The questions are:

* How would you define the variables in the model?
* Are you dealing with employees’ hours or employee assignments?

As a data scientist supporting staffing decisions for a new manufacturing division, defining the variables accurately is critical to building an effective optimization model.

**Defining the Variables**

In this scenario, we decided that the model revolves around **employee assignments**, such as:

* Which employee is assigned to which shift or task
* Whether a role is staffed (binary: 0 = no, 1 = yes)
* Number of employees assigned per task or shift

These decisions can be represented using **integer or binary variables**, especially since partial assignments (e.g., assigning 0.6 of an employee) do not make sense in real-world staffing.

**Integer vs. Linear Optimization**

* **Linear Programming (LP**) enables continuous decision variables, which is particularly advantageous when dealing with factors such as employee hours or resource levels that can be **fractional**.
* **Integer Programming (IP)** is necessary when decisions are **discrete**, such as whether to assign an employee or not (0/1) or how many whole individuals to assign to a task.

**Conclusion**

The data scientist's decision about how to approach the problem and what techniques to use depends on acquiring as much information about the scenario at hand as possible. The assumption here is that assigning people is appropriate, even though we do not have much more information about the scenario.

In our assumed scenario, staffing decisions involve assigning people (whole units) to tasks or shifts, so integer programming is a more appropriate method. If we were instead allocating work hours across employees (and partial hours were acceptable), then linear programming could be used.