1. What is the *orthogonal projection* of y on w?
Answer: It is the scalar multiple of w that is as close as possible to y. In other words, it is the element of SPAN(w) that is closest to y.

   So what does this mean? Well, draw the line through w and the origin, lets call that line $W = \text{SPAN}(w)$. If y is on that line, then the orthogonal projection of y on w is y itself. If y is not on that line, then pick the point on that line that is as close as possible to y, and then that point is the orthogonal projection of y on w
(same as: orthogonal projection of y on W).

2. How do you compute the orthogonal projection of vector y on w?
Answer: Compute these two numbers: $y \cdot w$ and $w \cdot w$. Then take the quotient. Multiply that by w and you get the orthogonal projection of y on w:
$$\text{proj}_w(y) \quad = \quad \frac{y \cdot w}{w \cdot w} w$$
Since this is a scalar (the quotient of those two dot-products) times w, we see that the projection of y on w is always on the line $W = \text{SPAN}(w)$

3. Let W be some subspace of $\mathbf{R}^n$ and let y be some element of $\mathbf{R}^n$. What is the *orthogonal projection* of y on W?
Answer: It is the element of W that is as close as possible to y. So if y is in W then the projection of y on W is just y itself. If y is not in W, then pick the point in W that is the closest to y, and then that point is the orthogonal projection of y on W.

4. How do you compute the orthogonal projection of vector y on W?
Answer: First you need an *orthogonal basis* of W. Suppose that $w_1, \ldots, w_k$ is an orthogonal basis of W (how to find an orthogonal basis of W is the subject of items 9,10). Then
$$\text{proj}_W(y) = \text{proj}_{w_1}(y) + \text{proj}_{w_2}(y) + \cdots + \text{proj}_{w_k}(y)$$
in other words, the projection of y on W is
$$\text{proj}_W(y) \;=\; \frac{y \cdot w_1}{w_1 \cdot w_1} w_1 \quad + \quad \frac{y \cdot w_2}{w_2 \cdot w_2} w_2 \quad + \cdots + \quad \frac{y \cdot w_k}{w_k \cdot w_k} w_k$$
This only works if $w_1, \ldots, w_k$ is an orthogonal basis of W.

5. What does $u \perp v$ mean?
Answer: $u \perp v$ means that u is orthogonal to v, which in turn means that the dot-product (the inner product) of u and v is zero, so $u \cdot v = 0$.
This happens when $u = 0$, or when $v = 0$, or when u, v are perpendicular (the angle between them is 90^o).

6. What's an orthogonal set?
   Answer: It's a set where every element is orthogonal to every other element.
   How do I check if $\{w_1, w_2, \ldots, w_k\}$ is an orthogonal set?
   Answer: You check that each of them is orthogonal to all the previous ones, so you check that $w_2 \cdot w_1 = 0$, then check that $w_3 \cdot w_1 = 0$ and $w_3 \cdot w_2 = 0$, then check that $w_4 \cdot w_1 = 0$, $w_4 \cdot w_2 = 0$, $w_4 \cdot w_3 = 0$, etc.

7. What's an orthogonal basis of a vector space W?
   Answer: a basis where every element is orthogonal to every other element.

8. If $w_1, \ldots, w_k$ are some vectors, what's the quickest way to see if they form an orthogonal basis of W?
   Answer: First of all, they must all be in W. Second, the zero-vector must not be among $w_1, \ldots, w_k$. Furthermore, k, the number of vectors in your set, must be equal to the dimension of V. Finally, check that they form an orthogonal set (see item 6).

   Don't I have to check that $w_1, \ldots, w_k$ are linearly independent to make sure that I have a basis of W?
   Answer: *an orthogonal set without zero-vectors* is automatically linearly independent.

9. How do I get an *orthogonal basis* of W?
   Answer: first, you need a basis (or a spanning set, that's OK too) for W. Say that $u_1, \ldots, u_k$ is a spanning set of W. Now you follow the following process, called the Gram-Schmidt process:

   Take $v_1 = u_1$.
   Take v_2 to be u_2 MINUS the projection of u_2 on all previous v's.
   Take v_3 to be u_3 MINUS the projection of u_3 on all previous v's.
   Take v_4 to be u_4 MINUS the projection of u_4 on all previous v's.
   etc.
   If any of these v's are zero, then just throw that one away (this only happens if the u's were linearly dependent).
   The remaining v's (the non-zero v's) will be an orthogonal basis of W.

10. Can you spell that out in some more detail, how to get an orthogonal basis of W if I have some spanning set $u_1, \ldots, u_k$ of W?
    Answer: Follow the previous item, and just plug in the these orthogonal projections. So you get:
    $v_1 = u_1$
    $v_2 = u_2 - \text{proj}_{v_1}(u_2)$
    $v_3 = u_3 - \text{proj}_{v_1,v_2}(u_3)$
    $v_4 = u_4 - \text{proj}_{v_1,v_2,v_3}(u_4)$, etc.
    If we spell this out with the formula for the orthogonal projection (see

items 2 and 4) then we get:
$v_1 = u_1$
$v_2 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1$
$v_3 = u_3 - \left(\frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 \right)$
$v_4 = u_4 - \left(\frac{u_4 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{u_4 \cdot v_2}{v_2 \cdot v_2} v_2 + \frac{u_4 \cdot v_3}{v_3 \cdot v_3} v_3 \right)$, etc.
In step 3, make sure that you use u_3 and the previous v's (not the previous u's). In step 4, use u_4 and the previous v's (not the previous u's).

11. Example, let $u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, u_3 = \begin{pmatrix} 0 \\ 1 \\ 4 \\ 9 \\ 16 \end{pmatrix}$ and $y = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 3 \end{pmatrix}$.
Let $W = \text{SPAN}(u_1, u_2, u_3)$. Find the orthogonal projection of y on W, i.e. find the vector in W that is as close as possible to y.
Answer: if u_1, u_2, u_3 were an orthogonal set, we could use the formula in item 4 (the w's in item 4 would then be the u's here). But, u_1, u_2, u_3 are not orthogonal, for example $u_1 \cdot u_2 \neq 0$. We'll have to fix that with Gram-Schmidt. We take:
$v_1 = u_1$
$v_2 = u_2 - \frac{0\cdot1+1\cdot1+2\cdot1+3\cdot1+4\cdot1}{1^2+1^2+1^2+1^2+1^2} u_1 = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$
$v_3 = u_3 - \left(\frac{0\cdot1+1\cdot1+4\cdot1+9\cdot1+16\cdot1}{1^2+1^2+1^2+1^2+1^2} u_1 + \frac{(-2)\cdot0+(-1)\cdot1+0\cdot4+1\cdot9+2\cdot16}{(-2)^2+(-1)^2+0^2+1^2+2^2} u_2 \right) = \begin{pmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{pmatrix}$
Now that we have an *orthogonal basis* v_1, v_2, v_3 of the vector space W, we are ready to compute the orthogonal projection of y on W with the formula from item 4 (the w's in item 4 are the v's here).
$\text{proj}_W(y) = \frac{5}{5} v_1 + \frac{5}{10} v_2 + \frac{7}{14} v_3$. If we compute that, we get y itself (this means that y was actually in W, so the vector in W closest to y is then of course y itself). Let's compute $\text{proj}_W(u)$ for another vector, say
$u = \begin{pmatrix} -2 \\ 0 \\ 3 \\ 2 \\ 2 \end{pmatrix}$. Then $\text{proj}_W(u) = \frac{5}{5} v_1 + \frac{10}{10} v_2 + \frac{-8}{14} v_3 = \begin{pmatrix} -15/7 \\ 4/7 \\ 15/7 \\ 18/7 \\ 13/7 \end{pmatrix}$.
Application: if $f(x)$ is a function that takes values $-2, 0, 3, 2, 2$ (the entries of u) at $x = 0, 1, 2, 3, 4$ then the quadratic function that best approximates this takes has values "the entries of $\text{proj}_W(u)$" at $x = 0, 1, 2, 3, 4$.