Aditya Singh

1. **Mutation Count and Fraction Genome Altered:** There is a negative correlation of approximately -0.23, suggesting a weak negative relationship. As one variable increases, the other tends to decrease, and vice versa.
2. **Mutation Count and Diagnosis Age:** There is a negative correlation of approximately -0.19, indicating a weak negative relationship. This suggests that, on average, as Mutation Count increases, Diagnosis Age tends to decrease, and vice versa.
3. **Mutation Count and MSI MANTIS Score:** There is a weak positive correlation of approximately 0.09 , suggesting a weak positive relationship. This implies that, on average, as Mutation Count increases, MSI MANTIS Score also tends to increase, and vice versa.
4. **Mutation Count and MSIsensor Score:** There is a weak positive correlation of approximately 0.09, similar to the correlation with MSI MANTIS Score. This indicates a weak positive relationship between Mutation Count and MSIsensor Score.
5. **Fraction Genome Altered and Diagnosis Age:** There is a positive correlation of approximately 0.26, indicating a weak positive relationship. This suggests that, on average, as Fraction Genome Altered increases, Diagnosis Age tends to increase, and vice versa.
6. **Fraction Genome Altered and MSI MANTIS Score:** There is a negative correlation of approximately -0.33, suggesting a moderate negative relationship. This implies that, on average, as Fraction Genome Altered increases, MSI MANTIS Score tends to decrease, and vice versa.
7. **Fraction Genome Altered and MSIsensor Score:** There is a negative correlation of approximately -0.35, similar to the correlation with MSI MANTIS Score. This indicates a moderate negative relationship between Fraction Genome Altered and MSIsensor Score.
8. **Diagnosis Age and MSI MANTIS Score:** There is a weak negative correlation of approximately -0.04, suggesting a very weak negative relationship. This implies that, on average, as Diagnosis Age increases, MSI MANTIS Score tends to decrease slightly, and vice versa.
9. **Diagnosis Age and MSIsensor Score:** There is a very weak negative correlation of approximately -0.03, similar to the correlation with MSI MANTIS Score. This indicates a very weak negative relationship between Diagnosis Age and MSIsensor Score.
10. **MSI MANTIS Score and MSIsensor Score:** There is a very strong positive correlation of approximately 0.95, indicating a very strong positive relationship. This suggests that MSI MANTIS Score and MSIsensor Score are highly positively correlated; as one increases, the other tends to increase, and vice versa.

The histogram shows that there are two peaks in the distribution of diagnosis ages. The first peak is around age 40, and the second peak is around age 60. This suggests that there are two age groups that are more likely to be diagnosed with this disease.

It is important to note that this is just a general overview of the distribution of diagnosis ages for this disease. The specific age at which someone is diagnosed can vary depending on a number of factors, such as genetics, lifestyle, and environmental exposures.

1. **Bottom and Top of Each Box:** These lines mark the 25th and 75th percentiles of the data, representing the interquartile range (IQR).
2. **Line in the Middle of Each Box (Median):** This line represents the median or the 50th percentile of the data.
3. **Whiskers:** The whiskers extend above and below the boxes and show the spread of the data, up to 1.5 times the interquartile range (IQR).
4. **Outliers:** Any data points beyond the whiskers are considered outliers and are plotted individually as small circles.

Here are some of the key things we can learn from the box plot:

Patients who are still living tend to be diagnosed at a younger age than patients who have died. The median diagnosis age for living patients is around 50, while the median diagnosis age for deceased patients is around 60.

There is a wider range of diagnosis ages for patients who have died. The IQR for the deceased patients is about 20 years, while the IQR for the living patients is about 15 years. This means that there is more variability in the ages at which patients who die are diagnosed with the disease.

There are a few outliers in the data. There are two data points for living patients that are diagnosed at a much older age than the rest of the group. There is also one data point for a deceased patient who was diagnosed at a younger age than the rest of the group.

It is important to note that this is just a general overview of the relationship between diagnosis age and overall survival status for patients with this disease. The specific age at which someone is diagnosed and their overall survival prognosis can vary depending on a number of factors, such as the type of disease, the stage of the disease at diagnosis, and the patient's individual health.