- Arrange your CSV into Excel file format
- Convert the 2nd column title with title of the LAPTOP Name
  - Actaul title entry is:"DELL Vostro Core i3 11th Gen - (8 GB/1 TB HDD/256 GB SSD/Windows 11 Home) Vostro 3510 Thin and Light L..." will be converted to DELL
  - Vostro Core i3 11th Gen

Source on Save

```r
# Load the readxl package
library(readxl)


# Use file.choose() to open a dialog box and select the Excel file
file_path <- file.choose()

# Read the Excel file into a data frame
df <- read_excel(file_path)

# Print the first few rows of the data frame to check if the file was read correctly
head(df)



# Get the column you want to modify
column_to_modify <- df$title

# Define a regular expression to match everything after "-"
regex <- "-.*$"

# Apply the regular expression to each value in the column using sub() and store the result in modified_column
modified_column <- sub(regex, "", column_to_modify)

# Replace the original column in the data frame with the modified column
df$title <- modified_column
head(df)

```

```
> regex <- "-.*$"
> # Apply the regular expression to each value in the column using sub() and store the result in modified_column
> modified_column <- sub(regex, "", column_to_modify)
> # Replace the original column in the data frame with the modified column
> df$title <- modified_column
> head(df)
# A tibble: 6 × 11
   ...1 title                                                  price     discount          Processor                                    RAM   OS    SSD   Display In_bu…¹ warra…²
  <dbl> <chr>                                                  <chr>     <chr>             <chr>                                        <chr> <chr> <chr> <chr>   <chr>   <chr>
1     0 "DELL Vostro Core i3 11th Gen "                        â,'37,990 â,'58,48935% off Processor: Intel i3-1115G4 (Base- 1.7 GHz… RAM … Grap… "Dis… Intel … NA      8 GB D…
2     1 "HP 14s Intel Core i3 11th Gen "                       â,'35,490 â,'47,20624% off Intel Core i3 Processor (11th Gen)           8 GB… 64 b… "256… 35.56 … NA      1 Year…
3     2 "Lenovo V15 G2 Core i3 11th Gen "                      â,'33,999 â,'59,76043% off Intel Core i3 Processor (11th Gen)           8 GB… 64 b… "512… 39.62 … No      1 Year…
4     3 "HP 15s Intel Core i3 12th Gen "                       â,'45,490 â,'56,26019% off Intel Core i3 Processor (12th Gen)           8 GB… 64 b… "512… 39.62 … Micros… 1 Year…
5     4 "ASUS VivoBook 15 (2022) Core i3 10th Gen "            â,'33,990 â,'45,99026% off Intel Core i3 Processor (10th Gen)           8 GB… 64 b… "512… 39.62 … Office… 1 Year…
6     5 "ASUS TUF Gaming A17 with 90Whr Battery Ryzen 7 Octa Core AMD R7" â,'67,990 â,'94,99028% off AMD Ryzen 7 Octa Core Processor    8 GB… 64 b… "512… 43.94 … NA      1 Year…
# … with abbreviated variable names ¹In_build_sw, ²warranty
> 
```

```r
library(readxl)


# Use file.choose() to open a dialog box and select the Excel file
file_path <- file.choose()

# Read the Excel file into a data frame
df <- read_excel(file_path)

# Print the first few rows of the data frame to check if the file was read correctly
#head(df)
column_name <- "price"

# Remove the special characters from the column
df[[column_name]] <- gsub("[^[:digit:]]", "", df[[column_name]])

# Print the modified data frame
print(df)

column_name <- "discount"

# Remove the special characters from the column and replace commas with decimal points
df[[column_name]] <- gsub(",", ".", gsub("[^[:digit:],]", "", df[[column_name]]))

# Print the modified data frame
print(df)
```

	price	discount	Processor	RAM	OS	SSD	Display	I
	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<
d Light L...	37990	58.48935	Processor: I…	RAM …	Grap…	"Dis…	Intel …	N
ht Laptop	35490	47.20624	Intel Core i…	8 GB…	64 b…	"256…	35.56 …	N
	33999	59.76043	Intel Core i…	8 GB…	64 b…	"512…	39.62 …	N
Laptop	45490	56.26019	Intel Core i…	8 GB…	64 b…	"512…	39.62 …	M
X515JA-E...	33990	45.99026	Intel Core i…	8 GB…	64 b…	"512…	39.62 …	O
ows 11 Ho...	67990	94.99028	AMD Ryzen 7 …	8 GB…	64 b…	"512…	43.94 …	N
ows 11 Ho...	52990	71.99026	AMD Ryzen 5 …	8 GB…	64 b…	"512…	43.94 …	N
ows 11 Ho...	47990	78.99039	AMD Ryzen 5 …	8 GB…	64 b…	"1 T…	39.62 …	O
Laptop	49123	59.10816	AMD Ryzen 5 …	16 G…	64 b…	"512…	39.62 …	M
Laptop	57499	70.23518	Intel Core i…	8 GB…	64 b…	"512…	39.62 …	N

```r
# Load the readxl package
library(readxl)


# Use file.choose() to open a dialog box and select the Excel file
file_path <- file.choose()

# Read the Excel file into a data frame
df <- read_excel(file_path)


column_name <- "warranty"

# Specify the word you want to exclude
word <- "warranty"

# Find the rows that do not contain the word in the specified column
rows_without_word <- !grepl(word, df[[column_name]], ignore.case = TRUE)

# Print the rows without the word
if (sum(rows_without_word) > 0) {
  print(df[rows_without_word, ])
} else {
  print(paste0("All rows contain '", word, "' in column '", column_name, "'"))
}









column_name <- "warranty"

# Specify the number you want to find
number_to_find <- 2

# Find the rows that contain the number in the specified column
rows_with_number <- grep(paste0("\\b", number_to_find, "\\b"), df[[column_name]])

# Print the rows with the number
if (length(rows_with_number) > 0) {
  print(df[rows_with_number, ])
  cat("There are", length(rows_with_number), "rows with the number", number_to_find)
} else {
  print(paste0("The number", number_to_find, "is not in column", column_name))
}
```

```
+ }
# A tibble: 63 × 11
   ...1 title                                                                                price      discount   Proce…¹ RAM   OS    SSD   Display In_bu…² warra…³
  <db1> <chr>                                                                              <chr>      <chr>      <chr>   <chr> <chr> <chr> <chr>   <chr>   <chr>
1    13 Lenovo Intel Celeron Dual Core - (8 GB/256 GB SSD/Windows 11 Home) 81WQ00MQIN|81WQ00NXIN Laptop          â‚¹27,799 â‚¹40,490… Intel … 8 GB… 64 b… 256 … 39.62 … NA      2 Year…
2    36 Lenovo Intel Core i3 11th Gen - (8 GB/512 GB SSD/Windows 11 Home) 81X800J3IN | 81X800LGIN Thin and Lig... â‚¹36,900 â‚¹59,890… Intel … 8 GB… 64 b… 512 … 39.62 … Office… 2 Year…
3    46 MSI Katana GF66 Core i7 11th Gen - (16 GB/512 GB SSD/Windows 10 Home/6 GB Graphics/NVIDIA GeForce RTX ... â‚¹85,990 â‚¹1,23,9… Intel … 16 G… 64 b… 512 … 39.62 … NA      2 Year…
4    52 Lenovo IdeaPad 3 Core i3 11th Gen - (8 GB/256 GB SSD/Windows 11 Home) 14ITL05 Thin and Light Laptop      â‚¹36,490 â‚¹60,890… Intel … 8 GB… 64 b… 256 … 35.56 … Office… 2 Year…
5    61 Lenovo Intel Core i3 11th Gen - (8 GB/256 GB SSD/Windows 11 Home) 15ITL05 Thin and Light Laptop         â‚¹33,990 â‚¹48,490… Intel … 8 GB… 64 b… 256 … 39.62 … NA      2 Year…
6    63 MSI Katana GF66 Core i5 12th Gen - (16 GB/512 GB SSD/Windows 11 Home/4 GB Graphics/NVIDIA GeForce RTX ... â‚¹76,990 â‚¹1,11,9… Intel … 16 G… 64 b… 512 … 39.62 … NA      2 Year…
7    70 Lenovo IdeaPad 3 Core i3 11th Gen - (8 GB/512 GB SSD/Windows 11 Home) 82H801L7IN | 82H802FJIN | 82H802... â‚¹41,490 â‚¹59,390… Intel … 8 GB… 64 b… 512 … 39.62 … Office… 2 Year…
8    72 Lenovo Intel Core i3 10th Gen - (8 GB/256 GB SSD/Windows 11 Home) 81WA00Q3IN|81WA00MEIN Thin and Light... â‚¹33,990 â‚¹50,290… Intel … 8 GB… 64 b… 256 … 35.56 … NA      2 Year…
9    81 Lenovo IdeaPad 3 Core i5 12th Gen - (8 GB/512 GB SSD/Windows 11 Home) 15IAU7 Thin and Light Laptop      â‚¹58,490 â‚¹76,890… Intel … 8 GB… 64 b… 512 … 39.62 … NA      2 Year…
10   91 MSI Creator M16 Core i7 12th Gen - (16 GB/512 GB SSD/Windows 11 Home/4 GB Graphics/NVIDIA GeForce RTX ... â‚¹99,990 â‚¹1,28,9… Intel … 16 G… 64 b… 512 … 40.64 … Window… 2 Year…
# … with 53 more rows, and abbreviated variable names ¹Processor, ²In_build_sw, ³warranty
# ℹ Use `print(n = ...)` to see more rows
There are 63 rows with the number 2
> |
```


```
# … with 53 more rows, and abbreviated variable na
# ℹ Use `print(n = ...)` to see more rows
There are 63 rows with the number 2
> |
```