```
!pip install -U diffusers
```

Local Inference on GPU

Model page: https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0

⚠️ If the generated code snippets do not work, please open an issue on either the model repo and/or on huggingface.js 🙏

```
pip install -U diffusers transformers accelerate
```

```
import torch
from diffusers import DiffusionPipeline

# switch to "mps" for apple devices
pipe = DiffusionPipeline.from_pretrained("stabilityai/stable-diffusion-xl-base-1.0", dtype=torch.bfloat16, device_map="cuda"

prompt = "Astronaut in a jungle, cold color palette, muted colors, detailed, 8k"
image = pipe(prompt).images[0]
```

Remote Inference via Inference Providers

Ensure you have a valid **HF_TOKEN** set in your environment. You can get your token from your settings page. Note: running this may incur charges above the free tier. The following Python example shows how to run the model remotely on HF Inference Providers, automatically selecting an available inference provider for you. For more information on how to use the Inference Providers, please refer to our documentation and guides.

```
import os
os.environ['HF_TOKEN'] = 'YOUR_TOKEN_HERE'
```

```
import os
from huggingface_hub import InferenceClient

client = InferenceClient(
    provider="auto",
    api_key=os.environ["HF_TOKEN"],
)

# output is a PIL.Image object
image = client.text_to_image(
    "Astronaut riding a horse",
    model="stabilityai/stable-diffusion-xl-base-1.0",
)
```