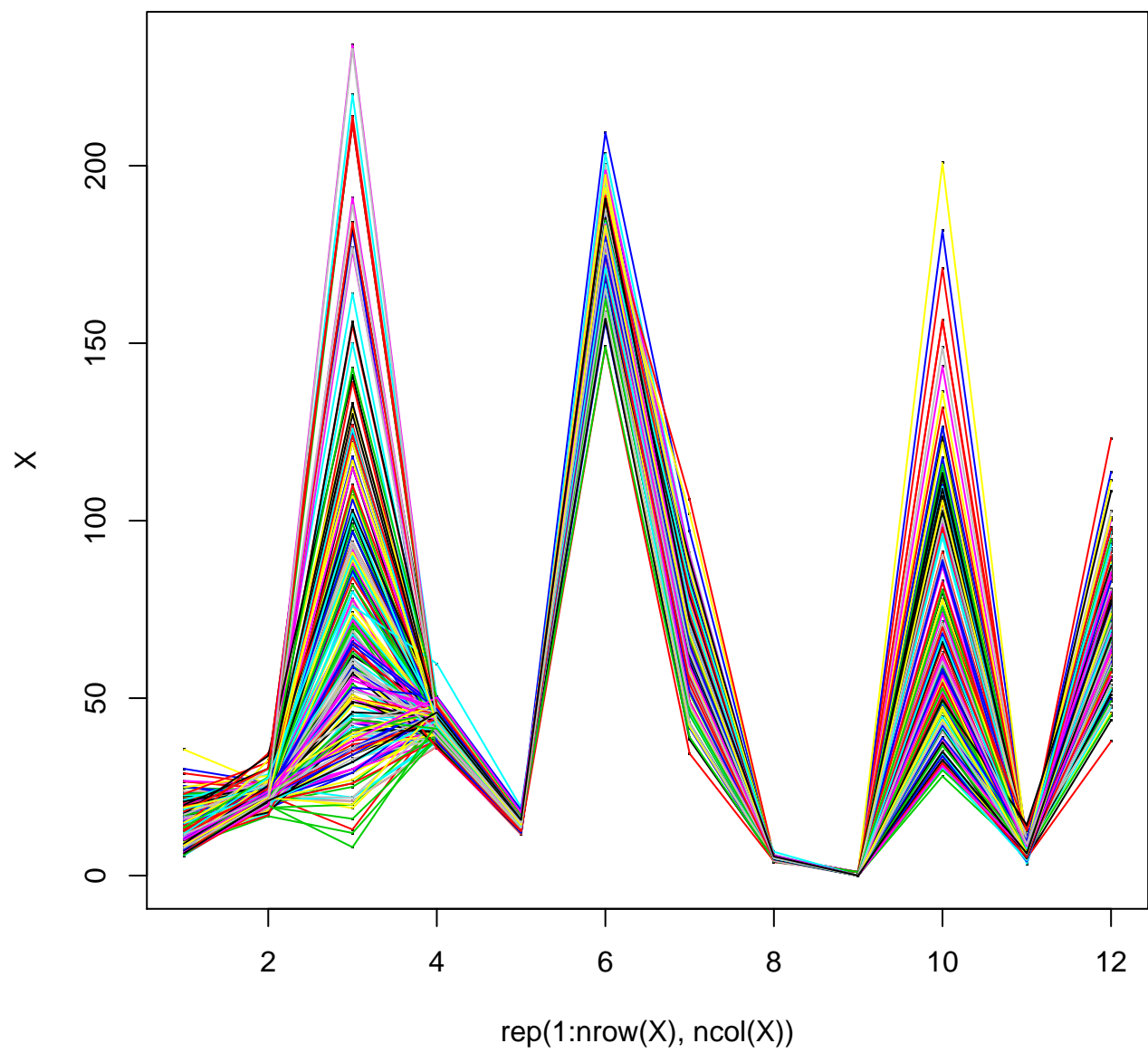FIGURE 1. Parallel plot of all variables - main

1. QUESTION 1

1.1. **a.**

FIGURE 2. Proportion of total variance explained by each eigenvalue - main
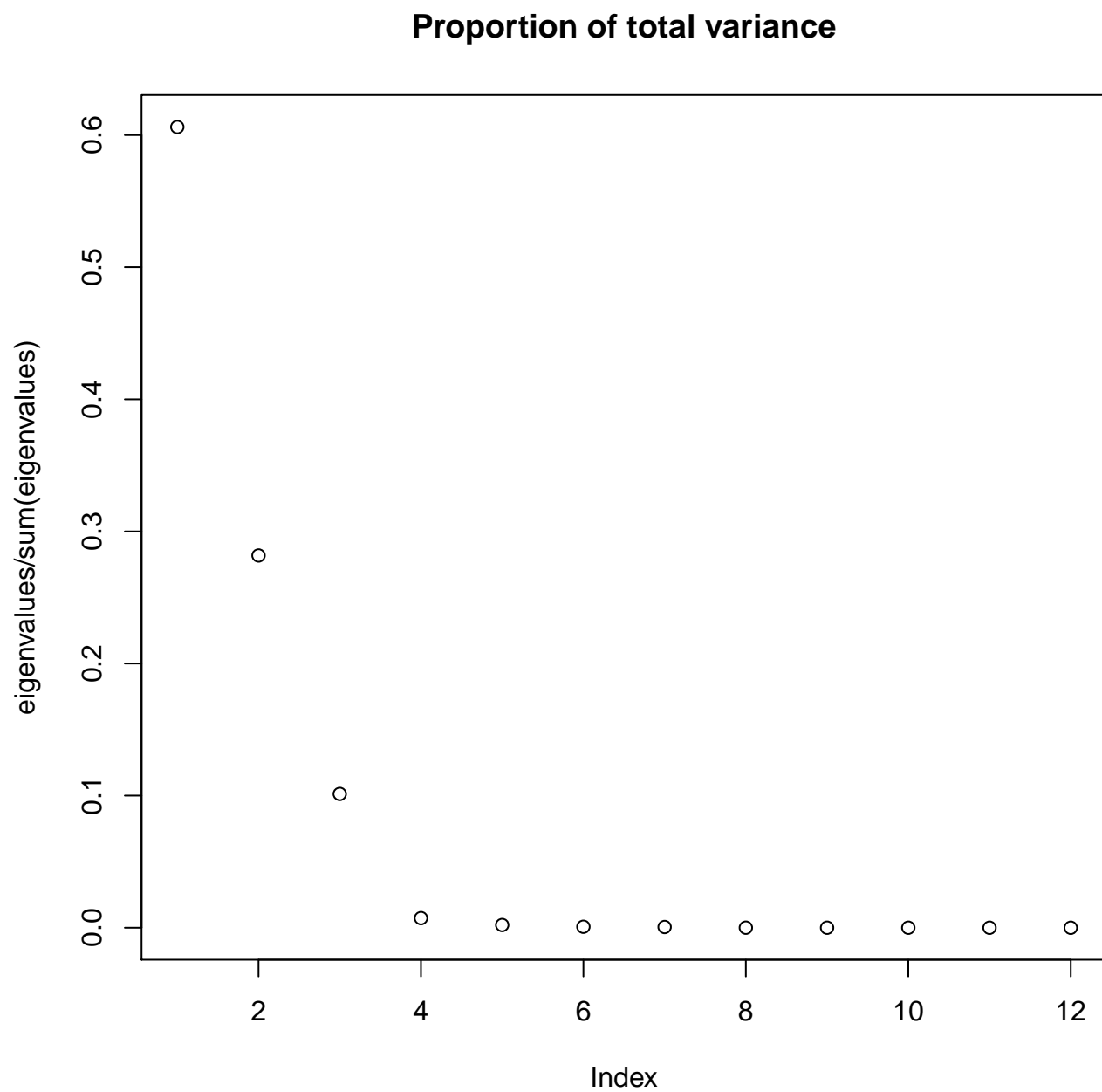

Proportion of total variance

FIGURE 3. Cumulative proportion of total variance explained by
each eigenvalue - main

Cumulative proportion of total variance
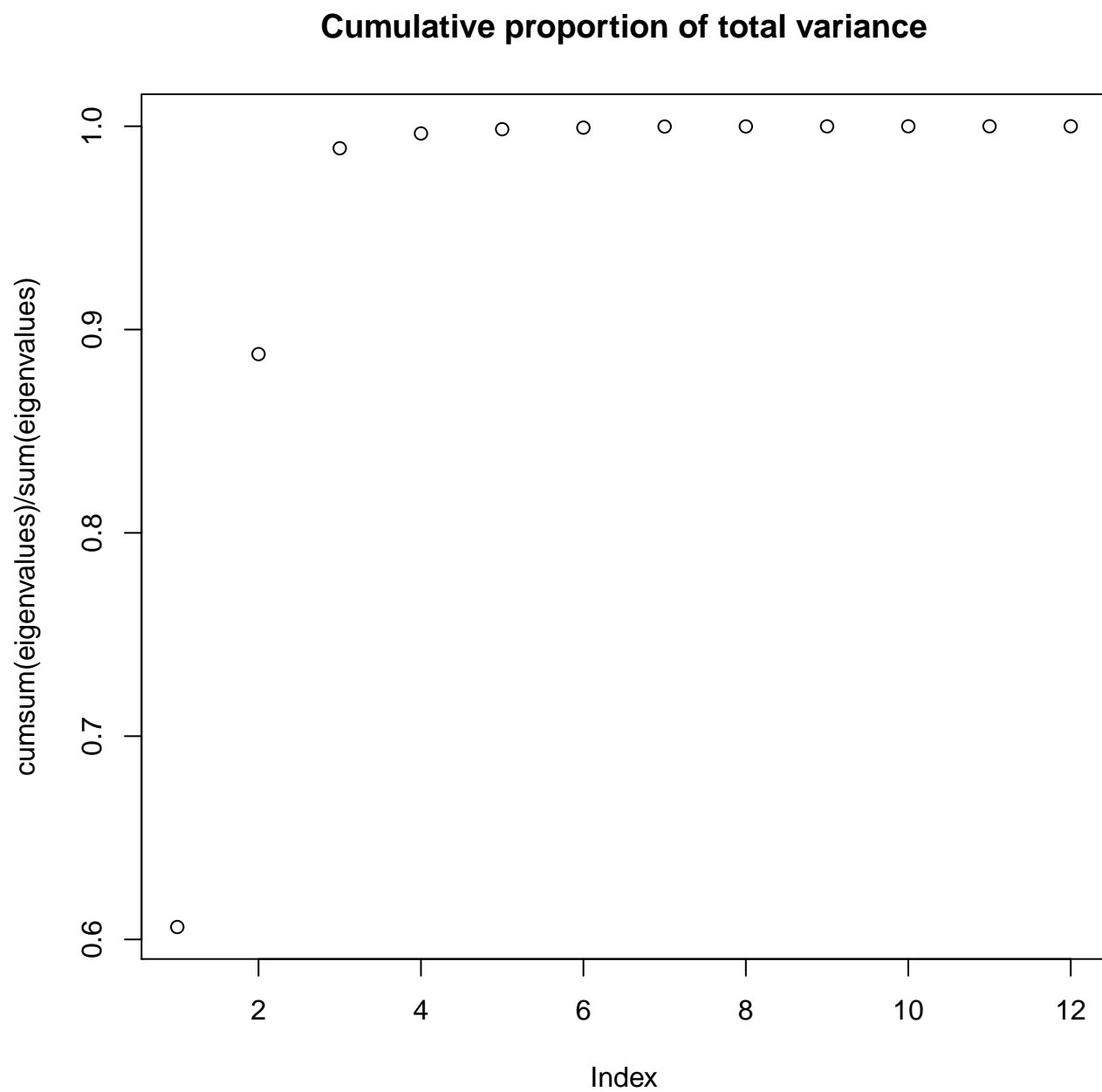
FIGURE 4. Parallel plot of first four PC scores - main
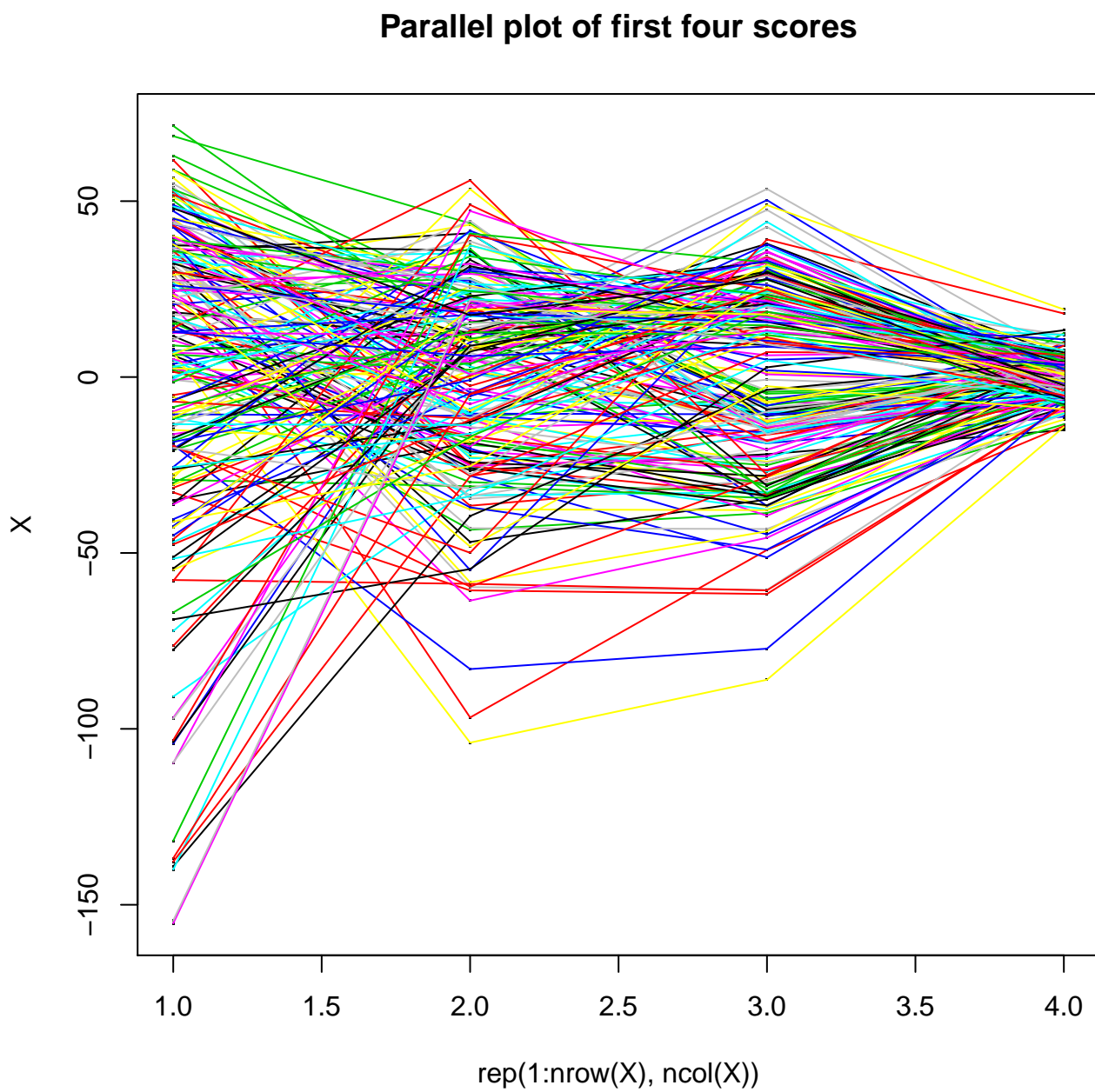

Parallel plot of first four scores

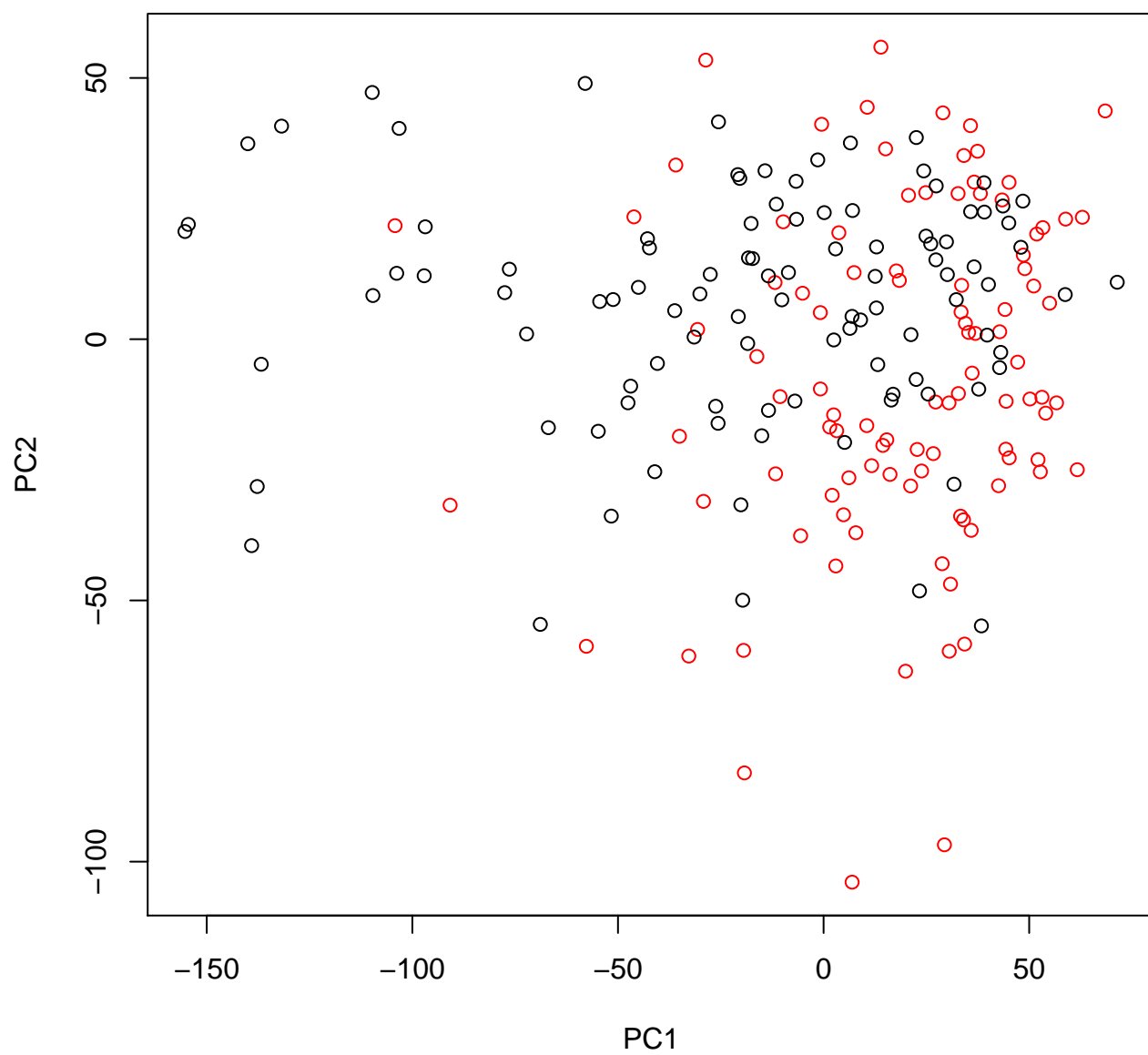Figure 5. Principal Component 1 versus Principal Component 2 - main

Figure 6. Principal Component 1 versus Principal Component 3 - main
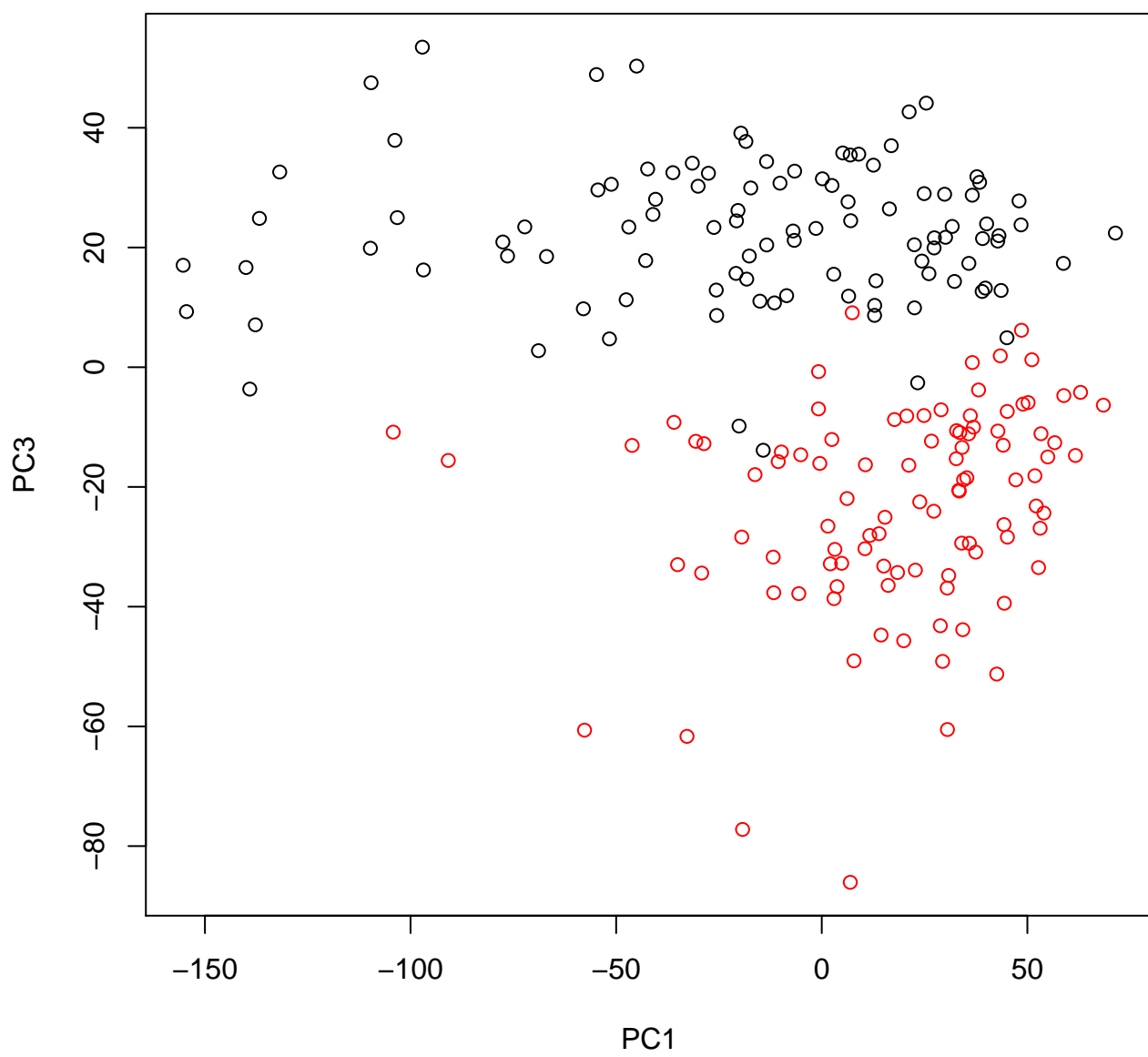
Figure 7. Principal Component 2 versus Principal Component 3 - main
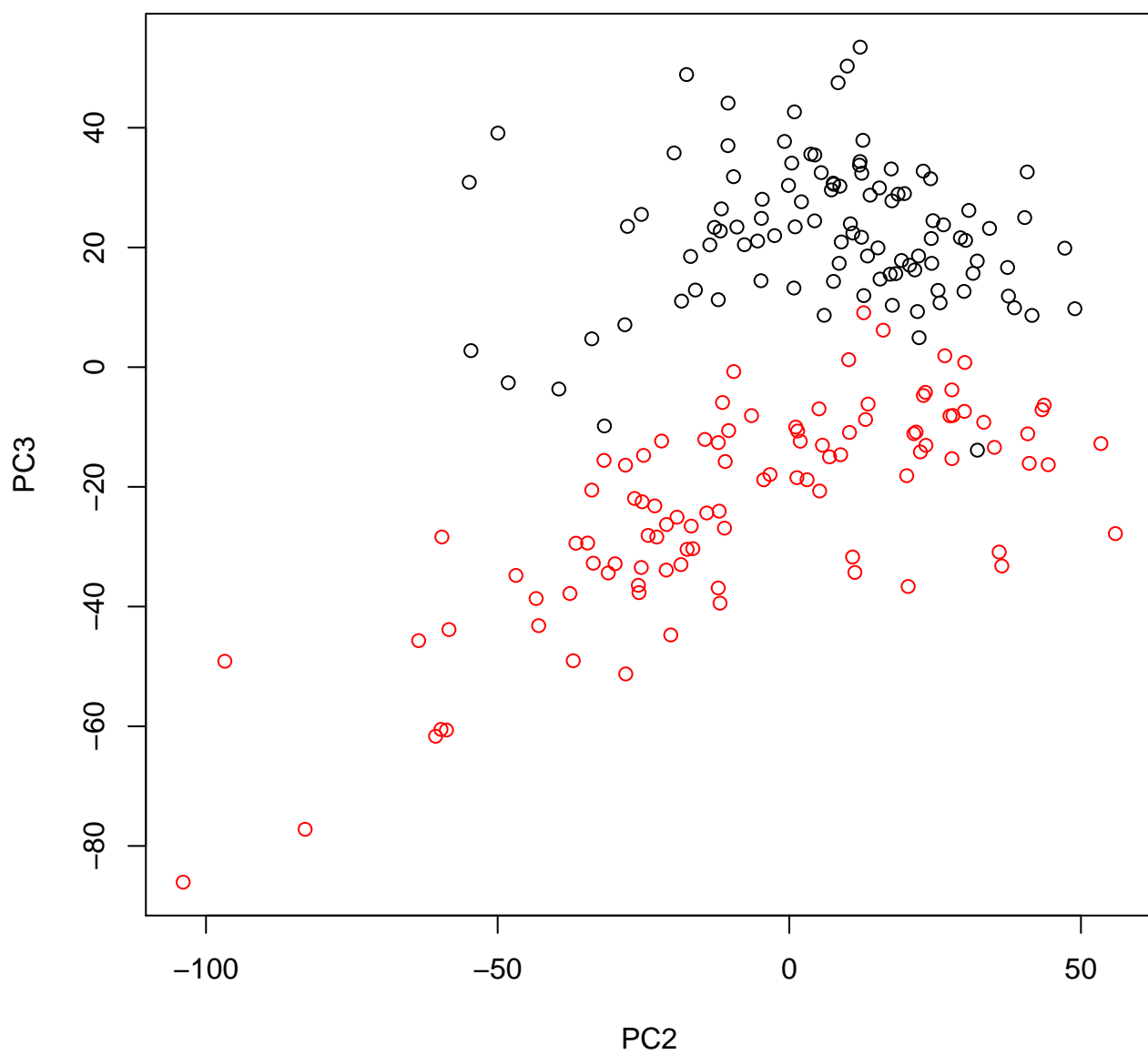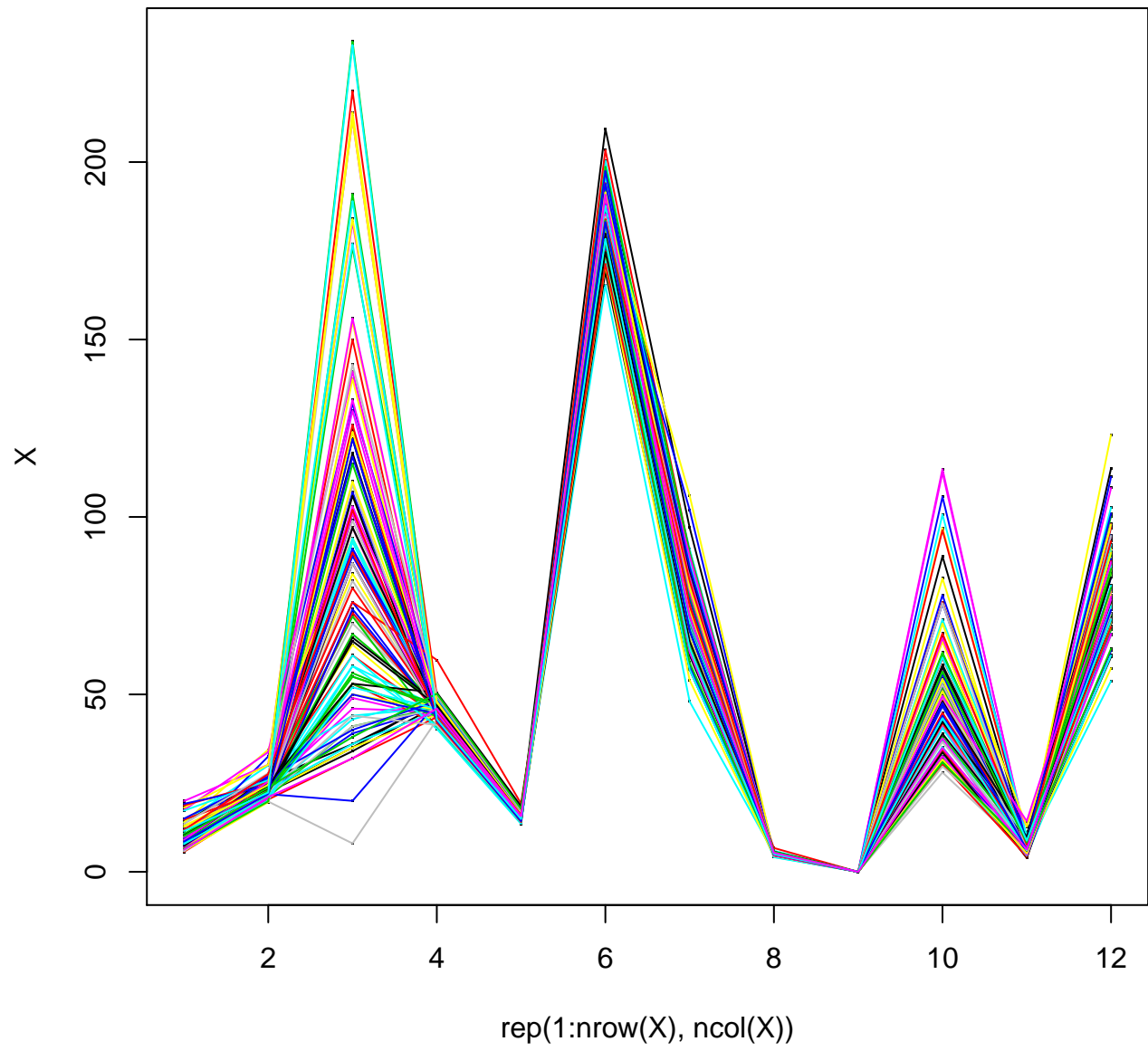
FIGURE 8. Parallel plot of all variables - male

1.2. **b.**

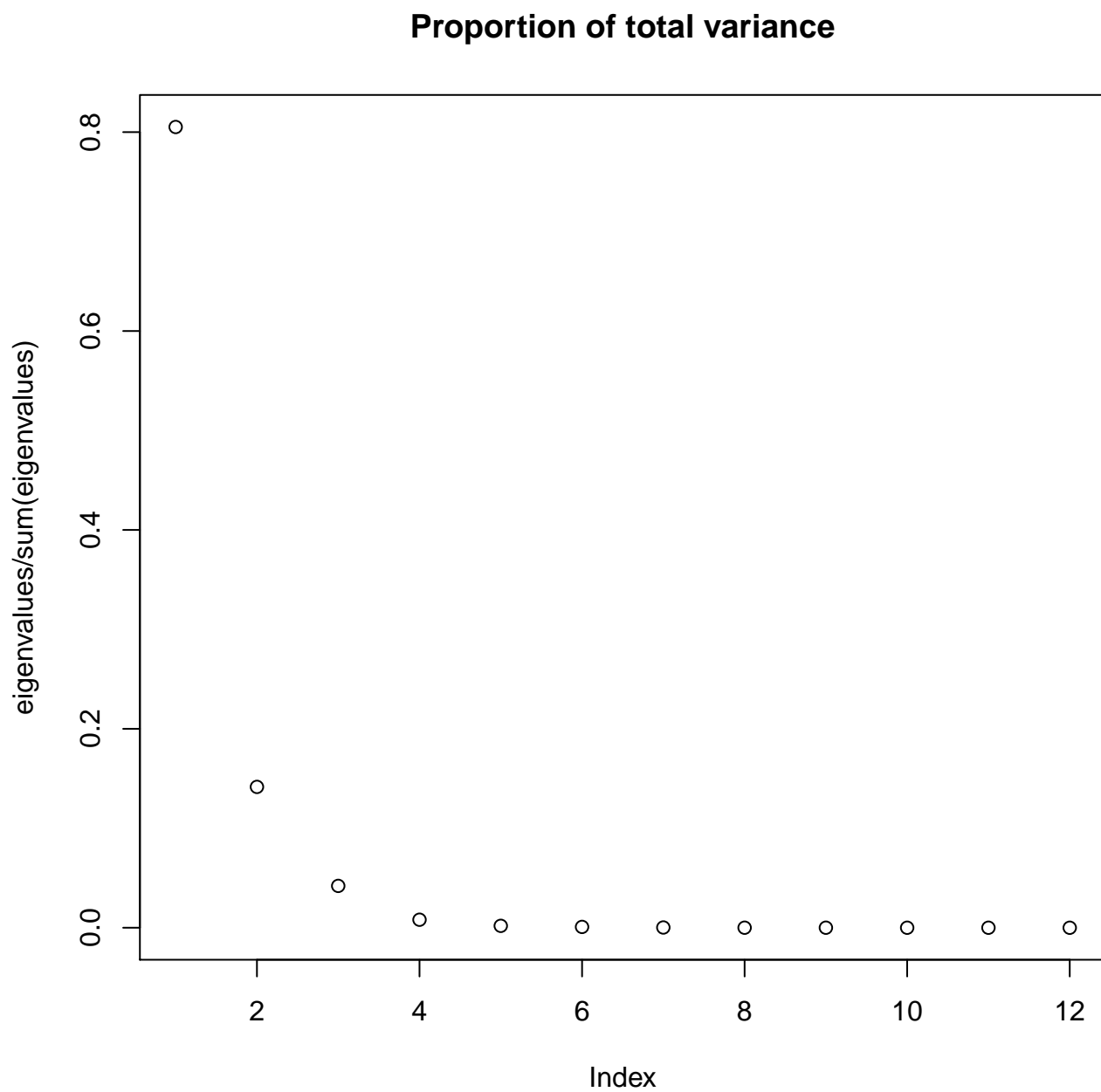FIGURE 9. Proportion of total variance explained by each eigen-value - male

Proportion of total variance

Figure 10. Cumulative proportion of total variance explained by
each eigenvalue - male

Cumulative proportion of total variance
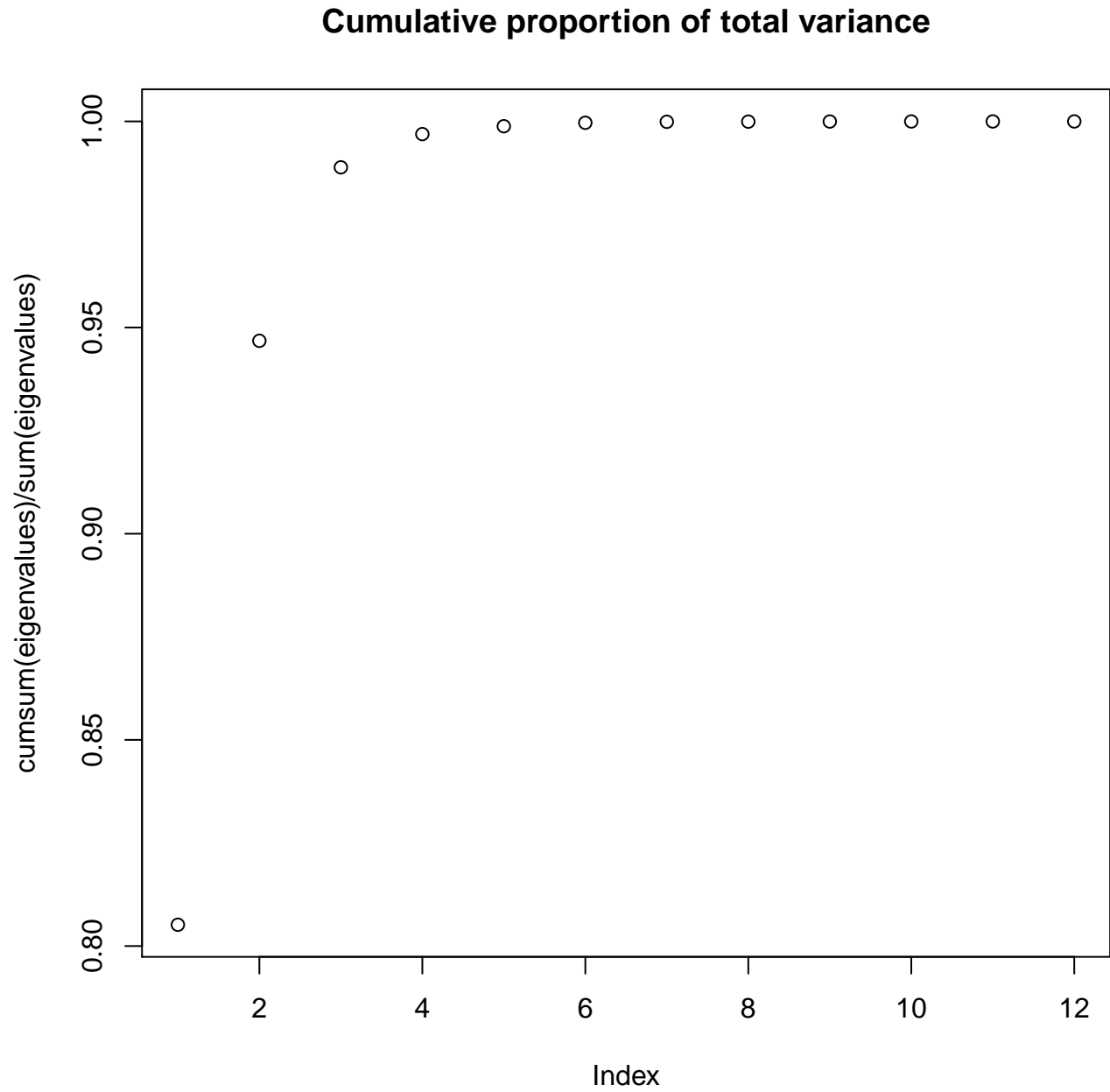
FIGURE 11. Parallel plot of first four PC scores - male
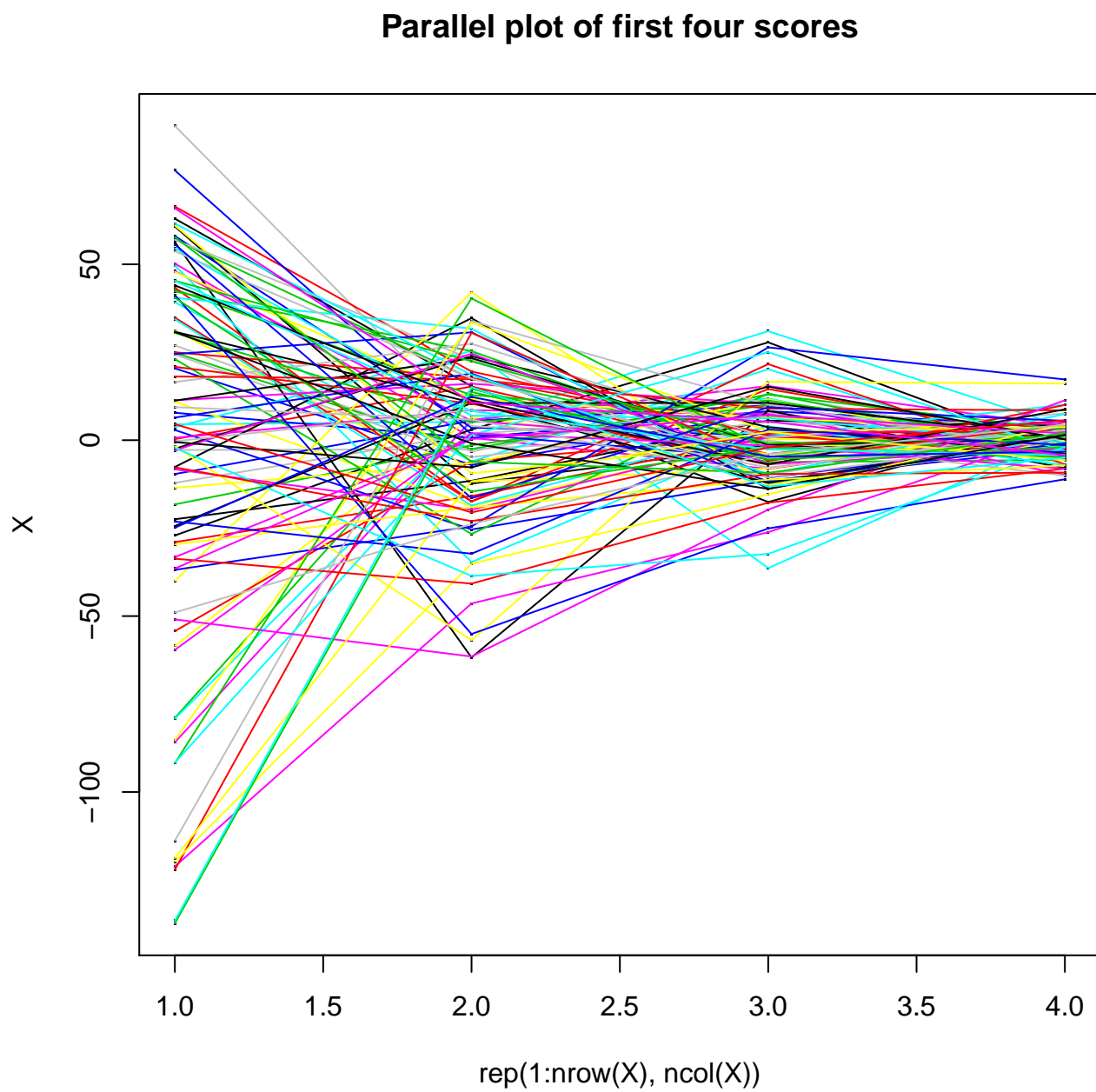

Parallel plot of first four scores

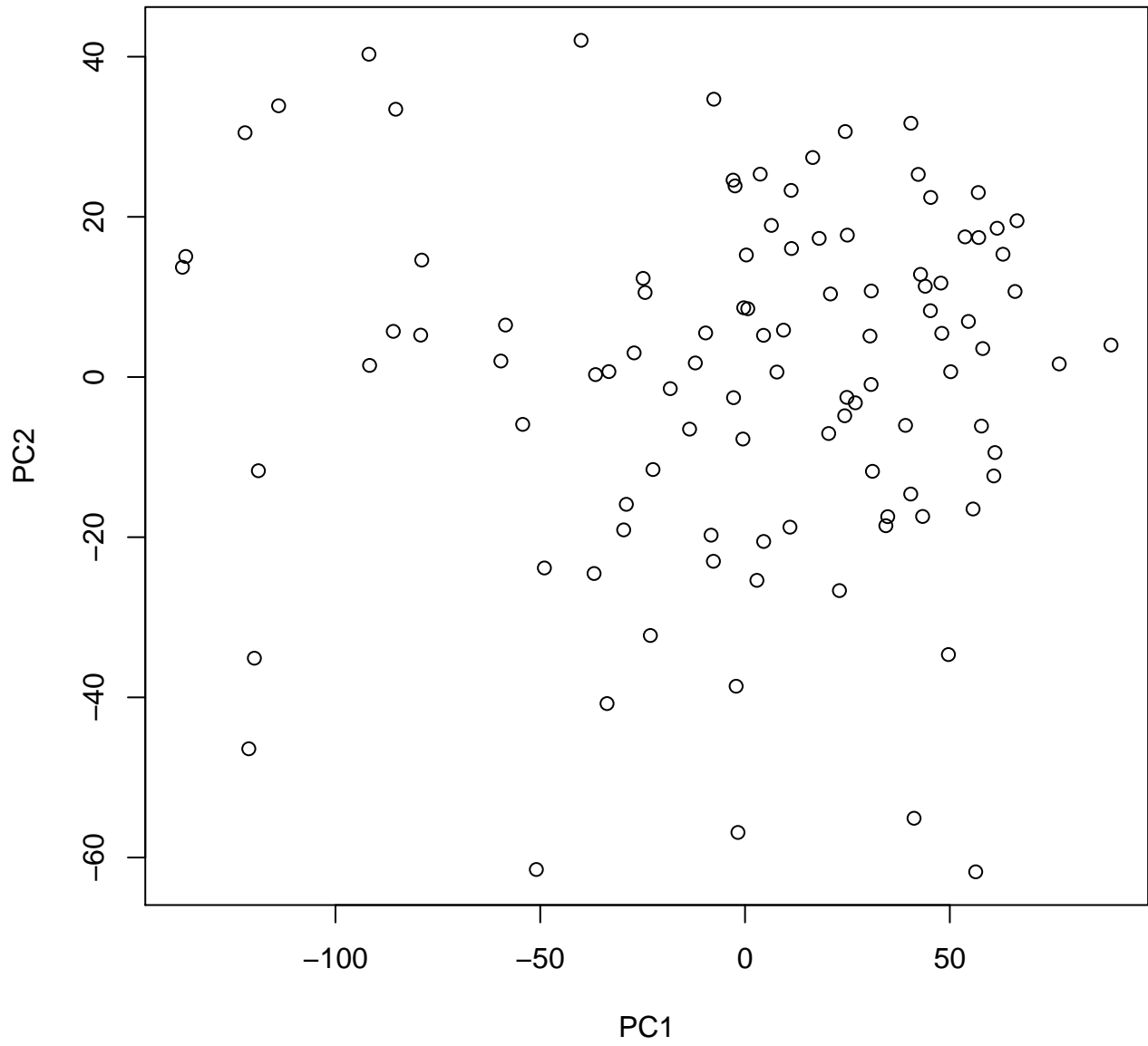Figure 12. Principal Component 1 versus Principal Component 2 - male

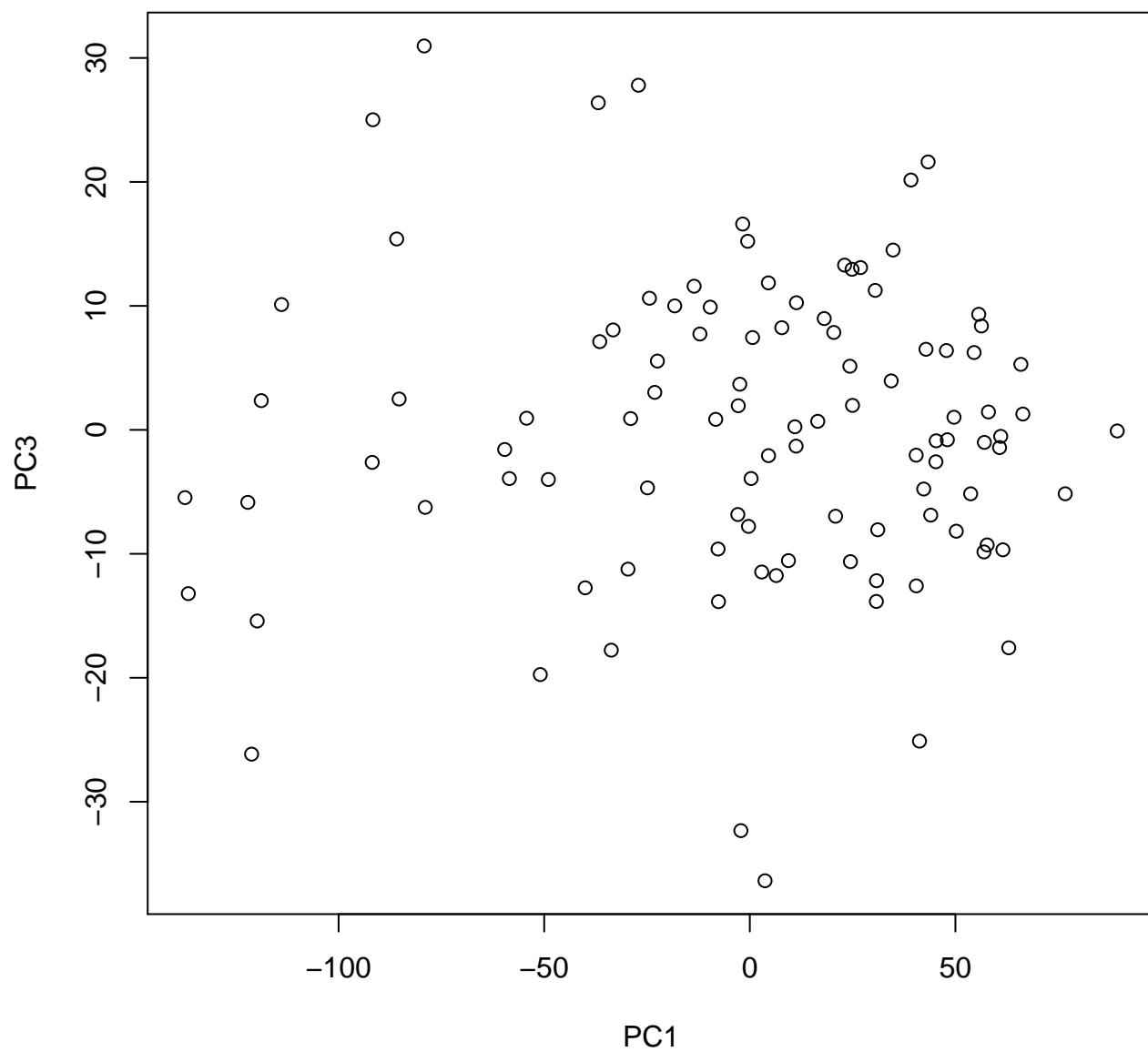FIGURE 13. Principal Component 1 versus Principal Component 3 - male

Figure 14. Principal Component 2 versus Principal Component 3 - male

FIGURE 15. Parallel plot of all variables - female
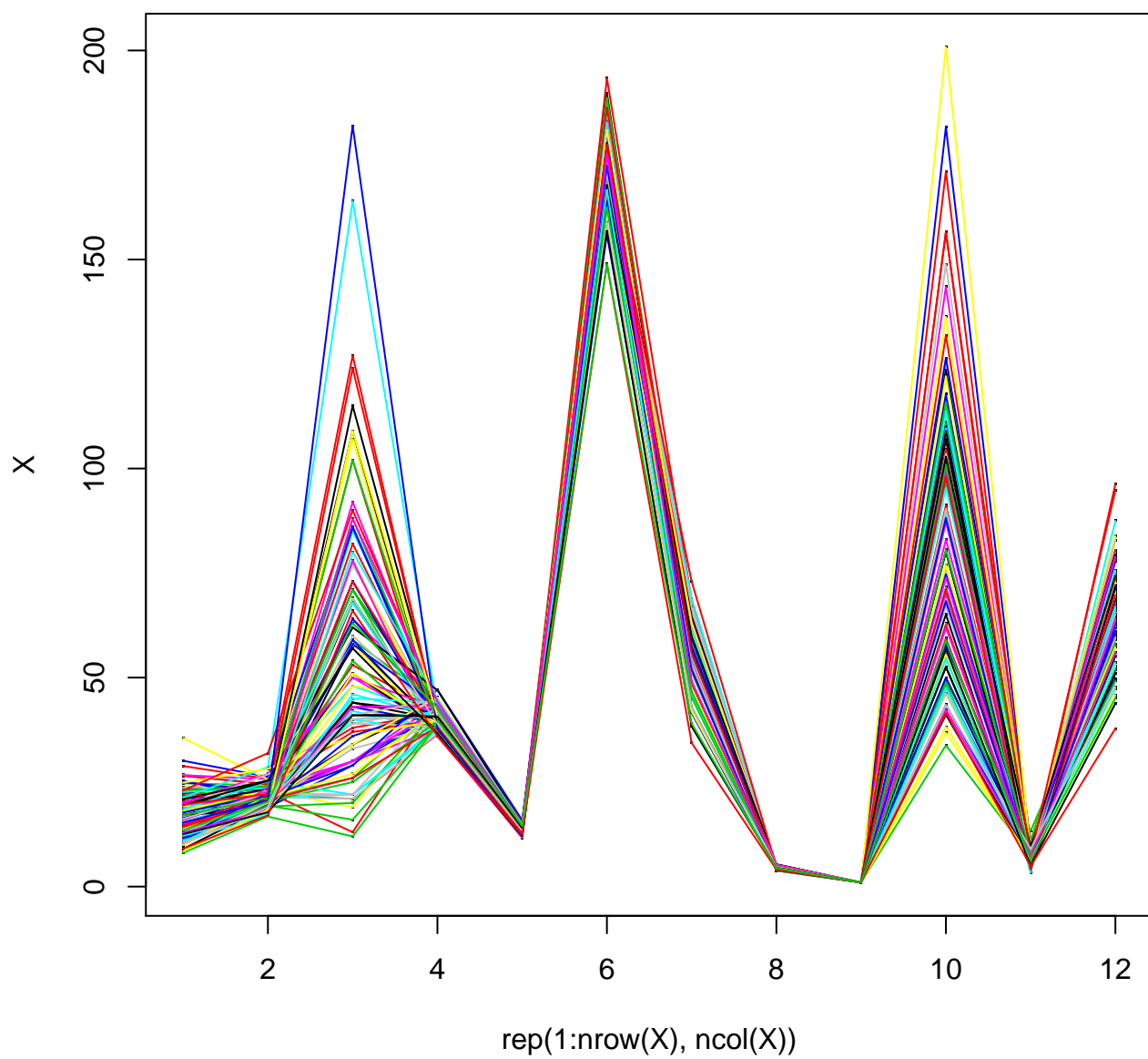
Figure 16. Proportion of total variance explained by each eigen-value - female

Proportion of total variance
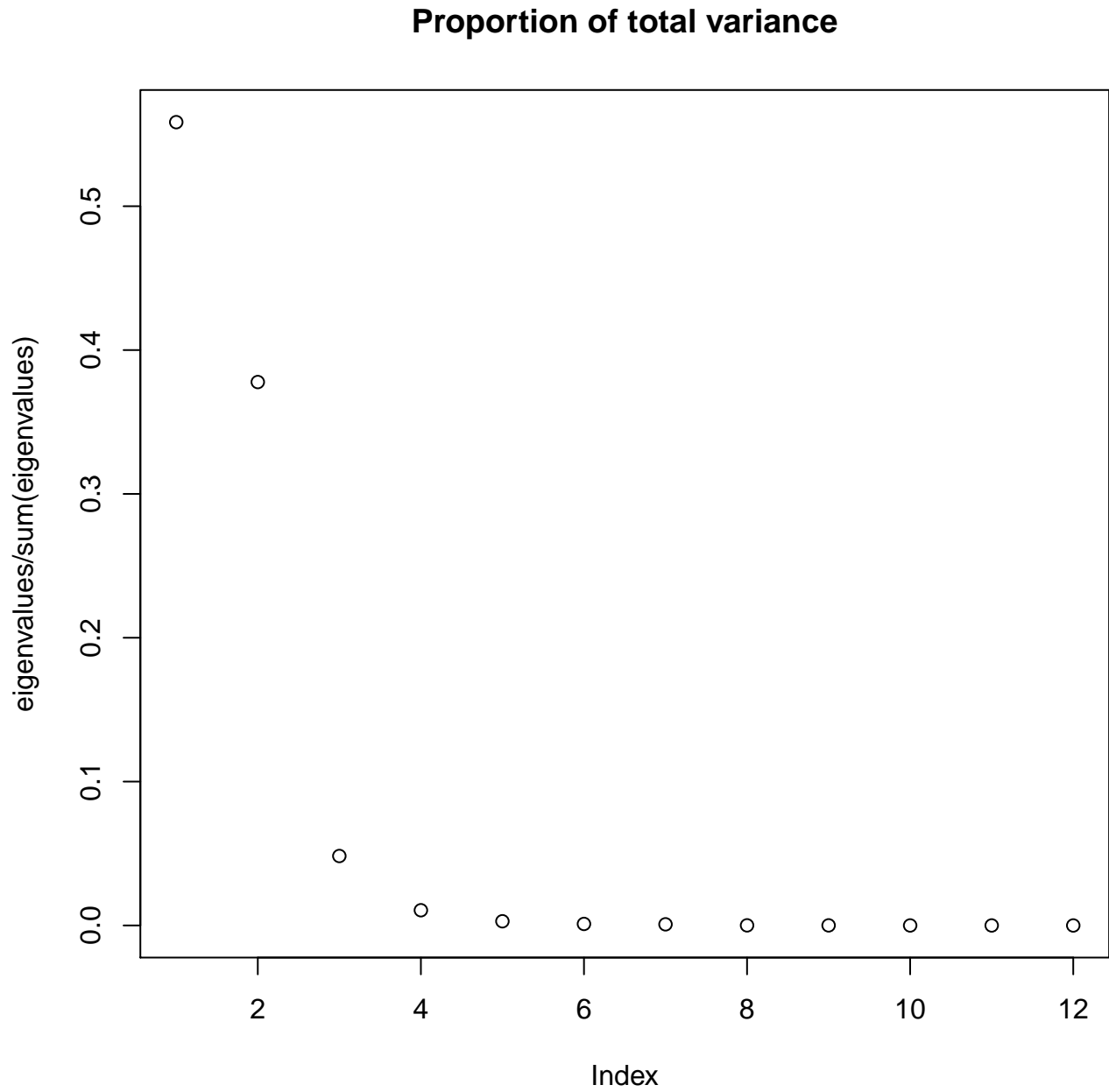
FIGURE 17. Cumulative proportion of total variance explained by each eigenvalue - female

Cumulative proportion of total variance

FIGURE 18. Parallel plot of first four PC scores - female

Parallel plot of first four scores

FIGURE 19. Principal Component 1 versus Principal Component 2 - female

Figure 20. Principal Component 1 versus Principal Component 3 - female

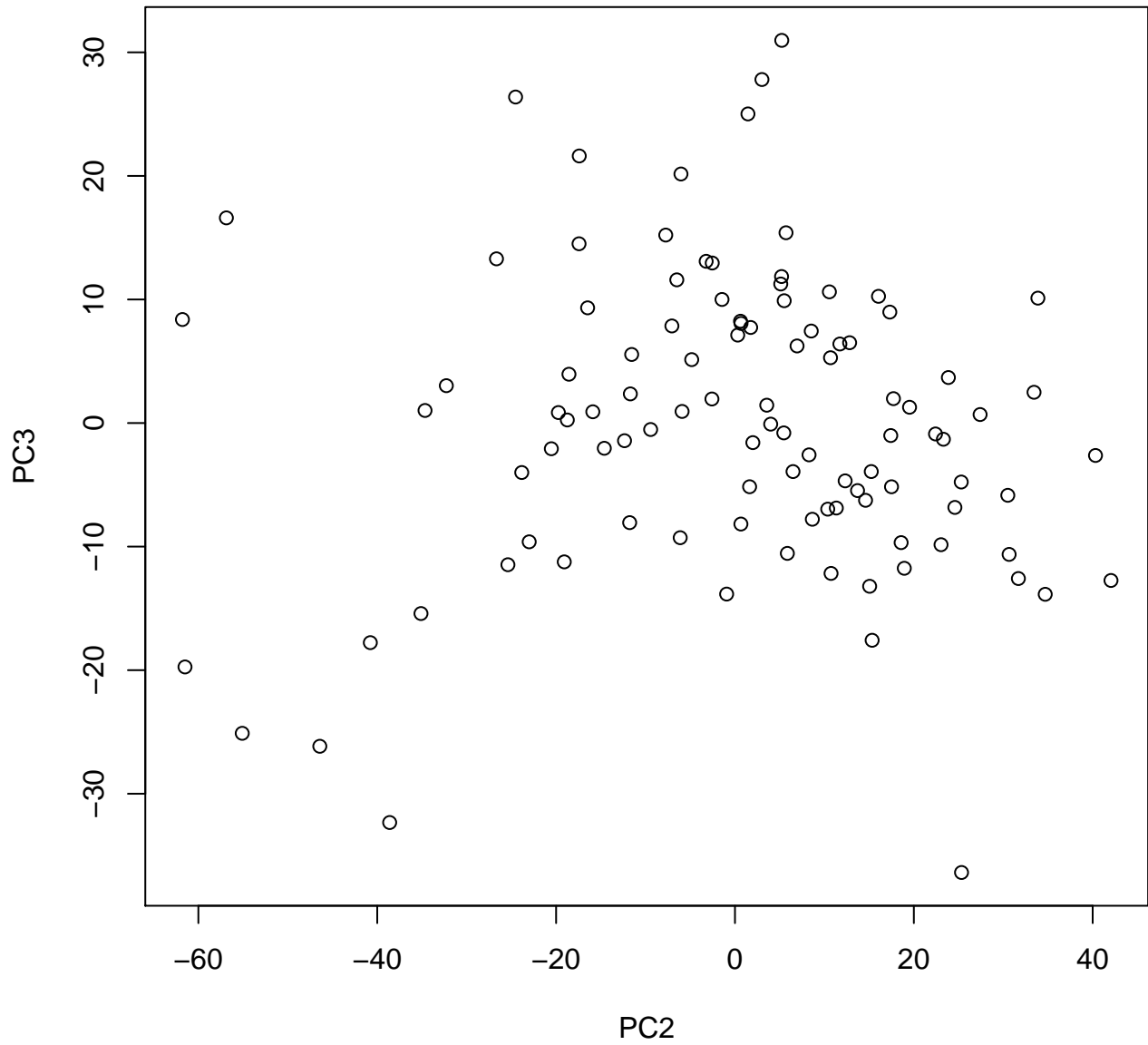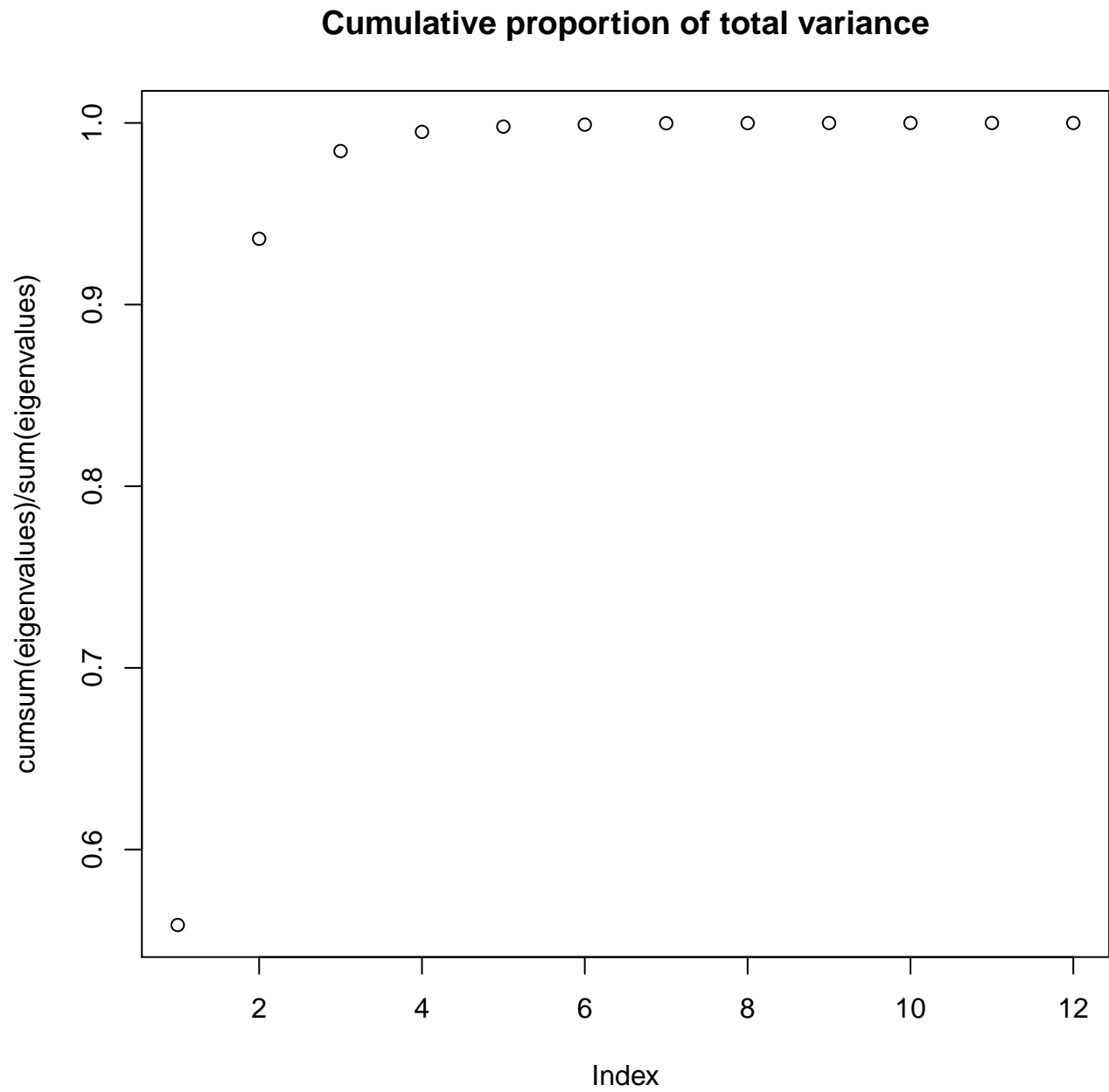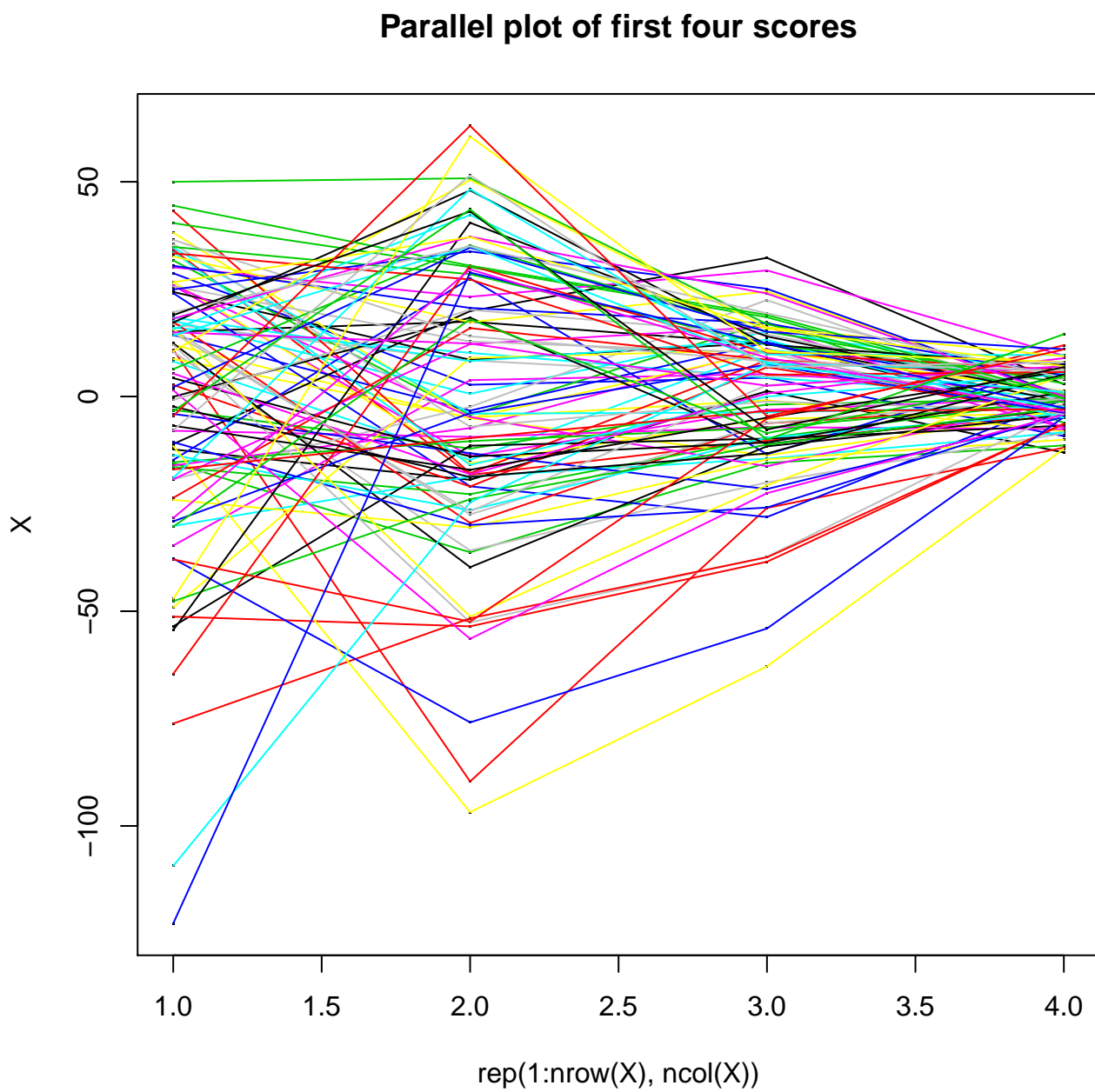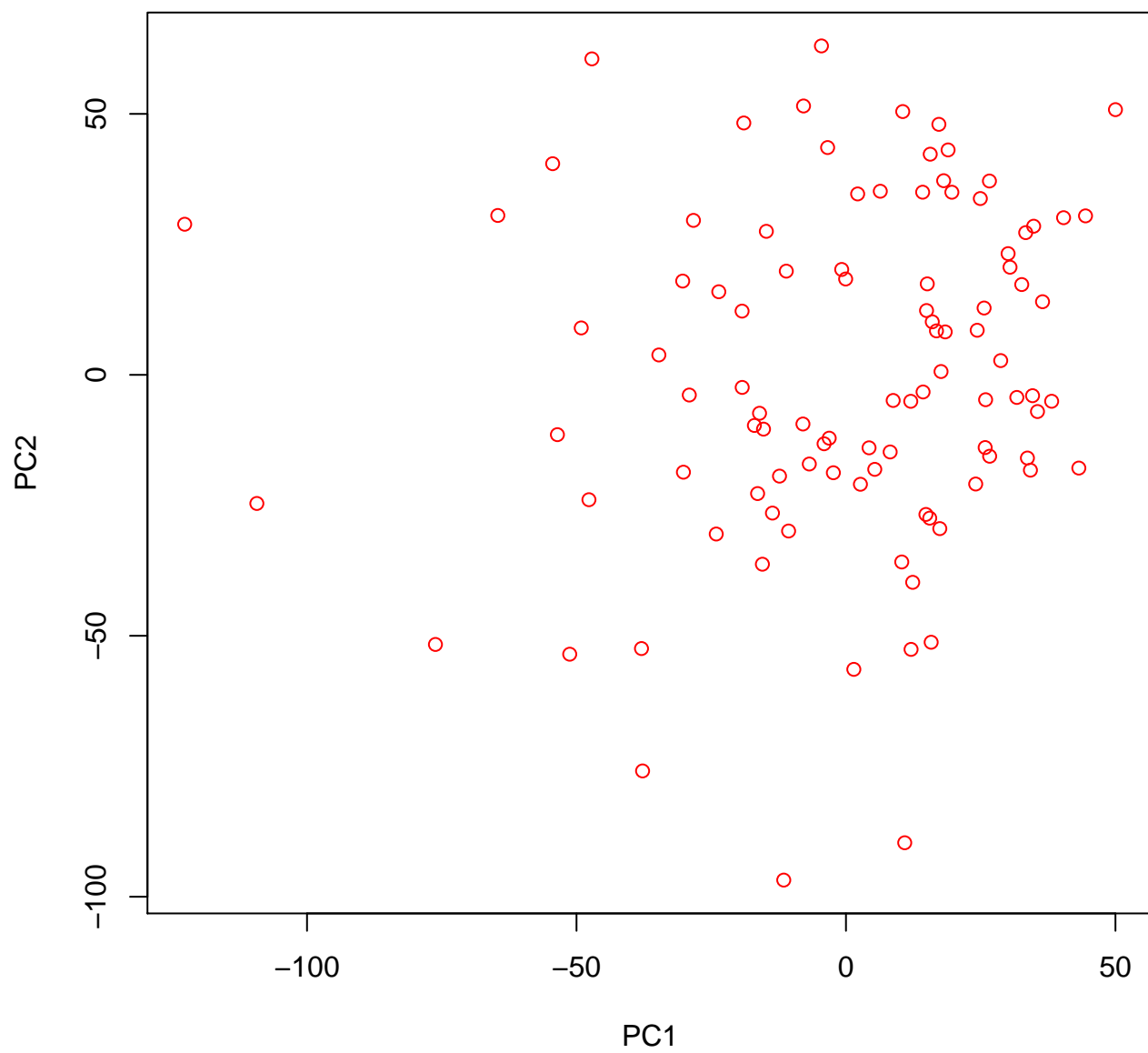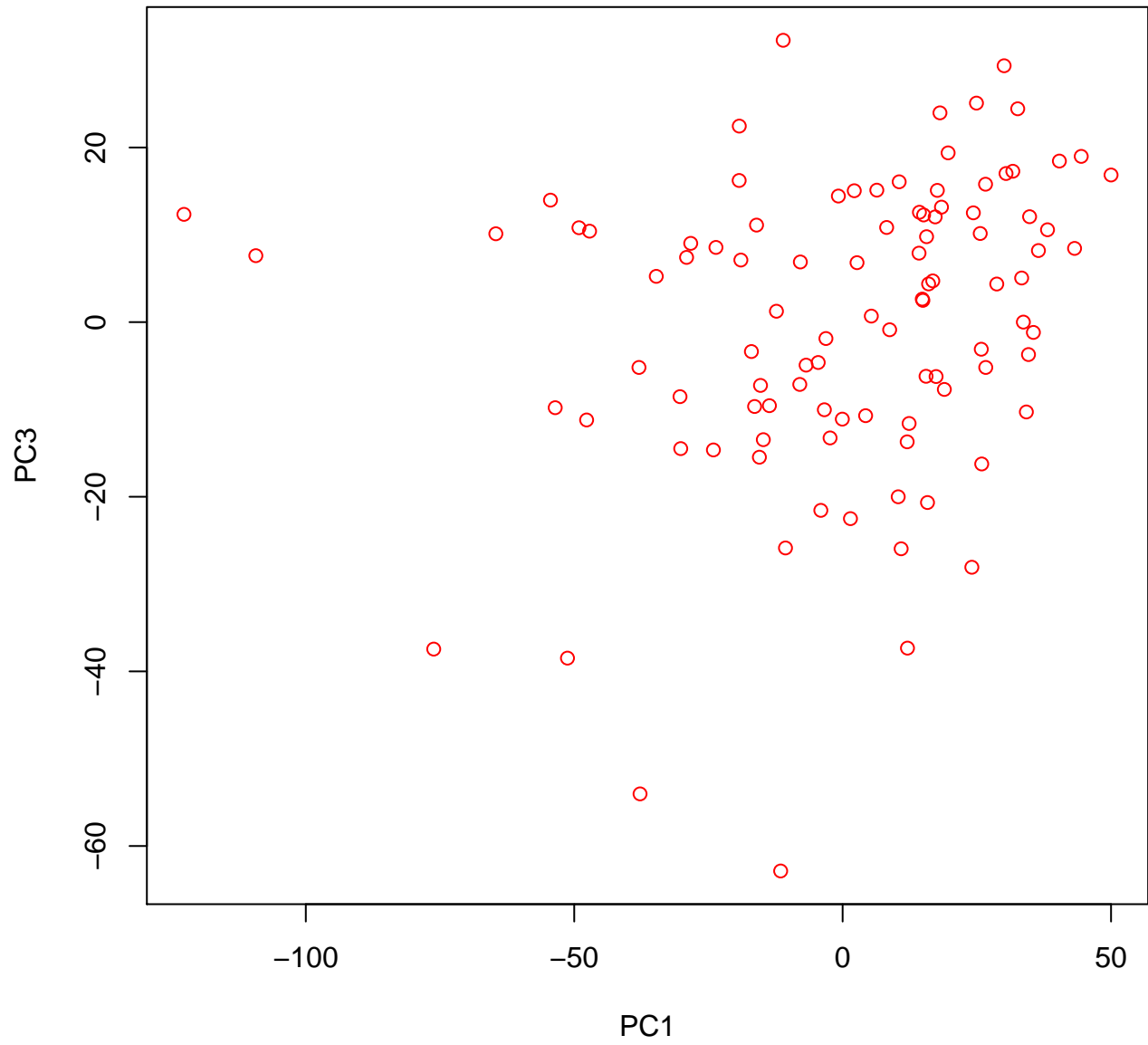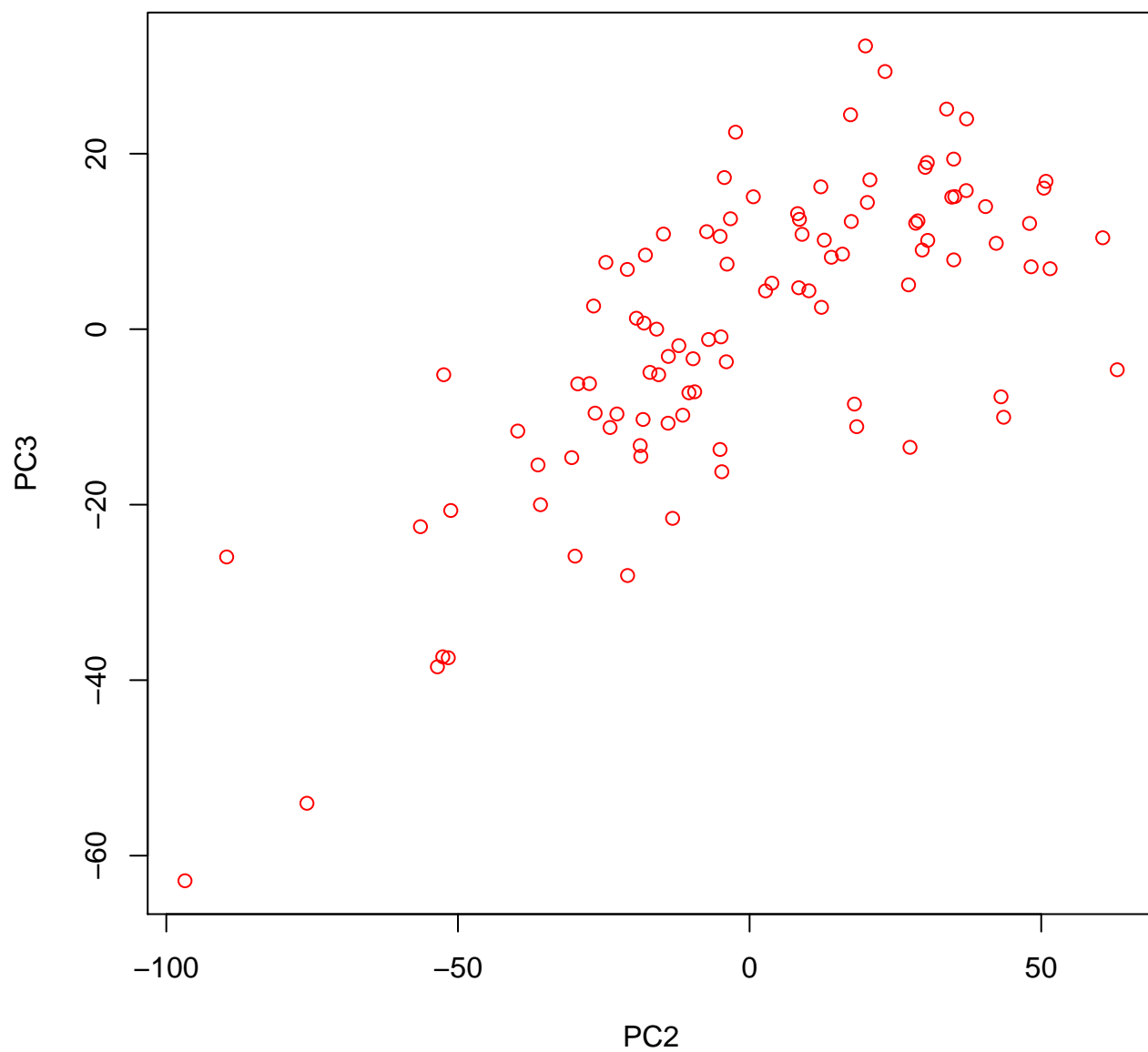FIGURE 21. Principal Component 2 versus Principal Component 3 - female

1.3. **c - Interpretation of results.** From the parallel plot for all variables of the entire data set, we see that some of the data sets are only very different scales. Most of the variance is explained by the first principle component. The parallel plot of the first four PC scores shows that the first few PC components distinguish the individuals well.

I have shown the males in the PC plots in black, and females in red. We can see that plotting the components against each other distinguishes the males from the females well - and one could almost place a linear decision boundary on the plots.

In the proportion of total variance plot for males, 80% of the variance is explained by the first PC component, and 95% is explained by the first two components. By comparison, only just over fifty percent of the variance in the females is explained by the first PC component, whereas just under 40% is explained by the second PC component. Comparing the male parallel PC score plot with the female parallel PC score plot, we see that for females the lines are more bunched at the first PC component, and more spread for the second component, consistent with what we observed above.

2. Code