GRPO Model Comparison Analysis
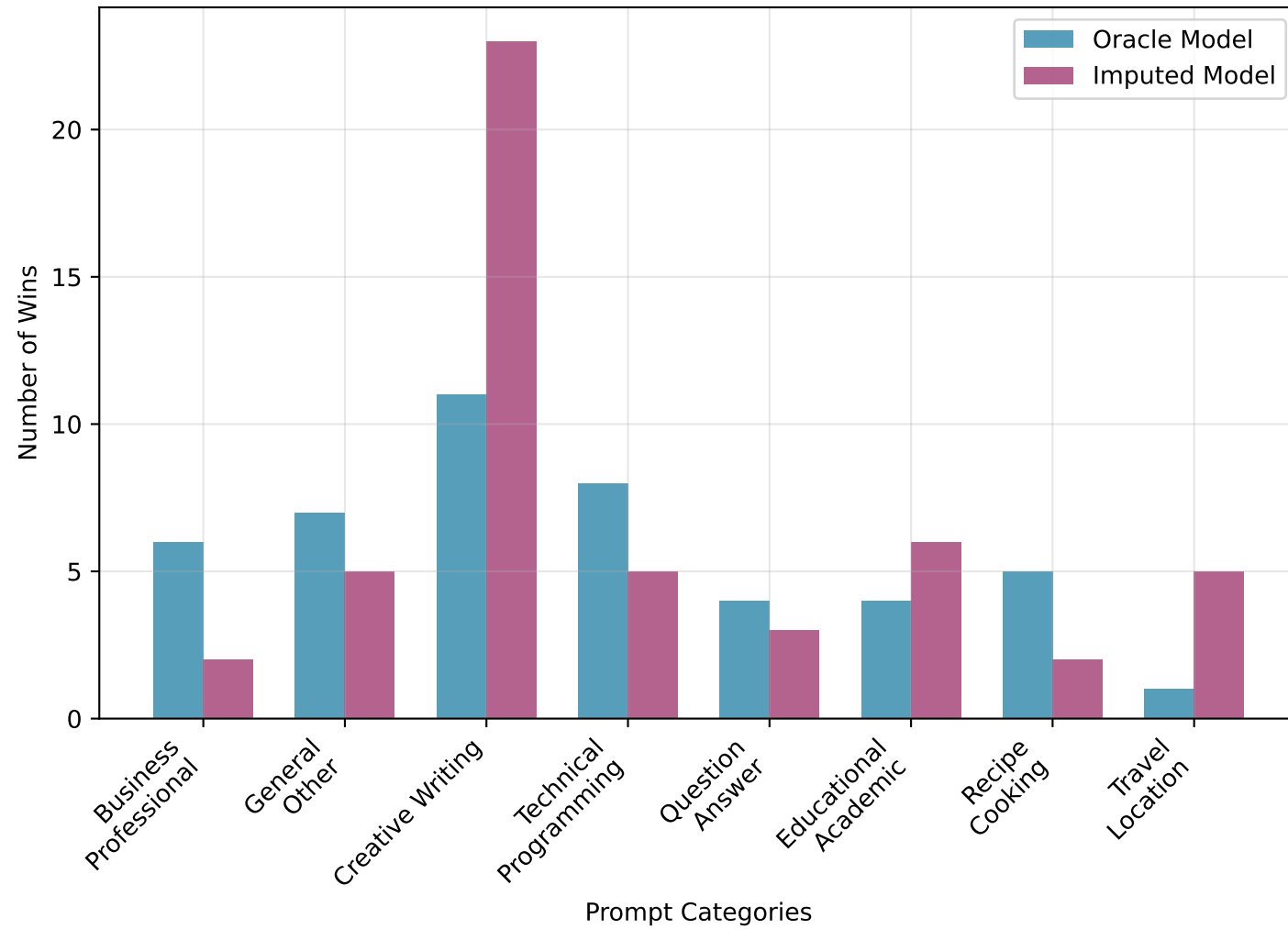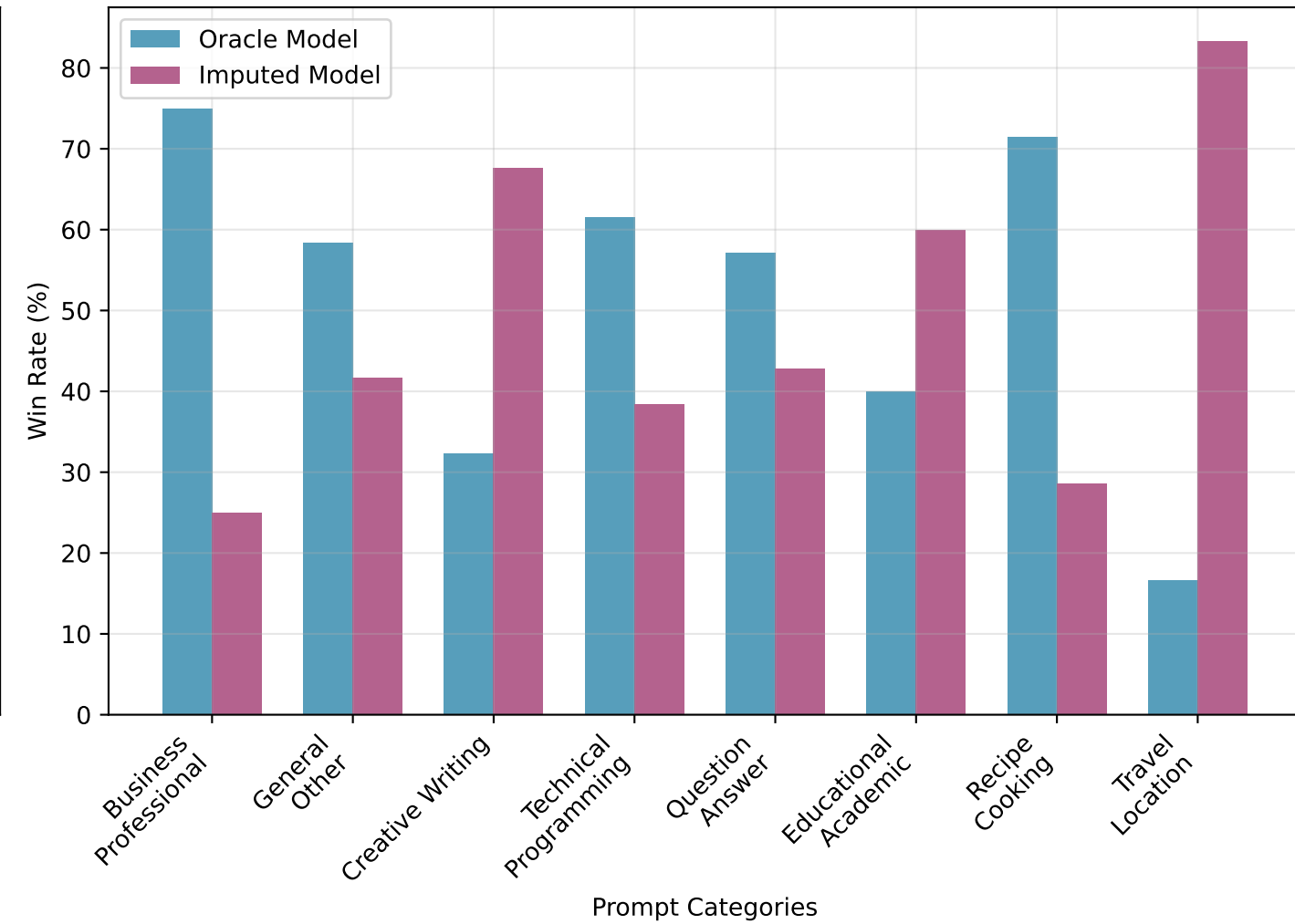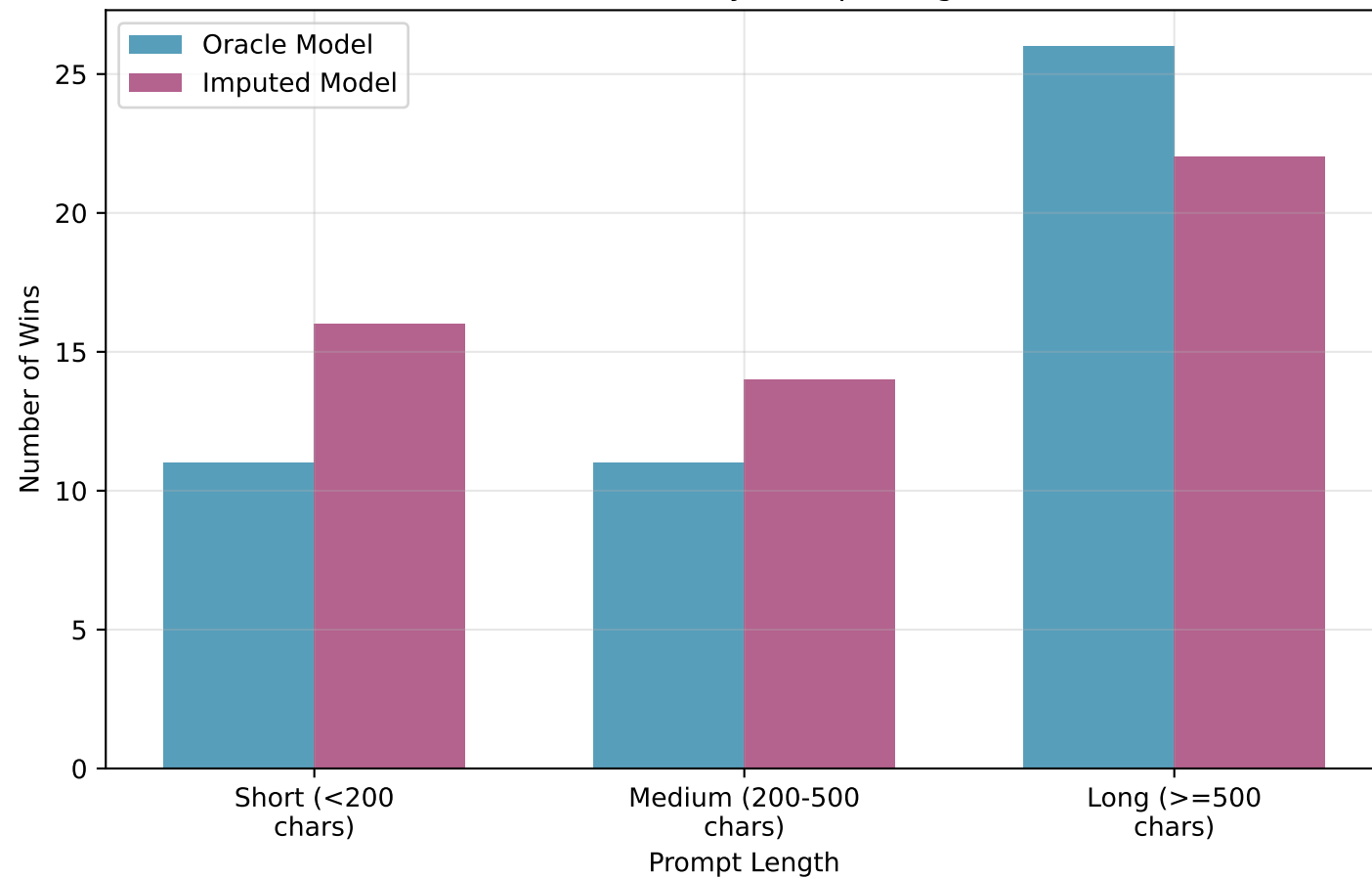
Performance by Prompt Category

Win Rate by Prompt Category

Performance by Prompt Length

Performance by Complexity Indicators