Theorem. If A is row equivalent to the identity matrix then A is a product of elementary matrices.

Proof. If $A \xrightarrow{\xi_1} B_1 \xrightarrow{\xi_2} B_2 \rightarrow \cdots \xrightarrow{\xi_r} B_r$ and
and $I \xrightarrow{\xi_i} E_i$ then $E_r \cdots E_2 E_1 A = B_r$.
If $B_r = I$ then $A^{-1} = E_r \cdots E_2 E_1$ and
$A = (E_r \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_r^{-1} = F_1 F_2 \cdots F_r$.
where F_i is the elementary matrix corresponding to the elementary operation inverse to ξ_i.

Corollary. A matrix is invertible $\Leftrightarrow$ it is a product of elementary matrices.

Example. $\begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

$$\Rightarrow \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Let A be an $n \times n$ matrix. If A is invertible then the system of equations $AX = Y$ is uniquely solvable for any $Y \in \mathbb{R}^{n\times 1}$. Since $AX = x_1 A_1 + \cdots + x_n A_n$ where $A_1, \dots, A_n$ are the columns of A and $X^T = [x_1, \dots, x_n]$ we see that A invertible $\Rightarrow$ every $Y \in \mathbb{R}^{n\times 1}$ can be uniquely written as a linear combinations of the columns of A. Taking $Y = 0$ we see that the columns of an invertible matrix are linearly independent as well as spanning $\mathbb{R}^{n\times 1}$.

Definition. The vectors $u_1, \dots, u_m$ in a subspace S of $\mathbb{R}^n$ is said to be a basis of S if

(1) $S = \text{span}(u_1, \dots, u_m)$

(2) $u_1, \dots, u_m$ are linearly independent

Exercise. Show that $u_1, \dots, u_m$ is a basis for $S \Leftrightarrow$ every $v \in S$ can be uniquely written in the form $x_1 u_1 + \cdots + x_m u_m$

Thus A invertible $\Rightarrow$ columns of A are a basis for $\mathbb{R}^{n \times 1}$.

Exercise. Show that A invertible $\Rightarrow$ rows of A are a basis for $\mathbb{R}^{1\times n}$.

Conversely, if the columns of A are a basis for $\mathbb{R}^{n\times 1}$ then the system of equations $AX = Y$ has a unique solution for any Y. In particular, taking $Y = 0$, we get $AX = 0 \Rightarrow X = 0$ which implies that A is row equivalent to I and hence is invertible.

Exercise. Show that a square matrix A is invertible if any one of the following holds

(1) the rows of A are a basis for $\mathbb{R}^{1\times n}$
(2) the rows of A span $\mathbb{R}^{1\times n}$
(3) the columns of A span $\mathbb{R}^{n\times 1}$

Theorem. Let S be a subspace of $\mathbb{R}^n$. Then
(1) S has a basis
(2) Any two bases for S have the same number of vectors. This number is called the dimension of S and is denoted by $\dim(S)$.

Proof. We prove (2) first. The proof needs the following auxilliary result

Lemma. If $v_1, \dots, v_p \in \operatorname{span}(u_1, \dots, u_m)$ and $p > m$ then $v_1, \dots, v_p$ are linearly dependent.

Proof. Let $v_i = a_{1i} u_1 + a_{2i} u_2 + \dots + a_{mi} u_m$

Then

$$x_1 v_1 + \dots + x_p v_p = (a_{11} x_1 + \dots + a_{1p} x_p) u_1 + \dots$$
$$+ (a_{m1} x_1 + \dots + a_{mp} x_p) u_m$$

Hence $x_1 v_1 + \dots + x_p v_p = 0$ if

$$a_{11} x_1 + \dots + a_{1p} x_p = 0$$
$$\vdots$$
$$a_{m1} x_1 + \dots + a_{mp} x_p = 0$$

But this system has a non-zero solution since there are more unknowns than equations.

QED (Lemma)

Now let's prove (2). If $u_1, \dots, u_m$ and $v_1, \dots, v_p$ are two bases for S then

$u_1, \dots, u_m$ lin. indep & $u_1, \dots, u_m \in \operatorname{span}(v_1, \dots, v_p) \Rightarrow m \leq p$

$v_1, \dots, v_p$ lin. indep & $v_1, \dots, v_p \in \operatorname{span}(u_1, \dots, u_m) \Rightarrow p \leq m$

Therefore $p = m$ QED (2)

To prove (1) we need the following

Lemma. If $u_1, \ldots, u_m$ are linearly independent vectors and $u \notin \text{span}(u_1, \ldots, u_m)$ then $u_1, \ldots, u_m, u$ are linearly independent

Proof. If $a_1 u_1 + \ldots + a_m u_m + a u = 0$ we must have $a = 0$; otherwise $u \in \text{span}(u_1, \ldots, u_m)$. But then $a_1 u_1 + \ldots + a = 0 \Rightarrow a_1 = \ldots = a_n = 0$ since $u_1, \ldots, u_m$ are linearly independent.

QED (Lemma)

Now let's prove (1). If $S = \{0\}$, the zero subspace then by convention we take the empty set as basis for S. Then $\dim(S) = 0$ which is consistent with the fact that S is a single point.
If $S \neq \{0\}$ let $u_1 \in S$, $u_1 \neq 0$. If $S = \text{span}(u_1)$ then $\{u_1\}$ is a basis for S and $\dim(S) = 1$. If $S \neq \text{span}(u_1)$ then there is $u_2 \in S$, $u_2 \notin \text{span}(u_1)$. If $S = \text{span}(u_1, u_2)$ then $\{u_1, u_2\}$ is a basis for S and $\dim(S) = 2$. Continuing this process, suppose we have found linearly independent vectors $u_1, \ldots, u_m \in S$.

If $S = \text{span}(u_1, \ldots, u_m)$ then $\{u_1, \ldots, u_m\}$ is a basis for S and $\dim(S) = m$.
If $S \neq \text{span}(u_1, \ldots, u_m)$ then there is a vector $u_{m+1} \in S$ with $u_{m+1} \notin \text{span}(u_1, \ldots, u_m)$.

Then, by the lemma, $u_1, \ldots, u_m, u_{m+1}$ are linearly independent. If these vectors span S then they are a basis for S and $\dim(S) = m+1$. Since $S \subseteq \mathbb{R}^n$ the largest number of independent vectors in S is n. It follows that the above process ends in at most n steps.

QED (Theorem)