Graphical user interface, table, Excel

Description automatically generated

Graphical user interface, application, table, Excel

Description automatically generated

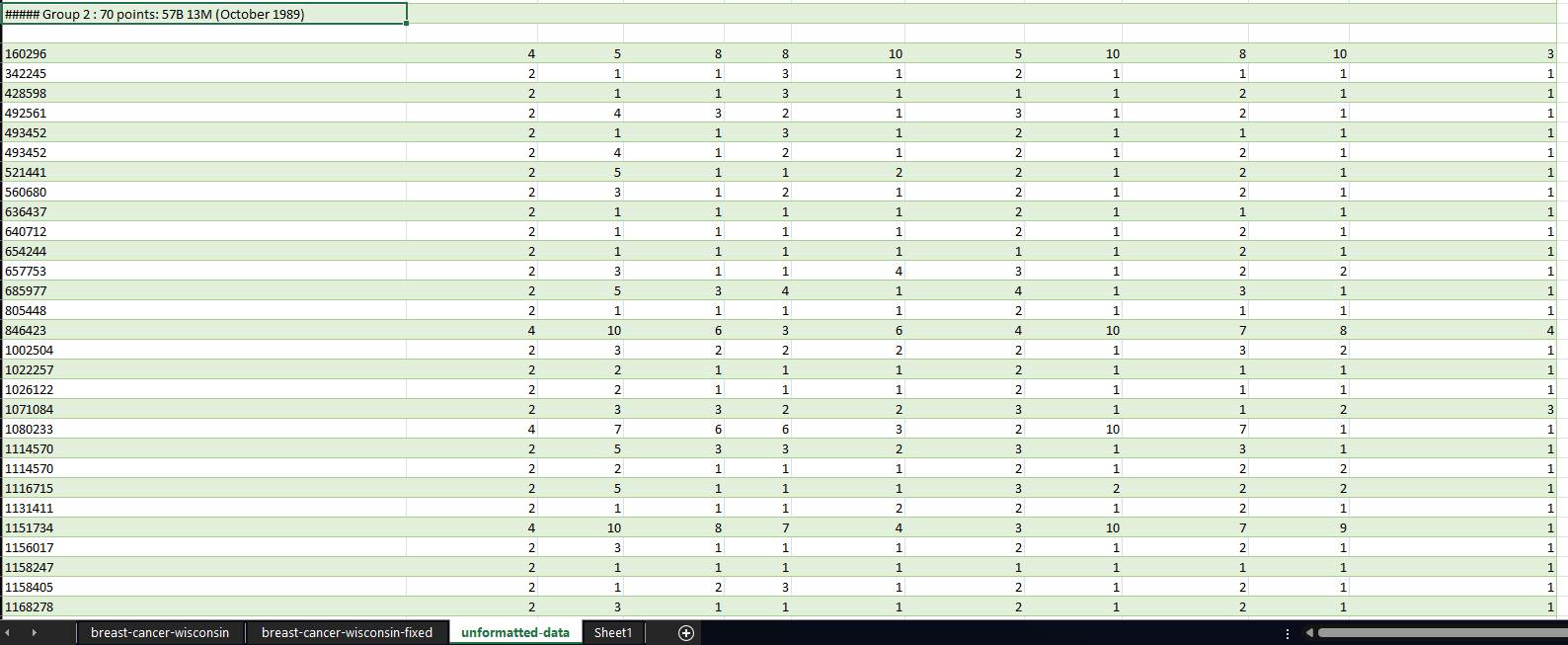
Added headers and cleaned up the columns

A picture containing text, building, window blind

Description automatically generated

Added a rule to check if there is data that doesn’t have numbers.

Removed any row with a question mark. The data went from 699 to 683.

Inside the unformatted Data: 

Here’s text that can get in the way. Going to remove.

Graphical user interface, application, table, Excel

Description automatically generated

Gave all their groups their sheet and specified the data from the unformatted data.

A screen shot of a computer

Description automatically generated with low confidence

This is the wdbc data and we need to clean this up before we can process it.

Graphical user interface, application, table, Excel

Description automatically generated

This is the cleaned up wdbc data. There is the wpbc data however I do now know the parameters and the value of the columns and it doesn’t say anywhere so I think we have to omit using that dataset.

Let me know if you guys want to add/take away from the data pre-processing. Thank you.