Week 4  
**Activity**

Open your preferred Data Mining Software (Weka, R, SAS, SPSS, etc), load the “cities” sample[[1]](#footnote-1).

Answer the following questions or follow the instructions. Answer using screenshots and paragraphs explaining what you do and why.

1. Don’t use or delete the “City” attribute in the next questions. Why do you have to do this?
2. Do you see any relation between pairs of attributes in the data?
3. Are there any correlations between pairs of attributes in the data?
4. Apply PCA to the dataset[[2]](#footnote-2) and,
   1. Explain every step that you follow
   2. Explain every output of your steps
5. Now that you have done all of the previous points, if you have to do a model[[3]](#footnote-3) with this data, how will PCA help you with the model?

1. Is in Webcursos [↑](#footnote-ref-1)
2. Did you standardize the attributes?, why? [↑](#footnote-ref-2)
3. We haven’t learned any data mining models, yet. Next class we will see our first model! [↑](#footnote-ref-3)