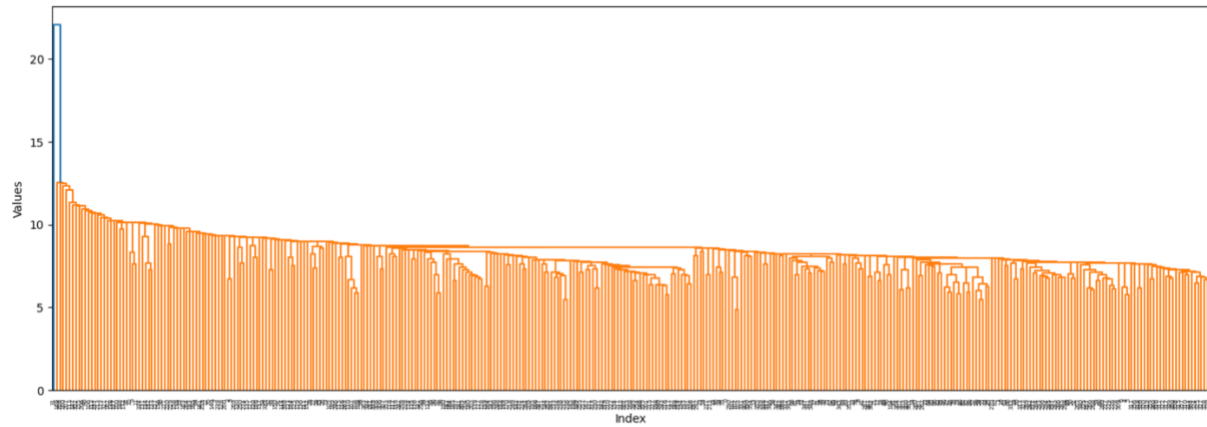


Hierarchical Clustering Analysis

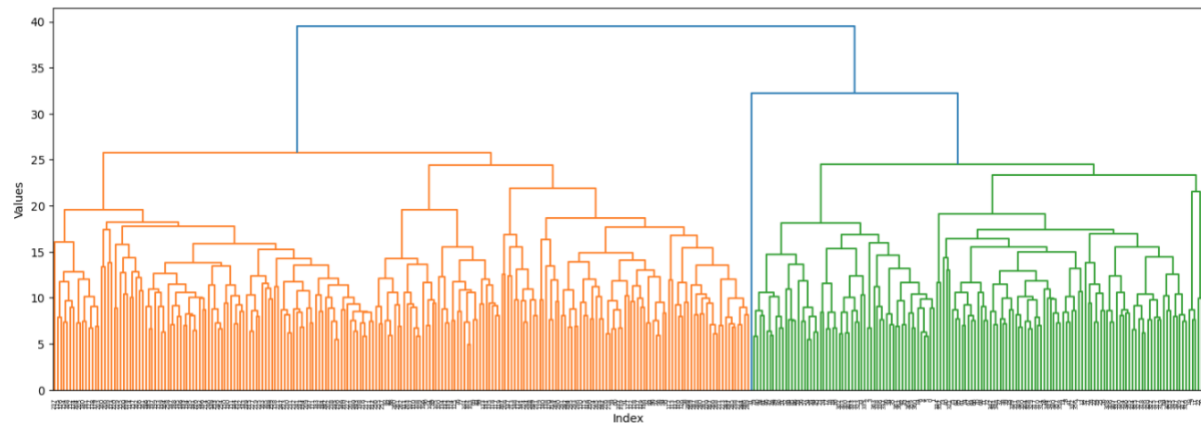
Laura DeCesare

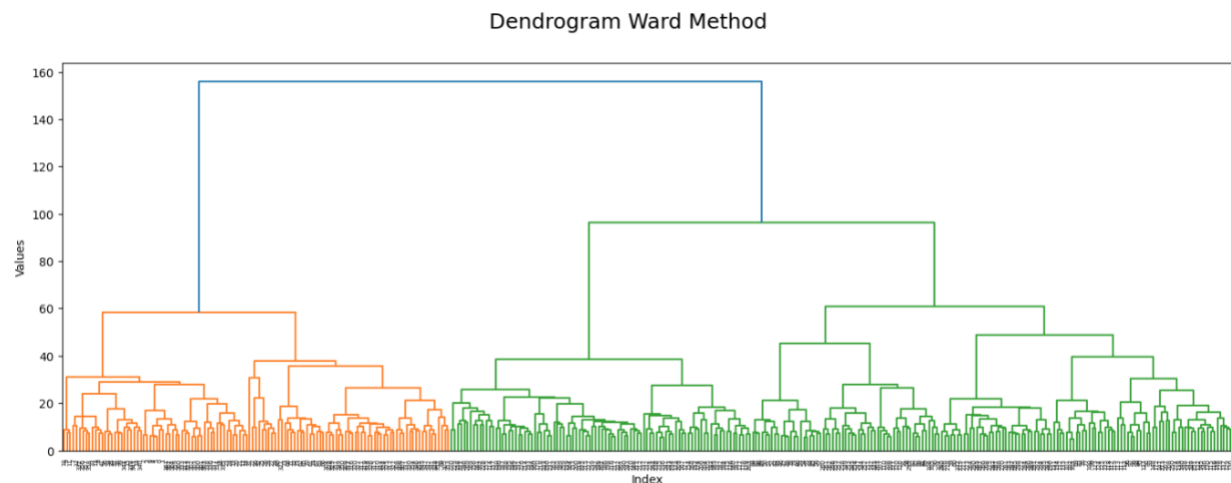
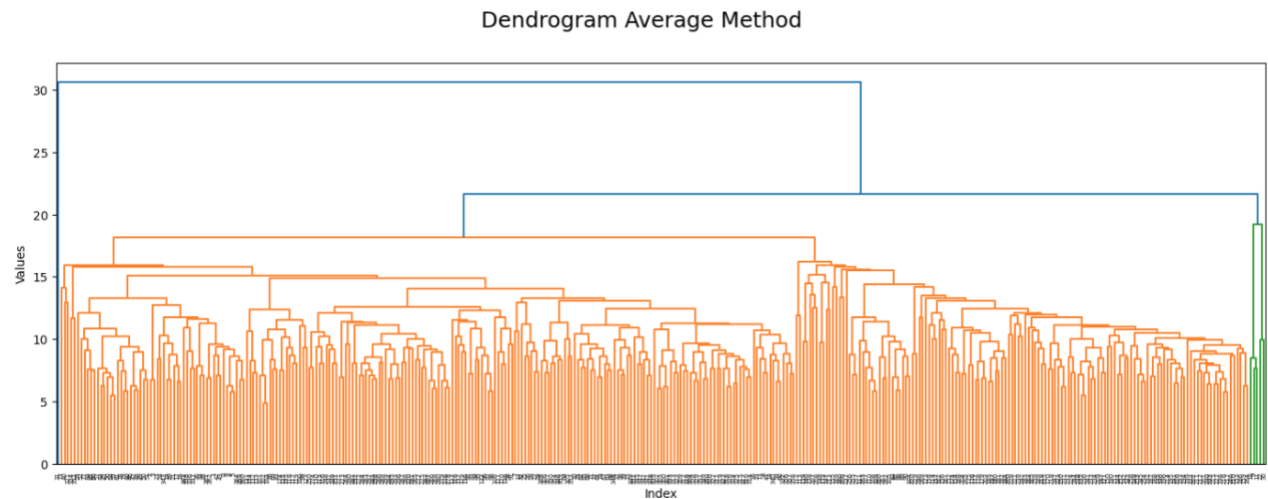
Part 1: Scaled Data, 2019

Dendrogram Single Method



Dendrogram Complete Method





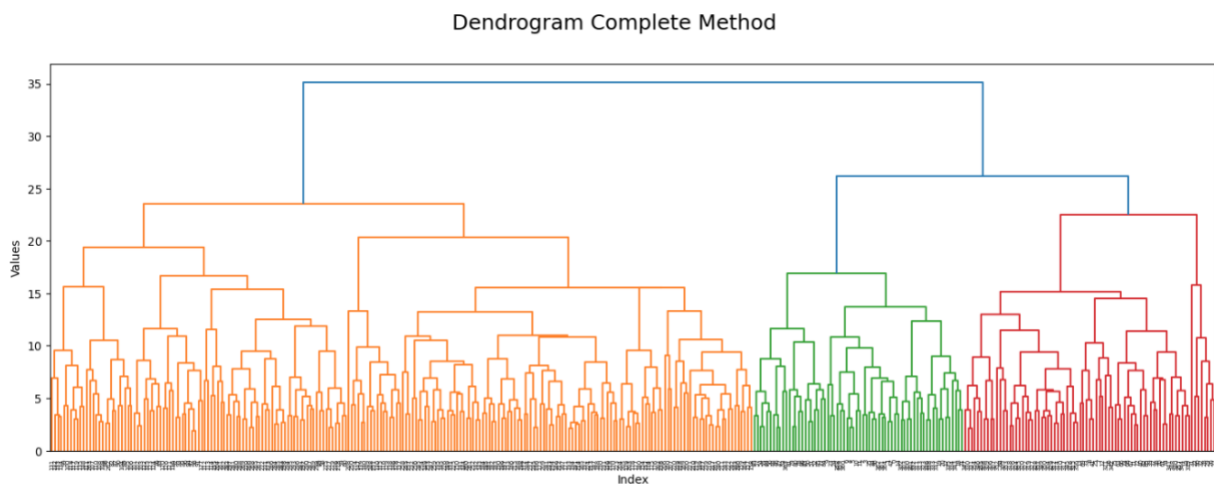
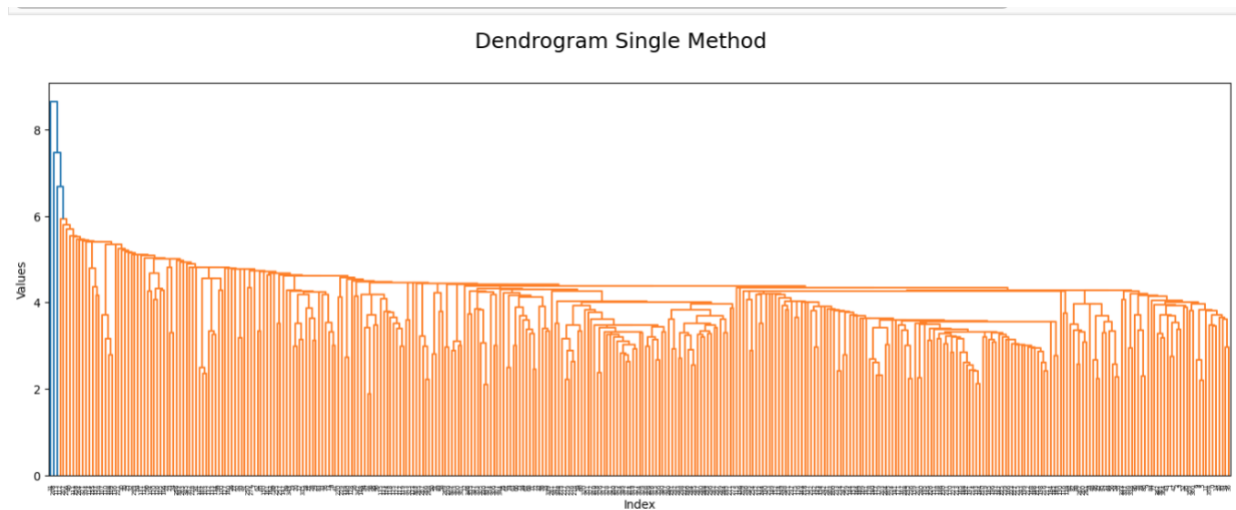
Insights:

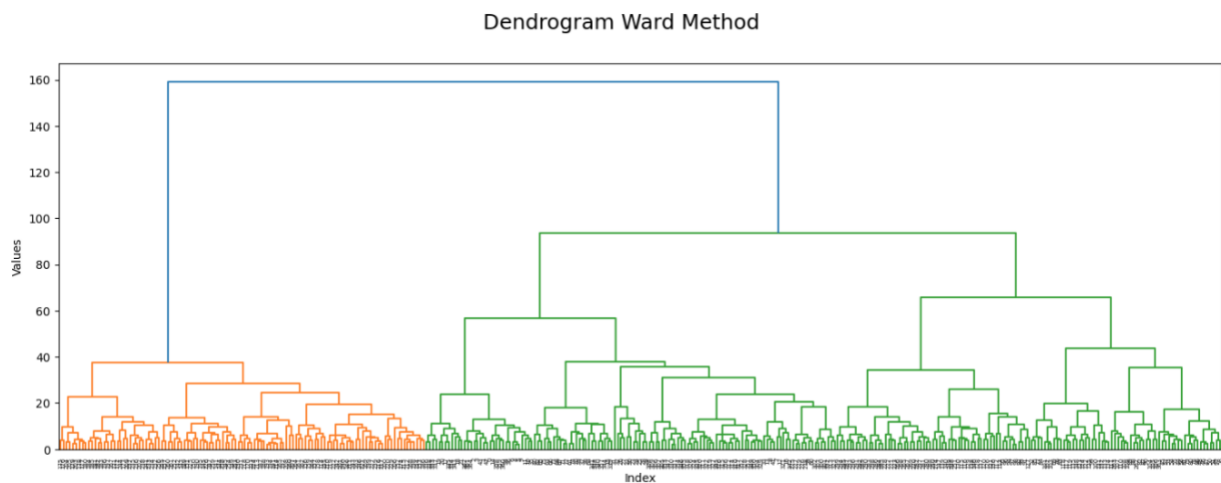
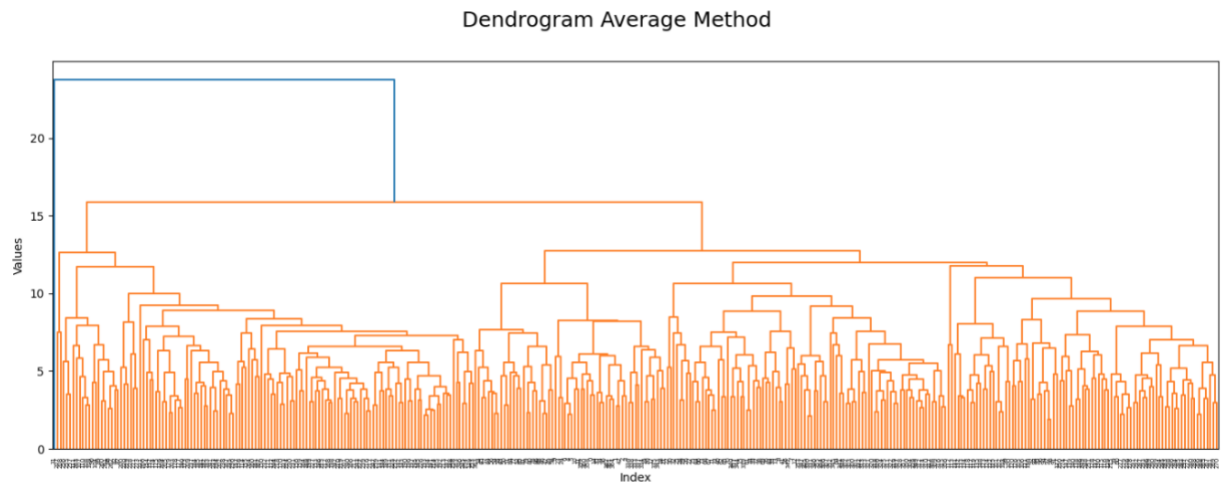
All three methods produce two distinct clusters. The Complete and Ward methods produce clusters that are closer in size, while one cluster is extremely small in both the Ward and Single methods.

Because pleasant vs. unpleasant weather is a binary classification, it's possible that either the Complete or Ward methods produce that result. But without further analysis, it's impossible to know for sure.

Part 2: Reduced Data

By performing a Principal Component Analysis (PCA) on weather data for 2019, we can reduce the data into nine categories. This corresponds with the number of factors analyzed and produces the following dendrograms:





This produces a varying number of clusters. The clearest difference is in the Complete method, where we have three clusters, one about twice as large as the other two. It's difficult to think of what that might represent from a weather classification standpoint. It may be creating three categories of temperature or livability, but without more information, we can't be more specific.