1. Consider a 3D transformation that does the following (in the order given):

   - translate by (3, 2, 1)
   - rotate by 90 degrees about the Z axis (bringing the X axis to the Y axis)
   - scale the X axis by a factor of 2
   - translate by (1, 2, 3)

   What are the elements of the first row of the transformation matrix?

   (a) (-3, 5, 4, 1)
   (b) (0 -2, 0, -3) **answer**
   (c) (0 -1, 0, -1)
   (d) (0 -3, 2, -1)

   Solution: Build four matrices $\mathbf{T_{1,2,3}S_2R_{Z,90}}, \mathbf{T_{3,2,1}}$ and multiply them together.

2. Consider a camera projection plane that is at a distance of 50 mm from the camera origin. Let the projection plane be sampled at 100 pixels per mm, and let the pixel grid be 4000 x 4000 and be centered on the camera Z axis.

   Consider a 3D point (X,Y,Z) = (2, 1, 5) in camera coordinates. What are the approximate pixel indices (x, y) of that point.

   (a) (4000, 3000) **answer**
   (b) (0, -1000)
   (c) (1000, 500)
   (d) (4000, 2000)
   (e) (2000, -1000)

   Solution: In the question, I neglected to write the units of the scene point XYZ. (I meant meters, not mm.) Fortunately the units for XYZ don't matter because the projection is the same whether the units *of the scene point* are mm or m.

   Here are the matrices that describe the scenario. The pixel pixel of the principal point (Z axis) is approximately (2000, 2000), namely there are 2000 pixels to the left (or above) and 2000 pixels to the right (or below).

$$\begin{bmatrix} 100 & 0 & 2000 \\ 0 & 100 & 2000 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 50 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 5 \\ 1 \end{bmatrix}$$

   Multiplying out gives the answer $(20000, 15000, 5)$ which is in homogeneous coordinates. To get pixel values, divide by the 5 to get $(4000, 3000, 1)$.

3. Let A be an 8×4 matrix whose columns are linearly dependent. Suppose none of the matrix entries of A are 0. What can you say about the singular values of A? Select all that apply.

   (a) There are no negative singular values. **answer**

   (b) The number of singular values that are 0 is exactly 1.

   (c) The number of singular values that are 0 is at most 1.

   (d) The number of singular values that are 0 is least 1. **answer**

   (e) The number of non-zero singular values is at most 4. **answer, but see below.**

   Solution: The number of non-zero singular values can be 1, 2, or 3.

   (a) is correct because singular values by definition are non-negative. (d) is correct because the columns are linearly *dependent*, which means that they span a space of dimension at most 3, so the number of non-zero singular values is at most 3, and the number of singular values that are 0 is at least 1.

   A few students pointed out to me that the wording of (e) is problematic since "at most 4" means less than or equal to 4, which is correct. However, "at most 4" also suggests that 4 is a possibility, which is incorrect in the case described (and contradicts (d)). In hindsight, it would have been better to drop (e) entirely! *Anyhow, I modified the grading scheme, such that anyone who chooses (a,d) only should get 1 point.*

   I set the grading scheme to "correct-incorrect".

4. To perform camera calibration, one is given a set of XYZ and xy values and one typically normalizes these data. Which of the following are true about data normalization? Select all that apply.

   (a) It shifts the camera position **C** to the mean of the XYZ values. ~~**answer**~~ **See below.**

   (b) It shifts the xy values so that their mean is centered approximately at the top left corner of the image. **answer**

   (c) It rescales the XYZ data so that the camera and world coordinates are aligned.

   (d) It removes the units of the data. **answer**

   Solution: [Updated:] (a) is incorrect. It shifts the origin of the world coordinate system (not camera position) to the mean XYZ position.

   (b) is correct because the top left corner of the image is approximately the origin (0,0) of the pixel space, and this is where the mean of the xy values are shifted to.

   (c) is incorrect. The statement makes no sense. "Rescaling" is a multiplicative effect. This has nothing to do with alignment of coordinate axes. Moreover, "alignment" of coordinate axes suggest that point in the same direction, but this would require a rotation and normalization does not perform a rotation.

   (d) is correct because scene units (m or mm) disappear when we normalize, and image pixel units disappear when we normalize.

   I set the grading scheme to "correct-incorrect".

5. Let $\mathbf{X}_0 = (-2, 1, 8)$ be the position of a 3D point in the camera coordinate system. Let the camera's projection plane be at depth $Z = 5$.

   Suppose the camera rotates 45 degrees about the Y axis. (Rotation by 90 degrees takes Z to X.)

   What is the approximate image y value of the 3D point? Round off your value to one decimal point.

   (a) 0.7 **answer**
   (b) 1.4
   (c) 0.6
   (d) 1.6
   (e) 0.5

   Solution:

$$\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix}$$

   We have $\theta = 45deg$, so

$$y = \frac{fY}{\frac{-X}{\sqrt{2}} + \frac{Z}{\sqrt{2}}} = \frac{fY\sqrt{2}}{-X + Z}$$

   and plugging for f, XYZ gives you 0.7.