



What is the average radius of cell nuclei for malignant tumours compared to benign tumours?

- **Benign tumours:** 12.15
- **Malignant tumours:** 17.46

What are the ranges (min-max) of the mean area for both classes?

- **Benign tumours:** (143.5, 992.1)
- **Malignant tumours:** (361.6, 2501.0)

What is the percentage of benign and malignant cases in the dataset?

- **Benign tumours:** 62.74%
- **Malignant tumours:** 37.26%

Which features have the highest correlation with the diagnosis (malignant/benign)?

The top 10 features most correlated with the diagnosis:

1. worst_concave_points: 0.79
2. worst_perimeter: 0.78
3. mean_concave_points: 0.78
4. worst_radius: 0.78
5. mean_perimeter: 0.74
6. worst_area: 0.73
7. mean_radius: 0.73
8. mean_area: 0.71
9. mean_concavity: 0.70

Are there any highly correlated features that might indicate redundancy?

Highly correlated feature pairs (correlation > 0.75):

- mean_perimeter and mean_radius: 0.998
- mean_area and mean_radius: 0.987
- mean_concavity and mean_compactness: 0.883
- mean_concave_points and mean_concavity: 0.921
- se_perimeter and se_radius: 0.973
- se_area and se_radius: 0.952
- se_concavity and se_compactness: 0.801
- se_concave_points and se_concavity: 0.772
- se_fractal_dimension and se_compactness: 0.803
- worst_radius and mean_radius: 0.970
- worst_texture and mean_texture: 0.912
- worst_perimeter and worst_radius: 0.994
- worst_area and worst_radius: 0.984
- worst_smoothness and mean_smoothness: 0.805
- worst_compactness and mean_compactness: 0.866
- worst_concavity and worst_compactness: 0.892
- worst_concave_points and mean_concave_points: 0.910
- worst_fractal_dimension and worst_compactness: 0.810

What do the histograms of the mean radius look like for benign and malignant tumours?

The histograms show that malignant tumours generally have a higher mean radius compared to benign tumours. The distribution for malignant tumours peaks at a higher radius value than for benign tumours.

How do the box plots of mean texture compare between the two classes?

The box plots indicate that malignant tumours have a higher median and a wider range of mean texture values compared to benign tumours.

Is the difference in the mean perimeter between benign and malignant tumours statistically significant?

- **T-statistic:** -22.94
- **P-value:** 1.02e-66

The extremely low p-value indicates that the difference in mean perimeter between benign and malignant tumours is statistically significant.