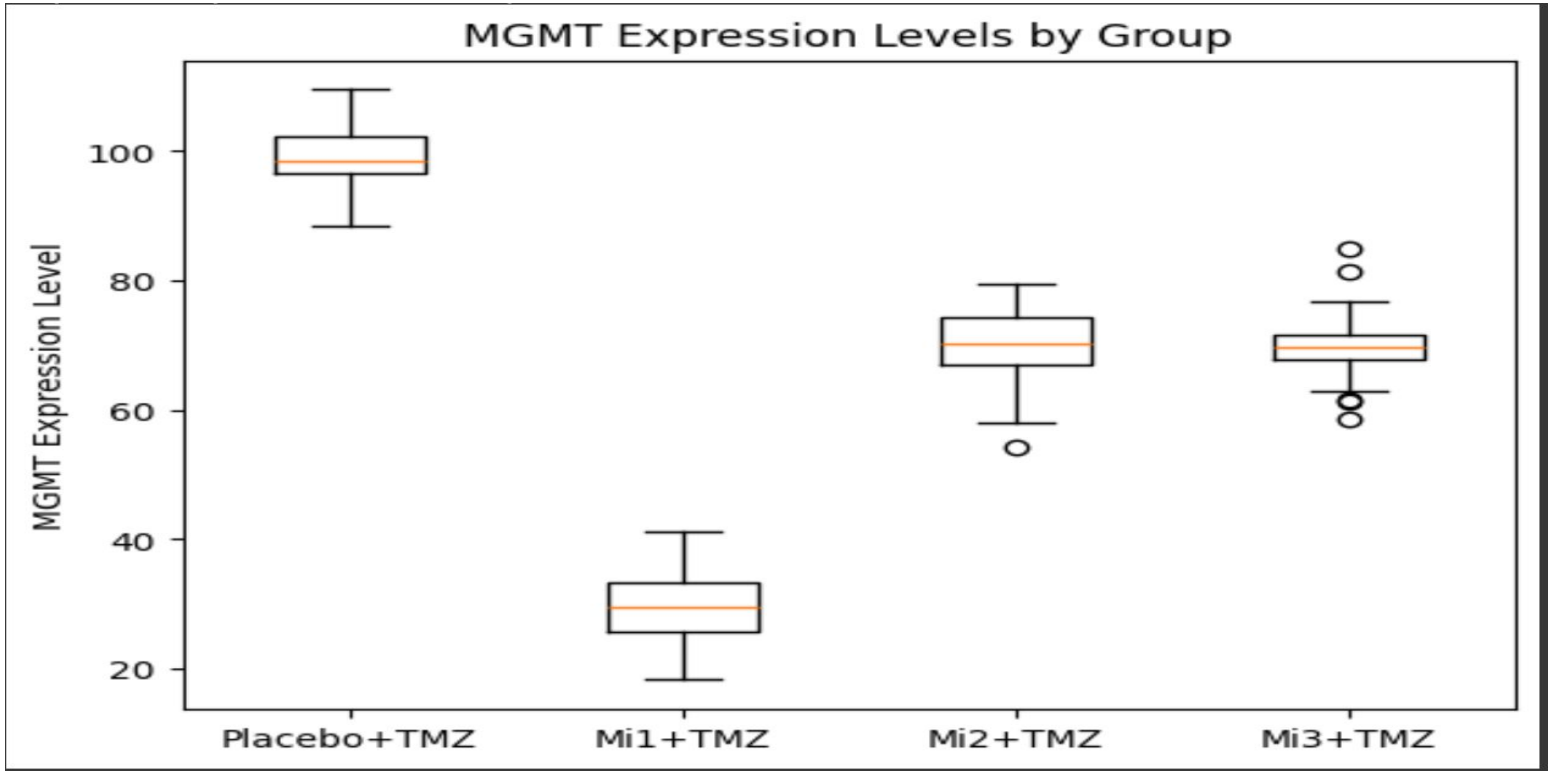
Elias Firisa
20220773
BCS410
HW#1 report

Problem 1: reproduce with problem1_sol.ipynb

Visual comparison of the MGMT expression levels using box plots



Statistical hypothesis

- Null hypothesis(H0): All four groups (Placebo+TMZ, Mi1+TMZ, Mi2+TMZ, Mi3+TMZ) have the **same** mean MGMT expression level.

- Alternative hypothesis(H1): At least one group has different mean MGMT expression level

Test1: What test did I perform?

1-way ANOVA: `F_onewayResult(statistic=782.8699434287794, pvalue=8.787050540043206e-77)`

- The p-value for the given F-statistic is very small, way below the threshold=0.05, I rejected the null hypothesis and concluded that there is a significant difference in mean MGMT expression level among the 4 groups

Tukey's HSD test

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Mi1+TMZ	Mi2+TMZ	40.8043	0.0	37.0339	44.5747	True
Mi1+TMZ	Mi3+TMZ	40.6055	0.0	36.8351	44.3758	True
Mi1+TMZ	Placebo+TMZ	69.5981	0.0	65.8277	73.3685	True
Mi2+TMZ	Mi3+TMZ	-0.1988	0.9991	-3.9692	3.5715	False
Mi2+TMZ	Placebo+TMZ	28.7938	0.0	25.0234	32.5642	True
Mi3+TMZ	Placebo+TMZ	28.9926	0.0	25.2223	32.763	True

Test2: What test did I perform?

We did ANOVA but it tells us just the groups have different mean expression level but not which groups are different. So, I run Tukey's HSD test and get the results given in the previous page

From Tukey's HSD table result we can say:

Mi1+TMZ has the **lowest** MGMT expression (significantly lower than all others).

Placebo+TMZ has the **highest** MGMT expression.

Mi2+TMZ and **Mi3+TMZ** fall in between and do **not** differ significantly from each other.

This aligns with the boxplot: Mi1+TMZ is visibly lowest, Placebo+TMZ highest, and Mi2+TMZ and Mi3+TMZ cluster together

Problem 2: reproduce with problem2_sol.ipynb

	sum_sq	df	F	PR(>F)
C(age_months)	5.366618e+06	2.0	276.759092	1.022155e-61
C(drug_candidate)	1.694063e+07	3.0	582.424396	1.442183e-106
C(age_months):C(drug_candidate)	4.440508e+06	6.0	76.333054	8.723920e-52
Residual	2.210567e+06	228.0	NaN	NaN

Statistical Hypothesis

H0: tumor volume doesn't vary among age groups

H1: tumor volume varies among age groups

H0: tumor volume doesn't vary among different drug treatments

H1: tumor volume varies among different drug treatments

H0: The effect of drug doesn't depend on age group (or vice-versa)

H1: The effect of drug depends on age group

Test1: What test did I perform?

Two-way ANOVA

	sum_sq	df	F	PR(>F)
C(age_months)	5.366618e+06	2.0	276.759092	1.022155e-61
C(drug_candidate)	1.694063e+07	3.0	582.424396	1.442183e-106
C(age_months):C(drug_candidate)	4.440508e+06	6.0	76.333054	8.723920e-52
Residual	2.210567e+06	228.0	NaN	NaN

Interpretation of the results from 2-way ANOVA test

- From the tabular values in the previous page, the p-value for all 3 rows are very small → We reject all the three null hypothesis.
- Both main effects- age and drug- matter for tumor size
- The interaction between age and drug is statistically significant → We can't just say this drug is the best or the worst since its effect is a function of age or we can't say younger mice have lower tumor size since this itself has influence from drug combination
- Therefore, we need to do Tukey's HSD on (age,drug) combination to see further

Test2: What test did I perform?

Tukey's HSD test: to find the group- (age,drug)- that has lowest mean

```
means = pd_data.groupby('group_combo')['tumor_volume_mm2'].mean().sort_values()
print(means)
```

```
group_combo
0~6_Mi1_TMZ            201.062240
18~24_Mi1_TMZ          341.245601
9~15_Mi1_TMZ           402.653095
18~24_Mi3_TMZ          452.476278
18~24_Mi2_TMZ          469.838157
9~15_Mi2_TMZ           594.340697
9~15_Mi3_TMZ           603.790906
18~24_placebo+TMZ      735.861963
0~6_Mi3_TMZ            896.232413
0~6_Mi2_TMZ            897.998222
9~15_placebo+TMZ       996.419203
0~6_placebo+TMZ       1461.457824
Name: tumor_volume_mm2, dtype: float64
```

Interpretation of the results from Tukey's HSD test

- From the previous page's output we can see each average tumor volume for each (age,drug) combinations in ascending order.
- That tells us 0–6_Mi1_TMZ (i.e., mice aged 0–6 months treated with Mi1+TMZ) had the smallest average tumor size among all groups
- Results from (prob2_sol.ipynb) shows that this same group has statistically significantly smaller tumor volume than most (or all) other combinations

Problem 3: reproduce with problem3_sol.ipynb

My approach:

Separate the dataset into two groups based on the “treatment” column:

- Group A (TMZ-untreated, treatment=0)
- Group B (TMZ-treated alone, treatment=1)

Categorical Variables (Sex, MGMT Methylation, IDH Mutation)

Test: Chi-square test of independence.

- Null Hypothesis: The proportion of each category (e.g., male/female) is the same in both groups.
- If p-value < 0.05, that indicates a significant difference in distribution (i.e., imbalance)

Results from Chi-Square Test (for categorical data)

	Group 0	Group 1
Female	139	165
Male	162	152

Chi-square statistic: 1.9009175206818725
p-value: 0.16797565523431188

	Group 0	Group 1
WT	179	182
Methylated	122	135

Chi-square statistic: 0.1905208357777805
p-value: 0.6624834835489024

	Group 0	Group 1
Mutant	181	136
WT	136	181

Chi-square statistic: 12.214511041009464
p-value: 0.0004741923702103594 5

- We've found **no evidence** of imbalance in MGMT methylation or gender (both $p > 0.05$).
- However, **IDH mutation does** appear to differ significantly between the two groups ($p < 0.05$)
- We might include this IDH status as a confounding variable (e.g covariate in Cox proportional model)

Determining data distribution for age (continuous) variable

Null hypothesis: The sample come from a normally distributed population (age is normally distributed in both groups)

Test: Shapiro-Wilk test

For untreated group: → `stat=0.994, p=0.245`

For treated group → `stat=0.994, p=0.294`

Given that the p-values (≈0.245 and 0.294) are both **greater than 0.05**, we fail to reject the null hypothesis of normality→age in both groups is not significantly deviating from normal.

Now let's test for variance equivalence | data is normal

Null Hypothesis: The variance in both groups are equal

Test: Now we've established normality, we'll perform Levene's test to check if the variance of age in the two groups are equal

Result: `stat=0.196, p=0.659`

- ➔ Interpretation: 0.659>>0.05 → There is no significant difference in the variance of age between the two groups.
- ➔ We failed to reject null hypothesis→ we'll assume equal variance
- ➔ Now we can proceed with two-samples t-test

Two-sample t-test

Null Hypothesis (H0):The mean age is the same in both groups, i.e., μTMZ-untreated=μTMZ-treated-alone
Alternative Hypothesis (H1):The mean age is different between the two groups, i.e., μTMZ-untreated≠μTMZ-treated-alone

Test: two-samples t-test

Result:

```
t_stat=-0.044, p_val=0.965
```

Interpretation:

- With p=0.965>>0.05 we fail to reject null hypothesis
- The groups have similar mean ages, indicating no age imbalance between them

Summary of findings for problem #3

The only imbalance between TMZ-untreated vs. TMZ-treated-alone groups is in **IDH mutation** status.

All other variables (age, sex, MGMT methylation) are balanced

Problem 4: reproduce with Problem4_sol.ipynb

Statistical Hypothesis:

- **Null Hypothesis (H0)**: There is **no difference** in overall survival (OS) between GBM patients who are **untreated with TMZ** and those **treated with TMZ alone**.
- **Alternative Hypothesis (H1)**: There **is** a difference in overall survival between the two groups (i.e., TMZ treatment improves OS)

Method

Kaplan-Meier Analysis

→ I fitted separate Kaplan-Meier survival curves for each group, plotting survival probability against time.

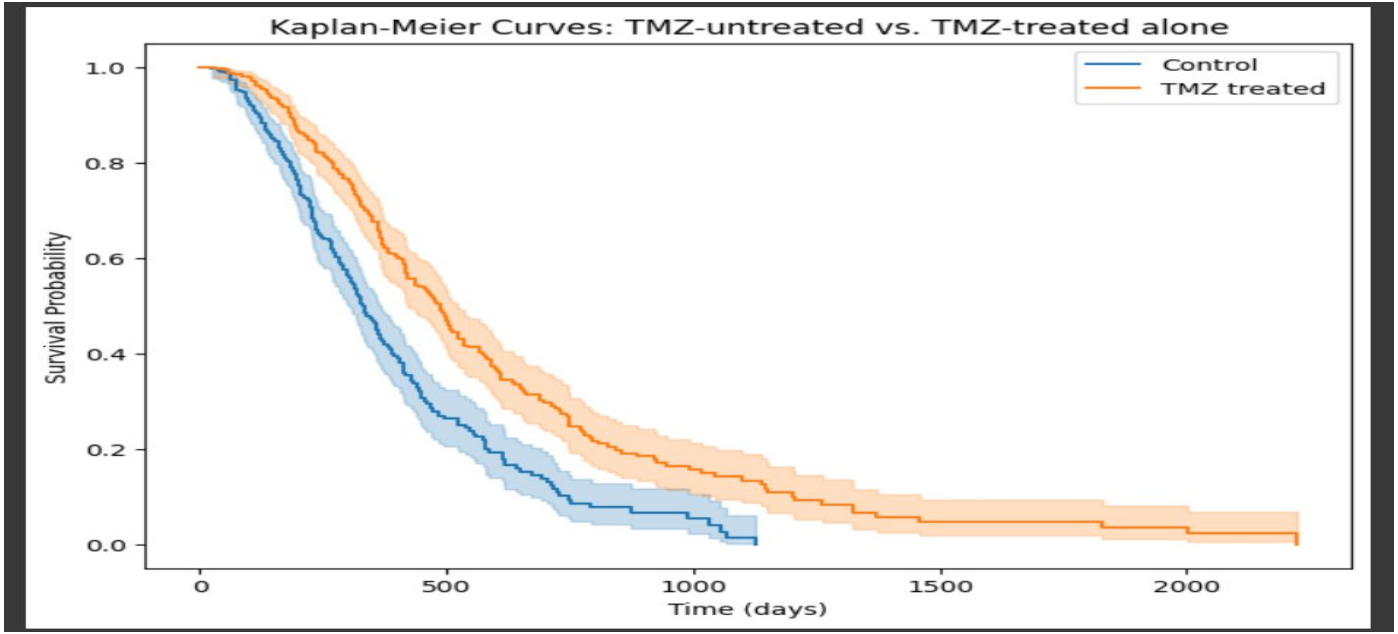
→ This step provides a visual comparison of how the survival function differs over time between the two groups.

Log-Rank Test

→ To determine whether the difference in survival curves is statistically significant, I performed a log-rank test.

→ Null Hypothesis for the log-rank test: Both groups share the same underlying survival distribution

Problem 4: results



Log-rank test result

```
P_value: 9.274697389915403e-10
```

Problem 4: Interpretations

Kaplan-Meier Curves

- The survival curve for the **TMZ-treated** group lies **above** that of the **TMZ-untreated** group, indicating a higher probability of survival at most time points.

Log-Rank p-value

- The p-value is **far below 0.05**, providing **strong evidence** against the null hypothesis.

Given the visually higher survival curve for the **TMZ-treated** group and the **extremely small** p-value from the log-rank test, we conclude:

TMZ administration is associated with a significantly improved overall survival compared to no TMZ treatment

Problem 5:reproduce with Problem5_sol.ipynb

Hypothesis

- **Null Hypothesis (H_0):** After adjusting for age, sex, MGMT promoter methylation status, and IDH mutation, there is **no difference** in overall survival (OS) between GBM patients who are **untreated with TMZ** (treatment=0) and those **treated with TMZ alone** (treatment=1).

- **Alternative Hypothesis (H_1):** Even after adjusting for these covariates, **TMZ-treated alone** is associated with a **significant difference** in OS compared to the untreated group (specifically, we expect a **lower hazard** if TMZ is beneficial)

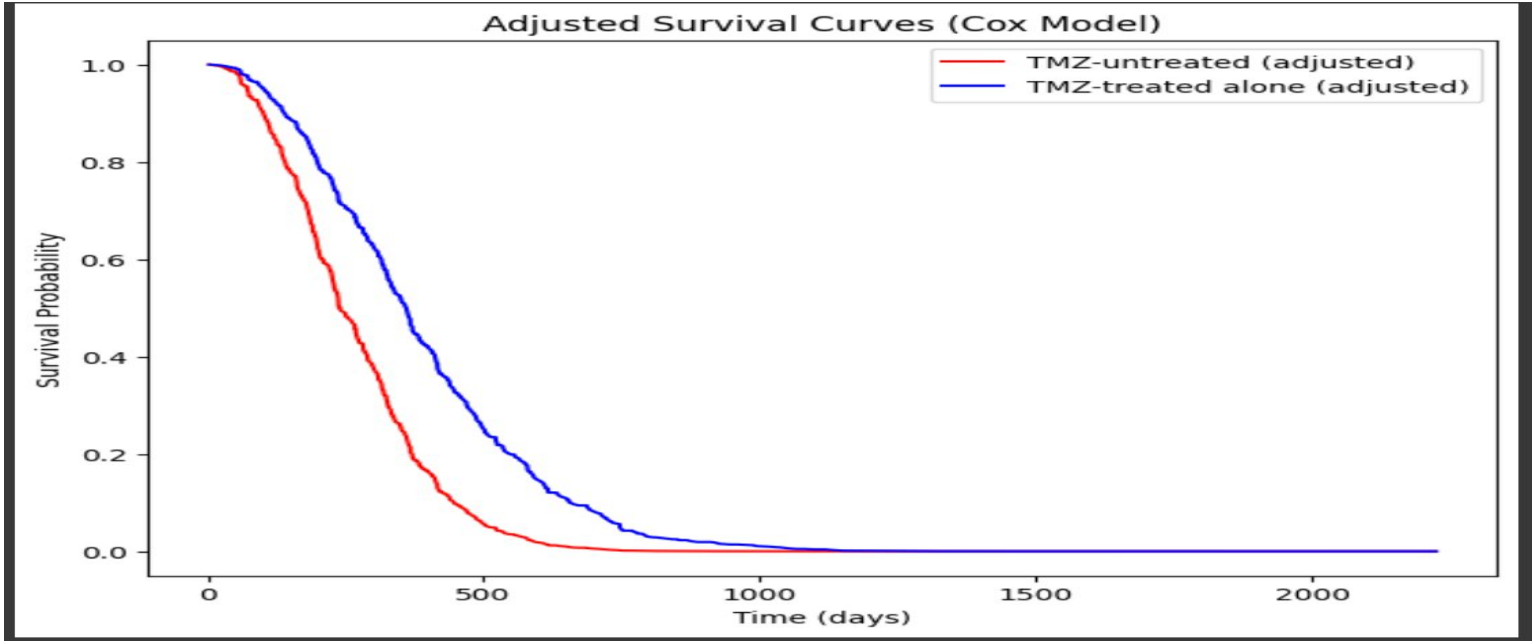
Method (Test):

- Statistical Model | Cox Proportional Hazards model (semiparametric survival analysis)

Problem 5: results

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
age (years)	0.03	1.03	0.01	0.02	0.04	1.02	1.04	0.00	5.68	<0.005	26.17
sex (0=Female, 1=Male)	0.08	1.09	0.10	-0.11	0.27	0.90	1.31	0.00	0.83	0.40	1.31
MGMT_promotor_methylated (0=WT, 1=Methylated)	-0.92	0.40	0.11	-1.13	-0.70	0.32	0.49	0.00	-8.44	<0.005	54.83
IDH_mutant (0=WT, 1=Mutant)	-0.43	0.65	0.10	-0.63	-0.23	0.53	0.79	0.00	-4.27	<0.005	15.64
treatment (0=TMZ-untreated, 1=TMZ-treated alone, 2=Mi1+TMZ)	-0.73	0.48	0.10	-0.93	-0.53	0.39	0.59	0.00	-7.18	<0.005	40.35

Problem 5: results



Reference categories: sex= male, MGMT=unmethylated,IDH_mutant=WT, age= average age

Problem 5: result interpretation

Interpreting the table

Effect of TMZ:

- A hazard ratio (HR) **significantly < 1** for treatment=1 implies that TMZ-treated patients have a **reduced hazard of death** compared to untreated patients, indicating **improved overall survival**.
- The **low p-value** for treatment confirms this difference is statistically significant even after adjusting for age, sex, MGMT, and IDH.

Other Covariates:

- **Age**: HR =1.03 > 1 → older age increases the hazard of death (supported by low p-value)
- **MGMT Methylation**: If HR=0.40 < 1, methylated MGMT improves survival (supported by low p-value)
- **IDH Mutation**: HR =0.65 < 1, IDH-mutant improves survival (supported by low p-value)
- **Sex**: HR=1.09 suggests that might have ~9% higher hazard than females, however confidence intervals (**lower 95% = 0.90**, **upper 95% = 1.31)** include HR=1 and p_value=0.4 >>0.5 also supports the null hypothesis
  - Difference in hazard for males vs. females is **not** statistically significant. Hence, I do **not** reject the null hypothesis that males and females have the same hazard after controlling for other factors.

Problem 5: result interpretation

Interpreting the plot

Survival Curve Positions:

- The blue curve (TMZ-treated alone) lies above the red curve (TMZ-untreated) for the entire follow-up period.
- This indicates that, at any given time, the adjusted probability of surviving is higher for patients receiving TMZ than for those not receiving it.

Interpretation:

- After controlling for age, sex, MGMT methylation, and IDH status, the model predicts that patients in the TMZ-treated group consistently have better survival outcomes compared to those in the untreated group

Problem 6: reproduce with Problem6_sol.ipynb

Statistical Hypothesis:

Null Hypothesis (H_0): Among GBM patients with **unmethylated MGMT**, there is **no difference** in overall survival (OS) between those treated with **TMZ alone** and those treated with **Mi1+TMZ.**

Alternative Hypothesis (H_1): Among GBM patients with **unmethylated MGMT**, **Mi1+TMZ** significantly improves OS compared to **TMZ alone**

Test performed:

Subsetting the Data:

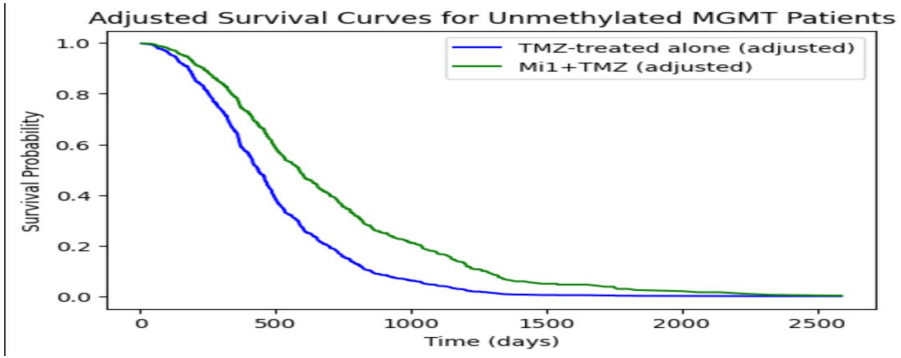
- I restricted our dataset to patients with **unmethylated MGMT** (MGMT_promotor_methylated=0).
- Within this subset, I selected two treatment groups:
  - **TMZ-treated alone** (treatment=1)
  - **Mi1+TMZ** (treatment=2)

Cox Proportional Hazards Model:

- I fit a Cox model using survival_time (days) and event (1=death, 0=censored) as our outcome.
- The main predictor of interest is **treatment** (TMZ alone vs. Mi1+TMZ).
- I also included **age**, **sex**, and **IDH_mutant** as additional covariates to adjust for confounding

Problem 6: results

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
age (years)	0.03	1.03	0.00	0.02	0.04	1.02	1.04	0.00	5.65	<0.005	25.89
sex (0=Female, 1=Male)	0.02	1.02	0.09	-0.16	0.20	0.85	1.22	0.00	0.20	0.84	0.25
IDH_mutant (0=WT, 1=Mutant)	-0.56	0.57	0.10	-0.75	-0.37	0.47	0.69	0.00	-5.75	<0.005	26.70
treatment (0=TMZ-untreated, 1=TMZ-treated alone, 2=Mi1+TMZ)	-0.58	0.56	0.09	-0.77	-0.40	0.46	0.67	0.00	-6.23	<0.005	31.00



Problem 6: Interpretations

Effect of Mi1+TMZ:

- Because the hazard ratio for treatment is significantly **< 1**, **Mi1+TMZ** is associated with **reduced hazard** of death compared to TMZ alone in patients with **unmethylated MGMT**.
- The p-value < 0.05 confirms a **statistically significant** survival benefit for Mi1+TMZ

Hence, in GBM patients with unmethylated MGMT, the **Mi1+TMZ** combination shows a **significant survival advantage** over TMZ alone, based on the **Cox proportional hazards model** adjusted for age, sex, and IDH status