

Sample clinical data for use in MSc project.

The data is analysis\_data.csv and it contains 27 attributes and 415 cases. The following columns are in the data:

Name	Type	Comments
id	identifier	
age	interval	
Gender	binary	
Performance.Status	nominal	
Site.in.mouth	nominal	
Lymphoscintigraphy.type	nominal	
Path.depth	interval	
T.Size	nominal	
Peri.neural.spread	binary	
Differentiation	nominal	
No.of.positive.nodes	interval	
No.of.negative.nodes	interval	
Sentinel.node.status	binary	
Recurrence.status	binary	
Type.of.recurrence	nominal	
Treatment.of.recurrence	nominal	
Metastasis.Type	nominal	
Current.status	nominal	
os	time to event	Overall Survival
os.censor	censoring indicator	
dfs	time to event	Disease Free Survival
dfs.censor	censoring indicator	
dss	time to event	Disease Specific Survival
dss.censor	censoring indicator	
false.negative	binary	
rec.3.yr	nominal	
fu.3yr	binary	

Potential Research Questions:

Is there a difference in survival (OS) by gender?

Is there a difference in survival (DFS) by sentinel node status?

These are based on the objective “time to event”.

Additional research objectives:

This is another type of objective for the SKB. It is relevant for cases where the target variable is of type interval. A research question of this type would be:

Is there a difference in ages (Interval) by Gender (nominal or binary)?