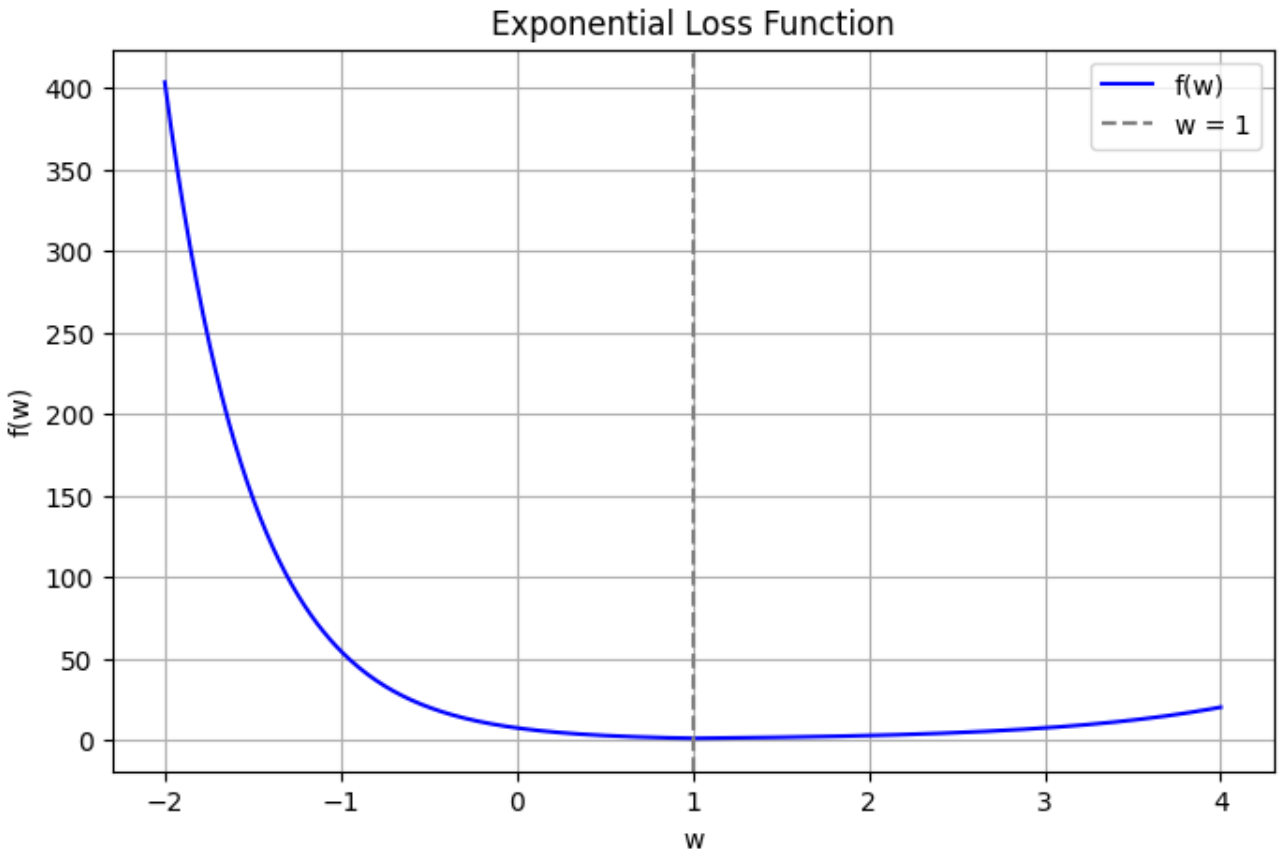
Question 1-a

In [1]:
```python
import numpy as np
import matplotlib.pyplot as plt

# Define the piecewise function
def f(w):
    return np.where(w < 1, np.exp(-2 * (w - 1)), np.exp(w - 1))

# Generate w values
w_values = np.linspace(-2, 4, 500)
f_values = f(w_values)

# Plotting
plt.figure(figsize=(8, 5))
plt.plot(w_values, f_values, label='f(w)', color='blue')
plt.axvline(x=1, color='gray', linestyle='--', label='w = 1')
plt.title("Exponential Loss Function")
plt.xlabel("w")
plt.ylabel("f(w)")
plt.legend()
plt.grid(True)
plt.show()
```



Looks convex to me

Question 1-b

To check if it differentiable lets check the derivative of both function as w = 1
$f(w) = e^{-2(w-1)}$

$$\frac{d}{dw} f(w) = \frac{d}{dw} e^{-2(w-1)} = e^{-2(w-1)} \cdot (-2) = -2e^{-2(w-1)}$$

$$f'(1) = -2e^{0} = -2$$

$f(w) = e^{w-1}$

$$\frac{d}{dw} f(w) = \frac{d}{dw} e^{w-1} = e^{w-1}$$

$$f'(1) = e^{0} = 1$$

The derivatives are not the same so its not a differentiable function

Question 1-c

$$\partial f(w) = \begin{cases} \left\{-2e^{-2(w-1)}\right\}, & \text{if } w < 1 \\ [-2, 1], & \text{if } w = 1 \\ \left\{e^{w-1}\right\}, & \text{if } w > 1 \end{cases}$$

Question 2-a

Gradient Descent

1. randomize weights
2. solve the loss function
3. take the gradient of the loss function
4. update the wieghts by taking a step in the negative direction of the gradient
5. repeat 2-3 until convergance

Question 2-b

if you label all point perfectly the gradient will become 0 and not more learning will take place

Question 3

In [ ]:
```
import numpy as np

X = np.array([
    [1, -1],
    [1, -2],
```

```
    [-1, 0],
    [-2, 1]
])
y = np.array([1, 2, -1, -2])
tau = 1
w = np.zeros(2)

def sign_subgrad(w):
    return np.sign(w)

updates = []
used_data_indices = []

for step in range(6):
    i = step % 4 # cyclical index formaula
    xi = X[i]
    yi = y[i]
    residual = yi - np.dot(xi, w)
    gradient = -2 * residual * xi + 2 * sign_subgrad(w)
    w = w - tau * gradient
    updates.append(w.copy())
    used_data_indices.append(i + 1)

# Print first two updates and data used for first 6 updates
print("First two updates:")
print("w(0):", [0,0])
print("w(1):", updates[0])
print("w(2):", updates[1])

print("\nData used for first 6 updates:")
for i, idx in enumerate(used_data_indices):
    print(f"Update {i+1}: Data point y{idx}, x{idx} = {y[idx-1]}, {X[idx-1]}
```

```
First two updates:
w(0): [0, 0]
w(1): [ 2. -2.]
w(2): [-8. 16.]

Data used for first 6 updates:
Update 1: Data point y1, x1 = 1, [ 1 -1]
Update 2: Data point y2, x2 = 2, [ 1 -2]
Update 3: Data point y3, x3 = -1, [-1  0]
Update 4: Data point y4, x4 = -2, [-2  1]
Update 5: Data point y1, x1 = 1, [ 1 -1]
Update 6: Data point y2, x2 = 2, [ 1 -2]
```