REPORT

2024-06-01

Summary Statistics

|  | year | sd\_hospital\_beds | sd\_paid\_employees | sd\_non\_paid\_employees | sd\_total\_cost | sd\_total\_revenues | sd\_medicare\_days | sd\_medicaid\_days | sd\_total\_discharge | sd\_medicare\_discharge | sd\_medicaid\_discharge |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sd | 2011 | 560.8998 | 1615.797 | 72.58805 | 304570722 | 323339811 | 19214.29 | 9207.699 | 10898.600 | 10898.600 | 1757.158 |
| Sd1 | 2012 | 579.8366 | 1961.637 | 81.29861 | 294143536 | 321273114 | 19765.74 | 9340.373 | 10994.170 | 10994.170 | 1740.423 |
| n1 | 2011 | 566796.0000 | 1853439.150 | 1199.19000 | 326394348921 | 344203010115 | 25092003.00 | 7866979.000 | 14238489.000 | 14238489.000 | 1674607.000 |
| n2 | 2012 | 574620.0000 | 2259049.058 | 1343.09280 | 327490734909 | 350717045576 | 25938976.38 | 8054865.540 | 14478325.000 | 14478325.000 | 1677082.000 |
| mean1 | 2011 | 376.6086 | 1237.276 | 39.97300 | 216873322 | 228706319 | 16739.16 | 5301.199 | 9492.326 | 9492.326 | 1130.727 |
| mean2 | 2012 | 376.8000 | 1491.121 | 44.76976 | 214748023 | 229978391 | 17110.14 | 5366.333 | 9544.051 | 9544.051 | 1119.547 |

| year | sd\_hospital\_beds | sd\_paid\_employees | sd\_non\_paid\_employees | sd\_total\_cost | sd\_total\_revenues | sd\_medicare\_days | sd\_medicaid\_days | sd\_total\_discharge | sd\_medicare\_discharge | sd\_medicaid\_discharge |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2011 | 560.8998 | 1615.797 | 72.58805 | 304570722 | 323339811 | 19214.29 | 9207.699 | 10898.600 | 10898.600 | 1757.158 |
| 2012 | 579.8366 | 1961.637 | 81.29861 | 294143536 | 321273114 | 19765.74 | 9340.373 | 10994.170 | 10994.170 | 1740.423 |
| 2011 | 566796.0000 | 1853439.150 | 1199.19000 | 326394348921 | 344203010115 | 25092003.00 | 7866979.000 | 14238489.000 | 14238489.000 | 1674607.000 |
| 2012 | 574620.0000 | 2259049.058 | 1343.09280 | 327490734909 | 350717045576 | 25938976.38 | 8054865.540 | 14478325.000 | 14478325.000 | 1677082.000 |
| 2011 | 376.6086 | 1237.276 | 39.97300 | 216873322 | 228706319 | 16739.16 | 5301.199 | 9492.326 | 9492.326 | 1130.727 |
| 2012 | 376.8000 | 1491.121 | 44.76976 | 214748023 | 229978391 | 17110.14 | 5366.333 | 9544.051 | 9544.051 | 1119.547 |

| Characteristics | X2011\_N | X2011\_Mean | X2011\_Std\_Dev | X2012\_N | X2012\_Mean | X2012\_Std\_Dev |
| --- | --- | --- | --- | --- | --- | --- |
| Hospital beds | 5.667960e+05 | 3.766086e+02 | 560.89980 | 5.746200e+05 | 3.768000e+02 | 579.83660 |
| Number of paid Employee | 1.853439e+06 | 1.237276e+03 | 1615.79700 | 2.259049e+06 | 1.491121e+03 | 1961.63700 |
| Number of non-paid Employee | 1.199190e+03 | 3.997300e+01 | 72.58805 | 1.343093e+03 | 4.476976e+01 | 81.29861 |
| Total hospital cost | 3.263943e+11 | 2.168733e+08 | NA | 3.274907e+11 | 2.147480e+08 | NA |
| Total hospital revenues | 3.442030e+11 | 2.287063e+08 | NA | 3.507170e+11 | 2.299784e+08 | NA |
| Available Medicare days | 2.509200e+07 | 1.673916e+04 | 19214.29000 | 2.593898e+07 | 1.711014e+04 | 19765.74000 |
| Available Medicaid days | 7.866979e+06 | 5.301199e+03 | 9207.69900 | 8.054866e+06 | 5.366333e+03 | 9340.37300 |
| Total Hospital Discharge | 1.423849e+07 | 9.492326e+03 | 10898.60000 | 1.447832e+07 | 9.544051e+03 | 10994.17000 |
| Medicare discharge | 1.423849e+07 | 9.492326e+03 | 10898.60000 | 1.447832e+07 | 9.544051e+03 | 10994.17000 |
| Medicaid discharge | 1.674607e+06 | 1.130727e+03 | 1757.15800 | 1.677082e+06 | 1.119547e+03 | 1740.42300 |

# Exercise 2

Descriptive statistics

## # A tibble: 4 × 3  
## # Groups: teaching\_hospital [2]  
## teaching\_hospital system\_member Count  
## <chr> <dbl> <int>  
## 1 Non-Teaching 0 885  
## 2 Non-Teaching 1 1209  
## 3 Teaching 0 309  
## 4 Teaching 1 627

Performing t.test of exercise 1

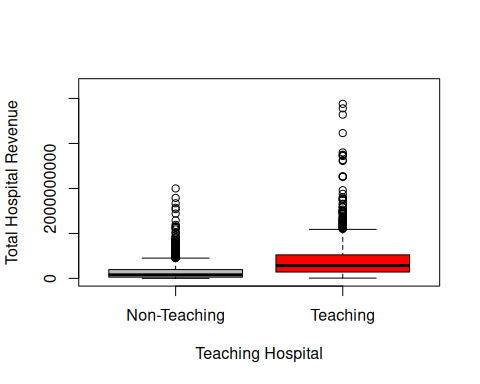
| Statistic | P\_Value |
| --- | --- |
| hospital\_beds | 0.0000 |
| total\_hospital\_employees\_on\_payr | 0.0000 |
| total\_hospital\_non\_paid\_workers | 0.1409 |
| interns\_and\_residents | 0.0000 |
| total\_hosp\_cost | 0.0000 |
| total\_hosp\_revenue | 0.0000 |
| net\_benefits | 0.0000 |
| total\_hospital\_medicare\_days | 0.0000 |
| total\_hospital\_medicaid\_days | 0.0000 |
| total\_hospital\_discharges | 0.0000 |
| total\_hospital\_medicare\_discharg | 0.0000 |
| total\_hospital\_medicaid\_discharg | 0.0000 |
| hospital\_beds | 0.9926 |
| total\_hospital\_employees\_on\_payr | 0.0001 |
| total\_hospital\_non\_paid\_workers | 0.8104 |
| total\_hosp\_cost | 0.8451 |
| total\_hosp\_revenue | 0.9135 |
| total\_hospital\_medicare\_days | 0.6014 |
| total\_hospital\_medicaid\_days | 0.8479 |
| total\_hospital\_discharges | 0.8968 |
| total\_hospital\_medicare\_discharg | 0.0050 |
| total\_hospital\_medicaid\_discharg | 0.8615 |

Comparing hospital net-benefit which hospitals has better performance? To answer this question first compute the hospital net benefits with subtracting hospital costs and revenues and then use ttest to compare the significant differences between teaching and non-teaching hospitals.

var.test to know if the varianccce is equal.

##   
## F test to compare two variances  
##   
## data: net\_benefits by teaching\_hospital  
## F = 0.34347, num df = 2093, denom df = 935, p-value < 2.2e-16  
## alternative hypothesis: true ratio of variances is not equal to 1  
## 95 percent confidence interval:  
## 0.3076147 0.3826102  
## sample estimates:  
## ratio of variances   
## 0.3434651

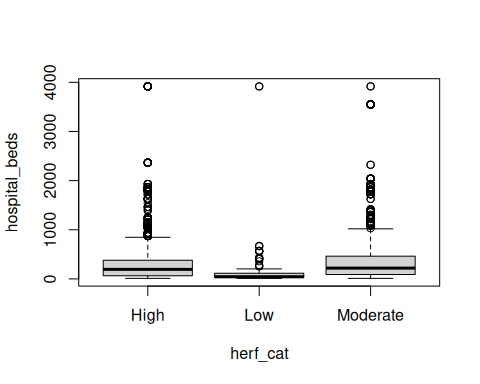
##   
## Welch Two Sample t-test  
##   
## data: net\_benefits by teaching\_hospital  
## t = 8.6723, df = 1231.2, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group Non-Teaching and group Teaching is not equal to 0  
## 95 percent confidence interval:  
## 12292896 19480912  
## sample estimates:  
## mean in group Non-Teaching mean in group Teaching   
## -8635257 -24522161

Use a box-plot and compare hospitals-cost and hospital-revenues between teaching and non-teaching hospitals. 

# Exercise 3

| herf\_cat | Mean\_hospital\_beds | Mean\_paid\_employees | Mean\_non\_paid\_employees | Mean\_internes\_residents | Mean\_system\_membership | Mean\_total\_hosp\_cost | Mean\_total\_revenues | Mean\_net\_benefit | Mean\_medicare\_days | Mean\_medicaid\_days | Mean\_total\_discharge | Mean\_medicare\_discharge | Mean\_medicaid\_discharge | Mean\_hhi |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| High | 373.6403 | 1308.9686 | 47.00928 | 86.55375 | 0.6315078 | 201077823 | 214430928 | -214430928 | 16792.697 | 4812.455 | 9120.806 | 3435.407 | 1066.6464 | 1.697769 |
| Low | 130.9178 | 499.8935 | NaN | 38.32182 | 0.4246575 | 73687086 | 77700145 | -77700145 | 5377.214 | 1416.413 | 2607.836 | 1067.938 | 309.8802 | 1.963470 |
| Moderate | 420.5188 | 1570.1115 | 35.87832 | 112.20558 | 0.6073574 | 255520655 | 270841127 | -270841127 | 18983.776 | 6553.995 | 11100.959 | 3781.610 | 1324.1560 | 1.668919 |

## One way Anova test



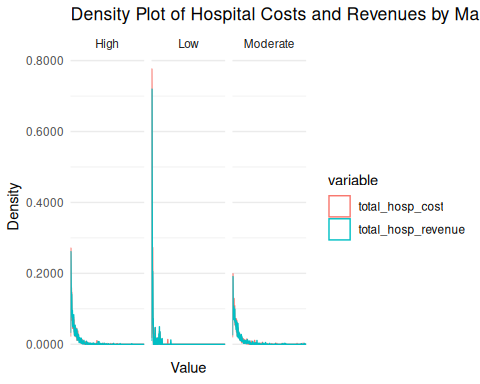
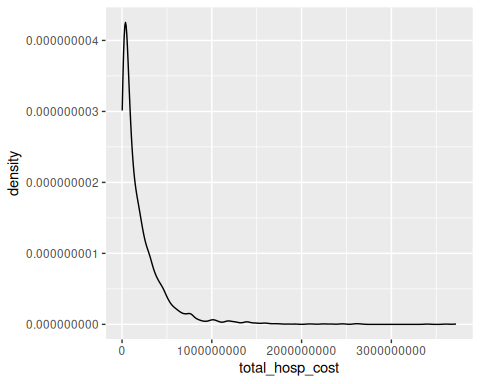
## [1] 0.00000000002378808

##   
## One-way analysis of means (not assuming equal variances)  
##   
## data: hospital\_beds and herf\_cat  
## F = 45.73, num df = 2.00, denom df = 700.48, p-value <  
## 0.00000000000000022

## Df Sum Sq Mean Sq F value Pr(>F)   
## herf\_cat 2 15800945 7900472 24.66 0.0000000000238 \*\*\*  
## Residuals 3027 969755074 320368   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

| p\_value\_hospital\_beds | p\_value\_paid\_employees | p\_value\_non\_paid\_employees | p\_value\_internes\_residents | p\_value\_system\_membership | p\_value\_total\_hosp\_cost | p\_value\_total\_revenues | p\_value\_medicare\_days | p\_value\_medicaid\_days | p\_value\_total\_discharge | p\_value\_medicare\_discharge | p\_value\_medicaid\_discharge | p\_value\_hhi |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0.5825674 | 0.0130616 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## Density Curves



# Linear Regression

##   
## Call:  
## lm(formula = net\_benefits ~ total\_hospital\_beds + own, data = hospital)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -654363322 -3961986 1635010 9567832 572394334   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -906253 1146471 -0.790 0.429   
## total\_hospital\_beds -73150 3220 -22.720 <0.0000000000000002 \*\*\*  
## own1 2409879 1837706 1.311 0.190   
## own2 3420199 2189969 1.562 0.118   
## own3 2224995 1661067 1.339 0.181   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 36300000 on 3010 degrees of freedom  
## (15 observations deleted due to missingness)  
## Multiple R-squared: 0.1474, Adjusted R-squared: 0.1463   
## F-statistic: 130.1 on 4 and 3010 DF, p-value: < 0.00000000000000022

##   
## Call:  
## lm(formula = net\_benefits ~ total\_hospital\_beds + own + system\_member,   
## data = hospital)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -654571181 -4329015 1786854 9697784 573220978   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1939586 1371589 1.414 0.157431   
## total\_hospital\_beds -71276 3251 -21.924 < 0.0000000000000002 \*\*\*  
## own1 2172134 1834796 1.184 0.236564   
## own2 3521355 2185368 1.611 0.107212   
## own3 2000968 1658522 1.206 0.227728   
## system\_member1 -5145510 1368152 -3.761 0.000173 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 36220000 on 3009 degrees of freedom  
## (15 observations deleted due to missingness)  
## Multiple R-squared: 0.1514, Adjusted R-squared: 0.15   
## F-statistic: 107.4 on 5 and 3009 DF, p-value: < 0.00000000000000022

##   
## Call:  
## lm(formula = lm(net\_benefits ~ total\_hospital\_beds + own + system\_member +   
## medicare\_dratio + medicaid\_dratio, data = hospital))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -654404007 -4428627 1789724 9810925 573249954   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1949233 3414651 0.571 0.568149   
## total\_hospital\_beds -71355 3661 -19.491 < 0.0000000000000002 \*\*\*  
## own1 2156294 1859286 1.160 0.246247   
## own2 3545406 2209990 1.604 0.108762   
## own3 1998577 1719083 1.163 0.245092   
## system\_member1 -5247414 1392923 -3.767 0.000168 \*\*\*  
## medicare\_dratio 305422 5298689 0.058 0.954038   
## medicaid\_dratio -300477 7626550 -0.039 0.968575   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 36450000 on 2967 degrees of freedom  
## (55 observations deleted due to missingness)  
## Multiple R-squared: 0.1511, Adjusted R-squared: 0.1491   
## F-statistic: 75.46 on 7 and 2967 DF, p-value: < 0.00000000000000022

## Df Sum Sq Mean Sq F value Pr(>F)  
## herf\_cat 2 0.041 0.020695 2.282 0.102  
## Residuals 2976 26.984 0.009067   
## 51 observations deleted due to missingness

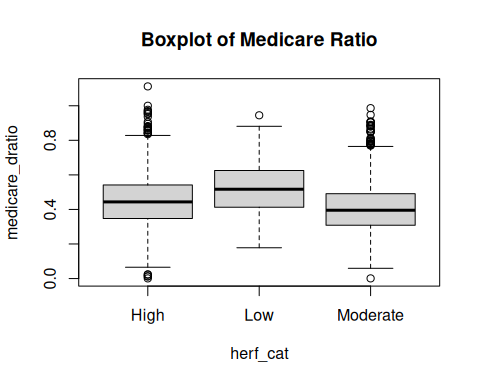
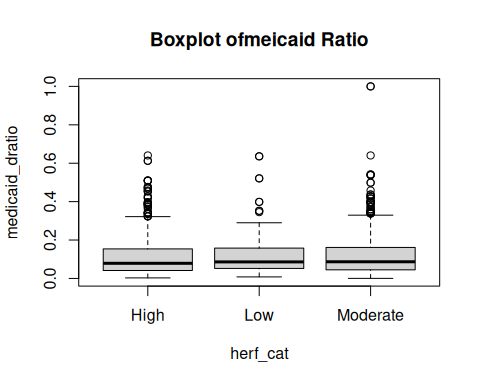
## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = medicaid\_dratio ~ herf\_cat, data = hospital)  
##   
## $herf\_cat  
## diff lwr upr p adj  
## Low-High 0.0069813532 -0.0092722013 0.02323491 0.5723549  
## Moderate-High 0.0075305776 -0.0009765952 0.01603775 0.0950725  
## Moderate-Low 0.0005492244 -0.0158119860 0.01691043 0.9965899

## Df Sum Sq Mean Sq F value Pr(>F)   
## herf\_cat 2 2.94 1.4686 65.66 <0.0000000000000002 \*\*\*  
## Residuals 3012 67.37 0.0224   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## 15 observations deleted due to missingness

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = medicare\_dratio ~ herf\_cat, data = hospital)  
##   
## $herf\_cat  
## diff lwr upr p adj  
## Low-High 0.06643762 0.04103817 0.09183707 0  
## Moderate-High -0.04452082 -0.05780391 -0.03123773 0  
## Moderate-Low -0.11095844 -0.13654108 -0.08537580 0

## Df Sum Sq Mean Sq F value Pr(>F)   
## herf\_cat 2 24151195817825168 12075597908912584 7.873 0.000389 \*\*\*  
## Residuals 3027 4642693347253927936 1533760603651777   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = net\_benefits ~ herf\_cat, data = hospital)  
##   
## $herf\_cat  
## diff lwr upr p adj  
## Low-High 9340047 2691026 15989069 0.0028719  
## Moderate-High -1967366 -5436246 1501514 0.3786309  
## Moderate-Low -11307414 -18003588 -4611239 0.0002266



Tukey Multiple Comparisons of Means

| Category | Difference | Lower | Upper | p.adj |
| --- | --- | --- | --- | --- |
| Low-High | 0.0070 | -0.0093 | 0.0232 | 0.5724 |
| Moderate-High | 0.0075 | -0.0010 | 0.0160 | 0.0951 |
| Moderate-Low | 0.0005 | -0.0158 | 0.0169 | 0.9966 |