Datasets

College data

College Data	
Variable Name	Description
Private	Private or public university
Apps*	Number of applicants received
Accept*	Number of applications accepted
Enroll*	Number of new students enrolled
Top10perc	Pct. new students from top 10% of high school class
Top25perc	Pct. new students from top 25% of high school class
F.Undergrad*	Number of full-time undergrads
P.Undergrad*	Number of part-time undergrads
Outstate*	Out-of-state tuition
Room.Board*	Room and board costs
Books*	Estimated book costs
Personal*	Estimated personal spending
PhD	Pct. of faculty with Ph.D.'s
Terminal	Pct. of faculty with terminal degree
S.F.Ratio	Student/faculty ratio
perc.alumni	Pct. alumni who donate
Expend*	Instructional expenditure per student
Grad.Rate	Graduation rate

Description of the college data. Variables marked with an asterisk have been log transformed.

Cervical cancer risk data

Cervical cancer Data	
Variable Name	Description
Age*	Age in years
No_sex_part*	Number of sexual partners
First_sex_inter*	Age, in years, of first sexual intercourse
No_preg*	Number of pregnancies
Smokes_yrs*	Number of years smoking
Smokes_pcks_yr*	Number of packs smoked in a year
Horm_Cont_yrs*	Hormonal contraceptives (in years)
IUD_yrs*	Number of years with an intrauterine device
STDs_No*	Number of STD diagnoses
STDs_condy	STDs: condylomatosis
STDs_syph	STDs: syphilis
STDs_HIV	STDs: HIV
Dx.Cancer	Previous cancer diagnosis
Dx.HPV	Previous HPV diagnosis
Biopsy	Biopsy results, either Healthy or Cancer

Description of the cervical cancer data. Variables marked with an asterisk have been $\log(x+1)$ transformed.