1. If \( f(x) = 3x^2 \), what is \( \frac{df}{dx} \)?

2. If \( g(x, y, z) = 3x^2 + 2xy + y^2 \), what are:
   (a) \( \frac{\partial g}{\partial x} \)?
   (b) \( \frac{\partial g}{\partial y} \)?
   (c) \( \frac{\partial g}{\partial z} \)?

3. Let \( h(w, x) = 3w^T x \) (where \( x \) and \( w \) are column vectors of length \( d \)). What are:
   (a) \( \frac{\partial h}{\partial w_1} \)?
   (b) \( \frac{\partial h}{\partial w_d} \)?
   (c) \( \nabla h \)?

4. You are trying to optimize a state-value approximation function of the form:
   \( \hat{v}(s, w) = w_2 s^2 + w_1 s + w_0 \).

   The state loss function is:
   \( VE_s(w, s) = [v_\pi(s) - \hat{v}(s, w)]^2 \).

   The overall loss function is: \( \nabla E(w) = \sum_{s \in S} \frac{1}{|S|} \nabla E_s(w, s) \).

   The learning rate, \( \alpha \), is .01. The current value of the column vector \( w \) is \([1, 3, 3]\).

   You are given a state \( t = 2 \) for which \( v_\pi(t) = 7 \).
   (a) What is the current value of the state loss function for \( t \)?

   (b) What is the new value of \( w \) after one step of stochastic gradient descent? Show your work.
(c) What is the new value of the state loss function for $t$? Did the loss go down?

(d) Could the overall loss have gone up? Why or why not?