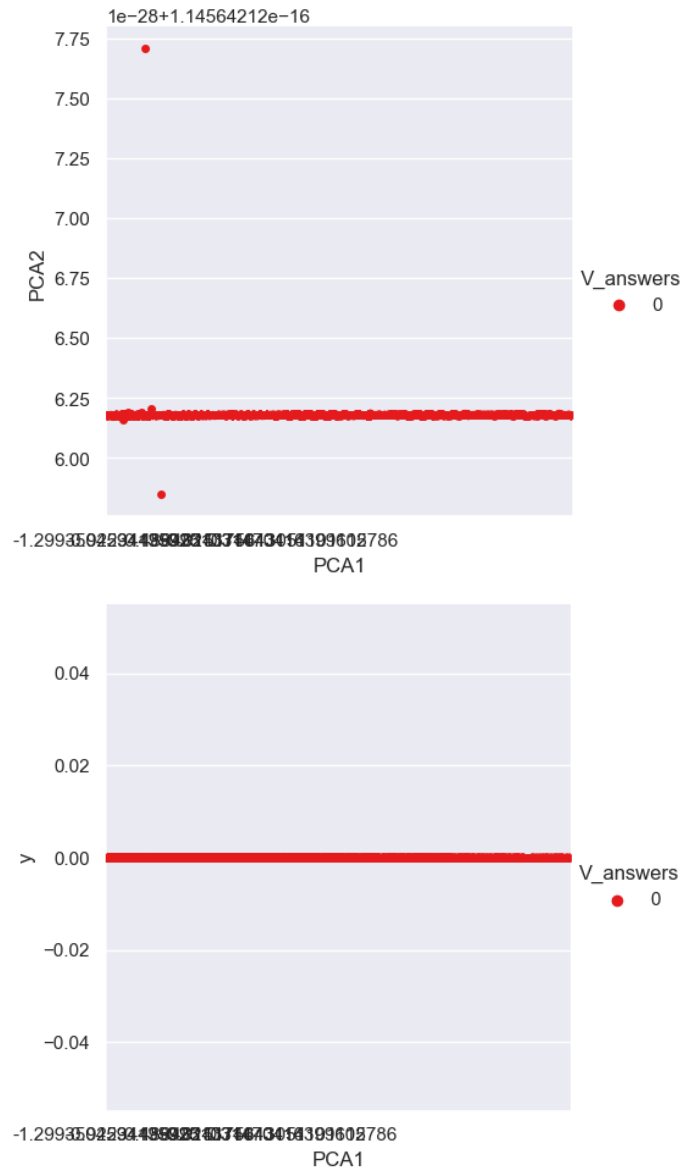## Real-World Applications of Machine Learning
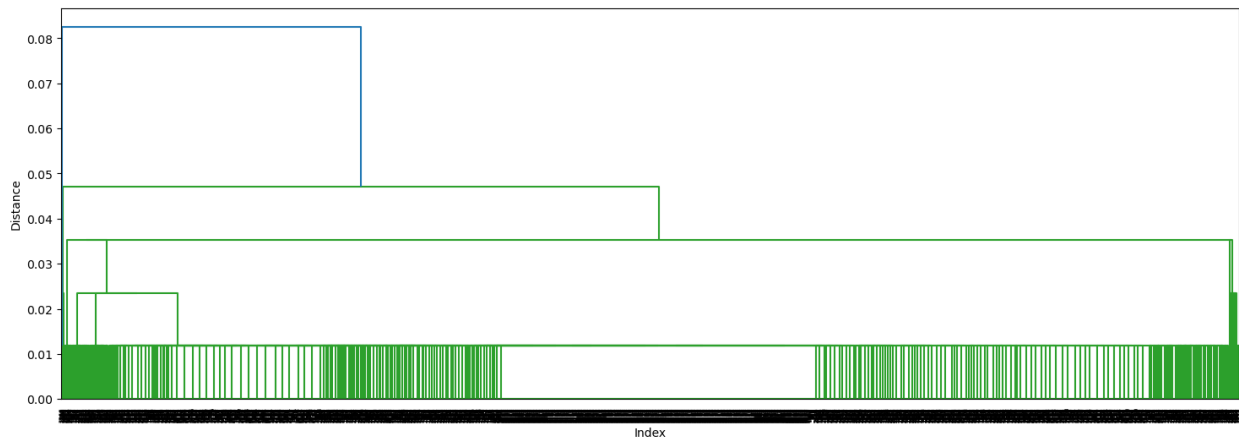## 2.1: Unsupervised Learning Algorithms

Cat plot of all data elements using dimensionality reduction for VALENTIA weather station categorized by its pleasant weather outcomes.
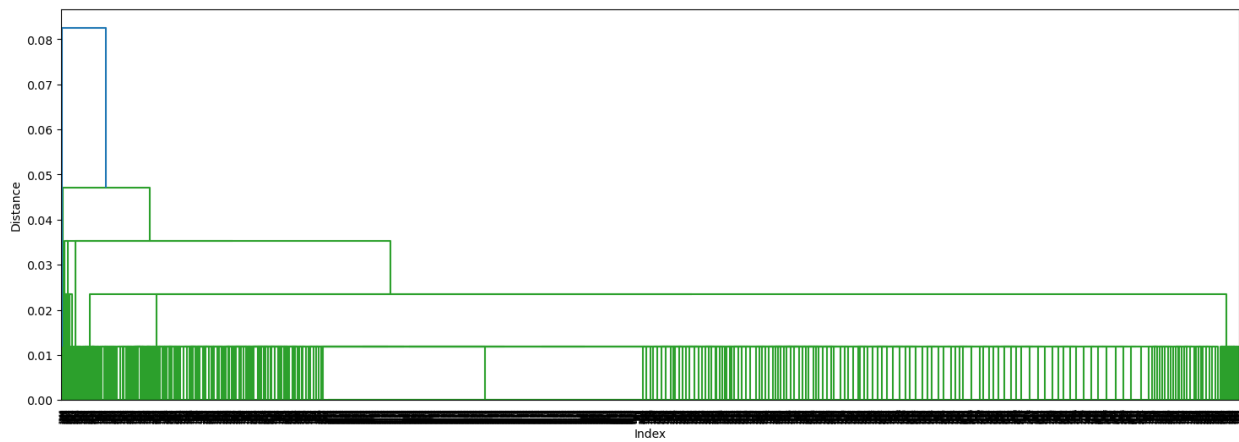

- As we can see, in the last decade, VALENTIA has had no pleasant weather outcomes. *[Yes = 1, No = 0]*
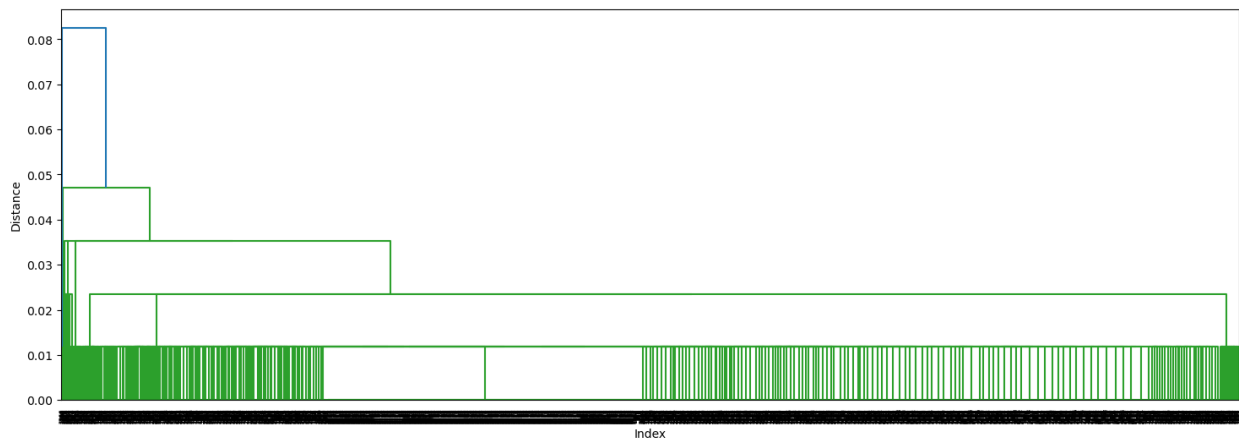
VALENTIA - Dendrogram Single Method

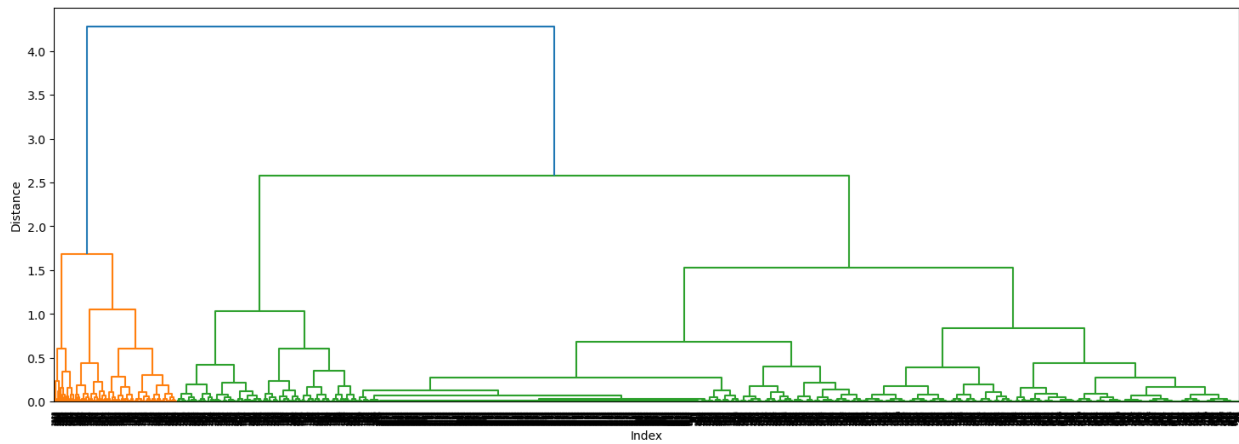VALENTIA - Dendrogram Single Method (PCA)

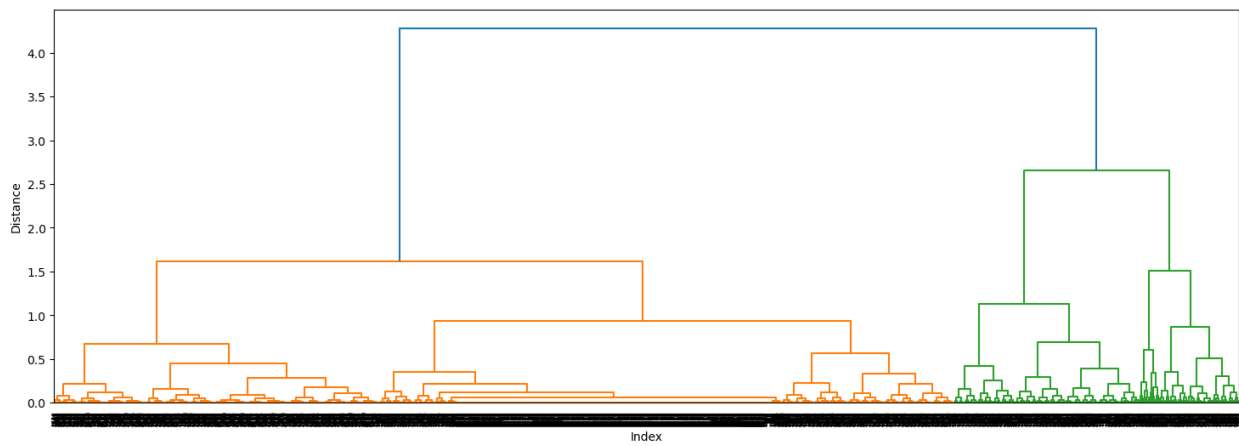VALENTIA - Dendrogram Single Method (PCA2)

- **Dendrogram Single Method** - Least optimal method with no discernable clusters.
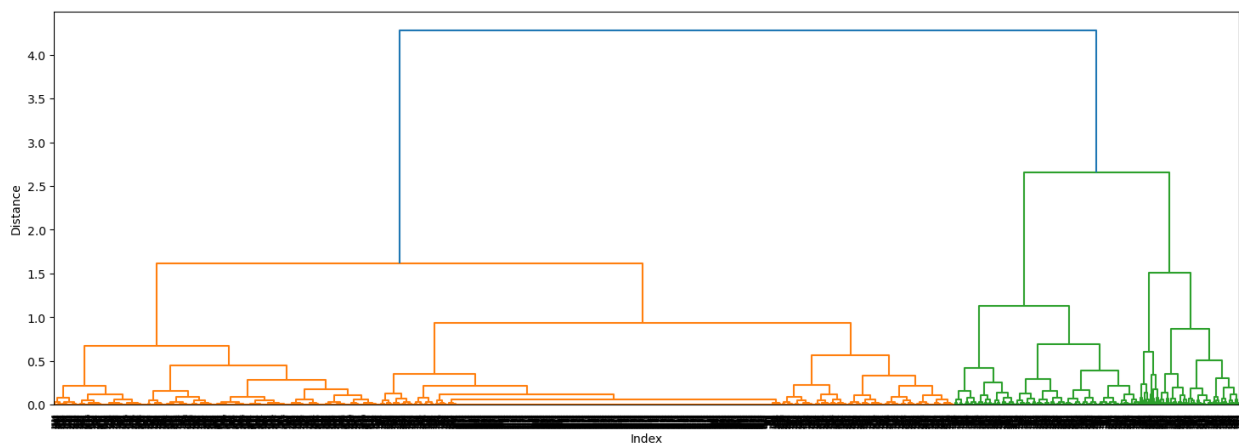
VALENTIA - Dendrogram Complete Method

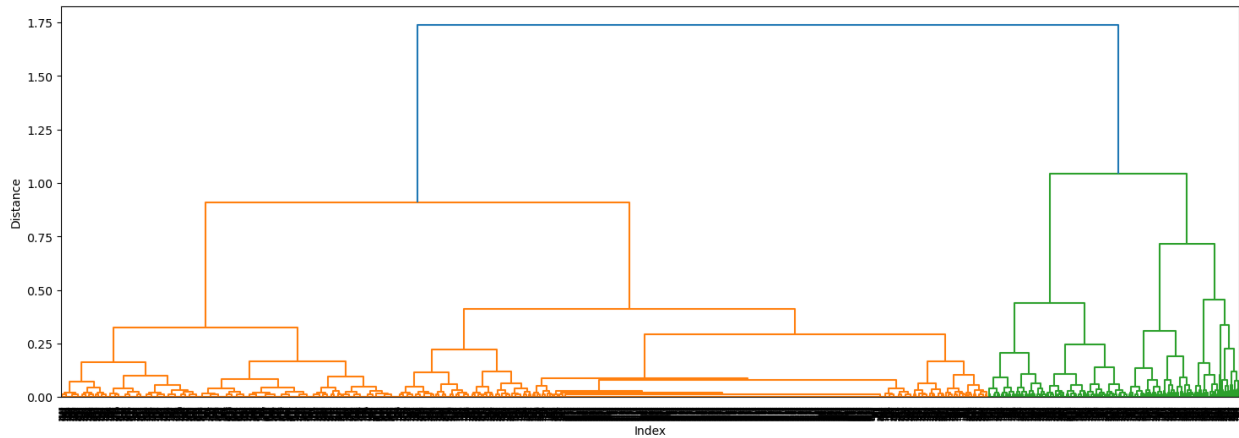VALENTIA - Dendrogram Complete Method (PCA)

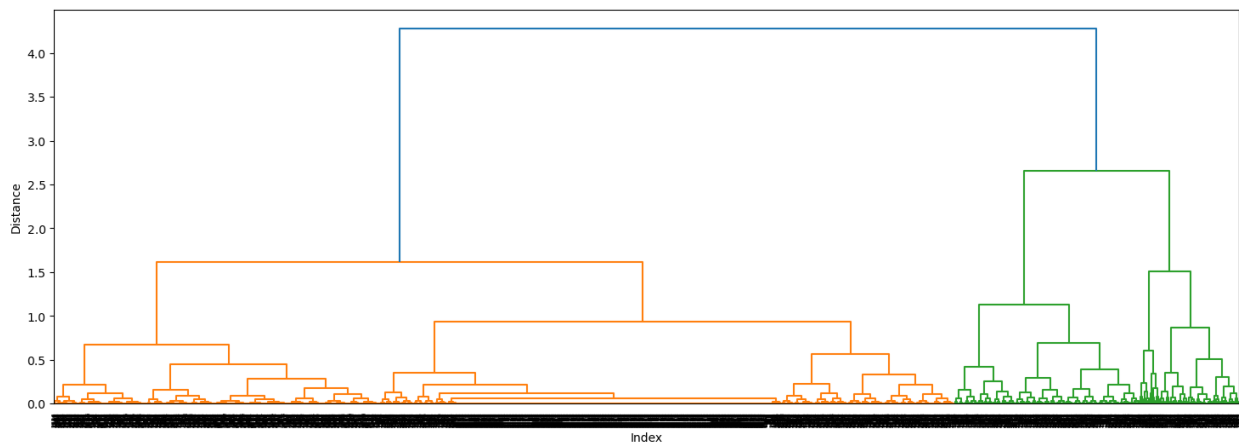VALENTIA - Dendrogram Complete Method (PCA2)

- **Dendrogram Complete Method** - Efficient method for scaled weather and reduced data, with PCA results yielding more discernable clusters.
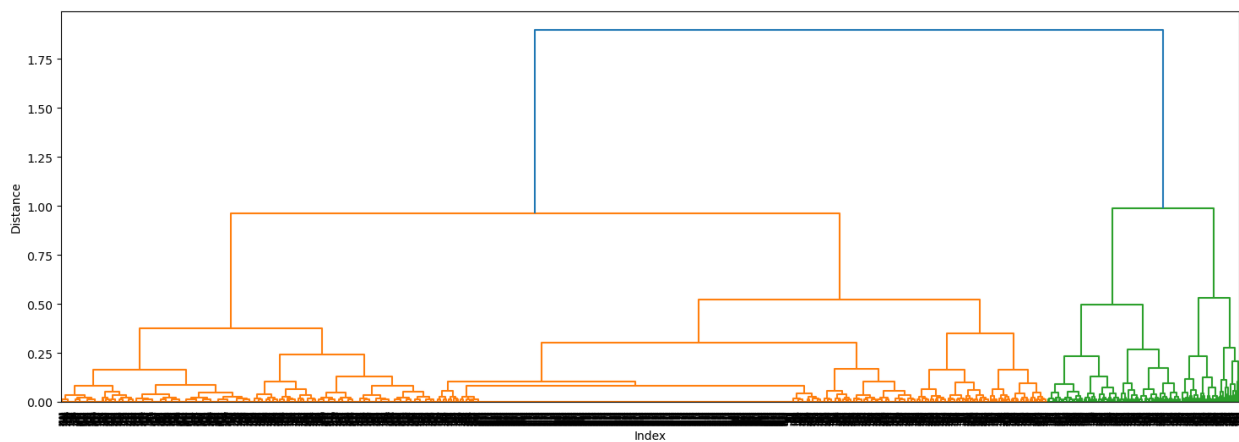
VALENTIA - Dendrogram Average Method

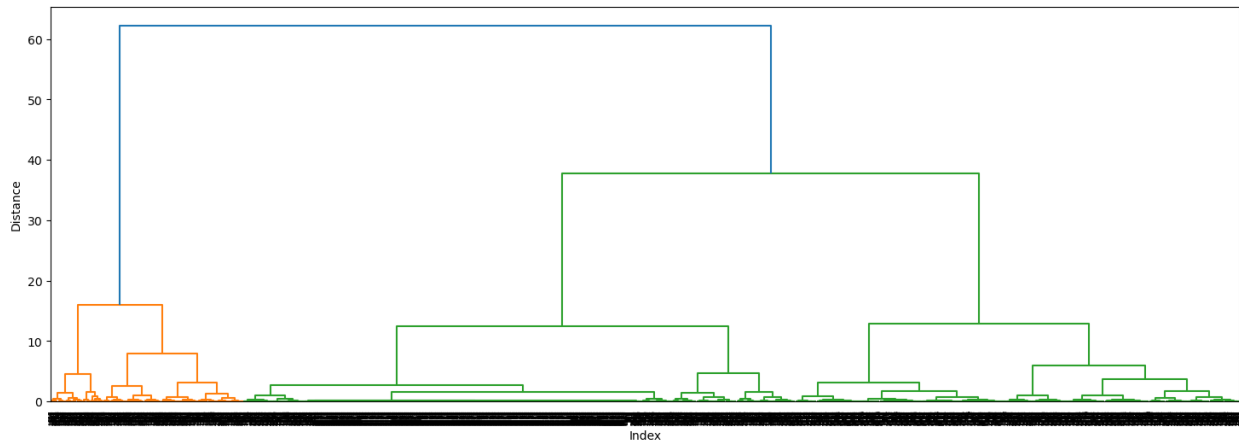VALENTIA - Dendrogram Complete Method (PCA)

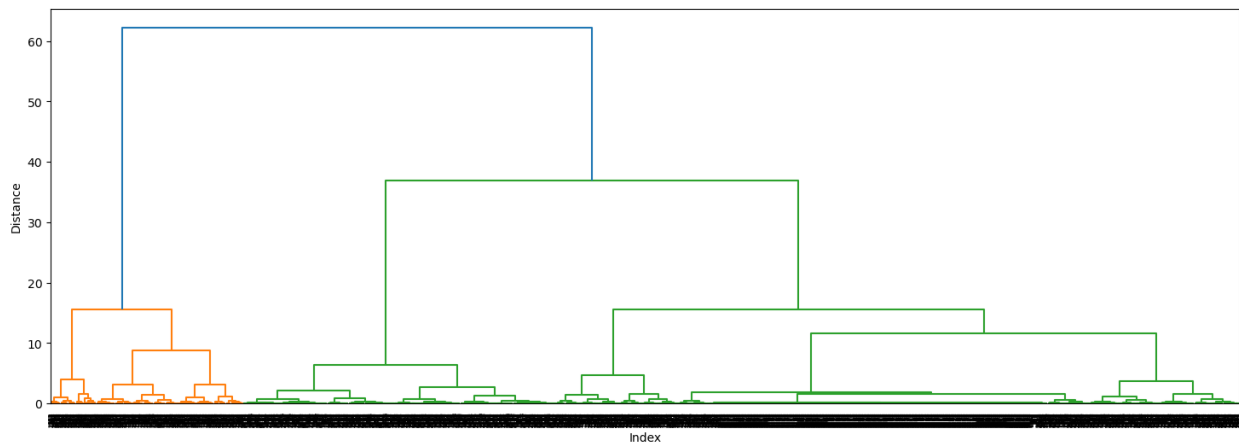VALENTIA - Dendrogram Average Method (PCA2)

- **Dendrogram Average Method** – produces similar clusters around 2-4. This is also an optimal method choice for clustering.
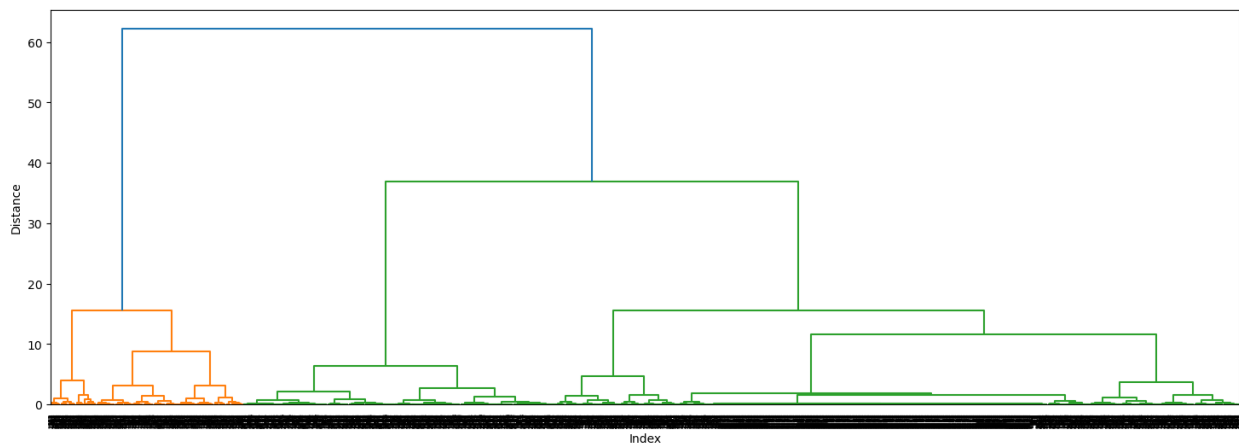
VALENTIA - Dendrogram Ward Method

VALENTIA - Dendrogram Ward Method (PCA)

VALENTIA - Dendrogram Ward Method (PCA2)

- **Dendrogram Ward Method** – produces the most distinguishable clustering around 3-4 clusters. This is the most optimal method choice for ClimatWins.

- **Overall Interpretation:** the *dendrogram ward method* yielded the most interpretable clusters while the *average* and *complete* method yield similar results to each other and just slightly less optimal than *ward* for clustering our weather data for ClimateWins.

- **Challenges:** overlapping data label points, categorical plot by pleasant weather outcomes from PCA model.

- **Final notes:** reducing the datasets by 5, 2, and 1 components yielded similar results.