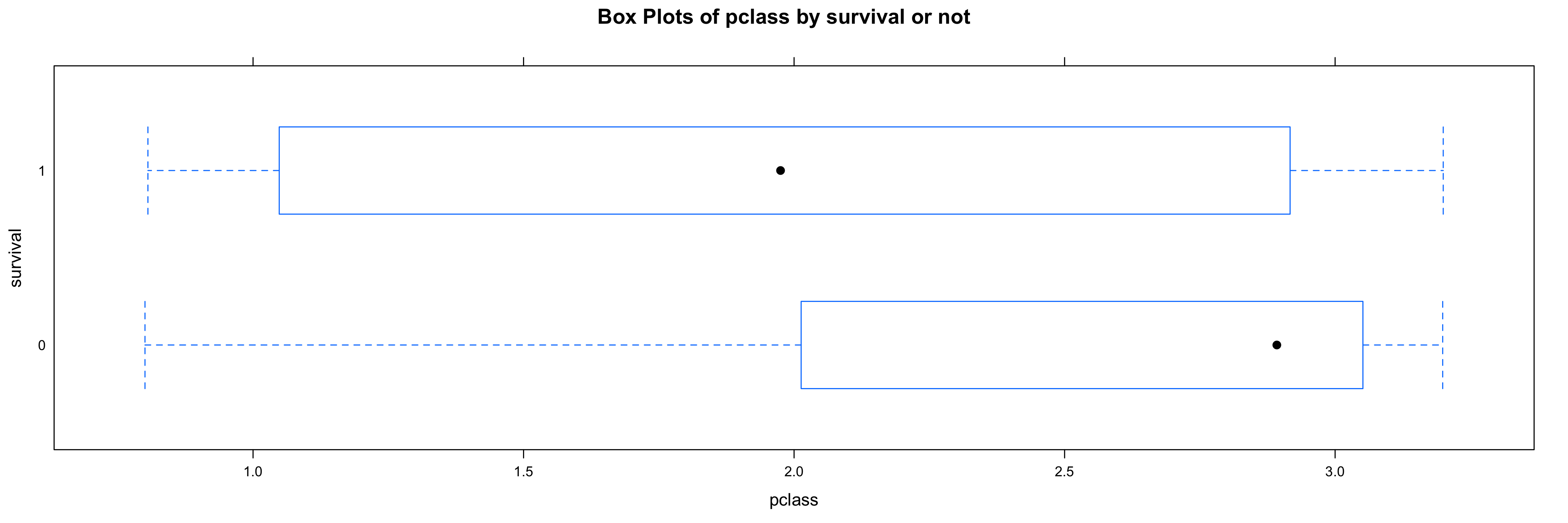
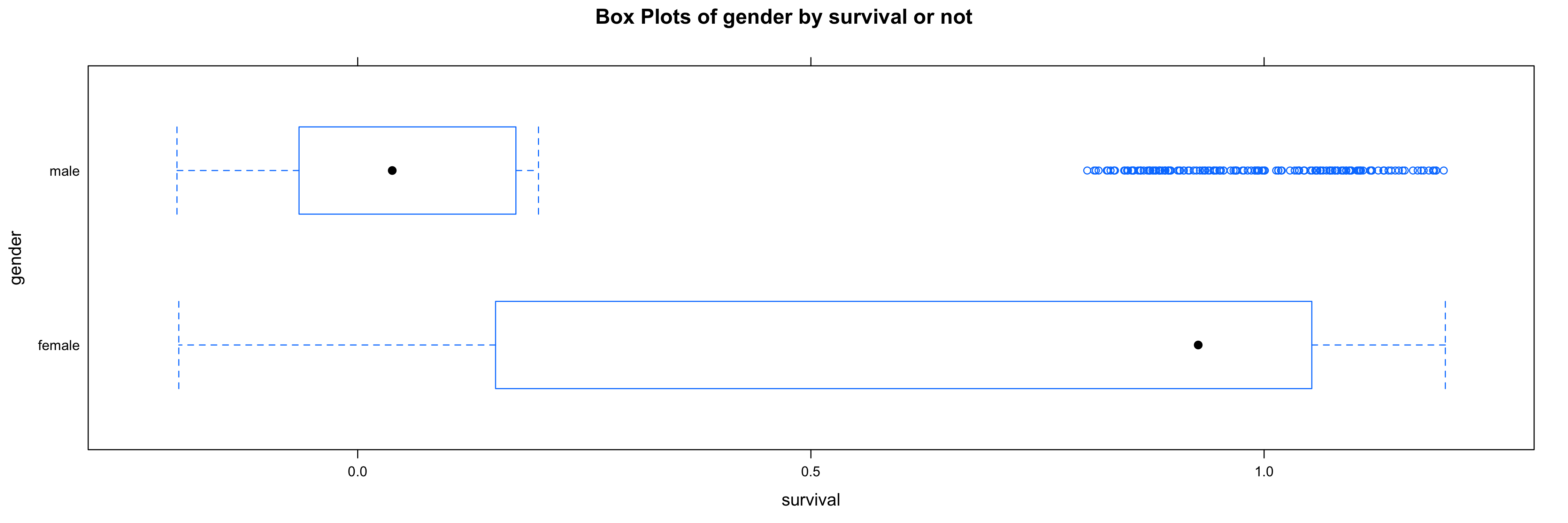
Data plot:

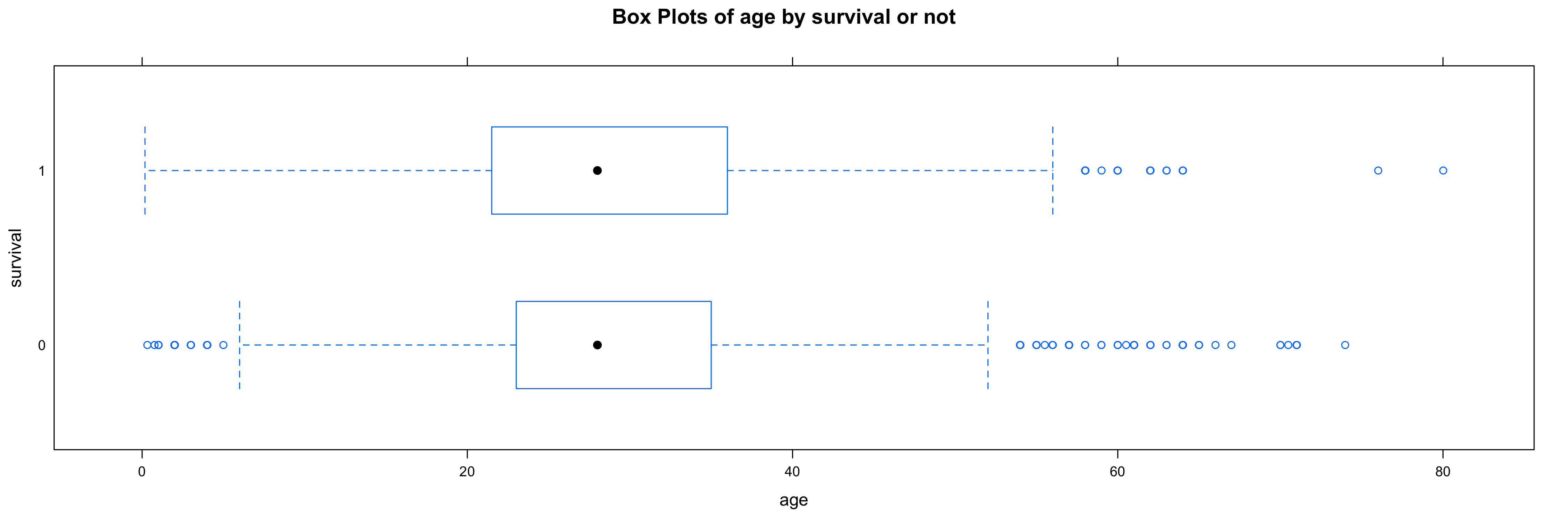
When we see the box plots of pclass by survive or not, though the people who are survived are separated in all the three pclasses, the people who are dead are mostly in the third class. The average number of dead people’s pclass is almost goes to 2.8. So if we see this plot we will find out that the pclass really has a big influence on the dead people.



Then we can check the relations between gender and survival. This box plot is pretty interesting. In the picture we can see that the male survival distribution is between 0 and 0.5, which means dead, and the average of the male survival is nearly 0. But for the female, the distribution is close from 0.5 and 1, and the average of the survival is near 1. This box plot shows that the gender could be a significant feature that influences the survival pattern.



The box plots of age by survival or not can give us an idea about the influence of age to the survival rate. In the picture we can see that the two boxes have a similar pattern. The average of survival people and dead people are nearly the same. The survival people’s age have a more board distribution, which is from 0 to 60. The dead people’s age distribution is mostly from 5 to 50. The age feature is less significant compared with gender and pclass.



When we check the feature of sibling numbers we can see that the pattern of dead and survival is nearly the same, but there are some discrete nodes in the picture. Maybe the distribution pattern is because of the small size of sibling data set and we are hard to find out by the box plots.



This picture shows the number of family member (except siblings) has influence on the survival rate. In the box plot we can see that the people who are survived have a boarder distribution than the people who dead. The average number is similar because many people have no family member. So we can make a hypothesis that people who have more family members will survive.



Then we shall see the fare. In the box plots we can see that the survival people have more fare in average than those who dead. So we can make a hypothesis that the people who pay more have higher chance to survive.

