Submission 1

Ruining Zhou

2025-07-05

```r
# read the datasets
genes <- read.csv("~/Desktop/QBS 103/Homework/Final Project/QBS103_GSE157103_genes.csv", row.names = 1)
series.matrix <- read.csv("~/Desktop/QBS 103/Homework/Final Project/QBS103_GSE157103_series_matrix-1.csv

# identify one gene, one continuous covariate, and two categorical covariates in the provided dataset.

# choose gene "AATF"
genes["AATF",]
```

```
##      COVID_01_39y_male_NonICU COVID_02_63y_male_NonICU COVID_03_33y_male_NonICU
## AATF                    45.83                    39.37                    42.35
##      COVID_04_49y_male_NonICU COVID_05_49y_male_NonICU COVID_06_.y_male_NonICU
## AATF                    41.92                    30.56                    36.3
##      COVID_07_38y_female_NonICU COVID_08_78y_male_ICU COVID_09_64y_female_ICU
## AATF                      44.16                 31.81                   30.44
##      COVID_10_62y_male_ICU COVID_11_52y_female_NonICU COVID_12_50y_male_ICU
## AATF                 30.99                      45.66                 35.24
##      COVID_13_37y_male_NonICU COVID_14_55y_male_ICU COVID_15_68y_male_ICU
## AATF                    41.33                 43.86                 31.54
##      COVID_16_48y_male_NonICU COVID_17_54y_male_NonICU
## AATF                    45.35                    44.33
##      COVID_18_70y_female_NonICU COVID_19_51y_male_NonICU COVID_20_62y_male_ICU
## AATF                       36.5                    54.87                 46.16
##      COVID_21_66y_male_ICU COVID_22_43y_male_ICU COVID_23_76y_male_ICU
## AATF                 33.95                 26.37                 34.65
##      COVID_24_55y_male_ICU COVID_25_55y_male_ICU COVID_26_41y_female_ICU
## AATF                 35.23                 39.65                   29.52
##      COVID_27_71y_female_ICU COVID_28_63y_male_ICU COVID_29_63y_female_ICU
## AATF                   33.25                 38.01                   31.84
##      COVID_30_54y_male_ICU COVID_31_50y_male_ICU COVID_32_72y_male_ICU
## AATF                 34.05                  28.4                 47.32
##      COVID_33_81y_male_NonICU COVID_34_64y_female_NonICU
## AATF                    40.75                      28.4
##      COVID_35_58y_female_NonICU COVID_36_68y_male_NonICU
## AATF                      35.31                     34.6
##      COVID_37_87y_male_NonICU COVID_38_68y_male_ICU COVID_39_80y_female_ICU
## AATF                    46.15                 28.34                   24.16
##      COVID_40_66y_male_ICU COVID_41_74y_male_ICU COVID_42_21y_female_ICU
## AATF                 37.86                 29.56                   39.06
##      COVID_43_83y_female_ICU COVID_44_46y_male_ICU COVID_45_62y_female_ICU
## AATF                   23.57                 37.64                   32.54
```

```
##      COVID_46_62y_male_ICU COVID_47_78y_male_ICU COVID_48_72y_female_ICU
## AATF                 35.15                 52.61                   25.55
##      COVID_49_73y_male_ICU COVID_50_37y_male_ICU COVID_51_58y_female_NonICU
## AATF                 40.75                   7.2                     49.57
##      COVID_52_71y_male_NonICU COVID_53_35y_female_NonICU
## AATF                    43.6                     56.99
##      COVID_55_62y_female_ICU COVID_56_33y_female_NonICU
## AATF                   42.51                     40.38
##      COVID_57_30y_female_NonICU COVID_58_62y_male_NonICU
## AATF                     49.31                     33.96
##      COVID_59_55y_male_NonICU COVID_60_49y_male_NonICU
## AATF                    45.74                     42.22
##      COVID_61_54y_female_NonICU COVID_62_78y_female_ICU COVID_63_39y_female_ICU
## AATF                     45.81                   32.03                   47.55
##      COVID_64_65y_male_ICU COVID_65_84y_male_NonICU COVID_66_66y_female_NonICU
## AATF                 36.76                    45.17                      51.42
##      COVID_67_57y_male_ICU COVID_68_79y_male_ICU COVID_69_77y_female_NonICU
## AATF                 29.55                 26.27                      49.5
##      COVID_70_81y_male_NonICU COVID_71_37y_male_ICU COVID_72_50y_female_NonICU
## AATF                    50.75                 30.91                      47.65
##      COVID_73_82y_male_NonICU COVID_74_55y_female_ICU COVID_75_55y_male_NonICU
## AATF                    55.87                   24.46                       44
##      COVID_76_73y_female_ICU COVID_77_55y_female_ICU COVID_78_80y_male_NonICU
## AATF                   24.91                   32.72                      49.67
##      COVID_79_27y_male_NonICU COVID_80_71y_male_ICU COVID_82_67y_male_NonICU
## AATF                    45.85                 27.99                    32.34
##      COVID_83_85y_female_NonICU COVID_84_75y_female_NonICU
## AATF                     35.35                     33.65
##      COVID_85_62y_male_ICU COVID_86_52y_female_NonICU COVID_87_61y_male_ICU
## AATF                 23.24                     38.6                 33.69
##      COVID_89_90y_female_NonICU COVID_90_86y_female_NonICU
## AATF                     29.41                     28.65
##      COVID_91_29y_female_NonICU COVID_92_82y_female_ICU COVID_93_81y_female_ICU
## AATF                     31.21                   22.38                   37.37
##      COVID_94_24y_female_NonICU COVID_95_49y_male_NonICU
## AATF                     43.17                     40.37
##      COVID_96_51y_male_NonICU COVID_97_76y_male_ICU COVID_98_81y_male_NonICU
## AATF                    27.71                 31.65                    32.06
##      COVID_99_71y_male_ICU COVID_100_74y_female_NonICU COVID_101_58y_male_ICU
## AATF                 36.85                      46.54                  47.01
##      COVID_102_84y_male_NonICU COVID_103_83y_male_NonICU
## AATF                     33.19                     27.96
##      NONCOVID_01_54y_female_NonICU NONCOVID_02_65y_male_ICU
## AATF                        37.33                    34.32
##      NONCOVID_03_65y_male_ICU NONCOVID_04_90y_male_NonICU
## AATF                    28.33                      33.81
##      NONCOVID_05_83y_female_NonICU NONCOVID_06_75y_female_ICU
## AATF                         32.3                      38.95
##      NONCOVID_07_50y_male_ICU NONCOVID_08_53y_female_ICU
## AATF                    40.42                      33.98
##      NONCOVID_09_49y_female_NonICU NONCOVID_10_67y_male_ICU
## AATF                        42.64                    27.9
##      NONCOVID_11_58y_female_NonICU NONCOVID_12_82y_male_ICU
## AATF                        43.38                    36.86
```

```
##      NONCOVID_13_65y_male_ICU NONCOVID_14_75y_female_ICU
## AATF                   34.24                      30.16
##      NONCOVID_15_83y_unknown_ICU NONCOVID_16_40y_female_ICU
## AATF                       49.21                      35.08
##      NONCOVID_17_84y_female_ICU NONCOVID_18_88y_male_ICU
## AATF                     42.71                    33.86
##      NONCOVID_19_66y_female_ICU NONCOVID_20_62y_female_ICU
## AATF                     27.21                      39.19
##      NONCOVID_21_71y_male_NonICU NONCOVID_22_63y_male_NonICU
## AATF                       44.55                       29.72
##      NONCOVID_23_42y_female_NonICU NONCOVID_24_32y_female_NonICU
## AATF                         37.58                         50.13
##      NONCOVID_25_62y_male_NonICU NONCOVID_26_36y_male_ICU
## AATF                       27.51                    18.31
```

```r
# check data set about to choose continuous covariate and categorical covariates
head(genes)
```

```
##         COVID_01_39y_male_NonICU COVID_02_63y_male_NonICU
## A1BG                        0.49                     0.29
## A1CF                        0.00                     0.00
## A2M                         0.21                     0.14
## A2ML1                       0.04                     0.00
## A3GALT2                     0.07                     0.00
## A4GALT                      0.00                     0.00
##         COVID_03_33y_male_NonICU COVID_04_49y_male_NonICU
## A1BG                        0.26                     0.45
## A1CF                        0.00                     0.01
## A2M                         0.03                     0.09
## A2ML1                       0.02                     0.07
## A3GALT2                     0.00                     0.00
## A4GALT                      0.00                     0.00
##         COVID_05_49y_male_NonICU COVID_06_.y_male_NonICU
## A1BG                        0.17                    0.21
## A1CF                        0.00                    0.00
## A2M                         0.00                    0.08
## A2ML1                       0.05                    0.04
## A3GALT2                     0.07                    0.00
## A4GALT                      0.00                    0.00
##         COVID_07_38y_female_NonICU COVID_08_78y_male_ICU
## A1BG                          0.49                  0.12
## A1CF                          0.01                  0.00
## A2M                           0.23                  0.08
## A2ML1                         0.03                  0.01
## A3GALT2                       0.07                  0.00
## A4GALT                        0.00                  0.00
##         COVID_09_64y_female_ICU COVID_10_62y_male_ICU
## A1BG                       0.51                  0.10
## A1CF                       0.01                  0.00
## A2M                        0.88                  0.13
## A2ML1                      0.02                  0.01
## A3GALT2                    0.79                  0.15
## A4GALT                     0.00                  0.00
##         COVID_11_52y_female_NonICU COVID_12_50y_male_ICU
```

```
## A1BG                                   0.38                   0.45
## A1CF                                   0.02                   0.00
## A2M                                    0.47                   0.16
## A2ML1                                  0.03                   0.00
## A3GALT2                                0.08                   1.75
## A4GALT                                 0.00                   0.00
##         COVID_13_37y_male_NonICU COVID_14_55y_male_ICU COVID_15_68y_male_ICU
## A1BG                        0.18                  0.23                  0.42
## A1CF                        0.00                  0.00                  0.00
## A2M                         0.07                  0.22                  0.07
## A2ML1                       0.01                  0.04                  0.00
## A3GALT2                     0.00                  0.93                  0.15
## A4GALT                      0.00                  0.00                  0.03
##         COVID_16_48y_male_NonICU COVID_17_54y_male_NonICU
## A1BG                        0.41                     0.63
## A1CF                        0.01                     0.02
## A2M                         0.58                     0.15
## A2ML1                       0.00                     0.02
## A3GALT2                     0.19                     0.00
## A4GALT                      0.00                     0.00
##         COVID_18_70y_female_NonICU COVID_19_51y_male_NonICU
## A1BG                          0.47                     0.33
## A1CF                          0.00                     0.02
## A2M                           0.30                     0.11
## A2ML1                         0.02                     0.02
## A3GALT2                       0.06                     0.00
## A4GALT                        0.03                     0.00
##         COVID_20_62y_male_ICU COVID_21_66y_male_ICU COVID_22_43y_male_ICU
## A1BG                     0.32                  0.18                  0.09
## A1CF                     0.00                  0.00                  0.00
## A2M                      0.07                  0.00                  0.06
## A2ML1                    0.00                  0.00                  0.00
## A3GALT2                  0.22                  0.37                  0.06
## A4GALT                   0.00                  0.03                  0.00
##         COVID_23_76y_male_ICU COVID_24_55y_male_ICU COVID_25_55y_male_ICU
## A1BG                     0.18                  0.22                  0.29
## A1CF                     0.01                  0.01                  0.00
## A2M                      0.03                  0.11                  0.09
## A2ML1                    0.00                  0.02                  0.03
## A3GALT2                  0.07                  0.15                  0.00
## A4GALT                   0.03                  0.00                  0.00
##         COVID_26_41y_female_ICU COVID_27_71y_female_ICU COVID_28_63y_male_ICU
## A1BG                       0.42                    0.16                  0.18
## A1CF                       0.00                    0.01                  0.00
## A2M                        0.18                    0.23                  0.18
## A2ML1                      0.00                    0.01                  0.05
## A3GALT2                    0.87                    0.18                  0.45
## A4GALT                     0.00                    0.00                  0.00
##         COVID_29_63y_female_ICU COVID_30_54y_male_ICU COVID_31_50y_male_ICU
## A1BG                       0.35                  0.23                  0.15
## A1CF                       0.00                  0.00                  0.00
## A2M                        0.03                  0.11                  0.47
## A2ML1                      0.03                  0.01                  0.00
## A3GALT2                    0.15                  0.00                  0.00
```

```
## A4GALT                              0.03                0.00                           0.03
##           COVID_32_72y_male_ICU COVID_33_81y_male_NonICU
## A1BG                        0.34                     0.35
## A1CF                        0.01                     0.00
## A2M                         0.04                     0.30
## A2ML1                       0.00                     0.06
## A3GALT2                     0.29                     0.26
## A4GALT                      0.00                     0.00
##           COVID_34_64y_female_NonICU COVID_35_58y_female_NonICU
## A1BG                         0.36                       0.26
## A1CF                         0.00                       0.00
## A2M                          0.11                       0.51
## A2ML1                        0.00                       0.02
## A3GALT2                      0.12                       0.16
## A4GALT                       0.00                       0.00
##           COVID_36_68y_male_NonICU COVID_37_87y_male_NonICU COVID_38_68y_male_ICU
## A1BG                       0.18                     0.20                  0.29
## A1CF                       0.01                     0.00                  0.00
## A2M                        0.09                     0.09                  0.10
## A2ML1                      0.00                     0.07                  0.02
## A3GALT2                    0.08                     0.31                  0.35
## A4GALT                     0.00                     0.00                  0.00
##           COVID_39_80y_female_ICU COVID_40_66y_male_ICU COVID_41_74y_male_ICU
## A1BG                      0.19                    0.22                  0.19
## A1CF                      0.00                    0.00                  0.00
## A2M                       0.27                    0.17                  0.14
## A2ML1                     0.00                    0.00                  0.00
## A3GALT2                   0.00                    0.08                  0.19
## A4GALT                    0.07                    0.00                  0.00
##           COVID_42_21y_female_ICU COVID_43_83y_female_ICU COVID_44_46y_male_ICU
## A1BG                      0.24                    0.29                  0.22
## A1CF                      0.01                    0.00                  0.00
## A2M                       0.33                    0.00                  0.14
## A2ML1                     0.01                    0.00                  0.00
## A3GALT2                   0.39                    0.11                  0.00
## A4GALT                    0.00                    0.00                  0.04
##           COVID_45_62y_female_ICU COVID_46_62y_male_ICU COVID_47_78y_male_ICU
## A1BG                      0.14                    0.53                  0.08
## A1CF                      0.00                    0.01                  0.01
## A2M                       0.15                    0.10                  0.04
## A2ML1                     0.03                    0.00                  0.03
## A3GALT2                   0.19                    0.06                  0.60
## A4GALT                    0.00                    0.00                  0.00
##           COVID_48_72y_female_ICU COVID_49_73y_male_ICU COVID_50_37y_male_ICU
## A1BG                      0.19                    0.48                  0.08
## A1CF                      0.00                    0.00                  0.00
## A2M                       0.06                    0.09                  0.01
## A2ML1                     0.01                    0.03                  0.00
## A3GALT2                   0.23                    0.00                  0.00
## A4GALT                    0.06                    0.00                  0.72
##           COVID_51_58y_female_NonICU COVID_52_71y_male_NonICU
## A1BG                         0.21                       0.25
## A1CF                         0.00                       0.01
## A2M                          0.13                       0.00
```

```
## A2ML1                                   0.00                        0.03
## A3GALT2                                 0.00                        0.00
## A4GALT                                  0.00                        0.00
##         COVID_53_35y_female_NonICU COVID_55_62y_female_ICU
## A1BG                                   0.25                        0.09
## A1CF                                   0.00                        0.00
## A2M                                    0.64                        0.09
## A2ML1                                  0.10                        0.01
## A3GALT2                                0.00                        0.00
## A4GALT                                 0.00                        0.00
##         COVID_56_33y_female_NonICU COVID_57_30y_female_NonICU
## A1BG                                   0.28                           0.42
## A1CF                                   0.00                           0.00
## A2M                                    0.16                           0.27
## A2ML1                                  0.09                           0.01
## A3GALT2                                0.23                           0.19
## A4GALT                                 0.00                           0.05
##         COVID_58_62y_male_NonICU COVID_59_55y_male_NonICU
## A1BG                                   0.39                         0.33
## A1CF                                   0.00                         0.00
## A2M                                    0.08                         0.10
## A2ML1                                  0.00                         0.00
## A3GALT2                                0.00                         0.07
## A4GALT                                 0.00                         0.00
##         COVID_60_49y_male_NonICU COVID_61_54y_female_NonICU
## A1BG                                   0.22                           0.25
## A1CF                                   0.00                           0.00
## A2M                                    0.14                           0.10
## A2ML1                                  0.00                           0.03
## A3GALT2                                0.00                           0.13
## A4GALT                                 0.02                           0.00
##         COVID_62_78y_female_ICU COVID_63_39y_female_ICU COVID_64_65y_male_ICU
## A1BG                                   0.21                        0.29                       0.38
## A1CF                                   0.00                        0.00                       0.01
## A2M                                    0.04                        0.01                       0.04
## A2ML1                                  0.00                        0.00                       0.02
## A3GALT2                                0.05                        0.14                       0.56
## A4GALT                                 0.00                        0.00                       0.00
##         COVID_65_84y_male_NonICU COVID_66_66y_female_NonICU
## A1BG                                   0.40                           0.64
## A1CF                                   0.01                           0.00
## A2M                                    0.07                           0.00
## A2ML1                                  0.00                           0.00
## A3GALT2                                0.58                           0.00
## A4GALT                                 0.00                           0.00
##         COVID_67_57y_male_ICU COVID_68_79y_male_ICU COVID_69_77y_female_NonICU
## A1BG                                   0.37                        0.58                           0.52
## A1CF                                   0.00                        0.00                           0.00
## A2M                                    0.35                        0.15                           0.29
## A2ML1                                  0.00                        0.01                           0.02
## A3GALT2                                0.00                        0.00                           0.00
## A4GALT                                 0.00                        0.05                           0.00
##         COVID_70_81y_male_NonICU COVID_71_37y_male_ICU
## A1BG                                   0.27                        0.07
```

```
## A1CF                                     0.00                      0.01
## A2M                                      0.07                      0.12
## A2ML1                                    0.00                      0.01
## A3GALT2                                  0.00                      0.00
## A4GALT                                   0.06                      0.00
##          COVID_72_50y_female_NonICU COVID_73_82y_male_NonICU
## A1BG                                     0.52                      0.46
## A1CF                                     0.00                      0.01
## A2M                                      0.10                      0.02
## A2ML1                                    0.01                      0.02
## A3GALT2                                  0.00                      0.17
## A4GALT                                   0.00                      0.04
##          COVID_74_55y_female_ICU COVID_75_55y_male_NonICU
## A1BG                                  0.24                    0.23
## A1CF                                  0.00                    0.01
## A2M                                   0.12                    0.14
## A2ML1                                 0.02                    0.00
## A3GALT2                               0.26                    0.00
## A4GALT                                0.00                    0.00
##          COVID_76_73y_female_ICU COVID_77_55y_female_ICU
## A1BG                                  0.17                   0.05
## A1CF                                  0.00                   0.00
## A2M                                   0.09                   0.01
## A2ML1                                 0.01                   0.00
## A3GALT2                               0.04                   0.00
## A4GALT                                0.00                   0.00
##          COVID_78_80y_male_NonICU COVID_79_27y_male_NonICU COVID_80_71y_male_ICU
## A1BG                                    0.19                      0.08                  0.28
## A1CF                                    0.00                      0.01                  0.00
## A2M                                     0.20                      0.03                  0.05
## A2ML1                                   0.00                      0.00                  0.00
## A3GALT2                                 0.00                      0.00                  0.05
## A4GALT                                  0.00                      0.00                  0.00
##          COVID_82_67y_male_NonICU COVID_83_85y_female_NonICU
## A1BG                                    0.39                        0.47
## A1CF                                    0.01                        0.00
## A2M                                     0.10                        0.18
## A2ML1                                   0.00                        0.05
## A3GALT2                                 0.00                        0.00
## A4GALT                                  0.00                        0.00
##          COVID_84_75y_female_NonICU COVID_85_62y_male_ICU
## A1BG                                      0.35                  0.29
## A1CF                                      0.00                  0.00
## A2M                                       0.03                  0.04
## A2ML1                                     0.00                  0.00
## A3GALT2                                   0.17                  0.00
## A4GALT                                    0.00                  0.00
##          COVID_86_52y_female_NonICU COVID_87_61y_male_ICU
## A1BG                                      0.60                  0.65
## A1CF                                      0.00                  0.00
## A2M                                       0.27                  0.15
## A2ML1                                     0.02                  0.00
## A3GALT2                                   0.00                  0.00
## A4GALT                                    0.00                  0.00
```

```
##         COVID_89_90y_female_NonICU COVID_90_86y_female_NonICU
## A1BG                          0.20                       0.40
## A1CF                          0.00                       0.00
## A2M                           0.07                       0.05
## A2ML1                         0.03                       0.01
## A3GALT2                       0.14                       0.31
## A4GALT                        0.00                       0.02
##         COVID_91_29y_female_NonICU COVID_92_82y_female_ICU
## A1BG                          0.60                    0.34
## A1CF                          0.00                    0.00
## A2M                           0.03                    0.02
## A2ML1                         0.02                    0.04
## A3GALT2                       0.05                    0.58
## A4GALT                        0.00                    0.00
##         COVID_93_81y_female_ICU COVID_94_24y_female_NonICU
## A1BG                       0.37                       0.81
## A1CF                       0.00                       0.00
## A2M                        0.11                       0.17
## A2ML1                      0.00                       0.02
## A3GALT2                    0.05                       0.00
## A4GALT                     0.00                       0.06
##         COVID_95_49y_male_NonICU COVID_96_51y_male_NonICU COVID_97_76y_male_ICU
## A1BG                        0.37                     1.61                  0.19
## A1CF                        0.01                     0.00                  0.00
## A2M                         0.20                     0.02                  0.02
## A2ML1                       0.02                     0.00                  0.05
## A3GALT2                     0.15                     0.00                  0.12
## A4GALT                      0.00                     0.00                  0.03
##         COVID_98_81y_male_NonICU COVID_99_71y_male_ICU
## A1BG                        0.78                  0.33
## A1CF                        0.00                  0.00
## A2M                         0.26                  0.02
## A2ML1                       0.00                  0.00
## A3GALT2                     0.37                  0.04
## A4GALT                      0.00                  0.00
##         COVID_100_74y_female_NonICU COVID_101_58y_male_ICU
## A1BG                           0.30                   0.33
## A1CF                           0.00                   0.00
## A2M                            0.09                   0.11
## A2ML1                          0.00                   0.03
## A3GALT2                        0.04                   0.05
## A4GALT                         0.00                   0.00
##         COVID_102_84y_male_NonICU COVID_103_83y_male_NonICU
## A1BG                         0.12                      0.20
## A1CF                         0.00                      0.00
## A2M                          0.01                      0.03
## A2ML1                        0.01                      0.03
## A3GALT2                      0.00                      0.04
## A4GALT                       0.07                      0.00
##         NONCOVID_01_54y_female_NonICU NONCOVID_02_65y_male_ICU
## A1BG                             0.89                     0.32
## A1CF                             0.00                     0.00
## A2M                              0.04                     0.01
## A2ML1                            0.00                     0.00
```

```
## A3GALT2                                   0.00                     0.04
## A4GALT                                    0.00                     0.00
##         NONCOVID_03_65y_male_ICU NONCOVID_04_90y_male_NonICU
## A1BG                        0.44                        0.21
## A1CF                        0.00                        0.00
## A2M                         0.05                        0.05
## A2ML1                       0.02                        0.00
## A3GALT2                     0.04                        0.21
## A4GALT                      0.00                        0.00
##         NONCOVID_05_83y_female_NonICU NONCOVID_06_75y_female_ICU
## A1BG                             0.31                       0.89
## A1CF                             0.00                       0.00
## A2M                              0.01                       0.14
## A2ML1                            0.01                       0.01
## A3GALT2                          0.00                       0.00
## A4GALT                           0.00                       0.06
##         NONCOVID_07_50y_male_ICU NONCOVID_08_53y_female_ICU
## A1BG                        0.45                        0.47
## A1CF                        0.00                        0.01
## A2M                         0.07                        0.04
## A2ML1                       0.02                        0.00
## A3GALT2                     0.00                        0.15
## A4GALT                      0.00                        0.00
##         NONCOVID_09_49y_female_NonICU NONCOVID_10_67y_male_ICU
## A1BG                             0.40                       0.33
## A1CF                             0.00                       0.00
## A2M                              0.04                       0.05
## A2ML1                            0.00                       0.01
## A3GALT2                          0.00                       0.23
## A4GALT                           0.00                       0.08
##         NONCOVID_11_58y_female_NonICU NONCOVID_12_82y_male_ICU
## A1BG                             0.58                       0.12
## A1CF                             0.00                       0.00
## A2M                              0.03                       0.02
## A2ML1                            0.00                       0.00
## A3GALT2                          0.00                       0.00
## A4GALT                           0.00                       0.02
##         NONCOVID_13_65y_male_ICU NONCOVID_14_75y_female_ICU
## A1BG                        0.31                        0.16
## A1CF                        0.00                        0.00
## A2M                         0.04                        0.08
## A2ML1                       0.01                        0.00
## A3GALT2                     0.32                        0.05
## A4GALT                      0.02                        0.02
##         NONCOVID_15_83y_unknown_ICU NONCOVID_16_40y_female_ICU
## A1BG                           0.59                       0.34
## A1CF                           0.00                       0.00
## A2M                            0.03                       0.07
## A2ML1                          0.04                       0.00
## A3GALT2                        0.00                       0.13
## A4GALT                         0.19                       0.00
##         NONCOVID_17_84y_female_ICU NONCOVID_18_88y_male_ICU
## A1BG                          0.37                       0.33
## A1CF                          0.00                       0.00
```

```
## A2M                                 0.07                     0.06
## A2ML1                               0.01                     0.00
## A3GALT2                             0.18                     0.00
## A4GALT                              0.00                     0.00
##          NONCOVID_19_66y_female_ICU NONCOVID_20_62y_female_ICU
## A1BG                                0.25                     0.20
## A1CF                                0.00                     0.00
## A2M                                 0.11                     0.01
## A2ML1                               0.00                     0.02
## A3GALT2                             0.04                     0.00
## A4GALT                              0.03                     0.07
##          NONCOVID_21_71y_male_NonICU NONCOVID_22_63y_male_NonICU
## A1BG                                 0.40                      0.30
## A1CF                                 0.00                      0.00
## A2M                                  0.04                      0.02
## A2ML1                                0.02                      0.02
## A3GALT2                              0.00                      0.00
## A4GALT                               0.00                      0.00
##          NONCOVID_23_42y_female_NonICU NONCOVID_24_32y_female_NonICU
## A1BG                                   0.70                        0.75
## A1CF                                   0.00                        0.00
## A2M                                    0.02                        0.27
## A2ML1                                  0.01                        0.00
## A3GALT2                                0.00                        0.06
## A4GALT                                 0.00                        0.00
##          NONCOVID_25_62y_male_NonICU NONCOVID_26_36y_male_ICU
## A1BG                                 2.80                    0.22
## A1CF                                 0.00                    0.00
## A2M                                  0.04                    0.28
## A2ML1                                0.00                    0.00
## A3GALT2                              0.00                    0.00
## A4GALT                               0.00                    0.00
```

str(genes)

```
## 'data.frame':    100 obs. of  126 variables:
##  $ COVID_01_39y_male_NonICU    : num  0.49 0 0.21 0.04 0.07 ...
##  $ COVID_02_63y_male_NonICU    : num  0.29 0 0.14 0 0 ...
##  $ COVID_03_33y_male_NonICU    : num  0.26 0 0.03 0.02 0 ...
##  $ COVID_04_49y_male_NonICU    : num  0.45 0.01 0.09 0.07 0 ...
##  $ COVID_05_49y_male_NonICU    : num  0.17 0 0 0.05 0.07 0 0 8.45 1.17 0 ...
##  $ COVID_06_.y_male_NonICU     : num  0.21 0 0.08 0.04 0 0 0.03 19.6 3.15 0 ...
##  $ COVID_07_38y_female_NonICU  : num  0.49 0.01 0.23 0.03 0.07 ...
##  $ COVID_08_78y_male_ICU       : num  0.12 0 0.08 0.01 0 0 0 10.5 2.1 0 ...
##  $ COVID_09_64y_female_ICU     : num  0.51 0.01 0.88 0.02 0.79 ...
##  $ COVID_10_62y_male_ICU       : num  0.1 0 0.13 0.01 0.15 ...
##  $ COVID_11_52y_female_NonICU  : num  0.38 0.02 0.47 0.03 0.08 ...
##  $ COVID_12_50y_male_ICU       : num  0.45 0 0.16 0 1.75 0 0 16 3.61 0 ...
##  $ COVID_13_37y_male_NonICU    : num  0.18 0 0.07 0.01 0 0 0 22.1 2.73 0 ...
##  $ COVID_14_55y_male_ICU       : num  0.23 0 0.22 0.04 0.93 0 0.07 10.3 2.16 0 ...
##  $ COVID_15_68y_male_ICU       : num  0.42 0 0.07 0 0.15 0.03 0 9.37 2.94 0 ...
##  $ COVID_16_48y_male_NonICU    : num  0.41 0.01 0.58 0 0.19 ...
##  $ COVID_17_54y_male_NonICU    : num  0.63 0.02 0.15 0.02 0 ...
##  $ COVID_18_70y_female_NonICU  : num  0.47 0 0.3 0.02 0.06 ...
```

```
##  $ COVID_19_51y_male_NonICU      : num  0.33 0.02 0.11 0.02 0 ...
##  $ COVID_20_62y_male_ICU         : num  0.32 0 0.07 0 0.22 ...
##  $ COVID_21_66y_male_ICU         : num  0.18 0 0 0 0.37 0.03 0 7.1 1.11 0 ...
##  $ COVID_22_43y_male_ICU         : num  0.09 0 0.06 0 0.06 0 0.06 5.17 1.05 0 ...
##  $ COVID_23_76y_male_ICU         : num  0.18 0.01 0.03 0 0.07 0.03 0.04 8.87 1.45 0 ...
##  $ COVID_24_55y_male_ICU         : num  0.22 0.01 0.11 0.02 0.15 ...
##  $ COVID_25_55y_male_ICU         : num  0.29 0 0.09 0.03 0 ...
##  $ COVID_26_41y_female_ICU       : num  0.42 0 0.18 0 0.87 ...
##  $ COVID_27_71y_female_ICU       : num  0.16 0.01 0.23 0.01 0.18 ...
##  $ COVID_28_63y_male_ICU         : num  0.18 0 0.18 0.05 0.45 ...
##  $ COVID_29_63y_female_ICU       : num  0.35 0 0.03 0.03 0.15 0.03 0.08 9.74 1.57 0 ...
##  $ COVID_30_54y_male_ICU         : num  0.23 0 0.11 0.01 0 ...
##  $ COVID_31_50y_male_ICU         : num  0.15 0 0.47 0 0 0.03 0 10.4 1.74 0 ...
##  $ COVID_32_72y_male_ICU         : num  0.34 0.01 0.04 0 0.29 0 0.04 8.96 1.88 0 ...
##  $ COVID_33_81y_male_NonICU      : num  0.35 0 0.3 0.06 0.26 ...
##  $ COVID_34_64y_female_NonICU    : num  0.36 0 0.11 0 0.12 ...
##  $ COVID_35_58y_female_NonICU    : num  0.26 0 0.51 0.02 0.16 ...
##  $ COVID_36_68y_male_NonICU      : num  0.18 0.01 0.09 0 0.08 ...
##  $ COVID_37_87y_male_NonICU      : num  0.2 0 0.09 0.07 0.31 ...
##  $ COVID_38_68y_male_ICU         : num  0.29 0 0.1 0.02 0.35 ...
##  $ COVID_39_80y_female_ICU       : num  0.19 0 0.27 0 0 ...
##  $ COVID_40_66y_male_ICU         : num  0.22 0 0.17 0 0.08 0 0 14.6 2.47 0 ...
##  $ COVID_41_74y_male_ICU         : num  0.19 0 0.14 0 0.19 0 0 6.63 1.21 0 ...
##  $ COVID_42_21y_female_ICU       : num  0.24 0.01 0.33 0.01 0.39 0 0 15.1 2.23 0 ...
##  $ COVID_43_83y_female_ICU       : num  0.29 0 0 0 0.11 0 0 5.78 1.44 0 ...
##  $ COVID_44_46y_male_ICU         : num  0.22 0 0.14 0 0 0.04 0 10.8 2.03 0 ...
##  $ COVID_45_62y_female_ICU       : num  0.14 0 0.15 0.03 0.19 0 0 5.36 1.26 0 ...
##  $ COVID_46_62y_male_ICU         : num  0.53 0.01 0.1 0 0.06 ...
##  $ COVID_47_78y_male_ICU         : num  0.08 0.01 0.04 0.03 0.6 ...
##  $ COVID_48_72y_female_ICU       : num  0.19 0 0.06 0.01 0.23 ...
##  $ COVID_49_73y_male_ICU         : num  0.48 0 0.09 0.03 0 ...
##  $ COVID_50_37y_male_ICU         : num  0.08 0 0.01 0 0 0.72 0 6.16 0.62 0 ...
##  $ COVID_51_58y_female_NonICU    : num  0.21 0 0.13 0 0 ...
##  $ COVID_52_71y_male_NonICU      : num  0.25 0.01 0 0.03 0 ...
##  $ COVID_53_35y_female_NonICU    : num  0.25 0 0.64 0.1 0 ...
##  $ COVID_55_62y_female_ICU       : num  0.09 0 0.09 0.01 0 ...
##  $ COVID_56_33y_female_NonICU    : num  0.28 0 0.16 0.09 0.23 ...
##  $ COVID_57_30y_female_NonICU    : num  0.42 0 0.27 0.01 0.19 ...
##  $ COVID_58_62y_male_NonICU      : num  0.39 0 0.08 0 0 ...
##  $ COVID_59_55y_male_NonICU      : num  0.33 0 0.1 0 0.07 ...
##  $ COVID_60_49y_male_NonICU      : num  0.22 0 0.14 0 0 ...
##  $ COVID_61_54y_female_NonICU    : num  0.25 0 0.1 0.03 0.13 0 0 19.8 3.67 0 ...
##  $ COVID_62_78y_female_ICU       : num  0.21 0 0.04 0 0.05 ...
##  $ COVID_63_39y_female_ICU       : num  0.29 0 0.01 0 0.14 ...
##  $ COVID_64_65y_male_ICU         : num  0.38 0.01 0.04 0.02 0.56 0 0.04 9.99 2.14 0 ...
##  $ COVID_65_84y_male_NonICU      : num  0.4 0.01 0.07 0 0.58 ...
##  $ COVID_66_66y_female_NonICU    : num  0.64 0 0 0 0 ...
##  $ COVID_67_57y_male_ICU         : num  0.37 0 0.35 0 0 ...
##  $ COVID_68_79y_male_ICU         : num  0.58 0 0.15 0.01 0 ...
##  $ COVID_69_77y_female_NonICU    : num  0.52 0 0.29 0.02 0 0 0 23.4 4.18 0 ...
##  $ COVID_70_81y_male_NonICU      : num  0.27 0 0.07 0 0 ...
##  $ COVID_71_37y_male_ICU         : num  0.07 0.01 0.12 0.01 0 ...
##  $ COVID_72_50y_female_NonICU    : num  0.52 0 0.1 0.01 0 ...
##  $ COVID_73_82y_male_NonICU      : num  0.46 0.01 0.02 0.02 0.17 ...
```

```
##  $ COVID_74_55y_female_ICU     : num  0.24 0 0.12 0.02 0.26 ...
##  $ COVID_75_55y_male_NonICU     : num  0.23 0.01 0.14 0 0 ...
##  $ COVID_76_73y_female_ICU     : num  0.17 0 0.09 0.01 0.04 0 0.04 7.88 0.83 0 ...
##  $ COVID_77_55y_female_ICU     : num  0.05 0 0.01 0 0 ...
##  $ COVID_78_80y_male_NonICU     : num  0.19 0 0.2 0 0 ...
##  $ COVID_79_27y_male_NonICU     : num  0.08 0.01 0.03 0 0 ...
##  $ COVID_80_71y_male_ICU       : num  0.28 0 0.05 0 0.05 ...
##  $ COVID_82_67y_male_NonICU     : num  0.39 0.01 0.1 0 0 0 0 17.1 2.31 0 ...
##  $ COVID_83_85y_female_NonICU   : num  0.47 0 0.18 0.05 0 ...
##  $ COVID_84_75y_female_NonICU   : num  0.35 0 0.03 0 0.17 ...
##  $ COVID_85_62y_male_ICU       : num  0.29 0 0.04 0 0 ...
##  $ COVID_86_52y_female_NonICU   : num  0.6 0 0.27 0.02 0 ...
##  $ COVID_87_61y_male_ICU       : num  0.65 0 0.15 0 0 ...
##  $ COVID_89_90y_female_NonICU   : num  0.2 0 0.07 0.03 0.14 0 0 14.8 1.67 0 ...
##  $ COVID_90_86y_female_NonICU   : num  0.4 0 0.05 0.01 0.31 ...
##  $ COVID_91_29y_female_NonICU   : num  0.6 0 0.03 0.02 0.05 ...
##  $ COVID_92_82y_female_ICU     : num  0.34 0 0.02 0.04 0.58 ...
##  $ COVID_93_81y_female_ICU     : num  0.37 0 0.11 0 0.05 ...
##  $ COVID_94_24y_female_NonICU   : num  0.81 0 0.17 0.02 0 ...
##  $ COVID_95_49y_male_NonICU     : num  0.37 0.01 0.2 0.02 0.15 ...
##  $ COVID_96_51y_male_NonICU     : num  1.61 0 0.02 0 0 ...
##  $ COVID_97_76y_male_ICU       : num  0.19 0 0.02 0.05 0.12 ...
##  $ COVID_98_81y_male_NonICU     : num  0.78 0 0.26 0 0.37 ...
##  $ COVID_99_71y_male_ICU       : num  0.33 0 0.02 0 0.04 0 0 9.76 1.11 0 ...
##  $ COVID_100_74y_female_NonICU  : num  0.3 0 0.09 0 0.04 0 0.02 18.4 1.84 0 ...
##  $ COVID_101_58y_male_ICU      : num  0.33 0 0.11 0.03 0.05 ...
##  $ COVID_102_84y_male_NonICU    : num  0.12 0 0.01 0.01 0 0.07 0 9.1 1.06 0 ...
##   [list output truncated]
```

```r
# from the str() we can find continuous covariate and categorical covariates
library(stringr)
library(dbplyr)
sample.names <- colnames(genes)

# continuous covariate is age
# extract age from names
# eg.name is COVUD_01_39y_male_NonICU
# age: we ask for find the "_"before age and y after age, so that we can extract ages from name
age <- as.numeric(str_extract(sample.names, "(?<=_)[0-9]+(?=y)"))
age
```

```
##   [1] 39 63 33 49 49 NA 38 78 64 62 52 50 37 55 68 48 54 70 51 62 66 43 76 55 55
##  [26] 41 71 63 63 54 50 72 81 64 58 68 87 68 80 66 74 21 83 46 62 62 78 72 73 37
##  [51] 58 71 35 62 33 30 62 55 49 54 78 39 65 84 66 57 79 77 81 37 50 82 55 55 73
##  [76] 55 80 27 71 67 85 75 62 52 61 90 86 29 82 81 24 49 51 76 81 71 74 58 84 83
## [101] 54 65 65 90 83 75 50 53 49 67 58 82 65 75 83 40 84 88 66 62 71 63 42 32 62
## [126] 36
```

```r
# two categorical covariates are sex (male/female) and status (ICU/NonICU)
# sex:  we ask for find the "_"before sex and status after sex to extract
sex <- str_extract(sample.names,"(?<=_)[a-zA-Z]+(?=_ICU|_NonICU)")
sex
```

```
##   [1] "male"    "male"    "male"    "male"    "male"    "male"    "female"
```

```
##     [8] "male"    "female"  "male"    "female"  "male"    "male"    "male"
##    [15] "male"    "male"    "male"    "female"  "male"    "male"    "male"
##    [22] "male"    "male"    "male"    "male"    "female"  "female"  "male"
##    [29] "female"  "male"    "male"    "male"    "male"    "female"  "female"
##    [36] "male"    "male"    "male"    "female"  "male"    "male"    "female"
##    [43] "female"  "male"    "female"  "male"    "male"    "female"  "male"
##    [50] "male"    "female"  "male"    "female"  "female"  "female"  "female"
##    [57] "male"    "male"    "male"    "female"  "female"  "female"  "male"
##    [64] "male"    "female"  "male"    "male"    "female"  "male"    "male"
##    [71] "female"  "male"    "female"  "male"    "female"  "female"  "male"
##    [78] "male"    "male"    "male"    "female"  "female"  "male"    "female"
##    [85] "male"    "female"  "female"  "female"  "female"  "female"  "female"
##    [92] "male"    "male"    "male"    "male"    "male"    "female"  "male"
##    [99] "male"    "male"    "female"  "male"    "male"    "male"    "female"
##   [106] "female"  "male"    "female"  "female"  "male"    "female"  "male"
##   [113] "male"    "female"  "unknown" "female"  "female"  "male"    "female"
##   [120] "female"  "male"    "male"    "female"  "female"  "male"    "male"
```

```r
# status:extract either NonICU or ICU
status <- ifelse(str_detect(sample.names,"NonICU$"), "ICU", "NonICU")
status
```

```
##     [1] "ICU"    "ICU"    "ICU"    "ICU"    "ICU"    "ICU"    "ICU"    "NonICU"
##     [9] "NonICU" "NonICU" "ICU"    "NonICU" "ICU"    "NonICU" "NonICU" "ICU"
##    [17] "ICU"    "ICU"    "ICU"    "NonICU" "NonICU" "NonICU" "NonICU" "NonICU"
##    [25] "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU"
##    [33] "ICU"    "ICU"    "ICU"    "ICU"    "ICU"    "NonICU" "NonICU" "NonICU"
##    [41] "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU"
##    [49] "NonICU" "NonICU" "ICU"    "ICU"    "ICU"    "NonICU" "ICU"    "ICU"
##    [57] "ICU"    "ICU"    "ICU"    "ICU"    "NonICU" "NonICU" "NonICU" "ICU"
##    [65] "ICU"    "NonICU" "NonICU" "ICU"    "ICU"    "NonICU" "ICU"    "ICU"
##    [73] "NonICU" "ICU"    "NonICU" "NonICU" "ICU"    "ICU"    "NonICU" "ICU"
##    [81] "ICU"    "ICU"    "NonICU" "ICU"    "NonICU" "ICU"    "ICU"    "ICU"
##    [89] "NonICU" "NonICU" "ICU"    "ICU"    "ICU"    "NonICU" "ICU"    "NonICU"
##    [97] "ICU"    "NonICU" "ICU"    "ICU"    "ICU"    "NonICU" "NonICU" "ICU"
##   [105] "ICU"    "NonICU" "NonICU" "NonICU" "ICU"    "NonICU" "ICU"    "NonICU"
##   [113] "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU" "NonICU"
##   [121] "ICU"    "ICU"    "ICU"    "ICU"    "ICU"    "NonICU"
```

```r
# gene.AATF: get the AATF value from the genes dataset
# since all the values are numbers, so we use as.numeric
gene.AATF <- as.numeric(unlist(genes["AATF", ]))
head(gene.AATF)
```

```
## [1] 45.83 39.37 42.35 41.92 30.56 36.30
```

```r
# use function data frame put them in to a new summary dataset covid.data
covid.data <- data.frame(
  sample = sample.names,
  sex = sex,
  age = age,
  status = status,
  gene.value = gene.AATF
```
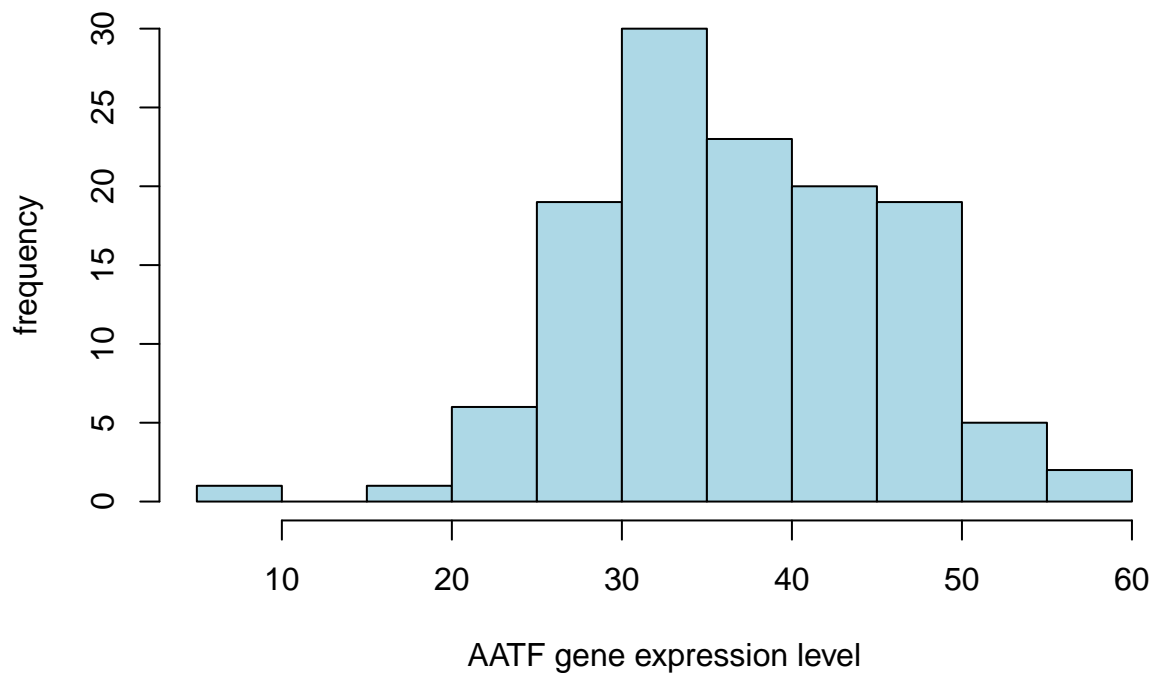
```
)

head(covid.data)
```

```
##                      sample   sex age status gene.value
## 1 COVID_01_39y_male_NonICU male  39    ICU      45.83
## 2 COVID_02_63y_male_NonICU male  63    ICU      39.37
## 3 COVID_03_33y_male_NonICU male  33    ICU      42.35
## 4 COVID_04_49y_male_NonICU male  49    ICU      41.92
## 5 COVID_05_49y_male_NonICU male  49    ICU      30.56
## 6  COVID_06_.y_male_NonICU male  NA    ICU      36.30
```

```
# Histogram for gene expression
histogram.gene <- hist(covid.data$gene.value, main = "Histogram for AATF gene expression",
                       xlab = "AATF gene expression level",
                       ylab = "frequency",
                       col = "lightblue")
```


Histogram for AATF gene expression

```
# Scatter plot for gene expression and continuous covariate
library(ggplot2)
scatter.plot.gene <- ggplot(covid.data, aes(x = age, y = gene.value)) +
                  # add points to our scatter plot
                  geom_point() +
                  # change labels
```

```
                    labs(title = "Scatter plot for gene expression and age",
                    x = "Age (years)",y = "AATF gene expression level") +
                    theme_minimal()
scatter.plot.gene
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## ('geom_point()').
```

Scatter plot for gene expression and age


```
# Box plot of gene expression separated by both categorical covariates
library(ggplot2)
box.plot.gene <- ggplot(covid.data, aes(x = interaction(sex, status), y = gene.value)) +
                    # add box plot
                    geom_boxplot() +
                    # change labels
                    labs(title = "Boxplot of gene expression separated by sex and status",
                    x = "sex&status",y = "AATF gene expression level") +
                    theme_minimal()
box.plot.gene
```

Boxplot of gene expression separated by sex and status