Russell's paradox: Let $\mathcal{A}$ be the set of all sets which do not contain themselves = $\{S \mid S \notin S\}$

Ex: $\{1\} \in \{\{1\}, \{1, 2\}\}$, but $\{1\} \notin \{1\}$

Is $\mathcal{A} \in \mathcal{A}$?

Suppose $\mathcal{A} \in \mathcal{A}$. Then by definition of $\mathcal{A}$, $\mathcal{A} \notin \mathcal{A}$.

Suppose $\mathcal{A} \notin \mathcal{A}$. Then by definition of $\mathcal{A}$, $\mathcal{A} \in \mathcal{A}$.

Thus we need axioms in order to create mathematical objects.

Principia Mathematica by Alfred North Whitehead and Bertrand Russell

From: http://plato.stanford.edu/entries/principia-mathematica/

Logicism is the view that (some or all of) mathematics can be reduced to (formal) logic. It is often explained as a two-part thesis. First, it consists of the claim that all mathematical truths can be translated into logical truths or, in other words, that the vocabulary of mathematics constitutes a proper subset of the vocabulary of logic. Second, it consists of the claim that all mathematical proofs can be recast as logical proofs or, in other words, that the theorems of mathematics constitute a proper subset of the theorems of logic. In Bertrand Russell's words, it is the logicist's goal "to show that all pure mathematics follows from purely logical premises and uses only concepts definable in logical terms."[1]

From: http://www.math.vanderbilt.edu/~schectex/ccc/choice.html

Axiom of Choice. Let C be a collection of nonempty sets. Then we can choose a member from each set in that collection. In other words, there exists a function f defined on C with the property that, for each set S in the collection, f(S) is a member of S.

Bertrand Russell was more famous for his work in philosophy and political activism, but he was also an accomplished mathematician. His book Introduction to Mathematical Philosophy includes some discussion of AC. Here is my paraphrasing of part of what he said:

To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed.

The idea is that the two socks in a pair are identical in appearance, and so we must make an arbitrary choice if we wish to choose one of them. For shoes, we can use an explicit algorithm – e.g., "always choose the left shoe." Why does Russell's statement mention infinitely many pairs? Well, if we only have finitely many pairs of socks, then AC is not needed – we can choose one member of each pair using the definition of "nonempty," and we can repeat an operation finitely many times using the rules of formal logic (not discussed here).

A few pure mathematicians and many applied mathematicians (including, e.g., some mathematical physicists) are uncomfortable with the Axiom of Choice. Although AC simplifies some parts of mathematics, it also yields some results that are unrelated to, or perhaps even contrary to, everyday "ordinary" experience; it implies the existence of some rather bizarre, counterintuitive objects. Perhaps the most bizarre is the Banach-Tarski Paradox: It is possible to take the 3-dimensional closed unit ball, $B = \{(x, y, z)R^3 : x_2 + y_2 + z_2 < 1\}$ and partition it into finitely many pieces, and move those pieces in rigid motions (i.e., rotations and translations, with pieces permitted to move through one another) and reassemble them to form two copies of B.

At first glance, the Banach-Tarski result seems to contradict some of our intuition about physics – e.g., the Law of Conservation of Mass, from classical Newtonian physics. If we assume that the ball has a uniform density, then the Banach-Tarski Paradox seems to say that we can disassemble a one-kilogram ball into pieces and rearrange them to get two one-kilogram balls. But actually, the contradiction can be explained away: Only a set with a defined volume can have a defined mass. A "volume" can be defined for many subsets of R^3 — spheres, cubes, cones, icosahedrons, etc. — and in fact a "volume" can be defined for nearly any subset of R^3 that we can think of. This leads beginners to expect that the notion of "volume" is applicable to every subset of R^3. But it's not. In particular, the pieces in the Banach-Tarski decomposition are sets whose volumes cannot be defined.

More precisely, Lebesgue measure is defined on some subsets of R^3, but it cannot be extended to all subsets of R^3 in a fashion that preserves two of its most important properties: the measure of the union of two disjoint sets is the sum of their measures, and measure is unchanged under translation and rotation. The pieces in the Banach-Tarski decomposition are not Lebesgue measurable. Thus, the Banach-Tarski Paradox gives as a corollary the fact that there exist sets that are not Lebesgue measurable. That corollary also has a much shorter proof (not involving the Banach-Tarski Paradox) which can be found in every introductory textbook on measure theory, but it too uses the Axiom of Choice.