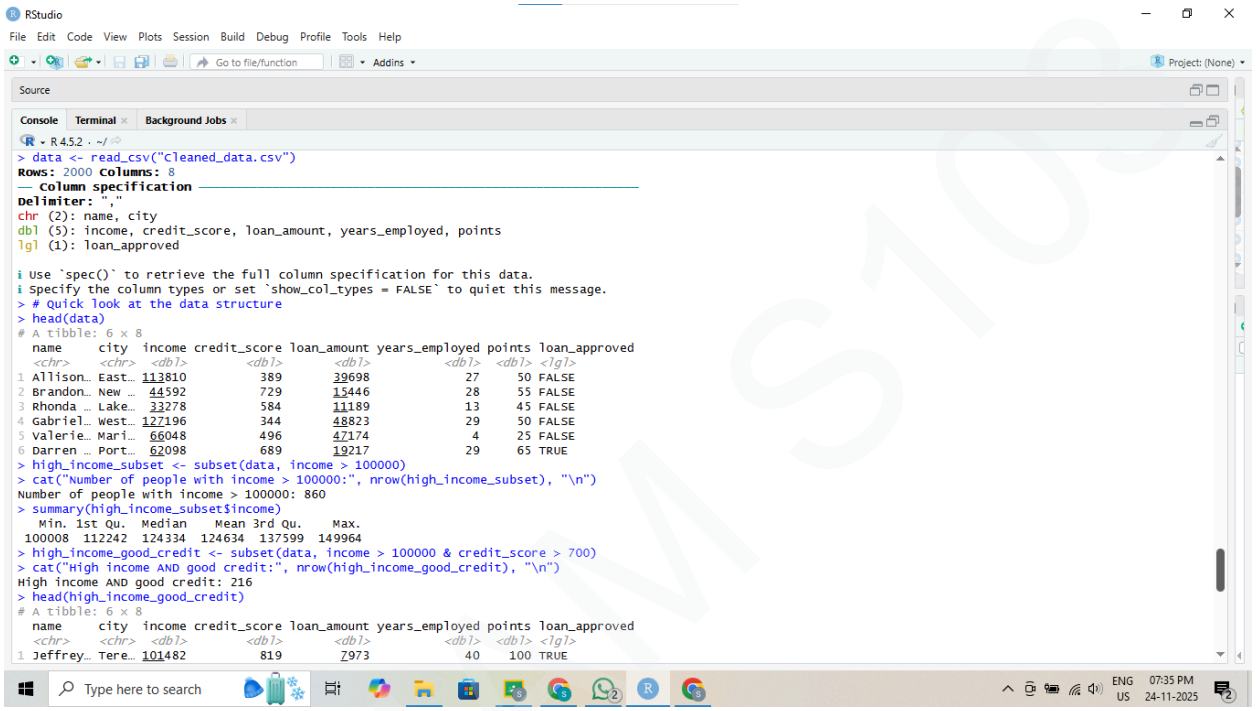
SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT:Data Analysis with R

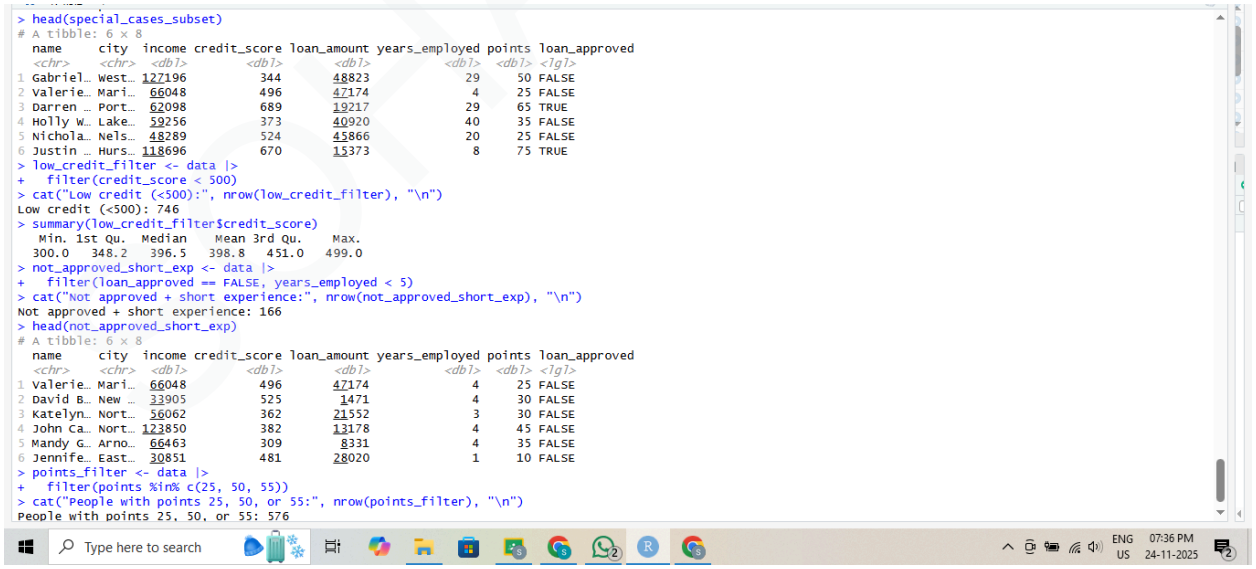
Aim: Applying conditional filters subset() or filter() in R

OUTPUTS



```
> data <- read_csv("Cleaned_data.csv")
Rows: 2000 Columns: 8
── Column specification ────────────────────────────────────────────────────────
Delimiter: ","
chr (2): name, city
dbl (5): income, credit_score, loan_amount, years_employed, points
lgl (1): loan_approved

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.
> # Quick look at the data structure
> head(data)
# A tibble: 6 × 8
  name      city  income credit_score loan_amount years_employed points loan_approved
  <chr>     <chr>  <dbl>        <dbl>       <dbl>          <dbl>  <dbl> <lgl>
1 Allison…  East… 113810          389       39698             27     50 FALSE
2 Brandon…  New …  44592          729       15446             28     55 FALSE
3 Rhonda …  Lake…  33278          584       11189             13     45 FALSE
4 Gabriel…  West… 127196          344       48823             29     50 FALSE
5 Valerie…  Mari…  66048          496       47174              4     25 FALSE
6 Darren …  Port…  62098          689       19217             29     65 TRUE
> high_income_subset <- subset(data, income > 100000)
> cat("Number of people with income > 100000:", nrow(high_income_subset), "\n")
Number of people with income > 100000: 860
> summary(high_income_subset$income)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 100008  112242  124334  124634  137599  149964
> high_income_good_credit <- subset(data, income > 100000 & credit_score > 700)
> cat("High income AND good credit:", nrow(high_income_good_credit), "\n")
High income AND good credit: 216
> head(high_income_good_credit)
# A tibble: 6 × 8
  name      city  income credit_score loan_amount years_employed points loan_approved
  <chr>     <chr>  <dbl>        <dbl>       <dbl>          <dbl>  <dbl> <lgl>
1 Jeffrey…  Tere… 101482          819        7973             40    100 TRUE
```



```
> head(special_cases_subset)
# A tibble: 6 × 8
  name      city  income credit_score loan_amount years_employed points loan_approved
  <chr>     <chr>  <dbl>        <dbl>       <dbl>          <dbl>  <dbl> <lgl>
1 Gabriel…  West… 127196          344       48823             29     50 FALSE
2 Valerie…  Mari…  66048          496       47174              4     25 FALSE
3 Darren …  Port…  62098          689       19217             29     65 TRUE
4 Holly W…  Lake…  59256          373       40920             40     35 FALSE
5 Nichola…  Nels…  48289          524       45866             20     25 FALSE
6 Justin …  Hurs… 118696          670       15373              8     75 TRUE
> low_credit_filter <- data |>
+   filter(credit_score < 500)
> cat("Low credit (<500):", nrow(low_credit_filter), "\n")
Low credit (<500): 746
> summary(low_credit_filter$credit_score)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  300.0   348.2   396.5   398.8   451.0   499.0
> not_approved_short_exp <- data |>
+   filter(loan_approved == FALSE, years_employed < 5)
> cat("Not approved + short experience:", nrow(not_approved_short_exp), "\n")
Not approved + short experience: 166
> head(not_approved_short_exp)
# A tibble: 6 × 8
  name      city  income credit_score loan_amount years_employed points loan_approved
  <chr>     <chr>  <dbl>        <dbl>       <dbl>          <dbl>  <dbl> <lgl>
1 Valerie…  Mari…  66048          496       47174              4     25 FALSE
2 David B…  New …  33905          525        1471              4     30 FALSE
3 Katelyn…  Nort…  56062          362       21552              3     30 FALSE
4 John Ca…  Nort… 123850          382       13178              4     45 FALSE
5 Mandy G…  Arno…  66463          309        8331              4     35 FALSE
6 Jennife…  East…  30851          481       28020              1     10 FALSE
> points_filter <- data |>
+   filter(points %in% c(25, 50, 55))
> cat("People with points 25, 50, or 55:", nrow(points_filter), "\n")
People with points 25, 50, or 55: 576
```

Soham Sycs
S103

SHETH L.U.J. AND SIR M.V. COLLEGE
SUBJECT:Data Analysis with R

	name	city	income	credit_score	loan_amount	years_employed	points	loan_approved	points_num	credit_score_num	years_employed_num	income_num	loan_amou
1	Allison Hill	East Jill	113810	389	39698	27	50	FALSE	50	389	27	113810	
2	Brandon Hall	New Jamesside	44592	729	15446	28	55	FALSE	55	729	28	44592	
3	Rhonda Smith	Lake Roberto	33278	584	11189	13	45	FALSE	45	584	13	33278	
4	Gabrielle Davis	West Melanieview	127196	344	48823	29	50	FALSE	50	344	29	127196	
5	Valerie Gray	Mariastad	66048	496	47174	4	25	FALSE	25	496	4	66048	
6	Darren Roberts	Port Jesseville	62098	689	19217	29	65	TRUE	65	689	29	62098	
7	Holly Wood	Lake Joseph	59256	373	40920	40	35	FALSE	35	373	40	59256	
8	Nicholas Martin	Nelsonside	48289	524	45866	20	25	FALSE	25	524	20	48289	
9	Patty Perez	Port Leslieview	126530	367	14826	36	55	FALSE	55	367	36	126530	
10	Emily Rios	Wilkersonmouth	43434	446	18359	8	20	FALSE	20	446	8	43434	
11	Justin Baker	Hurstfurt	118696	670	15373	8	75	TRUE	75	670	8	118696	
12	Ann Williams	East Courtneychester	127080	365	26216	24	55	FALSE	55	365	24	127080	
13	Julie King	Lake Jenniferside	146939	573	43006	21	50	FALSE	50	573	21	146939	
14	Jeffrey Chavez	Teresaburgh	101482	819	7973	40	100	TRUE	100	819	40	101482	
15	Mark Lynch	West Kathryn	41395	843	1037	38	80	TRUE	80	843	38	41395	
16	Alec Hickman	Johnbury	107397	380	6613	31	60	TRUE	60	380	31	107397	
17	Vanessa Patel	East Donna	85302	469	18370	33	50	FALSE	50	469	33	85302	
18	Jenny Lewis	Lake Larry	34165	682	29711	33	50	FALSE	50	682	33	34165	
19	David Brown	New Angelashire	33905	525	1471	4	30	FALSE	30	525	4	33905	
20	Amy Jones	Port Markhaven	42280	543	19250	31	25	FALSE	25	543	31	42280	
21	Carlos Walls	Sanchezfort	58657	716	13478	6	65	TRUE	65	716	6	58657	

Showing 1 to 21 of 2,000 entries, 14 total columns

S103 4TH.R × data ×

Filter

	income	credit_score	loan_amount	years_employed	points	loan_approved	points_num	credit_score_num	years_employed_num	income_num	loan_amount_num	loan_approved_logical
	113810	389	39698	27	50	FALSE	50	389	27	113810	39698	FALSE
e	44592	729	15446	28	55	FALSE	55	729	28	44592	15446	FALSE
	33278	584	11189	13	45	FALSE	45	584	13	33278	11189	FALSE
iew	127196	344	48823	29	50	FALSE	50	344	29	127196	48823	FALSE
	66048	496	47174	4	25	FALSE	25	496	4	66048	47174	FALSE
	62098	689	19217	29	65	TRUE	65	689	29	62098	19217	TRUE
	59256	373	40920	40	35	FALSE	35	373	40	59256	40920	FALSE
	48289	524	45866	20	25	FALSE	25	524	20	48289	45866	FALSE
	126530	367	14826	36	55	FALSE	55	367	36	126530	14826	FALSE
th	43434	446	18359	8	20	FALSE	20	446	8	43434	18359	FALSE
	118696	670	15373	8	75	TRUE	75	670	8	118696	15373	TRUE
chester	127080	365	26216	24	55	FALSE	55	365	24	127080	26216	FALSE
de	146939	573	43006	21	50	FALSE	50	573	21	146939	43006	FALSE
	101482	819	7973	40	100	TRUE	100	819	40	101482	7973	TRUE
	41395	843	1037	38	80	TRUE	80	843	38	41395	1037	TRUE
	107397	380	6613	31	60	TRUE	60	380	31	107397	6613	TRUE
	85302	469	18370	33	50	FALSE	50	469	33	85302	18370	FALSE
	34165	682	29711	33	50	FALSE	50	682	33	34165	29711	FALSE
ire	33905	525	1471	4	30	FALSE	30	525	4	33905	1471	FALSE
n	42280	543	19250	31	25	FALSE	25	543	31	42280	19250	FALSE
	58657	716	13478	6	65	TRUE	65	716	6	58657	13478	TRUE

Showing 1 to 21 of 2,000 entries, 14 total columns

Console

Type here to search

ENG US 07:34 PM 24-11-2025

Soham Sycs
S103