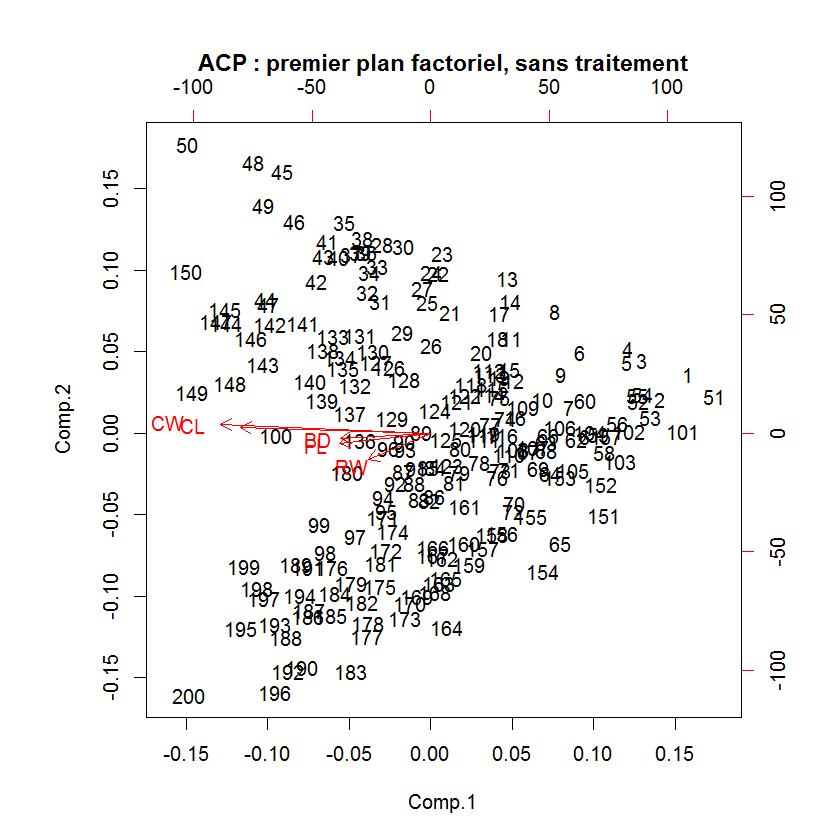
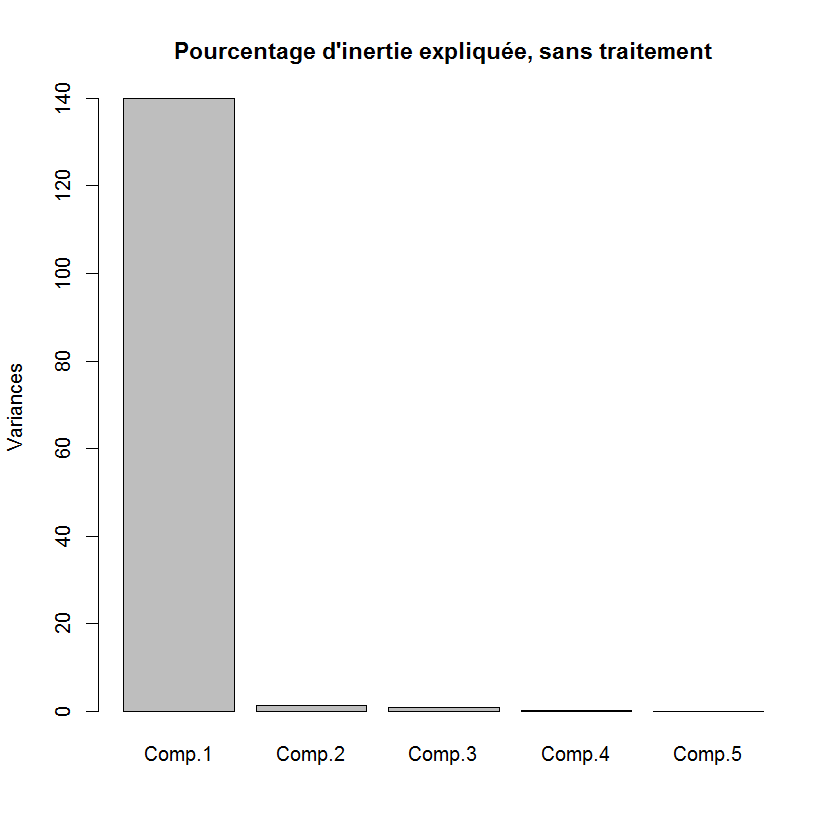
*Cette étude vise à utiliser l’ACP pour trouver une représentation des crabes qui permettent de distinguer visuellement différents groupes, liés à l’espèce et au sexe.*

Afin de compléter l’étude réalisée à l’exercice 1.2. sur le jeu de données CRABS, nous nous intéressons à l’utilisation de l’ACP pour visualiser les différents groupes.

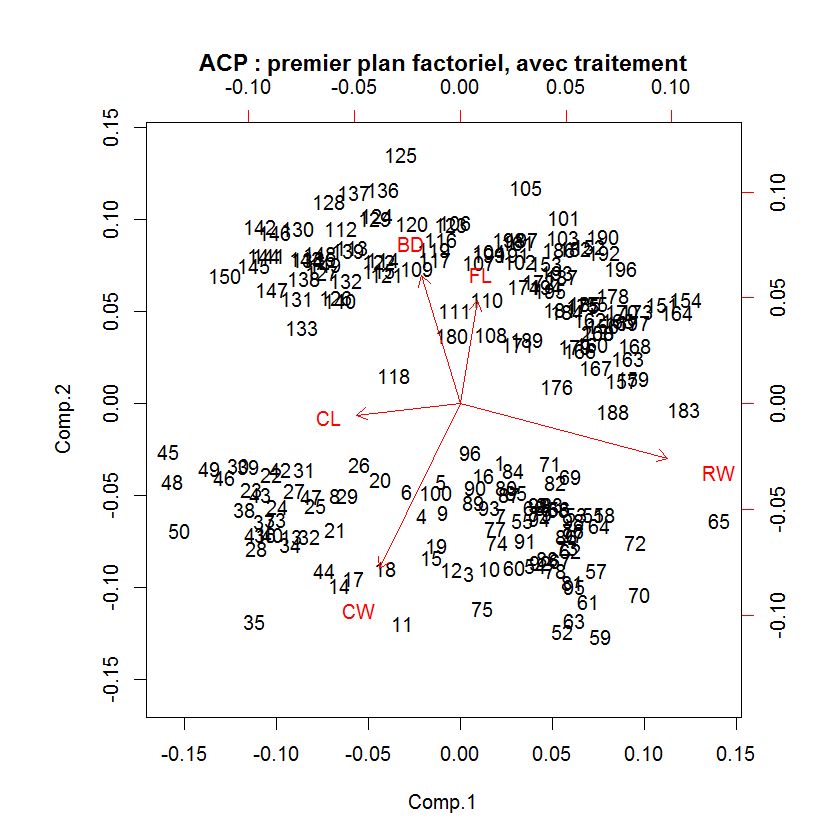
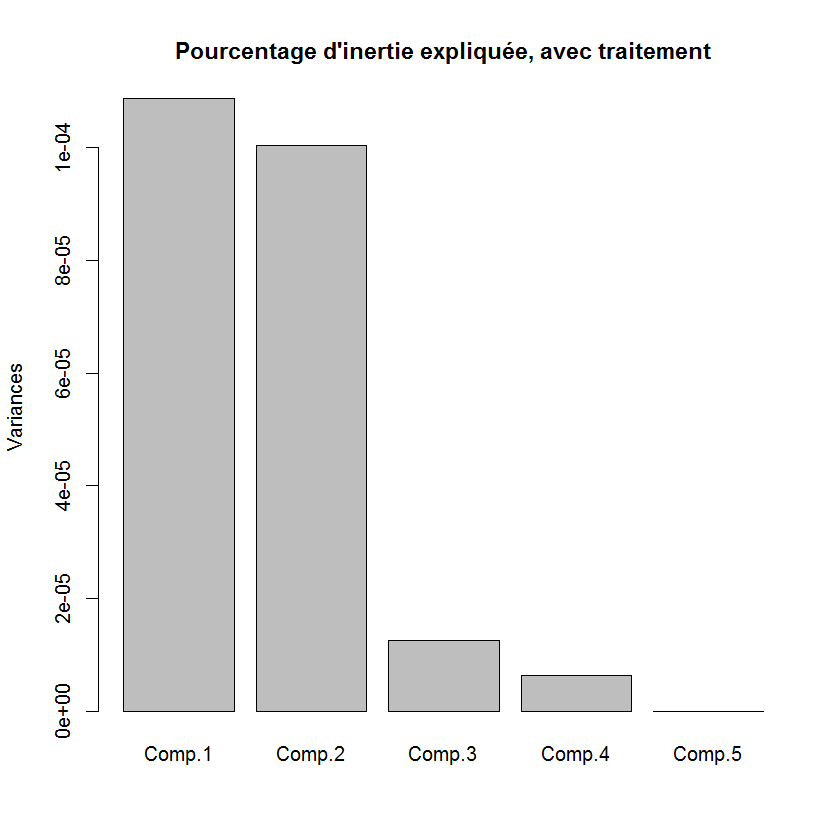
*1. Tester tout d’abord l’ACP sur crabsquant sans traitement préalable. Que constatez-vous ? Comment pouvez-vous expliquer ce phénomène à la lumière des analyses menées au paragraphe 1.2 ?*

L’ACP sur les données non traitées est réalisée sur R grâce à la fonction princomp préalablement introduite. La représentation biplot dans le premier plan factoriel et les pourcentages d’inertie expliquée pour chaque axe sont tracés dans la figure suivante.

Le biplot ne nous permet pas de distinguer de groupes, de plus les variables ont toutes la même direction. Ces observations indiquent que l’information n’est pas bien répartie sur les axes factoriels et que les variables sont très corrélées. Ce constat est confirmé par la valeur très importante du pourcentage d’inertie expliquée du premier axe factoriel et rejoint les observations réalisées lors de l’étude de l’exercice 1.2.

*2. Trouver une solution pour améliorer la qualité de votre représentation en termes de visualisation des différents groupes.*

Afin de rendre visible la présence de groupes parmi le jeu de données CRABS avec l’ACP il est possible d’appliquer à crasbsquant le même traitement que lors de l’analyse dans l’exercice 2.1, c’est-à-dire diviser la taille de chaque membre par la somme de tous les membres d’un individu. Une nouvelle ACP est réalisée sur les données traitées. La représentation biplot et les pourcentages d’inertie expliquée pour chaque axe sont tracés dans la figure suivante.

Les pourcentages d’inertie expliquée par les deux premiers axes factoriels sont respectivement de 47.65 et 44.00, ce qui représente 91.65 d’inerte expliquée par le premier plan factoriel. On distingue 4 groupes d’individus sur le biplot représentant le premier plan factoriel.