**Identifying the Problem:**

* When a patient is diagnosed with a tumor it does not mean that one has cancer. A tumor is described as an abnormal growth of tissue which will need further examination to determine the diagnosis. This will usually lead to tissue samples for testing to determine if they are Malignant (Cancerous) or Benign (Non-cancerous). A false-negative on these tissue samples can be catastrophic and prevent timely care to stop the spread and save the patient’s life.

**The Data:**

* The following data was pulled from Kaggle: <https://www.kaggle.com/datasets/erdemtaha/cancer-data>

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

**Outcome:**

* I believe that using this data we will be creating a classification model and our target data will be the “diagnosis” column. Using this data we should be able to create a model that will classify the tissue sample sizes as Malignant or Benign.