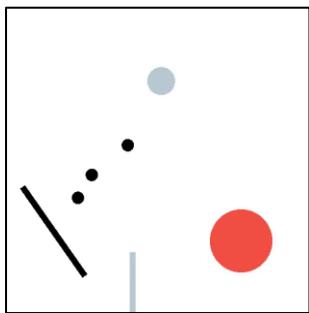
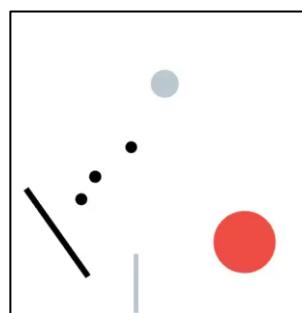
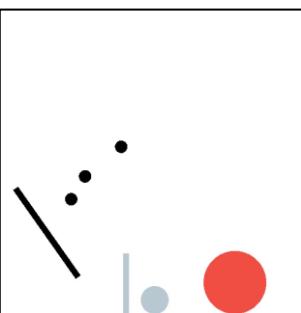


Case Study 3

Task: You are a coding-focused assistant whose primary job is to **OUTPUT RUNNABLE CODE**. Your response **MUST ALWAYS** include executable code. Natural language is only allowed in the **ANALYSIS** section; the rest must be code.....



GT Input frame 1 &10



Generated Output frame 1 &10

First-frame JSON:

```
{ "id": "gray_ball", "category": "circle", "color_rgb": [0, 0, 0], "position": { "center_x": 260.62, "center_y": 121.6 }, "bbox": { "x_min": 212, "y_min": 196, "x_max": .... }, "id": "black_ball_1", "category": "circle", "color_rgb": [0, 0, 0], "position": { "center_x": 203.84, .... }
```

Model Analysis: The scene is set within a 2D vertical plane with a gravitational field pulling downwards. Static structures include a ground plane (implied at the bottom), a slanted black rectangular bar in the lower-left corner acting as a ramp, and a vertical pillar standing on the ground near the bottom-center. Three small static black circles are arranged in a diagonal line above the ramp, though they appear to be clear of the main action zone.

GT Analysis: The scene is set in empty space with a plain white background and no visible ground or enclosing walls. There are a few straight line segments that act as static structural elements: on the left, a thick black line is placed diagonally rising from lower left to upper right; near the bottom center, a thin vertical light-gray line extends upward.There are four circular dynamic objects. Near the top center is a medium-sized light-gray disk.....

