

The figure shows a screenshot of the RStudio interface. The left pane contains an R script editor with code for analyzing hospital costs, including a histogram of age. The right pane shows the Global Environment and a histogram titled "Histogram of hospitalcost\$AGE".

R Script Editor:

```
16 t <- table(hospitalcost$APRDRG)
17 d <- as.data.frame(t)
18 names(d)[1] = 'Diagnosis Group'
19 d
20 which.max(table(hospitalcost$APRDRG))
21 which.max(t)
22 which.max(d)
23 res <- aggregate(TOTCHG ~ APRDRG, data = hospitalcost, sum)
24 res
25 which.max(res$TOTCHG)
26 res[which.max(res$TOTCHG),]
27
28 #To make sure that there is no malpractice, the agency needs to analyze if the race
29 table(hospitalcost$RACE)
30 hospitalcost$RACE <- as.factor(hospitalcost$RACE)
31 fit <- lm(TOTCHG ~ RACE, data=hospitalcost)
32 fit
33 summary(fit)
34
35
```

Console:

```
R 4.4.0 --> 
63 952 4833
> which.max(res$TOTCHG)
[1] 44
> res[which.max(res$TOTCHG),]
APRDRG TOTCHG
44 640 437978
> #To make sure that there is no malpractice, the agency needs to analyze if the race o
    the patient is related to the hospitalization costs
> table(hospitalcost$RACE)
```

Age	Frequency
1	280
2	10
3	10
4	5
5	5
6	6
7	10
8	10
9	10
10	10
11	10
12	10
13	10
14	10
15	10
16	10
17	10
18	10
19	10
20	10
21	10
22	10
23	10
24	10
25	10
26	10
27	10
28	10
29	10
30	10
31	10
32	10
33	10
34	10
35	10

Global Environment:

Object	Type	Size
d	63 obs. of 2 variables	152 MB
fit	List of 14	
fit1	List of 14	
hospitalcost	500 obs. of 6 variables	
mod	List of 13	
res	63 obs. of 2 variables	
X1555054100_hospic...	500 obs. of 6 variables	

Histogram of hospitalcost\$AGE

The histogram displays the frequency distribution of hospital costs by age group. The x-axis is labeled "hospitalcost\$AGE" and ranges from 0 to 25. The y-axis is labeled "Frequency" and ranges from 0 to 300. The distribution is highly right-skewed, with the highest frequency occurring at age 1.

The screenshot shows the RStudio interface with two main panes. The left pane contains the R Script editor with the following code:

```
19 d
20 which.max(table(hospitalcost$APRDRG))
21 which.max(t)
22 which.max(d)
23 res <- aggregate(TOTCHG ~ APRDRG, data = hospitalcost, sum)
24 res
25 which.max(res$TOTCHG)
26 res[which.max(res$TOTCHG),]
27
28 #To make sure that there is no malpractice, the agency needs to analyze if the race
29 table(hospitalcost$RACE)
30 hospitalcost$RACE <- as.factor(hospitalcost$RACE)
31 fit <- lm(TOTCHG ~ RACE,data=hospitalcost)
32 fit
33 summary(fit)
34 fit1 <- aov(TOTCHG ~ RACE,data=hospitalcost)
35 summary(fit1)
36 hospitalcost <- na.omit(hospitalcost)
37
```

The right pane shows the Global Environment and a Histogram of hospitalcost\$AGE.

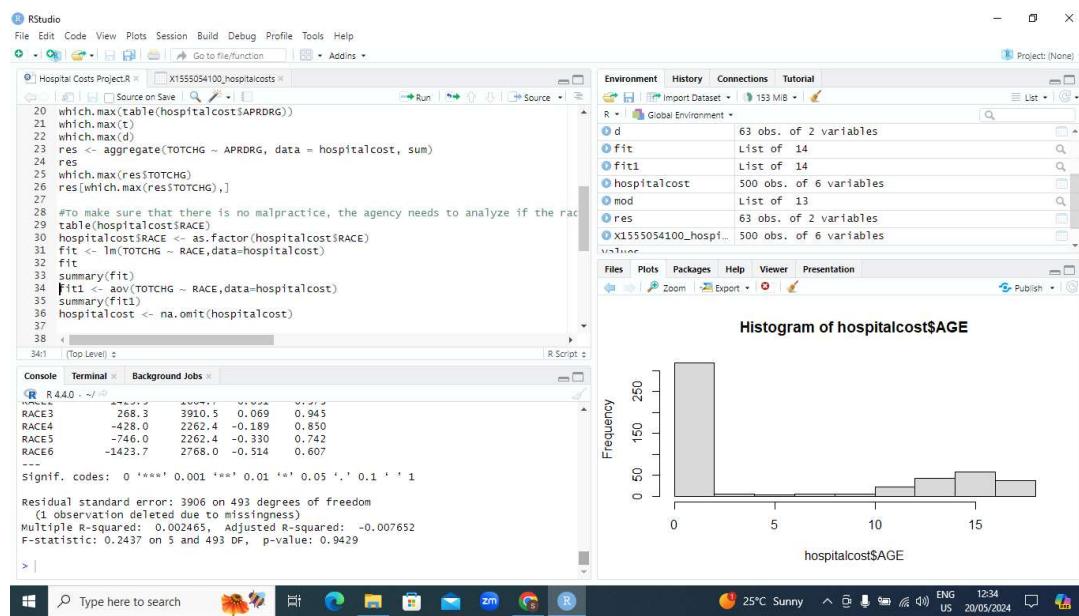
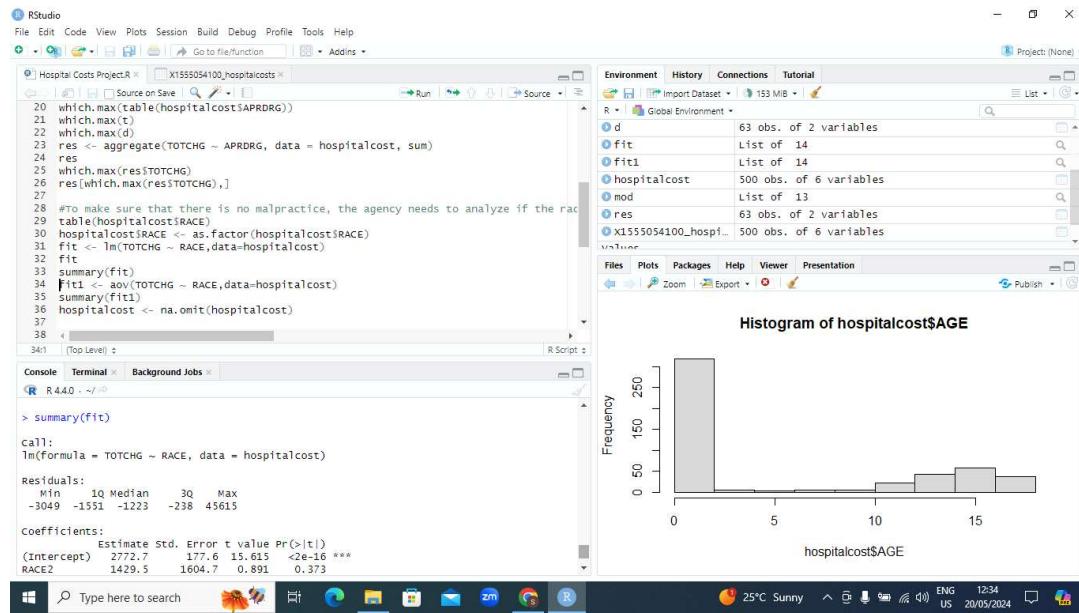
Global Environment:

- d: 63 obs. of 2 variables
- fit: List of 14
- fit1: List of 14
- hospitalcost: 500 obs. of 6 variables
- mod: List of 13
- res: 63 obs. of 2 variables
- X155054100_hospit...: 500 obs. of 6 variables

Histogram of hospitalcost\$AGE:

The histogram displays the frequency distribution of hospitalcost\$AGE. The x-axis is labeled "hospitalcost\$AGE" and ranges from 0 to 15. The y-axis is labeled "Frequency" and ranges from 0 to 250. The distribution is highly right-skewed, with the highest frequency occurring at age 1.

Age Group	Frequency
0	~280
1	~280
2	~10
3	~10
4	~10
5	~10
6	~20
7	~20
8	~20
9	~20
10	~20
11	~20
12	~20
13	~20
14	~20
15	~20
16	~20



RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)

Hospital Costs Project.R x X:\555054100_hospitalcosts
Source on Save Run Source
22 which.max(d)
23 res <- aggregate(TOTCHG ~ APRDRG, data = hospitalcost, sum)
24 res
25 which.max(res$TOTCHG)
26 res[which.max(res$TOTCHG),]
27
28 #to make sure that there is no malpractice, the agency needs to analyze if the race
29 table(hospitalcost$RACE)
30 hospitalcost$RACE <- as.factor(hospitalcost$RACE)
31 fit <- lm(TOTCHG ~ RACE,data=hospitalcost)
32 fit
33 summary(fit)
34 fit1 <- aov(TOTCHG ~ RACE,data=hospitalcost)
35 summary(fit1)
36 hospitalcost <- na.omit(hospitalcost)
37
38 #to properly utilize the costs, the agency has to analyze the severity of the hospital
39 table(hospitalcost$FEMALE)
40
41
Residual standard error: 3906 on 493 degrees of freedom
(1 observation deleted due to missingness)
Multiple R-squared:  0.002465, Adjusted R-squared: -0.007652
F-statistic: 0.2437 on 5 and 493 DF, p-value: 0.9429

> fit1 <- aov(TOTCHG ~ RACE,data=hospitalcost)
> summary(fit1)
   Df Sum Sq Mean Sq F value Pr(>F)
RACE      5 1.859e+07 3718656  0.244  0.943
Residuals 493 7.524e+09 15260687
1 observation deleted due to missingness
>

```

Environment History Connections Tutorial

R | Global Environment

- d 63 obs. of 2 variables
- fit List of 14
- fit1 List of 14
- hospitalcost 500 obs. of 6 variables
- mod List of 13
- res 63 obs. of 2 variables
- X:\555054100_hospitalcost 500 obs. of 6 variables

Files Plots Packages Help Viewer Presentation

Histogram of hospitalcost\$AGE

Frequency

0 5 10 15

0 50 150 250

hospitalcost\$AGE

Console Terminal Background Jobs

R 4.40 . ~/R/

25°C Sunny 12:34 ENG US 20/05/2024

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)

Hospital Costs Project.R x X:\555054100_hospitalcosts
Source on Save Run Source
27
28 #to make sure that there is no malpractice, the agency needs to analyze if the race
29 table(hospitalcost$RACE)
30 hospitalcost$RACE <- as.factor(hospitalcost$RACE)
31 fit <- lm(TOTCHG ~ RACE,data=hospitalcost)
32 fit
33 summary(fit)
34 fit1 <- aov(TOTCHG ~ RACE,data=hospitalcost)
35 summary(fit1)
36 hospitalcost <- na.omit(hospitalcost)
37
38 #to properly utilize the costs, the agency has to analyze the severity of the hospital
39 table(hospitalcost$FEMALE)
40 a <- aov(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
41 summary(a)
42 b <- lm(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
43 summary(b)
44
45
DF Sum Sq Mean Sq F value Pr(>F)
RACE      5 1.859e+07 3718656  0.244  0.943
Residuals 493 7.524e+09 15260687
1 observation deleted due to missingness
> hospitalcost <- na.omit(hospitalcost)
> #properly utilize the costs, the agency has to analyze the severity of the hospital
> costs by age and gender for proper allocation of resources.
> table(hospitalcost$FEMALE)

0 1
244 255
> a <- aov(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
>
```

Environment History Connections Tutorial

R | Global Environment

- data List of 12
- a List of 12
- age List of 2 variables
- b List of 12
- cat List of 13
- d 63 obs. of 2 variables
- fit List of 14
- fit1 List of 14

Files Plots Packages Help Viewer Presentation

Histogram of hospitalcost\$AGE

Frequency

0 5 10 15

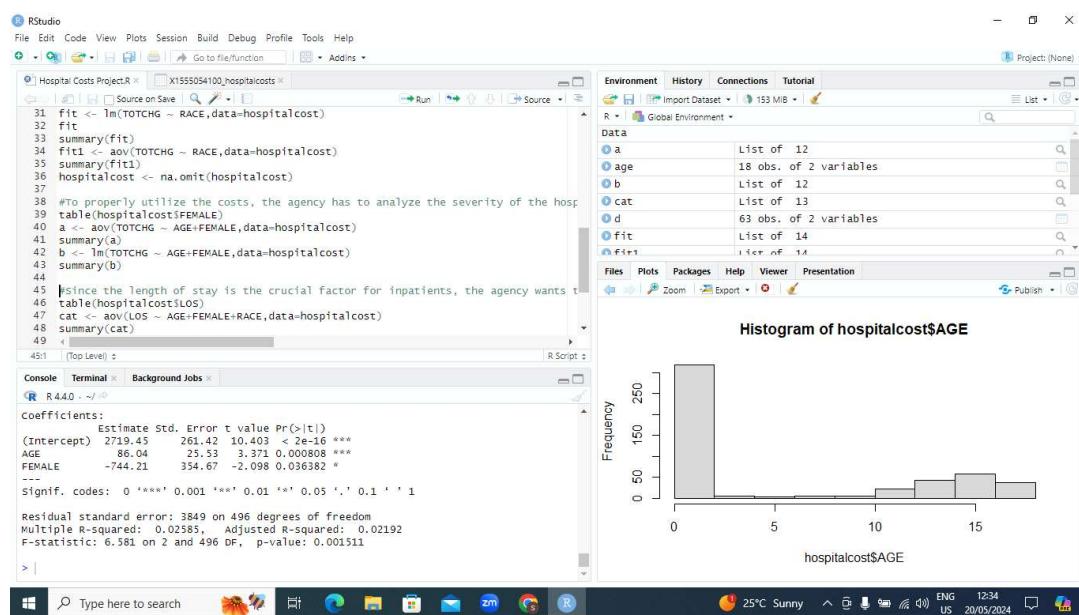
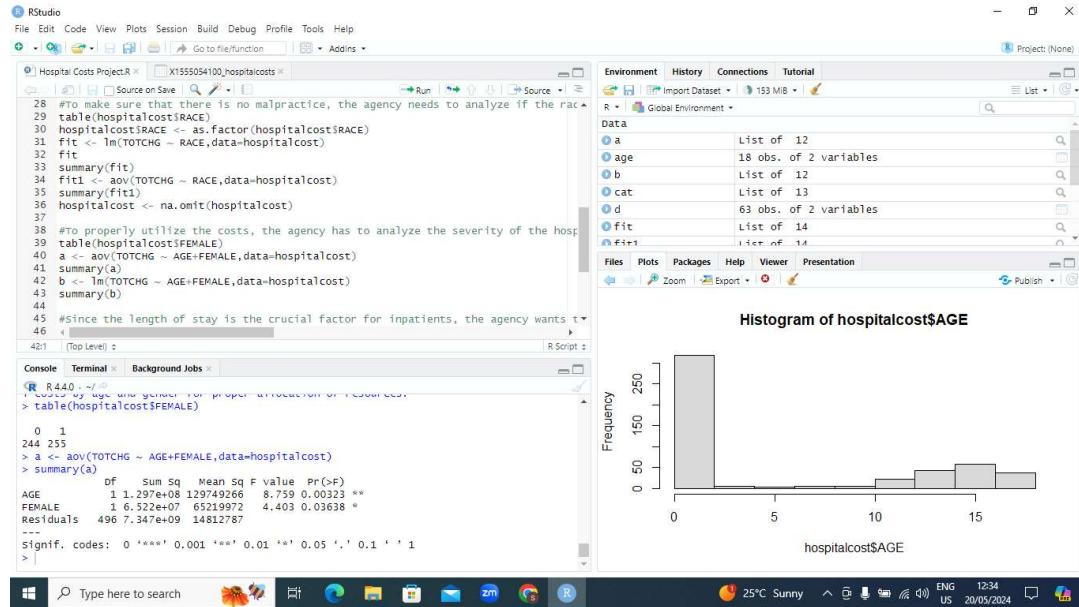
0 50 150 250

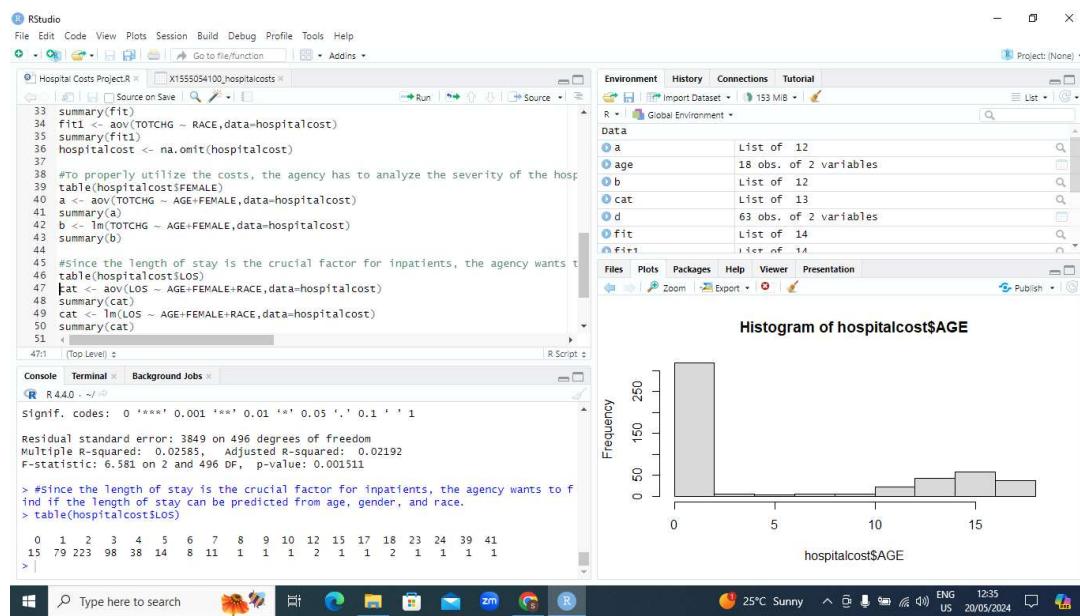
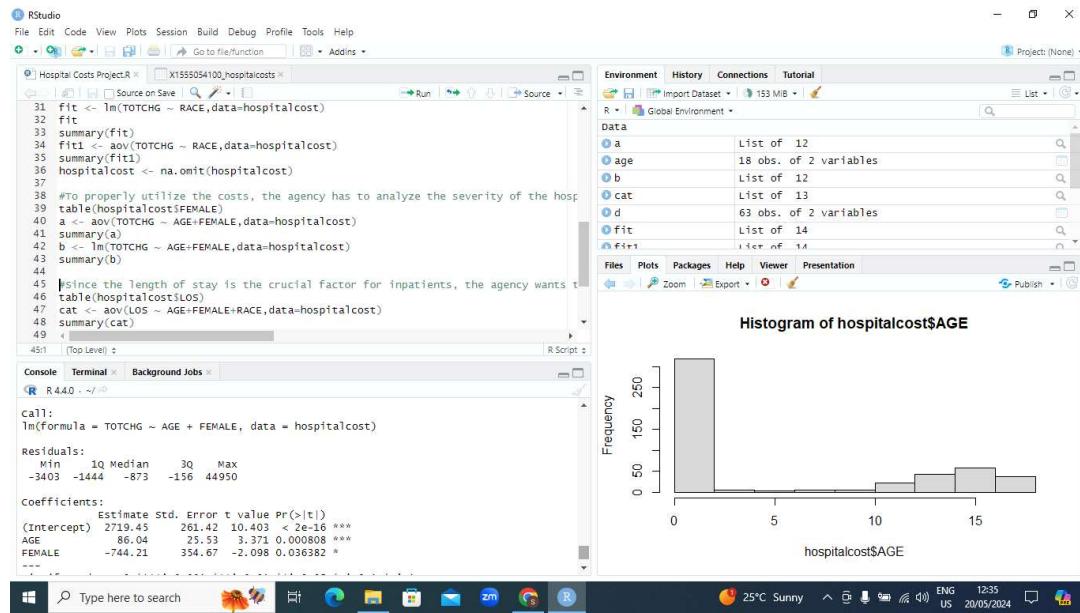
hospitalcost\$AGE

Console Terminal Background Jobs

R 4.40 . ~/R/

25°C Sunny 12:34 ENG US 20/05/2024





RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

Hospital Costs ProjectR X:\1555054100\hospitalcosts

```

34 fit1 <- aov(TOTCHG ~ RACE,data=hospitalcost)
35 summary(fit1)
36 hospitalcost <- na.omit(hospitalcost)
37
38 #To properly utilize the costs, the agency has to analyze the severity of the hospital
39 table(hospitalcost$FEMALE)
40 a <- aov(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
41 summary(a)
42 b <- lm(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
43 summary(b)
44
45 #since the length of stay is the crucial factor for inpatients, the agency wants to find if the length of stay can be predicted from age, gender, and race.
46 table(hospitalcost$LOS)
47 cat <- aov(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
48 summary(cat)
49 cat <- lm(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
50 summary(cat)
51
52 <-

```

(Top Level) R Script

Console Terminal Background Jobs

R 4.4.0

```

Residual standard error: 3849 on 496 degrees of freedom
Multiple R-squared:  0.02585, Adjusted R-squared:  0.02192
F-statistic: 6.584 on 2 and 496 DF, p-value: 0.000511

> #since the length of stay is the crucial factor for inpatients, the agency wants to find if the length of stay can be predicted from age, gender, and race.
> table(hospitalcost$LOS)

0 1 2 3 4 5 6 7 8 9 10 12 15 17 18 23 24 39 41
15 79 223 98 38 14 8 11 1 1 1 2 1 1 2 1 1 1 1 1 1 1
> cat <- aov(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
>

```

Frequency

Histogram of hospitalcost\$AGE

0 5 10 15

hospitalcost\$AGE

Windows Taskbar: Type here to search, 25°C Sunny, ENG US, 12:35, 20/05/2024

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

Hospital Costs ProjectR X:\1555054100\hospitalcosts

```

38 #To properly utilize the costs, the agency has to analyze the severity of the hospital
39 table(hospitalcost$FEMALE)
40 a <- aov(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
41 summary(a)
42 b <- lm(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
43 summary(b)
44
45 #since the length of stay is the crucial factor for inpatients, the agency wants to find the variable that mainly
46 aov(TOTCHG ~ ,data=hospitalcost)
47 mod <- lm(TOTCHG ~ ,data=hospitalcost)
48 summary(mod)
49 summary(mod)
50
51
52 #to perform a complete analysis, the agency wants to find the variable that mainly
53 aov(TOTCHG ~ ,data=hospitalcost)
54 mod <- lm(TOTCHG ~ ,data=hospitalcost)
55 summary(mod)
56

```

(Top Level) R Script

Console Terminal Background Jobs

R 4.4.0

```

Residuals 491 5595 11.396
> cat <- lm(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
> summary(cat)

Call:
lm(formula = LOS ~ AGE + FEMALE + RACE, data = hospitalcost)

Residuals:
Min 1Q Median 3Q Max
-3.211 -1.211 -0.857 0.143 37.789

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.85687 0.23160 12.335 <2e-16 ***

```

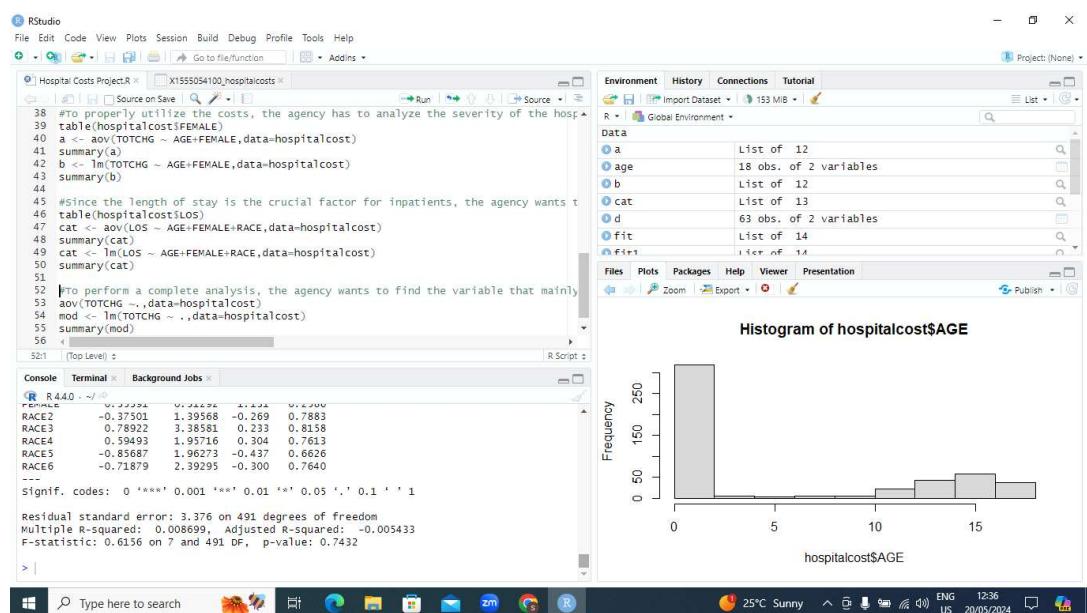
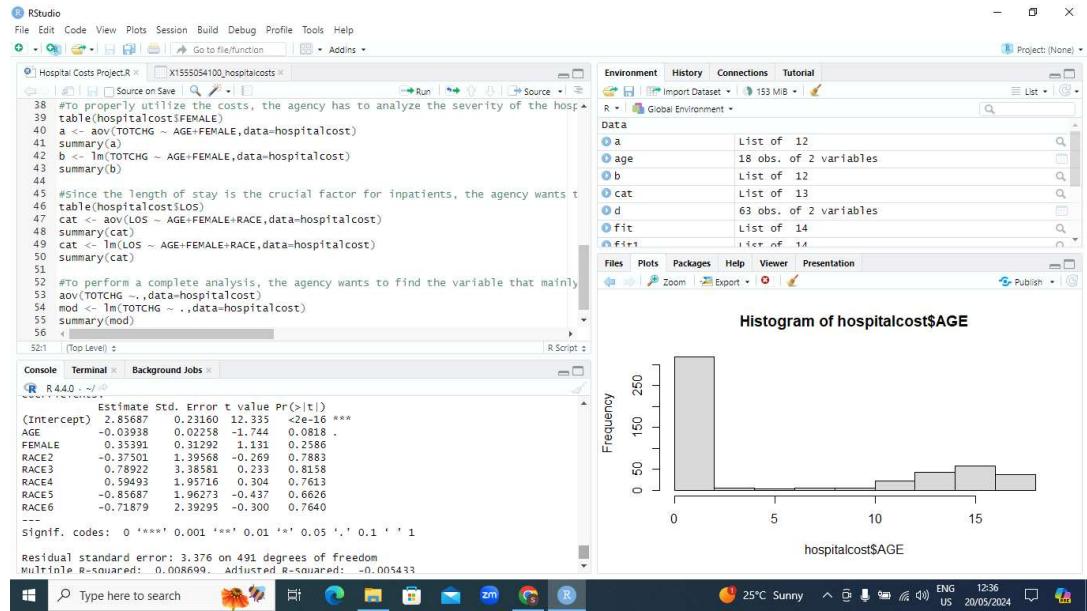
Frequency

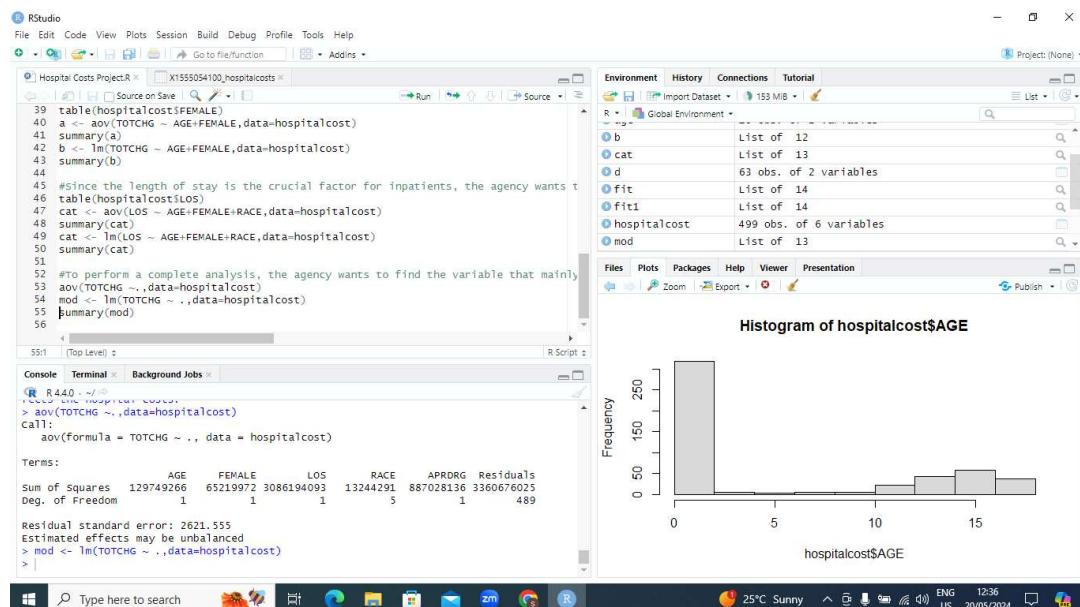
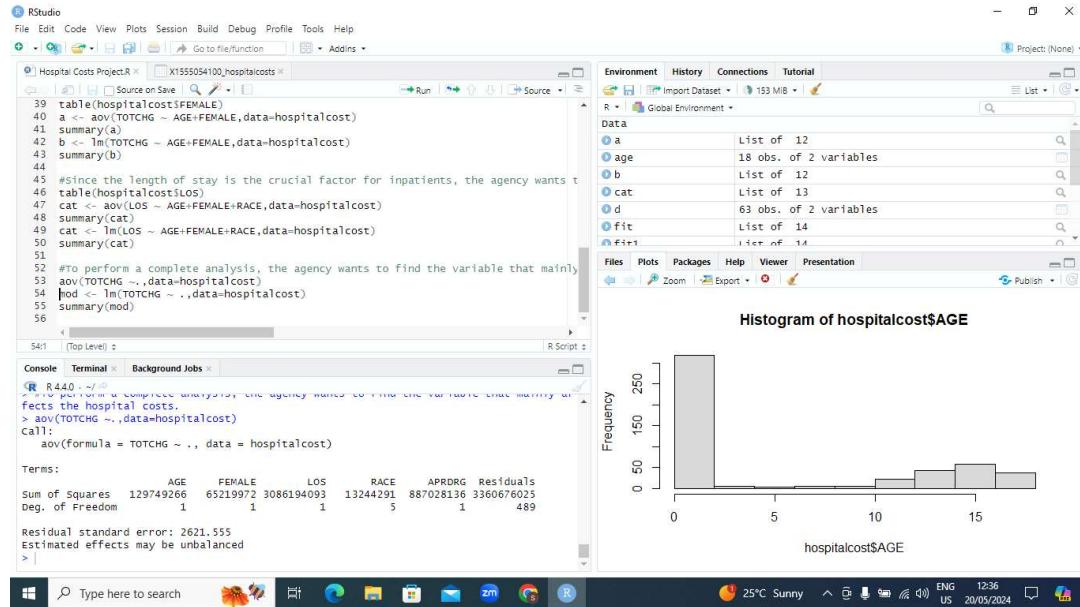
Histogram of hospitalcost\$AGE

0 5 10 15

hospitalcost\$AGE

Windows Taskbar: Type here to search, 25°C Sunny, ENG US, 12:35, 20/05/2024





RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Hospital Costs Project.R X:\1555054100_hospitalcosts
Source on Save Run Source Environment History Connections Tutorial
Project: (None)
R Global Environment
b List of 12
cat List of 13
d 63 obs. of 2 variables
fit List of 14
fit1 List of 14
hospitalcost 499 obs. of 6 variables
mod List of 13

Files Plots Packages Help Viewer Presentation
Zoom Export Publish

Histogram of hospitalcost$AGE

Frequency
0 50 150 250
0 5 10 15
hospitalcost$AGE

R 4.40 . ~/ ~
> mod <- lm(TOTCHG ~ ., data=hospitalcost)
> summary(mod)

Call:
lm(formula = TOTCHG ~ ., data = hospitalcost)

Residuals:
Min 1Q Median 3Q Max
-6367 -691 -186 121 43412

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 5024.9610 440.1366 11.417 < 2e-16 ***
AGE 133.2207 17.6662 7.541 2.29e-13 ***

```

(Top Level) R Script

Console Terminal Background Jobs

Type here to search 25°C Sunny ENG 12:36 US 20/05/2024

RStudio

```

File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins

Hospital Costs Project.R X:\1555054100_hospitalcosts
Source on Save Run Source Environment History Connections Tutorial
Project: (None)
R Global Environment
b List of 12
cat List of 13
d 63 obs. of 2 variables
fit List of 14
fit1 List of 14
hospitalcost 499 obs. of 6 variables
mod List of 13

Files Plots Packages Help Viewer Presentation
Zoom Export Publish

Histogram of hospitalcost$AGE

Frequency
0 50 150 250
0 5 10 15
hospitalcost$AGE

R 4.40 . ~/ ~
> mod <- lm(TOTCHG ~ ., data=hospitalcost)
> summary(mod)

Call:
lm(formula = TOTCHG ~ ., data = hospitalcost)

Residuals:
AGE FEMALE LOS RACE APRDRG Residuals
Sum of squares 129749266 65219972 3086194093 13244291 887028136 3360676025
Deg. of Freedom 1 1 1 5 1 489

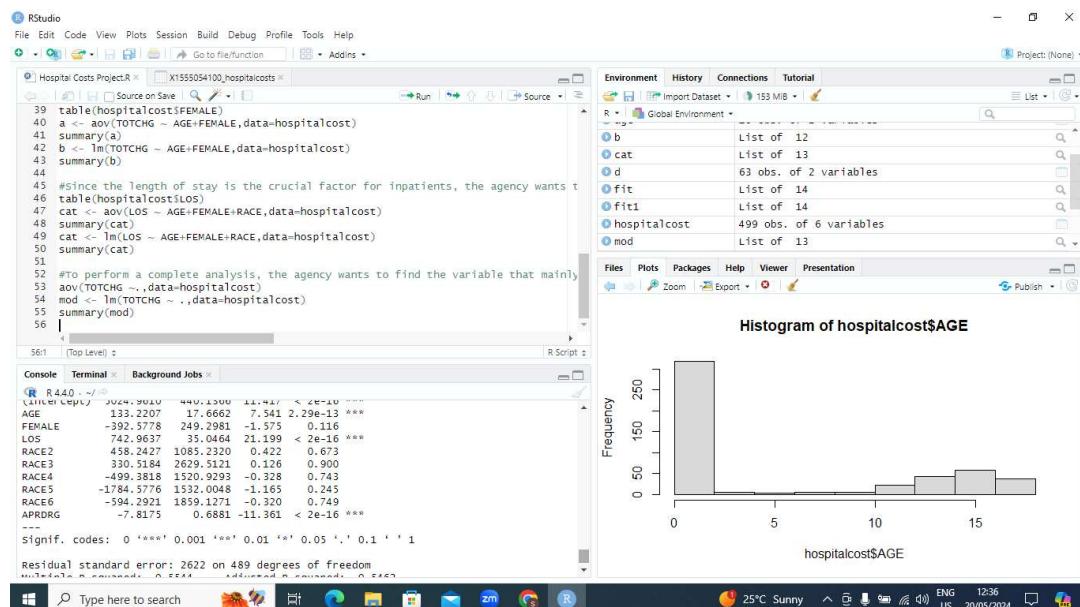
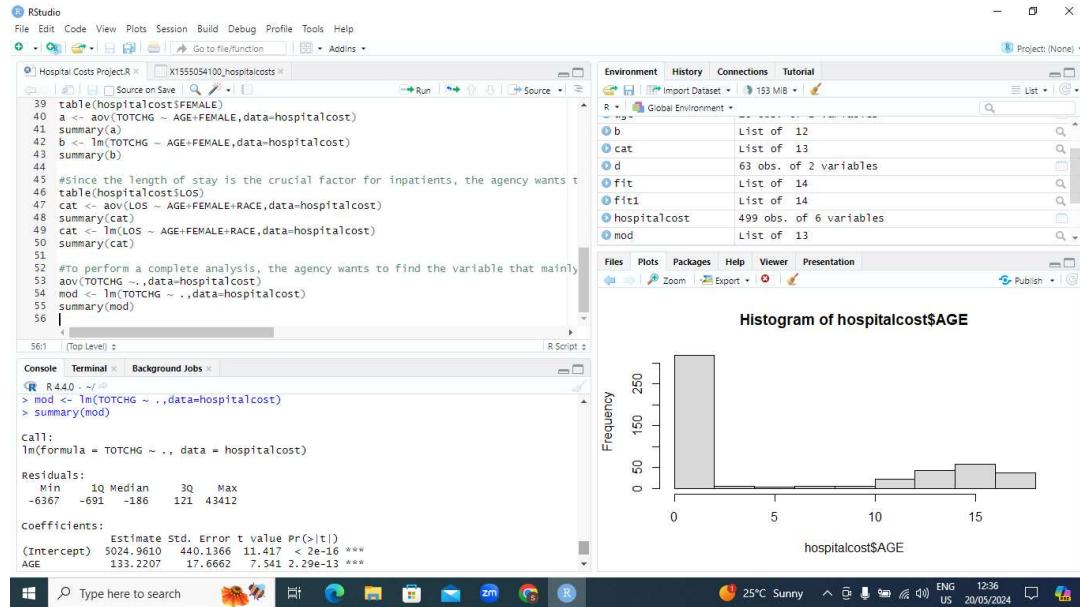
Residual standard error: 2621.555
Estimated effects may be unbalanced
> mod <- lm(TOTCHG ~ ., data=hospitalcost)
> |

```

(Top Level) R Script

Console Terminal Background Jobs

Type here to search 25°C Sunny ENG 12:36 US 20/05/2024



RStudio

Hospital Costs Project.R

```

39: table(hospitalcost$FEMALE)
40: a <- aov(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
41: summary(a)
42: b <- lm(TOTCHG ~ AGE+FEMALE,data=hospitalcost)
43: summary(b)
44:
45: #since the length of stay is the crucial factor for inpatients, the agency wants to
46: table(hospitalcost$LOS)
47: cat <- aov(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
48: summary(cat)
49: cat <- lm(LOS ~ AGE+FEMALE+RACE,data=hospitalcost)
50: summary(cat)
51:
52: #To perform a complete analysis, the agency wants to find the variable that mainly
53: aov(TOTCHG ~ .,data=hospitalcost)
54: mod <- lm(TOTCHG ~ .,data=hospitalcost)
55: summary(mod)
56

```

(Top Level) R Script

Console Terminal Background Jobs

```

R 4.4.0 - /~/
RACE3   330.5184 2629.5121 0.126 0.900
RACE4   -499.3818 1520.9293 -0.328 0.743
RACE5   -1784.5776 1532.0048 -1.165 0.245
RACE6   -594.2921 1859.1271 -0.320 0.749
APRDRG  -7.8175 0.6881 -11.361 < 2e-16 ***

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2622 on 489 degrees of freedom
Multiple R-squared: 0.5544, Adjusted R-squared: 0.5462
F-statistic: 67.6 on 9 and 489 DF, p-value: < 2.2e-16

Plots

Environment History Connections Tutorial

Global Environment

- b List of 12
- cat List of 13
- d 63 obs. of 2 variables
- fit List of 14
- fit1 List of 14
- hospitalcost 499 obs. of 6 variables
- mod List of 13

Files Plots Packages Help Viewer Presentation

Histogram of hospitalcost\$AGE

Frequency

0 50 100 150 200 250

0 5 10 15

hospitalcost\$AGE

Windows Taskbar

Type here to search

25°C Sunny 12:36 ENG US 20/05/2024

