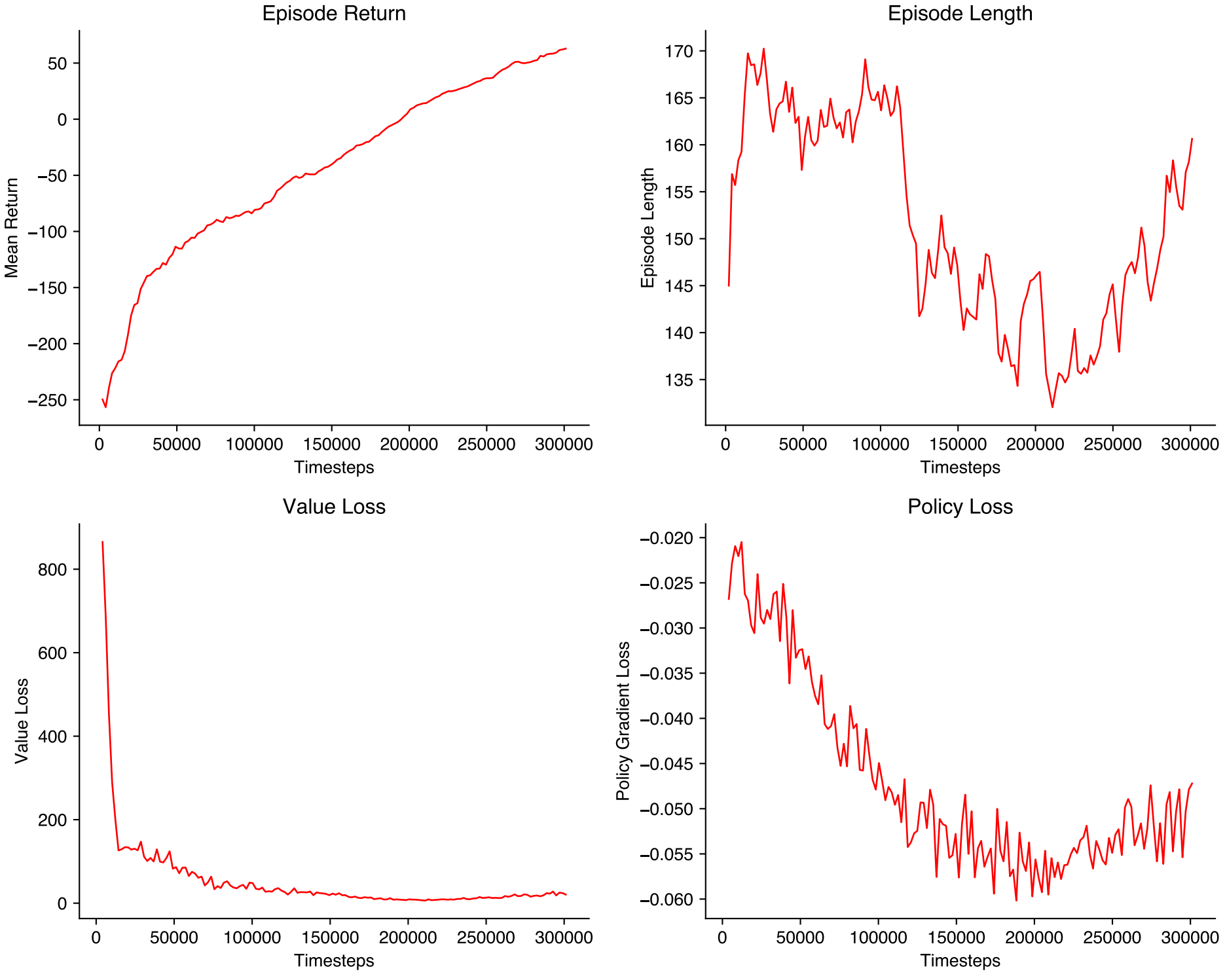
Episode Return
Mean Return
50
0
−50
−100
−150
−200
−250
0
50000
100000
150000
200000
250000
300000
Timesteps
Episode Length
Episode Length
170
165
160
155
150
145
140
135
0
50000
100000
150000
200000
250000
300000
Timesteps
Value Loss
Value Loss
800
600
400
200
0
0
50000
100000
150000
200000
250000
300000
Timesteps
Policy Loss
Policy Gradient Loss
−0.020
−0.025
−0.030
−0.035
−0.040
−0.045
−0.050
−0.055
−0.060
0
50000
100000
150000
200000
250000
300000
Timesteps