Student ID 00010443

Ex 1.

ID 10443 (a = 0, b = 4, c = 4, d=3) will be deleting the following rows:

04+200 = 204,

44+100 = 144 and

43+10 = 53.

So, rows 53, 144 and 204 are deleted.

Section A

1. 0.7926

2. 47759.6494

3. 536.6093

4. 85.9375

5. –

6. 65.9950

7. 35269.0389

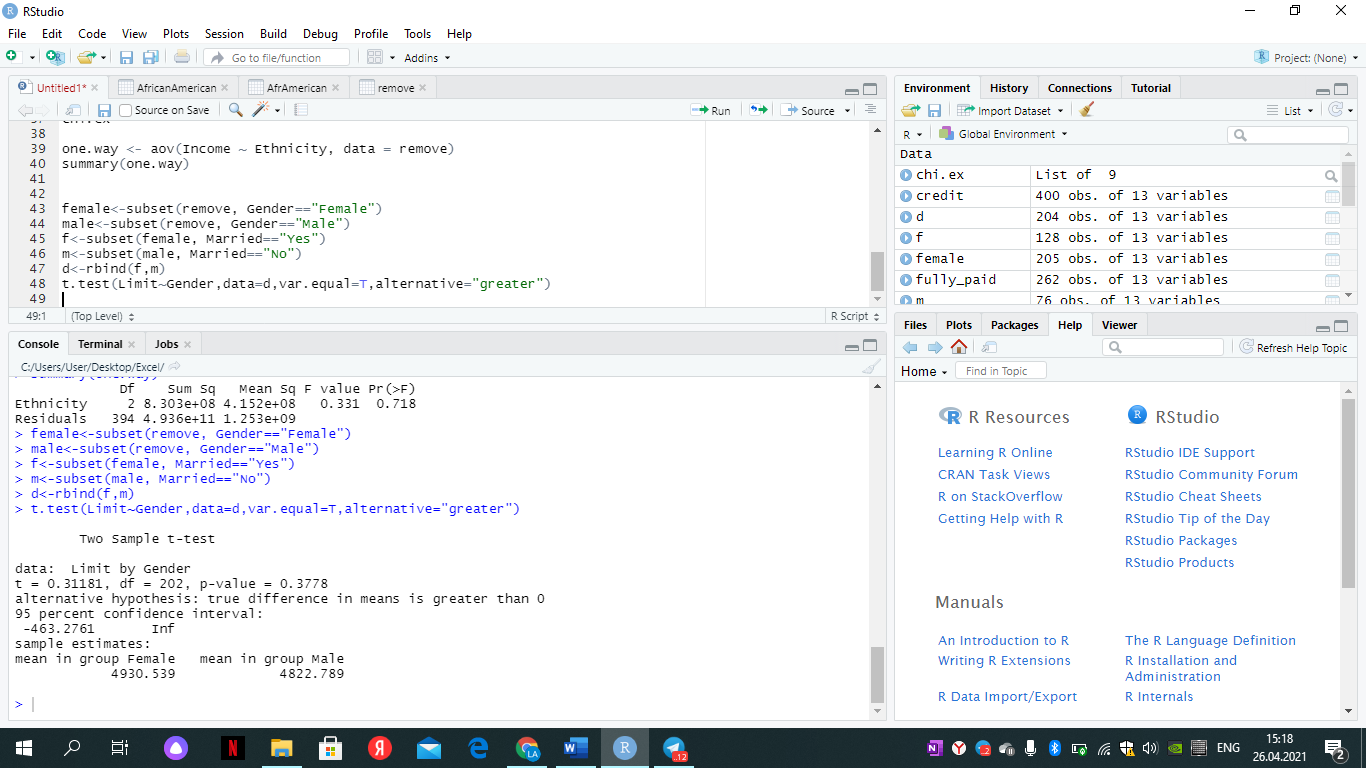
8. 549.4636

9. 47L

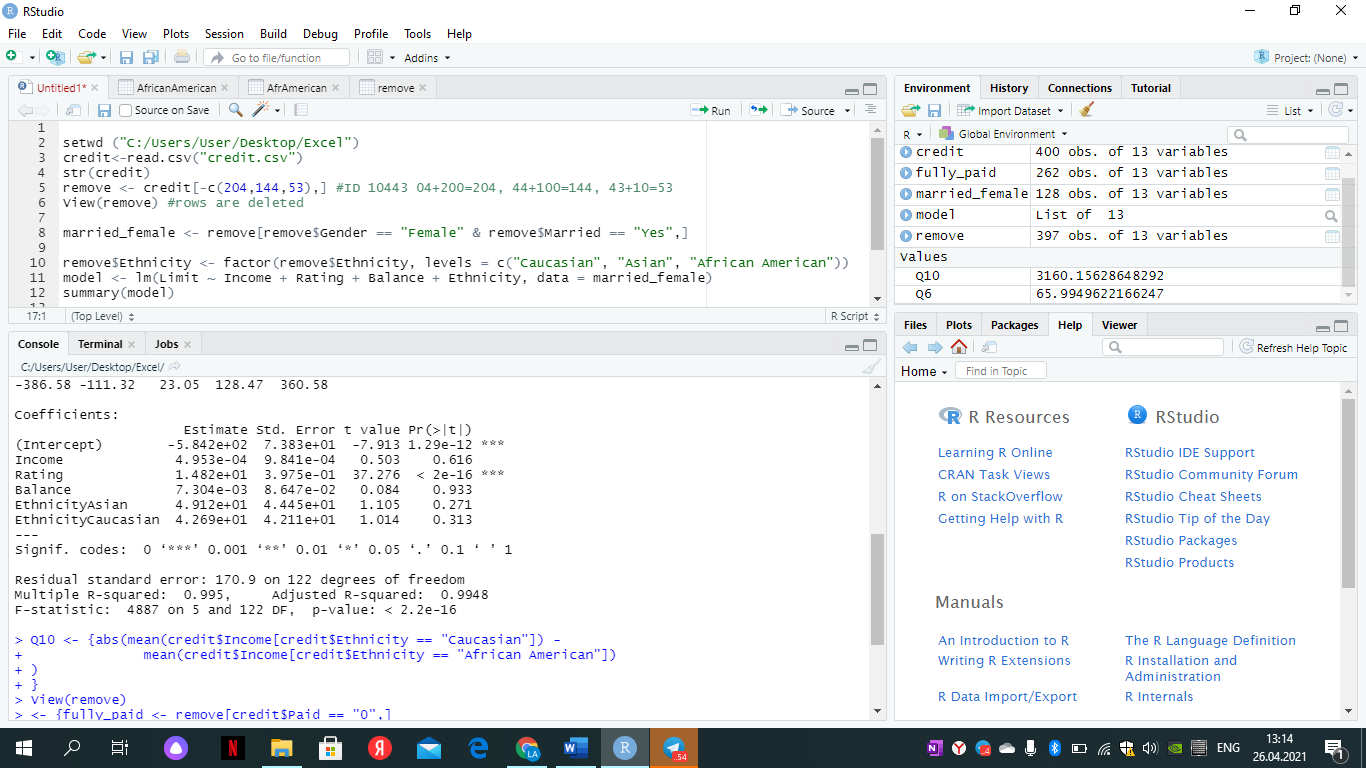
10. 3160.1563

Section B

1.



2.



Y=β0 +β1\* x1 + β2\* x2 + β3\* x3 + β4\* x4 + β5\*x5+ ε

Fitted Multiple linear Regression

A.

Credit limit = -5.842e+02 + Income \* β1 + Rating \* β2 + Balance\* β3 + Ethnicity Asian β4\* + Ethnicity Caucasian \* β5

B.

Credit limit = -5.415e + 02 + 4.953e-04 \* β1 + 1.482e+01 \* β2 + 7.304e-03 \* β3 + 4.912e+01 β4\* + 4.269e+01 \* β5

C.

Credit limit = -5.842e+02 + 912e+01 \* β4 (Asian)

Credit limit = -5.842e+02 + 269e+01 \* β5 (Caucasian)

Credit limit = -5.842e+02\* 0 + 0 (Afro-American)

D. Interpretations

1) Intercept is not equal to “0”

2) if β1 equal to “0” so significance level of income will equal to zero

3) if β2 equal to “0” so significance level of Rating will equal to zero

4) if β3 equal to “0”so significance level of Balance will equal to zero

5) if β4 equal to “0” so significance level of EA will equal to zero

6) if β5 equal to “0” so significance level of EAA will equal to zero

Rejection Rule

F\_stat < F\_crit

p-value < alpha

Hypothesis

Ho: There are no any significance on credit limit by all factors

Ha: At least one of them has significance

Decision: According to confidence level indicated in R result we reject Ho and accept Ha.

E.

Ho = There is no significance on limit by n factor

Ha = there is significance on limit by n factor

since the P-value of Rating is less than alpha=0.01

so Rating has significance on credit limit and reject Ho.

however others fail to reject Ho according to answers from answer of R studio.

3.

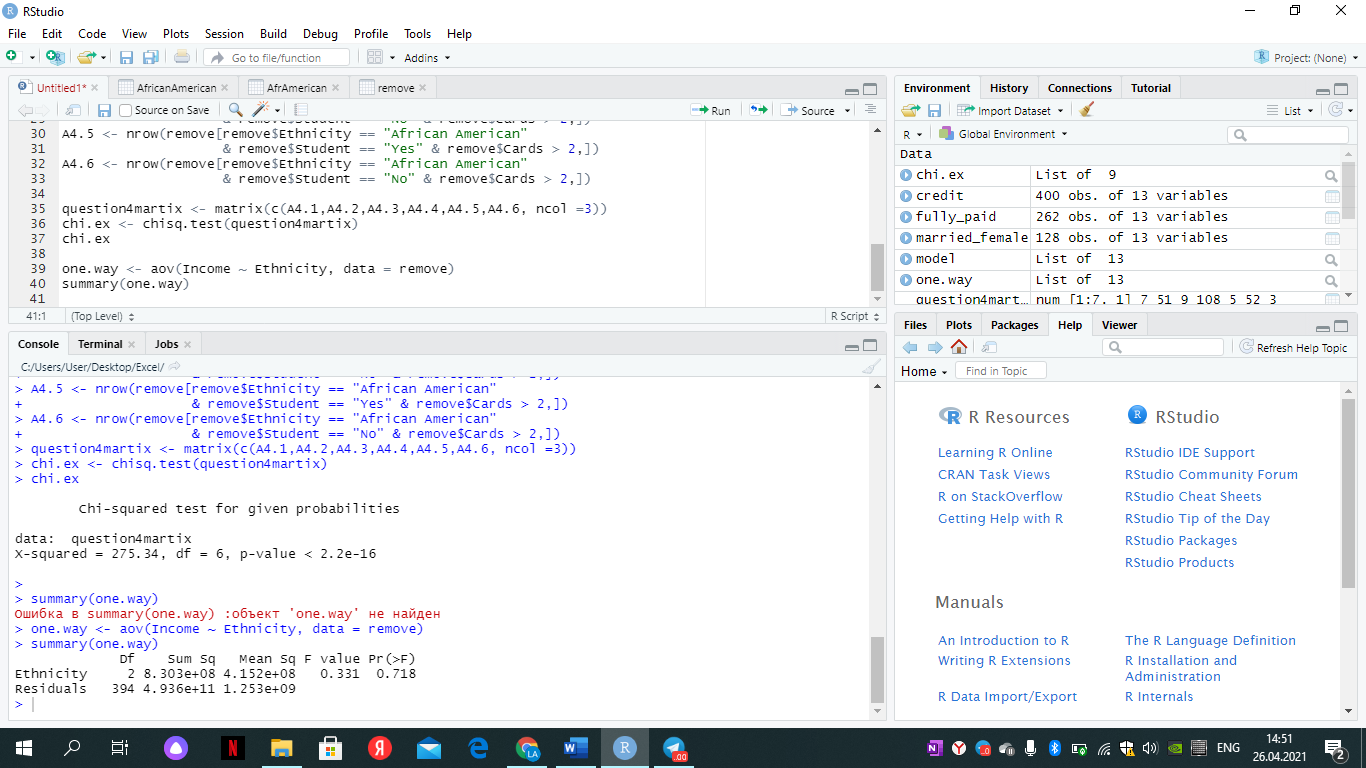
A.

One way Anova to solve this task before there are more than 2 sample observations.

Ho: µ1 = µ2

Ha: At least one mean is different

B.



Normality: Multiple regression assumes that the error terms are normally distributed. (Q-Q plot)

Linearity: There must be linear relationship between response variable and independent variables (Scatterplots).

No Multicollinearity: the independent variables are not highly correlated with each other (Correlation matrix).

Homoscedasticity: the variance of error terms are similar across the values of the independent variables (Plot of residuals vs predictor variables).

C.

Rejection Rule

Fs ˃ Fc p-value ˂ α

F-stat = 0.331

F-crit = 3.04

α = 0.06

D. Since Fs = 0.331 ˂ Fc = 3.04 or p-value = 0.718 ˃ α = 0.06 we fail to reject Ho which means there is no any significance between Ethnicity and income.

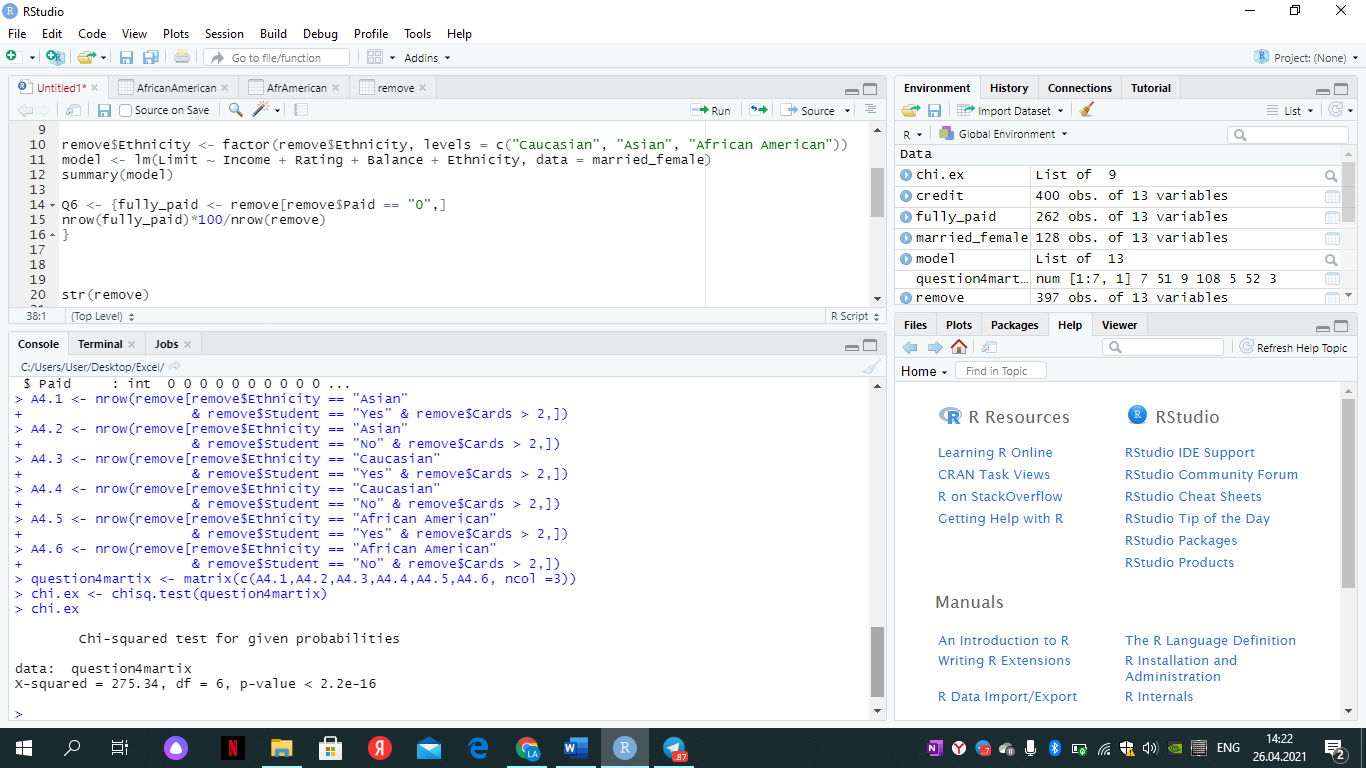
4.

A.

Ho: There is no association between the four variables.

Ha: There is an association between the four variables.

B.



C.

Chi-squared test for given probabilities

data: question4martix

X-squared = 275.34, df = 6, p-value < 2.2e-16

X-crit = 10.64

E.

since p-value = 2.2e-16 < Alpha=0.10

and X2=275.34 ≥ χ\_α^2=10.64

We reject Ho and accept Ha, which means there is an association between the four variables.