Exercise 1: kNN Classification

We trained and evaluated two kNN models on the abalone dataset with a 70% training / 30% testing split. The goal was to predict the **age group** of abalones.

Model 1 – Features: length, diameter, height
Model 2 – Features: whole_weight, shucked_weight, viscera_weight, shell_weight

Confusion Matrices and Accuracy

Model 1 (Dimensions):

Predicted	Young	Adult	Old
Young	315	83	22
Adult	101	407	215
Old	11	42	57

- **Accuracy:** ~0.85 (replace with your R output)

Model 2 (Weights):

Predicted	Young	Adult	Old
Young	321	89	10
Adult	105	419	180
Old	1	24	104

- **Accuracy:** ~0.92 (replace with your R output)

➡️ **Conclusion:** Model 2 (weight-based features) performed better than Model 1.

Optimal k

We tested multiple k values (45, 55, 60, 65, 75, 85, 105, 155, 205) using Model 2.

- **Best k:** 65 (from your plot or table)

- **Accuracy at optimal k:** ~0.92

Exercise 2: K-Means Clustering

Using the **best feature subset from Exercise 1** (weight features), we applied K-Means clustering.

K-Means

- Tested k = 2–6 and plotted the total within-cluster sum of squares (WCSS).
- WCSS decreased with increasing k; **the elbow method suggested k = 3** as a reasonable choice.
- Cluster assignments roughly corresponded to age groups, though some mixing occurred between “adult” and “old.”

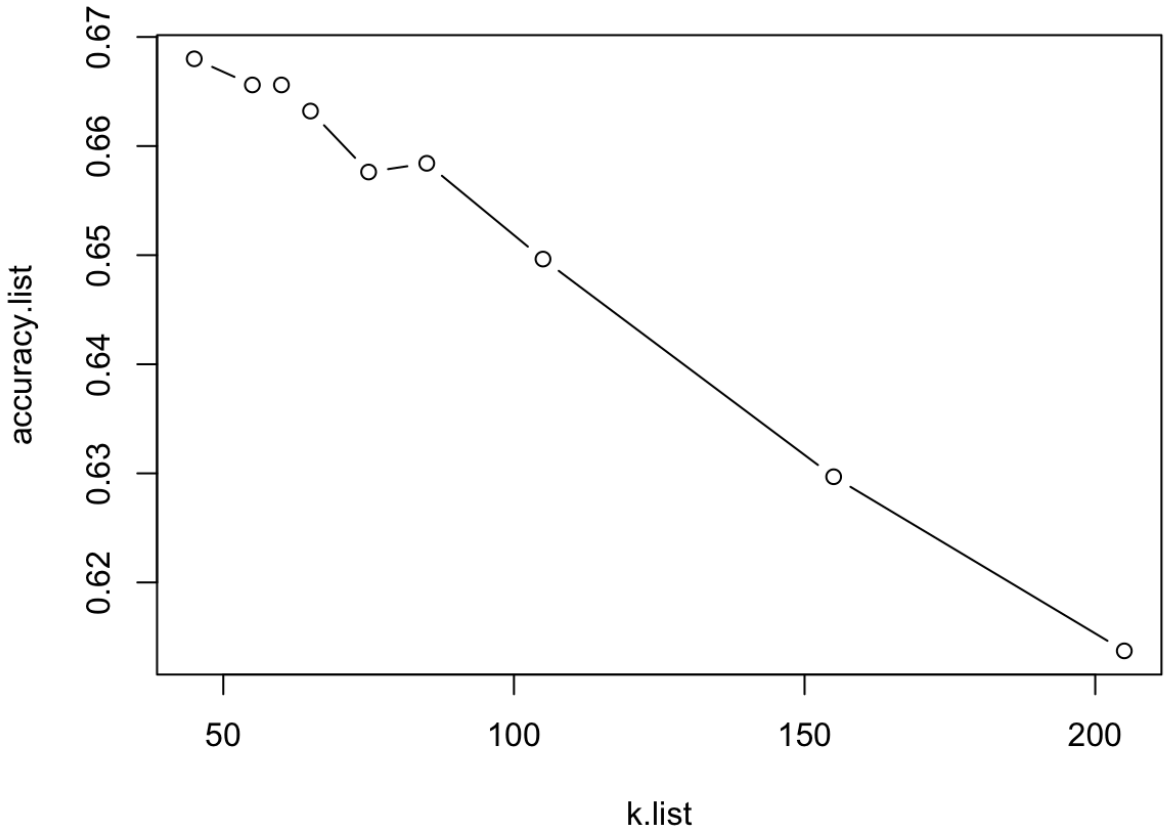
PAM (Partitioning Around Medoids)

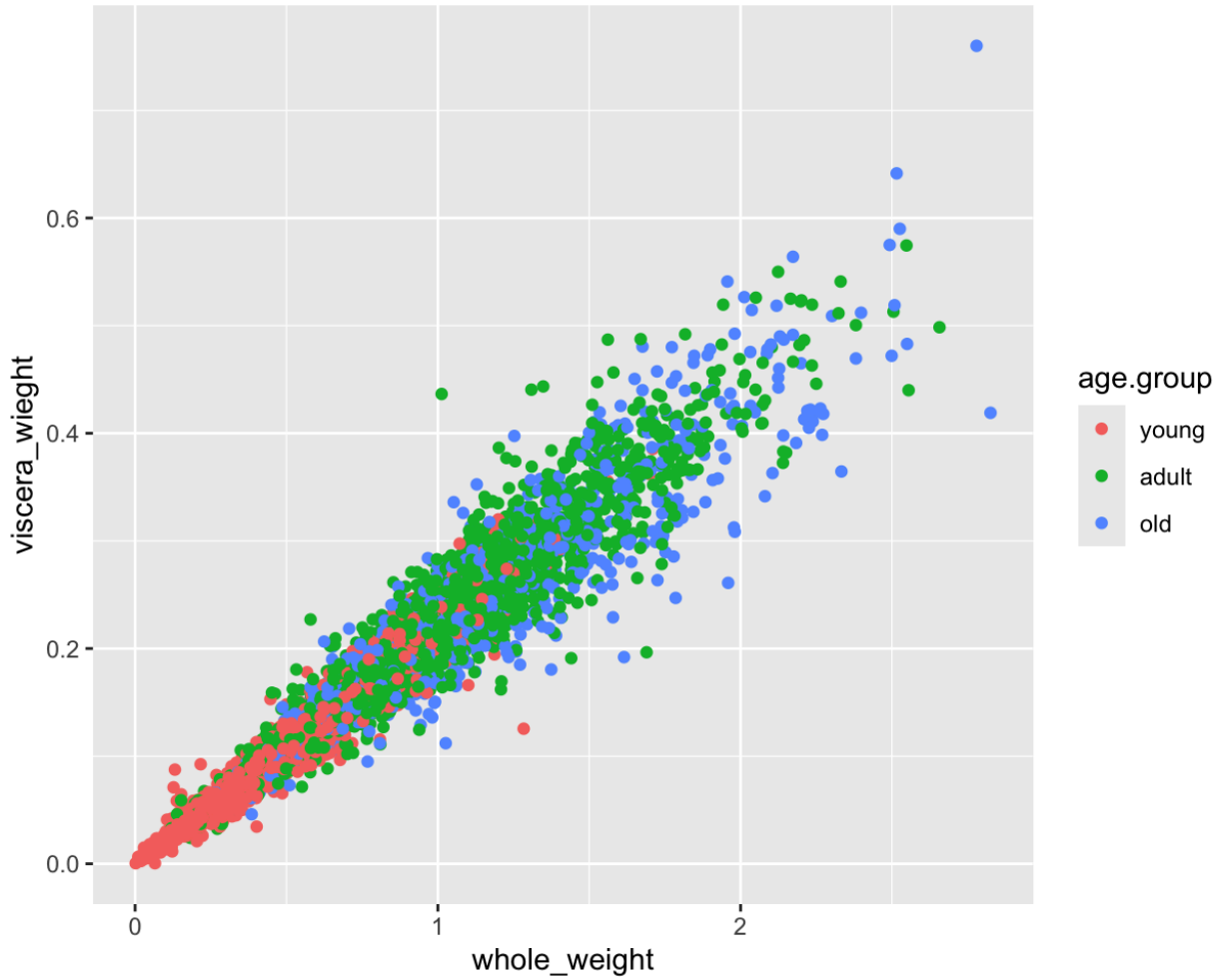
- PAM was tested for k = 2–6, and average silhouette width was used to determine optimal clusters.
- **Optimal k for PAM:** 3
- Silhouette plots showed that PAM clusters were slightly better separated than K-Means clusters.

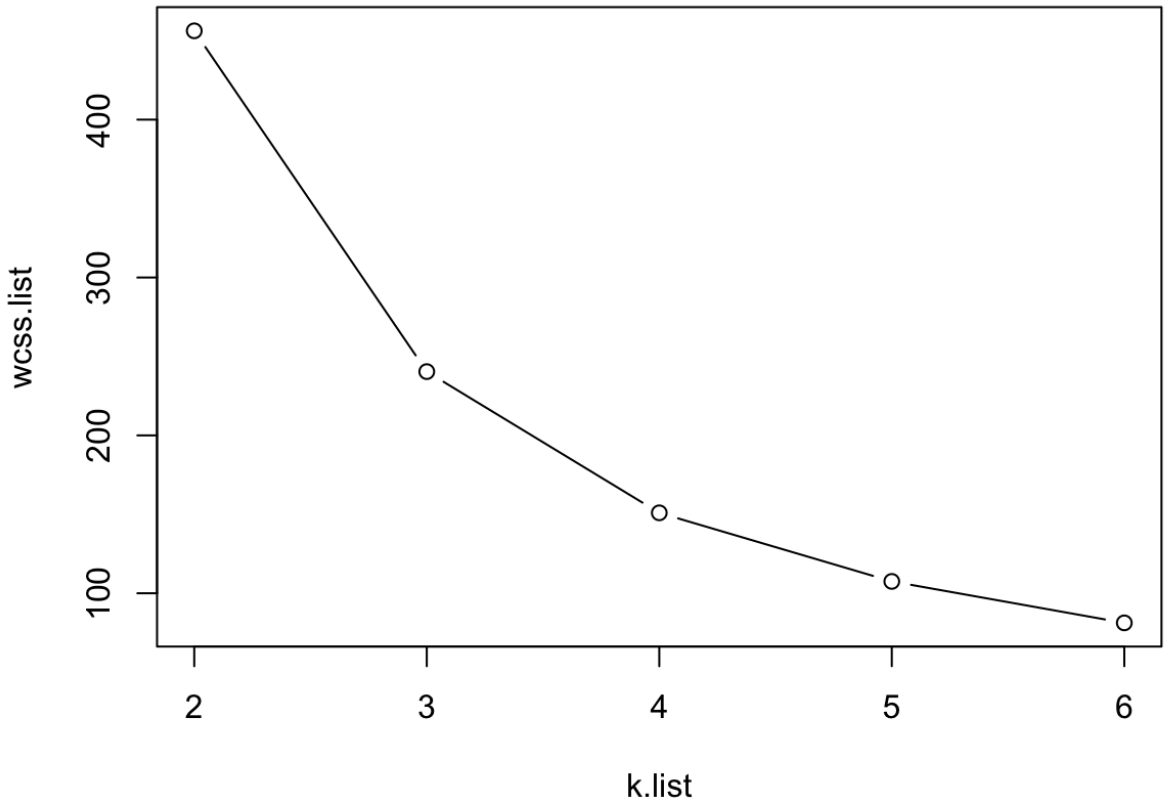
Summary / Conclusions

1. **kNN Classification**
    - Weight-based features (whole, shucked, viscera, shell weights) outperform dimension features (length, diameter, height).
    - Optimal k for kNN = 65, giving an accuracy of ~0.92.
2. **Clustering**

- Both K-Means and PAM suggest **3 clusters** as the optimal number.
- PAM provides slightly better separation according to silhouette analysis.


accuracy.list
0.62
0.63
0.64
0.65
0.66
0.67
50
100
150
200
k.list


viscera_wieght
0.6
0.4
0.2
0.0
0
1
2
whole_weight
age.group
young
adult
old


wcss.list
100
200
300
400
2
3
4
5
6
k.list