I have not given any references. You can consult any book on set theory and read the topics we are discussing like Halmos, Naive Set Theory. One of the best sources is the book Real Analysis by Hewitt and Stromberg. Generally most of the books contain much more (or much less) than what I intend to do. That is why I have not given any references so far. But please do consult any book you like from the library.

Axiom of Choice is abbreviated to AC and last time we introduced four statements.

The first one is existence of choice set: Given any family of nonempty sets which are pairwise disjoint, you can make a set which contains exactly one point from of each of the given sets. you would say 'is it not obviously true? why make an axiom?' You do not think an axiom is necessary at all. Let us name the second statement as existence of choice function; given a family of (non-empty) sets there is a function which associates with each set in the family one point of that set.

The content of the statement is brought out very beautifully by Bertrand Russel: Given lots of pairs of shoes you can provide an algorithm to select one shoe from each pair; but given lots of pairs of new socks you have no algorithm to select one from each pair.

Let us name the fourth statement: every set can be well-ordered as well-ordering principle. This would in particular imply that the set of real numbers can be well ordered. That is, you can prescribe an linear order on R such that every non-empty subset has a first point. You would probably say 'Is it not obviously false? why discuss this statement?'. You believe that this statement is blatantly false. The third statement: partially ordered set has maximal element if every chain has upper bound; is called Zorn's lemma (due, in another form, to Hausdorff and later Kuratowski, but popularised by Max Zorn).

However all are just the same.

Theorem: (1) existence of choice set (2) existence of choice function (3) Zorn's Lemma and (4) Well-ordering principle; are equivalent statements.

We have seen already an application of Zorn, to show existence of Basis for any vector space. You have learnt (finite dimensional vector spaces) in the first semester. Second semester you learnt groups. If you travel a little more on one of its roads, you will reach the Neilsen-Schreier theorem which says that every subgroup of a free-group is itself a free-group. This is proved using AC.

In this third semester you are learning rings. The fact that a non-trivial unital ring has non-trivial maximal ideals needs AC. Next semester you will learn fields. You will learn that very field can be put inside a bigger field F which is algebraically closed. This last phrase means every nontrivial polynomial in one variable x with coefficients coming from F will assume the value zero for at least one value of $x \in F$. This is achieved with AC.

So much for algebra. There are several applications in Analysis which is our interest. They will unfold in your later courses. We shall see one or two. But first few useful things are to be observed.

(1) Countable union of countable sets is countable.

Of course we have already done this rather quietly. For a moment assume that all the sets are infinite. You see that if $(A_n : n \geq 1)$ are the sets, then by definition of 'countably infinite', there exists a bijection between A_n and N. But what you need is to pick *one* bijection for *each* n (so that you can produce a map of the union to N). AC allows you this.

(2) For any two sets X and Y, either $|X| \leq |Y|$ or $|Y| \leq |X|$.

In other words there is an injection from X to Y or from Y to X. equivalently, any two sets can be compared. To prove this, let P be the set of all pairs (S, f) where $S \subset X$ and f is an injection on S to Y. This is non-empty because you can take any $x \in X$ and any $y \in Y$; put $S = \{x\}$ and $f(x) = y$. Then $(S, f) \in P$. Well, if one of them is empty you need not prove any thing (why?), so assume both are non-empty sets.

Define $(S, f) \leq (T, g)$ if $S \subset T$ and g extends f. The last phrase means that for $x \in S$ we have $f(x) = g(x)$. This is a partial order and so P is a poset. Every chain has an upper bound. Indeed if $\{(S_\alpha, f_\alpha) : \alpha \in \Delta\}$ is a chain then here is upper bound: take S as union of all the sets S_α. For $x \in S$, put $f(x) = f_\alpha(x)$ in case $x \in S_\alpha$. This is a good definition because if x is in two of the sets then the fact that we have a chain implies that you get the same values whichever f_α you use.

So by Zorn, get a maximal element (T, g). In case $T = X$ then you got an injection on X to Y. In case range g is all of Y, then g^{-1} provides an injection on Y to X. If neither happens, then you pick $x^* \in X - T$ and $y^* \in Y -$ range g. Put $T_1 = T \cup \{x^*\}$; put $g_1(x) = g(x)$ for $x \in T$ and $g_1(x^*) = y^*$. You see that $(T, g) < (T_1, g_1)$ contradicting maximality of (T, g).
This completes proof.

(3) Every infinite set contains a copy of N, that is, contains a countably infinite set.

Fix a choice function f for subsets of X, the given infinite set. This means for each non-empty subset $S \subset X$ we have $f(S) \in S$. Define

$$x_1 = f(X)$$

$$x_2 = f(X - \{x_1\})$$

$$x_3 = f(X - \{x_1, x_2\})$$

and in general

$$x_{k+1} = f(X - \{x_1, \cdots, x_k\}).$$

The set $(x_i : i \geq 1)$ does the job. Carefully note that these are distinct elements.

(4) If X is an infinite set and $F \subset X$ is a finite set, then $|X| = |X - F|$. That is, X and $X - F$ have the same cardinality.

Denote your finite set by

$$F = \{x_1, \cdots x_k\}.$$

Fix a choice function as above. Put

$$x_{k+1} = f(X - \{x_1, \cdots, x_k\})$$

$$x_{k+2} = f(X - \{x_1, \cdots, x_k, x_{k+1}\})$$

and in general for $n \in N$

$$x_{k+n} = f(X - \{x_1, \cdots, x_{k+n-1}\})$$

Here is a bijection on X onto $X - F$. If $x \notin (x_i : i \geq 1)$ put $h(x) = x$. Put $h(x_i) = x_{i+k}$ for $i = 1, 2, \cdots$. This does the job.

You only need to note that X is partitioned into two sets $A_1 = \{x_1, x_2, \cdots\}$ and $A_2 = X - A_1$. Similarly $X - F$ is partitioned into two sets $B_1 = \{x_{k+1}, x_{k+2}, \cdots\}$ and $B_2 = X - A_1$ (watch out, not $X - B_1$). The map h is an bijection of A_i and B_i.

(5) If X is an infinite set then $|X \times \{0, 1\}| = |X|$.

That is, two copies of X has the same potency as one copy of X. Remember $X \times \{0, 1\}$ consists of all pairs $(x, 0)$ for $x \in X$ and also all pairs $(x, 1)$ for $x \in X$.

To prove the statement, let P consist of all pairs (S, f) where $S \subset X$ and f is a bijection on $S \times \{0, 1\}$ to S. This is non-empty. Indeed using (3) you get S which is a copy of N. Since $S \times \{0, 1\}$ is also countably infinite there is a bijection f of $S \times \{0, 1\}$ to S. This (S, f) is in P.

Define $(S, f) \leq (T, g)$ if $S \subset T$ and g extends f. This makes sense because $S \times \{0, 1\} \subset T \times \{0, 1\}$. This makes P a poset.

Every chain has an upper bound. If (S_α, f_α) is a chain then take S as the union of these S_α. If $(x, i) \in S \times \{0, 1\}$ put $h(x, i) = f_\alpha(x, i)$ in case $x \in S_\alpha$. again the fact that we have a chain tells us there is no conflict in case x is in two of the sets S_α. It is a bijection too. If you take two distinct points (x, i) and (y, j) then there is one S_α in which both x and y are present etc.

Let (T, g) be a maximal element, exists by Zorn. Can $X - T$ be infinite? No because of the following reason. If it were infinite, use (3) and get a copy of N in $X - T$, denote it by A. Since $A \times \{0, 1\}$ is a countably infinite set, fix a bijection φ on $A \times \{0, 1\}$ onto A.

Take $T_1 = T \cup A$. Define g_1 on $T_1 \times \{0, 1\}$ as follows. If $x \in T$, then $g_1(x, i) = g(x, i)$ and if $x \in A$, then $g_1(x, i) = \varphi(x, i)$. Note that φ takes you to A whereas g takes you to T. This helps you to show that g_1 is bijection. But then $(T, g) < (T_1, g_1)$ contradicting maximality of (T, g).

Thus $X - T$ is finite and hence by (4), T has the same cardinality as X. Now

$$X \times \{0, 1\} \sim T \times \{0, 1\} \sim T \sim X.$$

(6) If S, T are disjoint infinite sets, $|S| < |X|$ and $|T| < |X|$ then $|S \cup T| < |X|$.

recall $|S| < |X|$ means there is an injection from S to X but there is no bijection.

In view of (2) either $|S| \leq |T|$ or $|T| \leq |S|$, there is no loss in assuming that $|S| \leq |T|$. So there is an injection $f : S \to T$. Define g from $S \cup T$ to $T \times \{0, 1\}$ by $g(x) = (f(x), 0)$ if $x \in S$ and $g(x) = (x, 1)$ if $x \in T$. Note that ranges of S and T under g are disjoint. Easy to see that g is an injection showing

$$|S \cup T| \leq |T \times \{0, 1\}| = |T| < |X|.$$

where we used (5) for the equality. The first inequality is witnessed by g and the last inequality is hypothesis. You only need to note that if $|A| = |B|$ and $|B| < |C|$ then $|A| < |C|$.

(7) if X is infinite then $X \times X \sim X$.

Let P be the set of all pairs (S, f) where $S \subset X$ and f is a bijection on $S \times S$ onto S. This is partially ordered by inclusion. That is $(S_1, f_1) \leq (S_2, f_2)$ if $S_1 \subset S_2$ and f_2 is an extension of f_1. This means $f_2(x, y) = f_1(x, y)$ for $x, y \in S_1$. This makes P a poset.

P is non-empty because from (3) you can take a countably infinite set $S \subset X$ and use the fact that $S \times S$ is also countably infinite to get a bijection $f : S \times S \to S$. Then $(S, f) \in P$.

Every chain has upper bound. Indeed if (S_α, f_α) is a chain, here is the upper bound: S is the union of all the S_α. For $x, y \in S$, if they both belong to S_α we put $f(x, y) = f_\alpha(x, y)$. Since we have a chain, there is one α such that both x and y are in S_α. Since we have a chain, this definition does not depend on which α we take. That this is a bijection is routine. if you take two pairs (x_1, y_1) and (x_2, y_2) then there is one S_α in which all these four points x_1, x_2, y_1, y_2 are available and f_α is a one-one map on this $S_\alpha \times S_\alpha$ so that the f_α values of these two pairs are different, but these are f values as well. Also given any $a \in S$, get an α such that $a \in S_\alpha$ and hence there is an $(x, y) \in S_\alpha \times S_\alpha$ such that $f_\alpha(x, y) = a$ so that $f(x, y) = a$.

Take a maximal element (T, g). Since $T \subset X$ there are only two possibilities, either $|T| = |X|$ or $|T| < |X|$ — the identity map shows that $|T| \leq |X|$.

If $|T| = |X|$ we are done because

$$|X \times X| = |T \times T| = |T| = |X|.$$

Assume $|T| < |X|$. we show contradiction for maximality by producing (T_1, g_1) such that $(T, g) < (T_1, g_1)$.

Towards this end we first note that $X - T$ can not be finite. If it were, then X and T differ by finitely many points so that (4) tells $|X| = |T|$ which is not the case now.

Since $T \times T \sim T$ we conclude that T is infinite; we saw that $X - T$ is infinite. We are assuming that $|T| < |X|$. If we also have $|X - T| < |X|$ then (6) leads to contradiction $|X| < |X|$. Thus we must have $X - T \sim X$.

We are now ready to contradict maximality of (T, g). Since $X - T$ and X are of the same cardinality, pick a subset $S \subset X - T$ with $|S| = |T|$. You only need to fix a bijection φ on X onto $X - T$ and take $S = \varphi(T)$. Let $T_1 = T \cup S$. We shall now define a bijection g_1 on $T_1 \times T_1$ onto T_1 which extends g. Let us make two observations.

Firstly, $T_1 \times T_1$ is disjoint union of the four sets $A_0 = T \times T$; $A_1 = T \times S$; $A_2 = S \times T$; $A_3 = S \times S$. And also these four sets have the same cardinality simply because S and T have the same cardinality. Since $T \times T \sim T$ (Remember g) we conclude that all these four sets are equipotent with T and also with S.

Second observation is the following. We can express S as disjoint union of three sets S_1, S_2, S_3 all having the same cardinality as S. This is seen as follows.

We know that T is infinite. So take two points s and t from T and denote $T_0 = T - \{s, t\}$. consider the sets $A = \{s\} \times T$; $B = \{t\} \times T$; $C = T_0 \times T$. You see that A has same power as T; B has same power as T; T_0 which differs from T by a finite set has same power as T and hence C has the same power as $T \times T$ which has same power of T (Remember g). Thus we decomposed $T \times T$ into three disjoint sets each of power of T. But $T \times T$ has same power as T (Remember g). so we could decompose T into three sets of the power of T. But S and T are of the same power so we can decompose S into three sets each of the same power as S. Denote $S = B_1 \cup B_2 \cup B_3$.

To complete the proof define g_1 on $T_1 \times T_1$ as follows. On $A_0 = T \times T$ follow g to take you to T; on the other three sets A_1, A_2, A_3 which makeup the remaining part of $T_1 \times T_1$ use any maps taking you in a bijective way to B_1, B_2, B_3 respectively.

This completes the proof (by showing an element larger than the alleged maximal element and thereby establishing that T must have same cardinality as X and thereby $X \times X$ must have a bijection to X).

(8) Let X be an infinite set. Let $seq(X)$ denote the set of finite sequences of points from X. That is, things of the form $(x_1, x_2, \cdots x_k)$ where $k \geq 1$ is an integer and each $x_i \in X$. Then $seq(X) \sim X$.

Since $T \times T$ is of power T, we conclude, in particular, that a countable union $S = \cup T_i$ of disjoint sets $T_1, T_2, \cdots$ each of power T has power T. Indeed T being infinite, get by (3), an injection $f : \{1, 2, \cdots\} \to T$ and for each $i \geq 1$ get a bijection $f_i : T_i \to T$. Let now $s \in S$, then there is unique i such that $s \in T_i$. Put $g(s) = (f(i), f_i(s))$ gives a injection from S to $T \times T \sim T$.

Now X, $X \times X$, $X \times X \times X$ $\cdots$ all have power of X and the earlier para tells you that their union is also of power X. But this union is precisely $seq(X)$.

Sometimes, for technical reasons one includes the empty sequence also in $seq(X)$. Empty sequence means sequence of length zero. There is only one such sequence, namely (). Even if you include this one element in the set $Seq(X)$ its power is still same as that of X.

Of course, some of you may have a psychological objection to include the empty sequence as a sequence at all. Do not worry, in that case, you do not have to do this. I only said: if sometime later somebody does some such thing you need not scratch your head. That is all.

(9) Let X be an infinite set. $Seq(Q) \times Seq(X) \sim X$.

Proof is already included in the above. for each fixed $a \in Seq(Q)$ the set of points in our set with first coordinate equal to a has power $Seq(X) \sim X$ and the number of possible a is countable.

(10) $f : X \to Y$ be surjection. Then $|Y| \leq |X|$.

Fix a choice function φ for non-empty subsets of X. For each $y \in Y$ the set $A_y = f^{-1}(y)$ is non-empty because f is surjection; so it makes sense to define $g(y) = \varphi(A_y)$. Then g is a one-one map of Y to X and $|Y| \leq |X|$.

Now we shall return to Analysis.

We showed that every vector space has a basis. We made no fuss about the underlying field. So let us use that freedom. Consider R as a vector space over Q, the field of rational numbers. Thus we have the following.

(11) The vector space R over the field Q has a basis. That is, there is a set $B \subset R$ such that every $x \in R$ can be uniquely expressed as a finite sum $\sum q_i b_i$ where $b_i \in B$ are distinct and $q_i \in Q$. Here uniqueness is interpreted as: if there are two such sums representing x the b s with non-zero coefficients are same in both expressions and those non-zero coefficients also agree.

In other words you can cheat by taking a finite sum and adding to it a term: $(0 \times b)$ with $b \in B$ which was not already there in the sum.

Such a Basis is called Hamel basis. How large is it?

(12) Any Hamel basis has the same cardinality as that of R.

Define a map $f : seq(Q) \times seq(B) \to R$ by
$f(q_1, \cdots, q_k; b_1, b_2, \cdots, b_l) = 0$ if $k \neq l$ and $= \sum q_i b_i$ if $k = l$.
Since B is a basis, this map is onto R. Thus (10) and (9) imply

$$|R| \leq |seq(Q) \times seq(B)| = |B|$$

But obviously $|B| \leq |R|$ and thus $|B| = |R|$. In using (9) we implicitly assumed that B is infinite. In fact, it is uncountable because if it were countable then we note that $seq(Q) \times seq(B)$ is also countable showing that the first inequality above itself is a contradiction. Just remember that the set of real numbers is not countable.

(13) There is a function $f : R \to R$ such that $f(x + y) = f(x) + f(y)$ for all $x, y \in R$ which is not continuous. In other words it is a (additive) group homomorphism.

You can make the range of f to be any non-trivial Q-subspace of R.

You can make f a bijection, so that it is a group isomorphism.

Fix a Hamel basis B.

Take $v \in B$. Here is f. For any $x \in R$, $f(x)$ is the coefficient of v in the expression of x as a finite rational linear combination of elements of B.

That this is a Q-linear map and hence additive is general vector space result. This is not continuous because it takes only rational values and is not the zero function. (Pause and think)

Take any Q-subspace of R. Take a basis for the subspace, say, H. Clearly $|H| \leq |B|$ and so fix a surjection $\varphi : B \to H \cup \{0\}$. if $x = \sum q_i b_i$ define, $f(x) = \sum q_i \varphi(b_i)$. Again the fact that f is additive is a genera vector space fact. It is not continuous because if $v \in B$ with $\varphi(v) = 0$ then f takes the value zero on the set of rational multiples of v, that is, on a dense set. Range is also the given subspace.

To get an additive isomorphism, you only need to take a bijection φ of B to itself and define f on R by: if $x = \sum q_i b_i$ then $f(x) = \sum q_i \varphi(b_i)$. Of course if you take the identity map as your bijection φ then f is identity too. Take four elements from the basis, x_1, y_1, x_2, y_2 such that $y_1/x_1 \neq y_2/x_2$ and take a bijection φ of B with $\varphi(x_i) = y_i$ for $i = 1, 2$. Note that the resulting f would map x_i to y_i. It is not continuous because for any continuous map $f(x)/x$ (?) is a constant.

(14) There is an (additive) group isomorphism from R to R^2.

Think of both R and R^2 as vector spaces over Q, take Hamel bases, observe that both have the same cardinality (namely c), set up a bijection between them and extend by Q-linearity. Exactly like the above.

There are several things you can say using AC. But we stop here. We shall make some comments about the equivalence of the four statements.

Choice set existence implies choice function existtence. Indeed, given any family of sets $(A_\alpha; \alpha \in \Delta)$ put

$$B_\alpha = \{(\alpha, x) : x \in A_\alpha\}; \qquad \alpha \in \Delta.$$

These are disjoint because every point in B_α is a pair and its first coordinate is α. Take a choice set S, thus for each α we have $f(\alpha) \in B_\alpha$. Define $f(\alpha)$ to be the second coordinate of the unique point in $S \cap B_\alpha$.

Choice function implies choice set. If $(A_\alpha : \alpha \in \Delta)$ are disjoint sets and f is a choice function, that is, $f(\alpha) \in A_\alpha$ for each α then we take $S =$ Range f. Then S is a choice set for the given family of disjoint sets.

Zorn implies choice set. In fact let $(A_\alpha : \alpha \in \Delta)$ be disjoint sets. Let

$$P = \{S \subset \cup A_\alpha : |S \cap A_\alpha| \leq 1 \quad \text{for each} \quad \alpha\}.$$

This is a poset by defining $S_1 \leq S_2$ if $S_1 \subset S_2$. It is non-empty because if you take one of the sets and a point x from that set then S consisting of this single point is in P. Every chain has upper bound, namely, union of sets in the chain. So let T be a maximal element. If there is an A_α such that S has no point of A_α, pick one point from this A_α and let T' be the set consisting of points in T along with this extra point. Then $T' \in P$ and contradicts maximality of T. Thus for each α, we have $|S \cap A_\alpha| = 1$ showing that S is a choice set.

Existence of choice function implies Zorn. Shall only outline but not carry out the full proof. Take a poset P where every chain has an upper bound. We need to exhibit a maximal element.

Let us say that a chain C is maximal if there is no element of P larger than every element of C. That is

$$\neg \ \exists p \in P \ \ \forall x \in C \ \ (x < p).$$

The idea is to show that there is a maximal chain. For this several methods are available. Simplest (not necessarily the best) method is the following. Fix a choice function for subsets of P. That is a function f which associates with each non-empty subset of P a point in that set. Consider

$$S_0 = P; \qquad p_0 = f(S_0);$$

$$S_1 = \{x : p_0 < x\} quad \quad p_1 = f(S_1)$$
$$S_2 = \{x : p_1 < x\}; \qquad p_2 = f(S_2)$$

and so on

$$S_\infty = \{x : p_n < x \ \text{ for all } \ n\}; \qquad p_\infty = f(S_\infty)$$
$$S_{\infty+1} = \{x : p_\infty < x\}; \qquad p_{\infty+1} = f(S_{\infty+1})$$

etc. The p s so collected form a chain because at any time we are selecting a point larger than what we already have. if at some stage the set S_α is empty then collect all the p got so far. That is a maximal chain. The fact that we have arrived at empty set signals that there is nothing larger than all the selected p s. And of course, at some stage you do arrive at empty set.

Now this maximal chain has a upper bound, let it be a. This is a maximal element of P. Indeed if there is an element b with $a < b$ we have $x \leq a < b$ for every $x \in C$ contradicting maximality of the chain.

The above procedure of selecting points can be formalized but needs some work we shall not undertake. It is not trivial. It is easy to say 'so on' but difficult to explain what is this 'so on'.

Well-ordering principle implies choice set. Indeed if (A_α) is a disjoint family of sets then well order their union, and pick the least (in that well order) lament of A_α for each α. This set of points so selected gives choice set. That choice set implies well ordering principle is carried out in a fashion similar to the above construction.

It is in this form of existence of maximal chains that Hausdorff and later, Kuratowski formulated. However Zorn's name got stuck because he popularised. Max Zorn himself does not like it to be called Zorn's lemma! It is too late to change things now.

There are several other beautiful results in set theory, but we need to return to Analysis. But before doing so few words on history. I have already mentioned that all this had origins in problems concerning Fourier series.

Unfortunately even in scientific investigations we have fundamentalism — after all we are human beings. Just as the famous Mach and other scientists opposed Boltzman's ideas; jut as the Church opposed Galileo's ideas; here too we have Georg Cantor being opposed from several quarters. the famous Kronecker and Poincare opposed rather very very vehemently. So did the Church. After all God is infinite and hence (!) infinity is God. How can the infinity that represents integers different from the infinity that represents real numbers? There can not be many Gods, there is only one God! Worse, you are saying given any God there is a bigger God! Definitely not acceptable.

It was left to Hilbert to say: No one can dislodge us from the Paradise created by Cantor. Paul Erdos used to say: Keep your brains open.

We shall now return to Analysis. One of our objectives is to construct the set of real numbers —- this means to show a set with some operations satisfying whatever we assumed last year. Of course, any construction work needs using cement, bricks, water and so on and more important, dirtying our hands.

After all, you can not construct something out of nothing. So the first question is what do we have to start with. To simplify life let us ask: how do you construct real numbers if I give you the set Q of rational numbers. It is indeed easy. After all we know enough to be able to say

$$x \in R \Rightarrow \sup\{q \in Q : q \leq x\} = x.$$

This identifies real numbers as sup of a set of rational numbers. the only nuisance above is that this set depends on x. We should be able to describe real numbers using *only* rationals. There is nothing but set of rationals before us. You can not pretend that you already have real numbers and you are only describing them using rationals.

Getting a clue from the above, let us put things differently. Every real number x, *cuts* Q into two non-empty parts: those rationals q that are not above it ($q \leq x$) and those rationals that are above it ($q > x$).

Also every cut determines a real number, namely, sup of the lower part of the cut.

So what is a cut and how do we construct a complex system of real numbers starting from Q. Remember you need to prescribe a set; need to prescribe addition and multiplication and order on your set; and then show you have a complete Archimedian ordered field.

Of course, you can also ask how do you construct rational numbers? We do so using the set Z of integers. So how do you construct Z, the set of integers. We do so using the set of natural numbers $N = \{1, 2, \cdots, \}$ with the only operation being the successor operation $Sn = n + 1$.

This brings us to: who gives you the set N and the operation S on this set which hopefully tells us 'adding one'.

You can get either terrified or excited on the attitude you take. But be assured you need not construct real numbers in the exam. then why are we doing this? Don't you want to reassure yourself that real number system does exist (in flesh and blood) and all the things you learnt in Calculus have meaning.