The objective here is to prove that, over a principal ideal domain, every submodule of a free is also a free module and that the rank of a free submodule is always at least as large of the ranks of its submodules.

So let $\mathbf{R}$ be a (nontrivial) principal ideal domain. We know that $\mathbf{R}$ is a free $\mathbf{R}$-module of rank 1. What about the submodules of $\mathbf{R}$? Suppose $\mathbf{E}$ is such a submodule. It is clear that E is an ideal and, in fact, that the ideals of $\mathbf{R}$ coincide with the submodules of $\mathbf{R}$. In case $\mathbf{E}$ is trivial (that is the sole element of E is 0) we see that $\mathbf{E}$ is the free $\mathbf{R}$-module of rank 0. So consider the case that $\mathbf{E}$ is nontrivial. Since $\mathbf{R}$ is a principal ideal domain we pick $w \neq 0$ so that $\mathbf{E}$ is generated by w. That is $E = \{rw \mid r \in R\}$. Since we know that $\mathbf{R}$ has $\{1\}$ as a basis, we see that the map that sends 1 to w extends to a unique module homomorphism from $\mathbf{R}$ onto $\mathbf{E}$. Indeed, notice $h(r \cdot 1) = r \cdot h(1) = rw$ for all $r \in R$. But the homomorphism h is also one-to-one since

$$\begin{aligned} h(r) &= h(s) \\ rh(1) &= sh(1) \\ rw &= sw \\ r &= s \end{aligned}$$

where the last step follows because integral domains satisfy the cancellation law and $w \neq 0$. In this way we see that $\mathbf{E}$ is isomorphic to the free $\mathbf{R}$-module of rank 1. We also see that $\{w\}$ is a basis for $\mathbf{E}$.

So we find that at least all the submodules of the free $\mathbf{R}$-module of rank 1 are themselves free and have either rank 0 or rank 1. We can also see where the fact that $\mathbf{R}$ is a principal ideal domain came into play.

The Freedom Theorem for Modules over a PID.
Let $\mathbf{R}$ be a principal ideal domain, let $\mathbf{F}$ be a free $\mathbf{R}$-module and let $\mathbf{E}$ be a submodule of $\mathbf{F}$. Then $\mathbf{E}$ is a free $\mathbf{R}$-module and the rank of $\mathbf{E}$ is no greater than the rank of $\mathbf{F}$.

Proof. Since trivial modules (those whose only element is 0) are free modules of rank 0, we suppose below that $\mathbf{E}$ is a nontrivial module. This entails that $\mathbf{F}$ is also nontrivial.

Let B be a basis for $\mathbf{F}$ and $C \subseteq B$. Because $\mathbf{F}$ is not the trivial module, we see that B is not empty. Let $\mathbf{F}_C$ be the submodule of $\mathbf{F}$ generated by C. Let $\mathbf{E}_C = \mathbf{E} \cap \mathbf{F}_C$. Evidently, C is a basis for $\mathbf{F}_C$. To see that $\mathbf{E}_C$ is free we will have to find a basis for it.

Suppose, for a moment, that C has been chosen so that $\mathbf{E}_C$ is known to be free and that $w \in B$ with $w \notin C$. Put $D := C \cup \{w\}$. Consider the map defined on D into R that sends all the elements of C to 0 and that sends w to 1. This map extends uniquely to a homomorphism φ from $\mathbf{F}_D$ onto $\mathbf{R}$ and it is easy to check (as hardworking graduate student will) that the kernel of φ is just F_C. By the Homomorphism Theorem, we draw the conclusion that $\mathbf{F}_D/F_C$ is isomorphic to $\mathbf{R}$ and that it is free of rank 1. What about $\mathbf{E}_D/E_C$? Observe that $E_C = E \cap F_C = E \cap F_D \cap F_C = E_D \cap F_C$. So we can apply the Second Isomorphism Theorem:

$$\mathbf{E}_D/E_C = \mathbf{E}_D/E_D \cap F_C \cong \mathbf{E}_D + \mathbf{F}_C/F_C.$$

But $\mathbf{E}_D + \mathbf{F}_C/F_C$ is a submodule of $\mathbf{F}_D/F_C$. This last is a free $\mathbf{R}$-module of rank 1. We saw above that every submodule of a free $\mathbf{R}$-module of rank 1 must be itself a free $\mathbf{R}$-module and have rank either 0 or 1. In this way, we find that either $\mathbf{E}_D = \mathbf{E}_C$ (in the rank 0 case) or else $\mathbf{E}_D/E_C$ is a free $\mathbf{R}$-module of rank 1. Let us take up this latter case. Let X be a basis for $\mathbf{E}_C$, which we assumed, for the moment, was free. Pick $u \in E_D$ so that $\{u/E_C\}$ is a basis for $\mathbf{E}_D/E_C$.

We contend that $X \cup \{u\}$ is a basis for $\mathbf{E}_D$. To establish linear independence, suppose $x_0, \dots, x_{n-1}$ are distinct element of X, that $r_0, \dots, r_n \in R$ and that

$$0 = r_0 x_0 + \cdots + r_{n-1}x_{n-1} + r_n u.$$

First notice that

$$r_n(u/E_C) = r_n u/E_C = (r_0x_0 + \cdots + r_{n-1}x_{n-1} + r_n u)/E_C = 0/E_C.$$

Since $\{u/E_C\}$ is a basis for $\mathbf{E}_D/E_C$, we must have $r_n = 0$. This leads to

$$0 = r_0x_0 + \cdots + r_{n-1}x_{n-1}.$$

But now since X is a basis for $\mathbf{E}_C$ we see that $0 = r_0 = \cdots = r_{n-1}$. So we find that $X \cup \{u\}$ is linearly independent.

To see that $X \cup \{u\}$ generates E_D, pick $z \in E_D$. Since $\{u/E_C\}$ is a basis for $\mathbf{E}_D/E_C$, pick $r \in R$ so that

$$z/E_C = ru/E_C.$$

This means that $z - ru \in E_C$. But X is a basis of $\mathbf{E}_C$. So pick $x_0, \dots, x_{n-1} \in X$ and $r_0, \dots, r_{n-1} \in R$ so that

$$z - ru = r_0x_0 + \cdots + r_{n-1}x_{n-1}.$$

Surely this is enough to see that z is in the submodule generated by $X \cup \{u\}$. So this set generates $\mathbf{E}_D$ and we conclude that it must be a basis of $\mathbf{E}_D$.

In this way we see that for $C \subseteq D \subseteq B$ where D arises from adding an element to C, if E_C is free, then so is $\mathbf{E}_D$ and that either $E_D = E_C$ or a basis for $\mathbf{E}_D$ can be produced by adding just one element to a basis for $\mathbf{E}_C$.

With this in mind, we can envision a procedure for showing that $\mathbf{E}$ is free and its rank cannot be larger than that of $\mathbf{F}$. Notice that $E = E \cap F = E \cap F_B$. So $\mathbf{E} = \mathbf{E}_B$. The idea is simple. We will start with $\varnothing \subseteq B$. We observe that $\mathbf{F}_\varnothing = \mathbf{E}_\varnothing$ is the module whose sole element is 0. It is free of rank 0. Next we select an element $w \in B$ and form $\varnothing \cup \{w\} = \{w\}$. We find that $\mathbf{E}_{\{w\}}$ is free of rank 0 or rank 1. We select another element and another and another... until finally all the elements of B have been selected. At this point we would have E_B is free and its rank can be no more than the total number of elements we selected, namely $|B|$ which is the rank of $\mathbf{F}$.

To carry out this program, in case B were finite or even countable, we could mount a proof by induction. You can probably see how it might be done. But we want to prove this for arbitary sets B. We could still pursue this inductive strategy openly by well-ordering B and using transfinite induction. By using the well-ordering we would always know what was meant by "pick the next element of B."

Instead, we will invoke Zorn's Lemma to short-circuit this rather long induction.

Let $\mathcal{F} = \{f \mid f \text{ is a function with } \operatorname{dom} f \subseteq B \text{ and } \operatorname{range} f \text{ a basis for } \mathbf{E}_{\operatorname{dom} f}\}$. Recalling that functions are certain kinds of sets of order pairs, we see that $\mathcal{F}$ is paritally ordered by set inclusion. Maybe it helps to realize that to assert $f \subseteq g$ is the same as asserting that g extends f. We note that $\mathcal{F}$ is not empty since the empty function (the function with empty domain) is a member of $\mathcal{F}$. To invoke Zorn's Lemma, let $\mathcal{C}$ be any chain included in $\mathcal{F}$. Let $h = \bigcup \mathcal{C}$. Evidently $f \subseteq h$ for all $f \in \mathcal{C}$. So h is an upper bound of $\mathcal{C}$. We contend that $h \in \mathcal{F}$. We ask the hard-working graduate students to check that the union of any chain of functions is itself a function. Once you do that bit of work, it should be evident that $\operatorname{dom} h = \bigcup\{\operatorname{dom} f \mid f \in \mathcal{C}\}$ and that $\operatorname{range} h = \bigcup\{\operatorname{range} f \mid f \in \mathcal{C}\}$. So it remains to show that $\operatorname{range} h$ is a basis for $E_{\operatorname{dom} h}$. To see that $\operatorname{range} h$ is a generating set, let z be an arbitrary element of $E_{\operatorname{dom} h} = E \cap F_{\operatorname{dom} h}$. Hence z must be generated by some finitely many elements belong in $\operatorname{dom} h$. This means there are finitely many functions $f_0, \dots, f_{n-1} \in \mathcal{C}$ so that z is generated by finitely many elements of $\operatorname{dom} f_0 \cup \cdots \cup \operatorname{dom} f_{n-1}$. But $\operatorname{dom} f_0, \dots, \operatorname{dom} f_{n-1}$, rearranged in some order, forms a chain under inclusion. So $z \in F_{\operatorname{dom} f_\ell}$ for some $\ell < n$. Hence

$z \in E_{\mathrm{dom}\, f_\ell}$. But range f_ℓ is a basis for $\mathbf{E}_{\mathrm{dom}\, f_\ell}$. Because range $f_\ell \subseteq$ range h we find that range h has enough elements to generate z. Since z was an arbitrary element of $E_{\mathrm{dom}\, h}$ we conclude that range h generates $E_{\mathrm{dom}\, h}$. It remains to show that range h is linearly independent. But range h is the union of the chain $\{\mathrm{range}\, f \mid f \in \mathcal{C}\}$. I ask the hard-working graduate students to prove that the union of any chain of linearly independent sets must also be linearly independent. Once you have done this you will be certain that h belongs to $\mathcal{F}$. By Zorn, let g be a maximal element of $\mathcal{F}$.

We would be done if $\mathrm{dom}\, g = B$, since then $E = E \cap F = E \cap F_B = E_B = E_{\mathrm{dom}\, g}$. In which case, range g would be a basis for $\mathbf{E}$ and rank $\mathbf{E} = |\,\mathrm{range}\, g| \leq |\,\mathrm{dom}\, g| = |B| = \mathrm{rank}\, \mathbf{F}$.

Consider the possibility that $\mathrm{dom}\, g$ is a proper subset of B. Put $C = \mathrm{dom}\, g$ and put $X = \mathrm{range}\, g$. Let $w \in B$ with $w \notin \mathrm{dom}\, g$. Put $D = C \cup \{w\}$. As we have seen above, either $E_D = E_C$ or $X \cup \{u\}$ is a basis for $\mathbf{E}_D$, for some appropriately chosen u. We can now extend g to a function g' by letting $g'(w)$ be any element of range g in the case when $E_D = E_C$ and by letting $g'(w) = u$ is the alternative case. In this way, $g' \in \mathcal{F}$, contradicting the maximality of g. So we reject this possibility.

This completes the proof. □

Corollary 0. *Let* $\mathbf{R}$ *be a principal ideal domain. Every submodule of a finitely generated* $\mathbf{R}$*-module must itself be finitely generated.*

Proof. Suppose $\mathbf{M}$ is an $\mathbf{R}$-module generated by n elements. Let $\mathbf{N}$ by a submodule of $\mathbf{M}$.

Now let $\mathbf{F}$ be the free $\mathbf{R}$-module with a basis of n elements. There is a function that matches this basis with the generating set of $\mathbf{M}$. So, appealing to freeness, there is a homomorphism h from $\mathbf{F}$ onto $\mathbf{M}$. Let $E = \{v \mid v \in F \text{ and } h(v) \in N\}$. It is straightforward to check (will you do it?) that E is closed under the module operations. So we get a submodule $\mathbf{E}$ of $\mathbf{F}$. Moreover, the restriction of h to E is a homomorphism from $\mathbf{E}$ onto $\mathbf{N}$. But by our theorem $\mathbf{E}$ is generated by a set with no more than n elements. Since the image, under a homomorphism, of any generating set for $\mathbf{E}$ must be a generating set of $\mathbf{N}$ (can you prove this?), we find that $\mathbf{N}$ is finitely generated. □