Some comments about HW2

1. One option is to show each has the same cardinality as $(1, \infty)$: use translation in one case and inversion (reciprocals) in the other.

2. Make sure you are always proving set statements like $X = Y$ by checking *both* $X \subseteq Y$ and $Y \subseteq X$ (separately). And check inclusions like $X \subseteq Y$ elementwise ("Well, if $x \in X$ then ... so $x \in Y$.")

Also note that inverse images are well defined whether there is an inverse function f^{-1} or not. (When f^{-1} *does* exist, it is true that an inverse image $f^{-1}(B)$ is identical to the forward image of applying the function f^{-1} to the subset B of its domain.)

3. This is the beginning of the wildness of set theory. There's a whole hierarchy of infinities here: $\mathbf{N}$, then $\mathcal{P}(\mathbf{N})$, then $\mathcal{P}(\mathcal{P}(\mathbf{N}))$, then .... Worse, there's more after that. My advice: don't go there.

5. When F is any field, functions $f : F \to F$ which preserve addition and multiplication are called *automorphisms* of F. I asked you to prove there aren't any automorphisms of $\mathbf{R}$ except the identity map. When $F = \mathbf{C}$, though, there are some, notably complex conjugation $f(a + bi) = a - bi$. The study of fields' automorphisms leads to a branch of math called Galois Theory.

Since $1 \cdot 1 = 1$ we must have $f(1) \cdot f(1) = f(1)$, so the element $z = f(1)$ makes $z^2 - z = 0$, i.e. $z \cdot (z - 1) = 0$. As we noted in class, this means one of those two factors must be zero. Having $z = 0$ isn't really a field automorphism: if $f(1) = 0$ then since $x = x \cdot 1$ for every $x \in F$ we would conclude $f(x) = f(x) \cdot f(1) = f(x) \cdot 0 = 0$, that is, f sends every element to zero! That function does indeed (trivially) preserve addition and multiplication but it's not very interesting, so I meant to exclude it. That leaves only $z = 1$. So now you know $f(1) = 1$.

Similarly you also know $f(0) = 0$: since $0 + 0 = 0$ and f preserves addition. (Actually $x = 0$ is the *only* element with $f(x) = 0$: if x is nonzero, it has an inverse with $x \cdot x^{-1} = 1$, which makes $f(x) \cdot f(x^{-1}) = f(1) = 1$, which precludes having $f(x) = 0$.)

Then $f(n) = n$ for every natural number n: we just proved it for $n = 0$ and $n = 1$, and if it's true for some value of n then $f(n + 1) = f(n) + f(1) = n + 1$.

Since $x + (-x) = 0$, it follows $f(x) + f(-x) = f(0) = 0$, so $f(-x) = -f(x)$ for every x. Thus for every natural number n, $f(-n) = -n$, meaning $f(z) = z$ has now been proved for every integer z.

Take the same line of thinking multiplicatively instead of additively and you prove $f(x^{-1}) = (f(x))^{-1}$ for every nonzero x, and in particular $f(1/n) = 1/n$ for every natural number. Then $f(m/n) = f(m \cdot (1/n)) = f(m) \cdot f(1/n) = m/n$, meaning $f(z) = z$ has now been proved for every rational number z, too.

Now all this reasoning applies to automorphisms of every field. To discuss the reals in particular, we need the other features that distinguish $\mathbf{R}$, namely the ordering and completeness. First note that if $h \geq 0$ then h is a square (Rudin's theorem 1.21), which gives $f(h) = f(z^2) = f(z)^2$ which is necessarily ≥ 0. Hence if $y \geq x$, we may let $h = y - x$ and then $f(y) = f(x + h) = f(x) + f(h) \geq f(x) + 0 = f(x)$, so f is increasing.

Now let x be any real number and ask if it's possible for $y = f(x)$ to be different from x. That would make $y < x$ or $y > x$. In either case, find a rational number a between x and y. So in the first case we have $y < a < x$; apply f to see $f(a) = a \leq f(x) = y$ since f is increasing. This is a contradiction. The second case is similar. So the only noncontradictory situation is $f(x) = x$.

So $f(x) = x$ for all real numbers x.

6. It's probably easier to write the elements of the field as $a + bi$ rather than (a, b) (where i stands for $(0, 1)$ and a, b lie in the underlying field $\mathbf{Z}_p$). Then you can use the same arithmetic you know from the complex numbers: $(a + bi)(a - bi) = a^2 + b^2$. As long as $a^2 + b^2$ isn't zero, this allows you to compute an inverse of $a + bi$, and if $a^2 + b^2$ *is* zero, you have a contradiction to the field axioms. (Recall that in a field the only time a product can be zero is if one of the factors is.) So the whole existence of inverses comes down to deciding whether it is possible to find two elements $a, b \in \mathbf{Z}_p$ with $a^2 + b^2 = 0$ i.e. $a^2 = -b^2$ or $(ab^{-1})^2 = -1$. You can resolve this question by simply squaring all the elements of $\mathbf{Z}_p$: in $\mathbf{Z}_3$ we have $0^2 = 0, 1^2 = 1, 2^2 = 1$ and so no square equals -1. But in $\mathbf{Z}_5$ we have $0^2 = 0, 1^2 = 1 = 4^2, 2^2 = 1 = 3^2$ and in particular 2 and 3 are already two square roots of -1!

More generally, -1 is a square in $\mathbf{Z}_p$ iff p is one larger than a multiple of 4 ($p = 5, 13, 17, \ldots$). That's an interesting theorem — not all that hard to prove but far from obvious — but it belongs in a Number Theory class, not Analysis, so I won't prove it here.