--------------------------------------------------------------------------

.

. use data/baseballsalaries.dta

.

. describe

Contains data from data/baseballsalaries.dta

Observations: 30

Variables: 5 18 Jan 2012 19:34

--------------------------------------------------------------------------

Variable Storage Display Value

name type format label Variable label

--------------------------------------------------------------------------

team str12 %12s Team

playoffs byte %8.0g 1=made playoffs, 0=did not make

playoffs

payroll double %12.0g Payroll in dollars

wins int %8.0g Wins

losses int %8.0g Losses

--------------------------------------------------------------------------

Sorted by:

. summarize

Variable | Obs Mean Std. dev. Min Max

-------------+---------------------------------------------------------

team | 0

playoffs | 30 .2666667 .4497764 0 1

payroll | 30 42170.28 17830.28 8317.5 73995.92

wins | 30 81 13.52902 54 114

losses | 30 81 13.45491 48 108

.

. \*\*\*\*Qa) What is the average payroll for these thirty teams? Is there a lot of variation in payroll or minimal variation in payroll?

. summarize payroll

Variable | Obs Mean Std. dev. Min Max

-------------+---------------------------------------------------------

payroll | 30 42170.28 17830.28 8317.5 73995.92

.

. display "Average payroll (in thousands) = " `r(mean)'

Average payroll (in thousands) = 42170.282

.

. \*\*\*\*Qb) What is the average percentage of games won by a team?

. generate total\_games = wins + losses

. generate win\_percentage = (wins/total\_games)\*100

.

. label var total\_games "Total games won by a team"

. label var win\_percentage "Average percentage of games won"

.

. sum win\_percentage

Variable | Obs Mean Std. dev. Min Max

-------------+---------------------------------------------------------

win\_percen~e | 30 49.99632 8.327284 33.33333 70.37037

. di "Percentage of games won by a team = " `r(mean)'

Percentage of games won by a team = 49.996319

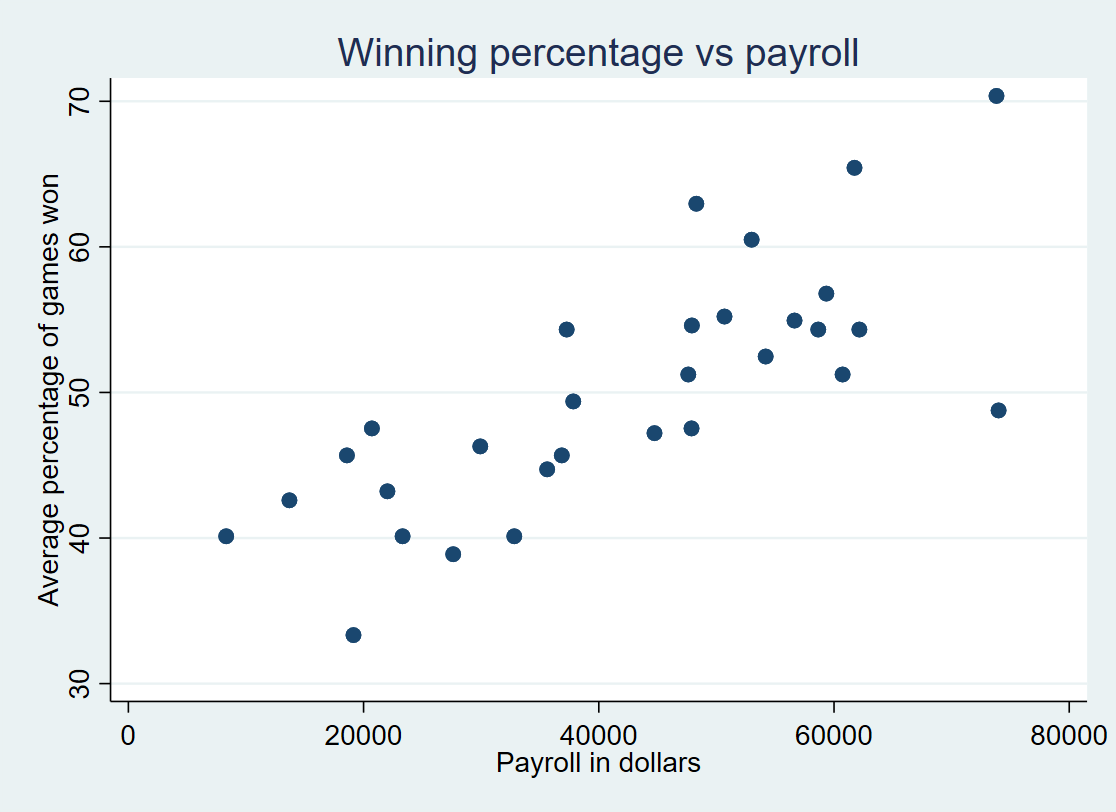
.

. \*\*\*\*Qc) Graph a scatterplot of the team's winning percentage versus payroll.

. twoway scatter win\_percentage payroll, title("Winning percentage vs payroll")

. graph export "output/graph.png",replace

file output/graph.png saved as PNG format



.

. \*\*\*\*Qd) What is the difference in average payroll for teams which make the playoffs compared to teams which do not? Is this difference (i) large in a real world sense, (ii) statistically significant (test this formally), and (iii) likely to be causal?

. ttest payroll, by(playoffs)

Two-sample t test with equal variances

--------------------------------------------------------------------------

Group | Obs Mean Std. err. Std. dev. [95% conf. interval]

---------+----------------------------------------------------------------

0 | 22 36337.84 3590.465 16840.77 28871.06 43804.62

1 | 8 58209.49 2857.861 8083.251 51451.72 64967.26

---------+----------------------------------------------------------------

Combined | 30 42170.28 3255.349 17830.28 35512.35 48828.22

---------+----------------------------------------------------------------

diff | -21871.65 6248.319 -34670.75 -9072.547

--------------------------------------------------------------------------

diff = mean(0) - mean(1) t = -3.5004

H0: diff = 0 Degrees of freedom = 28

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0

Pr(T < t) = 0.0008 Pr(|T| > |t|) = 0.0016 Pr(T > t) = 0.9992

i)The difference in average payroll for teams which make the playoffs compared to teams which do not is 21871.65.

ii)The p-value is really small which is less than 0.05 in two-tails t stat. In other words, we can reject the null at 5% significant level. Therefore, it is statistically significant.

iii)I think it is causal because the null hypothesis which there’s no difference has been rejected by the t test. This means the payroll could be a cause to the playoffs. However, there are many other factors which may affect the playoffs, such as coaches and training environment. Therefore, we cannot conclude that the payrolls are the only cause of the playoffs.

.

end of do-file

. exit, clear