

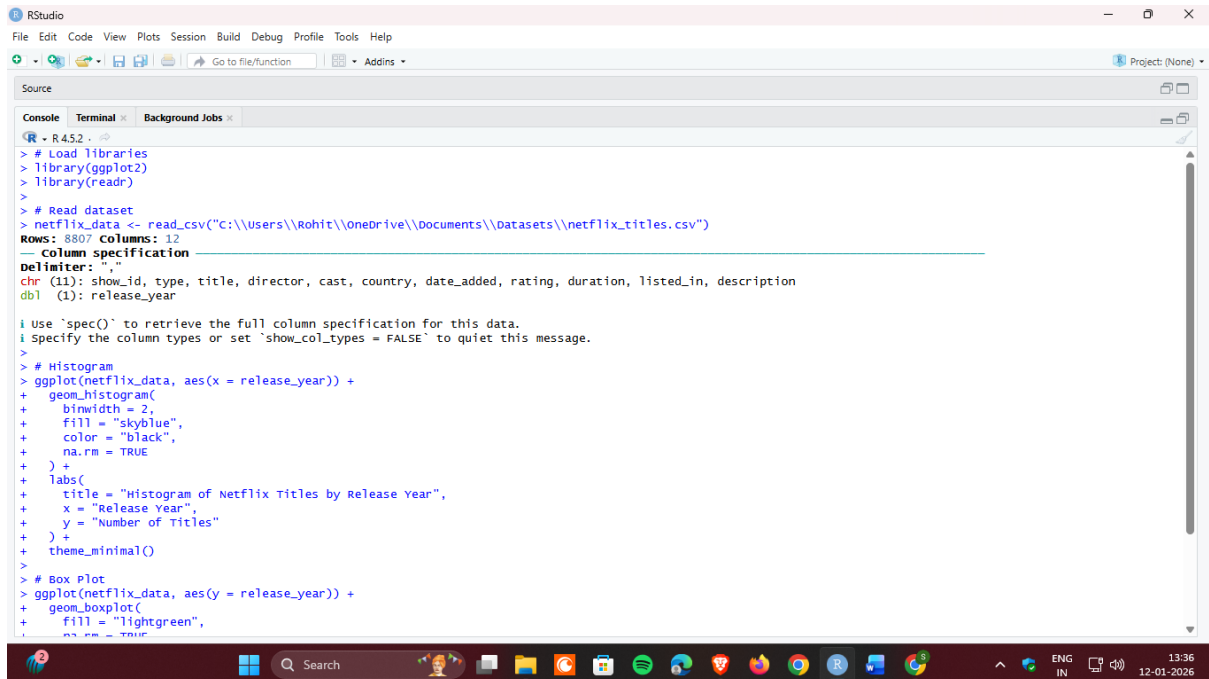
SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- DATA ANALYSIS WITH SAS/SPSS/R

PRACTICAL – 11

AIM:- Generating histograms and box plots using ggplot2 (R).

OUTPUT:-

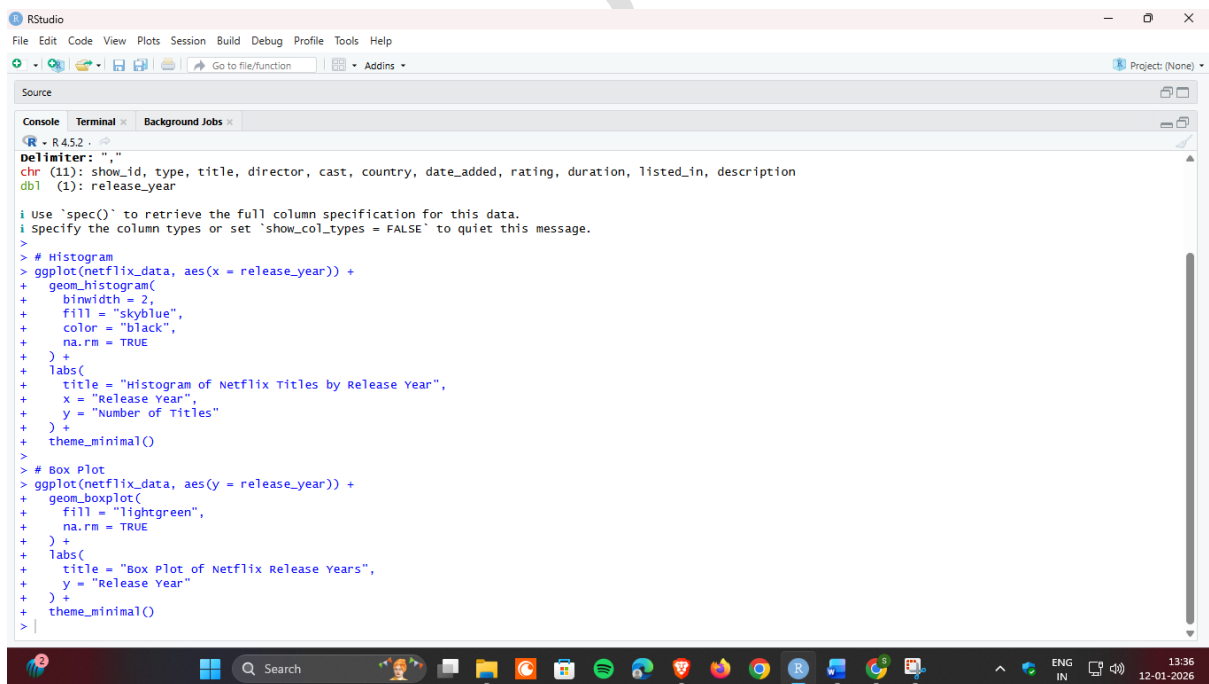


```
R - R4.5.2
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)

Source
Console Terminal Background Jobs

> # Load libraries
> library(ggplot2)
> library(readr)
>
> # Read dataset
> netflix_data <- read_csv("C:\\Users\\Rohit\\OneDrive\\Documents\\datasets\\netflix_titles.csv")
Rows: 8807 Columns: 12
Column specification
-----
Delimiter: ","
chr (11): show_id, type, title, director, cast, country, date_added, rating, duration, listed_in, description
dbl (1): release_year

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
>
> # Histogram
> ggplot(netflix_data, aes(x = release_year)) +
+   geom_histogram(
+     binwidth = 2,
+     fill = "skyblue",
+     color = "black",
+     na.rm = TRUE
+   ) +
+   labs(
+     title = "Histogram of Netflix Titles by Release Year",
+     x = "Release Year",
+     y = "Number of Titles"
+   ) +
+   theme_minimal()
>
> # Box Plot
> ggplot(netflix_data, aes(y = release_year)) +
+   geom_boxplot(
+     fill = "lightgreen",
+     na.rm = TRUE
+   ) +
+   labs(
+     title = "Box Plot of Netflix Release Years",
+     y = "Release Year"
+   ) +
+   theme_minimal()
>
```



```
R - R4.5.2
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)

Source
Console Terminal Background Jobs

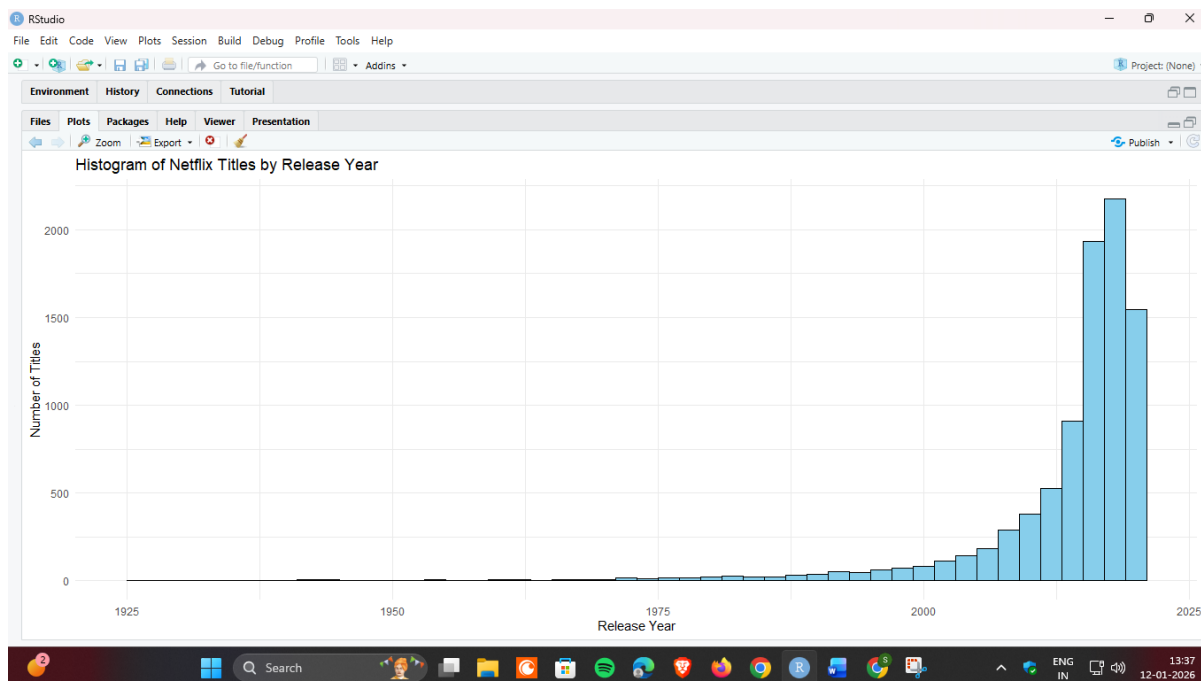
Delimiter: ","
chr (11): show_id, type, title, director, cast, country, date_added, rating, duration, listed_in, description
dbl (1): release_year

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
>
> # Histogram
> ggplot(netflix_data, aes(x = release_year)) +
+   geom_histogram(
+     binwidth = 2,
+     fill = "skyblue",
+     color = "black",
+     na.rm = TRUE
+   ) +
+   labs(
+     title = "Histogram of Netflix Titles by Release Year",
+     x = "Release Year",
+     y = "Number of Titles"
+   ) +
+   theme_minimal()
>
> # Box Plot
> ggplot(netflix_data, aes(y = release_year)) +
+   geom_boxplot(
+     fill = "lightgreen",
+     na.rm = TRUE
+   ) +
+   labs(
+     title = "Box Plot of Netflix Release Years",
+     y = "Release Year"
+   ) +
+   theme_minimal()
>
```

SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT :- DATA ANALYSIS WITH SAS/SPSS/R

#HISTOGRAM



#BOX PLOT

