Econometrics Assingment
name: Anamika Kumari
Roll no: IED/10022/21

1.Reading and exporting files into R :

code:

```
###Complete the codes.The codes were written for a different data set.
##Make appropriate changes to extract the data we are using in the class.
##Install the necessary packages

rm(list = ls())
# Define the widths based on the provided layout
widths <- c(
  4,  # File Identification
  3,  # Schedule
  2,  # Quarter
  2,  # Visit
  1,  # Sector
  2,  # State/Ut Code
  2,  # District Code
  3,  # NSS-Region
  2,  # Stratum
  2,  # Sub-Stratum
  1,  # Sub-Sample
  4,  # Fod Sub-Region
  5,  # FSU
  1,  # Sample Sg/Sb No.
  1,  # Second Stage Stratum No.
  2,  # Sample Household Number
  2,  # Month of Survey
  1,  # Response Code
  1,  # Survey Code
  1   # Reason for Substitution of original household
)

# Define the column names
col.names <- c(
  "File_Identification",
  "Schedule",
  "Quarter",
  "Visit",
  "Sector",
  "State_Ut_Code",
  "District_Code",
  "NSS_Region",
  "Stratum",
  "Sub_Stratum",
  "Sub_Sample",
  "Fod_Sub_Region",
```

```
 "FSU",
 "Sample_Sg_Sb_No",
 "Second_Stage_Stratum_No",
 "Sample_Household_Number",
 "Month_of_Survey",
 "Response_Code",
 "Survey_Code",
 "Reason_for_Substitution"
)

# Read the data
level1 <- read.fwf(
 file = "C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT",
 widths = widths,
 col.names = col.names, n= 102063)

level1

## By including 'common-id'
level2<-read.fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FPER_FV.TXT",
          widths=c(35,2,5,2,5,3,1,1,1,1,1,1,2,8,8,50,3,3,10),
          col.names=c("common-id", "level","filler", "hhsize", "NIC", "NCO",
                  "hhtype","religion", "socialgrp", "latrinetype", "drainagetype",
                  "source_of_drink_water", "primary_source_of_cookingenergy",
"amt_med_insurance",
                  "hh_cons_exp", "blank","NSS", "NSC","MLT"),
          n=65932)
level2

#install.packages("readr")
library(readr)
## using the 'readr' package
level3<-read_fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT",
fwf_cols(fsuslno=c(4,8), sector=c(15, 15),
subblockno=c(32,32),sssno=c(33,33),hhno=c(34,35),level=c(36,37),filler=c(38,40),
personid=c(41,42), sex=c(43,43),ageatdeath=c(44,46),

medicalatn=c(47,47),hospitalised=c(48,48),numhospitalised=c(49,50),pregnant=c(51,51),timeof
death=c(52,52), nss=c(127,129), nsc=c(130,132), mlt=c(133, 142)),
          col_types =
cols(fsuslno=col_character(),sector=col_character(),subblockno=col_character(),sssno=col_char
acter(),hhno=col_character(),level = col_character(),filler =
col_character(),personid=col_character(), sex=col_integer(),ageatdeath=col_integer(),
                    medicalatn=col_integer(), hospitalised=col_integer(),
numhospitalised=col_integer(), pregnant=col_integer(), timeofdeath=col_integer(),
nss=col_character(), nsc=col_character(),mlt=col_character()))

level3
## using the 'readr' package
level4<-
read_fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FPER_FV.TXT",fwf_cols(fsuslno=c(4,8),
sector=c(15, 15),
subblockno=c(32,32),sssno=c(33,33),hhno=c(34,35),level=c(36,37),filler=c(38,40),
personid=c(41,42),rltntohead=c(43,43),sex=c(44,44), age=c(45,47),
```

maritalstatus=c(48,48),education=c(49,50),
resident_hostel=c(51,51),hospitalised=c(52,52),
numhospitalised=c(53,54),chronicailment=c(55,55),othrailment15days=c(56,56),othrailmentdaybefore=c(57,57),
healthscheme=c(58,58),reporting=c(59,59), nss=c(127,129),
nsc=c(130,132),mlt=c(133,142)),
col_types =
cols(fsuslno=col_character(),sector=col_character(),subblockno=col_character(),sssno=col_character(),hhno=col_character(),level = col_character(),filler =
col_character(),personid=col_character(), rltntohead=col_integer(),
sex=col_integer(),age=col_integer(),

maritalstatus=col_integer(),education=col_integer(),resident_hostel=col_integer(),hospitalised=col_integer(),numhospitalised=col_integer(),chronicailment=col_integer(),othrailment15days=col_integer(),othrailmentdaybefore=col_integer(),
healthscheme=col_integer(), reporting=col_integer(),nss=col_character(),
nsc=col_character(), mlt=col_character()))

```
level4


output:
> ###Complete the codes.The codes were written for a different data set.> ##Make
appropriate changes to extract the data we are using in the class.> ##Install the
 necessary packages> > rm(list = ls())> # Define the widths based on the provided
 layout> widths <- c(+   4,  # File Identification+   3,  # Schedule+   2,  # Qua
rter+   2,  # Visit+   1,  # Sector+   2,  # State/Ut Code+   2,  # District Code
+   3,  # NSS-Region+   2,  # Stratum+   2,  # Sub-Stratum+   1,  # Sub-Sample+
 4,  # Fod Sub-Region+   5,  # FSU+   1,  # Sample Sg/Sb No.+   1,  # Second Stag
e Stratum No.+   2,  # Sample Household Number+   2,  # Month of Survey+   1,  #
Response Code+   1,  # Survey Code+   1   # Reason for Substitution of original h
ousehold+ )> > # Define the column names> col.names <- c(+   "File_Identification
",+   "Schedule",+   "Quarter",+   "Visit",+   "Sector",+   "State_Ut_Code",+   "
District_Code",+   "NSS_Region",+   "Stratum",+   "Sub_Stratum",+   "Sub_Sample",
+   "Fod_Sub_Region",+   "FSU",+   "Sample_Sg_Sb_No",+   "Second_Stage_Stratum_No
",+   "Sample_Household_Number",+   "Month_of_Survey",+   "Response_Code",+   "Su
rvey_Code",+   "Reason_for_Substitution"+ )> > # Read the data> level1 <- read.fw
f(+   file = "C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT",+   widths = widths,+
 col.names = col.names+ )> level1   File_Identification Schedule Quarter Visit Se
ctor State_Ut_Code District_Code NSS_Region
1                 FVH1      104      Q1    V1      1             1            21
        11
2                 FVH1      104      Q1    V1      1             1            21
        11
3                 FVH1      104      Q1    V1      1             1            21
        11
4                 FVH1      104      Q1    V1      1             1            21
        11
5                 FVH1      104      Q1    V1      1             1            21
        11
6                 FVH1      104      Q1    V1      1             1            21
        11
7                 FVH1      104      Q1    V1      1             1            21
        11
8                 FVH1      104      Q1    V1      1             1            21
        11
9                 FVH1      104      Q1    V1      1             1            21
        11
10                FVH1      104      Q1    V1      1             1            21
        11
11                FVH1      104      Q1    V1      1             1            21
        11
12                FVH1      104      Q1    V1      1             1            21
        11
13                FVH1      104      Q1    V1      1             1            21
        11
14                FVH1      104      Q1    V1      1             1            21
        11
15                FVH1      104      Q1    V1      1             1            21
        11
16                FVH1      104      Q1    V1      1             1            21
        11
```

17		FVH1	104	Q1	V1	1	1	7
	11							
18		FVH1	104	Q1	V1	1	1	7
	11							
19		FVH1	104	Q1	V1	1	1	7
	11							
20		FVH1	104	Q1	V1	1	1	7
	11							
21		FVH1	104	Q1	V1	1	1	7
	11							
22		FVH1	104	Q1	V1	1	1	7
	11							
23		FVH1	104	Q1	V1	1	1	7
	11							
24		FVH1	104	Q1	V1	1	1	7
	11							
25		FVH1	104	Q1	V1	1	1	21
	11							
26		FVH1	104	Q1	V1	1	1	21
	11							
27		FVH1	104	Q1	V1	1	1	21
	11							
28		FVH1	104	Q1	V1	1	1	21
	11							
29		FVH1	104	Q1	V1	1	1	21
	11							
30		FVH1	104	Q1	V1	1	1	21
	11							
31		FVH1	104	Q1	V1	1	1	21
	11							
32		FVH1	104	Q1	V1	1	1	21
	11							
33		FVH1	104	Q1	V1	1	1	21
	11							
34		FVH1	104	Q1	V1	1	1	21
	11							
35		FVH1	104	Q1	V1	1	1	21
	11							
36		FVH1	104	Q1	V1	1	1	21
	11							
37		FVH1	104	Q1	V1	1	1	21
	11							
38		FVH1	104	Q1	V1	1	1	21
	11							
39		FVH1	104	Q1	V1	1	1	21
	11							
40		FVH1	104	Q1	V1	1	1	21
	11							
41		FVH1	104	Q1	V1	1	1	21
	11							
42		FVH1	104	Q1	V1	1	1	21
	11							
43		FVH1	104	Q1	V1	1	1	21
	11							
44		FVH1	104	Q1	V1	1	1	21
	11							
45		FVH1	104	Q1	V1	1	1	21
	11							
46		FVH1	104	Q1	V1	1	1	21
	11							
47		FVH1	104	Q1	V1	1	1	21
	11							
48		FVH1	104	Q1	V1	1	1	21
	11							
49		FVH1	104	Q1	V1	1	1	21
	11							
50		FVH1	104	Q1	V1	1	1	21
	11							

	Stratum	Sub_Stratum	Sub_Sample	Fod_Sub_Region	FSU	Sample_Sg_Sb_No	Second_Stage_Stratum_No
1	1	1	1	110	52742	1	1
2	1	1	1	110	52742	1	2
3	1	1	1	110	52742	1	2
4	1	1	1	110	52742	1	2
5	1	1	1	110	52742	1	2
6	1	1	1	110	52742	1	2

7	1	3	1	1	110	52742	1
8	1	3	1	1	110	52742	1
9	1	1	1	2	110	52871	1
10	1	1	1	2	110	52871	1
11	1	2	1	2	110	52871	1
12	1	2	1	2	110	52871	1
13	1	2	1	2	110	52871	1
14	1	2	1	2	110	52871	1
15	1	3	1	2	110	52871	1
16	1	3	2	1	110	50959	1
17	1	1	2	1	110	50959	1
18	1	1	2	1	110	50959	1
19	1	2	2	1	110	50959	1
20	1	2	2	1	110	50959	1
21	1	2	2	1	110	50959	1
22	1	2	2	1	110	50959	1
23	1	3	2	1	110	50959	1
24	1	3	2	2	110	56460	1
25	1	1	2	2	110	56460	1
26	1	1	2	2	110	56460	1
27	1	2	2	2	110	56460	1
28	1	2	2	2	110	56460	1
29	1	2	2	2	110	56460	1
30	1	2	2	2	110	56460	1
31	1	3	2	2	110	56460	1
32	1	3	3	1	110	59153	1
33	1	1	3	1	110	59153	1
34	1	2	3	1	110	59153	1
35	1	2	3	1	110	59153	1
36	1	3	3	1	110	59153	2
37	1	1	3	1	110	59153	2
38	1	2	3	1	110	59153	2
39	1	2	3	1	110	59153	2
40	1	3	3	2	110	53097	1
41	1	1	3	2	110	53097	1
42	1	2	3	2	110	53097	1
43	1	2	3	2	110	53097	1
44	1	3	3	2	110	53097	2
45	1	1	3	2	110	53097	2
46	1	2	3	2	110	53097	2
47	1	2	3	2	110	53097	2

```
48       1              3          2            110 53097                2
         3
49       1              4          1            110 56221                1
         1
50       1              4          1            110 56221                1
         2
   Sample_Household_Number Month_of_Survey Response_Code Survey_Code Reason_for_S
ubstitution
1                         1              8             1           1
        NA
2                         2              8             1           1
        NA
3                         1              8             1           1
        NA
4                         2              8             1           1
        NA
5                         3              8             3           1
        NA
6                         4              8             4           1
        NA
7                         1              8             3           1
        NA
8                         2              8             2           1
        NA
9                         1              8             2           1
        NA
10                        2              8             2           2
         2
11                        1              8             2           1
        NA
12                        2              8             2           1
        NA
13                        3              8             2           1
        NA
14                        4              8             2           1
        NA
15                        1              8             2           1
        NA
16                        2              8             1           1
        NA
17                        1              8             1           2
         2
18                        2              8             1           1
        NA
19                        1              8             1           1
        NA
20                        2              8             1           1
        NA
21                        3              8             2           1
        NA
22                        4              8             4           1
        NA
23                        1              8             2           1
        NA
24                        2              8             1           1
        NA
25                        1              8             1           1
        NA
26                        2              8             1           1
        NA
27                        1              8             1           1
        NA
28                        2              8             1           1
        NA
29                        3              8             1           1
        NA
30                        4              8             1           1
        NA
31                        1              8             2           1
        NA
32                        2              8             1           1
        NA
33                        1              8             1           1
        NA
34                        1              8             1           1
        NA
35                        2              8             1           1
        NA
36                        1              8             1           1
        NA
37                        1              8             2           1
        NA
```

```
38                      1               8               1           1
        NA
39                      2               8               1           1
        NA
40                      1               8               1           1
        NA
41                      1               8               2           1
        NA
42                      1               8               2           1
        NA
43                      2               8               2           1
        NA
44                      1               8               2           1
        NA
45                      1               8               2           1
        NA
46                      1               8               2           1
        NA
47                      2               8               2           1
        NA
48                      1               8               2           1
        NA
49                      1               8               2           1
        NA
50                      1               8               4           1
        NA
 [ reached 'max' / getOption("max.print") -- omitted 102063 rows ]> # Read the da
ta> level1 <- read.fwf(+   file = "C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT",+
  widths = widths,+   col.names = col.names, n= 102063)> )Error: unexpected ')' i
n ")"> level1   File_Identification Schedule Quarter Visit Sector State_Ut_Code D
istrict_Code NSS_Region
1                 FVH1      104      Q1    V1      1            1            21
        11
2                 FVH1      104      Q1    V1      1            1            21
        11
3                 FVH1      104      Q1    V1      1            1            21
        11
4                 FVH1      104      Q1    V1      1            1            21
        11
5                 FVH1      104      Q1    V1      1            1            21
        11
6                 FVH1      104      Q1    V1      1            1            21
        11
7                 FVH1      104      Q1    V1      1            1            21
        11
8                 FVH1      104      Q1    V1      1            1            21
        11
9                 FVH1      104      Q1    V1      1            1            21
        11
10                FVH1      104      Q1    V1      1            1            21
        11
11                FVH1      104      Q1    V1      1            1            21
        11
12                FVH1      104      Q1    V1      1            1            21
        11
13                FVH1      104      Q1    V1      1            1            21
        11
14                FVH1      104      Q1    V1      1            1            21
        11
15                FVH1      104      Q1    V1      1            1            21
        11
16                FVH1      104      Q1    V1      1            1            21
        11
17                FVH1      104      Q1    V1      1            1             7
        11
18                FVH1      104      Q1    V1      1            1             7
        11
19                FVH1      104      Q1    V1      1            1             7
        11
20                FVH1      104      Q1    V1      1            1             7
        11
21                FVH1      104      Q1    V1      1            1             7
        11
22                FVH1      104      Q1    V1      1            1             7
        11
23                FVH1      104      Q1    V1      1            1             7
        11
24                FVH1      104      Q1    V1      1            1             7
        11
25                FVH1      104      Q1    V1      1            1            21
        11
```

26	FVH1	104	Q1	V1	1	1	21	11
27	FVH1	104	Q1	V1	1	1	21	11
28	FVH1	104	Q1	V1	1	1	21	11
29	FVH1	104	Q1	V1	1	1	21	11
30	FVH1	104	Q1	V1	1	1	21	11
31	FVH1	104	Q1	V1	1	1	21	11
32	FVH1	104	Q1	V1	1	1	21	11
33	FVH1	104	Q1	V1	1	1	21	11
34	FVH1	104	Q1	V1	1	1	21	11
35	FVH1	104	Q1	V1	1	1	21	11
36	FVH1	104	Q1	V1	1	1	21	11
37	FVH1	104	Q1	V1	1	1	21	11
38	FVH1	104	Q1	V1	1	1	21	11
39	FVH1	104	Q1	V1	1	1	21	11
40	FVH1	104	Q1	V1	1	1	21	11
41	FVH1	104	Q1	V1	1	1	21	11
42	FVH1	104	Q1	V1	1	1	21	11
43	FVH1	104	Q1	V1	1	1	21	11
44	FVH1	104	Q1	V1	1	1	21	11
45	FVH1	104	Q1	V1	1	1	21	11
46	FVH1	104	Q1	V1	1	1	21	11
47	FVH1	104	Q1	V1	1	1	21	11
48	FVH1	104	Q1	V1	1	1	21	11
49	FVH1	104	Q1	V1	1	1	21	11
50	FVH1	104	Q1	V1	1	1	21	11

	Stratum	Sub_Stratum	Sub_Sample	Fod_Sub_Region	FSU	Sample_Sg_Sb_No	Second_Stage_Stratum_No
1	1	1	1	110	52742	1	1
2	1	1	1	110	52742	1	1
3	1	1	1	110	52742	1	2
4	1	1	1	110	52742	1	2
5	1	1	1	110	52742	1	2
6	1	1	1	110	52742	1	2
7	1	1	1	110	52742	1	3
8	1	1	1	110	52742	1	3
9	1	1	2	110	52871	1	1
10	1	1	2	110	52871	1	1
11	1	1	2	110	52871	1	2
12	1	1	2	110	52871	1	2
13	1	1	2	110	52871	1	2
14	1	1	2	110	52871	1	2
15	1	1	2	110	52871	1	3

```
16       1              1          2           110 52871               1
                3
17       1              2          1           110 50959               1
                1
18       1              2          1           110 50959               1
                1
19       1              2          1           110 50959               1
                2
20       1              2          1           110 50959               1
                2
21       1              2          1           110 50959               1
                2
22       1              2          1           110 50959               1
                2
23       1              2          1           110 50959               1
                3
24       1              2          1           110 50959               1
                3
25       1              2          2           110 56460               1
                1
26       1              2          2           110 56460               1
                1
27       1              2          2           110 56460               1
                2
28       1              2          2           110 56460               1
                2
29       1              2          2           110 56460               1
                2
30       1              2          2           110 56460               1
                2
31       1              2          2           110 56460               1
                3
32       1              2          2           110 56460               1
                3
33       1              3          1           110 59153               1
                1
34       1              3          1           110 59153               1
                2
35       1              3          1           110 59153               1
                2
36       1              3          1           110 59153               1
                3
37       1              3          1           110 59153               2
                1
38       1              3          1           110 59153               2
                2
39       1              3          1           110 59153               2
                2
40       1              3          1           110 59153               2
                3
41       1              3          2           110 53097               1
                1
42       1              3          2           110 53097               1
                2
43       1              3          2           110 53097               1
                2
44       1              3          2           110 53097               1
                3
45       1              3          2           110 53097               2
                1
46       1              3          2           110 53097               2
                2
47       1              3          2           110 53097               2
                2
48       1              3          2           110 53097               2
                3
49       1              4          1           110 56221               1
                1
50       1              4          1           110 56221               1
                2
   Sample_Household_Number Month_of_Survey Response_Code Survey_Code Reason_for_S
ubstitution
1                       1              8             1           1
         NA
2                       2              8             1           1
         NA
3                       1              8             1           1
         NA
4                       2              8             1           1
         NA
5                       3              8             3           1
         NA
```

6	NA	4	8	4	1
7	NA	1	8	3	1
8	NA	2	8	2	1
9	NA	1	8	2	1
10	2	2	8	2	2
11	NA	1	8	2	1
12	NA	2	8	2	1
13	NA	3	8	2	1
14	NA	4	8	2	1
15	NA	1	8	2	1
16	NA	2	8	1	1
17	2	1	8	1	2
18	NA	2	8	1	1
19	NA	1	8	1	1
20	NA	2	8	1	1
21	NA	3	8	2	1
22	NA	4	8	4	1
23	NA	1	8	2	1
24	NA	2	8	1	1
25	NA	1	8	1	1
26	NA	2	8	1	1
27	NA	1	8	1	1
28	NA	2	8	1	1
29	NA	3	8	1	1
30	NA	4	8	1	1
31	NA	1	8	2	1
32	NA	2	8	1	1
33	NA	1	8	1	1
34	NA	1	8	1	1
35	NA	2	8	1	1
36	NA	1	8	1	1
37	NA	1	8	2	1
38	NA	1	8	1	1
39	NA	2	8	1	1
40	NA	1	8	1	1
41	NA	1	8	2	1
42	NA	1	8	2	1
43	NA	2	8	2	1
44	NA	1	8	2	1
45	NA	1	8	2	1
46	NA	1	8	2	1

```
47                          2               8             2           1
        NA
48                          1               8             2           1
        NA
49                          1               8             2           1
        NA
50                          1               8             4           1
        NA
 [ reached 'max' / getOption("max.print") -- omitted 102013 rows ]> ## By includi
ng 'common-id'> level2<-read.fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FPER_FV.TX
T", +                 widths=c(35,2,5,2,5,3,1,1,1,1,1,2,8,8,50,3,3,10), +
           col.names=c("common-id", "level","filler", "hhsize", "NIC", "NCO", "
hhtype","religion", "socialgrp", "latrinetype", "drainagetype","source_of_drink_w
ater", "primary_source_of_cookingenergy", "amt_med_insurance", "hh_cons_exp", "bl
ank","NSS", "NSC","MLT"), +                   n=65932)> level2
         common.id level filler hhsize   NIC NCO hhtype religion socialgrp
1  FVP1104Q1V110121011010110105274211     1   1116      5 20801  10      NA
 NA        NA
2  FVP1104Q1V110121011010110105274211     1   2226      2 20601   8      NA
 NA        NA
3  FVP1104Q1V110121011010110105274211     1   3323      1 21001  12      NA
  6        NA
4  FVP1104Q1V110121011010110105274211     1   4413      4 21001  12      NA
  6        NA
5  FVP1104Q1V110121011010110105274211     2   1116      5 20801  10      NA
 NA        NA
6  FVP1104Q1V110121011010110105274211     2   2226      0 20601   6      NA
 NA        NA
7  FVP1104Q1V110121011010110105274211     2   3313      8 21201  15      NA
  6        NA
8  FVP1104Q1V110121011010110105274211     2   4423      2 21001  12      NA
  6        NA
9  FVP1104Q1V110121011010110105274211     2   5616     NA 10501  12       4
 NA        NA
10 FVP1104Q1V110121011010110105274211     2   6314      2 40701   8      NA
  6        NA
11 FVP1104Q1V110121011010110105274212     1   1115      5 20101   0      NA
  6        NA
12 FVP1104Q1V110121011010110105274212     1   2225      0 20101   0      NA
  6        NA
13 FVP1104Q1V110121011010110105274212     1   3511      3 10701  82       6
  6        NA
14 FVP1104Q1V110121011010110105274212     1   4511      0 10501  42       4
 NA        NA
15 FVP1104Q1V110121011010110105274212     1   5521      8 10701  91       4
  6        NA
16 FVP1104Q1V110121011010110105274212     1   6521      6 10701  92       6
  6        NA
17 FVP1104Q1V110121011010110105274212     1   7521      3 10601  72       5
  6        NA
18 FVP1104Q1V110121011010110105274212     2   1114      5 20101   0      NA
  6        NA
19 FVP1104Q1V110121011010110105274212     2   2223      8 20701   9      NA
  6        NA
20 FVP1104Q1V110121011010110105274212     2   3511      5 10701  82       6
  6        NA
21 FVP1104Q1V110121011010110105274212     2   4521      3 10601  72       5
  6        NA
22 FVP1104Q1V110121011010110105274212     2   5521      8 11001 123       1
  6        NA
23 FVP1104Q1V110121011010110105274212     3   1119      2 20101   0      NA
 NA        NA
24 FVP1104Q1V110121011010110105274212     3   2228      5 20101   0      NA
 NA        NA
25 FVP1104Q1V110121011010110105274212     3   3315      0 20701   9      NA
  6        NA
26 FVP1104Q1V110121011010110105274212     3   4424      4 20701   9      NA
  6        NA
27 FVP1104Q1V110121011010110105274212     3   5511      9 10801 112       7
  6        NA
28 FVP1104Q1V110121011010110105274212     3   6511      4 10601  82       5
  6        NA
29 FVP1104Q1V110121011010110105274212     4   1118      5 30601   6      NA
 NA        NA
30 FVP1104Q1V110121011010110105274212     4   2314      4 20701   8      NA
  6        NA
31 FVP1104Q1V110121011010110105274212     4   3424      8 20701   8      NA
  6        NA
32 FVP1104Q1V110121011010110105274212     4   4611      5 10701  82       6
  6        NA
33 FVP1104Q1V110121011010110105274212     4   5611      7 10701  92       6
  6        NA
```

```
34 FVP1104Q1V110121011010111011052742124    6621       9 11001 123       1
  6         NA
35 FVP1104Q1V110121011010110110052742134    1113       6 20601   6      NA
  3         NA
36 FVP1104Q1V110121011010110110052742134    2223       3 20701   8      NA
  6         NA
37 FVP1104Q1V110121011010110110052742134    3511       6 10701  92       6
  6         NA
38 FVP1104Q1V110121011010110110052742134    4521       4 10701  92       6
  6         NA
39 FVP1104Q1V110121011010110110052742134    5521       2 10601  72       5
  6         NA
40 FVP1104Q1V110121011010110110052742134    1117       6 20101   0      NA
 NA         NA
41 FVP1104Q1V110121011010110110052742134    2226       7 20101   0      NA
 NA         NA
42 FVP1104Q1V110121011010110110052742134    3313       6 21001  12      NA
  6         NA
43 FVP1104Q1V110121011010110110052742134    4423       0 21001  12      NA
  6         NA
44 FVP1104Q1V110121011010110110052742134    5626      NA 10501  12       4
 NA         NA
45 FVP1104Q1V110121011010120110052871114    1115       4 20801  10      NA
  6         NA
46 FVP1104Q1V110121011010120110052871114    2225       2 20701   9      NA
  6         NA
47 FVP1104Q1V110121011010120110052871114    3522       7 11301 171       5
  6         NA
48 FVP1104Q1V110121011010120110052871114    4512       3 11201 151       5
  6         NA
49 FVP1104Q1V110121011010120110052871114    5522       1 11201 151       5
  6         NA
50 FVP1104Q1V110121011010120110052871114    6729       0 30101   0      NA
 NA         NA
51 FVP1104Q1V110121011010120110052871114    1116       0 20601   5      NA
 NA         NA
52 FVP1104Q1V110121011010120110052871114    2225       8 20101   0      NA
  6         NA
   latrinetype drainagetype source_of_drink_water primary_source_of_cookingenergy
1           NA           NA                    NA                                NA
2           NA           NA                    NA                                NA
3           NA           NA                    NA                                NA
4           NA           NA                    NA                                NA
5           NA           NA                    NA                                NA
6           NA           NA                    NA                                NA
7           NA           NA                    NA                                NA
8           NA           NA                    NA                                NA
9           NA           NA                    NA                                NA
10          NA           NA                    NA                                NA
11          NA           NA                    NA                                NA
12          NA           NA                    NA                                NA
13          NA           NA                    NA                                NA
14          NA           NA                    NA                                NA
15          NA           NA                    NA                                NA
16          NA           NA                    NA                                NA
17          NA           NA                    NA                                NA
18          NA           NA                    NA                                NA
19          NA           NA                    NA                                NA
20          NA           NA                    NA                                NA
21          NA           NA                    NA                                NA
22          NA           NA                    NA                                NA
23          NA           NA                    NA                                NA
24          NA           NA                    NA                                NA
25          NA           NA                    NA                                NA
26          NA           NA                    NA                                NA
27          NA           NA                    NA                                NA
28          NA           NA                    NA                                NA
29          NA           NA                    NA                                NA
30          NA           NA                    NA                                NA
31          NA           NA                    NA                                NA
32          NA           NA                    NA                                NA
33          NA           NA                    NA                                NA
34          NA           NA                    NA                                NA
35          NA           NA                    NA                                NA
36          NA           NA                    NA                                NA
37          NA           NA                    NA                                NA
38          NA           NA                    NA                                NA
39          NA           NA                    NA                                NA
40          NA           NA                    NA                                NA
41          NA           NA                    NA                                NA
42          NA           NA                    NA                                NA
43          NA           NA                    NA                                NA
```

```
44          NA           NA                     NA                                  NA
45          NA           NA                     NA                                  NA
46          NA           NA                     NA                                  NA
47          NA           NA                     NA                                  NA
48          NA           NA                     NA                                  NA
49          NA           NA                     NA                                  NA
50          NA           NA                     NA                                  NA
51          NA           NA                     NA                                  NA
52          NA           NA                     NA                                  NA
   amt_med_insurance hh_cons_exp                                                  bla
nk NSS NSC
1           11011116    112                    1101 6    0     0     0 6 01
10 1 6  NA
2                 93      2                    93   0    0     0     0 0 09
3    0  NA
3                 93      2                    93   0    0     0     0 0 09
3    0  NA
4           31110459    32226084 128           3111 8    0     0     0 8 03
11 1 8  NA
5                 94      2                    94   0    0     0     0 0 09
4    0  NA
6                 92      2                    92   0    0     0     0 0 09
2    0  NA
7           31851023    31226052 417           3185 8    0     0     0 8 03
18 5 8  NA
8                 93      2                    93   0    0     0     0 0 09
3    0  NA
9                 91      2                    91   0    0     0     0 0 09
1    0  NA
10          11011216    112                    1101 6    0     0     0 6 01
10 1 6  NA
11          11259327    22111011    1101121611 1125 6    0     0     0 6 01
12 5 6  NA
12                93      2                    93   0    0     0     0 0 09
3    0  NA
13                91      2                    91   0    0     0     0 0 09
1    0  NA
14                91      2                    91   0    0     0     0 0 09
1    0  NA
15                93      2                    93   0    0     0     0 0 09
3    0  NA
16                91      2                    91   0    0     0     0 0 09
1    0  NA
17                91      2                    91   0    0     0     0 0 09
1    0  NA
18          11310017    12214011               1131 8    0     0     0 8 01
13 1 8  NA
19                93      2                    93   0    0     0     0 0 09
3    0  NA
20                91      2                    91   0    0     0     0 0 09
1    0  NA
21                91      2                    91   0    0     0     0 0 09
1    0  NA
22                91      2                    91   0    0     0     0 0 09
1    0  NA
23                95      2                    95   0    0     0     0 0 09
5    0  NA
24                95      2                    95   0    0     0     0 0 09
5    0  NA
25          11011216    112                    1101 6    0     0     0 6 01
10 1 6  NA
26                93      2                    93   0    0     0     0 0 09
3    0  NA
27                91      2                    91   0    0     0     0 0 09
1    0  NA
28                91      2                    91   0    0     0     0 0 09
1    0  NA
29                97      2                    97   0    0     0     0 0 09
7    0  NA
30          51410019    31118011 1281101121611 5141 8  400     0     0 8 05
14 1 8   4
31                93      2                    93   0    0     0     0 0 09
3    0  NA
32                91      2                    91   0    0     0     0 0 09
1    0  NA
33                91      2                    91   0    0     0     0 0 09
1    0  NA
34                91      2                    91   0    0     0     0 0 09
1    0  NA
35          11310017    12214011               1131 8    0     0     0 8 01
13 1 6  NA
```

```
36                  93       2                                93   0    0      0     0 0 09
3    0  NA
37                  91       2                                91   0    0      0     0 0 09
1    0  NA
38                  91       2                                91   0    0      0     0 0 09
1    0  NA
39                  91       2                                91   0    0      0     0 0 09
1    0  NA
40               11141057    43214011                         1114 8    0      0     0 8 01
11 4 8  NA
41                  93       2                                93   0    0      0     0 0 09
3    0  NA
42               31842205    16226054 417                     3184 8    0      0     0 8 03
18 4 8  NA
43                  93       2                                93   0    0      0     0 0 09
3    0  NA
44                  91       2                                91   0    0      0     0 0 09
1    0  NA
45               11011216    112                              1101 5    0      0     0 5 01
10 1 5  NA
46                  92       2                                92   0    0      0     0 0 09
2    0  NA
47               11854913    31210021                         1185 2    0      0     0 2 01
18 5 2  NA
48               31692013    44226011 128                     3169 6    0      0     0 6 03
16 9 6  NA
49               11854913    31210021                         1185 2    0      0     0 2 01
18 5 2  NA
50                  95       2                                95   0    0      0     0 0 09
5    0  NA
51               11011216    112                              1101 5    0      0     0 5 01
10 1 5  NA
52                  92       2                                92   0    0      0     0 0 09
2    0  NA
          MLT
1    0      0
2    0      0
3    0      0
4    0      0
5    0      0
6    0      0
7    0      0
8    0      0
9    0      0
10   0      0
11   0      0
12   0      0
13   0      0
14   0      0
15   0      0
16   0      0
17   0      0
18   0      0
19   0      0
20   0      0
21   0      0
22   0      0
23   0      0
24   0      0
25   0      0
26   0      0
27   0      0
28   0      0
29   0      0
30 00      0
31   0      0
32   0      0
33   0      0
34   0      0
35   0      0
36   0      0
37   0      0
38   0      0
39   0      0
40   0      0
41   0      0
42   0      0
43   0      0
44   0      0
45   0      0
46   0      0
47   0      0
```

```
48  0     0
49  0     0
50  0     0
51  0     0
52  0     0
 [ reached 'max' / getOption("max.print") -- omitted 65880 rows ]> View(level2)>
View(level1)> View(level2)> View(level1)> ## using the 'readr' package> level3<-r
ead_fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT", fwf_cols(fsuslno=c(4,
8), sector=c(15, 15), subblockno=c(32,32),sssno=c(33,33),hhno=c(34,35),level=c(36,
37),filler=c(38,40), personid=c(41,42), sex=c(43,43),ageatdeath=c(44,46),+
                                                           medicalatn=c(47,47),ho
spitalised=c(48,48),numhospitalised=c(49,50),pregnant=c(51,51),timeofdeath=c(52,5
2), nss=c(127,129), nsc=c(130,132), mlt=c(133, 142)),+                 col_types
 = cols(fsuslno=col_character(),sector=col_character(),subblockno=col_character(),
sssno=col_character(),hhno=col_character(),level = col_character(),filler = col_c
haracter(),personid=col_character(), sex=col_integer(),ageatdeath=col_integer(),+
                                   medicalatn=col_integer(), hospitalised=col_int
eger(), numhospitalised=col_integer(), pregnant=col_integer(), timeofdeath=col_in
teger(), nss=col_character(), nsc=col_character(),mlt=col_character()))Error in r
ead_fwf(file = "C:/Users/hp/Downloads/Lab1_7Aug/FHH_FV.TXT",   :
  could not find function "read_fwf"> install.packages("readr")WARNING: Rtools is
 required to build R packages but is not currently installed. Please download and
 install the appropriate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into ‘C:/Users/hp/AppData/Local/R/win-library/4.3’
(as ‘lib’ is unspecified)trying URL 'https://cran.rstudio.com/bin/windows/contrib
/4.3/readr_2.1.4.zip'Content type 'application/zip' length 1167480 bytes (1.1 MB)
downloaded 1.1 MB
package ‘readr’ successfully unpacked and MD5 sums checked

The downloaded binary packages are in
         C:\Users\hp\AppData\Local\Temp\Rtmpai5ZZ1\downloaded_packages> library(r
eadr)Warning message:
package ‘readr’ was built under R version 4.3.2 > level3<-read_fwf(file="C:/Users
/hp/Downloads/Lab1_7Aug/FHH_FV.TXT", fwf_cols(fsuslno=c(4,8), sector=c(15, 15), s
ubblockno=c(32,32),sssno=c(33,33),hhno=c(34,35),level=c(36,37),filler=c(38,40), p
ersonid=c(41,42), sex=c(43,43),ageatdeath=c(44,46),+
                                     medicalatn=c(47,47),hospitalised=c(48,48),nu
mhospitalised=c(49,50),pregnant=c(51,51),timeofdeath=c(52,52), nss=c(127,129), ns
c=c(130,132), mlt=c(133, 142)),+                  col_types = cols(fsuslno=col_ch
aracter(),sector=col_character(),subblockno=col_character(),sssno=col_character(),
hhno=col_character(),level = col_character(),filler = col_character(),personid=co
l_character(), sex=col_integer(),ageatdeath=col_integer(),+
             medicalatn=col_integer(), hospitalised=col_integer(), numhospitalise
d=col_integer(), pregnant=col_integer(), timeofdeath=col_integer(), nss=col_chara
cter(), nsc=col_character(),mlt=col_character()))> > level3# A tibble: 102,113 ×
18
   fsuslno sector subblockno sssno hhno  level filler personid   sex ageatdeath m
edicalatn
   <chr>   <chr>  <chr>      <chr> <chr> <chr> <chr>  <chr>    <int>      <int>
    <int> 1 1104Q   2       4          2     11     01     081     1          NA
    411                   2  2 1104Q    2       4          2     11     02     081     1
   NA         634              9  3 1104Q    2       4          2     12     01     081
1          NA         721              2  4 1104Q    2       4          2     12     02
  081     1          NA         521              2  5 1104Q    2       4          2     1
2     03     083     1          NA         614              9  6 1104Q    2       4
  2     12     04     084     1          NA         651              2  7 1104Q    2
4          2     13     01     083     1          NA         521              2  8 1104Q
  2       4          2     13     02     082     1          NA         531              2
 9 1104Q    2       7          1     11     01     082     1          NA         631
       910 1104Q    2       7          1     11     02     082    22          NA
    619             9# ℹ 102,103 more rows# ℹ 7 more variables: hospitalised <int>,
 numhospitalised <int>, pregnant <int>,#   timeofdeath <int>, nss <chr>, nsc <ch
r>, mlt <chr># ℹ Use `print(n = ...)` to see more rows> ## using the 'readr' pac
kage> level4<-read_fwf(file="C:/Users/hp/Downloads/Lab1_7Aug/FPER_FV.TXT",fwf_col
s(fsuslno=c(4,8), sector=c(15, 15), subblockno=c(32,32),sssno=c(33,33),hhno=c(34,
35),level=c(36,37),filler=c(38,40), personid=c(41,42),rltntohead=c(43,43),sex=c(4
4,44), age=c(45,47),+
     maritalstatus=c(48,48),education=c(49,50), resident_hostel=c(51,51),hospital
ised=c(52,52), numhospitalised=c(53,54),chronicailment=c(55,55),othrailment15days
=c(56,56),othrailmentdaybefore=c(57,57),+
                          healthscheme=c(58,58),reporting=c(59,59), nss=c(127,129),
 nsc=c(130,132),mlt=c(133,142)),+                  col_types = cols(fsuslno=col_c
haracter(),sector=col_character(),subblockno=col_character(),sssno=col_character
(),hhno=col_character(),level = col_character(),filler = col_character(),personid
=col_character(), rltntohead=col_integer(), sex=col_integer(),age=col_integer(),+
                                   maritalstatus=col_integer(),education=col_inte
ger(),resident_hostel=col_integer(),hospitalised=col_integer(),numhospitalised=co
l_integer(),chronicailment=col_integer(),othrailment15days=col_integer(),othrailm
entdaybefore=col_integer(),+                                     healthscheme=col_i
```

```
nteger(), reporting=col_integer(),nss=col_character(), nsc=col_character(), mlt=c
ol_character()))> > level4# A tibble: 433,339 × 24
   fsuslno sector subblockno sssno hhno  level filler personid rltntohead   sex
 age
   <chr>   <chr>  <chr>      <chr> <chr> <chr> <chr>  <chr>          <int> <int> <
int> 1 1104Q   2      4          2     11    01    011    16                 5
NA   208 2 1104Q   2      4          2     11    01    022    26                 2
    NA   206 3 1104Q   2      4          2     11    01    033    23
   1    NA   210 4 1104Q   2      4          2     11    01    044    13
       4    NA   210 5 1104Q   2      4          2     11    02    011    16
           5    NA   208 6 1104Q   2      4          2     11    02    022    26
               0    NA   206 7 1104Q   2      4          2     11    02    033
 13                8    NA   212 8 1104Q   2      4          2     11    02    04
4    23                2    NA   210 9 1104Q   2      4          2     11    02
  056    16                NA    NA   10510 1104Q   2      4          2     11
02    063    14                2    NA   407# ⓘ 433,329 more rows# ⓘ 13 more va
riables: maritalstatus <int>, education <int>, resident_hostel <int>,#   hospital
ised <int>, numhospitalised <int>, chronicailment <int>, othrailment15days <int>,
#   othrailmentdaybefore <int>, healthscheme <int>, reporting <int>, nss <chr>, n
sc <chr>,#   mlt <chr># ⓘ Use `print(n = ...)` to see more rows
```

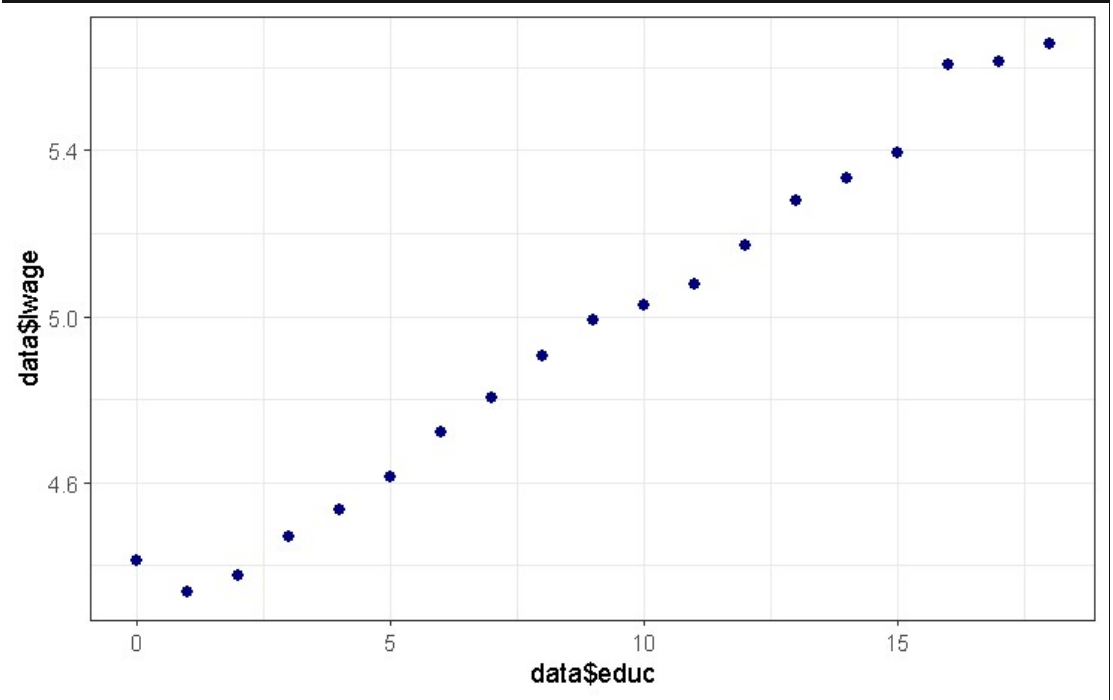
2.Instrumental Variables Analysis with Binscatter: An Empirical Study on Wage Determinants:

code:

```
#install.packages("colorspace")
library(haven) # Read .dta files
library(data.table) # For working with data
library(fixest) # For regressions
library(binsreg) # For binscatter

## Load data
# data <- haven::read_dta("https://github.com/Mixtape-Sessions/Instrumental-
Variables/blob/main/Exercises/Exercise1/angrist_krueger_91.dta?raw=true")
data <- read_dta("C:/Users/hp/Downloads/angrist_krueger (1).dta")
data <- as.data.table(data)
data
data[, qob_1 := (qob == 1)]
data[, qob_2 := (qob == 2)]
data[, qob_3 := (qob == 3)]
data[, qob_4 := (qob == 4)]

# ---- OLS and Binscatter ---------------------------------------------------
feols(
  lwage ~ educ, # Regression formula
  data,
  vcov = "hc1" # ,r
)
binscatter <- binsreg(data$lwage, data$educ)
binscatter
library(ggplot2)
# plot and add labels
binscatter$bins_plot +
  labs(y = "Log wages", x = "Years of Completed Schooling")
# ---- Simple (Wald) IV Estimator -------------------------------------------

# Formula y ~ exogenous | fixed effects | endogenous ~ instrument
# 1 = constant, 0 = no fixed effects
feols(
  lwage ~ 1 | 0 | educ ~ qob_1,
  data,
  vcov = "hc1"
)
data[,
```

```
   .(n = .N, mean = mean(lwage), sd = sd(lwage), min = min(lwage), max = max(lwage)),
   by = qob_1
]
data[,
   .(n = .N, mean = mean(educ), sd = sd(educ), min = min(educ), max = max(educ)),
   by = qob_1
]
# ---- Overidentified IV Estimator -------------------------------------------

feols(
 lwage ~ 1 | 0 | educ ~ qob_1 + qob_2 + qob_3,
 data,
 vcov = "hc1"
)

# collapse data by qob
collapsed <- data[,
           .(lwage = mean(lwage), educ = mean(educ)),
           by = qob
]
collapsed
# plot means
plot(collapsed$educ, collapsed$lwage)

# add regression line
abline(feols(lwage ~ educ, collapsed))
# ---- Putting the 2S in 2SLS -----------------------------------------------

feols(
 lwage ~ 1 | yob | educ ~ qob_1 + qob_2 + qob_3,
 data,
 vcov = "hc1"
)

first_stage <- feols(educ ~ i(qob) | yob, data)
first_stage

data[, educ_hat := predict(first_stage)]
feols(
 lwage ~ educ_hat | yob,
 data,
 vcov = "hc1"
)
# ---- Many IV Bias ----------------------------------------------------------

feols(
 lwage ~ 1 | yob | educ ~ i(yob, qob_1) + i(yob, qob_2) + i(yob, qob_3),
 data,
 vcov = "hc1"
)


output:
> library(haven) # Read .dta files> library(data.table) # For working with data>
library(fixest) # For regressions> library(binsreg) # For binscatter> ## Load dat
a> # data <- haven::read_dta("https://github.com/Mixtape-Sessions/Instrumental-Va
riables/blob/main/Exercises/Exercise1/angrist_krueger_91.dta?raw=true")> data <-
read_dta("C:/Users/hp/Downloads/angrist_krueger (1).dta")> data <- as.data.table
(data)> data
        educ    lwage qob  yob
     1:   11 5.023558   3 1929
     2:   12 5.061540   1 1929
     3:   12 5.378315   3 1928
     4:   12 5.178639   4 1923
     5:   16 6.378776   1 1924
    ---
247195:   15 5.167527   2 1920
247196:    5 3.357738   3 1920
```

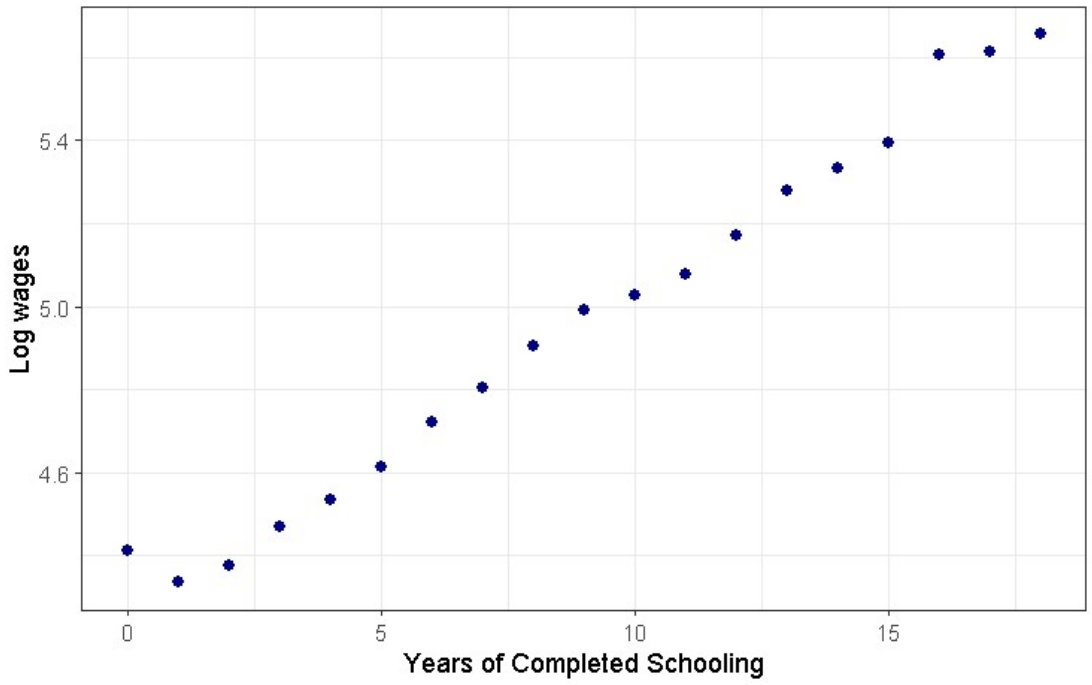
```
247197:     6 4.775905    2 1923
247198:    11 4.595219    3 1923
247199:    11 4.374523    4 1924> data[, qob_1 := (qob == 1)]> data[, qob_2 := (qob
 == 2)]> data[, qob_3 := (qob == 3)]> data[, qob_4 := (qob == 4)]> # ---- OLS and
 Binscatter ----------------------------------------------------> > feols(+   l
wage ~ educ, # Regression formula+   data,+   vcov = "hc1" # ,r+ )OLS estimation,
 Dep. Var.: lwage
Observations: 247,199
Standard-errors: Heteroskedasticity-robust
             Estimate Std. Error t value  Pr(>|t|)
(Intercept) 4.234430   0.004760 889.660 < 2.2e-16 ***
educ        0.080111   0.000394 203.323 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.593028   Adj. R2: 0.17088> > binscatter <- binsreg(data$lwage, data$educ)
Warning messages:
1: In binsreg(data$lwage, data$educ) :
  Too small effective sample size for bin selection. # of mass of points or clust
ers used and by option ignored.
2: In binsreg(data$lwage, data$educ) : dots=c(0,0) used.> binscatterCall: binsreg

Binscatter Plot
Bin/Degree selection method (binsmethod)  =  IMSE direct plug-in (select # of bin
s)
Placement (binspos)                       =  Quantile-spaced
Derivative (deriv)                        =  0

Group (by)                          =  Full Sample
Sample size (n)                     =  247199
# of distinct values (Ndist)        =  19
# of clusters (Nclust)              =  NA
dots, degree (p)                    =  0
dots, smoothness (s)                =  0
# of bins (nbins)                   =  19


Plot:
```



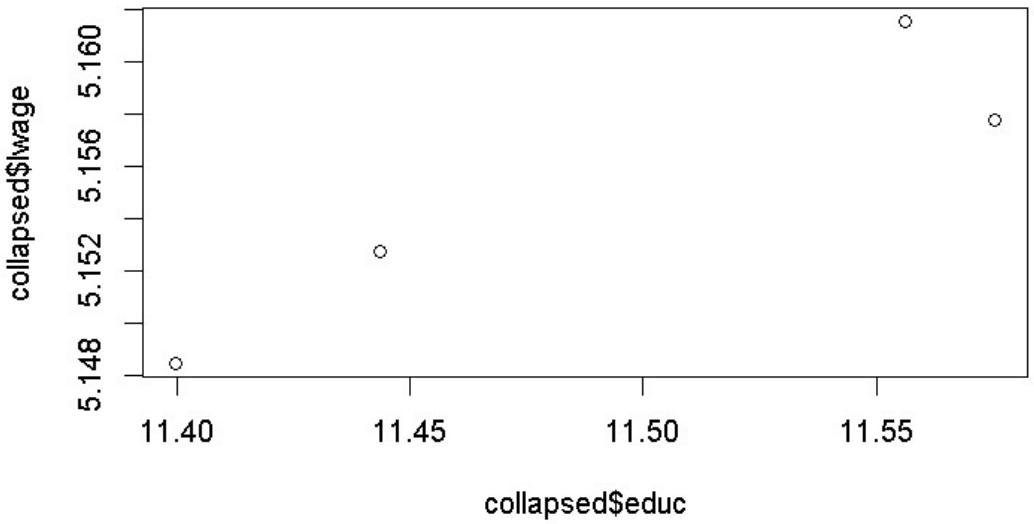

```
output:
> binscatterCall: binsreg

Binscatter Plot
Bin/Degree selection method (binsmethod)  =  IMSE direct plug-in (select # of bin
s)
Placement (binspos)                       =  Quantile-spaced
```

```
Derivative (deriv)                        =  0

Group (by)                          =  Full Sample
Sample size (n)                     =  247199
# of distinct values (Ndist)       =  19
# of clusters (Nclust)             =  NA
dots, degree (p)                    =  0
dots, smoothness (s)                =  0
# of bins (nbins)                  =  19
> library(ggplot2)> # plot and add labels> binscatter$bins_plot ++   labs(y = "Lo
g wages", x = "Years of Completed Schooling")> # ---- Simple (Wald) IV Estimator
----------------------------------------------> > # Formula y ~ exogenous | fixed
 effects | endogenous ~ instrument> # 1 = constant, 0 = no fixed effects> feols(+
   lwage ~ 1 | 0 | educ ~ qob_1,+   data,+   vcov = "hc1"+ )TSLS estimation, Dep.
 Var.: lwage, Endo.: educ, Instr.: qob_1
Second stage: Dep. Var.: lwage
Observations: 247,199
Standard-errors: Heteroskedasticity-robust
            Estimate Std. Error  t value  Pr(>|t|)
(Intercept) 4.333248   0.252253 17.17816 < 2.2e-16 ***
fit_educ    0.071513   0.021947  3.25848 0.0011203 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.593731   Adj. R2: 0.168912
F-test (1st stage), educ: stat = 65.3     , p = 6.508e-16, on 1 and 247,197 DoF.
               Wu-Hausman: stat =  0.154986, p = 0.693816 , on 1 and 247,196 DoF.>
 > > data[,+      .(n = .N, mean = mean(lwage), sd = sd(lwage), min = min(lwage),
 max = max(lwage)),+      by = qob_1+ ]   qob_1      n     mean        sd
 min      max
1: FALSE 184571 5.157450 0.6500542 -0.01980263 8.947976
2:  TRUE  62628 5.148471 0.6548401 -0.01980263 8.503235> data[,+      .(n = .N, m
ean = mean(educ), sd = sd(educ), min = min(educ), max = max(educ)),+      by = qo
b_1+ ]   qob_1      n     mean       sd min max
1: FALSE 184571 11.52515 3.350032   0  18
2:  TRUE  62628 11.39960 3.390094   0  18>  ---- Overidentified IV Estimator ----
-----------------------------------------Error: unexpected symbol in " ---- Overi
dentified IV"> # ---- Overidentified IV Estimator ------------------------------
--------------> > feols(+   lwage ~ 1 | 0 | educ ~ qob_1 + qob_2 + qob_3,+   data,
+   vcov = "hc1"+ )TSLS estimation, Dep. Var.: lwage, Endo.: educ, Instr.: qob_1,
 qob_2, qob_3
Second stage: Dep. Var.: lwage
Observations: 247,199
Standard-errors: Heteroskedasticity-robust
            Estimate Std. Error  t value   Pr(>|t|)
(Intercept) 4.432248   0.186406 23.77734  < 2.2e-16 ***
fit_educ    0.062900   0.016218  3.87837 0.00010519 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.595842   Adj. R2: 0.162992
F-test (1st stage), educ: stat = 40.0     , p < 2.2e-16 , on 3 and 247,195 DoF.
               Wu-Hausman: stat =  1.14254, p = 0.285117, on 1 and 247,196 DoF.
                   Sargan: stat =  2.35052, p = 0.308739, on 2 DoF.> # collapse da
ta by qob> collapsed <- data[,+                    .(lwage = mean(lwage), educ = m
ean(educ)),+                    by = qob+ ]> collapsed   qob    lwage     educ
1:   3 5.161561 11.55605
2:   1 5.148471 11.39960
3:   4 5.157816 11.57543
```

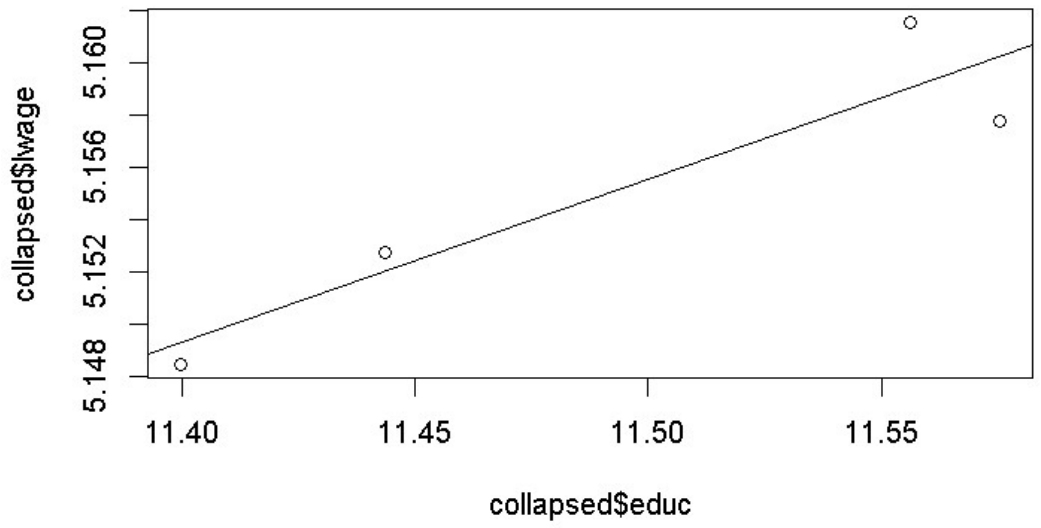
```
4:  2 5.152764 11.44342> # plot means> plot(collapsed$educ, collapsed$lwage)
Plot:
```




```
Output:
> # plot means> plot(collapsed$educ, collapsed$lwage)> # add regression line> abl
ine(feols(lwage ~ educ, collapsed))
Plot:
```




```
Plot: Addition of regression line:
```




```
Output:
> abline(feols(lwage ~ educ, collapsed))> # ---- Putting the 2S in 2SLS ---------
----------------------------------------> > feols(+   lwage ~ 1 | yob | educ ~ q
ob_1 + qob_2 + qob_3,+   data,+   vcov = "hc1"+ )TSLS estimation, Dep. Var.: lwag
e, Endo.: educ, Instr.: qob_1, qob_2, qob_3
Second stage: Dep. Var.: lwage
Observations: 247,199
Fixed-effects: yob: 10
Standard-errors: Heteroskedasticity-robust
         Estimate Std. Error t value   Pr(>|t|)
fit_educ 0.063351   0.016574 3.82223 0.00013229 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.595649     Adj. R2: 0.163504
                 Within R2: 0.163319
F-test (1st stage), educ: stat = 38.4     , p < 2.2e-16 , on 3 and 247,186 DoF.
             Wu-Hausman: stat =  1.04281, p = 0.307169, on 1 and 247,187 DoF.
                 Sargan: stat =  2.31607, p = 0.314102, on 2 DoF.> > first_stage
 <- feols(educ ~ i(qob) | yob, data)> first_stageOLS estimation, Dep. Var.: educ
Observations: 247,199
Fixed-effects: yob: 10
Standard-errors: Clustered (yob)
       Estimate Std. Error t value   Pr(>|t|)
qob::2 0.040859   0.015392 2.65459 2.6278e-02 *
qob::3 0.152733   0.021477 7.11147 5.5964e-05 ***
qob::4 0.171028   0.019760 8.65536 1.1738e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 3.35688     Adj. R2: 0.002197
                Within R2: 4.655e-4> data[, educ_hat := predict(first_stage)]> >
feols(+   lwage ~ educ_hat | yob,+   data,+   vcov = "hc1"+ )OLS estimation, Dep.
 Var.: lwage
Observations: 247,199
Fixed-effects: yob: 10
Standard-errors: Heteroskedasticity-robust
         Estimate Std. Error t value   Pr(>|t|)
educ_hat 0.063351   0.018107 3.49868 0.00046764 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.651177     Adj. R2: 2.717e-4
                 Within R2: 4.967e-5> # ---- Many IV Bias ----------------------
----------------------------------------> > feols(+   lwage ~ 1 | yob | educ ~ i(yob,
 qob_1) + i(yob, qob_2) + i(yob, qob_3),+   data,+   vcov = "hc1"+ )TSLS estimati
on, Dep. Var.: lwage, Endo.: educ, Instr.: i(yob, qob_1), i(yob, qob_2), i(yob, q
ob_3)
Second stage: Dep. Var.: lwage
Observations: 247,199
Fixed-effects: yob: 10
Standard-errors: Heteroskedasticity-robust
```

```
         Estimate Std. Error t value   Pr(>|t|)
fit_educ 0.076856   0.015123 5.08209 3.7358e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.593073     Adj. R2: 0.170724
                 Within R2: 0.17054
F-test (1st stage), educ: stat =  4.59855 , p = 8.844e-16, on 30 and 247,159 DoF.
             Wu-Hausman: stat =  0.048286, p = 0.826073 , on 1 and 247,187 DoF.
                 Sargan: stat = 36.0     , p = 0.172908 , on 29 DoF
```

.

3."Exploratory Data Analysis and Regression Analysis on mtcars Dataset" "Heteroskedasticity and Ridge Regression Analysis: An Empirical Study" "Fixed-Effects Estimation with the 'fixest' Package and PLFS 2017-18 Data Analysis"

Code:

```
mtcars
#dimension of the dataframe
dim(mtcars)
#structure of data
str(mtcars)

#install.packages("pastecs")
library(pastecs)
#summary statistics
stat.desc(mtcars)

#OLS
model <- lm(mpg ~ disp + hp + wt + drat, data = mtcars)
summary(model)

model2 <- lm(mpg ~  hp + wt , data = mtcars)
summary(model2)

####Heteroskedasticity######
###Create residual vs fitted plot
plot(fitted(model),resid(model),xlab='Fitted values',ylab = 'Residuals', abline(0,
0))

#install.packages("lmtest")
#install.packages("sandwich")
library(lmtest)
library(sandwich)

##Goldfeld Quandt test- change the number of central observations and see what h
appens
gqtest(model, order.by = ~disp+hp, data = mtcars, fraction = 7)

##Bruesch Pagan test
bptest(model)
##White test
bptest(model, ~ disp*hp + I(disp^2) + I(hp^2), data = mtcars)

###Perform weighted least squares/feasible GLS
#define weights
# estimating the variance of y for different values of x
wgt<-1/lm(abs(model$residuals)~model$fitted.values)$fitted.values^2
wls_model<-lm(mpg ~ disp + hp + wt + drat, data = mtcars, weights=wgt)

summary(wls_model)

####Heteroskedasticity robust standard errors
coeftest(model, vcov = vcovHC(model, type = "HC0"))



####Lab Date:28 August,2023####

####Multicollinearity#####
##Correlation matrix
install.packages("corrplot")
library(corrplot)
corrplot(cor(mtcars))
corrplot(cor(mtcars),method="color")
corrplot(cor(mtcars),method="number",type="upper")
```

```
#####Tolerances and variance inflation factor

library(olsrr)
ols_vif_tol(model)
ols_eigen_cindex(model)
##An conditional index value greater than 15 indicates presence of multicollinea
rity and
#greater than 30 indicates severe multicollinearity. Associated with conditional
index
#is output of variance decomposition for each principal component into intercept
#and regressors. For each component where conditional index exceeds 15,
#one should look for presence of variance concentration


####Ridge Regression####
install.packages("glmnet")
library(glmnet)
#Getting the independent variable
x_var<-data.matrix(mtcars[,c("hp", "wt", "drat")])

#Getting the dependent variable
y_var<-mtcars[, "mpg"]

#Setting the range of lambda values

lambda_seq<-10^seq(2,-2, by=-.1)
#Using glmnet function to build the ridge regression in r
fit<-glmnet(x_var, y_var, alpha=0, lambda = lambda_seq)
summary(fit)

#Next task is to identify the optimal value of lambda that will result in a mini
mum error
#This can be obtained by using cv.glmnet( ) function

ridge_cv<-cv.glmnet(x_var,y_var,alpha=0, lambda = lambda_seq)

best_lambda<-ridge_cv$lambda.min
best_lambda


###Building the final model with the best lambda
best_ridge<-glmnet(x_var,y_var,alpha=0,lambda = 0.5011872)
coef(best_ridge)

library(haven)
finaldata_ihds2 <- read_dta("C:/Users/hp/Downloads/finaldata_ihds2.dta")
View(finaldata_ihds2)
###Task 1: For each regression get the heteroskedasticity robust SE and compare
the standard errors of coefficients of 'science_eng' ####
m1<-lm(log_earnings~science_eng ,data=finaldata_ihds2)
summary(m1)
m2<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl ,da
ta=finaldata_ihds2)
summary(m2)
m3<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+ ie
du_level1+iedu_level4+iedu_level2+iedu_level3 ,data=finaldata_ihds2)
summary(m3)
finaldata_ihds2$District<-as.character(finaldata_ihds2$district)
m4<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+ ie
du_level1+iedu_level4+iedu_level2+iedu_level3+District ,data=finaldata_ihds2)
summary(m4)
m5<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+ ie
du_level1+iedu_level4+iedu_level2+age_lab +sq_age+ married +sc +st +obc +muslim+
christ +av_edu_min_i+District ,data=finaldata_ihds2)
summary(m5)

####Fixed-Effects Estimation in R with the fixest Package
#install.packages("fixest")
install.packages("AER")
library(fixest)
library(AER)

data(Grunfeld)
feols_model<- feols(invest ~ value + capital | firm + year , data = Grunfeld)
# one-way cluster by firm
feols_model<- feols(invest ~ value + capital | firm + year , data = Grunfeld, cl
uster = ~firm)

# two-way clustering by firm and year
feols_model<- feols(invest ~ value + capital | firm + year , data = Grunfeld, cl
uster = ~firm + year)
```

```
# estimate linear two-way fixed effect model with two-way clusting
feols_model<- feols(invest ~ value + capital | firm + year , data = Grunfeld, cl
uster = ~firm + year)

# get variance-covariance matrix with heteroskedasticity robust standard errors
hetero = vcov(feols_model, se = "hetero")

summary(feols_model)
# Alternatively, use etable:
etable(feols_model, tex = TRUE)

summary(feols_model, .vcov = hetero) # hetero is the var-cov matrix that was pre
viously computed using the vcov function
# OR
etable(feols_model, se = "white")




####Task2: Import the following .dta file to R

Class28_8 <- read_dta("C:/Users/hp/Downloads/Class28_8.dta")
View(Class28_8)

#This data is from PLFS 2017-18. The data includes employed males not working as
 casual labour.
#It includes males working in household enterprises/self employed or as salaried
 individuals.
#Q1: You want to examine whether social group-ie, caste identity affects earning
s.
#What should be your hypothesis?
#What should your dependent and main explanatory variables be?
#Q2: What variables should you control for, given your data?
#Q3: Do you think you can estimate causal effect from this exercise?If yes why?I
f not why?
#Q4: Interpret the results you get by running the regressions in Q1 and Q2.




#Now suppose you want to examine that self employment is not as rewarding as sal
aried employment.
#Q5: What will be your hypothesis?
#What should your dependent and main explanatory variables be?
#Q6: What variables should you control for, given your data?
#Q7: Do you think any variable can potentially cause hetroskedasticity?
#Q8: Can you say from your specification about existence of multicollinearity?Wh
at should you do?
#Q7: Do you think you can estimate causal effect from this exercise?If yes why?I
f not why?
#Q8: Interpret the results you get by running the OLS regressions in Q5 and Q6.



feols_model<- feols(GrossMonthlyEarnings ~ 0+ SC |
                      StateCode , data = Class28_8,
                    weights=Weight,cluster = ~Fsu)


Output:

> mtcars                    mpg cyl  disp  hp drat    wt  qsec vs am gear carb
Mazda RX4           21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag       21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
Datsun 710          22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive      21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout   18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
Valiant             18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
Duster 360          14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
Merc 240D           24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
Merc 230            22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
Merc 280            19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
Merc 280C           17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
Merc 450SE          16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
```

```
Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
Toyota Corona       21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
AMC Javelin         15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
Camaro Z28          13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
Pontiac Firebird    19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
Fiat X1-9           27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
Porsche 914-2       26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
Lotus Europa        30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
Ford Pantera L      15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
Ferrari Dino        19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
Maserati Bora       15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
Volvo 142E          21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2> #dimensi
on of the dataframe> dim(mtcars)[1] 32 11> #structure of data> str(mtcars)'data.f
rame':   32 obs. of  11 variables:
 $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
 $ cyl : num  6 6 4 6 8 6 8 4 4 6 ...
 $ disp: num  160 160 108 258 360 ...
 $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
 $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
 $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
 $ qsec: num  16.5 17 18.6 19.4 17 ...
 $ vs  : num  0 0 1 1 0 1 0 1 1 1 ...
 $ am  : num  1 1 1 0 0 0 0 0 0 0 ...
 $ gear: num  4 4 4 3 3 3 3 4 4 4 ...
 $ carb: num  4 4 1 1 2 1 4 2 2 4 ...> > #install.packages("pastecs")> library(pa
stecs)> #summary statistics> stat.desc(mtcars)                      mpg          cy
l         disp           hp         drat           wt
nbr.val       32.0000000  32.0000000 3.200000e+01   32.0000000  32.00000000  32.0
000000
nbr.null       0.0000000   0.0000000 0.000000e+00    0.0000000   0.00000000   0.0
000000
nbr.na         0.0000000   0.0000000 0.000000e+00    0.0000000   0.00000000   0.0
000000
min           10.4000000   4.0000000 7.110000e+01   52.0000000   2.76000000   1.5
130000
max           33.9000000   8.0000000 4.720000e+02  335.0000000   4.93000000   5.4
240000
range         23.5000000   4.0000000 4.009000e+02  283.0000000   2.17000000   3.9
110000
sum          642.9000000 198.0000000 7.383100e+03 4694.0000000 115.09000000 102.9
520000
median        19.2000000   6.0000000 1.963000e+02  123.0000000   3.69500000   3.3
250000
mean          20.0906250   6.1875000 2.307219e+02  146.6875000   3.59656250   3.2
172500
SE.mean        1.0654240   0.3157093 2.190947e+01   12.1203173   0.09451874   0.1
729685
CI.mean.0.95   2.1729465   0.6438934 4.468466e+01   24.7195501   0.19277224   0.3
527715
var           36.3241028   3.1895161 1.536080e+04 4700.8669355   0.28588135   0.9
573790
std.dev        6.0269481   1.7859216 1.239387e+02   68.5628685   0.53467874   0.9
784574
coef.var       0.2999881   0.2886338 5.371779e-01    0.4674077   0.14866382   0.3
041285
                   qsec          vs          am        gear        carb
nbr.val       32.0000000 32.00000000 32.00000000  32.0000000 32.0000000
nbr.null       0.0000000 18.00000000 19.00000000   0.0000000  0.0000000
nbr.na         0.0000000  0.00000000  0.00000000   0.0000000  0.0000000
min           14.5000000  0.00000000  0.00000000   3.0000000  1.0000000
max           22.9000000  1.00000000  1.00000000   5.0000000  8.0000000
range          8.4000000  1.00000000  1.00000000   2.0000000  7.0000000
sum          571.1600000 14.00000000 13.00000000 118.0000000 90.0000000
median        17.7100000  0.00000000  0.00000000   4.0000000  2.0000000
mean          17.8487500  0.43750000  0.40625000   3.6875000  2.8125000
SE.mean        0.3158899  0.08909831  0.08820997   0.1304266  0.2855297
CI.mean.0.95   0.6442617  0.18171719  0.17990541   0.2660067  0.5823417
var            3.1931661  0.25403226  0.24899194   0.5443548  2.6088710
std.dev        1.7869432  0.50401613  0.49899092   0.7378041  1.6152000
coef.var       0.1001159  1.15203687  1.22828533   0.2000825  0.5742933> #OLS> mo
del <- lm(mpg ~ disp + hp + wt + drat, data = mtcars)> summary(model)
Call:
lm(formula = mpg ~ disp + hp + wt + drat, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-3.5077 -1.9052 -0.5057  0.9821  5.6883

Coefficients:
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 29.148738   6.293588   4.631  8.2e-05 ***
disp         0.003815   0.010805   0.353  0.72675
hp          -0.034784   0.011597  -2.999  0.00576 **
wt          -3.479668   1.078371  -3.227  0.00327 **
drat         1.768049   1.319779   1.340  0.19153
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.602 on 27 degrees of freedom
Multiple R-squared:  0.8376,        Adjusted R-squared:  0.8136
F-statistic: 34.82 on 4 and 27 DF,  p-value: 2.704e-10
> > model2 <- lm(mpg ~  hp + wt , data = mtcars)> summary(model2)
Call:
lm(formula = mpg ~ hp + wt, data = mtcars)

Residuals:
   Min     1Q Median     3Q    Max
-3.941 -1.600 -0.182  1.050  5.854

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 37.22727    1.59879  23.285  < 2e-16 ***
hp          -0.03177    0.00903  -3.519  0.00145 **
wt          -3.87783    0.63273  -6.129 1.12e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.593 on 29 degrees of freedom
Multiple R-squared:  0.8268,        Adjusted R-squared:  0.8148
F-statistic: 69.21 on 2 and 29 DF,  p-value: 9.109e-12
> ####Heteroskedasticity######> ###Create residual vs fitted plot> plot(fitted(mo
del),resid(model),xlab='Fitted values',ylab = 'Residuals', abline(0,0))> #install.
packages("lmtest")> #install.packages("sandwich")> library(lmtest)> library(sandw
ich)> > ##Goldfeld Quandt test- change the number of central observations and see
 what happens> gqtest(model, order.by = ~disp+hp, data = mtcars, fraction = 7)
        Goldfeld-Quandt test

data:  model
GQ = 0.64357, df1 = 8, df2 = 7, p-value = 0.7256
alternative hypothesis: variance increases from segment 1 to 2
> #Bruesch Pagan test> bptest(model)
        studentized Breusch-Pagan test

data:  model
BP = 1.4406, df = 4, p-value = 0.8371
> ##White test> bptest(model, ~ disp*hp + I(disp^2) + I(hp^2), data = mtcars)
        studentized Breusch-Pagan test

data:  model
BP = 5.8536, df = 5, p-value = 0.3207
> > ###Perform weighted least squares/feasible GLS> #define weights> # estimating
 the variance of y for different values of x> wgt<-1/lm(abs(model$residuals)~mode
l$fitted.values)$fitted.values^2> wls_model<-lm(mpg ~ disp + hp + wt + drat, data
 = mtcars, weights=wgt)> summary(wls_model)
Call:
lm(formula = mpg ~ disp + hp + wt + drat, data = mtcars, weights = wgt)

Weighted Residuals:
    Min      1Q  Median      3Q     Max
-1.8038 -0.9371 -0.2982  0.5279  2.7285

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 26.810063   5.980278   4.483 0.000122 ***
disp         0.001268   0.010186   0.124 0.901853
hp          -0.032125   0.010281  -3.125 0.004222 **
wt          -2.920608   1.012028  -2.886 0.007586 **
drat         1.967974   1.295556   1.519 0.140382
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.339 on 27 degrees of freedom
Multiple R-squared:  0.8352,        Adjusted R-squared:  0.8107
F-statistic:  34.2 on 4 and 27 DF,  p-value: 3.308e-10
> ####Heteroskedasticity robust standard errors> coeftest(model, vcov = vcovHC(mo
del, type = "HC0"))
t test of coefficients:

              Estimate Std. Error t value  Pr(>|t|)
(Intercept) 29.1487376  5.5510404  5.2510 1.552e-05 ***
disp         0.0038152  0.0075726  0.5038 0.6184723
```

```
hp          -0.0347835  0.0093390 -3.7246 0.0009129 ***
wt          -3.4796675  1.0098460 -3.4457 0.0018775 **
drat         1.7680488  0.9606607  1.8405 0.0767186 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> ####Lab Date:28 August,2023####> > ####Multicollinearity#####> ##Correlation ma
trix> #install.packages("corrplot")> library(corrplot)> corrplot(cor(mtcars))> co
rrplot(cor(mtcars),method="color")> corrplot(cor(mtcars),method="number",type="up
per")> #####Tolerances and variance inflation factor> > library(olsrr)> ols_vif_t
ol(model)  Variables Tolerance      VIF
1      disp 0.1218116 8.209402
2        hp 0.3454979 2.894373
3        wt 0.1962092 5.096601
4      drat 0.4386836 2.279547> ols_eigen_cindex(model)   Eigenvalue Condition In
dex    intercept        disp          hp          wt        drat
1 4.692806914        1.000000 0.0002323252 0.001106455 0.002566185 0.0007172086 0.
0003775503
2 0.240308641        4.419078 0.0036813894 0.034132904 0.031334562 0.0009394254 0.
0148250672
3 0.052153430        9.485821 0.0009192095 0.058394262 0.735003722 0.0700789813 0.
0026259361
4 0.011406889       20.283026 0.0014476535 0.885725642 0.207337511 0.7179834661 0.
0568226912
5 0.003324127       37.573144 0.9937194224 0.020640737 0.023758021 0.2102809185 0.
9253487552> ####Ridge Regression####> install.packages("glmnet")Error in install.
packages : Updating loaded packages> ####Ridge Regression####> #install.packages
("glmnet")> library(glmnet)> #Getting the independent variable> x_var<-data.matri
x(mtcars[,c("hp", "wt", "drat")])> #Getting the independent variable> x_var<-data.
matrix(mtcars[,c("hp", "wt", "drat")])> x_var                     hp    wt drat
Mazda RX4            110 2.620 3.90
Mazda RX4 Wag        110 2.875 3.90
Datsun 710            93 2.320 3.85
Hornet 4 Drive       110 3.215 3.08
Hornet Sportabout    175 3.440 3.15
Valiant              105 3.460 2.76
Duster 360           245 3.570 3.21
Merc 240D             62 3.190 3.69
Merc 230              95 3.150 3.92
Merc 280             123 3.440 3.92
Merc 280C            123 3.440 3.92
Merc 450SE           180 4.070 3.07
Merc 450SL           180 3.730 3.07
Merc 450SLC          180 3.780 3.07
Cadillac Fleetwood   205 5.250 2.93
Lincoln Continental  215 5.424 3.00
Chrysler Imperial    230 5.345 3.23
Fiat 128              66 2.200 4.08
Honda Civic           52 1.615 4.93
Toyota Corolla        65 1.835 4.22
Toyota Corona         97 2.465 3.70
Dodge Challenger     150 3.520 2.76
AMC Javelin          150 3.435 3.15
Camaro Z28           245 3.840 3.73
Pontiac Firebird     175 3.845 3.08
Fiat X1-9             66 1.935 4.08
Porsche 914-2         91 2.140 4.43
Lotus Europa         113 1.513 3.77
Ford Pantera L       264 3.170 4.22
Ferrari Dino         175 2.770 3.62
Maserati Bora        335 3.570 3.54
Volvo 142E           109 2.780 4.11> #Getting the dependent variable> y_var<-mtcar
s[, "mpg"]> y_var [1] 21.0 21.0 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 17.8 16.4
 17.3 15.2 10.4 10.4 14.7 32.4
[19] 30.4 33.9 21.5 15.5 15.2 13.3 19.2 27.3 26.0 30.4 15.8 19.7 15.0 21.4> #Sett
ing the range of lambda values> > lambda_seq<-10^seq(2,-2, by=-.1)> #Using glmnet
 function to build the ridge regression in r> fit<-glmnet(x_var, y_var, alpha=0,
lambda = lambda_seq)> summary(fit)           Length Class     Mode
a0          41    -none-    numeric
beta       123    dgCMatrix S4
df          41    -none-    numeric
dim          2    -none-    numeric
lambda      41    -none-    numeric
dev.ratio   41    -none-    numeric
nulldev      1    -none-    numeric
npasses      1    -none-    numeric
jerr         1    -none-    numeric
offset       1    -none-    logical
call         5    -none-    call
nobs         1    -none-    numeric> #Next task is to identify the optimal value o
f lambda that will result in a minimum error> #This can be obtained by using cv.g
lmnet( ) function> > ridge_cv<-cv.glmnet(x_var,y_var,alpha=0, lambda = lambda_se
q)> > best_lambda<-ridge_cv$lambda.min> best_lambda[1] 0.6309573> ###Building the
```

```
 final model with the best lambda> best_ridge<-glmnet(x_var,y_var,alpha=0,lambda
= 0.5011872)> coef(best_ridge)4 x 1 sparse Matrix of class "dgCMatrix"
                    s0
(Intercept) 27.09643213
hp          -0.03164043
wt          -2.88955100
drat         1.92735603> ###Building the final model with the best lambda> best_r
idge<-glmnet(x_var,y_var,alpha=0,lambda = 0.5011872)> coef(best_ridge)4 x 1 spars
e Matrix of class "dgCMatrix"
                    s0
(Intercept) 27.09643213
hp          -0.03164043
wt          -2.88955100
drat         1.92735603> library(haven)> finaldata_ihds2 <- read_dta("C:/Users/hp
/Downloads/finaldata_ihds2.dta")> View(finaldata_ihds2)> ###Task 1: For each regr
ession get the heteroskedasticity robust SE and compare the standard errors of co
efficients of 'science_eng' ####> m1<-lm(log_earnings~science_eng ,data=finaldata
_ihds2)> summary(m1)
Call:
lm(formula = log_earnings ~ science_eng, data = finaldata_ihds2)

Residuals:
    Min      1Q  Median      3Q     Max
-5.9442 -0.5849  0.0795  0.6758  3.6811

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.65394    0.01634   284.8   <2e-16 ***
science_eng  0.37401    0.03282    11.4   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9781 on 4761 degrees of freedom
Multiple R-squared:  0.02655,       Adjusted R-squared:  0.02635
F-statistic: 129.9 on 1 and 4761 DF,  p-value: < 2.2e-16
> m2<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl ,
data=finaldata_ihds2)> summary(m2)
Call:
lm(formula = log_earnings ~ science_eng + first_div + second_div +
    repeated + eng_vfl + eng_lfl, data = finaldata_ihds2)

Residuals:
    Min      1Q  Median      3Q     Max
-5.6903 -0.5240  0.0483  0.6230  3.9697

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.29894    0.05056  85.023  < 2e-16 ***
science_eng  0.20413    0.03286   6.213 5.65e-10 ***
first_div    0.37267    0.04886   7.627 2.89e-14 ***
second_div   0.13971    0.04431   3.153 0.001627 **
repeated    -0.30828    0.04293  -7.180 8.03e-13 ***
eng_vfl      0.48285    0.04186  11.536  < 2e-16 ***
eng_lfl      0.13125    0.03903   3.363 0.000776 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9385 on 4756 degrees of freedom
Multiple R-squared:  0.1048,        Adjusted R-squared:  0.1036
F-statistic: 92.76 on 6 and 4756 DF,  p-value: < 2.2e-16
> m3<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+
iedu_level1+iedu_level4+iedu_level2+iedu_level3 ,data=finaldata_ihds2)> summary(m
3)
Call:
lm(formula = log_earnings ~ science_eng + first_div + second_div +
    repeated + eng_vfl + eng_lfl + iedu_level1 + iedu_level4 +
    iedu_level2 + iedu_level3, data = finaldata_ihds2)

Residuals:
    Min      1Q  Median      3Q     Max
-5.6676 -0.5188  0.0668  0.6172  3.8541

Coefficients: (1 not defined because of singularities)
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.45686    0.05678  78.494  < 2e-16 ***
science_eng  0.18661    0.03276   5.696 1.30e-08 ***
first_div    0.30553    0.04939   6.186 6.67e-10 ***
second_div   0.11490    0.04421   2.599  0.00938 **
repeated    -0.28053    0.04301  -6.522 7.66e-11 ***
eng_vfl      0.37310    0.04407   8.467  < 2e-16 ***
eng_lfl      0.09285    0.03919   2.369  0.01787 *
iedu_level1 -0.27371    0.05324  -5.141 2.84e-07 ***
```

```
iedu_level4  0.09024    0.03865   2.335  0.01958 *
iedu_level2 -0.16784    0.03355  -5.002 5.87e-07 ***
iedu_level3       NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9332 on 4753 degrees of freedom
Multiple R-squared:  0.1155,        Adjusted R-squared:  0.1138
F-statistic: 68.94 on 9 and 4753 DF,  p-value: < 2.2e-16
> finaldata_ihds2$District<-as.character(finaldata_ihds2$district)> m4<-lm(log_ea
rnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+ iedu_level1+ied
u_level4+iedu_level2+iedu_level3+District ,data=finaldata_ihds2)> summary(m4)
Call:
lm(formula = log_earnings ~ science_eng + first_div + second_div +
    repeated + eng_vfl + eng_lfl + iedu_level1 + iedu_level4 +
    iedu_level2 + iedu_level3 + District, data = finaldata_ihds2)

Residuals:
    Min      1Q  Median      3Q     Max
-5.4616 -0.4619  0.0383  0.5342  3.4829

Coefficients: (1 not defined because of singularities)
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.02746    0.24778  16.254  < 2e-16 ***
science_eng   0.24242    0.03238   7.486 8.51e-14 ***
first_div     0.19695    0.04932   3.994 6.61e-05 ***
second_div    0.04180    0.04409   0.948 0.343099
repeated     -0.22348    0.04271  -5.232 1.75e-07 ***
eng_vfl       0.31605    0.04663   6.778 1.37e-11 ***
eng_lfl       0.11030    0.04062   2.716 0.006642 **
iedu_level1  -0.24386    0.05234  -4.659 3.27e-06 ***
iedu_level4   0.17188    0.03828   4.490 7.29e-06 ***
iedu_level2  -0.16350    0.03260  -5.016 5.49e-07 ***
iedu_level3        NA         NA      NA       NA
District1017 -0.10732    0.33149  -0.324 0.746135
District102   1.41960    0.38015   3.734 0.000191 ***
District1022  0.36675    0.35813   1.024 0.305856
District1026  0.51532    0.49973   1.031 0.302508
District1027 -0.02549    0.38005  -0.067 0.946527
District1028  0.01038    0.30878   0.034 0.973198
District103   0.38365    0.27791   1.380 0.167513
District1030 -0.06603    0.32242  -0.205 0.837749
District1032  0.06071    0.27828   0.218 0.827325
District1035 -0.06518    0.36819  -0.177 0.859503
District1100  0.60482    0.33776   1.791 0.073417 .
District113   0.62336    0.26505   2.352 0.018723 *
District1200  1.78263    0.31186   5.716 1.16e-08 ***
District1300  1.32078    0.35929   3.676 0.000240 ***
District1400  0.50830    0.29233   1.739 0.082143 .
District1500  0.93679    0.43152   2.171 0.029989 *
District1600  0.63963    0.41009   1.560 0.118893
District1700  0.45413    0.32294   1.406 0.159716
District1803  0.77120    0.32716   2.357 0.018454 *
District1806  0.58135    0.28791   2.019 0.043525 *
District1809  0.74770    0.49988   1.496 0.134788
District1814  0.93989    0.34310   2.739 0.006179 **
District1817  0.85185    0.30866   2.760 0.005807 **
District1819  1.43904    0.34287   4.197 2.76e-05 ***
District1821  0.72683    0.32297   2.250 0.024467 *
District1902 -0.21785    0.43161  -0.505 0.613760
District1904  0.95168    0.35803   2.658 0.007887 **
District1906  0.31118    0.43134   0.721 0.470686
District1909  0.36301    0.26470   1.371 0.170330
District1910  1.17237    0.31856   3.680 0.000236 ***
District1911  0.56344    0.26671   2.113 0.034692 *
District1912  0.44027    0.29400   1.498 0.134323
District1913  0.39653    0.30347   1.307 0.191395
District1916  0.38023    0.27751   1.370 0.170703
District1917  0.54960    0.26213   2.097 0.036081 *
District1918  0.32208    0.27653   1.165 0.244198
District2002 -0.66055    0.31823  -2.076 0.037977 *
District2012  0.28594    0.31158   0.918 0.358813
District2013  0.27642    0.31148   0.887 0.374880
District2014 -0.14665    0.29024  -0.505 0.613398
District2018  0.77412    0.29805   2.597 0.009428 **
District202   0.71249    0.29232   2.437 0.014833 *
District204   0.49829    0.33691   1.479 0.139211
District205   0.47584    0.32210   1.477 0.139662
District209   0.70726    0.30589   2.312 0.020817 *
District210   0.76469    0.33154   2.306 0.021128 *
District2102  0.82165    0.37957   2.165 0.030463 *
District2103  1.51636    0.90694   1.672 0.094602 .
```

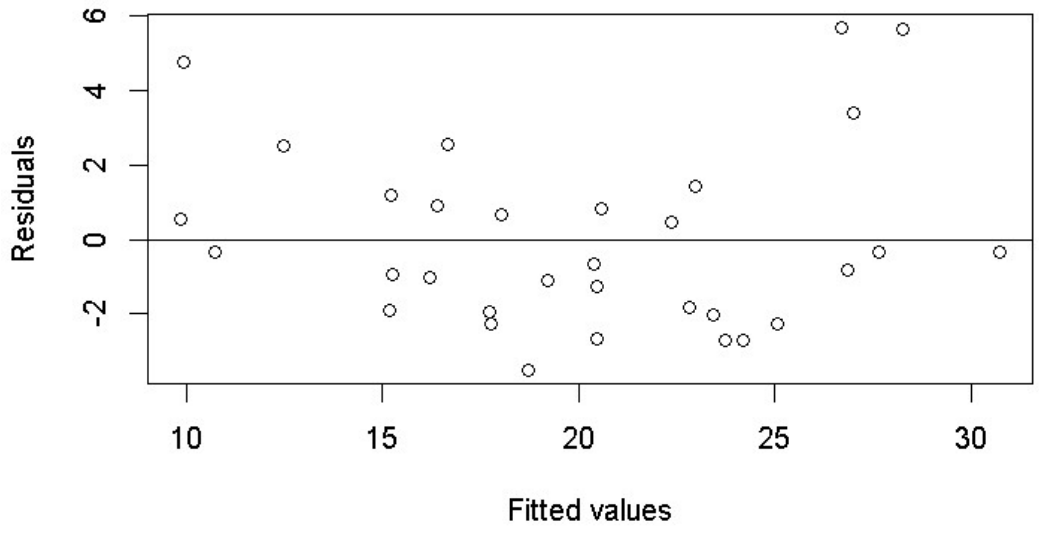
```
District2105  0.30703    0.37964   0.809 0.418707
District2108  0.22494    0.28153   0.799 0.424330
District211   0.50734    0.30326   1.673 0.094401 .
District2112  0.59616    0.28404   2.099 0.035886 *
District2113  0.09688    0.29999   0.323 0.746752
District2115  0.40891    0.37987   1.076 0.281781
District2116 -0.19295    0.33772  -0.571 0.567800
District2117  0.78440    0.29401   2.668 0.007660 **
District2120 -0.68758    0.50080  -1.373 0.169831
District2126 -0.07983    0.39331  -0.203 0.839170
District2127  0.53774    0.31903   1.686 0.091950 .
District2201  0.54151    0.43247   1.252 0.210579
District2202  1.01623    0.26938   3.773 0.000164 ***
District2204  0.62742    0.46019   1.363 0.172825
District2210  0.29314    0.28644   1.023 0.306176
District2211  0.53625    0.28787   1.863 0.062555 .
District2302  0.13677    0.32718   0.418 0.675947
District2304  0.17819    0.37919   0.470 0.638441
District2316  0.42062    0.33156   1.269 0.204641
District2321 -0.13664    0.36820  -0.371 0.710584
District2322 -0.17613    0.29201  -0.603 0.546444
District2325  0.18841    0.32659   0.577 0.564026
District2326  0.40826    0.29198   1.398 0.162105
District2327  0.07270    0.35840   0.203 0.839270
District2329  0.15832    0.90656   0.175 0.861369
District2330 -0.21695    0.35849  -0.605 0.545089
District2332  0.45003    0.36758   1.224 0.220908
District2338  0.34899    0.31173   1.120 0.262974
District2339  0.25359    0.29941   0.847 0.397055
District2406  0.52040    0.25992   2.002 0.045324 *
District2408  0.30669    0.90737   0.338 0.735376
District2409  0.28961    0.43191   0.671 0.502549
District2410  1.36027    0.39333   3.458 0.000549 ***
District2412  0.09418    0.36755   0.256 0.797788
District2413  0.13977    0.35841   0.390 0.696586
District2414  0.64752    0.56006   1.156 0.247675
District2415  0.18725    0.28404   0.659 0.509785
District2419  0.42617    0.28391   1.501 0.133401
District2421  0.26548    0.35826   0.741 0.458723
District2422  0.69436    0.27542   2.521 0.011734 *
District2600  0.30413    0.35046   0.868 0.385548
District2702  0.36203    0.29897   1.211 0.225979
District2707 -0.97259    0.34305  -2.835 0.004601 **
District2708  0.45460    0.33656   1.351 0.176851
District2709  0.17779    0.40969   0.434 0.664333
District2711 -0.52687    0.49997  -1.054 0.292032
District2713  0.77018    0.66405   1.160 0.246180
District2715 -0.04905    0.30199  -0.162 0.870985
District2718  0.50272    0.32211   1.561 0.118654
District2721  0.81486    0.25779   3.161 0.001583 **
District2722  0.92473    0.24988   3.701 0.000218 ***
District2725  0.71253    0.30876   2.308 0.021060 *
District2729  0.75066    0.30592   2.454 0.014175 *
District2730  0.58350    0.36769   1.587 0.112595
District2731  0.31226    0.33116   0.943 0.345764
District2734  0.33586    0.30135   1.114 0.265125
District2802  0.22252    0.32230   0.690 0.489966
District2803 -0.34430    0.33675  -1.022 0.306632
District2805  0.24850    0.27660   0.898 0.369008
District2806  0.26019    0.30673   0.848 0.396328
District2807 -0.16461    0.35828  -0.459 0.645941
District2813  0.71741    0.46064   1.557 0.119443
District2814  0.08097    0.43185   0.187 0.851284
District2815 -0.18291    0.40985  -0.446 0.655420
District2816  0.15407    0.40995   0.376 0.707054
District2817 -0.36249    0.46032  -0.787 0.431047
District2819  0.70248    0.43175   1.627 0.103799
District2820  0.16881    0.39311   0.429 0.667633
District2821 -0.19490    0.32280  -0.604 0.546006
District2823  0.05101    0.34313   0.149 0.881822
District2901  0.20516    0.31506   0.651 0.514963
District2902 -1.80874    0.40983  -4.413 1.04e-05 ***
District2903  0.51970    0.35832   1.450 0.147021
District2907 -1.00709    0.39324  -2.561 0.010468 *
District2908  0.63976    0.36812   1.738 0.082300 .
District2909  0.32686    0.40988   0.797 0.425224
District2910  0.37966    0.55960   0.678 0.497529
District2912  0.23430    0.31141   0.752 0.451867
District2915  0.04515    0.35010   0.129 0.897393
District2917 -0.16016    0.35872  -0.446 0.655265
District2918  0.06292    0.33148   0.190 0.849472
District2920  0.69871    0.25627   2.726 0.006426 **
```

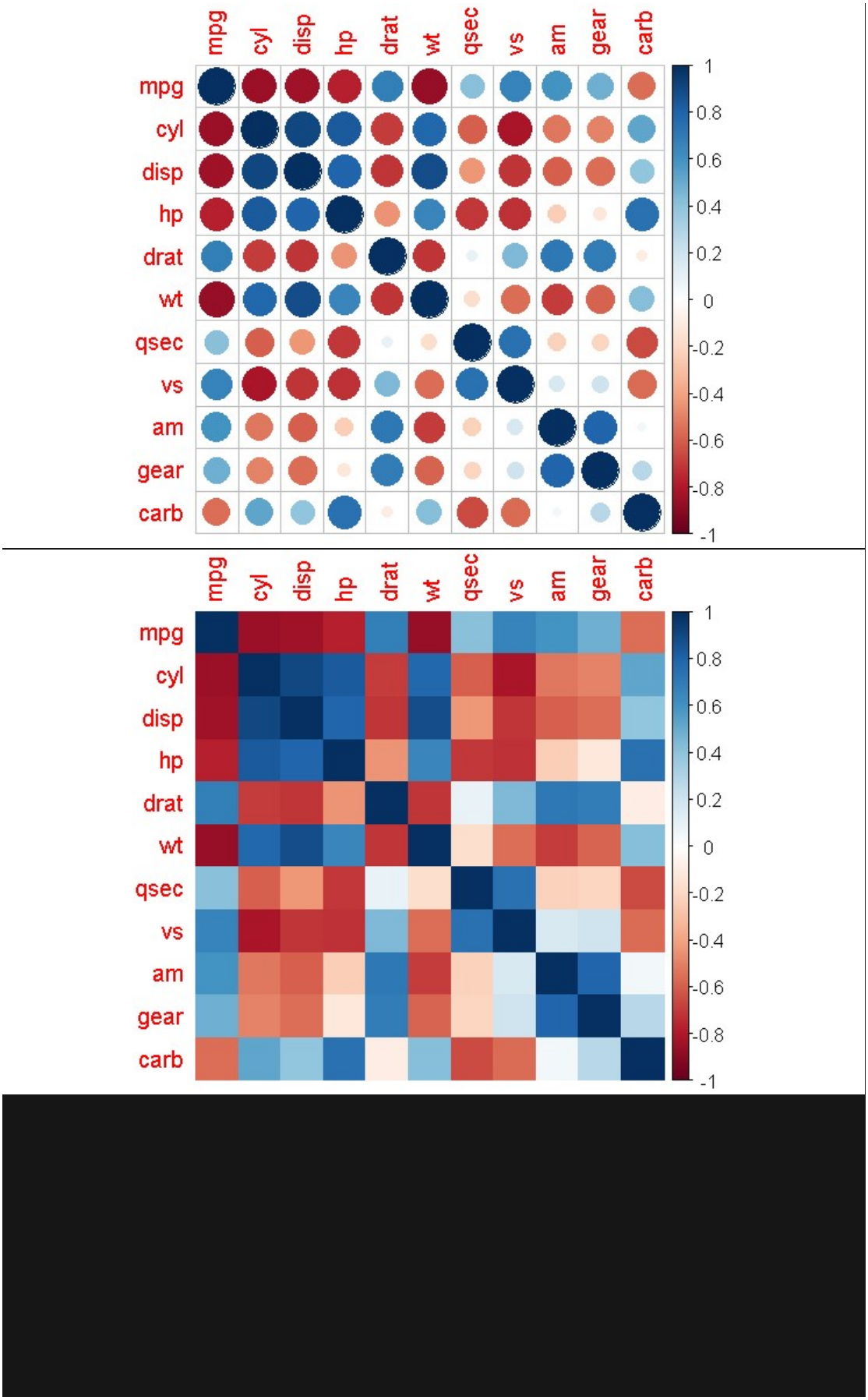
```
District2922  0.24446    0.33154   0.737 0.460951
District2924  0.60904    0.29349   2.075 0.038027 *
District2926 -0.16556    0.29753  -0.556 0.577940
District3001  0.76899    0.29909   2.571 0.010168 *
District301   0.43849    0.28682   1.529 0.126380
District302   0.31163    0.33733   0.924 0.355632
District303   0.25561    0.30592   0.836 0.403461
District304   0.71307    0.27581   2.585 0.009759 **
District307   0.48650    0.43157   1.127 0.259686
District309   0.54526    0.27906   1.954 0.050767 .
District310   0.18313    0.33110   0.553 0.580231
District311   0.34147    0.30558   1.117 0.263854
District317   0.39205    0.30582   1.282 0.199920
District3202  0.31654    0.32228   0.982 0.326059
District3204  0.66207    0.29919   2.213 0.026957 *
District3205  0.54677    0.28907   1.891 0.058622 .
District3206  0.33874    0.56010   0.605 0.545357
District3207  0.67057    0.33736   1.988 0.046908 *
District3208  0.27787    0.28544   0.973 0.330380
District3210  0.73544    0.66386   1.108 0.267996
District3211  0.31196    0.28118   1.109 0.267296
District3213  0.51808    0.43208   1.199 0.230575
District3214  0.66439    0.36839   1.803 0.071378 .
District3302  0.67793    0.26747   2.535 0.011291 *
District3303 -0.04816    0.41005  -0.117 0.906513
District3304 -0.13859    0.32262  -0.430 0.667522
District3305  0.79827    0.90691   0.880 0.378792
District3306  0.27915    0.46013   0.607 0.544103
District3309  0.09604    0.35839   0.268 0.788721
District3312  0.24666    0.30932   0.797 0.425251
District3313  0.28085    0.43163   0.651 0.515300
District3314  0.22297    0.34337   0.649 0.516148
District3317  0.42758    0.46145   0.927 0.354190
District3323  0.93127    0.37991   2.451 0.014271 *
District3324  0.37339    0.35882   1.041 0.298110
District3325  0.19646    0.56004   0.351 0.725760
District3327  0.01572    0.39292   0.040 0.968082
District3328  0.67989    0.43199   1.574 0.115590
District3329 -0.09282    0.49959  -0.186 0.852620
District3330  0.13282    0.39290   0.338 0.735349
District3400  0.65291    0.30650   2.130 0.033209 *
District400   0.95761    0.28041   3.415 0.000643 ***
District505   0.43138    0.29408   1.467 0.142480
District511   0.56734    0.31868   1.780 0.075098 .
District512   0.16201    0.46013   0.352 0.724778
District513   1.87912    0.55948   3.359 0.000790 ***
District602   0.81796    0.28766   2.843 0.004482 **
District603   1.86211    0.90694   2.053 0.040112 *
District608   0.70470    0.36790   1.915 0.055497 .
District609   0.74983    0.32253   2.325 0.020125 *
District612   0.99511    0.30895   3.221 0.001287 **
District617  -0.44974    0.39280  -1.145 0.252287
District618   1.01781    0.29537   3.446 0.000574 ***
District619   0.25358    0.29944   0.847 0.397127
District700   0.86754    0.24734   3.507 0.000457 ***
District803   0.57764    0.32293   1.789 0.073728 .
District806   0.19739    0.31883   0.619 0.535876
District812   0.86662    0.27361   3.167 0.001548 **
District813   0.52551    0.31478   1.669 0.095093 .
 [ reached getOption("max.print") -- omitted 31 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8732 on 4533 degrees of freedom
Multiple R-squared:  0.2614,       Adjusted R-squared:  0.224
F-statistic: 7.004 on 229 and 4533 DF,  p-value: < 2.2e-16
> m5<-lm(log_earnings~science_eng+first_div+second_div+repeated+eng_vfl+eng_lfl+
iedu_level1+iedu_level4+iedu_level2+age_lab +sq_age+ married +sc +st +obc +muslim
+ christ +av_edu_min_i+District ,data=finaldata_ihds2)> summary(m5)
Call:
lm(formula = log_earnings ~ science_eng + first_div + second_div +
    repeated + eng_vfl + eng_lfl + iedu_level1 + iedu_level4 +
    iedu_level2 + age_lab + sq_age + married + sc + st + obc +
    muslim + christ + av_edu_min_i + District, data = finaldata_ihds2)

Residuals:
    Min      1Q  Median      3Q     Max
-5.5068 -0.4319  0.0561  0.5083  3.1125

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.0116509  0.3245258   6.199 6.21e-10 ***
```

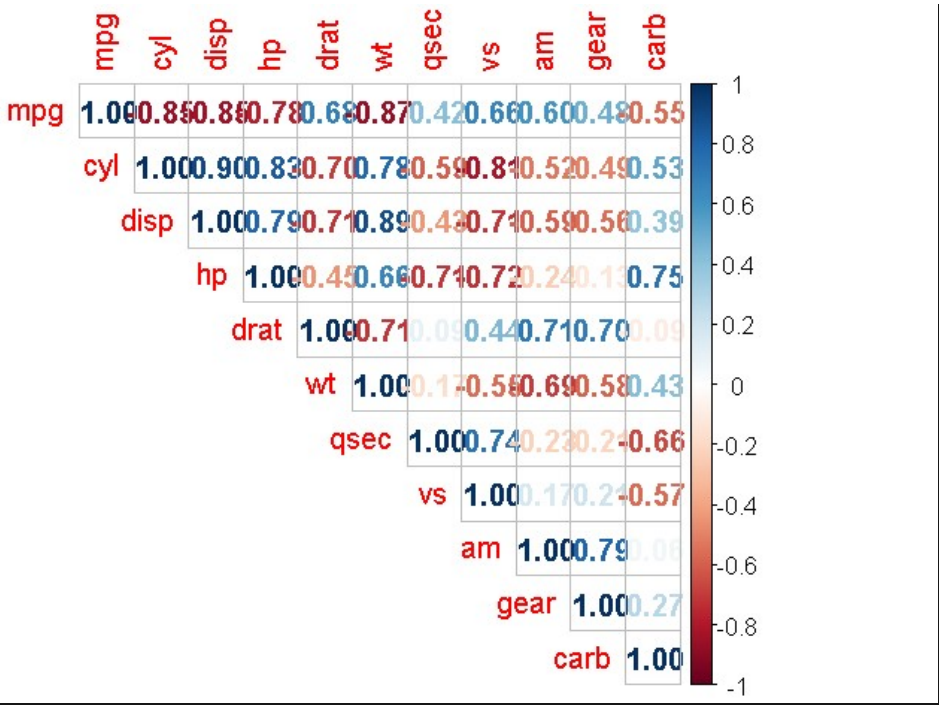
```
science_eng    0.2130854  0.0316660   6.729 1.92e-11 ***
first_div      0.1997498  0.0482394   4.141 3.53e-05 ***
second_div     0.0432737  0.0430666   1.005 0.315044
repeated      -0.1832815  0.0419026  -4.374 1.25e-05 ***
eng_vfl        0.2715172  0.0456326   5.950 2.89e-09 ***
eng_lfl        0.0975558  0.0396445   2.461 0.013902 *
iedu_level1   -0.1839891  0.0520069  -3.538 0.000408 ***
iedu_level4    0.1653622  0.0374017   4.421 1.00e-05 ***
iedu_level2   -0.1151947  0.0323195  -3.564 0.000369 ***
age_lab        0.0670939  0.0112546   5.961 2.69e-09 ***
sq_age        -0.0005984  0.0001308  -4.574 4.92e-06 ***
married        0.0871741  0.0399317   2.183 0.029082 *
sc            -0.0424111  0.0444613  -0.954 0.340193
st             0.0704631  0.0904790   0.779 0.436151
obc           -0.0387882  0.0326375  -1.188 0.234718
muslim        -0.0246251  0.0547552  -0.450 0.652927
christ         0.0863640  0.0896526   0.963 0.335440
av_edu_min_i   0.0255506  0.0038739   6.596 4.73e-11 ***
District1017   0.0349911  0.3204989   0.109 0.913067
District102    1.6777158  0.3719988   4.510 6.65e-06 ***
District1022   0.3103167  0.3471639   0.894 0.371444
District1026   0.5374901  0.4831439   1.112 0.265990
District1027  -0.0860261  0.3806459  -0.226 0.821212
District1028   0.0525924  0.2988184   0.176 0.860301
District103    0.5524804  0.2743641   2.014 0.044104 *
District1030  -0.0317452  0.3118539  -0.102 0.918924
District1032   0.1642534  0.2700520   0.608 0.543067
District1035   0.0427087  0.3561041   0.120 0.904541
District1100   0.6699525  0.3269686   2.049 0.040522 *
District113    0.6673059  0.2578070   2.588 0.009674 **
District1200   1.5567140  0.3096145   5.028 5.15e-07 ***
District1300   1.0623591  0.3670758   2.894 0.003821 **
District1400   0.4295826  0.2838001   1.514 0.130178
District1500   0.7525240  0.4339762   1.734 0.082984 .
District1600   0.6132625  0.3970963   1.544 0.122571
District1700   0.3798678  0.3247931   1.170 0.242237
District1803   0.7819690  0.3175090   2.463 0.013822 *
District1806   0.6470874  0.2791456   2.318 0.020489 *
District1809   1.1124798  0.4838701   2.299 0.021544 *
District1814   1.0096152  0.3324545   3.037 0.002404 **
District1817   0.7556999  0.2991863   2.526 0.011576 *
District1819   1.5083475  0.3387112   4.453 8.67e-06 ***
District1821   0.7491415  0.3129726   2.394 0.016724 *
District1902  -0.1095190  0.4192209  -0.261 0.793916
District1904   0.7967166  0.3471095   2.295 0.021763 *
District1906   0.4907765  0.4176434   1.175 0.240014
District1909   0.4411959  0.2572705   1.715 0.086431 .
District1910   1.1039143  0.3087776   3.575 0.000354 ***
District1911   0.5166856  0.2589290   1.995 0.046052 *
District1912   0.5256471  0.2851455   1.843 0.065332 .
District1913   0.4114402  0.2938797   1.400 0.161574
District1916   0.3834987  0.2697941   1.421 0.155256
District1917   0.5808366  0.2552386   2.276 0.022914 *
District1918   0.3101758  0.2685152   1.155 0.248090
District2002  -0.6408237  0.3085758  -2.077 0.037885 *
District2012   0.3556016  0.3014405   1.180 0.238193
District2013   0.2896735  0.3044074   0.952 0.341353
District2014  -0.0750616  0.2850167  -0.263 0.792286
District2018   0.7149159  0.2892455   2.472 0.013486 *
District202    0.8038299  0.2840355   2.830 0.004675 **
District204    0.5686451  0.3269327   1.739 0.082045 .
District205    0.5763646  0.3125935   1.844 0.065277 .
District209    0.6665498  0.2970007   2.244 0.024864 *
District210    0.6600942  0.3269969   2.019 0.043583 *
District2102   0.8949695  0.3673964   2.436 0.014891 *
District2103   1.5504713  0.8766984   1.769 0.077040 .
District2105   0.3608089  0.3684426   0.979 0.327494
District2108   0.2656260  0.2732994   0.972 0.331142
District211    0.5058276  0.2943546   1.718 0.085788 .
District2112   0.6217995  0.2751369   2.260 0.023872 *
District2113   0.2246408  0.2902013   0.774 0.438921
District2115   0.5609266  0.3683191   1.523 0.127846
District2116  -0.2491610  0.3327161  -0.749 0.453976
District2117   0.7459238  0.2853469   2.614 0.008977 **
District2120  -0.7761670  0.4846176  -1.602 0.109314
District2126  -0.0974523  0.4182042  -0.233 0.815752
District2127   0.5389627  0.3128422   1.723 0.084995 .
District2201   0.5677843  0.4183181   1.357 0.174754
District2202   0.9528274  0.2615185   3.643 0.000272 ***
District2204   0.6135873  0.4452229   1.378 0.168224
District2210   0.4094375  0.2803806   1.460 0.144281
District2211   0.5673793  0.2787447   2.035 0.041862 *
```

```
District2302  0.1593138  0.3177925   0.501 0.616175
District2304  0.1373767  0.3668960   0.374 0.708103
District2316  0.5942334  0.3208804   1.852 0.064109 .
District2321 -0.1719223  0.3559455  -0.483 0.629118
District2322 -0.1545374  0.2846816  -0.543 0.587265
District2325  0.2382064  0.3167535   0.752 0.452076
District2326  0.4213428  0.2846173   1.480 0.138842
District2327  0.2843092  0.3570208   0.796 0.425878
District2329  0.3204192  0.8762697   0.366 0.714634
District2330  0.0333083  0.3562401   0.093 0.925511
District2332  0.4703148  0.3562943   1.320 0.186897
District2338  0.2150214  0.3019557   0.712 0.476443
District2339  0.2666524  0.2946217   0.905 0.365479
District2406  0.5480960  0.2520855   2.174 0.029739 *
District2408  0.4815746  0.8773909   0.549 0.583121
District2409  0.1971051  0.4180786   0.471 0.637339
District2410  1.2654740  0.3813071   3.319 0.000911 ***
District2412  0.1883522  0.3561744   0.529 0.596957
District2413  0.2043369  0.3471251   0.589 0.556123
District2414  0.9976232  0.5417196   1.842 0.065602 .
District2415  0.2114568  0.2753760   0.768 0.442597
District2419  0.3918923  0.2765711   1.417 0.156562
District2421  0.2915185  0.3467367   0.841 0.400534
District2422  0.6542478  0.2682155   2.439 0.014756 *
District2600  0.3990167  0.3482987   1.146 0.252016
District2702  0.3017760  0.2919206   1.034 0.301304
District2707 -0.9247120  0.3318321  -2.787 0.005348 **
District2708  0.4754105  0.3259609   1.458 0.144776
District2709  0.2098187  0.3988213   0.526 0.598847
District2711 -0.3429470  0.4835861  -0.709 0.478253
District2713  1.0712973  0.6419924   1.669 0.095246 .
District2715  0.0470982  0.2929086   0.161 0.872262
District2718  0.6138172  0.3127555   1.963 0.049754 *
District2721  0.8417883  0.2501299   3.365 0.000771 ***
District2722  0.9649667  0.2424760   3.980 7.01e-05 ***
District2725  0.8242362  0.2992930   2.754 0.005912 **
District2729  0.7605974  0.2990531   2.543 0.011013 *
District2730  0.6869482  0.3800779   1.807 0.070769 .
District2731  0.4997824  0.3218925   1.553 0.120581
District2734  0.3170805  0.2969271   1.068 0.285636
District2802  0.3636498  0.3161112   1.150 0.250047
District2803 -0.1900615  0.3259126  -0.583 0.559810
District2805  0.4177335  0.2681780   1.558 0.119382
District2806  0.3587985  0.2972505   1.207 0.227474
District2807 -0.0504283  0.3469174  -0.145 0.884432
District2813  0.9563885  0.4457358   2.146 0.031956 *
District2814 -0.0087911  0.4460274  -0.020 0.984276
District2815 -0.0503185  0.3966338  -0.127 0.899054
District2816  0.1466679  0.3966998   0.370 0.711609
District2817 -0.2662408  0.4452910  -0.598 0.549935
District2819  0.6911713  0.4189805   1.650 0.099085 .
District2820  0.2578133  0.3807545   0.677 0.498370
District2821 -0.0342185  0.3215820  -0.106 0.915265
District2823  0.1309328  0.3320187   0.394 0.693339
District2901  0.1705660  0.3053666   0.559 0.576489
District2902 -1.6850370  0.3989313  -4.224 2.45e-05 ***
District2903  0.8181716  0.3488882   2.345 0.019067 *
District2907 -0.9611650  0.3809590  -2.523 0.011670 *
District2908  0.6106838  0.3561811   1.715 0.086501 .
District2909  0.4222662  0.3968086   1.064 0.287316
District2910  0.5747893  0.5418510   1.061 0.288844
District2912  0.4408689  0.3049223   1.446 0.148293
District2915  0.0193313  0.3388898   0.057 0.954513
District2917 -0.1116159  0.3473045  -0.321 0.747939
District2918  0.1646685  0.3207306   0.513 0.607685
District2920  0.7578848  0.2484437   3.051 0.002298 **
District2922  0.3370172  0.3213080   1.049 0.294285
District2924  0.6232234  0.2843729   2.192 0.028462 *
District2926 -0.0993217  0.2939020  -0.338 0.735423
District3001  0.7637087  0.2903976   2.630 0.008571 **
District301   0.4766653  0.2786953   1.710 0.087272 .
District302   0.3962808  0.3265781   1.213 0.225029
District303   0.3432103  0.2967952   1.156 0.247585
District304   0.7198355  0.2675955   2.690 0.007172 **
District307   0.5561953  0.4181985   1.330 0.183593
District309   0.6357688  0.2728678   2.330 0.019853 *
District310   0.1267947  0.3210440   0.395 0.692903
District311   0.3906234  0.2966350   1.317 0.187957
District317   0.3830590  0.2966171   1.291 0.196623
District3202  0.3825221  0.3119276   1.226 0.220144
District3204  0.6347995  0.2896178   2.192 0.028442 *
District3205  0.7279007  0.2819750   2.581 0.009871 **
```

```
District3206  0.7564096  0.5429542   1.393 0.163648
District3207  0.7019222  0.3267979   2.148 0.031777 *
District3208  0.3309982  0.2789767   1.186 0.235499
District3210  0.4350599  0.6482638   0.671 0.502182
District3211  0.3348950  0.2727543   1.228 0.219577
District3213  0.6077330  0.4190203   1.450 0.147027
District3214  0.6467128  0.3565490   1.814 0.069774 .
District3302  0.7348003  0.2600626   2.825 0.004742 **
District3303  0.1352199  0.3968563   0.341 0.733325
District3304 -0.0100705  0.3121067  -0.032 0.974261
District3305  1.3603687  0.8779569   1.549 0.121340
District3306  0.5528296  0.4465732   1.238 0.215805
District3309  0.5145680  0.3485660   1.476 0.139950
District3312  0.3972401  0.2992609   1.327 0.184443
District3313  0.3574569  0.4181444   0.855 0.392672
District3314  0.4359714  0.3324797   1.311 0.189833
District3317  0.4796694  0.4460790   1.075 0.282298
District3323  0.8519727  0.3673285   2.319 0.020420 *
District3324  0.4741096  0.3469715   1.366 0.171876
District3325  0.3606914  0.5416143   0.666 0.505474
District3327  0.0873308  0.3816804   0.229 0.819030
District3328  0.5768883  0.4203942   1.372 0.170053
District3329 -0.0298319  0.4832894  -0.062 0.950783
District3330  0.0415791  0.3813529   0.109 0.913183
District3400  0.7195087  0.2966766   2.425 0.015338 *
District400   0.9167556  0.2729963   3.358 0.000791 ***
District505   0.4980810  0.2852858   1.746 0.080897 .
District511   0.5818551  0.3088162   1.884 0.059610 .
District512   0.4324332  0.4456692   0.970 0.331949
District513   1.8673223  0.5411193   3.451 0.000564 ***
District602   0.8940013  0.2787147   3.208 0.001348 **
District603   1.6329533  0.8767397   1.863 0.062594 .
District608   0.8142791  0.3565887   2.284 0.022446 *
District609   0.8642943  0.3132008   2.760 0.005812 **
District612   1.0113189  0.2992829   3.379 0.000733 ***
 [ reached getOption("max.print") -- omitted 39 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8437 on 4448 degrees of freedom
  (76 observations deleted due to missingness)
Multiple R-squared:  0.3133,        Adjusted R-squared:  0.2766
F-statistic: 8.527 on 238 and 4448 DF,  p-value: < 2.2e-16
```




mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1



output:

```
> library(fixest)> library(AER)> data(Grunfeld)> feols_model<- feols(invest ~ val
ue + capital | firm + year , data = Grunfeld)> # one-way cluster by firm> feols_m
odel<- feols(invest ~ value + capital | firm + year , data = Grunfeld, cluster =
~firm)> > # two-way clustering by firm and year> feols_model<- feols(invest ~ val
ue + capital | firm + year , data = Grunfeld, cluster = ~firm + year)> # estimate
 linear two-way fixed effect model with two-way clusting> feols_model<- feols(inv
est ~ value + capital | firm + year , data = Grunfeld, cluster = ~firm + year)> >
 # get variance-covariance matrix with heteroskedasticity robust standard errors>
 hetero = vcov(feols_model, se = "hetero")> summary(feols_model)OLS estimation, D
ep. Var.: invest
Observations: 220
Fixed-effects: firm: 11,  year: 20
Standard-errors: Clustered (firm & year)
        Estimate Std. Error  t value   Pr(>|t|)
value   0.116681   0.011348 10.28247 1.2307e-06 ***
capital 0.351436   0.043537  8.07210 1.0884e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 45.7     Adj. R2: 0.944898
              Within R2: 0.725267> # Alternatively, use etable:> etable(feols_mode
l, tex = TRUE)\begingroup
\centering
\begin{tabular}{lc}
   \tabularnewline \midrule \midrule
   Dependent Variable: & invest\\
   Model:              & (1)\\
   \midrule
   \emph{Variables}\\
   value               & 0.1167$^{***}$\\
                       & (0.0114)\\
   capital             & 0.3514$^{***}$\\
                       & (0.0435)\\
   \midrule
   \emph{Fixed-effects}\\
   firm                & Yes\\
   year                & Yes\\
   \midrule
   \emph{Fit statistics}\\
   Observations        & 220\\
   R$^2$               & 0.95270\\
   Within R$^2$        & 0.72527\\
   \midrule \midrule
   \multicolumn{2}{l}{\emph{Clustered (firm \& year) standard-errors in parenthes
es}}\\
   \multicolumn{2}{l}{\emph{Signif. Codes: ***: 0.01, **: 0.05, *: 0.1}}\\
\end{tabular}
```

```
\par\endgroup> summary(feols_model, .vcov = hetero) # hetero is the var-cov matri
x that was previously computed using the vcov functionOLS estimation, Dep. Var.:
invest
Observations: 220
Fixed-effects: firm: 11,  year: 20
Standard-errors: Custom
        Estimate Std. Error t value   Pr(>|t|)
value   0.116681   0.019101 6.10872 5.6519e-09 ***
capital 0.351436   0.052870 6.64715 3.1464e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 45.7     Adj. R2: 0.944898
             Within R2: 0.725267> # OR> etable(feols_model, se = "white")
               feols_model
Dependent Var.:             invest

value           0.1167*** (0.0191)
capital         0.3514*** (0.0529)
Fixed-Effects:  ------------------
firm                           Yes
year                           Yes
_______________ __________________
S.E. type       Heteroskedas.-rob.
Observations                   220
R2                         0.95270
Within R2                  0.72527
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1> ####Task2: Import
the following .dta file to R> > Class28_8 <- read_dta("C:/Users/hp/Downloads/Clas
s28_8.dta")> View(Class28_8)> > #This data is> #This data is from PLFS 2017-18. T
he data includes employed males not working as casual labour.> #It includes males
 working in household enterprises/self employed or as salaried individuals.> #Q1:
 You want to examine whether social group-ie, caste identity affects earnings. >
#What should be your hypothesis? > #What should your dependent and main explanato
ry variables be?> #Q2: What variables should you control for, given your data?> #
Q3: Do you think you can estimate causal effect from this exercise?If yes why?If
not why?> #Q4: Interpret the results you get by running the regressions in Q1 and
 Q2.> > > > > #Now suppose you want to examine that self employment is not as rew
arding as salaried employment.> #Q5: What will be your hypothesis?> #What should
your dependent and main explanatory variables be?> #Q6: What variables should you
 control for, given your data?> #Q7: Do you think any variable can potentially ca
use hetroskedasticity?> #Q8: Can you say from your specification about existence
of multicollinearity?What should you do?> #Q7: Do you think you can estimate caus
al effect from this exercise?If yes why?If not why?> #Q8: Interpret the results y
ou get by running the OLS regressions in Q5 and Q6.> > > > feols_model<- feols(Gr
ossMonthlyEarnings ~ 0+ SC |+                         StateCode , data = Class28_8,
+                     weights=Weight,cluster = ~Fsu)
```

4. Autocorrelation:

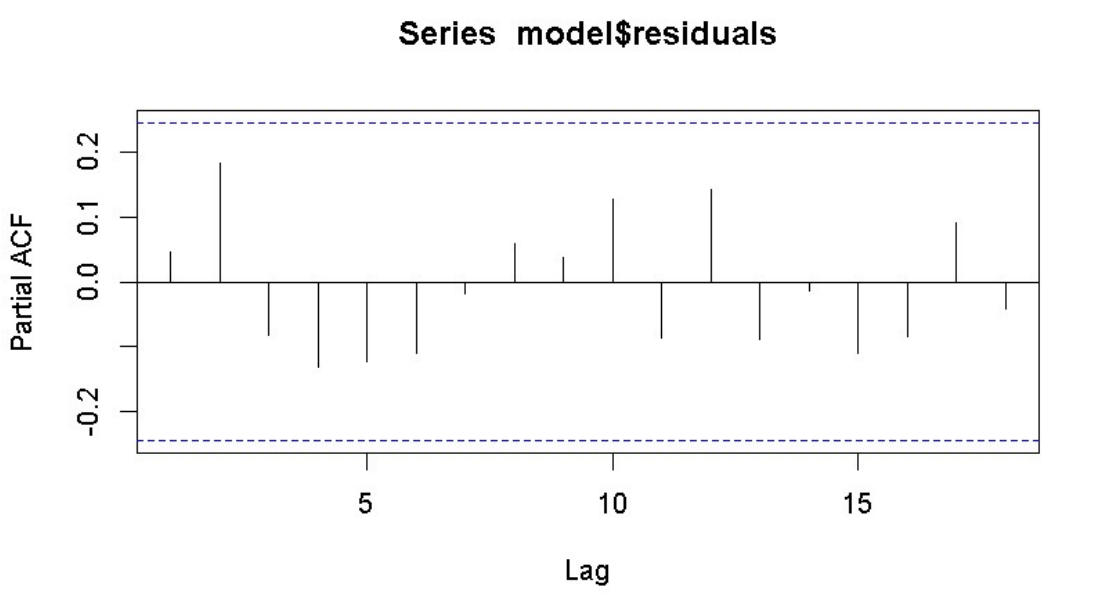
```
Code:
rm(list = ls())
# Install and load the readxl package
install.packages("readxl")
library(readxl)

# Read the Excel file
ExportImportOil <- read_excel("C:/Users/hp/Downloads/ExportImportOil.xlsx")
View(ExportImportOil)
ExportImportOil
library(stats)
# Rename the variables with spaces using backticks
ExportImportOil$`Oil Exports` <- ExportImportOil$`Oil Exports`
ExportImportOil$`Oil Imports` <- ExportImportOil$`Oil Imports`

# Fit the linear regression model
model <- lm(`Oil Exports` ~ `Oil Imports`, data = ExportImportOil)
summary(model)
acf(model$residuals, type = "correlation")
pacf(model$residuals)
library(lmtest)
dwtest(model)
bgtest(model, order = 2)
bgtest(model, order = 3)
bgtest(model, order = 4)

model1<-lm(Non.Oil.Imports~Oil.Imports, data = ExportImportOil)
summary(model1)
acf(model1$residuals, type = "correlation")
pacf(model1$residuals)
```

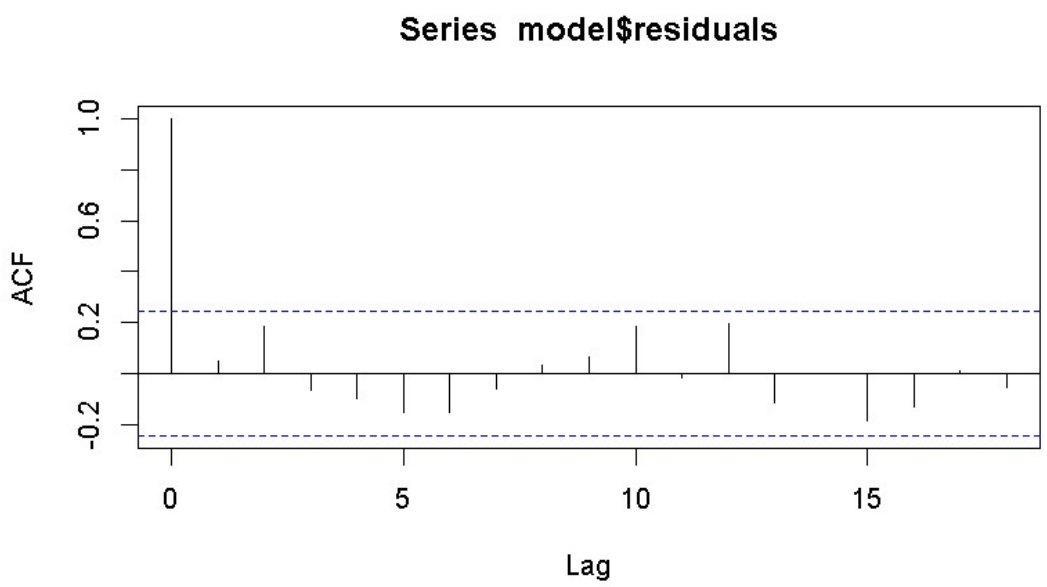
```
dwtest(model1)
bgtest(model, order = 2)
bgtest(model, order = 3)
bgtest(model, order = 4)


#Time series Classical Decomposition
data.ts <- ts(ExportImportOil$'Oil Imports', frequency = 12)
data.ts
ts.plot(data.ts, xlab="Time Period", ylab="Oil Imports", main="Monthly Oil Import
s")
plot(pacf(ExportImportOil$Oil.Imports,plot=FALSE),main="Partial Autocorrelation P
lot")
plot(acf(ExportImportOil$Oil.Imports,plot=FALSE),main="Autocorrelation Plot")
decomp<-decompose(data.ts)
plot(decomp)
decomp$seasonal
decomp$trend
decomp$random
seasadj <- data.ts - decomp$seasonal
plot(seasadj)

#decomposition by loess method.
decomp1<-stl(data.ts,s.window="periodic")
plot(decomp1)
seasonal_stl_model1 <- decomp1$time.series[,1]
trend_stl_model1 <- decomp1$time.series[,2]
random_stl_model1 <- decomp1$time.series[,3]


seasadj1 <- data.ts - seasonal_stl_model1
trendadj1<-data.ts-trend_stl_model1
plot(trendadj1)

plot(pacf(ExportImportOil$Oil.Imports,plot=FALSE),main="Partial Autocorrelation P
lot")
plot(acf(ExportImportOil$Non.Oil.Imports,plot=FALSE),main="Partial Autocorrelatio
n Plot")

Diff1<-diff(ExportImportOil$Non.Oil.Exports, differences=1)
plot.ts(Diff1)

Diff21<-diff(trendadj1, differences=2)
plot.ts(Diff21)

Diff31<-diff(trendadj1, differences=3)
plot.ts(Diff31)


####Forecasting-Holt Winters

HW1 <- HoltWinters(data.ts)

HW2 <- HoltWinters(data.ts, alpha=0.2, beta=0.1, gamma=0.1)

HW1.pred <- predict(HW1, 6, prediction.interval = TRUE, level=0.95)


Output:
 View(ExportImportOil)> library(stats)> model<-lm(Oil.Exports~Oil.Imports, data =
 ExportImportOil)Error in eval(predvars, data, env) : object 'Oil.Exports' not fo
und> View(ExportImportOil)> rm(list = ls())> ExportImportOil <- read.csv("C:/User
s/hp/Downloads/ExportImportOil.xlsx")Error in read.table(file = file, header = he
ader, sep = sep, quote = quote,  :
  more columns than column namesIn addition: Warning messages:
1: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 2 appears to contain embedded nulls
2: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 4 appears to contain embedded nulls
3: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  incomplete final line found by readTableHeader on 'C:/Users/hp/Downloads/Export
ImportOil.xlsx'> rm(list = ls())> ExportImportOil <- read.csv("C:/Users/hp/Downlo
ads/ExportImportOil.xlsx")Error in read.table(file = file, header = header, sep =
 sep, quote = quote,  :
  more columns than column names
In addition: Warning messages:
1: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 2 appears to contain embedded nulls
2: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 4 appears to contain embedded nulls
3: In read.table(file = file, header = header, sep = sep, quote = quote,  :
```

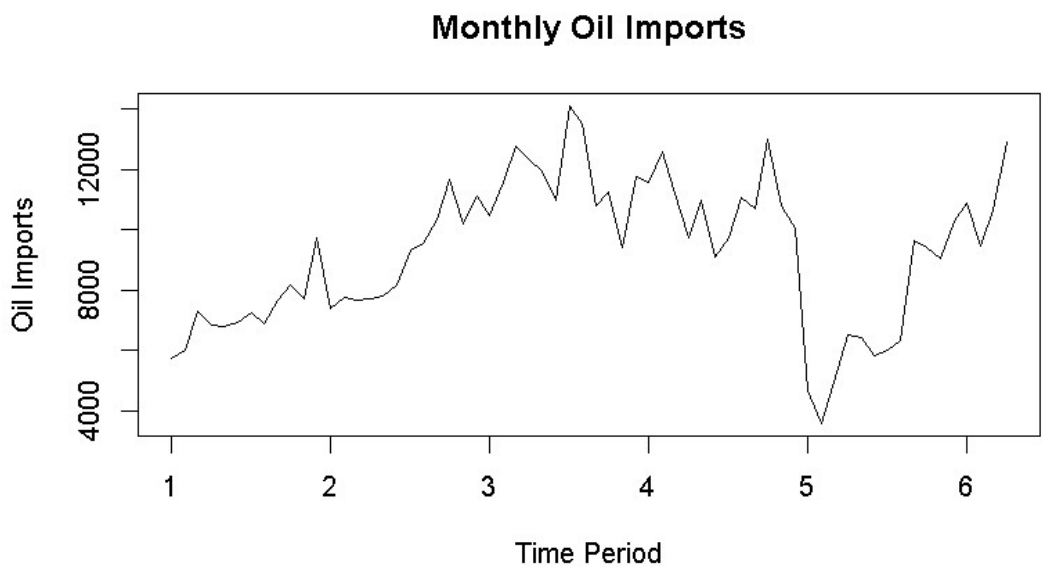
```
  incomplete final line found by readTableHeader on 'C:/Users/hp/Downloads/Export
ImportOil.xlsx'> ExportImportOil <- read_fwf("C:/Users/hp/Downloads/ExportImportO
il.xlsx")Multiple files in zip: reading '[Content_Types].xml'Error: `file` must b
e a regular file, not a connection> ExportImportOil <- read.fwf("C:/Users/hp/Down
loads/ExportImportOil.xlsx")Error in read.fwf("C:/Users/hp/Downloads/ExportImport
Oil.xlsx") :
  argument "widths" is missing, with no default> ExportImportOil <- read.csv("C:/
Users/hp/Downloads/ExportImportOil.xlsx")Error in read.table(file = file, header
= header, sep = sep, quote = quote,  :
  more columns than column names
In addition: Warning messages:
1: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 2 appears to contain embedded nulls
2: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  line 4 appears to contain embedded nulls
3: In read.table(file = file, header = header, sep = sep, quote = quote,  :
  incomplete final line found by readTableHeader on 'C:/Users/hp/Downloads/Export
ImportOil.xlsx'> install.packages("readxl")WARNING: Rtools is required to build R
 packages but is not currently installed. Please download and install the appropr
iate version of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/hp/AppData/Local/R/win-library/4.3'
(as 'lib' is unspecified)trying URL 'https://cran.rstudio.com/bin/windows/contrib
/4.3/readxl_1.4.3.zip'Content type 'application/zip' length 1197336 bytes (1.1 MB)
downloaded 1.1 MB
package 'readxl' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\hp\AppData\Local\Temp\Rtmpai5ZZ1\downloaded_packages> library(r
eadxl)Warning message:
package 'readxl' was built under R version 4.3.2 > ExportImportOil <- read_excel
("C:/Users/hp/Downloads/ExportImportOil.xlsx")> View(ExportImportOil)> library(st
ats)> model<-lm(Oil.Exports~Oil.Imports, data = ExportImportOil)Error in eval(pre
dvars, data, env) : object 'Oil.Exports' not found> model<-lm(Oil.Exports~Oil Imp
orts, data = ExportImportOil)Error: unexpected symbol in "model<-lm(Oil.Exports~O
il Imports"> ExportImportOil# A tibble: 64 × 7
   Year    Month     Period `Oil Exports` `Non-Oil Exports` `Oil Imports` `Non-Oi
l Imports`
   <chr>   <chr>      <dbl>         <dbl>             <dbl>         <dbl>
     <dbl> 1 2016-17 April          1         2048.             18815.         571
6.            19973. 2 2016-17 May            2         2125.             20282.
       5996.            22290. 3 2016-17 June           3         2601
  20055.         7291.          23626. 4 2016-17 July           4         2573.
           19119.         6823.           22482. 5 2016-17 August         5
     2488.            19109.         6790.           22514. 6 2016-17 September
     6         2574            20195.         6911.            24929. 7 2016-17
October        7         2776            20585.         7261.            27234.
8 2016-17 November       8         2431            17635.         6864.
   26598. 9 2016-17 December       9         2887.             21169         7667
           26936.10 2016-17 January       10         2747.             19610.
      8174            24087.# ⓘ 54 more rows# ⓘ Use `print(n = ...)` to see mo
re rows> model<-lm(Oil Exports~Oil Imports, data = ExportImportOil)Error: unexpec
ted symbol in "model<-lm(Oil Exports"> library(stats)> model<-lm(Oil Exports~Oil
Imports, data = ExportImportOil)Error: unexpected symbol in "model<-lm(Oil Export
s"> # Rename the variables with spaces using backticks> ExportImportOil$`Oil Expo
rts` <- ExportImportOil$`Oil Exports`> ExportImportOil$`Oil Imports` <- ExportImp
ortOil$`Oil Imports`> > # Fit the linear regression model> model <- lm(`Oil Expor
ts` ~ `Oil Imports`, data = ExportImportOil)> summary(model)
Call:
lm(formula = `Oil Exports` ~ `Oil Imports`, data = ExportImportOil)

Residuals:
     Min       1Q   Median       3Q      Max
-1093.42  -366.52    -5.21   240.70  2118.91

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   340.81173  306.54054   1.112    0.271
`Oil Imports`   0.30253    0.03201   9.451 1.26e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 609.8 on 62 degrees of freedom
Multiple R-squared:  0.5903,        Adjusted R-squared:  0.5837
F-statistic: 89.32 on 1 and 62 DF,  p-value: 1.263e-13
> acf(model$residuals, type = "correlation")> pacf(model$residuals)> acf(model$re
siduals, type = "correlation")> pacf(model$residuals)> library(lmtest)> dwtest(mo
del)
        Durbin-Watson test

data:  model
```

```
DW = 1.7948, p-value = 0.175
alternative hypothesis: true autocorrelation is greater than 0
> bgtest(model, order = 2)
        Breusch-Godfrey test for serial correlation of order up to 2

data:  model
LM test = 2.6885, df = 2, p-value = 0.2607
> bgtest(model, order = 3)
        Breusch-Godfrey test for serial correlation of order up to 3

data:  model
LM test = 3.0016, df = 3, p-value = 0.3914
> bgtest(model, order = 4)
        Breusch-Godfrey test for serial correlation of order up to 4

data:  model
LM test = 4.0479, df = 4, p-value = 0.3996
> model1<-lm(Non.Oil.Imports~Oil.Imports, data = ExportImportOil)Error in eval(pr
edvars, data, env) : object 'Non.Oil.Imports' not found> #Time series Classical D
ecomposition> data.ts <- ts(ExportImportOil$Oil.Imports, frequency = 12)Error in
ts(ExportImportOil$Oil.Imports, frequency = 12) :
  'ts' object must have one or more observationsIn addition: Warning message:
Unknown or uninitialised column: `Oil.Imports`. > data.ts <- ts(ExportImportOil$O
il.Imports, frequency = 12)Error in ts(ExportImportOil$Oil.Imports, frequency = 1
2) :
  'ts' object must have one or more observationsIn addition: Warning message:
Unknown or uninitialised column: `Oil.Imports`. > #Time series Classical Decompos
ition> data.ts <- ts(ExportImportOil$Oil.Imports, frequency = 12)Error in ts(Expo
rtImportOil$Oil.Imports, frequency = 12) :
  'ts' object must have one or more observationsIn addition: Warning message:
Unknown or uninitialised column: `Oil.Imports`. > #Time series Classical Decompos
ition> data.ts <- ts(ExportImportOil$Oil Imports, frequency = 12)Error: unexpecte
d symbol in "data.ts <- ts(ExportImportOil$Oil Imports"> #Time series Classical D
ecomposition> data.ts <- ts(ExportImportOil$'Oil Imports', frequency = 12)> data.
ts      Jan     Feb     Mar     Apr     May     Jun     Jul     Aug     Sep     O
ct      Nov
1  5716.1  5996.3  7291.4  6822.7  6789.5  6911.3  7261.2  6864.3  7667.0  8174.0
  7719.8
2  7396.2  7734.8  7677.2  7691.6  7802.6  8169.1  9309.3  9548.5 10345.0 11659.0
 10196.0
3 10464.0 11573.6 12764.6 12331.3 11939.1 10991.2 14105.9 13518.3 10783.9 11253.1
  9414.0
4 11564.3 12592.8 11205.0  9748.6 10999.4  9086.3  9728.6 11069.1 10718.9 13009.2
 10780.5
5  4649.6  3568.2  4967.5  6530.6  6440.5  5825.6  6000.3  6314.8  9629.0  9424.3
  9031.4
6 10871.4  9452.0 10676.2 12894.6

      Dec
1  9750.3
2 11117.7
3 11780.2
4 10047.6
5 10271.4
6        > ts.plot(data.ts, xlab="Time Period", ylab="Oil Imports", main="Monthly
 Oil Imports")> plot(pacf(ExportImportOil$Oil.Imports,plot=FALSE),main="Partial A
utocorrelation Plot")
plot:
```

Plot:




Series model$residuals
ACF
1.0
0.6
0.2
-0.2
0
5
10
15
Lag


Monthly Oil Imports
Oil Imports
12000
8000
4000
1
2
3
4
5
6
Time Period