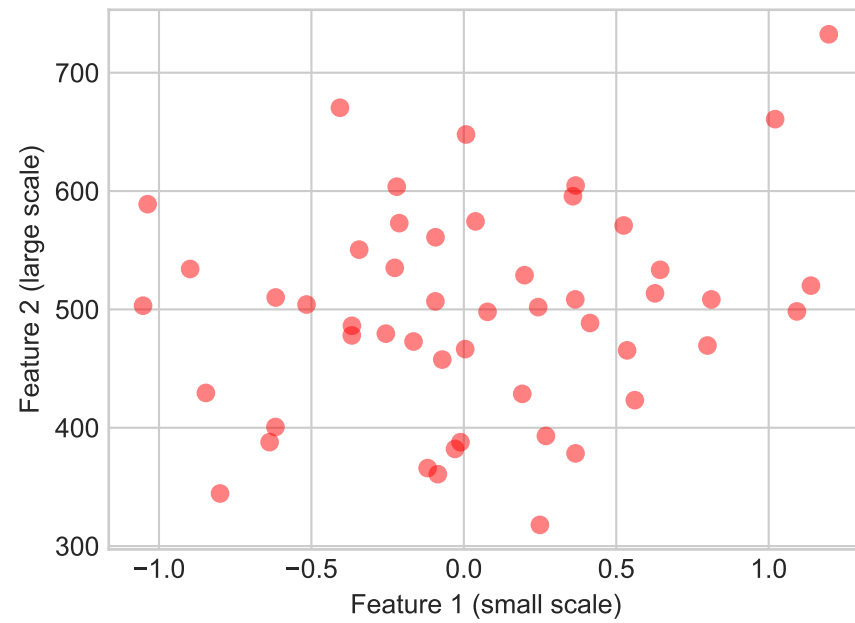
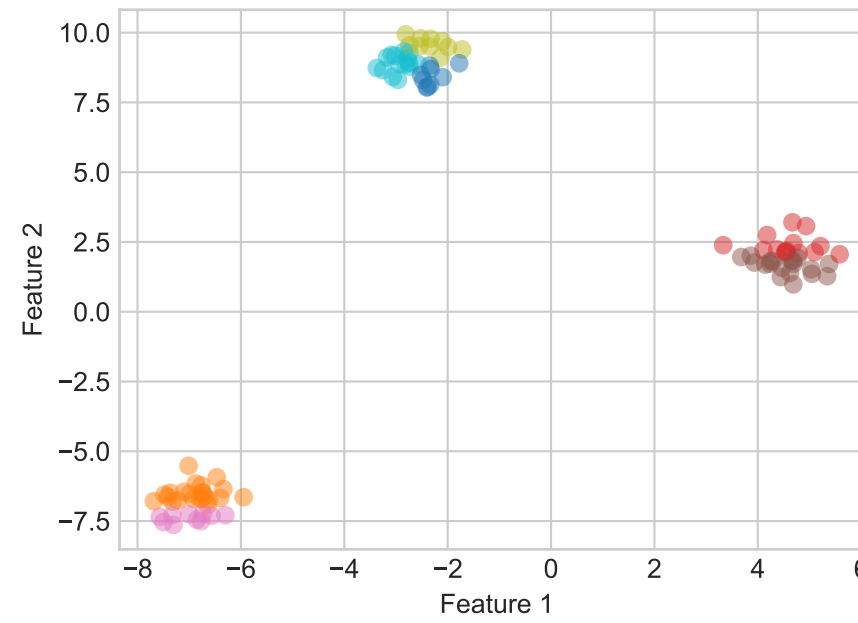


# Common Clustering Mistakes to Avoid

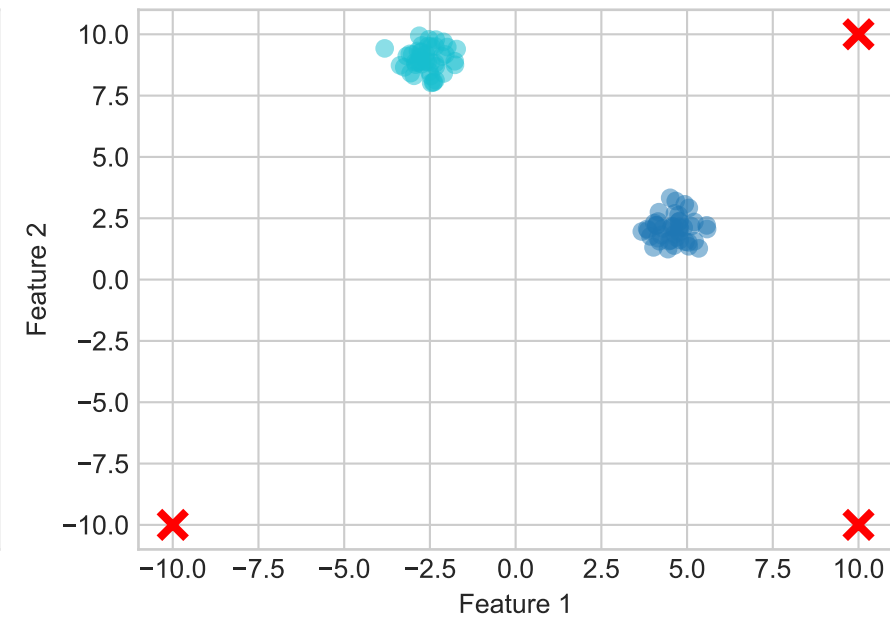
**No Scaling**  
Features have  
different scales



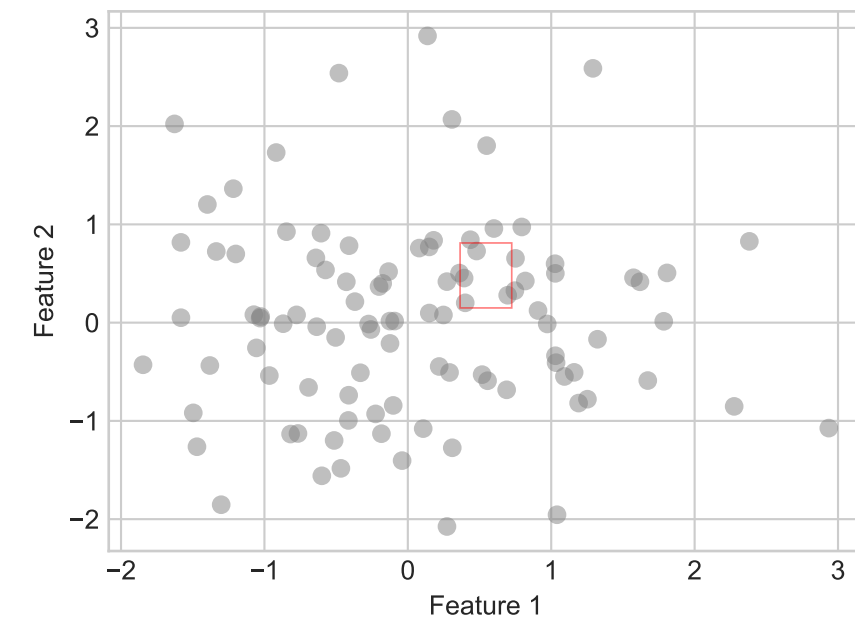
**Wrong K**  
Too many/few  
clusters



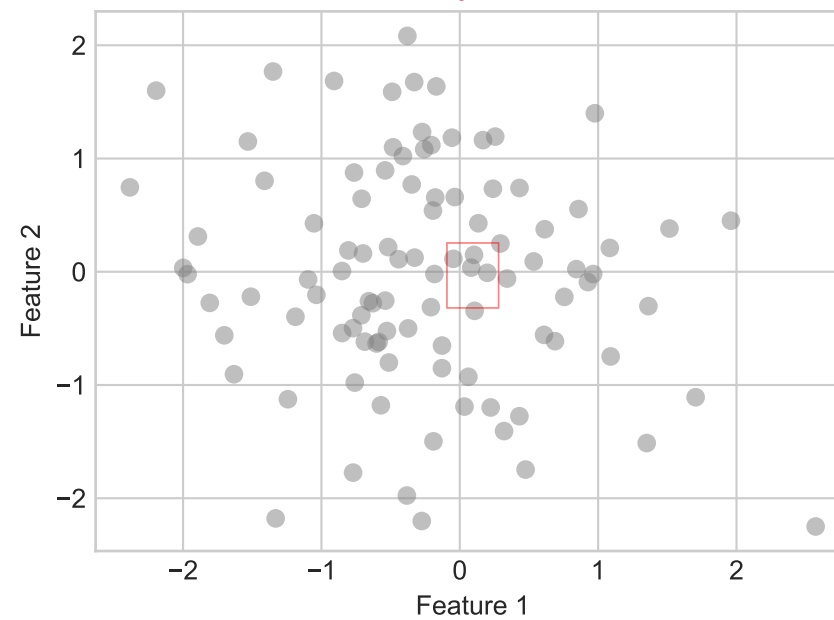
**Ignoring Outliers**  
Outliers skew  
centers



**No Validation**  
No quality  
metrics used



**Wrong Algorithm**  
K-means for  
non-spherical



**Overfitting**  
Too complex  
for data

