Andrew Abbott

Unit 8 Live Session HW

June 27, 2016

1.



Looking at the scatter plot I expect to see a positive correlation close to zero. The relationship does not appear to be strong and its linearity is unclear. My guess is 0.25

2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Team | Payroll | Wins | Zx | Zy | ZxZy |
| NYY | 206 | 95 | 3.03 | 1.27 | 3.86 |
| BOS | 162 | 89 | 1.88 | 0.73 | 1.37 |
| CHC | 146 | 75 | 1.46 | -0.55 | -0.80 |
| PHI | 142 | 97 | 1.36 | 1.45 | 1.97 |
| NYM | 134 | 79 | 1.15 | -0.18 | -0.21 |
| DET | 123 | 81 | 0.86 | 0.00 | 0.00 |
| CHW | 106 | 88 | 0.41 | 0.64 | 0.26 |
| LAA | 105 | 80 | 0.39 | -0.09 | -0.04 |
| SF | 99 | 92 | 0.23 | 1.00 | 0.23 |
| MIN | 98 | 94 | 0.20 | 1.18 | 0.24 |
| LAD | 95 | 80 | 0.12 | -0.09 | -0.01 |
| HOU | 92 | 76 | 0.05 | -0.45 | -0.02 |
| SEA | 86 | 61 | -0.11 | -1.82 | 0.20 |
| STL | 86 | 86 | -0.11 | 0.45 | -0.05 |
| ATL | 84 | 91 | -0.16 | 0.91 | -0.15 |
| COL | 84 | 83 | -0.16 | 0.18 | -0.03 |
| BAL | 82 | 66 | -0.22 | -1.36 | 0.29 |
| MIL | 81 | 77 | -0.24 | -0.36 | 0.09 |
| TB | 72 | 96 | -0.48 | 1.36 | -0.65 |
| CIN | 71 | 91 | -0.50 | 0.91 | -0.46 |
| KC | 71 | 67 | -0.50 | -1.27 | 0.64 |
| TOR | 62 | 85 | -0.74 | 0.36 | -0.27 |
| ARZ | 61 | 65 | -0.77 | -1.45 | 1.11 |
| CLE | 61 | 69 | -0.77 | -1.09 | 0.84 |
| WAS | 61 | 69 | -0.77 | -1.09 | 0.84 |
| FA | 57 | 80 | -0.87 | -0.09 | 0.08 |
| TEX | 55 | 90 | -0.92 | 0.82 | -0.75 |
| OAK | 52 | 81 | -1.00 | 0.00 | 0.00 |
| SD | 38 | 90 | -1.37 | 0.82 | -1.12 |
| PIT | 35 | 57 | -1.45 | -2.18 | 3.16 |
| mean | 90.23 | 81 |  | r = | 0.36623 |
| S.D. | 38.17 | 11.00 |  |  |  |

Doing the calculation for Pearson’s r in excel using formulas I found a correlation of 0.36623.

Using SAS I also found 0.36623

| **Pearson Correlation Coefficients, N = 30 Prob > |r| under H0: Rho=0** | | |
| --- | --- | --- |
|  | **PAYROLL** | **WINS** |
| **PAYROLL** | 1.00000 | 0.36623 0.0465 |
| **WINS** | 0.36623 0.0465 | 1.00000 |

3. Removing the Padres increases the correlation quite a bit to 0.42555.

4. There is a positive correlation albeit weak. Even removing those 2 outliers as well as S.D. because who likes the padres, the correlation is still only 0.519. many of the highest payroll teams do perform the best and the lowest payroll teams are among the worst though. His examples themselves do not prove his point.

5. The population for this data is all Major league baseball teams for all years. This data is taken from only 2010 payrolls and we do not know why that year was chosen. This is not a random sample.