SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE
SUBJECT - Data Analysis with SAS / SPSS / R

Aim : Selecting and dropping variables using select() in R. import dataset.

```
> library(dplyr)  # select() comes from dplyr
> # ========================================
> # 1. IMPORT THE DATASET
> # ========================================
> movies <- read.csv("tmdb_5000_movies.csv")
> print("--- Original Dataset (First 3 rows) ---")
[1] "--- Original Dataset (First 3 rows) ---"
> print(head(movies, 3))
     budget
1 237000000
2 300000000
3 245000000

genres
1 [{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "nam
e": "Fantasy"}, {"id": 878, "name": "Science Fiction"}]
2                                          [{"id": 12, "name": "Adventure"}, {"i
d": 14, "name": "Fantasy"}, {"id": 28, "name": "Action"}]
3                                           [{"id": 28, "name": "Action"}, {"id":
12, "name": "Adventure"}, {"id": 80, "name": "Crime"}]
                                       homepage     id
1                  http://www.avatarmovie.com/  19995
2 http://disney.go.com/disneypictures/pirates/    285
3  http://www.sonypictures.com/movies/spectre/ 206647

keywords
1 [{"id": 1463, "name": "culture clash"}, {"id": 2964, "name": "future"}, {"id":
3386, "name": "space war"}, {"id": 3388, "name": "space colony"}, {"id": 3679, "n
ame": "society"}, {"id": 3801, "name": "space travel"}, {"id": 9685, "name": "fut
uristic"}, {"id": 9840, "name": "romance"}, {"id": 9882, "name": "space"}, {"id":
9951, "name": "alien"}, {"id": 10148, "name": "tribe"}, {"id": 10158, "name": "al
ien planet"}, {"id": 10987, "name": "cgi"}, {"id": 11399, "name": "marine"}, {"i
d": 13065, "name": "soldier"}, {"id": 14643, "name": "battle"}, {"id": 14720, "na
me": "love affair"}, {"id": 165431, "name": "anti war"}, {"id": 193554, "name":
"power relations"}, {"id": 206690, "name": "mind and soul"}, {"id": 209714, "nam
e": "3d"}]
2
```


Name - Mithil Kadam
Roll No - S083

```
[{"id": 270, "name": "ocean"}, {"id": 726, "name": "drug abuse"}, {"id": 911, "na
me": "exotic island"}, {"id": 1319, "name": "east india trading company"}, {"id":
2038, "name": "love of one's life"}, {"id": 2052, "name": "traitor"}, {"id": 258
0, "name": "shipwreck"}, {"id": 2660, "name": "strong woman"}, {"id": 3799, "nam
e": "ship"}, {"id": 5740, "name": "alliance"}, {"id": 5941, "name": "calypso"},
{"id": 6155, "name": "afterlife"}, {"id": 6211, "name": "fighter"}, {"id": 12988,
"name": "pirate"}, {"id": 157186, "name": "swashbuckler"}, {"id": 179430, "name":
"aftercreditsstinger"}]
3
[{"id": 470, "name": "spy"}, {"id": 818, "name": "based on novel"}, {"id": 4289,
"name": "secret agent"}, {"id": 9663, "name": "sequel"}, {"id": 14555, "name": "m
i6"}, {"id": 156095, "name": "british secret service"}, {"id": 158431, "name": "u
nited kingdom"}]
  original_language                          original_title
1                en                                  Avatar
2                en Pirates of the Caribbean: At World's End
3                en                                 Spectre

overview
1                                                                                In the 22nd ce
ntury, a paraplegic Marine is dispatched to the moon Pandora on a unique mission,
but becomes torn between following orders and protecting an alien civilization.
2                                                                              Captain Barboss
a, long believed to be dead, has come back to life and is headed to the edge of t
he Earth with Will Turner and Elizabeth Swann. But nothing is quite as it seems.
3 A cryptic message from Bond’s past sends him on a trail to uncover a sinister o
rganization. While M battles political forces to keep the secret service alive, B
ond peels back the layers of deceit to reveal the terrible truth behind SPECTRE.
  popularity
1   150.4376
2   139.0826
3   107.3768

production_companies
1 [{"name": "Ingenious Film Partners", "id": 289}, {"name": "Twentieth Century Fo
x Film Corporation", "id": 306}, {"name": "Dune Entertainment", "id": 444}, {"nam
e": "Lightstorm Entertainment", "id": 574}]
```


Name - Mithil Kadam
Roll No - S083

```
isney Pictures", "id": 2}, {"name": "Jerry Bruckheimer Films", "id": 130}, {"nam
e": "Second Mate Productions", "id": 19936}]
3
[{"name": "Columbia Pictures", "id": 5}, {"name": "Danjaq", "id": 10761}, {"nam
e": "B24", "id": 69434}]

production_countries
1 [{"iso_3166_1": "US", "name": "United States of America"}, {"iso_3166_1": "GB",
"name": "United Kingdom"}]
2                                                [{"iso_3166_1": "US", "name":
"United States of America"}]
3 [{"iso_3166_1": "GB", "name": "United Kingdom"}, {"iso_3166_1": "US", "name":
"United States of America"}]
  release_date    revenue runtime
1   2009-12-10 2787965087     162
2   2007-05-19  961000000     169
3   2015-10-26  880674609     148

spoken_languages
1
[{"iso_639_1": "en", "name": "English"}, {"iso_639_1": "es", "name": "Espa\\u00f1
ol"}]
2
[{"iso_639_1": "en", "name": "English"}]
3 [{"iso_639_1": "fr", "name": "Fran\\u00e7ais"}, {"iso_639_1": "en", "name": "En
glish"}, {"iso_639_1": "es", "name": "Espa\\u00f1ol"}, {"iso_639_1": "it", "nam
e": "Italiano"}, {"iso_639_1": "de", "name": "Deutsch"}]
    status                                tagline
1 Released                    Enter the World of Pandora.
2 Released At the end of the world, the adventure begins.
3 Released                          A Plan No One Escapes
                                     title vote_average vote_count
1                                   Avatar          7.2      11800
2 Pirates of the Caribbean: At World's End          6.9       4500
3                                  Spectre          6.3       4466
> # Method A: Select specific columns by name
> # Scenario: Keep title, budget, revenue
```


Name - Mithil Kadam
Roll No - S083

```
2 Pirates of the Caribbean: At World's End 300000000  961000000
3                                  Spectre 245000000  880674609
> # Method B: Select a range of adjacent columns
> # Example range: from 'budget' to 'popularity'
> range_cols <- movies %>%
+   select(budget:popularity)
> print("--- Selected Range of Columns (budget to popularity) ---")
[1] "--- Selected Range of Columns (budget to popularity) ---"
> print(head(range_cols, 3))
     budget
1 237000000
2 300000000
3 245000000

genres
1 [{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}, {"id": 14, "nam
e": "Fantasy"}, {"id": 878, "name": "Science Fiction"}]
2                                          [{"id": 12, "name": "Adventure"}, {"i
d": 14, "name": "Fantasy"}, {"id": 28, "name": "Action"}]
3                                            [{"id": 28, "name": "Action"}, {"id":
12, "name": "Adventure"}, {"id": 80, "name": "Crime"}]
                                        homepage     id
1                  http://www.avatarmovie.com/  19995
2 http://disney.go.com/disneypictures/pirates/    285
3  http://www.sonypictures.com/movies/spectre/ 206647

keywords
1 [{"id": 1463, "name": "culture clash"}, {"id": 2964, "name": "future"}, {"id":
3386, "name": "space war"}, {"id": 3388, "name": "space colony"}, {"id": 3679, "n
ame": "society"}, {"id": 3801, "name": "space travel"}, {"id": 9685, "name": "fut
uristic"}, {"id": 9840, "name": "romance"}, {"id": 9882, "name": "space"}, {"id":
9951, "name": "alien"}, {"id": 10148, "name": "tribe"}, {"id": 10158, "name": "al
ien planet"}, {"id": 10987, "name": "cgi"}, {"id": 11399, "name": "marine"}, {"i
d": 13065, "name": "soldier"}, {"id": 14643, "name": "battle"}, {"id": 14720, "na
me": "love affair"}, {"id": 165431, "name": "anti war"}, {"id": 193554, "name":
"power relations"}, {"id": 206690, "name": "mind and soul"}, {"id": 209714, "nam
```


Name - Mithil Kadam
Roll No - S083

```
2
[{"id": 270, "name": "ocean"}, {"id": 726, "name": "drug abuse"}, {"id": 911, "na
me": "exotic island"}, {"id": 1319, "name": "east india trading company"}, {"id":
2038, "name": "love of one's life"}, {"id": 2052, "name": "traitor"}, {"id": 258
0, "name": "shipwreck"}, {"id": 2660, "name": "strong woman"}, {"id": 3799, "nam
e": "ship"}, {"id": 5740, "name": "alliance"}, {"id": 5941, "name": "calypso"},
{"id": 6155, "name": "afterlife"}, {"id": 6211, "name": "fighter"}, {"id": 12988,
"name": "pirate"}, {"id": 157186, "name": "swashbuckler"}, {"id": 179430, "name":
"aftercreditsstinger"}]
3
[{"id": 470, "name": "spy"}, {"id": 818, "name": "based on novel"}, {"id": 4289,
"name": "secret agent"}, {"id": 9663, "name": "sequel"}, {"id": 14555, "name": "m
i6"}, {"id": 156095, "name": "british secret service"}, {"id": 158431, "name": "u
nited kingdom"}]
  original_language                          original_title
1                en                                  Avatar
2                en Pirates of the Caribbean: At World's End
3                en                                 Spectre

overview
1                                                                    In the 22nd ce
ntury, a paraplegic Marine is dispatched to the moon Pandora on a unique mission,
but becomes torn between following orders and protecting an alien civilization.
2                                                                  Captain Barboss
a, long believed to be dead, has come back to life and is headed to the edge of t
he Earth with Will Turner and Elizabeth Swann. But nothing is quite as it seems.
3 A cryptic message from Bond’s past sends him on a trail to uncover a sinister o
rganization. While M battles political forces to keep the secret service alive, B
ond peels back the layers of deceit to reveal the terrible truth behind SPECTRE.
  popularity
1   150.4376
2   139.0826
3   107.3768
> # Method C: Select using helper functions
> # Scenario: Select columns that start with "vote" (vote_count, vote_average)
> starts_with_vote <- movies %>%
+   select(starts_with("vote"))
```


Name - Mithil Kadam
Roll No - S083

```
[1] "--- Selected columns starting with 'vote' ---"
> print(head(starts_with_vote, 3))
  vote_average vote_count
1          7.2      11800
2          6.9       4500
3          6.3       4466
> # Method A: Drop a single specific column
> # Scenario: Remove 'homepage'
> dropped_one <- movies %>%
+   select(-homepage)
> print("--- Dataset with 'homepage' dropped ---")
[1] "--- Dataset with 'homepage' dropped ---"
> print(names(dropped_one))
 [1] "budget"              "genres"              "id"
 [4] "keywords"            "original_language"   "original_title"
 [7] "overview"            "popularity"          "production_companies"
[10] "production_countries" "release_date"        "revenue"
[13] "runtime"             "spoken_languages"    "status"
[16] "tagline"             "title"               "vote_average"
[19] "vote_count"
> # Method B: Drop multiple columns
> # Scenario: Remove 'overview' and 'tagline'
> dropped_multiple <- movies %>%
+   select(-overview, -tagline)
> print("--- Dataset with 'overview' and 'tagline' dropped ---")
[1] "--- Dataset with 'overview' and 'tagline' dropped ---"
> print(names(dropped_multiple))
 [1] "budget"              "genres"              "homepage"
 [4] "id"                  "keywords"            "original_language"
 [7] "original_title"      "popularity"          "production_companies"
[10] "production_countries" "release_date"        "revenue"
[13] "runtime"             "spoken_languages"    "status"
[16] "title"               "vote_average"        "vote_count"
> # Method C: Drop a range of columns
> # Scenario: Remove everything from 'production_companies' to 'popularity'
> # (Adjust depending on your dataset's column order)
```


Name - Mithil Kadam
Roll No - S083