

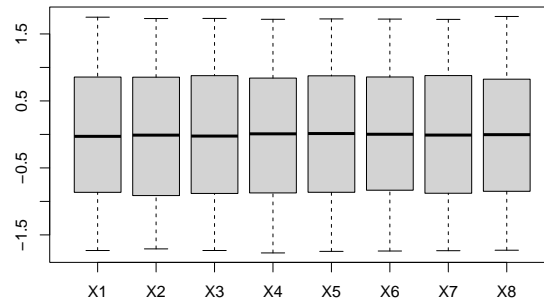
# Robotics

## Robotics dataset

Le jeu de données d'apprentissage de ce problème représente la cinématique du bras d'un robot. Les prédictors ainsi la réponse données sont toutes de type numérique. Le but est de trouver la relation de la réponse  $y$  et les 8 prédictors  $X1$ ,  $X2$ ,  $X3$ ,  $X4$ ,  $X5$ ,  $X6$ ,  $X7$ ,  $X8$ . Il s'agit bien d'un problème de régression.

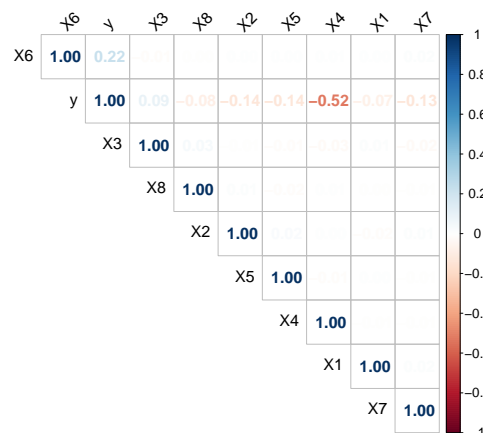
## Analyse exploratoire

Dans un premier temps, on essaie de regarder la plage de toutes les données.

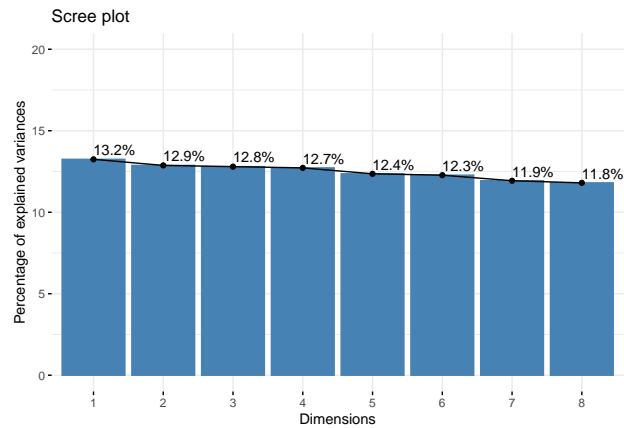


On observe que la plage de données est très homogène, ce qui nous donne la possibilité d'explorer nos données sans faire un scaling.

Ensuite on vérifie la corrélation entre les variables.



On constate qu'il existe pas de corrélations significatives entre les variables, seule de faibles corrélations entre la réponse et les prédicteurs. Cette graphe exclut le besoin d'enlever certaines variables puisqu'elles sont peu corrélées. On confirme cette observation en faisant une ACP



On constate avec cette graphe que la variance est expliquée par toutes les variables. Et on en déduit que toutes les méthodes concernant **Subset Selection** eront pas utiles vu que les variables devraient toutes être incluses.

## Sélection de modèle

On test sur le jeu de données **Robotics** les méthodes suivantes, que l'on a vues en cours:

- Modèle linéaire - La méthode des **KNN**