I think the dependent variable would be either the RSDI or the CT (or endoscopy) depending on what particular questions we are asking.  If it’s a patient centered questions RSDI makes more sense because that is their experience.  If it is a disease severity (CT or endoscopy) are more objective endpoints but may not correspond with the patients perception of their disease.

Variables:

* (subjective) survey score (RSDI)
* (subjective) number of allergies
* (objective) endoscopy score – right, left, total
* (objective) CT score

Explore relationships among variables:

* PCA:
  + <https://www.youtube.com/watch?v=xKl4LJAXnEA&ab_channel=HefinRhys>
* Clustering: hierarchical clustering (heatmap) 🡪 all variables are X
* Multi-regression analysis + prep 🡪 one of them is Y
  + Relationship between each X and Y: explanation? (scatterplot)
  + Relationship between each pair of X: collinearity? (scatterplot)
  + Correlation table: Pearson correlation with p-value
  + Ridge, Lasso & Elastic Net Regression <https://www.youtube.com/watch?v=_3xMSbIde2I&ab_channel=Dr.BharatendraRai>

Text

Description automatically generatedGraphical user interface, text

Description automatically generated

Other Notes:

* PCA: only useful for visual

Collinearity diagnostics