☰ ← **Back**   **Graded Quiz**
                Graded Quiz • 30 min

coursera

Due   Oct 30, 11:59 PM PDT

☰ Item Navigation

✓ **Congratulations! You passed!**

Grade received 100%

Graded Quiz
Latest Submission Grade 100%

Quiz • 30 min **To pass** 80% or higher

Go to next item

✓ **Submit your assignment**

Due   Oct 30, 11:59 PM PDT

Try again

1. Which approach ensures continual (never-ending) exploration? (**Select all that apply**)   **1 / 1 point**

✓ **Receive grade**
☑   Exploring starts

To Pass   80% or higher

⊘   **Correct**

Your grade   Correct! Exploring starts guarantee that all state-action pairs are visited an
                 infinite number of times in the limit of an infinite number of episodes.

100%

☐   On-policy learning with a **deterministic** policy

View Feedback   On-policy learning with an ϵ-soft policy
☑

We keep your highest score

⊘   **Correct**

         Correct! ϵ-soft policies assign non-zero probabilities to all state-action pairs.

👍 Like ☑ 👎 Dislike Off-Policy Report an issue behavior policy and a **deterministic** target
💡   **Congratulations!**   ✕
                         policy
         You just completed the most difficult
⊘ assignment in the course! You should feel
   Correct
proud of all of your hard work and success.
         Correct! ϵ-soft policies have non-zero probabilities for all actions in all states.
         The behavior policy is used to generate samples and should be exploratory.

     Keep learning

☐   Off-Policy learning with an ϵ-soft target policy and a **deterministic** behavior