1. **Project Description**

The client Michelle Paret, an avid baker, wants to improve her skills to bake the perfect chocolate cookie. This perfect cookie is described by our client as an almost symmetrical cookie with a short diameter and a very crunchy texture. Our client had tried to make chocolate cookies by experimenting with different factors and keeping others constant to have the least possible amount of variation.

The main goal of the study is to find through statistical analysis a way to minimize the diameter of the chocolate cookies with the elements our client used. She had 2 identical baking sheets, same scooper, same bowl, same kitchen, same temperature, and she baked 24 chocolate cookies in the same day, baking each at a time following the same recipe from the bag of chocolate cookies.

This experimental study was randomized, and it was replicated because each test was repeated 3 times. She collected 24 observations, of which 8 were unique tests and they were replicated 3 times each (as mentioned before). Our population of interest is any chocolate cookie she baked in her oven, where our parameter of interest is μ.

* 1. **Research Questions**

*Question1: Do fat, flour or chill time affect average cookie diameter?  
Question 2: For any factors (fat, flour, and chill time) that are significant, what are the best settings to minimize cookie diameter*?

* 1. **Variables**   
     We have 4 different variables. Our independent variables are 3: *ChillTime, Fat, and Flour*.
* *ChillTime*, as the name says it, is the amount of time the dough was left chilling in the fridge before baking it. Even though is numerical, in this case it is a categorical variable because there are two levels: 135 minutes and 360 minutes.
* *Fat* is the type of fat that was used in the experiment. This variable is categorical, and our client used two types of fat: margarine and butter*.*
* *Flour* is our last independent variable, and it consists of the amount of flour in ounces that our client used for the chocolate cookie. Even though this variable (as *ChillTime*) is naturally numerical, we only have two factors: 1 ounce and 1.75 ounces, therefore is a categorical variable.
* Our response variable is *Diameter,* which ranges from 5.4 centimeters to 10 centimeters. The four variables in the dataset are summarized in Table 1 below.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Type | Description | Levels and Ranges |
| ChillTime | Categorical | Number of minutes that the dough was left chilling (minutes) | Two levels: 135 and 360 |
| Fat | Categorical | Type of fat that was used (type of fat) | Two levels: Margarine and Butter |
| Flour | Categorical | Amount of flour (oz) | Two levels: 1 and 1.75 |
| Diameter | Numerical | Diameter in centimeters of the diameter (cm) | 5.4 to 10 |

*Table 1: Summary of variables for the chocolate cookie study*

1. **Exploratory Data Analysis (EDA)**