

**Aim:** Write Detailed Problem statement and design dimensional modeling (creation of star and snowflake schema)

## **Theory:**

### **Data Warehouse:**

A data warehouse is a centralized repository that stores integrated data from multiple sources. It supports analytical reporting, structured and ad hoc queries, and decision-making. Data warehouses are optimized for read access and data analysis, providing a historical perspective of the data.

### **Data Mart:**

A data mart is a subset of a data warehouse, focused on a specific business area or department. It provides a more focused and faster access to relevant data for specific user groups. Data marts can be dependent (sourced from a data warehouse) or independent (sourced directly from operational systems).

### **Dimensional Modeling:**

Dimensional modeling is a design technique used to structure data for easy querying and reporting. It involves creating schemas that represent the data in a way that supports intuitive analysis. The two common types of schemas are star schema and snowflake schema.

### **Information package:**

An information package is a foundational document in data warehousing that captures the business requirements and outlines the data needs for reporting and analysis. It serves as a blueprint for designing the data warehouse, ensuring that it meets the specific needs of the organization.

### **Star Schema:**

A star schema is the simplest form of a dimensional model, where a central fact table is connected to one or more dimension tables. It is called a star schema because the diagram resembles a star, with the fact table at the center and the dimension tables radiating out from it.

### **Snowflake Schema:**

A snowflake schema is a more complex form of a dimensional model. It is a variation of the star schema where dimension tables are normalized into multiple related tables. This results in a structure that resembles a snowflake.

## Problem Statement:

Design and implement a data warehouse for the Garden Management System (GMS) to consolidate and analyze data on plant activities, health, soil conditions, and time to assess their impact on overall plant health.

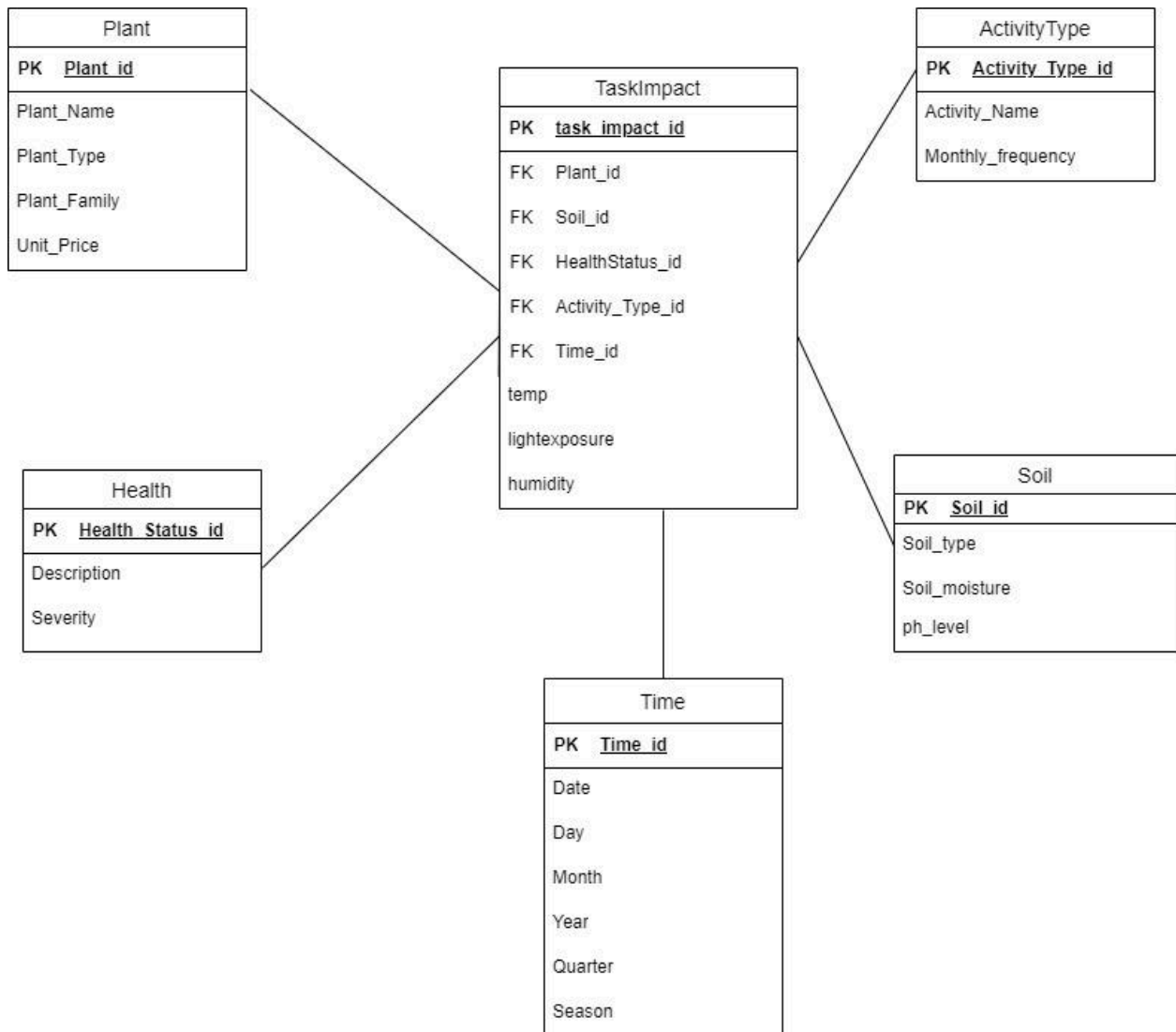
## Information Package:

### Dimensions

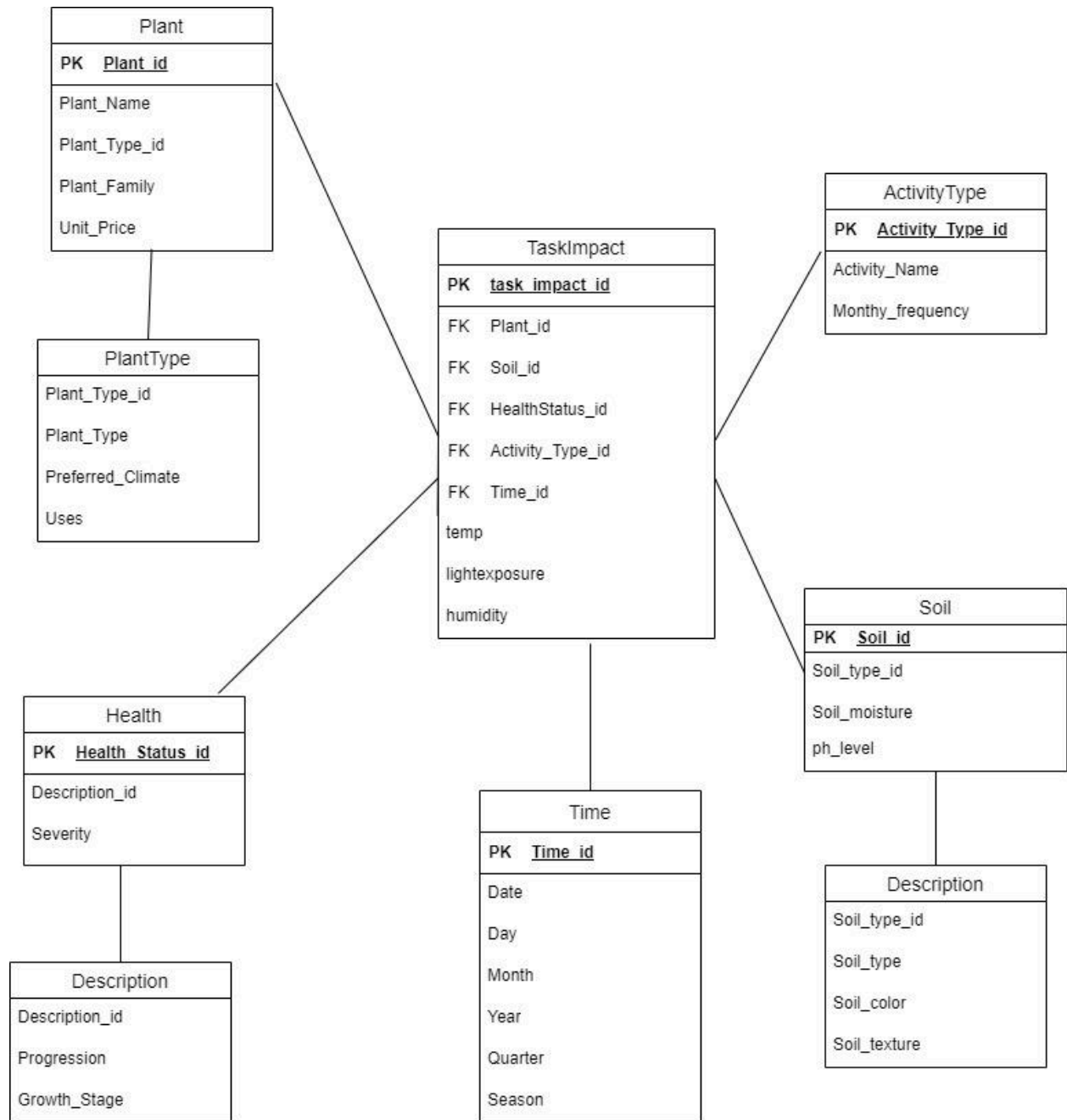
Plant	Activity_Type	Health	Soil	Time
Plant_id	Activity_Type_id	Health_Status_id	Soil_id	Time_id
Plant_Name	Activity_Name	Description	Soil_type	Date
Plant_Type	Monthy_frequency	Severity	Soil_moisture	Day
Plant_Family			ph_level	Month
Unit_Price				Year
				Quarter
				Season

**Facts:** temp , lightexposure, humidity

## Star schema:



## Snowflake Schema:



## **Conclusion:**

Implementing a data warehouse for the Garden Management System (GMS) will provide a unified platform to consolidate and analyze diverse data on plant activities, health, soil conditions, and time. This integrated approach will enable more accurate assessments of how various factors impact plant health, leading to improved garden management practices, enhanced decision-making, and optimized plant care strategies. The data warehouse will empower stakeholders with valuable insights, ultimately contributing to the overall productivity and sustainability of garden operations.