

**Name:** Alpanshu Kataria

**Reg. No.:** 18BCE1267

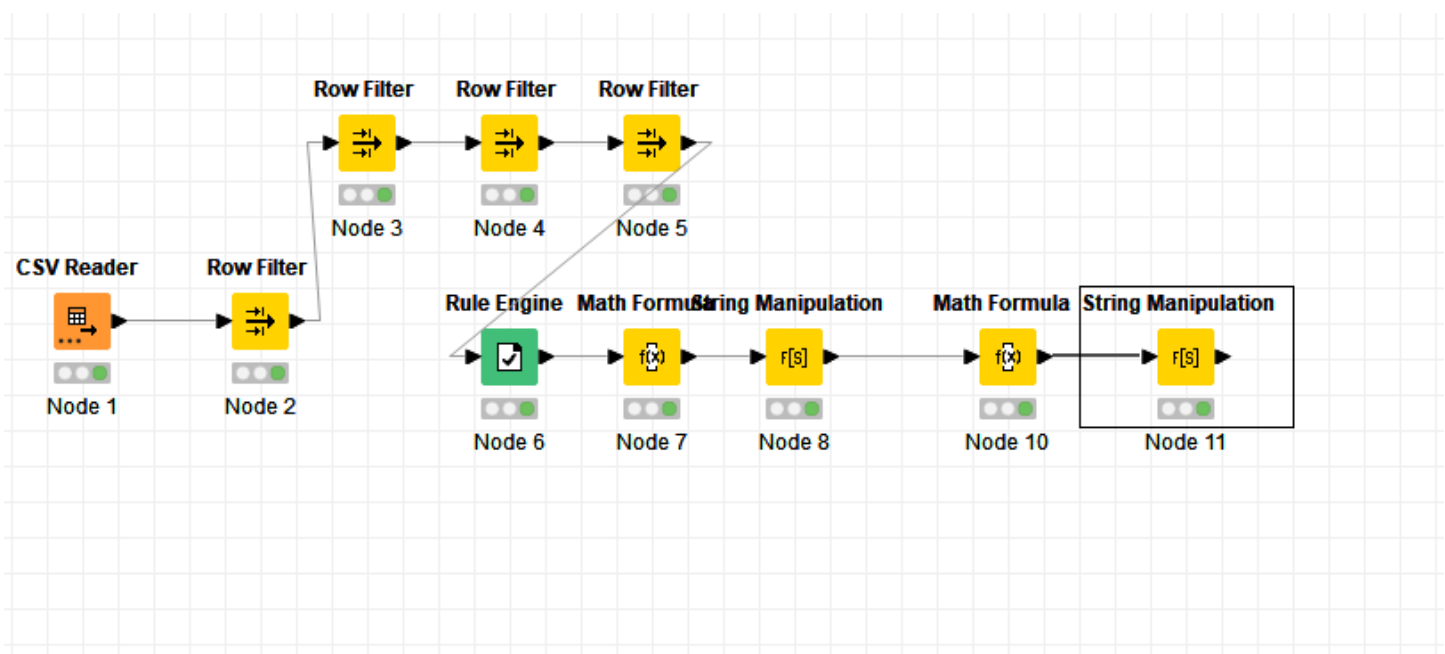
**Faculty:** Prof. Maheshwari R

## *Lab 9*

### **Task:**

Netflix data manipulation in Kyme analytics platform.

### **Flow:**



### **Objective:**

The dataset already consists of a column named duration, that shows duration of movie in minutes and each datapoint is appended with string 'min'. Using this column I'll be creating a new column that will show duration of movie in hour and each datapoint will be appended with string 'hr'. I'll also try to find out which is the oldest release year for any movie in the dataset.

## CSV Reader:

File Table - 3:1 - CSV Reader

File Edit Hilite Navigation View

Table "default" - Rows: 7787 Spec - Columns: 12 Properties Flow Variables					
Row ID	show_id	type	title	director	cast
Row0	s1	TV Show	3%	?	João Miguel, Bianca Comparato, Michel Gomes, Rodolfo Valente, Vaneza Oliveira, Rafael Lozano,
Row1	s2	Movie	7:19	Jorge Michel...	Demián Bichir, Héctor Bonilla, Oscar Serrano, Azalia Ortiz, Octavio Michel, Carmen Beato
Row2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence Koh, Tommy Kuan, Josh Lai, Mark Lee, Susan Leor
Row3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly, Christopher Plummer, Crispin Glover, Martin Landau
Row4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aaron Yoo, Liza Lapira, Jacob Pitts, Laurence Fishbur
Row5	s6	TV Show	46	Serdar Akar	Erdal Beşikçioğlu, Yasemin Allen, Melis Birkan, Saygıno Soysal, Berkan Altın, Metin Belgin, Ay.
Row6	s7	Movie	122	Yasir Al Yasiri	Amina Khalil, Ahmed Dawood, Tarek Lotfy, Ahmed El Fishawy, Mahmoud Hijazi, Jihane Khalil, Asma
Row7	s8	Movie	187	Kevin Reyno...	Samuel L. Jackson, John Heard, Kelly Rowan, Clifton Collins Jr., Tony Plana
Row8	s9	Movie	706	Shravan Ku...	Divya Dutta, Atul Kulkarni, Mohan Agashe, Anupam Shyam, Raayo S. Bakhirta, Yashvit Sancheti, i
Row9	s10	Movie	1920	Vikram Bhatt	Rajneesh Duggal, Adah Sharma, Indraneil Sengupta, Anjori Alagh, Rajendranath Zutshi, Vipin Sha
Row10	s11	Movie	1922	Zak Hilditch	Thomas Jane, Molly Parker, Dylan Schmid, Kaitlyn Bernard, Bob Frazer, Brian d'Arcy James, Neal M
Row11	s12	TV Show	1983	?	Robert Wiackiewicz, Maciej Musiała, Michalina Olszowska, Andrzej Chyra, Clive Russell, Zofia W
Row12	s13	TV Show	1994	Diego Enriqu...	?
Row13	s14	Movie	2,215	Nottapon Bo...	Artiwara Kongmalai
Row14	s15	Movie	3022	John Suits	Omar Epps, Kate Walsh, Miranda Cosgrove, Angus Macfadyen, Jorja Fox, Enver Gjokaj, Haaz Sle
Row15	s16	Movie	Oct-01	Kunle Afolayan	Sadiq Daba, David Bailie, Kayode Olaiya, Kehinde Bankole, Fabian Adeoye Lojede, Nick Rhys, Kuni
Row16	s17	TV Show	Feb-09	?	Shahd El Yaseen, Shaila Sabt, Hala, Hanadi Al-Kandari, Salma Salem, Ibrahim Al-Harbi, Mahmoud E
Row17	s18	Movie	22-Jul	Paul Greeng...	Anders Danielsen Lie, Jon Årnes, Jonas Strand Gravli, Ola G. Furuseth, Maria Bock, Thorbjørn, r
Row18	s19	Movie	15-Aug	Swapnaneel...	Rahul Pethe, Mrunmayee Deshpande, Adinath Kothare, Vaibhav Mangale, Jaywant Wadkar, Satis
Row19	s20	Movie	'89	?	Lee Dixon, Ian Wright, Paul Merson
Row20	s21	Movie	â€œKuch ...	Onir	Geetanjali Thapa, Zain Khan Durrani, Shray Rai Tiwari, Mona Ambegaonkar, Chandreyee Ghosh, E
Row21	s22	Movie	â€œGoli Soda 2	Vijay Milton	Samuthirakani, Bharath Seeni, Vinoth, Esakki Barath, Chemban Vinod Jose, Gautham Menon, Krish
Row22	s23	Movie	â€œMaj Rati ...	Santwana B...	Adil Hussain, Shakil Imtiaz, Mahendra Rabha, Sulakshana Baruah, Rahul Gautam Sarma, Kulada Bh
Row23	s24	Movie	â€œMayurakshi	Atanu Ghosh	Soumitra Chatterjee, Prasenjit Chatterjee, Indrani Halder, Sudipta Chakraborty, Suman Banerjee,

Then using row filter nodes to remove all empty data points

## Rule Engine:

Dialog - 3:6 - Rule Engine

File

Rule Editor | Flow Variables | Memory Policy

Column List

- ROWID
- ROWINDEX
- ROWCOUNT
- S show\_id
- S type
- S title
- S director
- S cast
- S country
- S date\_added
- I release\_year
- S rating
- S duration
- S listed\_in
- S description

Flow Variable List

- S knime.workspace

Category: All

Function

- ? < ?
- ? <= ?
- ? = ?
- ? > ?
- ? >= ?
- ? AND ?
- ? IN ?
- ? LIKE ?
- ? MATCHES ?
- ? OR ?
- ? XOR ?
- FALSE
- MISSING ?
- NOT ?

Description

Expression

```
1 // enter ordered set of rules, e.g.:
2 // $double column name$ > 5.0 => "large"
3 // $string column name$ LIKE "*blue*" => "small and blue"
4 // TRUE => "default outcome"
5 $release_year$ < 2000 => "Old Movie"
6 TRUE => "New Movie"
```

Append Column: Old or New S

Replace Column: S description

Math Formula Node:

File

Math Expression Flow Variables Memory Policy

Column List

ROWINDEX  
ROWCOUNT  
I release\_year

Flow Variable List

Category

All

Function

ROWCOUNT  
ROWINDEX  
pi  
e  
COL\_MIN(col\_name)  
COL\_MAX(col\_name)  
COL\_MEAN(col\_name)  
COL\_MEDIAN(col\_name)  
COL\_SUM(col\_name)  
COL\_STDDEV(col\_name)  
COL\_VAR(col\_name)  
ln(x)  
log(x)  
log10(x)

Description

Expression

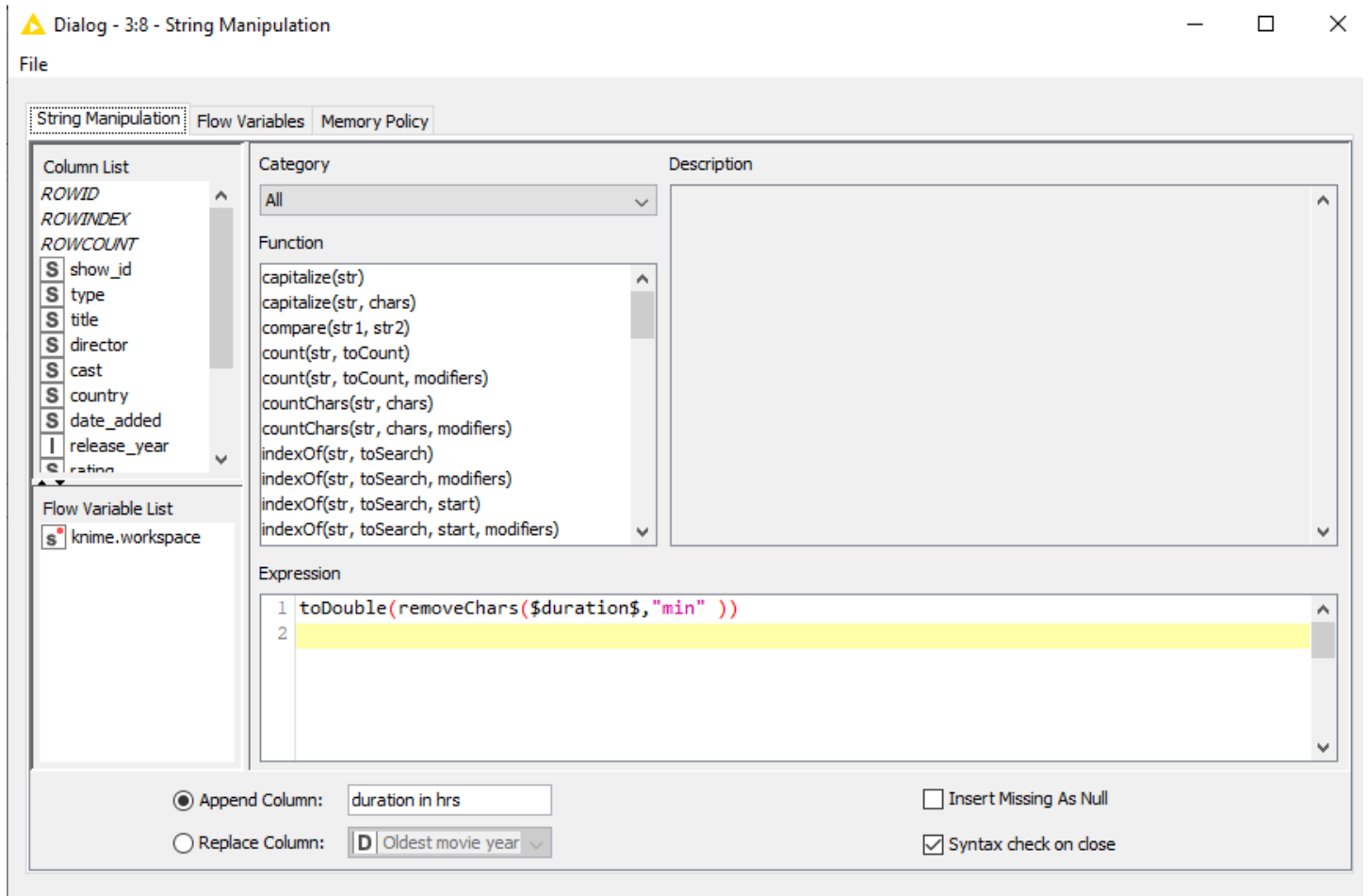
1 COL\_MIN(\$release\_year\$)

☒ Append Column: Oldest movie year

☐ Replace Column: S Old or New

☐ Convert to Int

String Manipulation Node 1:



Math Formula 2 Node:

File

Math Expression

Flow Variables

Memory Policy

Column List

ROWINDEX

ROWCOUNT

I release\_year

D Oldest movie year

D duration in hrs

Flow Variable List

Category

All

Description

Function

ROWCOUNT

ROWINDEX

pi

e

COL\_MIN(col\_name)

COL\_MAX(col\_name)

COL\_MEAN(col\_name)

COL\_MEDIAN(col\_name)

COL\_SUM(col\_name)

COL\_STDDEV(col\_name)

COL\_VAR(col\_name)

ln(x)

log(x)

log10(x)

Expression

1

\$duration in hrs\$/60

Append Column:

Replace Column:

D duration in hrs

Convert to Int

String Manipulation 2:

String Manipulation | Flow Variables | Memory Policy

Column List

- ROWID
- ROWINDEX
- ROWCOUNT
- show\_id
- type
- title
- director
- cast
- country
- date\_added
- release\_year
- rating

Flow Variable List

- knime.workspace

Category: All

Function

- capitalize(str)
- capitalize(str, chars)
- compare(str 1, str 2)
- count(str, toCount)
- count(str, toCount, modifiers)
- countChars(str, chars)
- countChars(str, chars, modifiers)
- indexOf(str, toSearch)
- indexOf(str, toSearch, modifiers)
- indexOf(str, toSearch, start)
- indexOf(str, toSearch, start, modifiers)

Description

Expression

```
1 join(substr(string($duration in hrs$),0 , 4)," hr")
2
```

☐ Append Column: 
☐ Insert Missing As Null

☒ Replace Column: 
☒ Syntax check on close

Final Table showing duration of movie in hour:

Row ID		S Old or ...	D Oldest ...	S duratio...
Row1	mits Mexico City, trapped survivors from all walks of life wait to be rescued while trying desperately to stay alive.	New Movie	1,942	1.55 hr
Row2	ead, his fellow soldiers are forced to confront a terrifying secret that's haunting their jungle island training camp.	New Movie	1,942	1.3 hr
Row3	oll robots hide in fear from dangerous machines out to exterminate them, until a brave newcomer joins the gr...	New Movie	1,942	1.33 hr
Row4	ome card-counting experts with the intent of swindling millions out of Las Vegas casinos by playing blackjack.	New Movie	1,942	2.05 hr
Row5	s with a treatment for his comatose sister that blends medical and shamanic cures, but unlocks a shocking side ...	New Movie	1,942	?
Row6	e admitted to a grisly hospital are separated and must find each other to escape " before death finds them.	New Movie	1,942	1.58 hr
Row7	ents attacks him, dedicated teacher Trevor Garfield grows weary of the gang warfare in the New York City sch...	Old Movie	1,942	1.98 hr
Row8	psychiatrist wife treats the bizarre medical condition of a psychic patient, who knows much more than he's lea...	New Movie	1,942	1.96 hr
Row9	ito a castle that is slated to become a luxury hotel. But something inside is determined to stop the renovation.	New Movie	1,942	2.38 hr
Row10	tting to his wife's murder, but her death is just the beginning of a macabre tale. Based on Stephen King's novella.	New Movie	1,942	1.71 hr
Row13	s rock star Artiwarar Kongmalai on his historic, 2,215-kilometer charity run across Thailand in 2017.	New Movie	1,942	1.48 hr
Row14	enly destroyed in a mysterious catadysm, the astronauts aboard a marooned space station slowly lose their m...	New Movie	1,942	1.51 hr
Row15	looming independence from Britain, detective Danladi Waziri races to capture a killer terrorizing local women.	New Movie	1,942	2.48 hr
Row17	in Norway, a young survivor, grieving families and the country rally for justice and healing. Based on a true st...	New Movie	1,942	2.4 hr
Row18	zany mishap in a Mumbai chawl disrupts a young love story while compelling the residents to unite in aid of a lit...	New Movie	1,942	2.06 hr
Row20	er the Internet, two strangers form a tight friendship " not knowing they already share a bond.	New Movie	1,942	1.83 hr
Row21	athlete struggle to better their lives despite obstacles like crooked politicians, evil dons and caste barriers.	New Movie	1,942	2.13 hr
Row22	e town that launched his career, encountering people who spark nostalgic, often painfully illuminating flashbacks.	New Movie	1,942	1.95 hr
Row23	turns to Kolkata to visit his ailing father, long-buried memories resurface, bringing new discoveries with them.	New Movie	1,942	1.66 hr
Row27	a lone man stays locked inside his apartment, digitally cut off from seeking help and desperate to find a way o...	New Movie	1,942	1.65 hr
Row28	story is retold alongside those of five Holocaust survivors in this poignant documentary from Oscar winner Hel...	New Movie	1,942	1.58 hr
Row31	or years, a young man puts up his best efforts to move out of the friend zone until she reveals she's getting m...	New Movie	1,942	1.7 hr
Row32	n from best friends to newlyweds, a quick pregnancy creates uncertainty for the future of their young marriage.	New Movie	1,942	1.73 hr
Row33	inally attracts the interest of her longtime crush, she lands in the cross hairs of his ex, a social media celebrity.	New Movie	1,942	1.65 hr