Multiple Regression

Kesav Adithya Venkidusamy

Bellevue university - Master of Science in Data Science

Course Name: DSC520-T301 Statistics for Data Science (2221-1)

Assignment: Week 8.2 Assignment

Instructor: Dr Richard Bushart

Due Date: 10/31/2021

Assignment 07

```
# Assignment: ASSIGNMENT 7
# Name: Venkidusamy, Kesav Adithya
# Date: 2021-10-24

> ## Set the working directory to the root of your DSC 520 directory
> setwd("E:/Personal/Bellevue University/Course/github/dsc520")
> ## Load the `data/r4ds/heights.csv` to
> heights_df <- read.csv("data/r4ds/heights.csv")
> summary(heights_df)
      earn            height          sex                ed             age
 Min.   :   200   Min.   :57.50   Length:1192        Min.   : 3.0   Min.   :18.00
 1st Qu.: 10000   1st Qu.:64.01   Class :character   1st Qu.:12.0   1st Qu.:29.00
 Median : 20000   Median :66.45   Mode  :character   Median :13.0   Median :38.00
 Mean   : 23155   Mean   :66.92                      Mean   :13.5   Mean   :41.38
 3rd Qu.: 30000   3rd Qu.:69.85                      3rd Qu.:16.0   3rd Qu.:51.00
 Max.   :200000   Max.   :77.05                      Max.   :18.0   Max.   :91.00
    race
 Length:1192
 Class :character
 Mode  :character

> # Fit a linear model
> earn_lm <-  lm(earn ~ height + sex + ed + age + race, data=heights_df)
> # View the summary of your model
> summary(earn_lm)
```

```
Call:
lm(formula = earn ~ height + sex + ed + age + race, data = heights_df)

Residuals:
   Min     1Q Median     3Q    Max
-39423  -9827  -2208   6157 158723

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -41478.4    12409.4  -3.342 0.000856 ***
height         202.5      185.6   1.091 0.275420
sexmale      10325.6     1424.5   7.249 7.57e-13 ***
ed            2768.4      209.9  13.190  < 2e-16 ***
age            178.3       32.2   5.537 3.78e-08 ***
racehispanic  -1414.3     2685.2  -0.527 0.598507
raceother       371.0     3837.0   0.097 0.922983
racewhite      2432.5     1723.9   1.411 0.158489
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17250 on 1184 degrees of freedom
Multiple R-squared:  0.2199,  Adjusted R-squared:  0.2153
F-statistic: 47.68 on 7 and 1184 DF,  p-value: < 2.2e-16

> predicted_df <- data.frame(
+   earn = predict(earn_lm, heights_df),
```

```
+   ed=heights_df$ed, race=heights_df$race, height=heights_df$height,
+   age=heights_df$age, sex=heights_df$sex
+   )
> ## Compute deviation (i.e. residuals)
> mean_earn <- mean(heights_df$earn)
> mean_earn
[1] 23154.77
> ## Corrected Sum of Squares Total
> sst <- sum((mean_earn - heights_df$earn)^2)
> sst
[1] 451591883937
> ## Corrected Sum of Squares for Model
> ssm <- sum((mean_earn - predicted_df$earn)^2)
> ssm
[1] 99302918657
> ## Residuals
> residuals <- heights_df$earn - predicted_df$earn
> residuals
  [1]  11333.890941  31140.911188   6698.099079  17810.164851  23192.609973
  [6] -11154.599443  13604.930235  -9263.321847 -25288.836877   3238.413948
 [11] -34707.558926   8431.415457 -10992.711766  -4515.869768  -2783.485318
 [16] -11560.980665  21278.601782  14792.164289  -7544.315052  27127.790515
 [21] -13665.656619  -2216.202925 -22071.699892  25242.760836   6882.092950
 [26]   -756.855658  -1969.892050  13145.215569   3993.937070  -2105.566159
 [31]  12023.211638  35208.801814 -10540.644416    151.524576 -18505.781671
 [36]  13631.453485   3437.419145   9780.867692 -24556.200216  12333.550089
 [41]   9396.930903   2491.492079 -17168.532548   4688.275674   2932.577058
```

```
 [46]  8138.885154 -13164.241773  -1608.481840  -5242.623862 -18777.168420
 [51]  3530.028252 -11951.073208  -6869.465800  10384.864604 -21204.898309
 [56] -9314.031509  -7777.883713  -9762.076217  -1336.643042  -7700.985598
 [61] -3881.294636  -6920.147978  -4861.751428  -3920.901310 -26567.929098
 [66]  2615.296998 -17179.884312  -2745.360279  -7127.812156 -10415.049466
 [71] -5079.059595 -17745.449056 -10997.799427   2617.000856   9810.940390
 [76]  5437.653742  -3464.926732  -9324.607282  -9502.945994   6232.131975
 [81]  5561.946112  -8297.251702  -9938.545623 -24114.190946 -14541.345088
 [86]  3368.442767   9980.022048  17315.770961  -8999.217483 -24387.262083
 [91] 24043.827014  -1646.421969   2922.664553  -1155.523997 -26362.143368
 [96] 14810.732583   4680.317692   1403.081485  -1876.620163  -5206.744098
[101] 14804.424337   8468.180326    186.099706  -4472.590442   9396.903318
[106]  1896.157339  -5564.141052  -3811.065444  -7150.198013  45515.404627
[111] -8972.549391 -28654.110131     -3.881121 -27203.209519  35523.420586
[116] -9887.229293   3667.206244  39506.929523  12718.966843  30197.936829
[121]  9652.365699 -14847.034004  -6775.360166   1190.096369 -15027.580600
[126] -22590.183672 -12186.746557  17210.410045  14393.215537  -2204.407779
[131]  2307.689974  16491.322829  14405.740607 -22189.889273  29726.679828
[136]  6292.518431 -19392.328343 -19883.861488   3897.640181  -6358.926987
[141]  1125.332677  66350.743919  -7099.263831   1490.648589   3755.459040
[146] -6800.007713  -2428.836828    151.254975   4343.760774 -17840.445775
[151] 15425.619379 -10202.574919   1957.452621  -5192.159347 -21850.923020
[156]  6067.345323  -5619.371308   3693.697496  -4268.954084   9259.740283
[161] -18810.082613  27411.694962 -10734.910921  -6194.017146 -14046.936173
[166] -14163.282694  -7792.991188   5147.208452   8732.104620  20963.152390
[171]  6707.461748 -12259.747084  -5616.919424   2806.472821  80814.141624
[176]  6965.225880  -6625.987683  12968.719173  39730.969576  11068.086031
```

```
[181]  -7417.904352  13030.230217    603.047386  -2025.530157 -23059.184763
[186]  -8301.762623  -9807.038184  20277.060148  -1669.370514  -5266.874689
[191]  -9401.148295   9832.703834   5416.167417   4466.091263  11774.641361
[196]   5633.487632  43529.298561  18901.180355   5032.119736   9526.455320
[201]   3613.667778  -5261.208831 126488.502286     53.850617   4782.832136
[206]   3653.152007   6671.499215  12886.994823  13792.561570  -5555.474955
[211] -10798.003560   5198.754987  14171.088005 -13592.345671  15644.822158
[216]  -6040.190791   6537.363821   5223.379132   8806.591687  -5174.730159
[221]  15933.657766 -12087.103976  -4671.912521  25083.497154  22249.041553
[226]  29248.418158  20023.996496 -25507.061331  -6126.668811   5497.557098
[231]  -2620.666663 -10993.418736  42440.129722  23695.271892  -3788.674776
[236]  -6530.054264   3816.055352 -14572.885984  17103.878588  -9621.339964
[241]   2025.019128 -13022.288326 -15606.846547  -1639.766246  -2194.087351
[246]  -2555.062694   1235.312859  -4843.068874   7653.393988 -11252.805968
[251] -20649.481287  -5954.503821 -10543.962114   6496.888274    -92.507946
[256]  -4711.146726  11179.144701  -1992.557102  -8762.663651  -5958.159879
[261] -17383.062470  -8471.587669  14994.212961  -9338.465895  -2473.083447
[266] -14808.150356  33108.456012  10421.409940  -9615.078888    743.012989
[271]   -611.347947   3016.116315 -10064.794753  -8566.711681  -1646.581418
[276]   2839.321495  19023.247647  -8432.895485   3667.208536   3707.559217
[281]  -3302.207564 -15398.405033  -2471.365770   8703.696682 -14951.327586
[286] -14908.899720 -16976.358141 -12064.577206    443.367133    164.312850
[291]  -3059.550404  -9670.758759  14162.596931   5338.129297  -6988.260841
[296]   7789.983357  -4470.182412  -2663.626562  -2511.548097  -9493.149022
[301] -16004.636575   2261.431152   1063.282225  11934.482241   8387.103912
[306]   3821.849964  -6244.433513  48194.467565   2046.069389 -27117.745856
[311]   3345.721926  -2738.135109   3333.736591   1197.463545 -15897.738249
```

[316] -10314.748262 -21187.218229 -1434.419379 -12486.262006 -18088.753900

[321] -7664.905252 -12700.375996 12180.583893   503.310543 -3772.437683

[326]   6564.747383 -25196.310916 -13594.036217 12860.605042   4058.252634

[331] -2896.986471 -10454.954118   7970.978764   696.508697   3668.609214

[336] -34210.400221 -12993.945709 -17269.237401 -8588.439567 136575.449371

[341] -39423.460319 -14042.053807   9763.058215 -6949.548796   3759.759459

[346]   -93.721424   4923.633903 14988.554140 21747.419002 54620.816888

[351]   8856.888870   9873.310125 -5957.725937 -5455.106202   2613.315836

[356] -5237.059139 106601.160960   5667.015766 -14415.995073 -21309.261105

[361] -24443.964818   482.706789 -1591.027833   450.963593   849.934309

[366] -4163.111779 -5009.601461   3443.689542 -13632.554068 -9282.917128

[371] -12672.202119 11734.858672   8502.735765 13029.107387 -5685.706765

[376] -16390.590672 68936.414159 17285.961093   -519.269279 -1861.508060

[381] -7963.645544   -235.502113 -13541.490778   568.332569 -7267.573565

[386] -1454.859944 19408.670257   4143.754872   9041.794515 25841.395537

[391] -17929.566866 -12111.731641 -2438.042180 -10760.913235 13509.489644

[396] -8583.633968 -10860.305787 -9971.173541 -4874.725895 -34741.037174

[401] -5399.532798   3273.769933   9937.445751   4536.984867 -18609.364954

[406]   1267.607399 -6910.024497 -9741.969417 -8449.536194 -11687.973666

[411]   -674.324404 -15334.799323   9414.638478 12853.264929   8237.027140

[416] -10076.900032 -15253.382123 15853.308519 -7394.082288 -17649.210850

[421]   3181.690693 50619.201983 10792.689708 -7934.373669 -15271.629244

[426]   4093.623924   7446.634207   -908.368889 26630.187127 14348.031542

[431] 16636.072688 -5051.628098   628.006405 -2120.369073 -10345.569542

[436] -2082.690742 -5497.550541 -11662.783902 -12761.984186 19774.192728

[441]   8734.861368 11008.880887 -8994.924004 -2583.213952 -14094.896455

[446] 13026.144602   2325.951216 -3491.057050   7586.773890 -15485.028710

[451] 18582.326831 2320.704414 12198.374323 -24784.966722 -8572.881158

[456] 11180.777239 12893.572416 -17350.967578 925.539695 5604.411399

[461] 6253.960101 -20849.311359 23310.695214 -4146.893302 24171.784127

[466] -10262.079289 1821.616701 -15927.640642 120.334978 -18389.983584

[471] 1489.913932 -2163.751120 9288.176903 8406.240711 5219.183401

[476] 19473.047363 -4689.215871 13160.820475 -3711.085223 3183.997902

[481] 373.687994 4389.113373 -3216.065580 -12883.606617 -3881.546897

[486] -8677.656511 25658.293013 -915.759846 3586.017549 -6157.705736

[491] -13812.447546 -5948.802135 4897.853649 -1281.305124 -8617.179894

[496] -15599.275922 -12691.848864 10361.196570 -9367.201806 170.726942

[501] 3990.934835 16955.038845 5.903462 -16384.463062 -7514.587734

[506] 8908.174379 16976.440355 -4448.931515 2650.757751 4576.197797

[511] 6369.201774 -1113.307518 8986.249032 355.705815 -4367.056738

[516] -5107.485842 36945.203820 7474.526042 -16621.592718 -9780.217216

[521] -13324.355484 20310.118284 5412.115461 5615.379814 -315.647346

[526] -5838.952401 -8477.252286 -17976.633733 2364.715466 -5028.869351

[531] 1364.529467 -8551.650990 4723.036084 -1144.823488 -9027.816495

[536] -5937.571415 20869.863749 5739.365220 -13695.960053 -3002.637851

[541] -5981.725883 -23196.557538 4741.178003 5696.968136 -1814.970902

[546] 8008.717215 27352.078240 -6549.300974 -2047.325463 11355.190901

[551] -16008.858646 -2211.898810 10757.141175 -3625.570515 10291.250507

[556] -3214.986231 4629.753605 7904.177075 -11030.182610 936.568124

[561] 8875.015538 8321.704123 -5865.534343 14943.000850 -5712.148871

[566] -486.111726 76662.610295 57679.850359 -6631.956204 -12658.017651

[571] 9298.816055 -13632.057010 44682.785022 -19995.897715 9007.368397

[576] -19666.410028 18568.570897 -13867.265555 -11559.978331 -804.991510

[581] -10761.429292 -19300.726649 -3264.701571 -31581.162399 8781.526421

[586] 1924.953815 1270.255333 12448.324879 -9446.160894 -15412.031456

[591] 14388.990580 3063.035646 -12561.280286 3021.944609 18707.281413

[596] -7935.138958 5248.375888 13464.684073 9309.140800 16123.854186

[601] 5909.076608 398.777426 1243.260872 9223.511015 -11105.422084

[606] -8103.012928 1772.519837 36231.714573 -23483.774543 -9918.689175

[611] -1028.353841 -28589.439829 -20436.099033 -4545.129299 4261.385539

[616] 57587.096491 27857.149469 875.776825 -12882.876680 -3909.511698

[621] -448.872535 -10294.287857 -16523.717832 -29283.272524 5235.236464

[626] -979.105366 -4622.007015 953.574914 6970.699612 19691.624891

[631] 10277.144164 -7263.036780 -11931.705356 -3335.945094 -2354.164406

[636] 5154.319662 1885.158868 7966.349873 -76.093656 -10537.639841

[641] -6982.243249 -4805.381433 -19485.564254 1063.585637 1579.304470

[646] -11878.330103 -24979.676190 2522.284624 2636.922462 2473.556960

[651] 13318.393327 -2141.939214 -7114.395885 -5213.653468 6458.226324

[656] 6475.028617 -8674.866226 -8625.202119 400.044275 -1055.649190

[661] 5132.098761 1682.534272 2074.275219 51317.640034 12900.839059

[666] 3162.497415 -2105.319522 -2286.019298 -13662.129240 12987.153116

[671] -615.186365 -2469.158301 -14840.343243 -1378.566252 -9490.708418

[676] -16954.408450 2233.989737 5836.043141 -17837.800409 -11510.787842

[681] 3435.316627 -3132.805245 -4034.979170 -15994.316125 -12982.931949

[686] -8048.351110 -11772.395082 2448.630097 -11012.985548 3811.184939

[691] 19792.290830 -2574.144955 -30051.311255 2318.103413 7997.983676

[696] 6611.794858 548.617457 3042.529945 -6428.676476 -8336.277749

[701] -11574.849995 -7343.182823 -15561.282411 -7024.221298 2850.124892

[706] 4577.862824 -3106.933959 12451.642348 4235.677418 -8522.014565

[711] -2643.046121 -2423.763153 -7659.812251 -14302.875236 2740.979396

[716] -5943.620254 -14646.003184 52366.044716 -11121.098914 16843.993822

[721] 6452.688792 1079.576442 -16355.439089 29958.968107 -780.601759

[726] -17644.009357 9681.700580 -1366.154988 -16217.546426 3480.975521

[731] 6139.794459 -13139.161427 -5868.096744 -15768.816300 -5084.216359

[736] -2334.920670 -5473.449146 7384.053307 4512.301872 -4029.551371

[741] 9056.423841 -482.670884 -709.064246 -22823.203450 -9685.378982

[746] 150.553305 -5284.586315 -11830.043851 -6856.507963 607.154861

[751] 100509.491540 1301.124226 -1359.776403 5406.803084 -1418.730682

[756] 7220.076509 -13789.782723 -8220.765786 4547.988080 -13396.217155

[761] -12035.374227 -6939.876394 -9101.647110 -5510.161632 -6678.615221

[766] -4222.743117 -8733.278328 -8386.713108 12695.185804 2590.275487

[771] 16537.163717 4932.091600 -7787.239779 4001.470796 -2356.581335

[776] 2755.736774 -13988.596481 -4900.181261 -6493.719128 6007.420976

[781] 5143.116701 12105.578646 -5275.696335 -361.564796 13838.366826

[786] -29772.245151 6745.494075 -11679.121642 -1974.323130 6803.773577

[791] -7727.221446 67030.580256 24856.998529 -20188.582625 -12133.518688

[796] -11151.467622 -10164.401378 7199.219675 -7456.456555 -15665.634927

[801] -14327.244375 -18609.690245 -7820.843515 -3658.666664 -12613.380918

[806] 3565.254329 -12628.317891 -15685.103119 738.933629 -10507.238418

[811] -14220.193195 -8832.546610 -19789.513857 -6926.770300 -13819.397669

[816] -4200.410263 8231.580515 -3342.532629 27986.190618 16425.195771

[821] 38741.858188 11208.494004 6281.129461 -5108.889537 13402.458717

[826] -4078.908544 2810.332751 78694.548550 -4250.905839 -24561.807430

[831] -13044.866539 22268.774675 -11928.078904 74328.889755 -1181.129266

[836] 332.422444 -13152.536590 -16486.938021 4047.764100 -3529.541180

[841] -8228.357717 -830.086290 -9959.471780 -4430.064250 15346.307677

[846] -4342.602246 -12038.004036 -7408.162603 -9392.078435 -6028.073252

[851] -14246.802466 3423.987859 -15353.395804 -5072.675137 -4480.573081

[856] 33791.805537 -5618.480247 -19969.118144 -20252.480710 -15424.467137

[861] -483.030017 -10951.584239 2176.264300 -14101.610366 -862.951866

[866] 12770.323232 1265.047347 -13054.899831 -5387.587119 28088.282364

[871] -1189.420315 22857.086364 11327.501656 -14664.757355 6666.208380

[876] -7248.377675 -12278.938308 -2866.938211 -14922.244732 -8106.497771

[881] -22056.264792 3618.855161 -11934.120911 3279.026518 -2037.629776

[886] -13846.924291 -9438.419201 16972.817873 -8406.955175 2432.131251

[891] -13629.433726 1601.371259 -21127.804433 -17051.667427 20322.758562

[896] 34138.174344 -2832.836212 31066.823272 705.141866 8729.096396

[901] -4535.726626 2668.585786 -12375.167703 -7349.403865 57462.397811

[906] -5136.384144 271.569469 -17022.647308 -18866.892444 -11800.639217

[911] 6209.214776 4238.876412 -2613.722772 18363.238196 861.117943

[916] -8559.685360 -7735.020616 15033.343843 -7104.952074 -18546.277292

[921] -63.116791 -11483.593349 -12155.506251 4918.034295 -6159.234883

[926] 5654.222419 -20757.669553 -7003.488486 -4728.407121 14428.202674

[931] -1737.334485 9888.944719 -7035.799541 -9115.005148 986.823891

[936] 7459.220612 -3026.459665 -24560.603693 -5799.517470 -3367.685477

[941] -3794.396376 -2977.958577 -18859.214629 135.672304 241.401965

[946] 4978.262778 -7864.033428 -7275.420498 -9996.597568 7777.219631

[951] -11486.661949 -7347.793938 -9195.115007 -121.923195 -11237.935185

[956] 2392.130877 -3349.787527 4445.090431 -7696.376411 -11130.375743

[961] -19113.594511 -12077.860738 -8454.682569 -10092.946296 -14393.113111

[966] -27844.193463 -6878.039480 -8682.948916 -19261.677398 -10541.480469

[971] 250.025650 -799.798297 -21770.041875 25641.423897 -2525.770185

[976] -13248.429962 -10544.958451 1423.253063 -15124.964139 -8275.057247

[981] -2202.414986 -9111.805064 -3745.760103 -14166.141448 -21858.959274

[986] 4671.197553 -8150.748695 30393.979577 2045.918123 -6354.693501

```
 [991] -30810.329845  15410.782012 -10789.511531  -5554.775891  13816.851878
 [996]  21583.950499   2670.886278 -15855.293195   -920.921637 -13503.184687
 [ reached getOption("max.print") -- omitted 192 entries ]
> ## Sum of Squares for Error
> sse <- sum(residuals^2)
> sse
[1] 3.52289e+11
> ## R Squared
> r_squared <- ssm/sst
> r_squared
[1] 0.2198953
> ## Number of observations
> n <- nrow(heights_df)
> n
[1] 1192
> ## Number of regression paramaters
> p <- 8
> ## Corrected Degrees of Freedom for Model
> dfm <- p-1
> dfm
[1] 7
> ## Degrees of Freedom for Error
> dfe <- n - p
> dfe
[1] 1184
> ## Corrected Degrees of Freedom Total:   DFT = n - 1
> dft <- n -1
```

```
> dft
[1] 1191
> ## Mean of Squares for Model:   MSM = SSM / DFM
> msm <- ssm/dfm
> msm
[1] 14186131237
> ## Mean of Squares for Error:   MSE = SSE / DFE
> mse <- sse/dfe
> mse
[1] 297541356
> ## Mean of Squares Total:   MST = SST / DFT
> mst <- sst/dft
> mst
[1] 379170348
> ## F Statistic
> f_score <- msm/mse
> f_score
[1] 47.67785
> ## Adjusted R Squared R2 = 1 - (1 - R2)(n - 1) / (n - p)
> adjusted_r_squared <- 1 - (1-r_squared)*(n-1)/(n-p)
> adjusted_r_squared
[1] 0.2152832
```