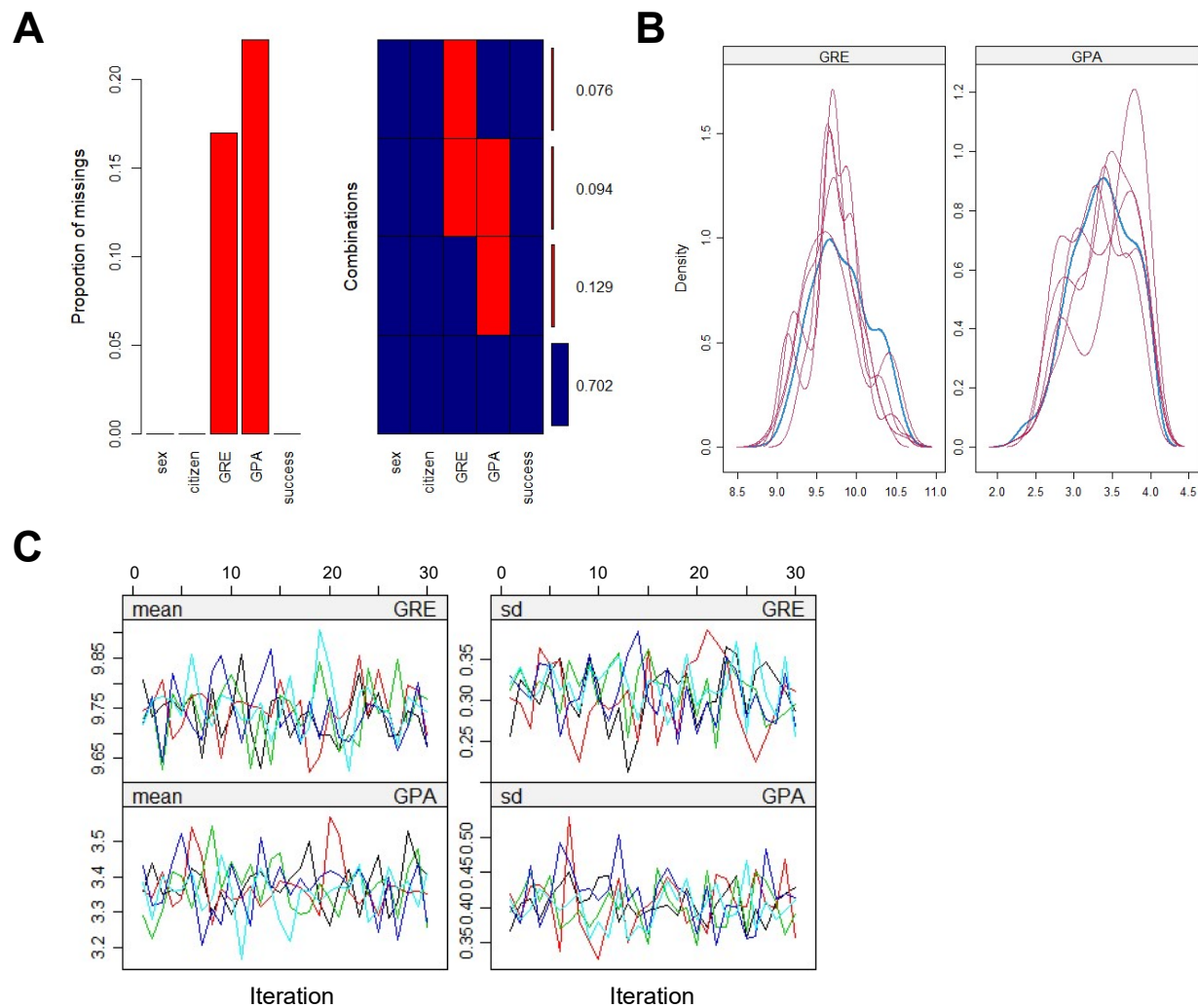
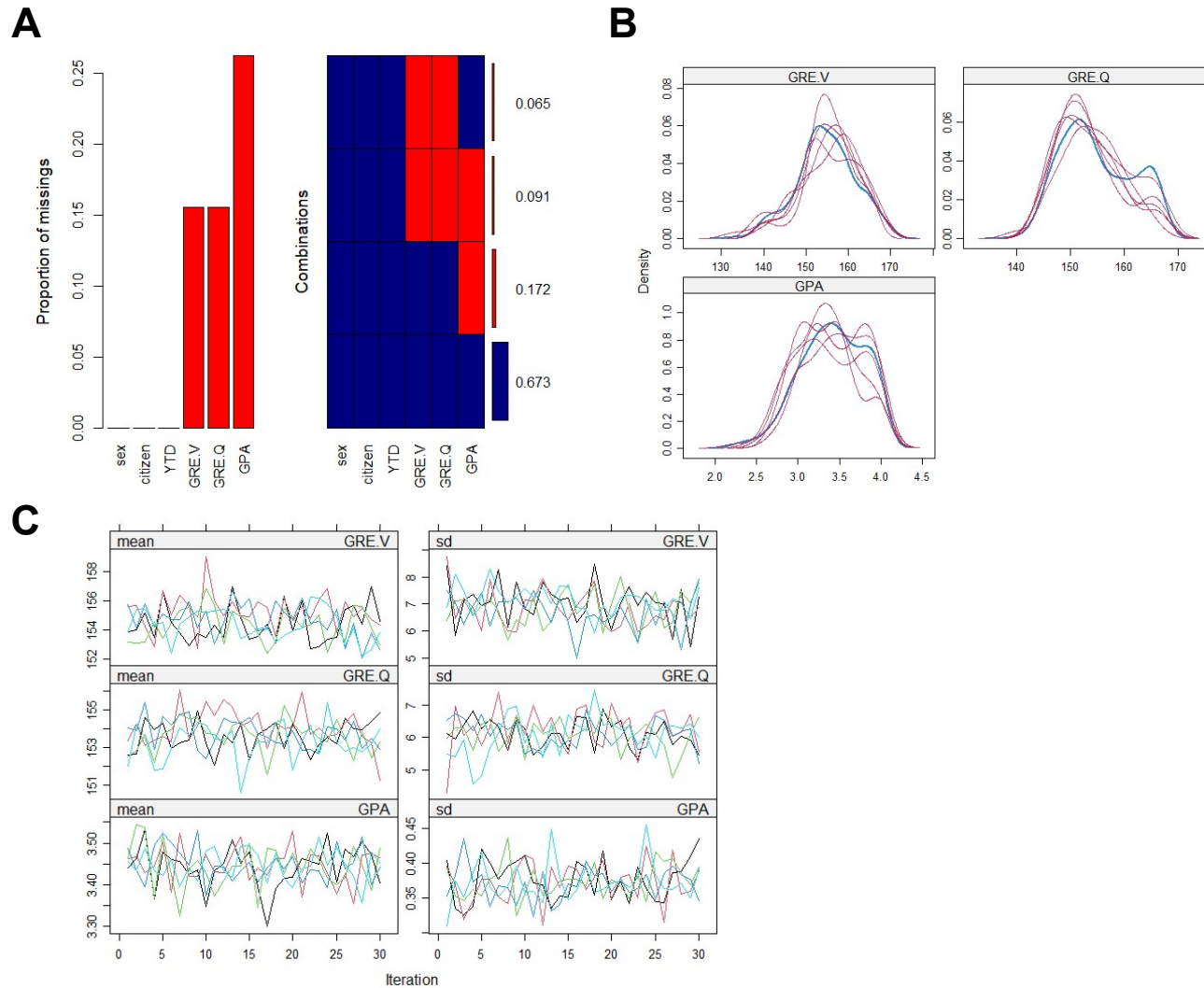


# Supplemental Material

*CBE—Life Sciences Education*  
Mendoza-Sanchez *et al.*



**Supplementary Figure 1.** Imputing missing data of all enrolled students, using the mice R language package. **A**, Fraction of students missing each of the admission variables we considered. The plot was generated with the VIM R language package. **B**, Density plots for each imputed dataset is in magenta while the density of the observed data is in blue. **C**, Mean and standard deviation (sd) of the synthetic values plotted against iteration number for each imputed dataset.



**Supplementary Figure 2.** Imputing missing data of all students that completed their Ph.D., using the mice R language package. **A**, Fraction of students missing each of the admission variables we considered. The plot was generated with the VIM R language package. **B**, Density plots for each imputed dataset is in magenta while the density of the observed data is in blue. **C**, Mean and standard deviation (sd) of the synthetic values plotted against iteration number for each imputed dataset.