Data Analytics Consulting Assignment 2

Unnikrishnan Sivakumaran Nair

17/02/2022

Analysis of the mercury levels in Maine

```
HgData_maine<-read.csv('Assignment2_2022_Data.txt', sep=" ", header = FALSE)
colnames(HgData_maine)<-c("NAME","HG","N","ELV","SA","Z","LT","ST","DA","RF","FR","DAM","LAT
1","LAT2","LAT3","LONG1","LONG2","LONG3")
HgData_maine
```

```
##                     NAME    HG N  ELV    SA   Z LT ST  DA   RF   FR DAM LAT1
## 1                 ALLEN.P 1.080 3  425    83  27  3  1   2 0.60  2.8   1   44
## 2             ALLIGATOR.P 0.025 2 1494    47  26  2  0   1 0.69  0.8   1   45
## 3       ANASAGUNTICOOK.L 0.570 5  402   568  54  2  1  15 0.56  1.1   0   44
## 4      BALCH&STUMP.PONDS 0.770 5  557   704  44  2  1  14 0.58  2.7   0   43
## 5           BASKAHEGAN.L 0.790 5  417  6944  22  2  0 123 0.57  2.0   1   45
## 6          BAUNEAG.BEG.L 0.750 4  205   200  29  2  1  18 0.51  9.6   0   43
## 7               BEAVER.P 0.270 5  397   128   8  3  0   2 0.61  7.9   1   43
## 8               BELDEN.P 0.660 3  350    24  30  3  1   1   NA   NA   1   44
## 9            BEN.ANNIS.P 0.180 5  122    25   9  2  0  10 0.51 58.8   1   44
## 10              BOTTLE.L 1.050 5  298   281  42  2  1   8 0.48  2.1   1   45
## 11            BRACKETT.L 0.310 5  446   576  25  2  0   7 0.56  1.1   1   45
## 12    BRADBURY(BARKER).L 0.810 2  449    38  45  1  1  17 0.51 24.5   1   46
## 13            BRAINARD.P 0.230 5  270    20  13  2  0   2   NA   NA   1   44
## 14       BRANCH.L(SOUTH) 0.580 5  227  2035  28  2  0  12 0.51  0.6   0   45
## 15        BRANCH.P(EAST) 0.570 5  910    45   9  2  0   2   NA   NA   1   46
## 16   BRANCH.P(UPPER.MID) 0.430 3  341   467  55  1  1   4 0.58  0.5   1   44
## 17              BUBBLE.P 0.100 2  331    32  39  3  0   1 0.64  1.9   0   44
## 18              BURDEN.P 0.490 5  639   197  32  3  1  17 0.58 10.5   1   45
## 19        BURNT.MEADOW.P 0.770 5  374    63  45  2  1   4 0.62  4.5   1   43
## 20               BURNT.P 0.410 5  328   315  27  3  0  NA 0.58  0.7   1   44
## 21        CANADA.FALLS.L 0.790 4 1235  2627  24  2  0 182 0.61 13.1   0   45
## 22     CARLTON.BOG(POND) 0.290 5  203   430   8  2  0  23 0.51 20.1   0   44
## 23               CEDAR.L 0.910 4  500   685  25  2  0   5 0.51  0.6   1   45
## 24        CHAIN.OF.PONDS 0.910 5 1273   700 106  1  1  65 0.62  4.3   0   45
## 25            CHANDLER.L 0.250 5  824   401  19  2  0   5 0.52  1.1   1   46
## 26               CHASE.L 0.430 5  819   403  31  3  1  47 0.58  9.5   1   46
## 27        CHASE.P(FIRST) 0.130 2  995    12  37  1  1   4 0.54 28.6   1   46
## 28                CHUB.P 0.180 5 1095    24  19  2  1   0 0.46  0.8   1   45
## 29           CHURCHILL.L 0.260 5  922  2923  62  1  1 298 0.51  5.8   0   46
## 30     COBBOSSEECONTEE.L 0.290 5  165  5543 100  2  1 131 0.51  1.1   0   44
## 31               CROSS.L 0.390 5  578  2515  46  3  0 164 0.50  3.3   1   47
## 32      CRYSTAL(BEALS).P 0.410 5  328    47  39  3  1   1 0.54  1.1   1   44
## 33        DAMARISCOTTA.L 0.210 4   54  4381 114  3  1  57 0.59  0.5   0   44
## 34     DEBSCONEAG.L(4TH) 0.430 2  634   227 150  1  1   6 0.53  0.5   1   45
## 35     DIMMICK.P(LITTLE) 0.050 4 1390    41  14  2  0   4 0.66 19.4   1   45
## 36                DUCK.L 0.220 5  519  1222  88  1  1   6   NA   NA   1   45
## 37               EAGLE.L 0.440 5  574  5581 136  1  1 762 0.50  3.2   1   47
## 38                EAST.P 0.940 4  263  1823  27  2 NA  NA 0.47  0.2   0   44
## 39              EMBDEN.P 0.570 3  416  1568 158  1  1  22 0.53  0.3   0   44
## 40              FIELDS.P 0.960 5  109   182  31  2  0   3 0.51  8.2   0   44
## 41                FISH.P 0.360 3 1503   211  58  1  1   6 0.56  1.8   1   45
## 42         FISHER.P(BIG) 0.360 5 1150    60  11  2  0   1 0.56  2.9   1   45
## 43              FLYING.P 0.350 4  345   360  80  3  1  15 0.51  1.7   0   44
## 44              FOLSOM.P 0.710 5  221   282  19  2  0  14 0.49  6.2   0   45
## 45              FOREST.L 1.220 5  276   210  38  2  1   3 0.53  1.4   0   43
## 46              GRAHAM.L 0.710 3  102  7865  47  2  0 499 0.58  5.8   0   44
## 47         GRAND.L(WEST) 0.280 2  298 14340 128  1  1 226 0.56  0.5   0   45
## 48             GRANGER.P 0.730 5  524   126  28  3  0   1 0.61  1.0   0   43
## 49   GREENWOOD.P(LITTLE) 0.240 3  683    61  38  1  1   1 0.61  1.1   1   45
## 50                 HAY.L 0.240 2  653   588  34  3  0   6 0.53  0.8   1   46
## 51               HICKS.P 0.900 5  683    93  18  3  0  10 0.59 18.9   0   44
## 52             HODGDON.P 2.500 4   50    35  22  2  1   1 0.63  3.9   1   44
## 53           HORSESHOE.L 0.800 3  454   202  20  2  0   2 0.58  1.1   1   45
## 54              HOSMER.P 0.071 3  212    53  16  3  0   2 0.62  6.8   0   44
```

```
## 55          INDIAN.P(BIG) 0.090 4 1209   280  68  1  1   5 0.76  0.9   1   45
## 56           JACOB.BUCK.P 0.770 4  205   190  52  1  1   3 0.53  0.8   1   44
## 57                JERRY.P 0.620 5  717   272  13  3  0   4 0.56  2.7   1   46
## 58                 JUMP.P 0.430 5  312    29  42  3  1   1 0.56  3.3   1   44
## 59                KEENE.L 0.350 3  195   115  37  3  1   1 0.61  1.3   1   45
## 60            KEEWAYDIN.L 0.890 2  676   307  52  2  1   9 0.06  0.3   0   44
## 61            KINGSBURY.P 0.340 5  929   390  62  3  1  13 0.61  2.2   0   45
## 62               KNIGHT.P 0.280 5  101    49  18  3  0   0 0.51  0.9   1   43
## 63              LAMBERT.L 0.450 5  419   605  60  3  1   6 0.58  0.7   1   45
## 64                 LILY.P 0.370 5  146    44  30  3  1   0 0.46  1.0   1   44
## 65                 LONG.P 0.210 4 1157  3053  44  3  0 558 0.46 20.3   1   45
## 66                 LONG.P 0.400 5  390   271  36  3  1   3 0.58  0.9   1   44
## 67             LOVEWELL.P 0.450 5  357  1120  45  2  1   9 0.06  0.1   1   44
## 68      MACHIAS.L(FOURTH) 1.120 5  311  1539  26  3  0  66 0.56  4.7   1   45
## 69           MEDDYBEMPS.L 0.320 5  170  6765  38  2  0  45 0.62  0.6   0   45
## 70             MOLUNKUS.L 1.120 5  354  1050  38  2  0  35 0.52  2.5   0   45
## 71            MOOSELEUK.L 0.480 5  846   422   6  2  0  92 0.58 45.2   1   46
## 72            NEQUASSET.P 0.370 3   17   392  63  3  1  21 0.58  2.3   0   43
## 73                NORTH.P 0.540 5  487   175  10  3  0   1 0.57  1.5   0   44
## 74                NORTH.P 0.620 4  510   164  50  2  1   2 0.56  0.7   0   44
## 75               ORANGE.L 0.860 5   76   234  24  3  0  19 0.66 12.6   0   44
## 76      OSSIPEE.L(LITTLE) 0.770 3  311   564  74  2  1   6 0.61  0.8   0   43
## 77                OTTER.P 0.160 4 1373    30   8  2  0   0 0.61  2.3   1   45
## 78                OTTER.P 0.130 3 1633    14  18  2  0   0 0.71  1.8   1   45
## 79    PASSAGASSAWAUKEAG.L 0.550 5  304   118  40  3  1   3 0.55  1.9   1   44
## 80               PATTEE.P 0.380 5  141   712  27  2  0  17 0.46  2.3   1   44
## 81                PEASE.P 0.360 5  377   109  19  2  0   2 0.58  2.2   1   44
## 82           PENNINGTON.P 0.080 2  904    45   5  2  0   1 0.51 17.5   1   46
## 83             PINE.P(BIG) 0.670 1 1097   164  33  3  1   5 0.51  3.3   1   45
## 84              PITCHER.P 0.670 5  204   367  38  3  0   9 0.58  2.3   0   44
## 85             PLEASANT.L 0.480 5  232   339  36  3  1   3 0.62  0.7   0   45
## 86             PLEASANT.L 0.410 5  319  1574  92  1  1  21 0.50  0.5  NA   45
## 87             PLEASANT.P 0.600 5  362   239  15  3  0  14 0.06  1.2   1   44
## 88             PORTLAND.L 0.560 5  446    41  53  3  1   1   NA   NA   1   46
## 89    PURGATORY.P(LITTLE) 0.230 5  177    44  20  2  0  NA   NA   NA   0   44
## 90         RANGE.P(LOWER) 1.250 3  306   290  41  2  1  14 0.51  3.7   0   44
## 91       ROACH.P(SECOND) 0.220 5 1271   970  46  3  0  25 0.66  2.1   0   45
## 92     ROBERTS&WADLEY.PDS 0.520 5  271   203  22  3  0   9 0.58 10.1   0   43
## 93                ROCKY.P 0.680 5  312   153  14  2  0   2 0.57  1.8   1   44
## 94          ROUND(GREY).P 0.510 4  269   134  30  2  1   3 0.47  2.1   1   44
## 95                ROUND.P 0.440 5  474   161  32  3  1   2 0.55  0.7   1   44
## 96                ROUND.P 0.570 1   34   250  34  2  1 116 0.61 43.7   1   44
## 97                 ROWE.P 0.220 5 1203   205  43  3  1   2 0.56  0.8   1   45
## 98   SANDY.RIVER.P(LOWER) 0.100 3 1690    17  21  2  1   4 0.56 64.1   1   44
## 99     SANDY.RIVER.P(MID) 0.370 5 1700    70  58  3  1   4 0.56  3.8   1   44
## 100              SECOND.L 0.580 5  247   102  NA NA NA   5   NA   NA  NA   45
## 101            SENNEBEC.P 0.410 3   87   532  57  3  1 106 0.60 14.2   0   44
## 102              SEWALL.P 0.190 5   15    46  11  2  0   0 0.60  1.7   0   43
## 103         SHIN.P(LOWER) 0.470 5  778   638  25  3  0  23 0.56  4.2   1   46
## 104  SLY.BROOK.L(SECOND) 0.370 5  637    13  21  3  1   3 0.50 25.2   1   47
## 105             SPENCER.P 0.140 3 1045   980  16  3  0  21 0.61  5.4   0   45
## 106           SQUAW.P(BIG) 0.260 5 1486    91  96  1  1   1 0.76  0.8   0   45
## 107              SUNDAY.P 0.410 5 1409    30  50  3  1   1 0.61  1.9   1   44
## 108              SYMMES.P 0.180 5  499    36  30  2  1   1 0.61  3.3   0   43
## 109               THIRD.L 0.360 2  751   474  37  1  1  32 0.56  7.4   0   46
## 110               TOGUE.P 0.110 5 1189   388  85  1  1   4 0.56  0.3   1   46
```

```
## 111              TOGUS.P 0.120 5  180   660  49  2  1   5 0.51  0.5   1   44
## 112             TRAVEL.P 0.820 4  204   102   6  2  0  14 0.57 50.9   1   44
## 113            UMBAGOG.L 0.290 5 1245  7850  48  2  0 600 0.56  9.2   0   44
## 114           UMCOLCUS.L 0.430 4  882   630  17  2  0  15 0.61  2.9   1   46
## 115             VARNUM.P 0.160 5  756   331  75  1  1   4 0.61  0.5   0   44
## 116           WADLEIGH.P 0.410 5  913   225  90  1  1  41 0.61  7.3   1   45
## 117             WEBBER.P 0.180 4  118  1201  41  2  1  28 0.51  1.6   0   44
## 118           WEYMOUTH.P 0.190 5  296    87  15  2  0   1 0.47  2.0   1   44
## 119              WIGHT.P 0.490 5   67   135  21  3  1  11 0.58  5.9   0   44
## 120   WOOD.P(LITTLE.BIG) 0.250 5 1244   713  80  1  1  39 0.46  2.0   1   45
##     LAT2 LAT3 LONG1 LONG2 LONG3
## 1     57   44    68     5     7
## 2     37   50    69    12    30
## 3     25   13    70    19    22
## 4     37    0    70    59     4
## 5     30   32    67    50     2
## 6     21   46    70    44    23
## 7     59   47    70    49    26
## 8     24   48    69    23     8
## 9     46   25    68    56    35
## 10    18   36    68     3    17
## 11    44   44    67    51    34
## 12     8   10    68     0    31
## 13    22   15    69    54    59
## 14    23   27    68    40    43
## 15    15   31    69     9    27
## 16    54    0    68    14    45
## 17    20   44    68    14    20
## 18    20   42    69    14    44
## 19    55   28    70    53     9
## 20    44   41    68    31    41
## 21    52   15    70     0     1
## 22    42   14    69    16    21
## 23    31   14    68    48    32
## 24    21    8    70    41    50
## 25    27   13    68    42    12
## 26    24   26    69     2    18
## 27    53   28    68    53    46
## 28    27   30    70    18    20
## 29    26   41    69    18    10
## 30    15   10    69    56    30
## 31     5   16    68    18    32
## 32    16   55    70    16    10
## 33    10   45    69    28    30
## 34    45    4    69     4    43
## 35    13   45    69    52    25
## 36     9    0    68     5    35
## 37     2   24    68    33    10
## 38    36   39    69    46    53
## 39    55   54    69    56    58
## 40    43   47    68    44     6
## 41    44   46    70     7    27
## 42    46   43    69    17     5
## 43    31   16    69    59    36
## 44    20   25    68    26    37
## 45    49   12    70    19    42
```

```
## 46    35   38   68   26   15
## 47    13   56   67   48    6
## 48    57    6   70   46   50
## 49    22    7   69   24   50
## 50     9   10   68   43   18
## 51    18   24   70   39   16
## 52    19   32   68   23   51
## 53     1    5   68    3   52
## 54    12   53   69    7   44
## 55    26   20   69   44   12
## 56    38   45   68   44   40
## 57     5   50   68   40   33
## 58    24    9   69   23   55
## 59     6   38   67   10   30
## 60    15   54   70   50   13
## 61     6   38   69   39   19
## 62    15   21   70   45   49
## 63    32   56   67   33   15
## 64    27   54   69   42   20
## 65    37   20   70    2    8
## 66    55   25   68   15   59
## 67     0    7   70   55   36
## 68     7   39   68    0   26
## 69     4   27   67   21   43
## 70    39   40   68   18   18
## 71    30   33   68   54   18
## 72    57    4   69   46   13
## 73    15   38   70   35   11
## 74    19   43   70   24    1
## 75    46    8   67   14   56
## 76    35   48   70   42   26
## 77    21   57   70   44   53
## 78    10   51   70   58   53
## 79    30   48   69    7   51
## 80    32    1   69   33   49
## 81    35   55   70   10   34
## 82    56   10   68   31   11
## 83    52    1   69   25   37
## 84    20   14   69    2   24
## 85     3   59   67   29   10
## 86    21   33   67   55   10
## 87     0   24   70   53   25
## 88    24    4   67   49   28
## 89    12   56   69   56   47
## 90     2   25   70   21   31
## 91    40   34   69   16   36
## 92    32    6   70   38   34
## 93    35   17   68   35   52
## 94    44   26   69   13   30
## 95    25   57   70   13   15
## 96    12    3   69   17   36
## 97     7   32   69   59   29
## 98    53   52   70   32   34
## 99    53   52   70   33   16
## 100    0   52   67   47   34
## 101   15   26   69   15   59
```

```
## 102   52    7    69    46    48
## 103    5    9    68    33    50
## 104    7   11    68    31    19
## 105   44   34    69    33    31
## 106   27   22    69    40    44
## 107   47   56    70    57    12
## 108   38   56    70    52    43
## 109   14   42    69     1    54
## 110   56    1    68    53    30
## 111   19   28    69    39    31
## 112   15   14    69    31    49
## 113   47   25    71     0    47
## 114   17   16    68    25    49
## 115   39   27    70    14    23
## 116   44   43    69    11    24
## 117   24   13    69    39    53
## 118   58    9    69    19    33
## 119   27   48    68    40    33
## 120   38   12    70    20    40
```

Test of hypothesis to test statistical significance of Hg levels in Maine

H0: The Hg levels in Maine <= 0.43 H1: The Hg levels in Maine > 0.43

```
t.test(HgData_maine$HG,mu=0.43,alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  HgData_maine$HG
## t = 1.8095, df = 119, p-value = 0.03645
## alternative hypothesis: true mean is greater than 0.43
## 95 percent confidence interval:
##  0.4346161       Inf
## sample estimates:
## mean of x
##   0.48505
```

From the analysis there is not enough statistical evidence to claim that the mean Hg levels in Maine is greater than 0.43

Test of hypothesis to test statistical significance of Hg levels in Maine

H0: The Hg levels in Maine <= 0.5 H1: The Hg levels in Maine > 0.5

```
t.test(HgData_maine$HG,mu=0.5,alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  HgData_maine$HG
## t = -0.49141, df = 119, p-value = 0.688
## alternative hypothesis: true mean is greater than 0.5
## 95 percent confidence interval:
##  0.4346161       Inf
## sample estimates:
## mean of x
##   0.48505
```

H0: The Hg levels in Maine <= 1 H1: The Hg levels in Maine > 1

```
t.test(HgData_maine$HG,mu=1,alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  HgData_maine$HG
## t = -16.926, df = 119, p-value = 1
## alternative hypothesis: true mean is greater than 1
## 95 percent confidence interval:
##  0.4346161       Inf
## sample estimates:
## mean of x
##   0.48505
```

From the p-value it can be inferred that the mercury levels in Maine is not as high as .5 which is the normal Hg range accepted in most of the other states in the US.

Checking if the presence of Dams have an influence on the Hg levels.

```
model<-lm(HG~DAM+ELV, data=HgData_maine)
summary(model)
```

```
##
## Call:
## lm(formula = HG ~ DAM + ELV, data = HgData_maine)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51651 -0.21115 -0.05976  0.14199  1.89209
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.408e-01  5.627e-02  11.387  < 2e-16 ***
## DAM         -2.028e-02  6.181e-02  -0.328 0.743386
## ELV         -2.511e-04  7.023e-05  -3.576 0.000512 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3196 on 115 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:   0.11,  Adjusted R-squared:  0.09452
## F-statistic: 7.106 on 2 and 115 DF,  p-value: 0.001231
```

```
HG_dam<-HgData_maine[HgData_maine$DAM==1,]
HG_nodam<-HgData_maine[HgData_maine$DAM==0,]
```

```
var.test(HG_dam$HG, HG_nodam$HG)
```

```
##
##  F test to compare two variances
##
## data:  HG_dam$HG and HG_nodam$HG
## F = 1.2558, num df = 70, denom df = 46, p-value = 0.4129
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.7270827 2.1058031
## sample estimates:
## ratio of variances
##           1.255785
```

```
t.test(HG_dam$HG,HG_nodam$HG,var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  HG_dam$HG and HG_nodam$HG
## t = -1.1387, df = 116, p-value = 0.2572
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.19679194  0.05311438
## sample estimates:
## mean of x mean of y
## 0.4562676 0.5281064
```

Checking if the lake type have an influence on the Hg levels.

```
HG_Eutrophic<-HgData_maine[HgData_maine$LT==2,]
HG_OtherLT<-HgData_maine[HgData_maine$LT!=0,]
```

```
var.test(HG_Eutrophic$HG, HG_OtherLT$HG)
```

```
##
##  F test to compare two variances
##
## data:  HG_Eutrophic$HG and HG_OtherLT$HG
## F = 1.6153, num df = 52, denom df = 118, p-value = 0.03426
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  1.035952 2.627901
## sample estimates:
## ratio of variances
##           1.615308
```

```
t.test(HG_Eutrophic$HG,HG_OtherLT$HG,alternative = "greater",var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  HG_Eutrophic$HG and HG_OtherLT$HG
## t = 0.96947, df = 81.885, p-value = 0.1676
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.0457964        Inf
## sample estimates:
## mean of x mean of y
## 0.5482075 0.4842521
```