiple policies.
2. Many Investigators can investigate many claims.

**PRODUCT:**

The Product is a superclass that has details of the Insurance plans that the company has to offer.

Graphical user interface, text

Description automatically generated with medium confidence

**Attributes:**

1. The Product\_id is the product’s primary key to uniquely identify a Product.
2. The type tells the type of the Product which is auto insurance in this case.
3. The name of the Product.
4. The Base Price of a particular Product.
5. The premium that needs to be paid on top of the base price that depends on the Customer’s History.
6. The Coverage which tells what is covered in the particular Product.
7. The benefits of the Product.
8. The Details of the Product.

**Specialization:**

There are 4 specialized classes of Products. We have used disjoint partial participation as insurance covers different things and is not overlapping. This set of insurances is non-exhaustive and hence partial participation.

Text

Description automatically generated with medium confidence

1. LOSS DAMAGE WAIVER: Covers the loss to the own vehicle in case of an accident.
2. THIRD-PARTY LIABILITY: Covers the loss done to other’s vehicle in case of an accident.
3. PROPERTY DAMAGE: Covers the loss to other’s properties in case of an accident.
4. BODILY INJURY: Covers the injuries done to other people in case of an accident.

**Cardinality:**

1. One Loss Damage waiver can be a part of 1 policy.
2. One Third-party Liability can be a part of 1 policy.
3. One Property Damage can be a part of 1 policy.
4. One Bodily Injury can be a part of 1 policy.

**POLICY:**

The entity which the Customer buys and the Agent sells. The policy covers and insures a Customer.

Diagram, text

Description automatically generated

**Attributes:**

1. The Policy\_id is the policy’s primary key to uniquely identify a policy.
2. The payment period of the policy like monthly, quarterly, yearly, etc.
3. The Start Date of the Policy.
4. The End Date of the Policy.
5. The Payment Status of the Policy like paid or unpaid.
6. The Price of the Policy.
7. Whether the policy is one-time or recurring.
8. The status of the policy depicting if it is active or not.

**Cardinality:**

1. Many policies can be sold by an agent.
2. Many policies can be bought by a Customer.
3. One policy has many Transactions associated with it.
4. One Policy can have 0 or 1 claim associated with it.
5. One policy can have a combination of different Products associated with it.

**TRANSACTION**:

The entity has the details of a particular Transaction that happened while claiming or buying a Policy.

A picture containing text, pool ball

Description automatically generated

**Attributes:**

1. The T\_id is the transaction’s primary key to uniquely identify a transaction.
2. The date when the transaction happens.

**Cardinality:**

1. Many Transactions interact with one Financial Institution in our case.
2. One Transaction can be associated with one Claim.
3. Many Transactions can be done by one Customer.
4. Many Transactions can be associated with a Policy.
5. One Transaction Generates one Invoice.

**FINANCIAL INST.:**

The Financial Institution which the transaction interacts with for finance.

Graphical user interface, text, application

Description automatically generated

**Attributes:**

1. The F\_id is the Financial Institution’s primary key to uniquely identify a Financial Institution.
2. The Financial Institution’s name.
3. The Address of the Financial Institution.
4. The Phone number of Financial Institution is a Multivalued attribute.

**Cardinality:**

1. One Financial institution contains many transactions.

**INVOICE:**

The Invoice is a superclass that has details of the Transaction

Graphical user interface

Description automatically generated with medium confidence

**Attributes:**

1. The Invoice\_id is the Invoice’s primary key to uniquely identify an Invoice.
2. The Billing Address associated with the Invoice.
3. The Payment\_type like cash, credit card, debit card, etc.

**Specialization:**

There are 2 specialized classes of Invoice. We have used disjoint total participation as each Invoice covers different things and is not overlapping. This set of Invoices is exhaustive in our case and hence partial participation.

Graphical user interface, application

Description automatically generated

1. CLAIMED: The Invoice generated for Claimed Transactions.
2. PURCHASED: The Invoice generated for Purchased Transactions for instance buying or renewing a Policy.

**Cardinality:**

1. One Invoice is generated by one transaction.

**DATA SCIENCE:**

1. We are planning to save some snapshots to understand how well the business is doing and also to get insights on what should be the new Premium for a Customer given their history!

**CUSTOMER SNAPSHOT:**

One snapshot would be of the Customer to easily query the details like Total Premium paid, Total money claimed, and the policies held and renewed.

**BUSINESS SNAPSHOT:**

One snapshot would be of the Business to easily query the details like Total Sales, Total Policies bought, and Total money claimed.