

K.R. MANGALAM UNIVERSITY, GURUGRAM-122103

SCHOOL OF ENGENIERRING AND TECHNOLOGY

ASSIGNMENT 3

Data Analysis with Power BI s KNIME

ETMMML174

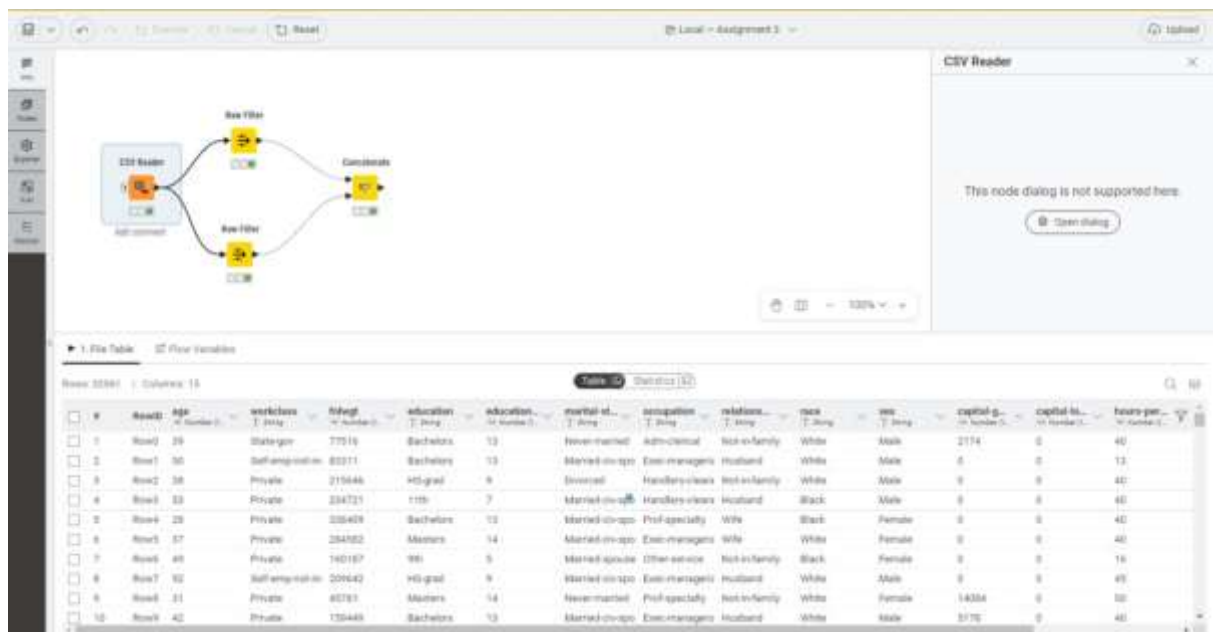


Department: SOET	Session: 2025-27
Program: MCA (AI & ML)	Semester: 1
Course Code: ETMMML174	College Roll no: 2501940069
Course Name: Data Analysis with Power BI & KNIME	
Submitted by: Tanya Bhatia	Faculty: Mr. Mohammad Ajaz

2 Power BI Assignment 3

- 1) Read the adult.csv file available in the **data** folder on the KNIME Hub. The data are provided by the **UCI Machine Learning Repository**.
- 2) Extract people with age between 20 and 40 (both included) and working in a workclass starting with "S"
- 3) Extract people with age between 40 and 60 (both included) and working in a workclass starting with "P"
- 4) Concatenate both subsets into a single data table

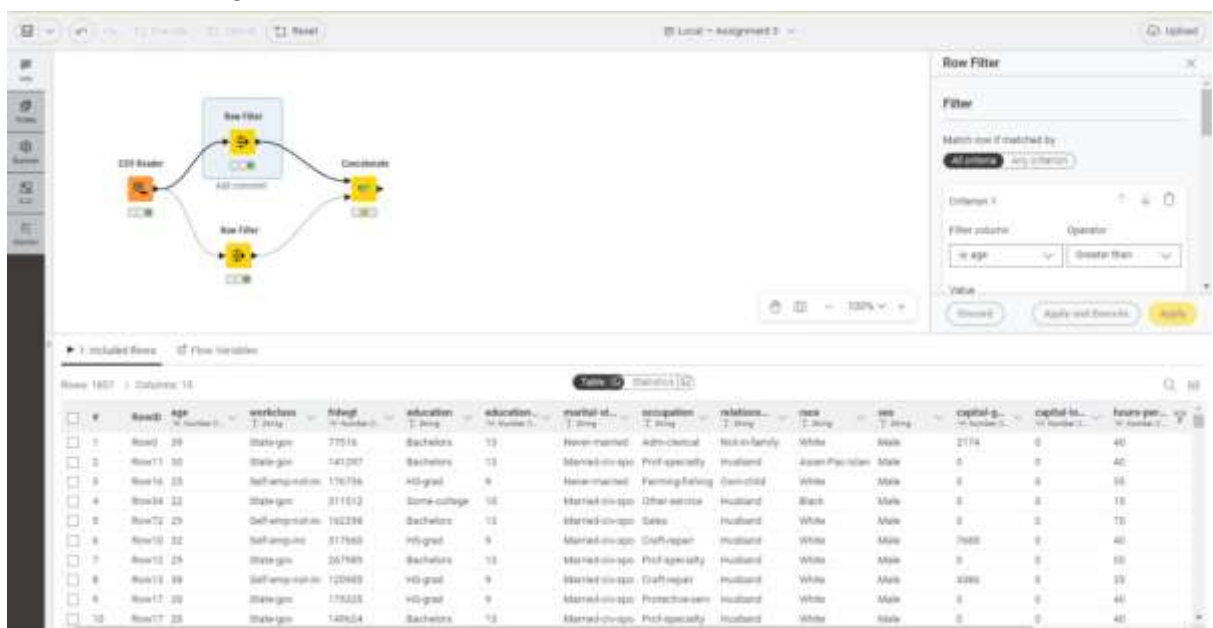
1) Read the adult.csv file



The screenshot shows a KNIME workflow with three nodes: 'CSV Reader', 'Row Filter', and 'Concatenate'. The 'CSV Reader' node is connected to the 'Row Filter' node, which is then connected to the 'Concatenate' node. The 'Row Filter' node is configured with the filter: 'age >= 20 AND age <= 40 AND workclass = S'. The 'Concatenate' node is configured with the filter: 'workclass = P'. The 'CSV Reader' node is also configured with the file path: 'C:\Users\Tanya\Documents\adult.csv'. The 'Row Filter' node is also configured with the filter: 'age >= 20 AND age <= 40 AND workclass = S'. The 'Concatenate' node is also configured with the filter: 'workclass = P'. The 'CSV Reader' node is also configured with the file path: 'C:\Users\Tanya\Documents\adult.csv'. The 'Row Filter' node is also configured with the filter: 'age >= 20 AND age <= 40 AND workclass = S'. The 'Concatenate' node is also configured with the filter: 'workclass = P'.

#	RowID	age	workclass	fnlwgt	education	education-yr	marital-st.	occupation	relations	race	sex	capital-g.	capital-l.	hours-per.
1	Row0	29	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40
2	Row1	35	Self-emp-not-inc	83211	Bachelors	13	Married-civ-sp	Exec-managerial	Husband	White	Male	0	0	13
3	Row2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	White	Male	0	0	40
4	Row3	58	Private	234721	11th	7	Married-civ-sp	Handlers-cleaners	Husband	Black	Male	0	0	40
5	Row4	29	Private	226409	Bachelors	13	Married-civ-sp	Prof-specialty	Wife	Black	Female	0	0	40
6	Row5	57	Private	254752	Masters	14	Married-civ-sp	Exec-managerial	Wife	White	Female	0	0	40
7	Row6	49	Private	140157	9th	5	Married-spouse	Other service	Not-in-family	Black	Female	0	0	16
8	Row7	52	Self-emp-not-inc	20842	HS-grad	9	Married-civ-sp	Exec-managerial	Husband	White	Male	0	0	49
9	Row8	31	Private	40761	Masters	14	Never-married	Prof-specialty	Not-in-family	White	Female	14354	0	52
10	Row9	42	Private	139449	Bachelors	13	Married-civ-sp	Exec-managerial	Husband	White	Male	5770	0	40

2) Extract people with age between 20 and 40 (both included) and working in a work class starting with "S"



The screenshot shows a KNIME workflow with three nodes: 'CSV Reader', 'Row Filter', and 'Concatenate'. The 'CSV Reader' node is connected to the 'Row Filter' node, which is then connected to the 'Concatenate' node. The 'Row Filter' node is configured with the filter: 'age >= 20 AND age <= 40 AND workclass = S'. The 'Concatenate' node is configured with the filter: 'workclass = P'. The 'CSV Reader' node is also configured with the file path: 'C:\Users\Tanya\Documents\adult.csv'. The 'Row Filter' node is also configured with the filter: 'age >= 20 AND age <= 40 AND workclass = S'. The 'Concatenate' node is also configured with the filter: 'workclass = P'.

#	RowID	age	workclass	fnlwgt	education	education-yr	marital-st.	occupation	relations	race	sex	capital-g.	capital-l.	hours-per.
1	Row0	29	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male	2174	0	40
2	Row1	35	State-gov	742207	Bachelors	13	Married-civ-sp	Prof-specialty	Husband	Asian-Pac-Islander	Male	0	0	40
3	Row16	25	Self-emp-not-inc	176736	HS-grad	9	Never-married	Farming-fishing	Own-child	White	Male	0	0	58
4	Row34	33	State-gov	211512	Some-college	10	Married-civ-sp	Other service	Husband	Black	Male	0	0	18
5	Row79	29	Self-emp-not-inc	162296	Bachelors	13	Married-civ-sp	Sales	Husband	White	Male	0	0	70
6	Row10	32	Self-emp-inc	317865	HS-grad	9	Married-civ-sp	Craft-repair	Husband	White	Male	7685	0	40
7	Row12	29	State-gov	267989	Bachelors	13	Married-civ-sp	Prof-specialty	Husband	White	Male	0	0	55
8	Row19	39	Self-emp-not-inc	122989	HS-grad	9	Married-civ-sp	Craft-repair	Husband	White	Male	3390	0	39
9	Row17	30	State-gov	176325	HS-grad	9	Married-civ-sp	Protective serv	Husband	White	Male	0	0	44
10	Row17	35	State-gov	148624	Bachelors	13	Married-civ-sp	Prof-specialty	Husband	White	Male	0	0	40

- 3) Extract People with age between 40 and 60 (both included) and working in a work class starting with “P”

The screenshot shows a Power BI Desktop interface. On the left, a data flow diagram illustrates the process: a 'Data Source' (orange icon) feeds into a 'Row Filter' (yellow icon), which then feeds into a 'Concatenate' (yellow icon). The 'Row Filter' is configured with the following criteria:

- Match row (if matched by): All columns, Any criterion
- Criterion 1: Filter column: age, Operator: Greater than, Value: 40
- Criterion 2: Filter column: workclass, Operator: Greater than, Value: P

The 'Concatenate' step is labeled 'All connected'. Below the diagram, a table titled 'Table 1' displays the filtered data. The table has 15 columns: #, RowID, age, workclass, bweight, education, marital st., occupation, relations, race, sex, capital gain, capital loss, and hours per. The data is as follows:

#	RowID	age	workclass	bweight	education	marital st.	occupation	relations	race	sex	capital gain	capital loss	hours per.
1	Row0	53	Private	234725	11th	7	Married-in-opp	Handlers-clean	Husband	Black	Male	0	40
2	Row6	49	Private	140187	9th	5	Married spouse	Other service	Not in family	Black	Female	0	18
3	Row9	42	Private	159449	Bachelors	13	Married-in-opp	Exec-manager	Husband	White	Male	1179	40
4	Row21	54	Private	302146	HS-grad	9	Separated	Other service	Unmarried	Black	Female	0	20
5	Row32	40	Private	117027	11th	7	Married-in-opp	Transportation	Husband	White	Male	0	2042
6	Row34	50	Private	109019	HS-grad	9	Divorced	Tech support	Unmarried	White	Female	0	40
7	Row29	44	Private	100394	HS-grad	9	Married-in-opp	Craft-repair	Husband	White	Male	0	40
8	Row52	45	Private	369445	Bachelors	13	Divorced	Exec-manager	Over-40s	White	Male	0	1438
9	Row42	46	Private	242405	11th	7	Never-married	Mach-op-mnt	Unmarried	White	Male	0	40
10	Row43	40	Private	94638	HS-grad	9	Separated	Admin-clerical	Unmarried	White	Female	0	40

- 4) Concatenate both subsets into a single data

The screenshot shows a Power BI Desktop interface. On the left, a data flow diagram illustrates the process: two 'Data Source' (orange icons) feed into two 'Row Filter' (yellow icons), which then feed into a 'Concatenate' (yellow icon). The 'Concatenate' step is configured with the following settings:

- How to combine input columns: Union
- RowID handling: Create new, Force existing

The 'Concatenate' step is labeled 'All connected'. Below the diagram, a table titled 'Table 1' displays the concatenated data. The table has 15 columns: #, RowID, age, workclass, bweight, education, marital st., occupation, relations, race, sex, capital gain, capital loss, and hours per. The data is as follows:

#	RowID	age	workclass	bweight	education	marital st.	occupation	relations	race	sex	capital gain	capital loss	hours per.
1	Row0	59	State-gov	77516	Bachelors	13	Never-married	Admin-clerical	Not in family	White	Male	2174	40
2	Row1	30	State-gov	141207	Bachelors	13	Married-in-opp	Prof-specialty	Husband	Asian-Pac-Islander	Male	0	40
3	Row2	23	Self-emp-not-inc	175756	HS-grad	9	Never-married	Farming-fishing	Over-40s	White	Male	0	33
4	Row3	22	State-gov	311932	Some college	10	Married-in-opp	Other service	Husband	Black	Male	0	18
5	Row4	29	Self-emp-not-inc	162298	Bachelors	13	Married-in-opp	Sales	Husband	White	Male	0	70
6	Row5	32	Self-emp-inc	217640	HS-grad	9	Married-in-opp	Craft-repair	Husband	White	Male	7680	40
7	Row6	34	State-gov	227689	Bachelors	13	Married-in-opp	Prof-specialty	Husband	White	Male	0	50
8	Row7	36	Self-emp-not-inc	122685	HS-grad	9	Married-in-opp	Craft-repair	Husband	White	Male	4389	35
9	Row8	38	State-gov	179335	HS-grad	9	Married-in-opp	Protective serv	Husband	White	Male	0	40
10	Row9	38	State-gov	149624	Bachelors	13	Married-in-opp	Prof-specialty	Husband	White	Male	0	40