UNIT 8 HW

On this upload page you will submit unit 8 homework for grading.

Live Session Problem

Please use SAS or R for the scatterplot in question 1 below. For question 2, find the correlation by hand and use SAS or R to check your calculation. Use SAS for question 3.

Americans love their baseball. Even so, there is concern that teams with more money to spend on players have more success than teams that have less money to spend. The payroll (X) (in millions of dollars) and the number of games won in the season (out of 162) (Y) are provided in the table below for the 30 major league teams. The numbers are for the 2010 regular season. We can use these data to illustrate statistical methods for drawing inferences about correlations.

Team	Payroll	Wins	Team	Payroll	Wins	Team	Payroll	Wins
NYY	206	95	LAD	95	80	KC	71	67
BOS	162	89	HOU	92	76	TOR	62	85
CHC	146	75	SEA	86	61	ARZ	61	65
PHI	142	97	STL	86	86	CLE	61	69
NYM	134	79	ATL	84	91	WAS	61	69
DET	123	81	COL	84	83	FA	57	80
CHW	106	88	BAL	82	66	TEX	55	90
LAA	105	80	MIL	81	77	OAK	52	81
SF	99	92	TB	72	96	SD	38	90
MIN	98	94	CIN	71	91	PIT	35	57

1. Draw a scatterplot of the data **using software (required!)**. Looking at the scatterplot, do you expect the correlation to be positive, negative, or close to 0? Why? Is the relationship between team payroll and number of wins strong, moderate or weak? Is the relationship linear? Take a guess at the value of the correlation coefficient.

2. Find the correlation between team payroll and the number of wins – no fair going back and changing your answer to the previous question! (SAS Note: you can get correlations from SAS using PROC CORR; VAR X1 X2 X3;, where X1, X2, and X3 are the variables for which you want correlations. This does all combinations of pairwise correlations.)

3. San Diego (SD) has a payroll of 38 million, yet has 90 wins – more than Boston does. Delete SD from the data and rerun the analysis. You can use SAS. How does the correlation change?

4. The league commissioner notes that the Texas Rangers with one of the lowest payrolls won 90 games (and were the American League Champions) and the Chicago Cubs with the third highest payroll won only 75 games. This, he argues, proves there is no advantage to teams with a higher payroll. Comment on his argument.

5. What is the population for these data? Can these data be considered as a random sample from that population?