Lab Assignment #9

This lab is due at 9:35 AM on Wednesday 2/21 and is worth 6 points. This may be done individually, or in a group of 2 or 3 people.

Use the attached data sheet from the 2022 MLB season, all American League teams, for all problems in this lab.

1) Use home runs (HR) for your *x*-variable and runs (R) for your *y*-variable.

a) Make a scatterplot.
b) Find the equation of best-fit line. It is not necessary to plot the line on your scatterplot, but you may if you want to.
c) Find the correlation coefficient *r*. Describe what your value of *r* means.
d) Find the coefficient of determination r^2. Write this number in a sentence along with its meaning.

2) Use batting average (AVG) for your *x*-variable and runs (R) for your *y*-variable. Use three-digit whole numbers for AVG instead of the decimal numbers provided.

a) Make a scatterplot.
b) Find the equation of best-fit line. It is not necessary to plot the line on your scatterplot, but you may if you want to.
c) Find the correlation coefficient *r*. Describe what your value of *r* means.
d) Find the coefficient of determination r^2. Write this number in a sentence along with its meaning.

3) Use on-base percentage (OBP) for your x-variable and runs (R) for your y-variable.

a) Make a scatterplot. Use three-digit whole numbers for OBP instead of the decimal numbers provided.
b) Find the equation of best-fit line. It is not necessary to plot the line on your scatterplot, but you may if you want to.
c) Find the correlation coefficient r. Describe what your value of r means.
d) Find the coefficient of determination r^2. Write this number in a sentence along with its meaning.

4) Use slugging percentage (SLG) for your *x*-variable and runs (R) for your *y*-variable.

a) Make a scatterplot. Use three-digit whole numbers for SLG instead of the decimal numbers provided.
b) Find the equation of best-fit line. It is not necessary to plot the line on your scatterplot, but you may if you want to.
c) Find the correlation coefficient *r*. Describe what your value of *r* means.
d) Find the coefficient of determination r^2. Write this number in a sentence along with its meaning.

5) Use the results from Q1. On average, how many runs is each home run worth? (See top of p.3 of lecture notes.)

6) Use the results from Q1-Q4. Of home runs, batting average, on-base percentage, and slugging percentage, which two of these correlate most closely with number of runs scored?

7) a) Use the results from Q3. A point of OBP is worth about _____ runs. (From p. 4 of the lecture notes.)

b) Use the results from Q4. A point of SLG is worth about _____ runs.

c) Use the results from Q7a and Q7b. A point of OBP is worth about _____ times as much as a point of SLG.

8) Did you notice that the 2022 MLB American League results once again confirm the results of Moneyball discussed in the lecture? That is, the two most important stats to predict runs are OBP and SLG, and that a point of OBP is worth about 2 or 3 times as much as a point of SLG. (This is a Y/N question.)

Team	R	HR	AVG	OBP	SLG
New York Yankees	807	254	0.241	0.325	0.426
Toronto Blue Jays	775	200	0.264	0.329	0.431
Houston Astros	737	214	0.248	0.319	0.424
Boston Red Sox	735	155	0.258	0.321	0.409
Texas Rangers	707	198	0.239	0.301	0.395
Cleveland Guardians	698	127	0.254	0.316	0.383
Minnesota Twins	696	178	0.248	0.317	0.401
Seattle Mariners	690	197	0.23	0.315	0.39
Chicago White Sox	686	149	0.256	0.31	0.387
Baltimore Orioles	674	171	0.236	0.305	0.39
Tampa Bay Rays	666	139	0.239	0.309	0.377
Kansas City Royals	640	138	0.244	0.306	0.38
Los Angeles Angels	623	190	0.233	0.297	0.39
Oakland Athletics	568	137	0.216	0.281	0.346
Detroit Tigers	557	110	0.231	0.286	0.346