**Samuel Slomowitz**

**A Dive into Machine Learning and Data Visualization on Breast Cancer Data**

Introduction

The incidence of breast cancer in women is the second highest behind skin cancer. The median age of diagnosis is 62. On average, 1 in 8 women will develop breast cancer in the USA. Death by breast cancer ranks second in women behind lung cancer. Further, 3.8 million patients overcome breast cancer as survivors.

Breast Cancer Dataset from Kaggle:

<https://www.kaggle.com/datasets/sarahvch/breast-cancer-wisconsin-prognostic-data-set>

Inspiration:

Chart, bar chart

Description automatically generated­­

A picture containing chart

Description automatically generated

Chart, treemap chart

Description automatically generated

Chart

Description automatically generated

Chart

Description automatically generated with low confidence

Colors:

Bar chart

Description automatically generated

Predictive Analytics:

I will be predicting prognosis of breast cancer, i.e. malignant or begin tumor in biopsies.

Basic design concepts for dashboard:

Timeline

Description automatically generated with medium confidence

Text

Description automatically generated with low confidence

A picture containing text, businesscard, screenshot

Description automatically generated

**References**

<https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html>

<https://medium.com/swlh/simple-machine-learning-model-on-breast-cancer-dataset-c1a013b594ad>