California Housing - Module 1 (Q2) Summary

Samples: 20640 | Target mean: $206,855.82

Top 3 Models (by Test RMSE)

Model	Train_RMSE	Test_RMSE	Test_MAE
GradientBoosting	$41,537	$50,557	$33,896
RandomForest_100	$41,405	$54,946	$37,056
DecisionTree_d10	$47,315	$65,909	$43,880

Underfitting / Overfitting Notes

- Linear models show underfitting (high train & test RMSE).
- Full decision tree shows overfitting (near-zero train RMSE, high test RMSE).
- Ensembles (RF/GB) balance bias and variance and give best test RMSEs.
- Consider spatial CV and robust losses in production to handle outliers & spatial bias.

Outputs saved under /content/outputs/. Figures: EDA, bias-variance plot, learning curve, residuals, feature importance. Use GitHub link in submission.