**Section 6.10: Unsupervised Learning**

**Duration:** 3 hours

**Concepts:**

* Principal Components Analysis
* Matrix Completion
* K-Means Clustering
* Hierarchical Clustering

**Textbook section:** An Introduction to Statistical Learning, Chapter 11

| **Materials and Resources** | **Learning Goals** |
| --- | --- |
| * Computers for students with R Studio * Unsupervised learning slides * Unsupervised learning exercises R Markdown file | * What is PCA and why is it useful * Clustering methods |

| **Duration** | **Lesson Section** | **Learning Objectives** |
| --- | --- | --- |
| 40 mins | Go through the “Principal Components Analysis” section. | * Finding the first principal component * Finding the next principal components * Loading vectors * Scores * Interpretation of principal components * Biplot * PVE * Scree plot |
| 15 mins | Go through the “Principal Components Analysis” section in the R Markdown file as a class. | * Use `prcomp()` to perform PCA * Make biplots * Interpret results * Make scree plot |
| 5 mins | Go through the “Missing Values and Matrix Completion” section. | * Basic idea of matrix completion |
| 10 mins | Go through the “Matrix Completion” section in the R Markdown file as a class. | * Use `softImpute()` to fill in missing matrix values * Compute the correlation between the true and filled-in values |
| 20 mins | Go through the “K-Means Clustering" section. | * K-means clustering algorithm |
| 20 mins | Go through the "K-Means Clustering" section in the R Markdown file as a class. | * Use `kmeans()` to cluster 2D and 4D data * Make visualisations of clustered data * Use PCA to visualise the 4D clusters |
| 20 mins | Go through the “Hierarchical Clustering" section. | * Dendrogram * Dissimilarity measures * Complete, average, single, and centroid linkage * Hierarchical clustering algorithm |
| 15 mins | Go through the "Hierarchical Clustering" section in the R Markdown file as a class. | * Use the `hclaust()` function to cluster * Plot dendrograms |