BACS - HW (Week 14)

Let's reconsider the security questionnaire from last week, where consumers were asked security related questions about one of the e-commerce websites they had recently used.

Question 1) Earlier, we examined a dataset from a security survey sent to customers of e-commerce websites. However, we only used the eigenvalue > 1 criteria and the screeplot "elbow" rule to find a suitable number of components. Let's perform a parallel analysis as well this week:

a. Show a single visualization with scree plot of data, scree plot of simulated noise (use average eigenvalues of ≥ 100 noise samples), and a horizontal line showing the eigenvalue = 1 cutoff.
b. How many dimensions would you retain if we used Parallel Analysis?

Question 2) Earlier, we treated the underlying dimensions of the security dataset as composites and examined their eigenvectors (weights). Now, let's treat them as factors and examine *factor loadings* (use the `principal()` method from the `psych` package)

a. Looking at the loadings of the first 3 principal components, to which components does each item seem to best belong?
b. How much of the total variance of the security dataset do the first 3 PCs capture?
c. Looking at commonality and uniqueness, which items are *less than adequately* explained by the first 3 principal components?
d. How many measurement items share similar loadings between 2 or more components?
e. Can you interpret a 'meaning' behind the first principal component from the items that load best upon it? (see the wording of the questions of those items)

Question 3) To improve interpretability of loadings, let's rotate our principal component axes using the *varimax* technique to get *rotated components* (extract and rotate only three principal components)

a. Individually, does each rotated component (RC) explain the same, or different, amount of variance than the corresponding principal components (PCs)?
b. Together, do the three rotated components explain the same, more, or less cumulative variance as the three principal components combined?
c. Looking back at the items that shared similar loadings with multiple principal components (#2d), do those items have more clearly differentiated loadings among rotated components?
d. Can you now more easily interpret the "meaning" of the 3 rotated components from the items that load best upon each of them? (see the wording of the questions of those items)
e. If we reduced the number of extracted and rotated components to 2, does the meaning of our rotated components change?

(ungraded) Looking back at all our results and analyses of this dataset (from this week and previous), how many components (1-3) do you believe we should extract and analyze to understand the security dataset? Feel free to suggest different answers for different purposes.