Exercise 2:

Problem Statement: In a conceptual iPhone factory, we are seeing a dip in yield due to a previous unseen issue. The engineering team has narrowed down the root cause analysis to three specific sub modules, and requested our help to see if there are any additional signals in the data related to these modules. We have gathered various data related to these phones and modules, and we are interested in providing a direction to the engineering team based on this data. The failed units are labeled, and use any methodology you'd like for this analysis.

There is no right or wrong answer to this question. What we would like to see is your approach and methodology to this problem, and how you present your findings to cross-functional teams.

About the data set: - The 'data_exercise2.csv' file contains the data we want to to study

Select variable explanation:

- STARTTIME: Time when the error was found

- TESTDURATION: How long the test lasted

- MACHINEID: A module related machine

- MACHINEID_TESTER: A module related machine tester

- LINE_ID: The line ID that the unit went through during assembly

- MODULE2_FACTORY: Specific factory location of module 2

- WoMs: The day when the module was made

Delivery expected:

You can use either Python or R to complete the analysis and final delivery is expected to be in a simple report format (formula, description and visualizations). We will set up time to have you walk us through all the

findings (you present findings to the team)