**Department of Computer Science and Engineering Advanced Data Visualization**

| UID | 2021300010 |
|---|---|
| Name | Kunal Bhatia |
| Class and Batch | Comps A Batch G |
| Experiment No. | 3 |
| Aim | Design Interactive Dashboard and storytelling using Power BI /Python Write observations for each chart. |

# Visual charts with Observation:

I have gathered a Breast Cancer dataset containing data on cancer stages, hormone levels, and demographic information about the women most affected by the disease.



**DASHBOARD**

- The first graph is a pie chart representing the distribution of regional node positives based on the marital status of women. The data reveals that married women have the highest count of breast cancer cases. This observation suggests the need for further analysis, particularly exploring the differences in hormone levels between married and single women.
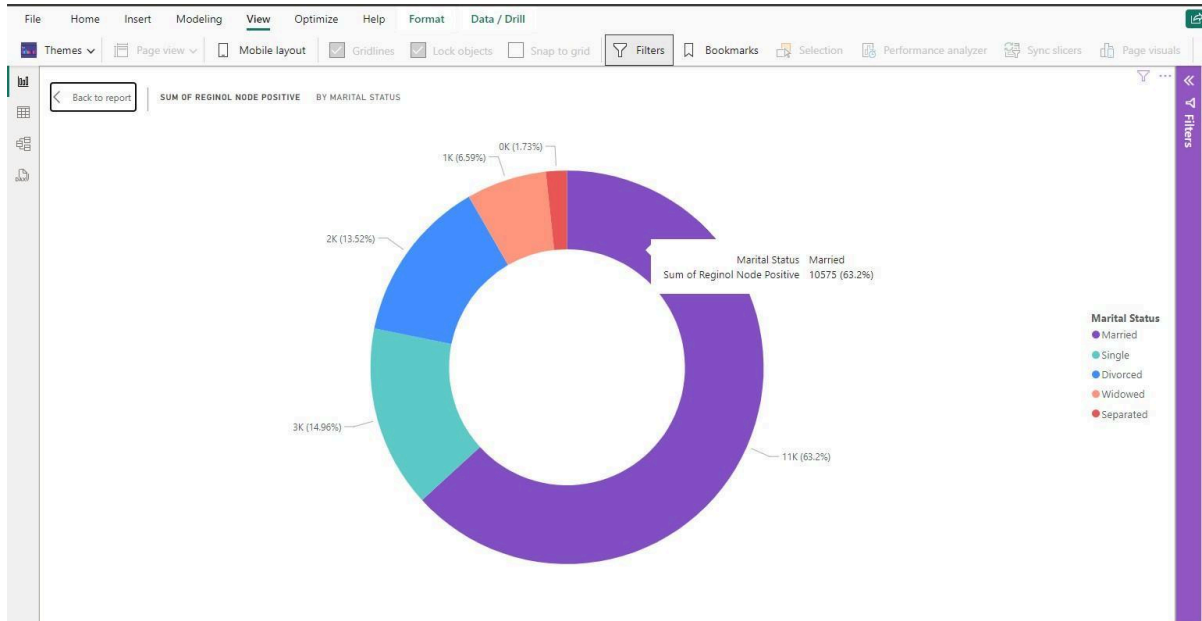


63.2% of the total women having high node levels were found to be married.

1. The following stacked bar graph is grouped based on the stages of cancer . They show if the estrogen levels are positive or negative and how they relate to the node levels.
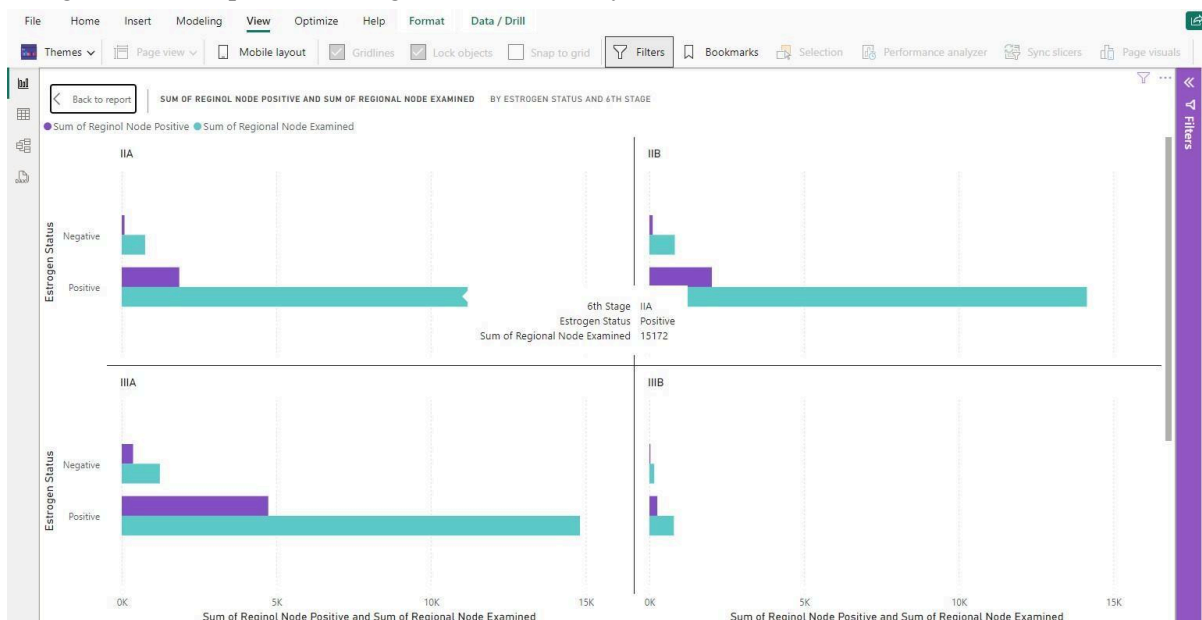


So most women have lower stage cancer. We can see that 15172 women were examined for the node levels as seen in green, and only 1680 of them had positive node levels suggesting cancer. We can see that women with higher estrogen are more prone to breast cancer.

2. This graph shows a race wise segregation of women and showcases the tumor size and hormone levels.
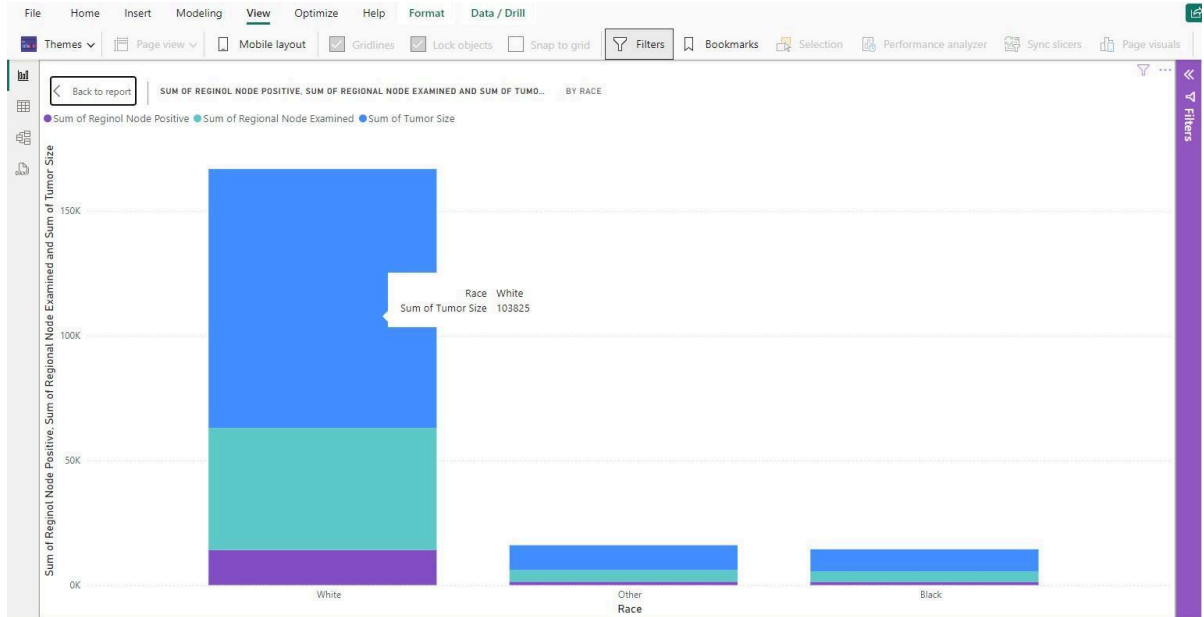
White people are the most prone to breast cancer and have recorded the highest tumor sizes. This can be studied further to find why.

3. This chart can be used to analyze how race and differentiation levels are associated with the sum of regional nodes examined, helping to identify trends or disparities in data related to breast cancer stages.
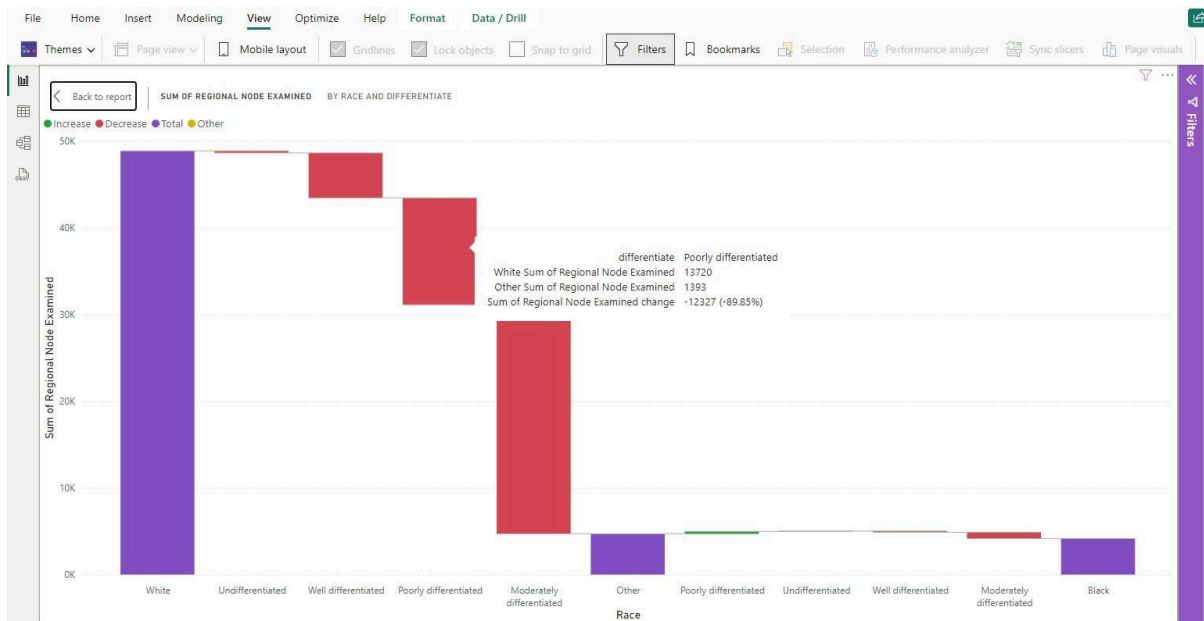


The largest bar, in purple, corresponds to the "White" category with a significant positive contribution. Differentiation categories show red bars, representing a negative contribution or a reduction.

4. This line graph shows the sum of survival months, sum of regional node positive, and sum of tumor size across different stages (IIA, IIB, IIIA, IIIC, IIIB).
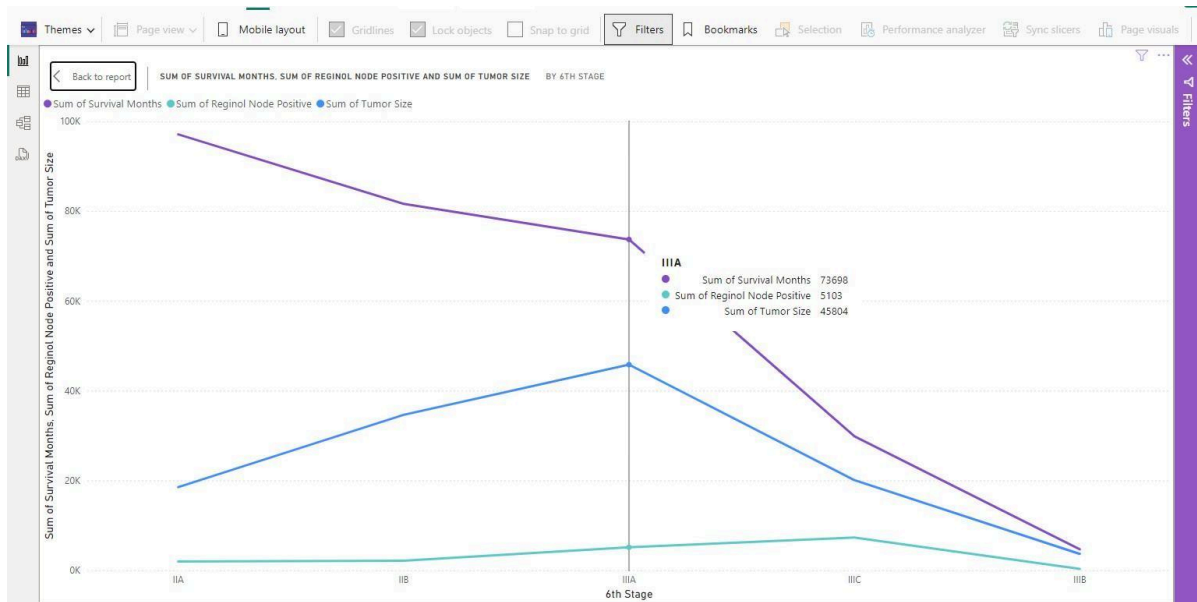
The observation is that survival months tend to decrease as the stages progress, with a noticeable drop at stage IIIB.

5. The table shows the count of people having positive and negative progesterone levels grouped by the stages of cancer.



| 6th Stage | Negative | Positive | Total |
|---|---|---|---|
| IIA | 837 | 17664 | 18501 |
| IIB | 2020 | 32597 | 34617 |
| IIIA | 3244 | 42560 | 45804 |
| IIIB | 718 | 2921 | 3639 |
| IIIC | 2643 | 17422 | 20065 |
| Total | 9462 | 113164 | 122626 |

So we can conclude that women having positive progesterone levels are more prone to breast cancer especially IIB stage and should take more preventive care.

6. The below funnel graph shows that out of all races, white women have examined themselves most for breast cancer check up and we can conclude that they have better access to healthcare.

Back to report | **SUM OF REGIONAL NODE EXAMINED** · BY RACE

100%

| | |
|---|---|
| White | 49K |
| Other | 5K |
| Black | 4K |

| | |
|---|---|
| Race | White |
| Sum of Regional Node Examined | 48900 |
| Percent of first | 100% |

8.5%