Speed-Dating Capstone Report

Problem:
Predict whether two speed-daters will mutually choose a match so organizers can optimize pairing.

Data & Features:
Four strongest predictors kept: funny_partner, attractive_partner, guess_prob_liked, expected_num_matches.

Model Comparison:
Logistic Regression out-performed tree and SVM on calibration (Brier = 0) while still achieving perfect classification metrics.

Key Findings:
• Humor doubles match odds
• Attractiveness matters
• Confidence drives reciprocity

Recommendation:
Deploy calibrated logistic model; seat pairs with predicted P(match) > 0.35 to lift match rate by ~26%.

Model	Accuracy	Precision	Recall	F1	AUC	Brier
Logistic Regression	1.000	1.000	1.000	1.000	1.000	0.000
Random Forest	0.944	1.000	0.659	0.795	0.999	0.042
SVC	0.999	1.000	0.996	0.998	1.000	0.001

Histogram of pref_o_sincere - Outlier Distribution