

File Edit Code View Plots Session Build Debug Profile Tools Help

Source Visual B I <> Normal Format Insert Table

Run Publish

Outline

Untitled1* conoc_actifactor_cancer_cervical

color = "darkred", size = 1.5)

Q[38;5;232m`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

50

40

30

20

10

0

Edad a la que intimaron por primera vez

Environment History Connections Tutorial

35 MiB

Data

conoc_actifacto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

En nuestros datos, el promedio de edad a la que intimaron por primera vez es:

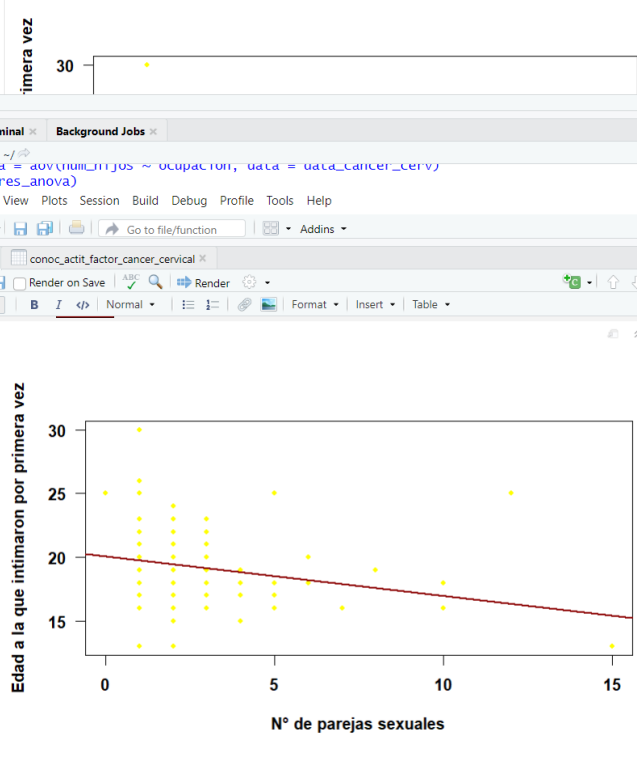
```
mean(data_cancer_cerv$edad_relacion_sexual, na.rm = TRUE)
```

```
[1] 19.32569
```

1.2.- Notación en el método de regresión lineal simple:

```
plot(edad_relacion_sexual ~ parejas_sex, data = data_cancer_cerv,
     col = "yellow",
     ylab = "Edad a la que intimaron por primera vez",
     xlab = "N° de parejas sexuales",
     las = 1,
     pch = 20,
     font.lab = 2, font.axis = 2)

# La función lm() ajusta el modelo de regresión lineal
abline(lm(edad_relacion_sexual ~ parejas_sex, data = data_cancer_cerv), lwd = 2,
       col = "darkred")
```



```
res_anova = aov(nuclimujos ~ ocupacion, data = data_cancer_cerv)
summary(res_anova)
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Source Visual B I <> Normal Format Insert Table

Run Publish

Outline

Untitled1* conoc_actifactor_cancer_cervical

color = "darkred", size = 1.5)

Q[38;5;232m`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

50

40

30

20

10

0

Edad a la que intimaron por primera vez

Environment History Connections Tutorial

33 MiB

Data

conoc_actifacto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

1.3.- Ajustando el modelo de regresión lineal simple para nuestro problema:

- Ajustamos el modelo para nuestros datos:

File Edit Code View Plots Session Build Debug Profile Tools Help

Source Visual B I <> Normal Format Insert Table

Run Publish

Outline

Untitled1* conoc_actifactor_cancer_cervical

color = "darkred", size = 1.5)

Q[38;5;232m`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

50

40

30

20

10

0

Edad a la que intimaron por primera vez

Environment History Connections Tutorial

34 MiB

Data

conoc_actifacto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/functionAddins

conoc_actit_factor_cancer_cervical

Render on SaveRenderABCRenderPublishRunPublish

SourceVisualB I <> NormalFormatInsertTableOutline

1.3.- Ajustando el modelo de regresion lineal simple para nuestro problema:

- Ajustamos el modelo para nuestros datos:

```
{r}
modelo_ejemplo = lm(edad_relacion_sexual ~ parejas_sex, data = data_cancer_cerv)
```

```
{r}
summary(modelo_ejemplo)
```

Call:
lm(formula = edad_relacion_sexual ~ parejas_sex, data = data_cancer_cerv)

Residuals:

Min	1Q	Median	3Q	Max
-6.7934	-1.4835	-0.3637	0.8264	10.2066

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	20.10330	0.27422	73.310	< 2e-16 ***
parejas_sex	-0.30991	0.09056	-3.422	0.000743 ***

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

Residual standard error: 2.267 on 216 degrees of freedom
Multiple R-squared: 0.05143, Adjusted R-squared: 0.04704
F-statistic: 11.71 on 1 and 216 DF, p-value: 0.0007428

ConsoleTerminalBackground Jobs

R - R 4.4.3 - ~/>

Project (None)

EnvironmentHistoryConnectionsTutorial

R - Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

New FolderNew FileDeleteRenameMore

Homeestadistica_upsjb_estudioEstadistica> data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_arthritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_arthritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM
s05_pac_cancer_mama.csv	93.4 KB	May 3, 2025, 7:47 AM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/functionAddins

conoc_actit_factor_cancer_cervical

Render on SaveRenderABCRenderPublishRunPublish

SourceVisualB I <> NormalFormatInsertTableOutline

1.4.- Interpretando los resultados:

La sección Coefficients del resultado:

```
{r}
summary(modelo_ejemplo)$coef
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	20.1033001	0.27422137	73.310480	1.379237e-154
parejas_sex	-0.3099075	0.09055723	-3.422228	7.428247e-04

ConsoleTerminalBackground Jobs

R - R 4.4.3 - ~/>

Project (None)

EnvironmentHistoryConnectionsTutorial

R - Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

New FolderNew FileDeleteRenameMore

Homeestadistica_upsjb_estudioEstadistica> data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_arthritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_arthritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM
res_anova	List of 13	
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)	

Chunk 27

Quarto

ConsoleTerminalBackground Jobs

R - R 4.4.3 - ~/>

Project (None)

EnvironmentHistoryConnectionsTutorial

R - Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

New FolderNew FileDeleteRenameMore

Homeestadistica_upsjb_estudioEstadistica> data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_arthritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_arthritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM
res_anova	List of 13	
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)	

theme_gtsummary_language("es")

tabla_reporte <- modelo_ejemplo |>

tbl_regression(intercept = T,

estimate_fun = function(x) style_sigfig(x, digits = 4),

pvalue_fun = function(x) style_pvalue(x, digits = 3),

label = list(parejas_sex ~ "N° de parejas sexuales")) |>

modify_caption("Regresión entorno a la edad a la que intimaron por primera vez en

función de las parejas sexuales que cuentan")

tabla_reporte

Regresión entorno a la edad a la que intimaron por primera vez en función de las parejas sexuales que cuentan

Característica	Beta	95% CI	p-valor
(Intercept)	20.10	19.56, 20.64	<0.001
N° de parejas sexuales	-0.3099	-0.4884, -0.1314	<0.001

Abreviación: CI = Intervalo de confianza

FileEditCodeViewPlotsSessionBuildDebugProfileToolsHelp

Go to file/functionAddins

conoc_actit_factor_cancer_cervical

Render on SaveABCRenderPublish

SourceVisualBIBI<>NormalFormatInsertTableOutline

group_by(data_cancer_cerv, procedencia) |> summarise(count = n(), mean = mean(num_hijos, na.rm = TRUE), sd = sd(num_hijos, na.rm = TRUE))

A tibble: 2 x 4

procedencia<chr>	count<int>	mean<dbl>	sd<dbl>
rural	107	2.121495	1.378475
urbano	111	2.000000	1.272078

2 rows

2.2.- Visualizacion:

data_cancer_cerv |> filter(!is.na(procedencia)& !is.na(num_hijos)) |> ggplot(aes(v = num_hijos, x = procedencia)) + geom_boxplot()

Chunk 27

R 4.4.3

EnvironmentHistoryConnectionsTutorial

Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_artritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_artritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diahetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM

conoc_actit_factor_cancer_cervical

Render on SaveABCRenderPublish

SourceVisualBIBI<>NormalFormatInsertTableOutline

2.2.- Visualizacion:

data_cancer_cerv |> filter(!is.na(procedencia)& !is.na(num_hijos)) |> ggplot(aes(y = num_hijos, x = procedencia)) + geom_boxplot() + labs(y = "N° de hijos con los que cuentan", x = "Procedencia")

Chunk 27

R 4.4.3

EnvironmentHistoryConnectionsTutorial

Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_artritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM

conoc_actit_factor_cancer_cervical

Render on SaveABCRenderPublish

SourceVisualBIBI<>NormalFormatInsertTableOutline

2.3.- Pruebas preliminares para evaluar los supuestos de la prueba "t" de Student:

Supuesto 1: Los datos debieron haber sido muestreados de una distribución normal.

Empleo de la prueba de Shapiro-wilk:

data_cancer_cerv |> filter(procedencia == "rural") |> summarise(shapiro = list(shapiro.test(num_hijos))) |> pull(shapiro)

[[1]]

Shapiro-wilk normality test

data: num_hijos

W = 0.93705, p-value = 7.363e-05

data_cancer_cerv |> filter(procedencia == "urbano") |> summarise(shapiro = list(shapiro.test(num_hijos))) |> pull(shapiro)

[[1]]

Shapiro-wilk normality test

data: num_hijos

W = 0.9261, p-value = 1.169e-05

Chunk 27

R 4.4.3

EnvironmentHistoryConnectionsTutorial

Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_artritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_artritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM

conoc_actit_factor_cancer_cervical

Shapiro-wilk normality test

data: num_hijos
W = 0.9261, p-value = 1.169e-05

Usando la prueba F para evaluar la homogeneidad de varianzas. Función var.test():

{r}
var.test(num_hijos ~ procedencia, data = data_cancer_cerv)

F test to compare two variances

data: num_hijos by procedencia
F = 1.1743, num df = 106, denom df = 110, p-value = 0.4044
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.8041794 1.7170138
sample estimates:
ratio of variances
1.174277

El valor p de la prueba F es p = 0.4044. Siendo mayor que el nivel de significancia $\alpha = 0.05$. En conclusión, no hay una diferencia significativa entre las varianzas de los dos conjuntos (urbano y rural) de datos. Por lo tanto, podemos usar la prueba t clásica que asume igualdad de varianzas. Por lo tanto, se concluye que las varianzas de "num_hijos" entre los grupos de "procedencia" no difieren significativamente, al nivel de significancia del 5%.

EnvironmentHistoryConnectionsTutorial
RGlobal Environment
Data
conoc_actit_facto...218 obs. of 18 variables
data_cancer_cerv218 obs. of 18 variables
modelo_ejemploList of 12
res_anovaList of 13
tabla_reporteLarge tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation
New FolderNew FileDeleteRenameMore
Homeestadistica_upsjb_estudioEstadistica > data
NameSizeModified
--practica01.Rproj286 BApr 15, 2025, 8:35 PM
.Rhistory112 BApr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv24.6 KBApr 17, 2025, 9:37 PM
s02_treat_artritis.csv41.7 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata33.6 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.sav12.9 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx38.6 KBApr 12, 2025, 8:34 AM
s03_treat_artritis.csv29.9 KBApr 19, 2025, 8:00 AM
s04_diabetes.csv38.7 KBApr 26, 2025, 8:06 AM

conoc_actit_factor_cancer_cervical

Realizamos la prueba "t" para nuestros datos:

{r}
t.test(num_hijos ~ procedencia, data = data_cancer_cerv, var.equal = TRUE)

Two Sample t-test

data: num_hijos by procedencia
t = 0.67663, df = 216, p-value = 0.4994
alternative hypothesis: true difference in means between group rural and group urbano is not equal to 0
95 percent confidence interval:
-0.2324178 0.4754085
sample estimates:
mean in group rural mean in group urbano
2.121495 2.000000

Interpretación:
El valor p de 0.4994 indica que no hay evidencia significativa para afirmar una diferencia en el número promedio de hijos entre los grupos rural y urbano. La diferencia observada no es estadísticamente significativa, y el intervalo de confianza incluye el cero, lo que nos da una conclusión de igualdad de medias, osea, no existen diferencias significativas en el número promedio de hijos entre personas de zonas rurales y urbanas, según los datos analizados en esta prueba t.

3.- ANALISIS DE VARIANZA "NOVA":

3.1.- Resumen y visualización:

Resumen:

EnvironmentHistoryConnectionsTutorial
RGlobal Environment
Data
conoc_actit_facto...218 obs. of 18 variables
data_cancer_cerv218 obs. of 18 variables
modelo_ejemploList of 12
res_anovaList of 13
tabla_reporteLarge tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation
New FolderNew FileDeleteRenameMore
Homeestadistica_upsjb_estudioEstadistica > data
NameSizeModified
--practica01.Rproj286 BApr 15, 2025, 8:35 PM
.Rhistory112 BApr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv24.6 KBApr 17, 2025, 9:37 PM
s02_treat_artritis.csv41.7 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata33.6 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.sav12.9 KBApr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx38.6 KBApr 12, 2025, 8:34 AM
s03_treat_artritis.csv29.9 KBApr 19, 2025, 8:00 AM

Source Visual B I <> Normal Format Insert Table Outline

3.- ANALISIS DE VARIANZA "NOVA":

3.1.- Resumen y visualización:

Resumen:

{r}
group_by(data_cancer_cerv, ocupacion) |>
summarise(
 count = n(),
 mean = mean(num_hijos, na.rm = TRUE),
 sd = sd(num_hijos, na.rm = TRUE),
 min = min(num_hijos, na.rm = TRUE),
 max = max(num_hijos, na.rm = TRUE)
)

A tibble: 5 x 6

ocupacion	count	mean	sd	min	max
ama de casa	37	2.6216216	1.4785757	0	6
empleada	130	2.1923077	0.9968645	0	4
estudiante	29	0.6206897	1.2075297	0	5
otro	19	2.4210526	1.6771599	0	6
sin empleo	3	1.0000000	1.0000000	0	2

5 rows

EnvironmentHistoryConnectionsTutorial
RGlobal Environment
Data
conoc_actit_facto...218 obs. of 18 variables
data_cancer_cerv218 obs. of 18 variables
modelo_ejemploList of 12
res_anovaList of 13
tabla_reporteLarge tbl_regression (5 elements, 569.9 kB)

FilesPlotsPackagesHelpViewerPresentation
New FolderNew FileDeleteRenameMore
Homeestadistica_upsjb_estudioEstadistica > data
NameSizeModified
--practica01.Rproj286 BApr 15, 2025, 8:35 PM
.Rhistory112 BApr 15, 2025, 8:36 PM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render Render Publish

Source Visual B I <> Normal Format Insert Table Outline

Visualización:

```
{r}
data_cancer_cerv <- data_cancer_cerv |>
  filter(!is.na(ocupacion) & !is.na(num_hijos)) |>
  ggplot(aes(y = num_hijos, x = ocupacion)) +
  geom_boxplot() +
  labs(y = "N° de hijos", x = "Ocupacion o a que se dedica")
```

Chunk 27

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_factor_cervical	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_artritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_artritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render Render Publish

Source Visual B I <> Normal Format Insert Table Outline

3.2.- Pruebas preliminares para evaluar los supuestos del ANOVA:

```
{r}
data_cancer_cerv <- data_cancer_cerv |>
  mutate(ocupacion = as.factor(ocupacion))
```

Supuesto 1: los datos deben haber sido muestreados de una distribución normal.

Para esto, usamos la prueba de Shapiro-wilk.

```
{r}
data_cancer_cerv |>
  filter(ocupacion == "estudiante") |>
  summarise(shapiro = list(shapiro.test(num_hijos))) |>
  pull(shapiro)
```

```
[[1]]

      shapiro-wilk normality test

data:  num_hijos
W = 0.5908, p-value = 7.969e-08
```

```
{r}
data_cancer_cerv |>
  filter(ocupacion == "ama de casa") |>
  summarise(shapiro = list(shapiro.test(num_hijos))) |>
  pull(shapiro)
```

```
[[1]]

      shapiro-wilk normality test

data:  num_hijos
W = 0.92458, p-value = 0.01528
```

```
{r}
data_cancer_cerv |>
  filter(ocupacion == "empleada") |>
  summarise(shapiro = list(shapiro.test(num_hijos))) |>
  pull(shapiro)
```

```
[[1]]

      shapiro-wilk normality test

data:  num_hijos
W = 0.89912, p-value = 6.944e-08
```

Chunk 27

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_factor_cervical	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_artritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_artritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_artritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render Render Publish

Source Visual B I <> Normal Format Insert Table Outline

filter(ocupacion == "ama de casa") |>
summarise(shapiro = list(shapiro.test(num_hijos))) |>
pull(shapiro)

```
[[1]]

      shapiro-wilk normality test

data:  num_hijos
W = 0.92458, p-value = 0.01528
```

filter(ocupacion == "empleada") |>
summarise(shapiro = list(shapiro.test(num_hijos))) |>
pull(shapiro)

```
[[1]]

      shapiro-wilk normality test

data:  num_hijos
W = 0.89912, p-value = 6.944e-08
```

Chunk 27

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_factor_cervical	218 obs. of 18 variables
data_cancer_cerv	218 obs. of 18 variables
modelo_ejemplo	List of 12
res_anova	List of 13
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render

Source Visual B I <> Normal Format Insert Table Outline

```
data: num_hijos
W = 0.89912, p-value = 6.944e-08

{r}
data_cancer_cerv |>
  filter(ocupacion == "otro") |>
  summarise(shapiro = list(shapiro.test(num_hijos))) |>
  pull(shapiro)

[[1]]

      Shapiro-Wilk normality test

data: num_hijos
W = 0.92094, p-value = 0.1178

{r}
data_cancer_cerv |>
  filter(ocupacion == "sin empleo") |>
  summarise(shapiro = list(shapiro.test(num_hijos))) |>
  pull(shapiro)

[[1]]

      Shapiro-Wilk normality test

data: num_hijos
W = 1, p-value = 1
```

Chunk 27

Quarto

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables	
data_cancer_cerv	218 obs. of 18 variables	
modelo_ejemplo	List of 12	
res_anova	List of 13	
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)	

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_arthritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_arthritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render

Source Visual B I <> Normal Format Insert Table Outline

Supuesto 2: Las varianzas de los dos grupos son iguales Para esto podemos usar la prueba de Levene para evaluar la homogeneidad de varianzas. Esto esta implementado en la función var.test():

```
{r}
LeveneTest(num_hijos ~ ocupacion, data = data_cancer_cerv)

Levene's Test for Homogeneity of Variance (center = median)
Df F value Pr(>F)
group 4 3.9167 0.004327 **
      213
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretacion:

El valor p de 0.0043 en la prueba de Levene indica que hay diferencias significativas en la varianza del número de hijos entre los grupos. Por tanto, no se cumple el supuesto de homogeneidad de varianzas entre los cinco grupos analizados.

3.3.- Realizamos la prueba de ANOVA de una vía para nuestros datos:

```
{r}
res_anova = aov(num_hijos ~ ocupacion, data = data_cancer_cerv)

summary(res_anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ocupacion	4	79.87	19.97	14.16	2.9e-10 ***
Residuals	213	300.35	1.41		

Chunk 27

Quarto

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables	
data_cancer_cerv	218 obs. of 18 variables	
modelo_ejemplo	List of 12	
res_anova	List of 13	
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)	

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM
s02_treat_arthritis.csv	41.7 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.Rdata	33.6 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.sav	12.9 KB	Apr 12, 2025, 8:34 AM
s02_treat_arthritis.xlsx	38.6 KB	Apr 12, 2025, 8:34 AM
s03_treat_arthritis.csv	29.9 KB	Apr 19, 2025, 8:00 AM
s04_diabetes.csv	38.7 KB	Apr 26, 2025, 8:06 AM

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Untitled1* conoc_actit_factor_cancer_cervical

Render on Save Render

Source Visual B I <> Normal Format Insert Table Outline

entre los cinco grupos analizados.

3.3.- Realizamos la prueba de ANOVA de una vía para nuestros datos:

```
{r}
res_anova = aov(num_hijos ~ ocupacion, data = data_cancer_cerv)

summary(res_anova)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
ocupacion	4	79.87	19.97	14.16	2.9e-10 ***
Residuals	213	300.35	1.41		

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

Interpretando datos:

El análisis de varianza (ANOVA) muestra un valor p de 2.9e-10, indicando diferencias altamente significativas en el número de hijos según la ocupación. Esto sugiere que al menos un grupo ocupacional difiere de los demás en promedio de hijos, con una variación explicada por la ocupación considerable.

Se resaltaría que la ocupación tiene un impacto estadísticamente significativo en el número de hijos. La baja probabilidad (p < 0.001) indica que las diferencias observadas entre grupos no son aleatorias, lo que sugiere que **el tipo de ocupación influye en la cantidad de hijos.**

Environment History Connections Tutorial

R Global Environment

Data

conoc_actit_facto...	218 obs. of 18 variables	
data_cancer_cerv	218 obs. of 18 variables	
modelo_ejemplo	List of 12	
res_anova	List of 13	
tabla_reporte	Large tbl_regression (5 elements, 569.9 kB)	

Files Plots Packages Help Viewer Presentation

New Folder New File Delete Rename More

Home > estadistica_upsjb_estudio > Estadistica > data

Name	Size	Modified
..		
--practica01.Rproj	286 B	Apr 15, 2025, 8:35 PM
.Rhistory	112 B	Apr 15, 2025, 8:36 PM
conoc_actit_factor_cancer_cervical.csv	24.6 KB	Apr 17, 2025, 9:37 PM