Brain Tumor Stage-Based Recurrence Data Analysis

Final Project Report
Neuroscience Data Analysis in Python
January 2026

Submitted by:
Miriam Chavay, Ayal Adler, Dana Sela, Tamar Krakover

1. Introduction

This report presents a computational analysis of clinical data from brain tumor patients. The primary objective was to investigate the relationships between various clinical and demographic factors- such as tumor location, type, and patient age- and their impact on prognosis and recurrence.

The study addresses four key research questions:

1. Predicting the site of tumor recurrence based on the initial tumor location.
2. Examining the correlation between patient age/gender and survival time.
3. Assessing the efficacy of different treatment protocols.
4. Determining whether the tumor type can reliably predict the tumor grade.

By applying statistical methods and data visualization tools in Python, common medical assumptions were validated against quantitative evidence.

2. Results

The following sections outline the findings for each research question.

2.1 Interactive Recurrence Prediction (GUI)

Research Question: *Can we predict the chances of a tumor recurrence and the recurrence site based on its initial location?*

To address this question effectively, an interactive Graphical User Interface (GUI) was developed. Instead of presenting a static chart, this tool allows the user to input a specific initial tumor location. The system then analyzes historical data to calculate and display the conditional probabilities for recurrence in various brain regions.

Example Usage:

The following text demonstrates the system's output logic. When 'Frontal lobe' is selected as the primary site, the algorithm retrieves relevant cases and computes the recurrence distribution:

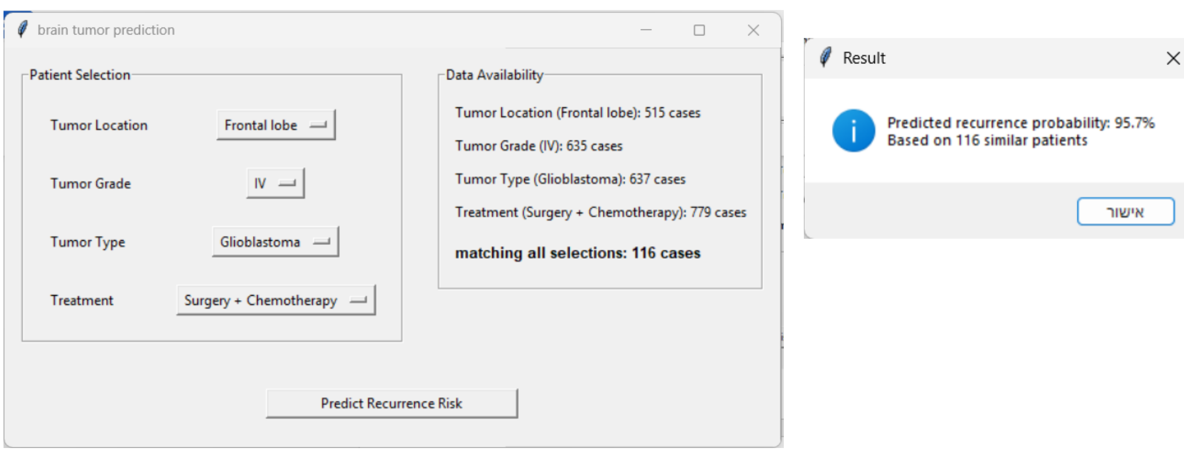

Figure: A simulated output example from the interactive GUI, demonstrating the conditional probability calculation for a specific tumor

2.2 Correlation: Age vs. Survival

Research Question: *Is there a correlation between the survival time and the age of the patient (at the time of diagnosis)?*

It is often hypothesized that older age at diagnosis correlates with shorter survival times. This hypothesis was analyzed by plotting patient age against survival duration (in months). As illustrated in the scatter plot below, the data points are widely dispersed, indicating that age alone is not a strong linear predictor of survival in this dataset.

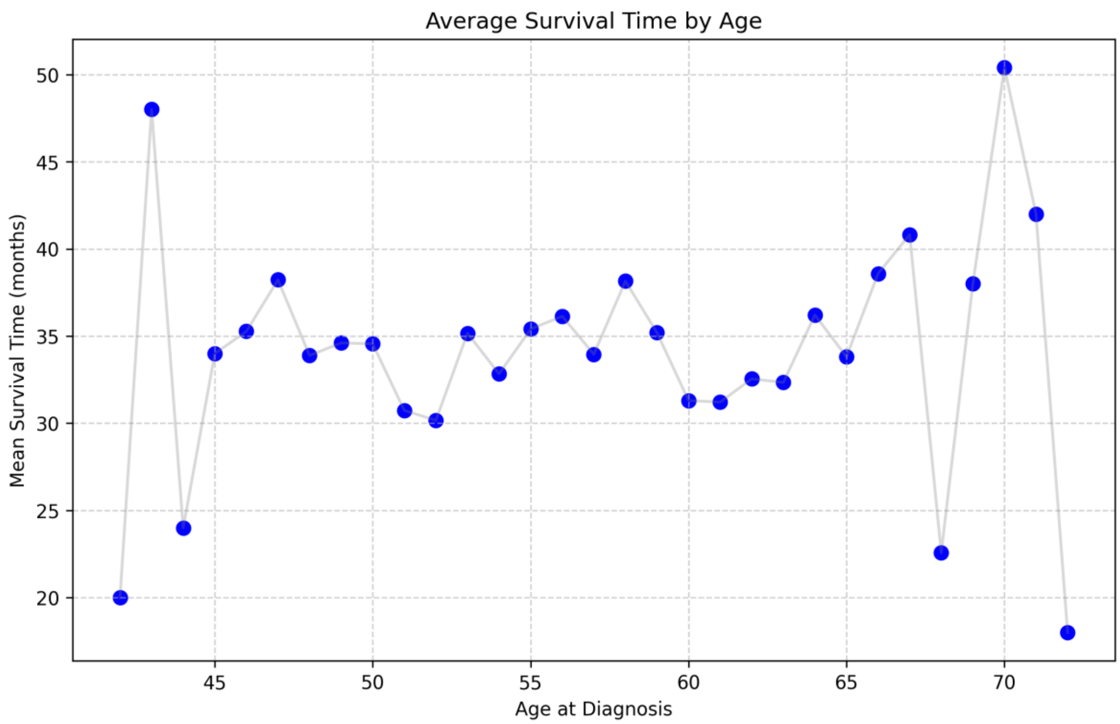

Figure: Scatter Plot: Patient Age vs. Survival Time

2.3 Gender Impact Analysis

Research Question: *Is there a connection between the survival time and the patient's gender?*

The influence of biological sex on survival outcomes was investigated. Using boxplot analysis to compare the distributions for male and female patients, it was observed that the median survival times and overall spread are largely similar between the groups, suggesting no significant gender-based disparity.

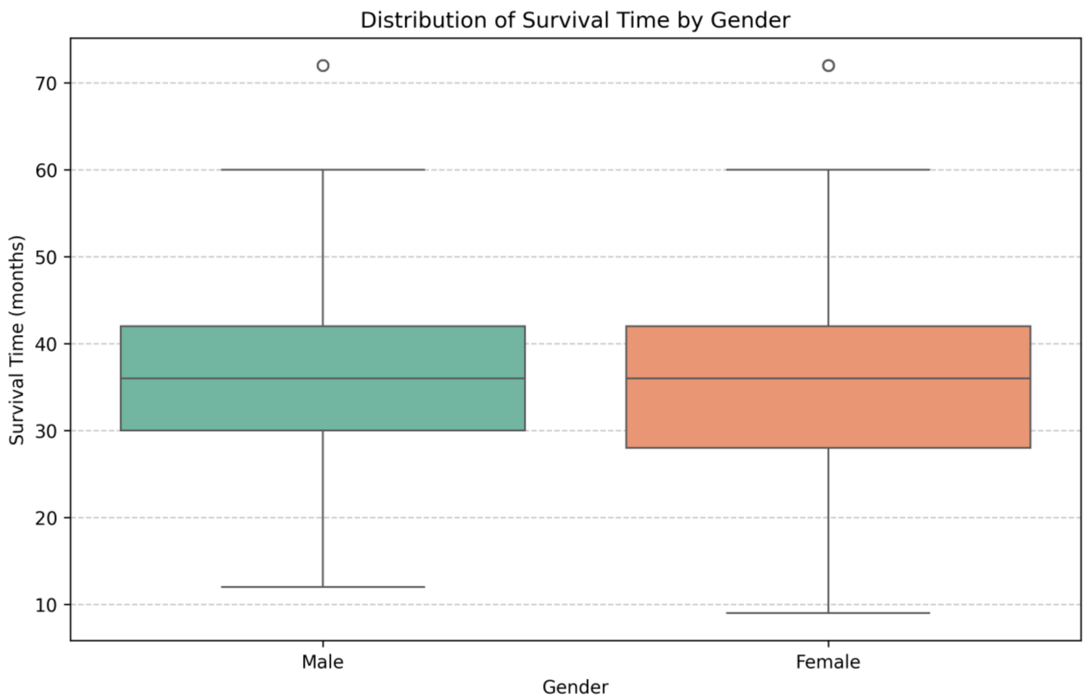

Figure: Boxplot: Survival Time Distribution by Gender

2.4 Treatment Efficacy

Research Question: *Can we predict the treatment outcome based on the chosen treatment?*

The analysis evaluated the success rates associated with different treatment protocols (e.g., Surgery, Radiation, Chemotherapy). The figure below provides a representative example of these outcome distributions. This visualization facilitates the identification of treatments that are statistically more likely to result in a 'Complete Response' versus a 'Progressive Disease'.

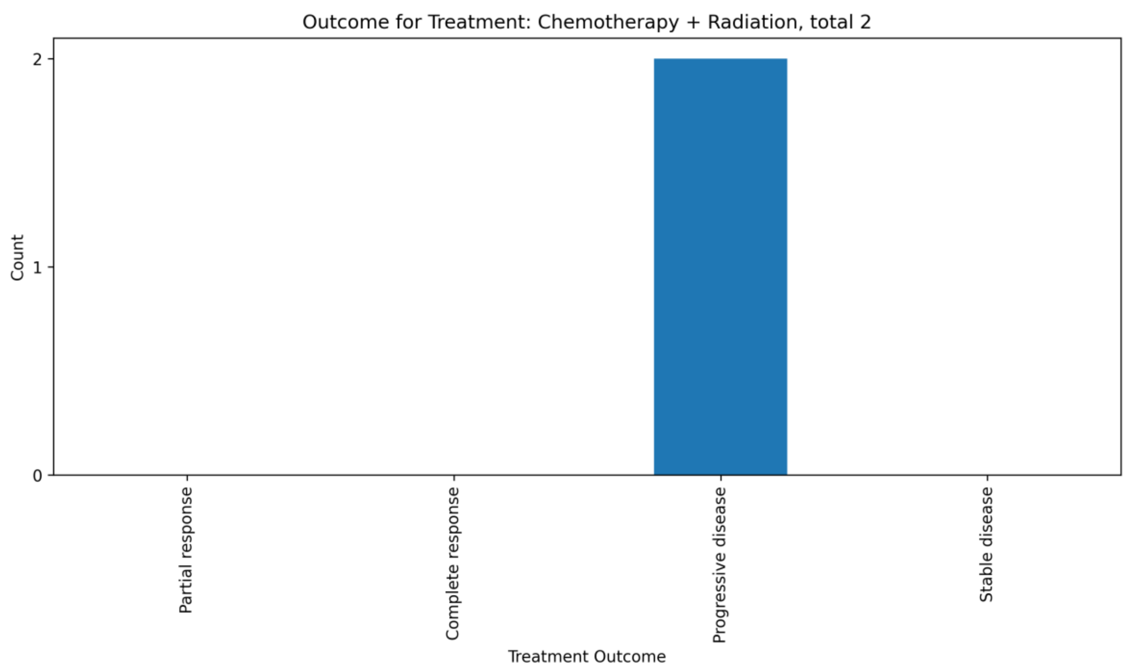

Figure: Representative Distribution of Treatment Outcomes

2.5 Tumor Type and Grade Association

Research Question: *Can tumor type predict the tumor grade at initial diagnosis?*

The relationship between the pathological tumor type and its severity grade was examined. Statistical analysis (Cramer's V) and the distribution plot below reveal a very strong association. Specifically, certain tumor types (such as Glioblastoma) are almost exclusively associated with higher grades, confirming a distinct biological pattern.

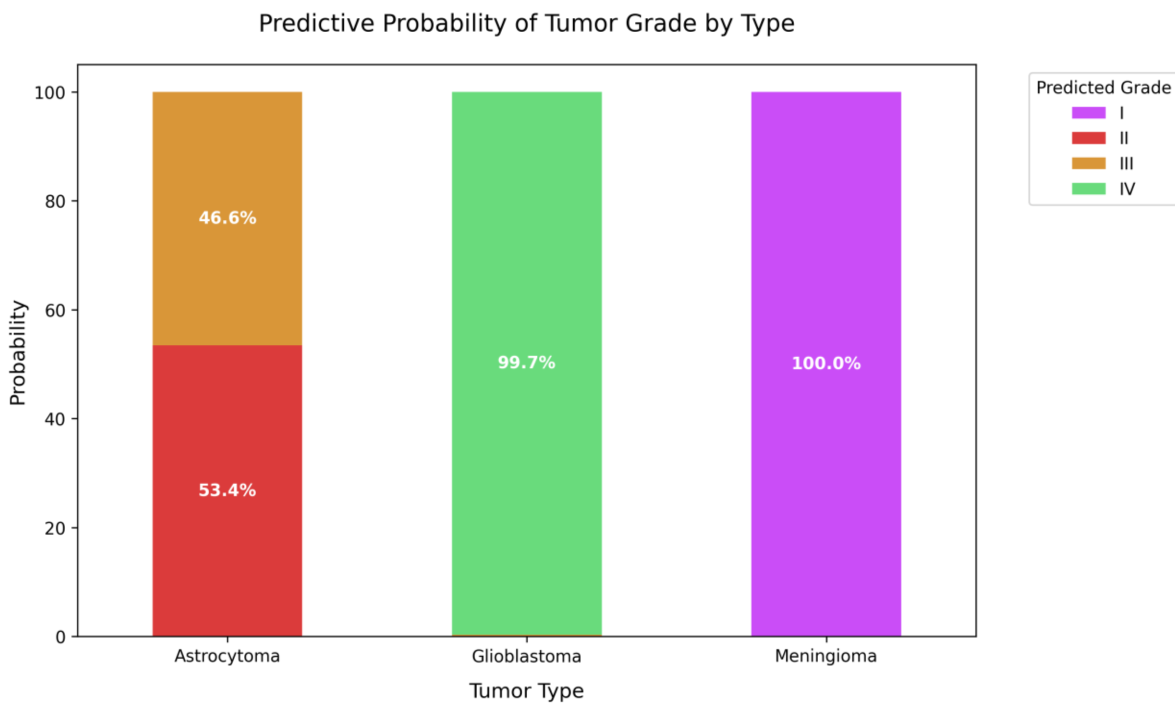

Figure: Distribution of Tumor Grade by Type

3. Discussion

The analysis of the dataset yields several important clinical insights:

1. *Tumor Characteristics Over Demographics: The strongest statistical signal in this study was the link between Tumor Type and Grade. In contrast, demographic factors like Age and Gender showed weak or negligible correlations with survival. This suggests that for brain tumors, the specific pathology is a far more critical determinant of prognosis than the patient's demographic profile.*
2. *Treatment Complexity: The variability in treatment outcomes highlights that there is no single 'best' protocol for all cases. The probabilistic data derived from this*

study can help set realistic expectations regarding the efficacy of surgery and radiation.

In conclusion, this project demonstrates the value of data-driven analysis in validating medical hypotheses. While some assumptions (like the age-survival link) were not supported by this specific dataset, others (like the type-grade link) were strongly confirmed.