**Module 6: Assignment Final Project Report**

Pruthvi Samir Patel

Master of Professional Studies in Analytics, Northeastern University

ITC6000: Database Management Systems

Dr. Farnaz Derakhshan

Dec 11, 2023

**Introduction**

This report offers a thorough overview of the "Plant Care Assistant System" project, including all the necessary information and a list of its main elements, including an in-depth examination of the project's track from Phase 1 to the implementation of ER diagrams using Crow's Foot Notation to visually show the many relationships and entities formed during each phase.

**About The Project**

The "Plant Care Assistant System" is a flexible system made to serve farmers and other agricultural professionals in addition to plant enthusiasts. This software's primary objective is to assist users in providing proper care and nurturing for plants and crops. Farmers confront a variety of difficulties, including both natural and man-made ones, such as soil erosion, climate change, depletion of water resources, and many others (M, 2023). Whereas, Pest control, underwatering problems, and unpredictable weather are major difficulties gardeners frequently encounter (ECO gardener, 2023). So, this system will help farmers and gardeners enhance crop and plant quality, provide guidance on soil and weather, make informed choices about harvesting, provide real-time pest and disease alerts and also provide watering guidelines.

The app's pricing structure provides a variety of choices. I'm currently considering that depending on user preferences and the app's monetization approach, it may adopt a subscription-based model or a one-time purchase model with a very affordable price rate.

**Personal Connection**: My curiosity about the "Plant Care Assistant System" project originates from my ability to integrate my technology skills with a quite good understanding of agriculture. While I was growing up in India, I observed the difficulties farmers had in preserving the production and quality of their crops. Additionally, having ownership of a farm built my personal bond and opened my eyes to the complexity and typically risky nature of farming, particularly in regard to unpredictability in weather and incomplete understanding of soil conditions. Therefore, this initiative truly motivates me since it addresses problems that I have seen firsthand in the real world.

**Business Analysis**

**User Personas**

***Hobbyist Gardener:*** This persona utilises the app to efficiently maintain indoor and outdoor plants, utilising features such as real-time sensor data and care advice for the proper development of the plants.

***Farmer:*** This user depends on the app to assure crop quality, maximize crop care, track plant health, boost agricultural production, and get recommendations for harvesting.

***Professional botanist researchers:*** use the app to collect plant data for educational research. Utilizing the technically advanced data and analysis features, they depend on the app to keep an eye on a variety of plant species and gather data for their research projects.

***Supporter of Sustainable Agriculture:*** This environmentally concerned user uses the app for eco-friendly farming, gaining access to vital information on soil health, water, and resource utilisation in order to reduce environmental effects while increasing crop growth.

**Business functionality**

***Plant Profiles*:** With features like plant name, species, watering schedule, and sunlight requirements, users may create and manage plant profiles.

***Sensor Data Analysis*:** Using sensor data, the app offers intelligent care advice, including suggestions for when to water, expose plants to sunshine, and harvest.

***Monitoring for Pests and Diseases*:** Users can get notifications about problems with plants as well as recommendations for remedies.

***Harvest Recommendations:*** Users receive immediate updates for optimal harvest seasons, as well as personalised recommendations on appropriate harvesting practices.

***Weather Integration:*** Plant care advice and harvesting window timings are based on weather data.

***Data Privacy:*** To safeguard user and plant data, the app implements robust confidentiality and safety protections.

That is what I had in mind for the application level. A powerful Machine Learning recommendation algorithm generates the values stored in the **HarvestRecommend** table attributes and the Treatment-Recommend attribute in the **PestDiseaseRecord** table. This ML model, created for a specific purpose, leverages powerful algorithms to analyse relevant data and provide optimised ideas for harvesting practices and treatment recommendations. However, because this subject is the foundation for database design, I will continue to concentrate on database design.

**Data Creation, Storage and retrieval**

1. ***Data Creation***

* By creating plant profiles, scheduling care, entering data from sensors, and reporting pest or disease problems, users generate data.
* Real-time data on plant properties is generated by sensors.
* The system will be connected with weather data from external data sources.

2. ***Data Storage***

* User-specific databases hold user data, such as schedules, problems, and plant profiles.
* Sensor data will be kept in a database, with each reading associated with the related plant.

3. ***Data retrieval***

* Using the app, users may access their plant profiles and maintenance information.
* Real-time updates are available for tracking the health of plants by retrieving sensor data.
* For weather-based guidance, will retrieve meteorological data, both past and present.
* The system will suggest ideal crop harvest dates to harvest crops based on data analysis.

**Business Logic and Rules**

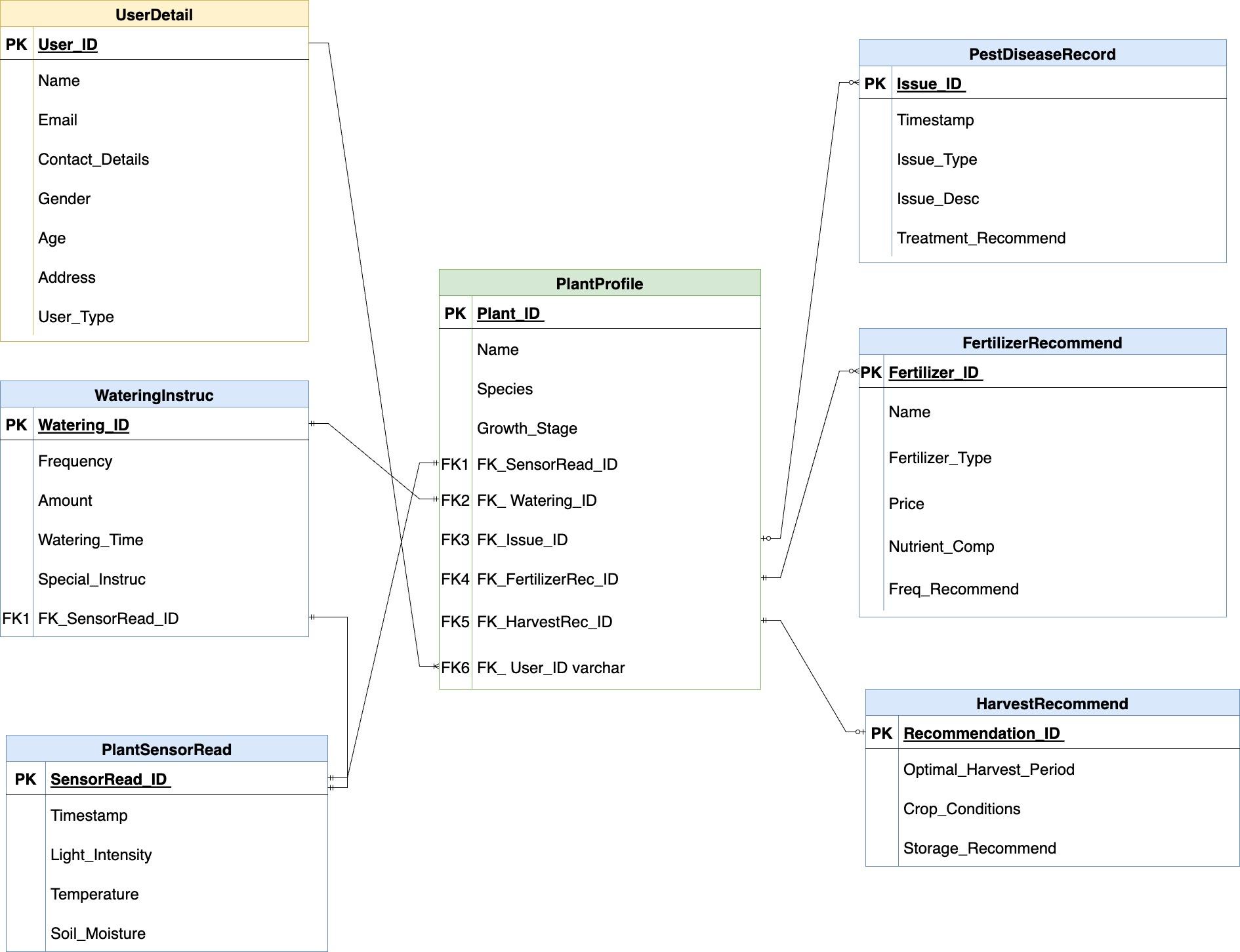
Business rules ensure tasks are completed consistently since particular criteria must be met in order for a task to be completed (*What Are Business Rules?*, n.d.). Here The Plant Care Assistant System's business rules enable uniform and successful task completion by providing defined criteria that must be met for each activity. These rules serve the system's goal by standardising actions and ensuring that plant care duties are completed consistently and optimally, resulting in healthier and more productive plant development.

1. The **PLANT PROFILE** is linked to several **SENSOR READINGS (PlantSensorRead)**, allowing various sensor readings to be associated with specific plants. This helps in the monitoring and study of plant situations in real time.
2. Each **PLANT PROFILE** is linked to **PEST AND DISEASE RECORDS (PestDiseaseRecord)**, allowing users to track diseases and pest problems and remedies for specific plants.
3. A **USER** can manage numerous **PLANT PROFILES**, allowing users to manage a variety of plants in their collection, all of which have their own set of care requirements.
4. **WATERING GUIDELINES (WateringInstruct)** provide particular directions for watering various **PLANT PROFILES**, taking into account aspects such as the development stage and environmental circumstances to ensure optimal care suited to each plant's needs.
5. **HARVEST RECOMMENDATIONS (HarvestRecommend)** suggest optimal harvest timing and conditions for specific **PLANT PROFILES**, supporting users in selecting the best time to harvest their crops.
6. **FERTILISER RECOMMENDATIONS (FertilizerRecommend)** is associated with each **PLANT PROFILE**, where the plant profile in the "PlantProfile" table can have zero or more associated fertilizer recommendations in the "FertilizerRecommend" table.  
   It provides several fertiliser solutions that suit the individual demands of different plant varieties.
7. The **WATERING GUIDELINES (WateringInstruct)** table has an association with the **SENSOR READINGS (PlantSensorRead)** table to continually alter watering recommendations depending on real-time sensor data, such as soil moisture and temperature. Each record in the Watering Instructions table corresponds to a distinctive Plant Sensor Read.

**ER Diagram**

An entity-relationship diagram, or ER diagram, is required for modelling database data. It is the foundation on which a database is created (*Why Do You Need an ER Diagram?*, 2021).

Through a clear, interconnected structure, the ER diagram visually defines the links between data elements in the Plant Care Assistant System, simplifying database architecture and boosting system functionality. It aids in comprehending how various components—such as plant profiles, sensor data, and task scheduling—interconnect, providing a clear representation of how information flows inside the system.



**Relationships between tables**  
  
***1. UserDetails - PlantProfile:***

**Relationship**: One-to-Many (1:N)

**Explanation**: Although a user may have more than one plant profile, only one user is associated with any given plant profile.

***2. PlantProfile - FertilizerRecommend:***

**Relationship**: One(mandatory)-to-Many(optional) (1:N)

**Explanation**: There must be at least one related fertiliser recommendation for every plant profile, and many fertilizer recommendations may be made for the same plant.

***3. PlantProfile - HarvestRecommend:***

**Relationship**: One(mandatory)-to-One(optional) (1:N)

**Explanation**: There can only be one related harvest recommendation per plant profile, but there must be exactly one. Every recommended harvest has a single plant profile associated with it.

***4. PlantProfile - PlantSensorRead:***

**Relationship**: One-to-One (1:N)

**Explanation**: Every plant profile has a unique plant sensor reading associated with it, and every plant sensor reading corresponds to a certain plant profile.

***5. PlantProfile - WateringInstruc:***

**Relationship**: One-to-One (1:N)

**Explanation**: Every plant profile is exclusively connected to a single set of watering instructions, and each set of watering instructions is connected to a particular plant profile. making certain that every plant has a unique set of guidelines for irrigation.

***6. PlantProfile - PestDiseaseRecord:***

**Relationship**: One(optional)-to-Many(optional) (1:N)

**Explanation**: Every plant profile may have zero or one pest and disease record, and many pest and disease records can be found for the same plant, or none at all.

***7. WateringInstruc - PlantSensorRead:***

**Relationship**: One-to-One (1:1)

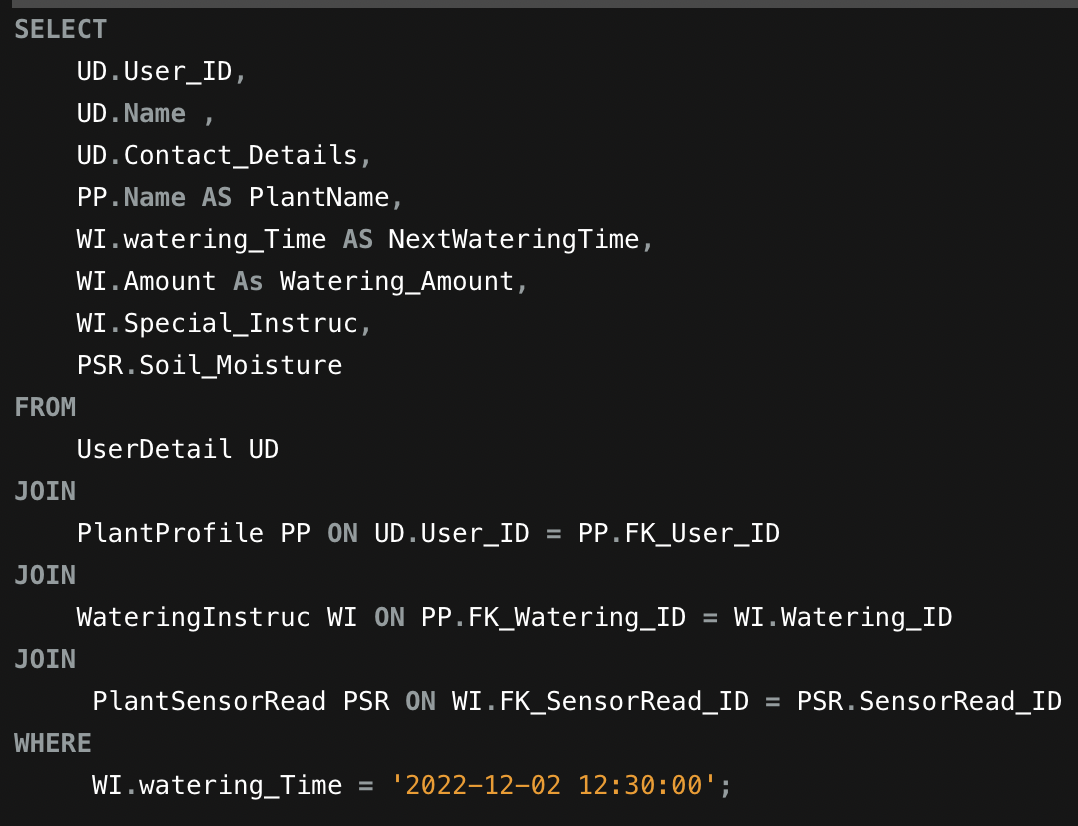
**Explanation**: Each set of watering instructions corresponds to a single plant sensor signal, indicating a direct and unique relationship between them.

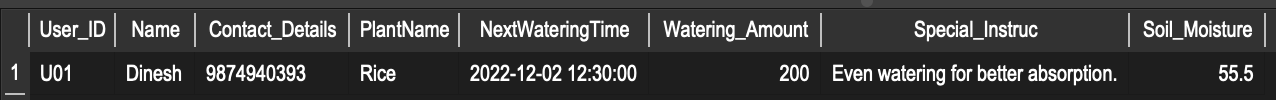
**Hands-on with Database**

**Statements to Pull Data from Database**

Every statement listed below includes multiple application-specific usage cases.

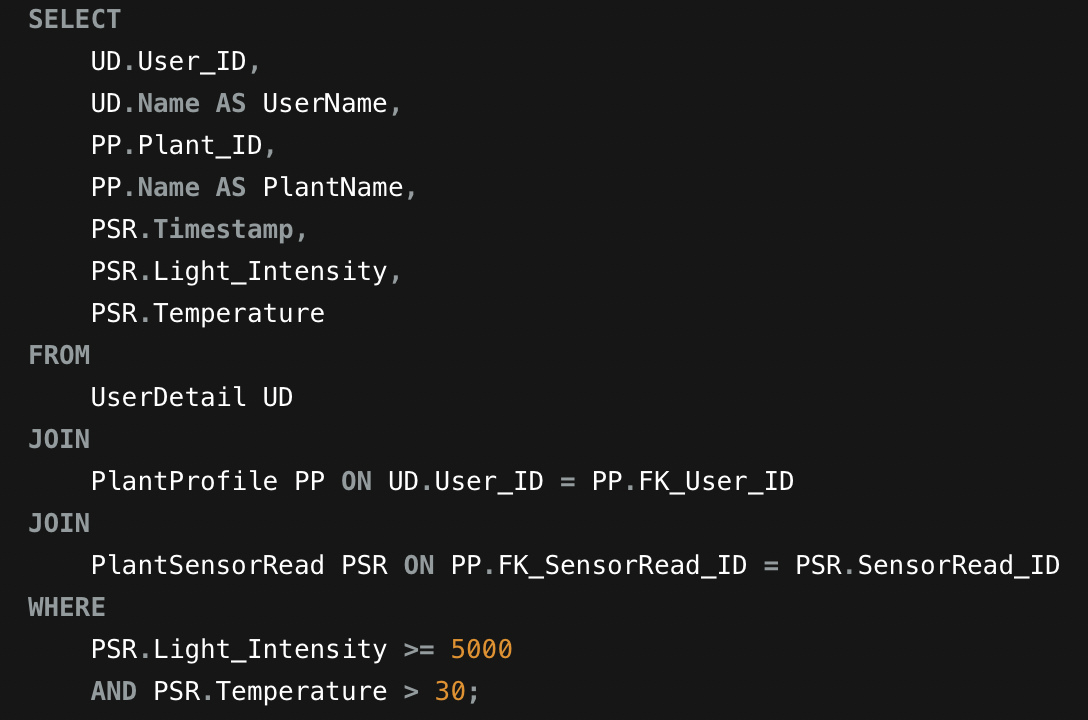
1. Assuming it's 12:25 PM on December 2, 2022, the system needs to identify all users who have a watering scheduled at 12:30 PM to notify 5 min before using the notification service.

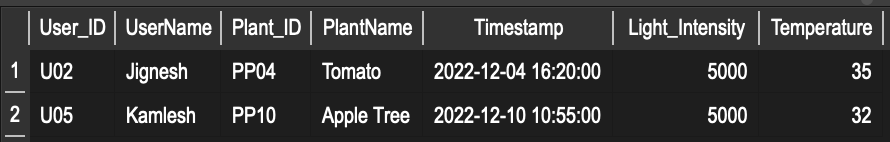




*Note.* The above-mentioned SQL query is meant to collect data regarding users who have upcoming watering schedules. I specified a fixed time in the WHERE clause, but we can also utilise SQL methods like GETDATE() to dynamically calculate and select records depending on the current date and time. I've combined four tables to accomplish this: UserDetails, PlantProfile, WateringInstructions, and PlantSensorRead. After receiving the filtered data, the list of users will be utilised to generate notification alerts for their upcoming watering plans.

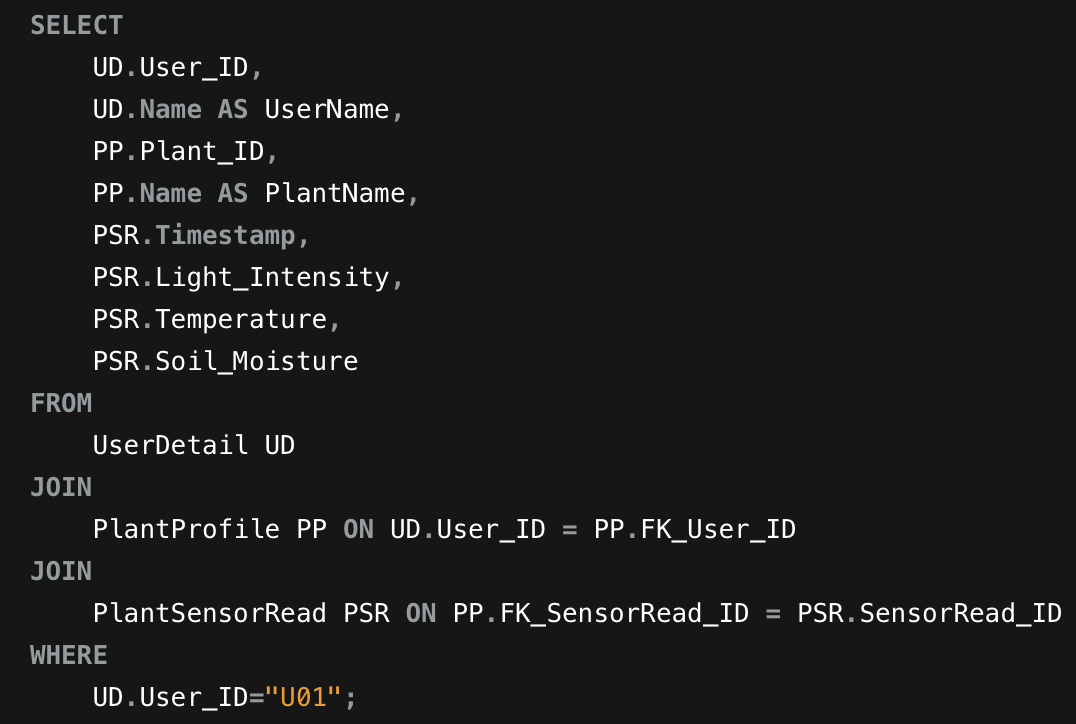
1. Alerting users when their area light intensity is 5000 or above and temp is more than 30 degrees.

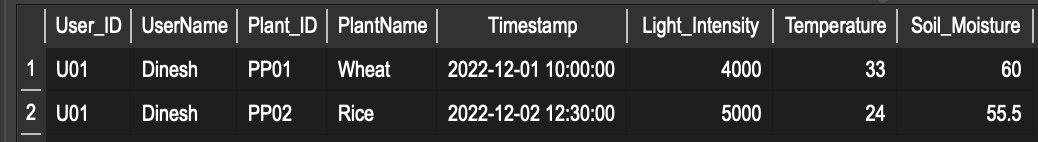




*Note.*  The SQL query mentioned above fulfil the use-case, which is to obtain user details where their plants' sensors detect that the temperature is above 30 and the light intensity is 5000 Lux or more, it joins three tables: UserDetails, PlantProfile, and PlantSensorRead for the same purpose. Additionally, the where clause is used for fulfilling the temperature and light intensity conditions. After obtaining user information, the system can help users take better care of their plants by providing more general guidance and, if practical, moving the plants to an area with lower light levels. It is only for plants whose growth is hampered by high temperatures or intense light.

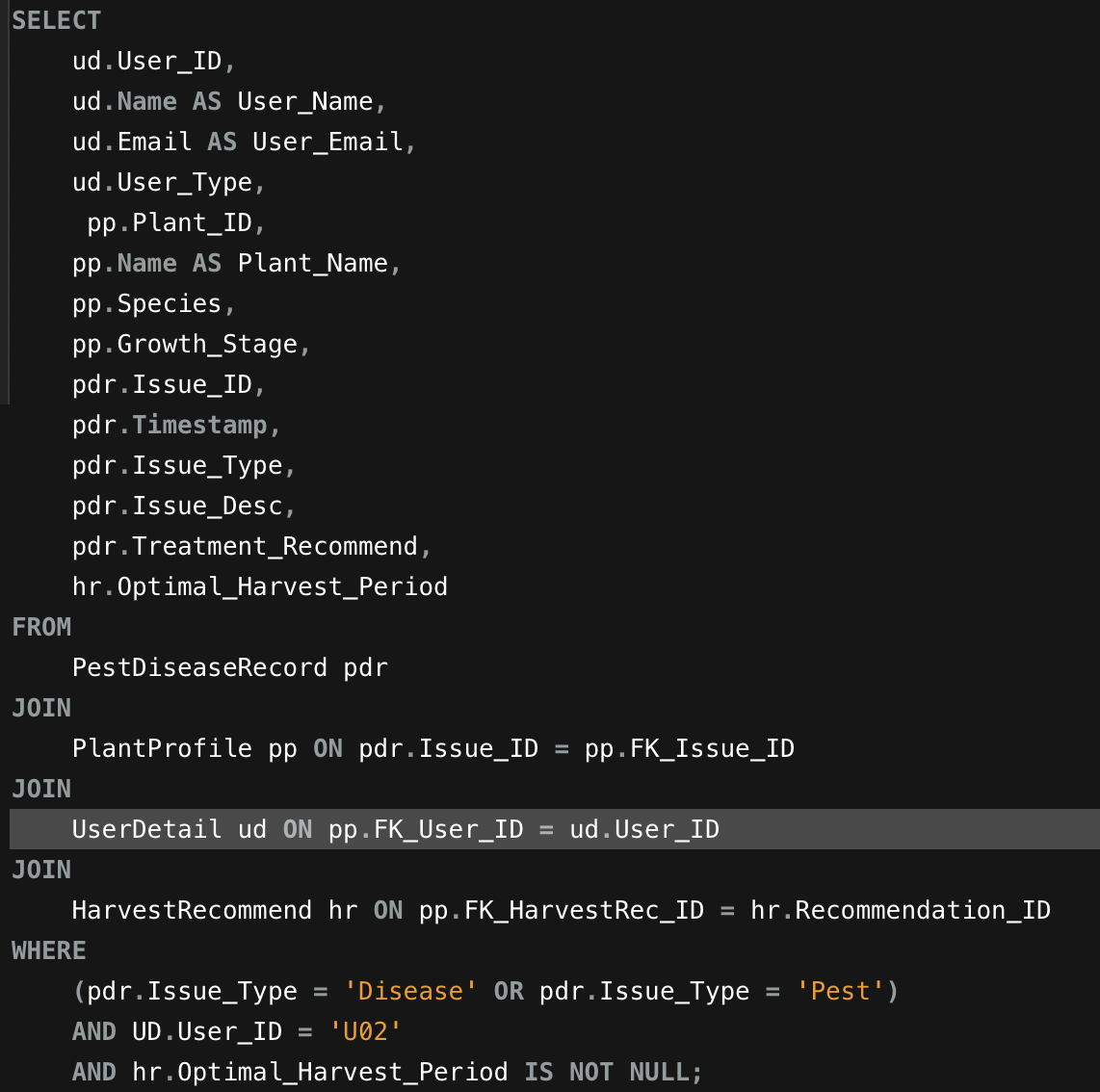
1. Assuming User U01 want to check the sensor reading of his/her plants.

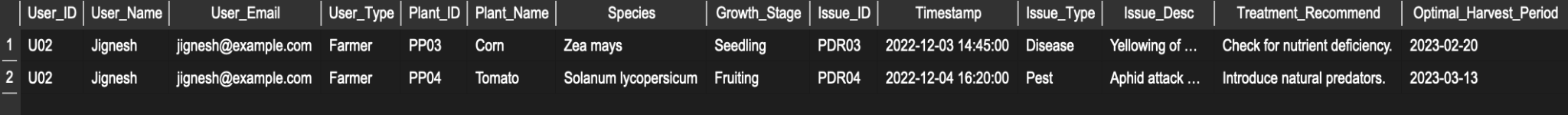




*Note.* The SQL statement shown above retrieves all the information a user would need to know about their planted plants' sensor reading. This SQL query will execute and deliver the necessary information based on the user ID specified in the Where clause when the user lands on the page where the plant's sensor record has been displayed. Using relevant IDs, this query joins three tables: UserDetails, PlantProfile, and PlantSensorRead.

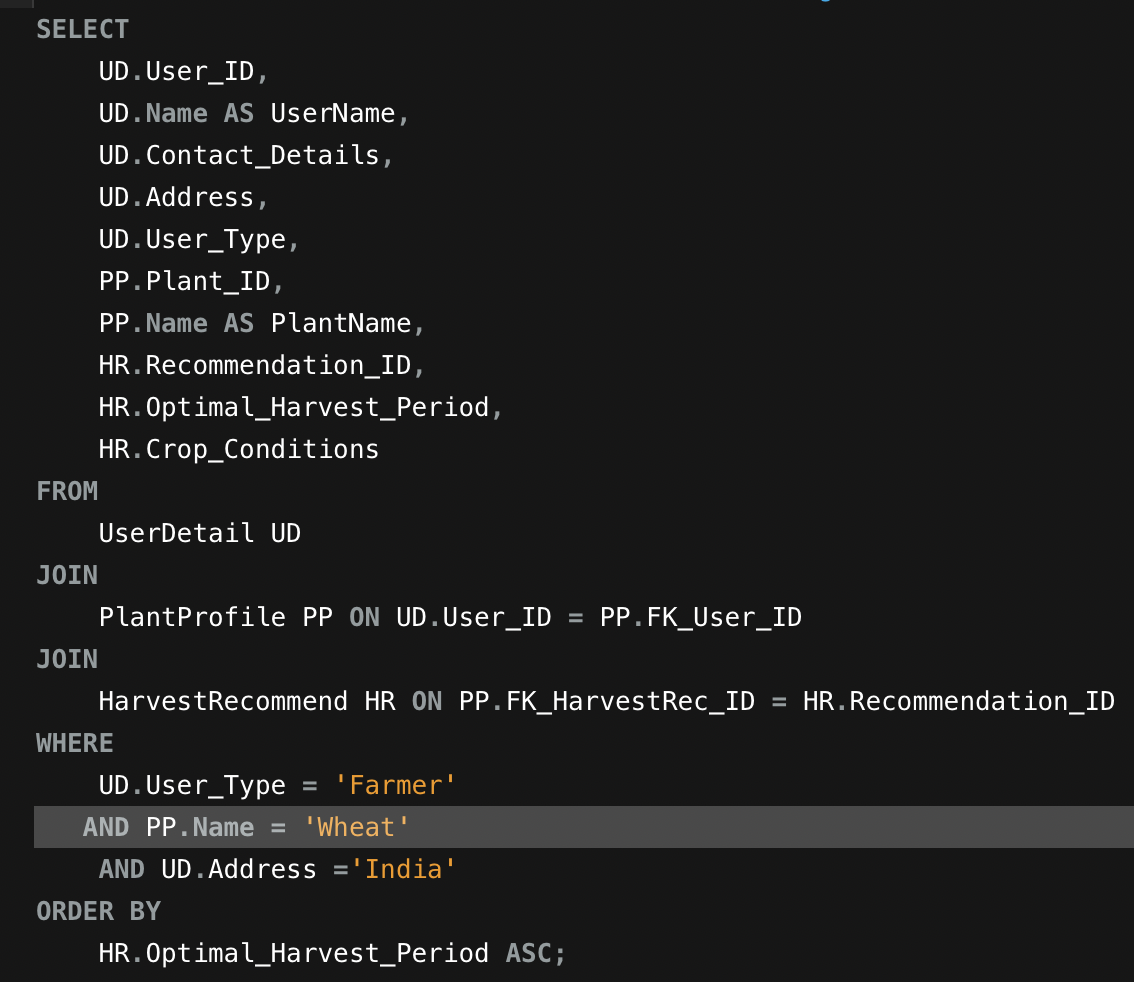
1. Assuming that the user with ID u02 has created pest and disease records and that the treatment recommendation is now ready, it is time to present the contents of the treatment recommendation within the harvest period time so that the user can take quick action.

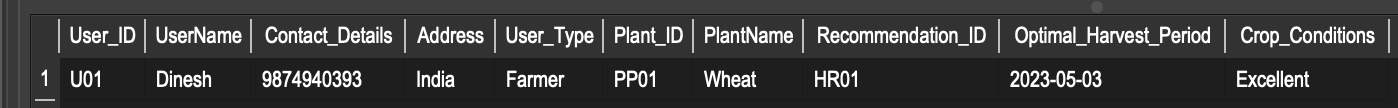




*Note.* The above-mentioned SQL query gets information that will give the user insight into suggested treatments for their plants. An ML model generates a recommendation when the user provides details about the pest or disease issue. The user is then provided with every relevant detail, including the appropriate harvest time period. This allows the user to do what has to be done in advance of harvest time. Four tables — PestDiseaseRecord, PlantProfile, UserDetails, and HarvestRecommend—are joined to do this. Also, I used the WHERE clause to filter records about diseases or pests for a specific user and to display only records with an appropriate harvest time period (using NOT NULL constrain).

1. The case where a botanist/researcher is looking for information on farmers who are growing wheat and have their next harvest planned for the nearest possible period.

****

****

*Note.* The above-mentioned SQL query gets data for all farmers in India growing wheat, taking into account the nearest harvest period indicated in the WHERE clause and sorting the results in an ascending order using the ORDER BY clause. Three tables — UserDetails, PlantProfile, and HarvestRecommend—are joined using relevant IDs to collect all necessary information. The collected data will be made available to botanists and researchers who might want the same report for their studies. This will allow them to contact farmers to talk about their needs.

**Suggested Metrics and Analysis**

This system would provide different readings and reports to each type of user, and stakeholders could also gain many insights from these metrics.  
  
1. Gardener Metrics

* Plant Profiles Created (count)
* Involvement with Sensor Data (frequency per week)
* Treatment Actions Taken (count)

2. Farmer Metrics

* Improvement in Crop Quality (% Improvement)
* Harvest Optimisation (yield per acre) [evaluates how well harvest planning is working]
* Watering Economy (amount of water used per plant) [monitors the farmer's methods for water conservation by calculating the quantity of water used for every crop.]

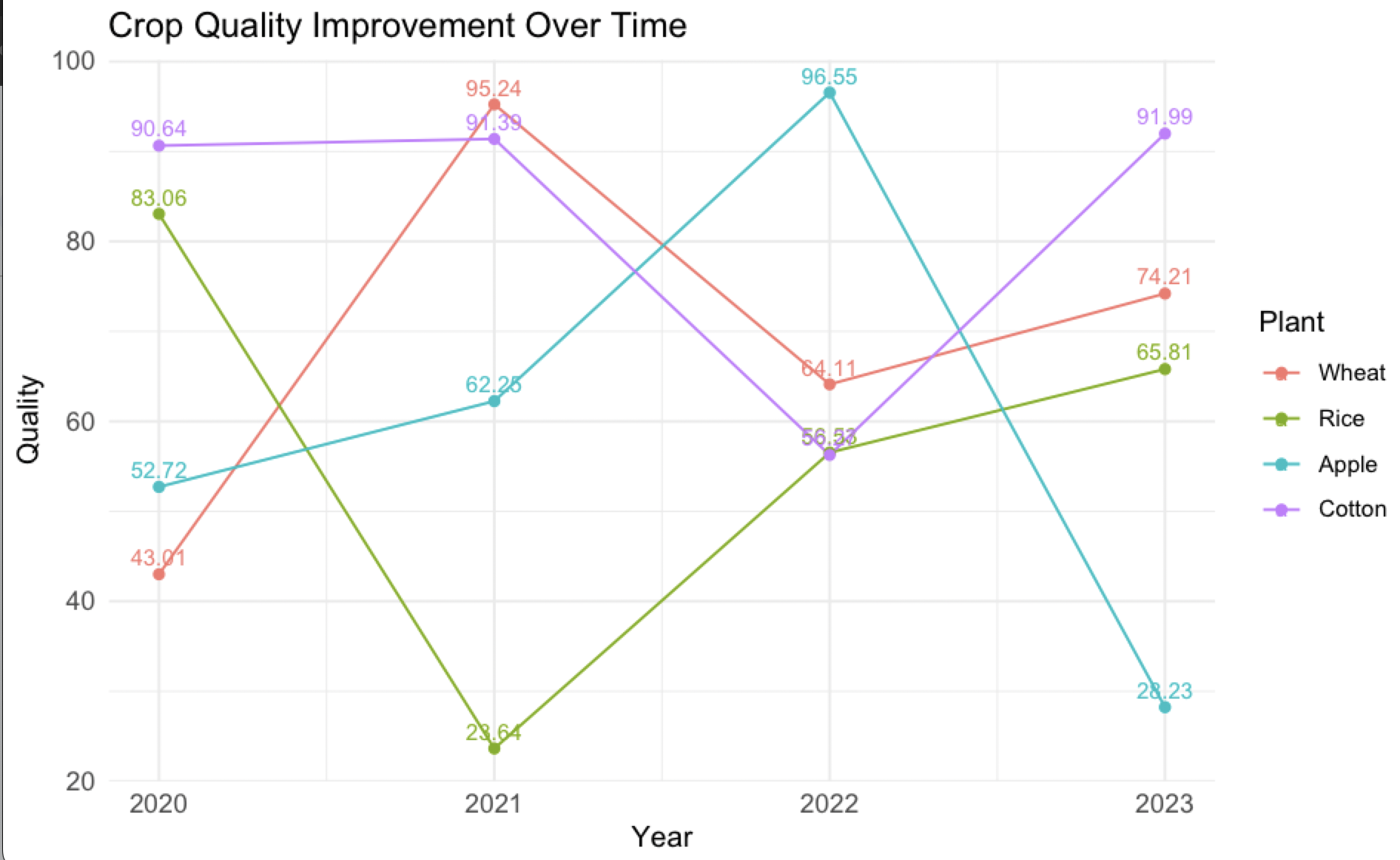
3. Botanist Researcher Metrics

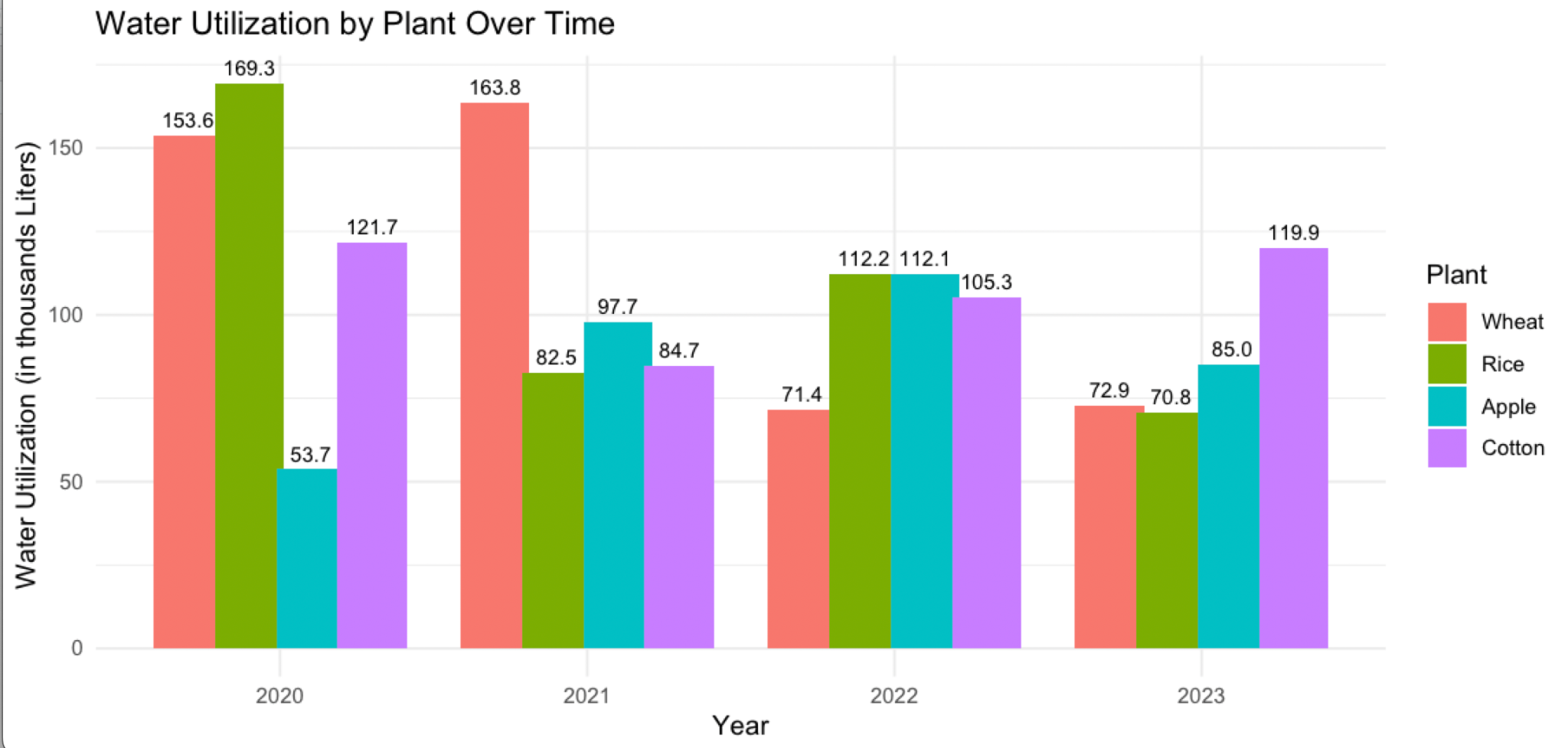
* Collaborative Field Studies (count)
* Variety of Plant Species Monitored (count)

4. Supporter of Sustainable Agriculture

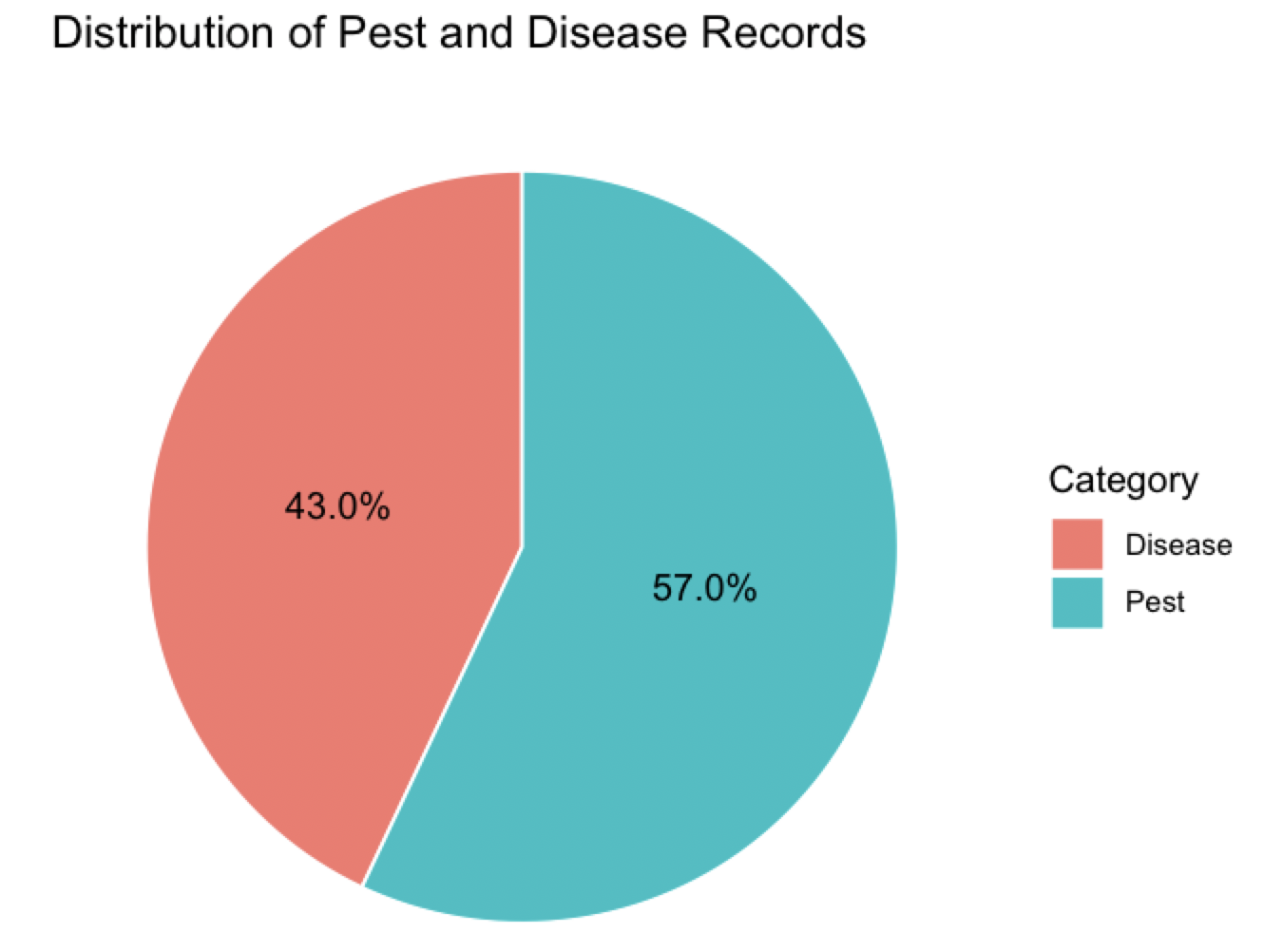
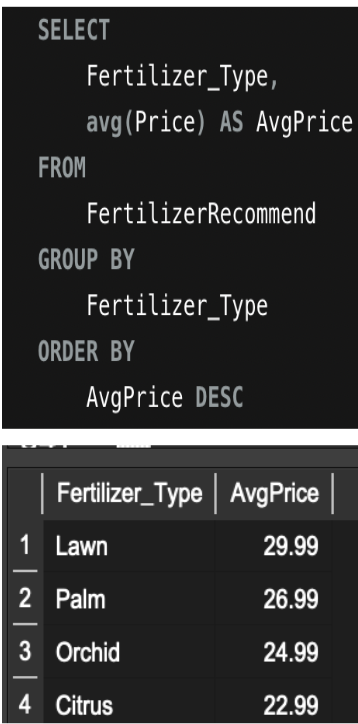
* Environmental Impact Reduction (% reduction) [demonstrating their dedication to environmentally friendly farming methods.]
* Sustainable Fertilizer Impact Rating (Scale 1-10) [user satisfaction with how recommendations for sustainable fertilisers affect the health of plants,]
* Resource Utilization Insights (count)

**Few Analysis Examples**

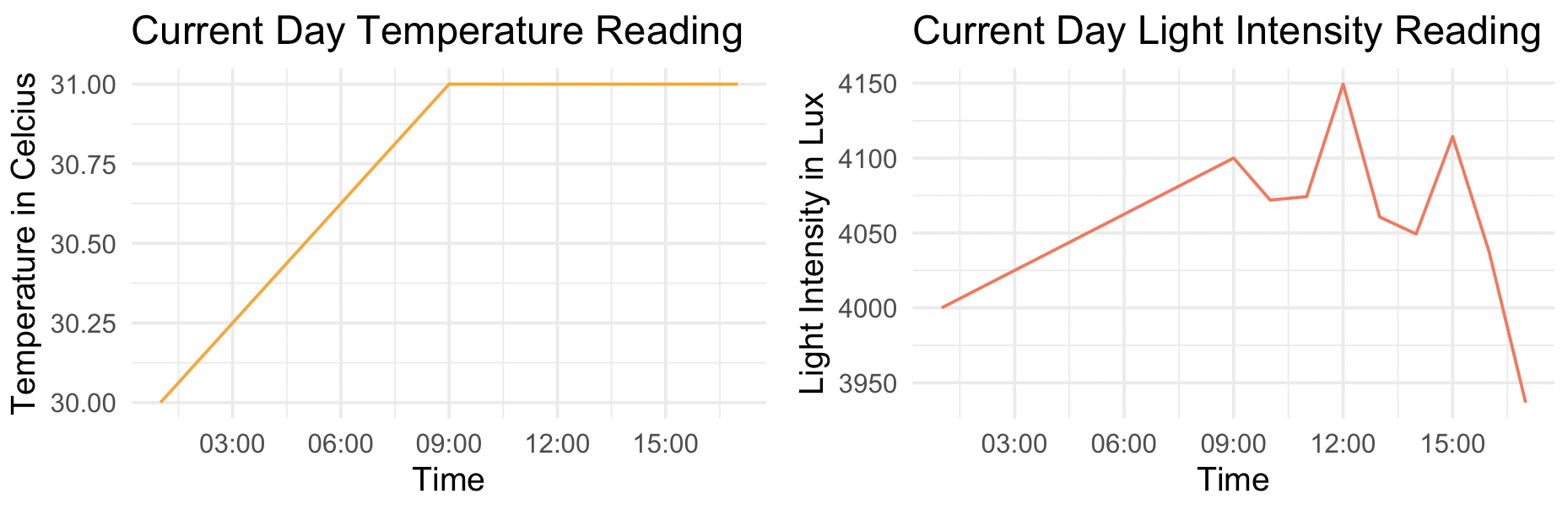
****

****

*Note.* With the help of the trend on the left-hand graph, users can quickly determine what crop type would be most advantageous for them based on crop quality reports. For instance, cotton consistently produces good-quality crops which continues to rank among the best in 2020 and 2023. Another graph discusses water use, and it is evident that cotton uses the least amount of water while maintaining the highest quality, making it a viable option. To provide users with a wide range of options for planting various crops in the near future, team members of this system can also concentrate on underperforming plant types and improve the recommendation criteria.

** **

*Note.* By examining the pie chart, researchers and botanists can learn a lot about how many other users are having trouble with their plants. For example, 57% of users report having problems with pests on their plants. By focusing on these cases, they can work with those users to find solutions that will help them introduce new research work into the market and provide support to farmers. Also, running an SQL query to identify all the records of fertiliser types with high average prices could assist the internal team in adding more affordable fertiliser types to any category where the average price is high.

****

*Note.* To protect light-sensitive plants, gardeners in particular can benefit from the temperature and light intensity readings as they can identify trends in these two variables and relocate their plants to areas with lower light levels. The user learns how the temperature rises from 30 to 31 degrees Celsius from the first graph. Additionally, observe how the light intensity fluctuates between 1:00 AM to 4:00 PM, going from 4000 lux to almost 3950 lux. Nevertheless, the system will generate suggestions based on this reading, so they won't need to make any decisions on their own.

Stakeholders have access to all analysis reports, allowing them to compare how the system is operating based on these reports rather than just monitoring data. This allows them to make suggestions and work towards improving the design pattern as needed to achieve the best possible system functioning. The analyses mentioned above are just a few examples of the many that could be done with the current system. For instance, how many plants each user has based on a particular location, when is the best time of year to harvest based on readings from previous years, and many more could be done using the metrics mentioned above.

**Security Concerns**

Regarding security concerns, several considerations are being made, including

1. Role-based access control in the UserDetail table can be used to limit sensitive data, like user contact information, and make sure that only individuals with the proper authorization can view and edit this information.
2. Focusing on Data integrity and making use of data validation techniques in the table, such as PestDiseaseRecord, to keep the accuracy and dependability of information about diseases and pests by preventing unauthorised or incorrect entries. This is relevant for other tables as well.
3. Continuous Monitoring and Troubleshooting will be adapted to enable quick identification and resolution of possible problems e.g. routinely checking the PlantSensorRead table for anomalies in temperature and light intensity, to find issues affecting plant health.
4. Prioritize to Create a careful data recovery and backup plan that will effectively minimise downtime and guarantee data integrity by returning the database to its pre-threat state whenever a security threat arises.

This system does contain a lot of sensitive data, which must be protected with security measures for the system to operate properly. Those data are:

1. User Details: To safeguard sensitive user information, encryption techniques will be used to store it in the UserDetail table, providing an extra degree of privacy. Volume: very little personal data of each user like e-mail, password, contact etc.
2. Security of Data for ML-generated Recommendation: Since tampering with these records would pose a significant risk, a secure, hashed format for data storage in the PestDiseaseRecord and HarvestRecommend tables will be implemented to preserve the privacy of harvest and treatment recommendation records.Volume: Applies to all of the entries in these tables.
3. Plant Senor Data: Encrypting the content of the PlantSensorRead table to protect plant-specific information is one method of protecting sensor-read data. If a hacker can manipulate those data, it will have a significant impact on the performance of ML models, and the users such as farmers, will suffer significant financial loss due to incorrect sensor values and inaccurate recommendations because they work with large crop numbers. Volume: Each plant entry will only have one record.

**Database Architecture Requirements**

**Client-Server Architecture Solution**

To enable communication between the user interface and the database, the project will adopt Three-tier a client/server architecture. In my opinion, it would be an appropriate solution for this system.

**Presentation Layer (Client):** The client handles the user interface, allowing interactions on mobile devices via a user-friendly interface.

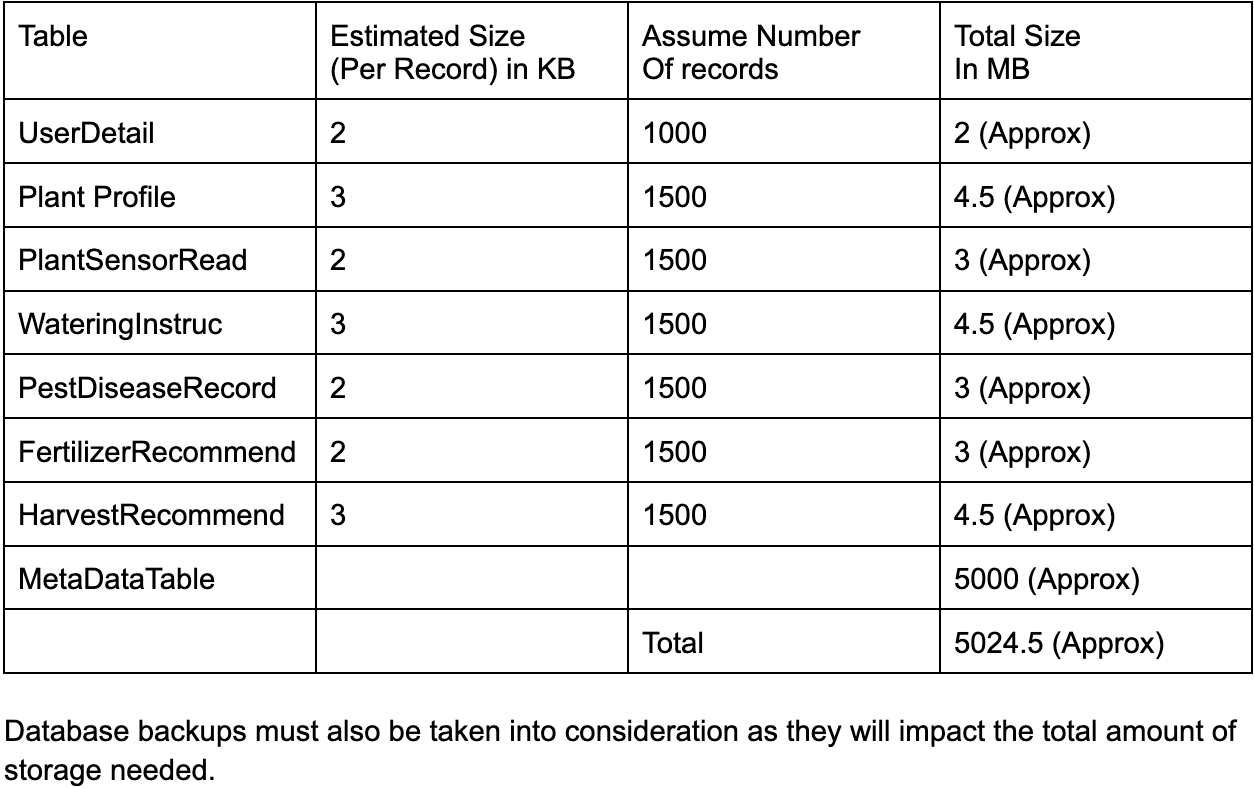
**Application Layer (Server or Middleware):** It handles application logic, evaluates real-time sensor data, provides automated suggestions, and manages user requests. It serves as a bridge between the presentation layer and the data layer.

**Data Layer (Database Server):** It includes the database server, which stores and manages plant profiles, sensor data, suggestions, and other essential information.

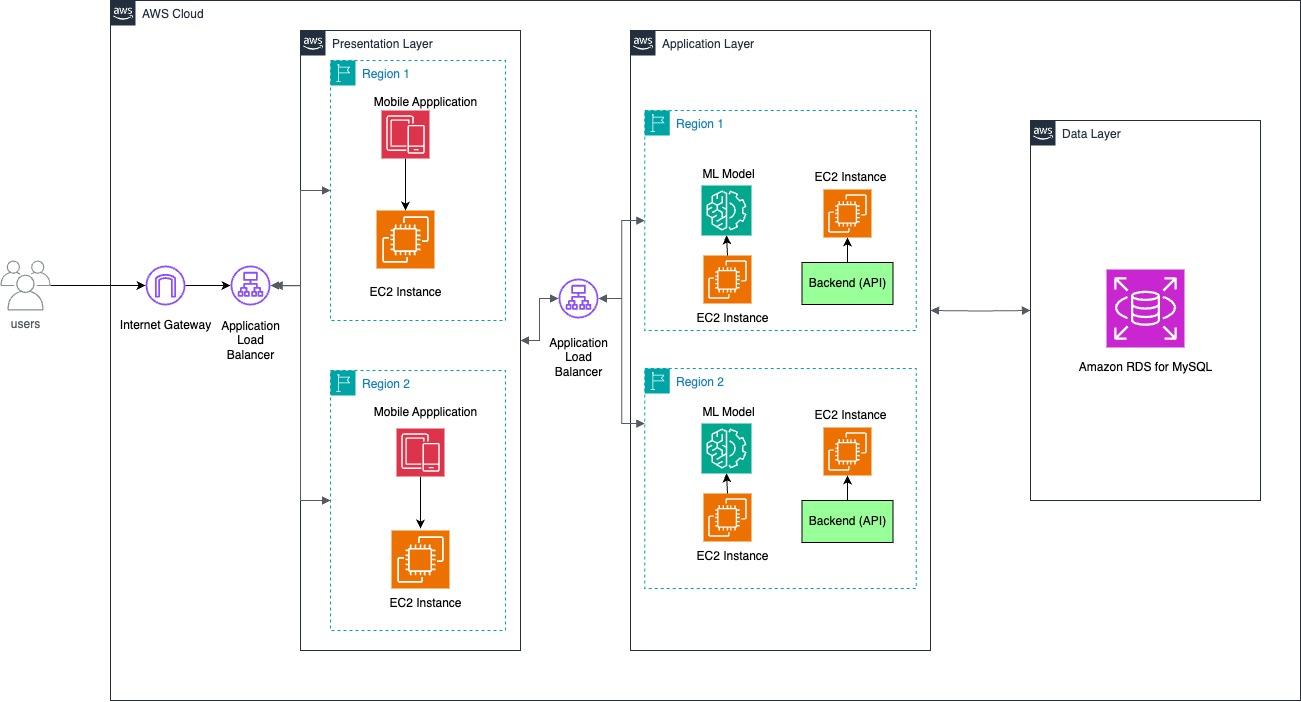
For this system, I make use of AWS cloud-based database hosting due to its inherent benefits in terms of scalability, accessibility, and ease of maintenance. AWS also provides benefits like robust security, automated backups, and has a worldwide reach of AWS data centres (*Purpose-Built Databases on AWS | Amazon Web Services*, n.d.). That could result in a comprehensive and dependable infrastructure for handling large-scale databases. The flexibility of scaling resources on-demand, as well as the ability to easily connect with other AWS services, add to the overall system architecture's efficiency and durability.

**Storage Plan:**

Based on the assumption the target audience will be approximately 1000 users for the first few months.



It might be necessary to scale up database storage space to the terabyte range as the system becomes more widely used. Given that there are an estimated 150 million farmers in India alone (Agarwal, 2021), and even just 5% of them will use the system, which will be equivalent to 7.5 million farmers would be a big number. Apart from this farmers from other nations would also add to the total number of users once the system reaches to international Market, as would botanists, researchers, and gardeners. Also considering data backups and additional logs, storage space would undoubtedly need to be increased.

****

*Note.* Suggested three-tier architecture for the Plant Care Assistant system.

**Key Learnings and Future Considerations**

This final project taught me important considerations both before and after the

development of a database system. Before beginning, it became clear that understanding

business logic and rules was necessary. I learned how to create an ER diagram, which

serves as a blueprint with defined table relations, using tools like draw.io. Following that, I learned about SQL queries which were used for real database development. Following development, I learned about the importance of client-server architecture, security, and privacy concepts, allowing for more effective database management.

In the future, I hope to work on machine learning (ML)-based recommendation

models that will be able to offer important advice on crop harvesting as well as appropriate

treatment recommendations for pests or diseased plants. Additionally, I will try to create a simple

interface prototype that, at the very least, conveys the concept of how it would function in real

life and could be helpful when pitching to investors down the line.

**References**

Agarwal, S. (2021, October 4). *Harish Damodaran, Samridhi Agarwal write: Revealing India's actual farmer population*. The Indian Express. Retrieved December 3, 2023, from <https://indianexpress.com/article/opinion/columns/revealing-indias-actual-farmer-population-7550159/>

ECO gardener. (2023, August 7). *8 Common Gardening Challenges and How to Solve Them*. ECOgardener. Retrieved November 20, 2023, from <https://ecogardener.com/blogs/news/8-gardening-challenges-and-how-to-solve-them>

M, D. (2023, March 13). *11 Major Problems Faced By Indian Farmers In Agriculture In 2023*. Bighaat. Retrieved November 20, 2023, from <https://kisanvedika.bighaat.com/news-updates/11-major-problems-faced-by-indian-farmers-in-agriculture-in-2023/>

*Primary and Foreign Key Constraints - SQL Server*. (2023, September 26). Microsoft Learn. Retrieved December 3, 2023, from <https://learn.microsoft.com/en-us/sql/relational-databases/tables/primary-and-foreign-key-constraints?view=sql-server-ver16>

*Purpose-Built Databases on AWS | Amazon Web Services*. (n.d.). AWS. Retrieved December 7, 2023, from <https://aws.amazon.com/products/databases/>

*What are business rules?* (n.d.). IBM. Retrieved November 20, 2023, from <https://www.ibm.com/topics/business-rules>

*Why Do You Need an ER Diagram?* (2021, February 19). Vertabelo. Retrieved November 20, 2023, from <https://vertabelo.com/blog/why-need-an-er-diagram/>