2.

| Column  Name | Feature  Type | Way to manage a feature | Example before and after the feature  transformation |
| --- | --- | --- | --- |
| Overall | Numeric | Standardize the feature to have zero mean and unit variance. | Before: Overall = 85  After: If the mean of Overall is 70 and the standard deviation is 10, then the standardized value would be (85 - 70) / 10 = 1.5. |
| Potencial | Numeric | Standardize the feature to have zero mean and unit variance. | Before: Potential = 90  After: If the mean of Potential is 75 and the standard deviation is 8, then the standardized value would be (90 - 75) / 8 = 1.875. |
| Age | Numeric | Standardize the feature to have zero mean and unit variance. | Before: Age = 30  After: If the mean of Age is 25 and the standard deviation is 5, then the standardized value would be (30 - 25) / 5 = 1.0. |
| Stamina | Numeric | Standardize the feature to have zero mean and unit variance. | Before: Stamina = 75  After: If the mean of Stamina is 65 and the standard deviation is 10, then the standardized value would be (75 - 65) / 10 = 1.0. |
| Dribbling | Numeric | Standardize the feature to have zero mean and unit variance. | Before: Dribbling = 80  After: If the mean of Dribbling is 70 and the standard deviation is 8, then the standardized value would be (80 - 70) / 8 = 1.25. |

4.

| Clusters ID | The number of players from the 10 players with biggest salary |
| --- | --- |
| 0 | 0 |
| 1 | 0 |
| 2 | 0 |
| 3 | 10 |

5. use PCA to reduce the dimensions of features vector. is it more helpful for clustering?

The number of players in each cluster remains relatively consistent across different numbers of PCA components (5, 4, 3). This suggests that PCA is maintaining the general structure of the data even as the dimensionality is reduced.

For 5 Features:

| Clusters ID | The number of players from the 10 players with biggest salary |
| --- | --- |
| 0 | 0 |
| 1 | 0 |
| 2 | 0 |
| 3 | 10 |

For 4 Features:

| Clusters ID | The number of players from the 10 players with biggest salary |
| --- | --- |
| 0 | 0 |
| 1 | 0 |
| 2 | 0 |
| 3 | 10 |

For 3 Features:

| Clusters ID | The number of players from the 10 players with biggest salary |
| --- | --- |
| 0 | 0 |
| 1 | 10 |
| 2 | 0 |
| 3 | 0 |

6.

