Scenario 1: Exploration Strategy Analysis

Methodology

Two distinct exploration strategies were implemented and compared:

High Exploration Strategy

- **Initial ε**: 1.0 (100% random actions initially)
- **Decay Rate**: 0.995 per episode
- **Minimum ε**: 0.01

Moderate Exploration Strategy

- **Initial ε**: 0.5 (50% random actions initially)
- **Decay Rate**: 0.99 per episode
- **Minimum ε**: 0.01

Key Findings

Learning Speed

The **moderate exploration strategy** demonstrates faster initial learning, reaching higher rewards sooner due to reduced random exploration. However, the **high exploration strategy** shows more thorough state space coverage, leading to more robust long-term performance.

Policy Quality

- **High Exploration**: Discovers more optimal paths through comprehensive exploration
- **Moderate Exploration**: Converges faster but may get trapped in suboptimal policies

Stability Across Runs

The high exploration strategy shows higher variance initially but greater stability in final performance, while moderate exploration exhibits more consistent intermediate performance but potential for suboptimal convergence.

Conclusion

The choice between exploration strategies involves a trade-off between learning speed and policy optimality. High exploration is recommended for complex environments where finding the global optimum is critical, while moderate exploration suits scenarios requiring faster convergence with acceptable suboptimality.