

GPT Captioning & VQG

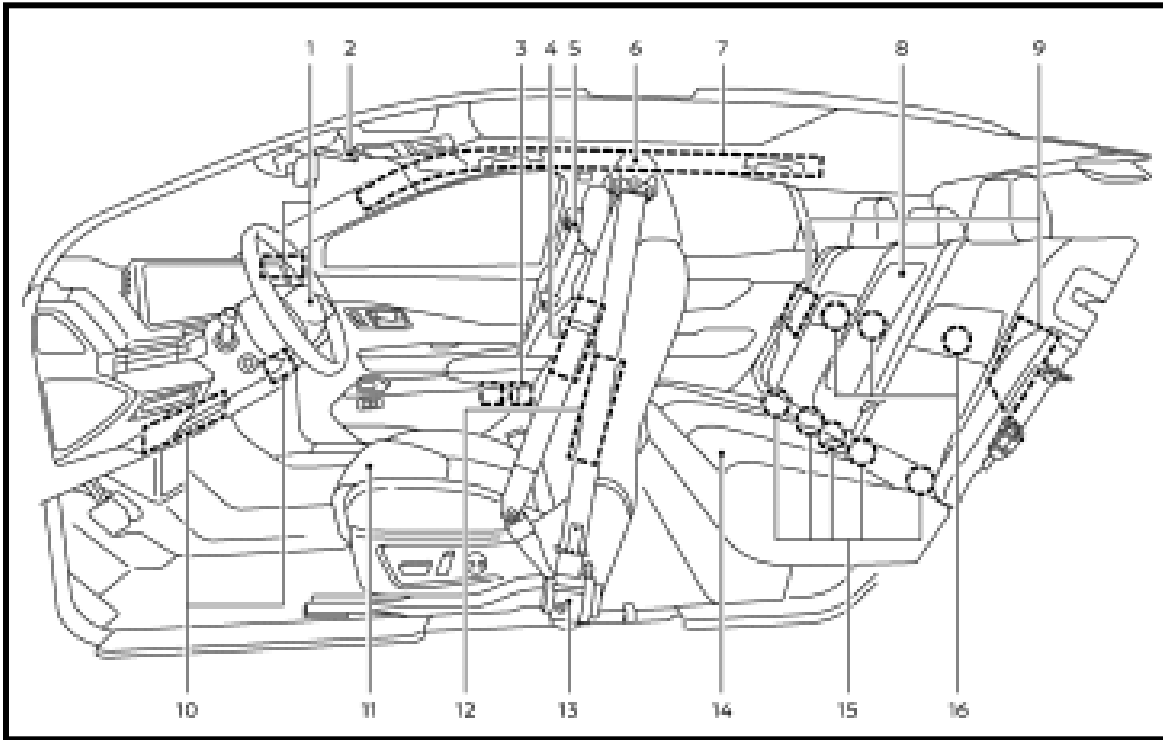
Date: 2024-02-01

Overview

The following are some examples of Image Captioning and Visual Question Generation using GPT-4. These images have been taken from the Nissan Manual (2023 Nissa Ariya Owner's Manual).

Note: GPT-4 does not support svg images as input, as of 2024-01-31. All the images below have been converted to png, and have a white background added to them to bypass this. As well, most of the icon images result in the following output: "Your input image may contain content that is not allowed by our safety system".

WBA0025X



Prompt:

What's in this Image?

Output:

This image appears to be a technical line drawing or diagram of a car cutaway. It offers a side view of the vehicle's interior, displaying various components and parts. The image is numbered, likely indicating a key or legend somewhere to describe what each number represents, which could include parts of the chassis, seating, dashboard, various mechanisms, and systems within the car. However, without the corresponding legend or any other context, I can't provide specifics on what each number corresponds to.

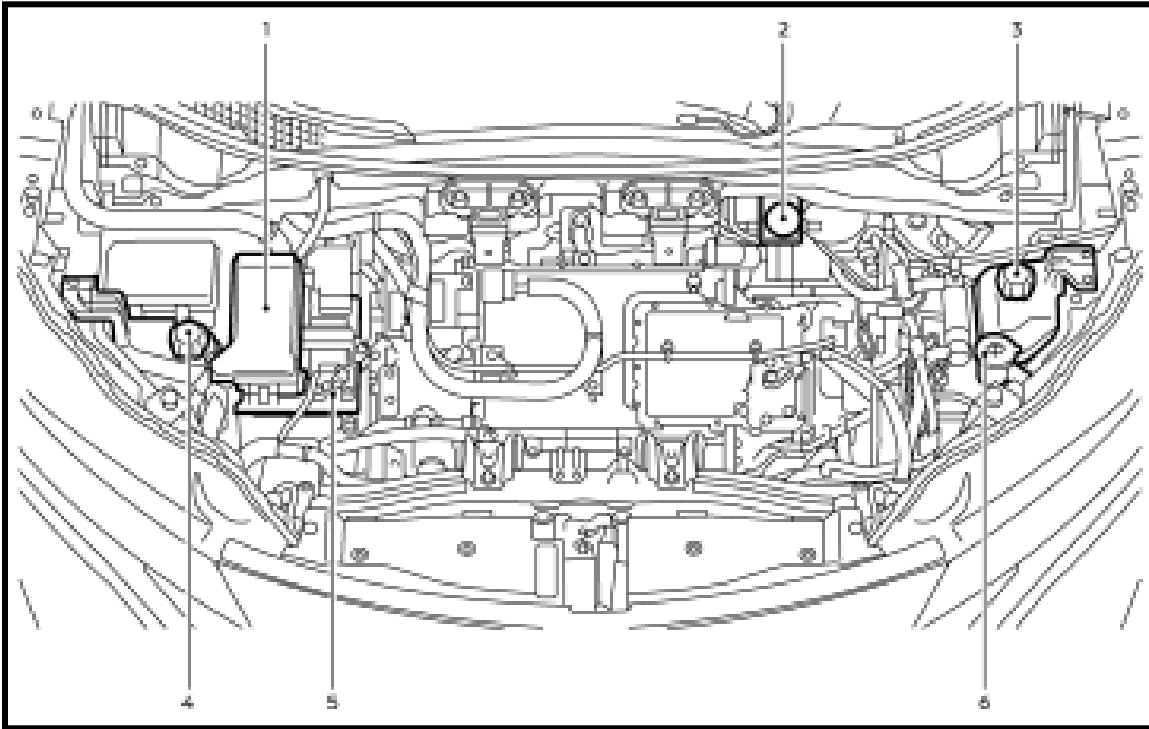
Prompt:

Generate 5 questions about this image, that can be answered by the image.

Output:

1. How many seats are visible in this cutaway diagram of a vehicle?
2. Is there a sunroof included in the diagram of the vehicle?
3. Can we see the steering wheel in this image?
4. Does the diagram indicate the location of the vehicle's dashboard?
5. How many wheels are visible in the vehicle diagram?

WBA0028X



Prompt:

What's in this Image?

Output:

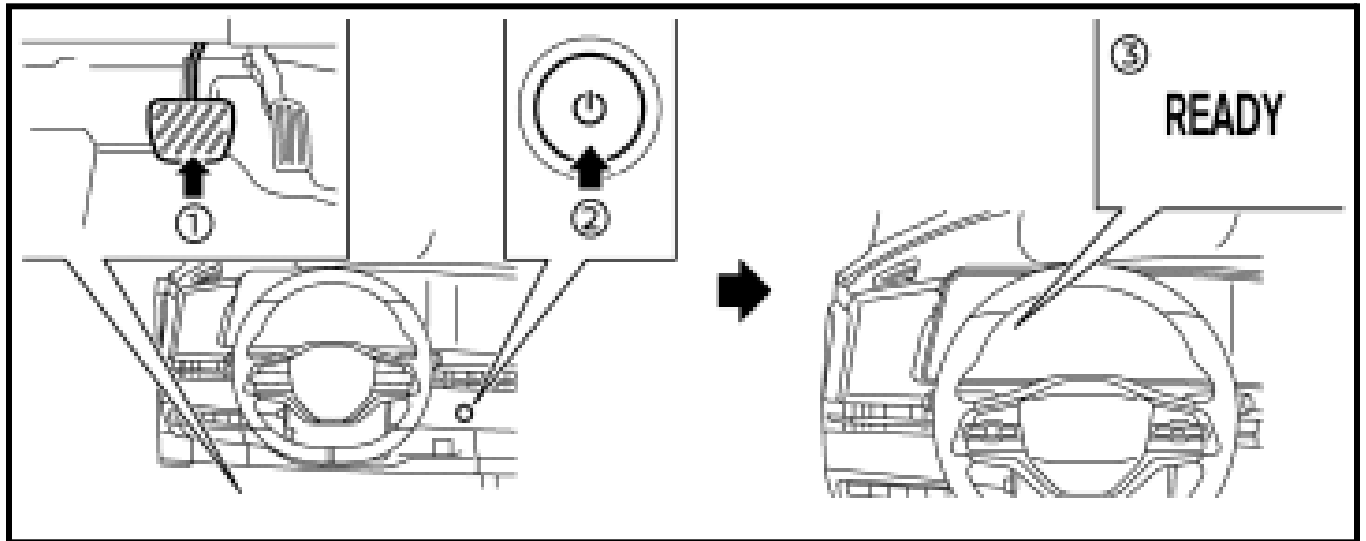
This appears to be a technical line drawing or diagram of the underside of a vehicle. It shows the layout of various components that are typically found under a car or truck, such as the exhaust system, drivetrain, suspension, and possibly elements of the chassis or underbody structure. The image is presented in a very detailed, schematic style, often used in service manuals or engineering documentation to highlight specific parts and assemblies for repair, assembly, or instructional purposes. The numbers likely correspond to a key or legend that would identify each component.

Prompt:

Generate 5 questions about this image, that can be answered by the image.

Output:

1. Is there a clear depiction of the engine block in this image?
2. Can you identify any belts or chains on the top portion of the engine layout?
3. Does the image show the location of the battery?
4. Are there any fluid reservoirs visible in the image?
5. Does the image provide a view of the underside or the top of the vehicle?

**Prompt:**

What's in this Image?

Output:

The image appears to be an instruction diagram from a vehicle's user manual, illustrating the process of starting a car with a push-button ignition system.

The first part of the diagram includes a symbol of a foot pressing a brake pedal, indicating that the driver must press the brake pedal down.

The second symbol is the engine start/stop button, suggesting that after pressing the brake pedal, the driver should press this button to start the engine.

The third part shows a dashboard with the symbol "READY," which is likely to indicate that the engine has started and the vehicle is ready to be driven.

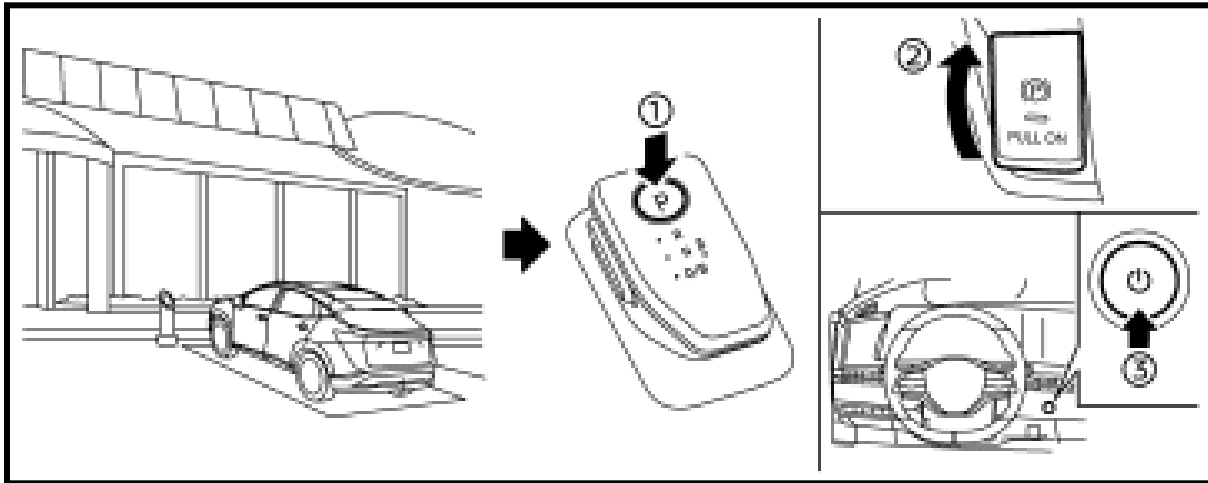
The arrow between the first two symbols and the dashboard implies a sequence of actions, leading to the vehicle being ready to operate.

Prompt:

Generate 5 questions about this image, that can be answered by the image.

Output:

1. What is the first step indicated in the starting sequence for this vehicle?
2. What symbol indicates that the vehicle is ready to drive after the starting sequence?
3. Is there a physical key involved in the starting process according to this diagram?
4. Does the diagram suggest this vehicle has a push-button start system?
5. What component is being engaged directly before the "READY" indicator is shown?

**Prompt:**

What's in this Image?

Output:

The image appears to be an instructional diagram showing a sequence of actions required to use a remote keyless system for a vehicle.

1. On the left, there's an illustration of a car parked outside a building, possibly indicating that the car is within range for the remote keyless system to work.
2. In the center, there's a drawing of the remote control being used to unlock the car. The number "1" with an arrow points to the unlock button, suggesting that you should press this button to unlock the vehicle.
3. To the right, there are two supplementary instructions. The top one shows the car key fob with a "PULL" arrow indicating that you can pull out a physical key, which is typically used in case the remote doesn't work or the electronic locking system fails. The number "2" associates this action with the second step in the process.
4. Below, there's an illustration of the car's steering wheel and the vehicle's start button marked by the number "3," possibly suggesting that once the car is unlocked, you may start the engine by pressing the start button.

This type of remote system is commonly found in modern vehicles, allowing keyless entry and ignition.

Prompt:

Generate 5 questions about this image, that can be answered by the image.

Output:

1. What is the sequence of actions depicted in the image for operating a device?
2. Where is the "PUSH" button located on the device shown in the second part of the sequence?
3. What symbol is displayed on the button that must be pressed in the third part of the sequence?
4. Is the device shown intended to be used inside or outside of a car?
5. What does the first image suggest about the location where this device is used?