

# The Impact on Job Interest of Beginner Programmers\*

TBD

TBD

08 April 2021

## Abstract

First sentence. Second sentence. Third sentence. Fourth sentence.

## 1 Introduction

## 2 Data

```
#### Message ####
# Show code for reviewers
#
# // Variables
# Y: MoneyForLearning
# X: Age, CityPopulation, CommuteTime, Income, MonthsProgramming, SchoolDegree
# c: Gender
#
# // Clean csv
# https://github.com/bonjwow/new-coders/blob/main/inputs/data/clean\_new-coders.csv
#
#
```

```
#### Get data ####
dfNewCoders <-
  readr::read_csv("../inputs/data/clean_new-coders.csv")
```

```
## Parsed with column specification:
## cols(
##   Gender = col_double(),
##   Age = col_double(),
##   CityPopulation = col_double(),
##   CommuteTime = col_double(),
##   Income = col_double(),
##   MonthsProgramming = col_double(),
##   SchoolDegree = col_double(),
##   MoneyForLearning = col_double()
## )
```

---

\*<https://github.com/bonjwow/new-coders>

```
#### Descriptive statistics ####
```

```
stargazer::stargazer(data.frame(dfNewCoders), type="text")
```

```
##
## =====
## Statistic      N      Mean      St. Dev.   Min   Pctl(25) Pctl(75)   Max
## -----
## Gender          7,022    0.174      0.379      0      0         0         1
## Age              7,022    29.774     7.684     13     25        33        71
## CityPopulation   7,022    1.221     0.777      0      1         2         2
## CommuteTime      7,022    2.199     1.451      0      1         3         5
## Income           7,022  42,966.890 59,162.290 6,000  17,000    55,000    1,000,000
## MonthsProgramming 7,022    23.986     46.496      0      3         26        744
## SchoolDegree     7,022    1.676     0.968      0      1         2         4
## MoneyForLearning 7,022  1,032.273  4,030.722    0      0        399      170,000
## -----
```

```
#### Correlation analysis ####
```

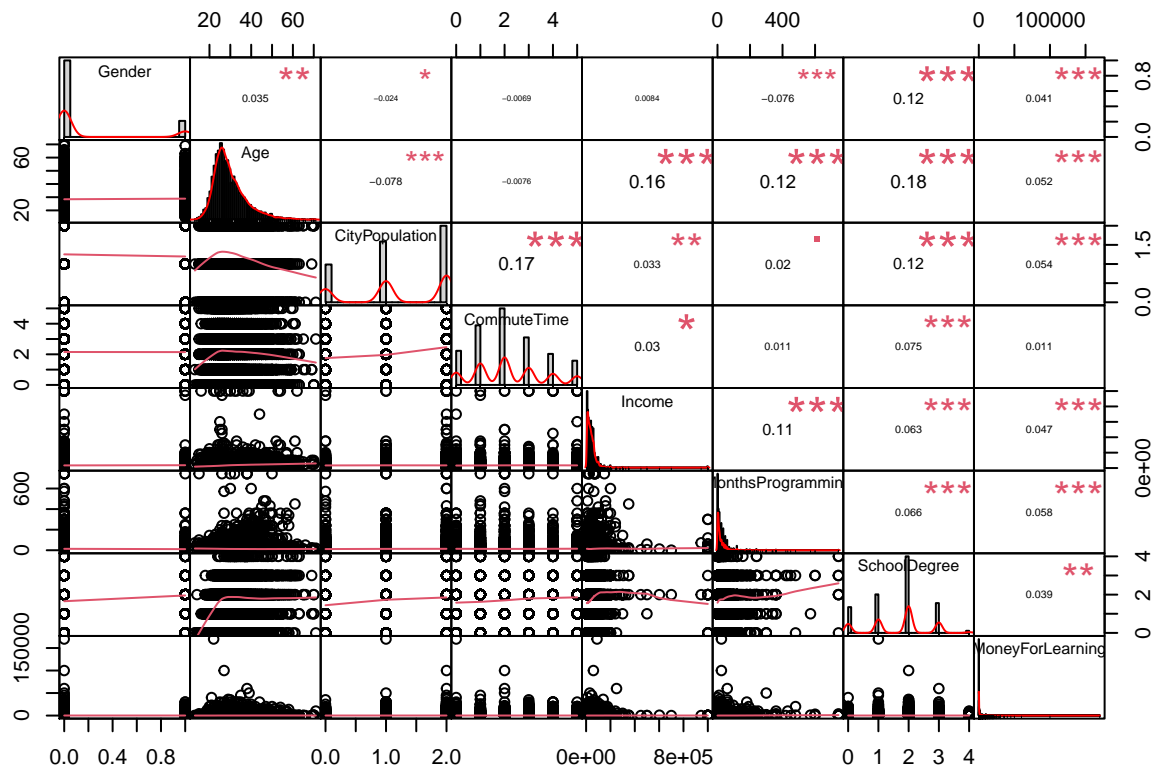
```
### Print correlation coefficient
```

```
round(cor(dfNewCoders), 3)
```

```
##          Gender    Age CityPopulation CommuteTime Income
## Gender          1.000  0.035      -0.024      -0.007  0.008
## Age              0.035  1.000      -0.078      -0.008  0.156
## CityPopulation   -0.024 -0.078       1.000       0.172  0.033
## CommuteTime      -0.007 -0.008       0.172       1.000  0.030
## Income            0.008  0.156       0.033       0.030  1.000
## MonthsProgramming -0.076  0.125       0.020       0.011  0.106
## SchoolDegree      0.115  0.181       0.119       0.075  0.063
## MoneyForLearning  0.041  0.052       0.054       0.011  0.047
##
##          MonthsProgramming SchoolDegree MoneyForLearning
## Gender          -0.076       0.115         0.041
## Age              0.125       0.181         0.052
## CityPopulation   0.020       0.119         0.054
## CommuteTime      0.011       0.075         0.011
## Income            0.106       0.063         0.047
## MonthsProgramming 1.000       0.066         0.058
## SchoolDegree      0.066       1.000         0.039
## MoneyForLearning  0.058       0.039         1.000
```

```
### Print correlation matrix
```

```
PerformanceAnalytics::chart.Correlation(dfNewCoders,
                                         histogram = TRUE,
                                         pch = 19)
```



### 3 Model

```
#### Multiple regression analysis ####
```

```
summary(fit.lm <- lm(formula =
    MoneyForLearning ~
    Gender +
    CityPopulation +
    CommuteTime +
    Income +
    MonthsProgramming +
    SchoolDegree,
    data = dfNewCoders))
```

```
##
```

```
## Call:
```

```
## lm(formula = MoneyForLearning ~ Gender + CityPopulation + CommuteTime +
##     Income + MonthsProgramming + SchoolDegree, data = dfNewCoders)
```

```
##
```

```
## Residuals:
```

```
##      Min      1Q  Median      3Q      Max
## -4831  -1081   -799   -438 169046
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.594e+02  1.339e+02   1.937 0.052798 .
## Gender         4.604e+02  1.278e+02   3.602 0.000318 ***
## CityPopulation  2.579e+02  6.294e+01   4.097 4.24e-05 ***
## CommuteTime    -5.922e-01  3.355e+01  -0.018 0.985917
## Income         2.553e-03  8.156e-04   3.130 0.001754 **
## MonthsProgramming 4.788e+00  1.041e+00   4.598 4.34e-06 ***
## SchoolDegree    9.240e+01  5.046e+01   1.831 0.067134 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4012 on 7015 degrees of freedom
## Multiple R-squared:  0.01027,    Adjusted R-squared:  0.009424
## F-statistic: 12.13 on 6 and 7015 DF,  p-value: 1.288e-13
```

#### Beta ####

```
QuantPsyc::lm.beta(fit.lm)
```

```
##              Gender    CityPopulation    CommuteTime    Income
##      0.0432618989    0.0497254551    -0.0002132173    0.0374740254
## MonthsProgramming    SchoolDegree
##      0.0552291292    0.0221842342
```

#### Check multicollinearity with Variance Inflation Factor (VIF) & Tolerance ####

### VIF

```
car::vif(fit.lm)
```

```
##              Gender    CityPopulation    CommuteTime    Income
##      1.022348    1.044332    1.034067    1.015859
## MonthsProgramming    SchoolDegree
##      1.022534    1.040387
```

### Tolerance

```
1/car::vif(fit.lm)
```

```
##              Gender    CityPopulation    CommuteTime    Income
##      0.9781404    0.9575497    0.9670552    0.9843882
## MonthsProgramming    SchoolDegree
##      0.9779630    0.9611803
```

## 4 Results

## 5 Discussion

# Appendix

## 6 References