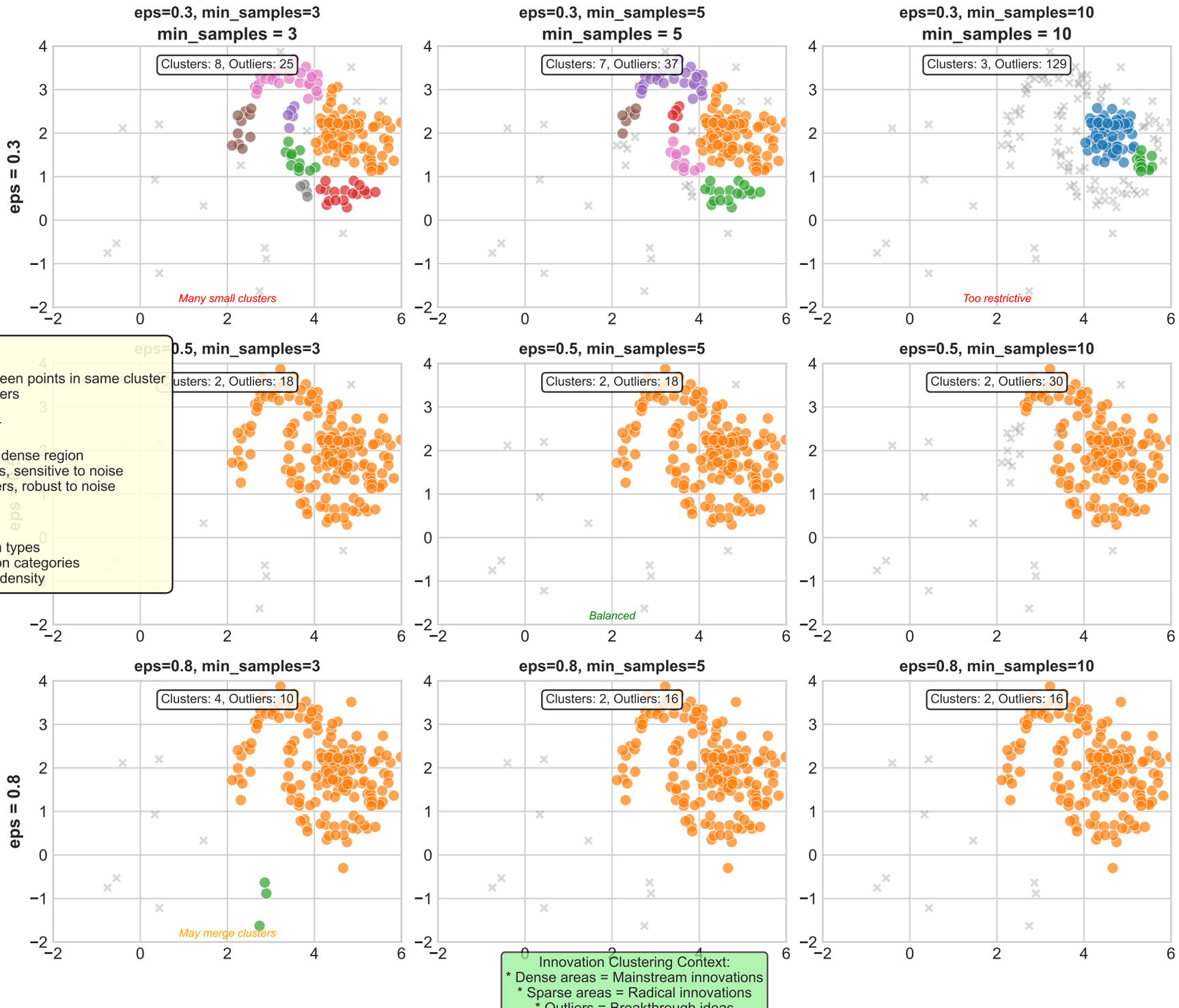
DBSCAN Parameter Tuning: Impact on Innovation Clustering



Parameter Guidelines:

eps (epsilon): Maximum distance between points in same cluster
* Small eps → Many small, tight clusters
* Large eps → Fewer, larger clusters
* Too large → All points in one cluster

min_samples: Minimum points to form dense region
* Small min_samples → More clusters, sensitive to noise
* Large min_samples → Fewer clusters, robust to noise
* Too large → Many outliers

For Innovation Data:
* Use small eps for distinct innovation types
* Use larger eps for broader innovation categories
* Adjust min_samples based on data density

Tuning Strategy:
1. Start with k-distance plot
2. Look for 'elbow' in plot
3. Set eps at elbow point
4. min_samples = 2*dimensions
5. Validate with domain knowledge

Innovation Clustering Context:
* Dense areas = Mainstream innovations
* Sparse areas = Radical innovations
* Outliers = Breakthrough ideas