



K.R. MANGALAM UNIVERSITY
THE COMPLETE WORLD OF EDUCATION

Data Analysis with Power BI & KNIME (ETSEDA115)

MCA (AI & ML)- Sem 1

Assignment 2

Roll No:-

2501940072

Submitted by:-

KULSUM BANOO

Submitted to:-

MR. MOHAMMAD AIJAZ

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) Calculate the average age and count for each one of the 4 groups defined by sex and income values

3) Join the two aggregated values to the original table

1) Read the adult.csv file

#	RowID	age	workclass	fnlwgt	education	education	marital-st...	occupation	relations...	race	sex
1	Row0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	White	Male
2	Row1	50	Self-emp-not-in	83311	Bachelors	13	Married-civ-spo	Exec-manager	Husband	White	Male
3	Row2	38	Private	215646	HS-grad	9	Divorced	Handlers-clean	Not-in-family	White	Male
4	Row3	53	Private	234721	11th	7	Married-civ-spo	Handlers-clean	Husband	Black	Male
5	Row4	28	Private	338409	Bachelors	13	Married-civ-spo	Prof-specialty	Wife	Black	Female
6	Row5	37	Private	284582	Masters	14	Married-civ-spo	Exec-manager	Wife	White	Female
7	Row6	49	Private	160187	9th	5	Married-spouse	Other-service	Not-in-family	Black	Female
8	Row7	52	Self-emp-not-in	209642	HS-grad	9	Married-civ-spo	Exec-manager	Husband	White	Male
9	Row8	31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family	White	Female
10	Row9	42	Private	159449	Bachelors	13	Married-civ-spo	Exec-manager	Husband	White	Male

2) Calculate the average age and count for each one of the 4 groups defined by sex and income values

#	RowID	sex	income	Mean(age)	Count*(age)
1	Row0	Female	<=50K	36.211	9592
2	Row1	Female	>50K	42.126	1179
3	Row2	Male	<=50K	37.147	15128
4	Row3	Male	>50K	44.626	6662

3) Join the two aggregated values to the original value

Screenshot of KNIME Analytics Platform interface showing a workflow for joining aggregated data.

Workflow Diagram:

```

graph LR
    CSVReader[CSV Reader] --> GroupBy[GroupBy]
    GroupBy --> Joiner[Joiner]
    Joiner --> Result[Join result]
  
```

Joiner Node Configuration:

- Matching Criteria:** All of the following
- Match:** Add matching criterion
- Compare values in join columns by:** Value and type

Input Ports (Left Table):

- Type: Left table
- Left input table

Input Ports (Right Table):

- Type: Right table
- Right input table

Output Port:

- Type: Join result

Result Table:

Index	sex	capital-g...	capital-lo...	hours-per...	native-co...	income	sex (Right)	income (...	Mean(age)	Count(a...)
te	Male	2174	0	40	United-States	<=50K	Female	>=50K	36.211	9592
te	Male	0	0	13	United-States	<=50K	Female	>50K	42.126	1179
te	Male	0	0	40	United-States	<=50K	Male	<=50K	37.147	15128
k	Male	0	0	40	United-States	<=50K	Male	>50K	44.626	6662