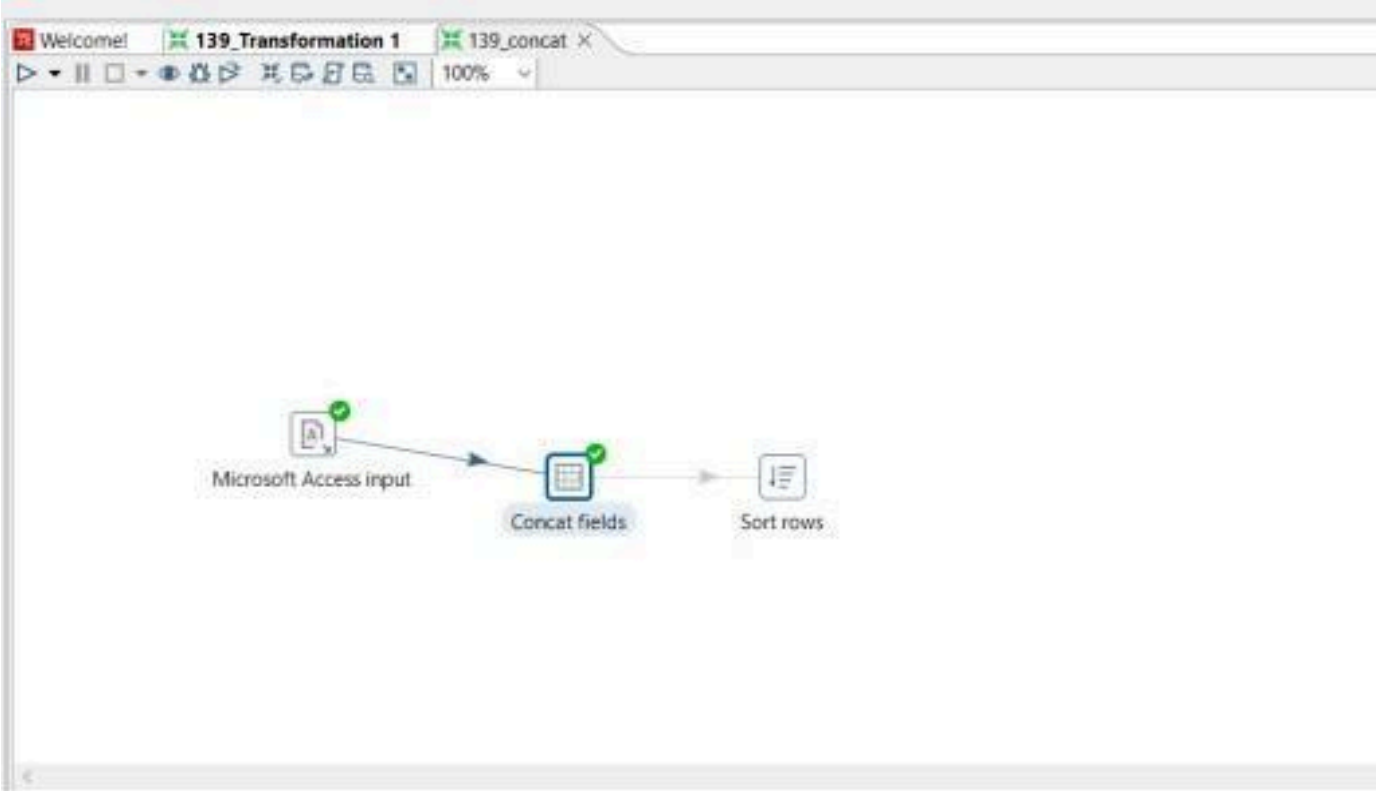


Name:- Akhilesh Singh

Roll No.186 / Batch : C

Lab. Assignment 3 Pentaho

1. Retrieve data from file staff.accdb and concat two columns firstname and lastname as fullname of teacher table and load data to data warehouse



The screenshot shows the Pentaho Data Integration (PDI) interface. At the top, there are tabs for 'Welcome!', '139_Transformation 1', and '139_concat'. Below the tabs is a toolbar with various icons and a '100%' zoom level. The main workspace displays a data flow diagram with three components: 'Microsoft Access input' (a document icon with a green checkmark), 'Concat fields' (a grid icon with a green checkmark), and 'Sort rows' (a document icon with a green checkmark). Arrows indicate the flow from 'Microsoft Access input' to 'Concat fields', and then from 'Concat fields' to 'Sort rows'.

Below the workspace, the 'Execution Results' section is visible. It includes tabs for 'Logging', 'Execution History', 'Step Metrics', 'Performance Graph', 'Metrics', and 'Preview data'. The 'Preview data' tab is selected, showing a table with 11 columns: '#', 'ID', 'First Name', 'Last Name', 'Street', 'City', 'State', 'ZipCode', 'Age', 'Sal', 'Bonus', and 'Full Name'. The table contains 5 rows of data.

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name
1	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	400101	5	32000	1200	Mary_Joe
2	2	Ram	Singh	LT Marg	Bardoa	Gujrat	400206	15	76000	1300	Ram_Singh
3	3	Akshay	Kumar	SV Road	Lucknow	UP	400207	25	67000	3400	Akshay_Kumar
4	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	400203	11	56000	2400	Viraj_Gupta
5	5	Samay	Khurana	WE highway	Nagpur	Maharastra	400102	18	34000	1200	Samay_Khurana

2. Sort column fullname in descending order and load data to data warehouse.

The screenshot displays the Alteryx Designer interface. At the top, there are tabs for 'Welcome!', '139_Transformation 1', and '139_concat'. Below the tabs is a toolbar with various icons. The main workspace shows a workflow diagram with three components: 'Microsoft Access input', 'Concat fields', and 'Sort rows'. Arrows indicate the flow from input to concat, and then to sort. Below the workflow, the 'Execution Results' section is visible, showing a table of data. The table has columns for ID, First Name, Last Name, Street, City, State, ZipCode, Age, Sal, Bonus, and Full Name. The data is sorted by Full Name in descending order.

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	400203	11	56000	2400	Viraj_Gupta
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	400102	18	34000	1200	Samay_Khurana
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	400206	15	76000	1300	Ram_Singh
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	400101	5	32000	1200	Mary_Joe
5	3	Akshay	Kumar	SV Road	Lucknow	UP	400207	25	67000	3400	Akshay_Kumar

3. Retrieve data from file and descritize age column of class table with <5, 5-10,10-15,15- 20, 20-25, >25 and load data to data warehouse

139_concat x

100%

```

graph LR
    A[Microsoft Access input] --> B[Concat fields]
    B --> C[Sort rows]
    C --> D[Number range]
  
```

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	400203	11	56000	2400	Viraj_Gupta	Less than 1
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	400102	18	34000	1200	Samay_Khurana	1-5
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	400206	15	76000	1300	Ram_Singh	Less than 1
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	400101	5	32000	1200	Mary_Joe	Less than 1
5	3	Akshay	Kumar	SV Road	Lucknow	UP	400207	25	67000	3400	Akshay_Kumar	Less than 1

4. Retrieve data from file and add sequence in the staff table.

139_concat x

100%

```

graph LR
    A[Microsoft Access input] --> B[Concat fields]
    B --> C[Sort rows]
    C --> D[Number range]
    D --> E[Add sequence]
  
```

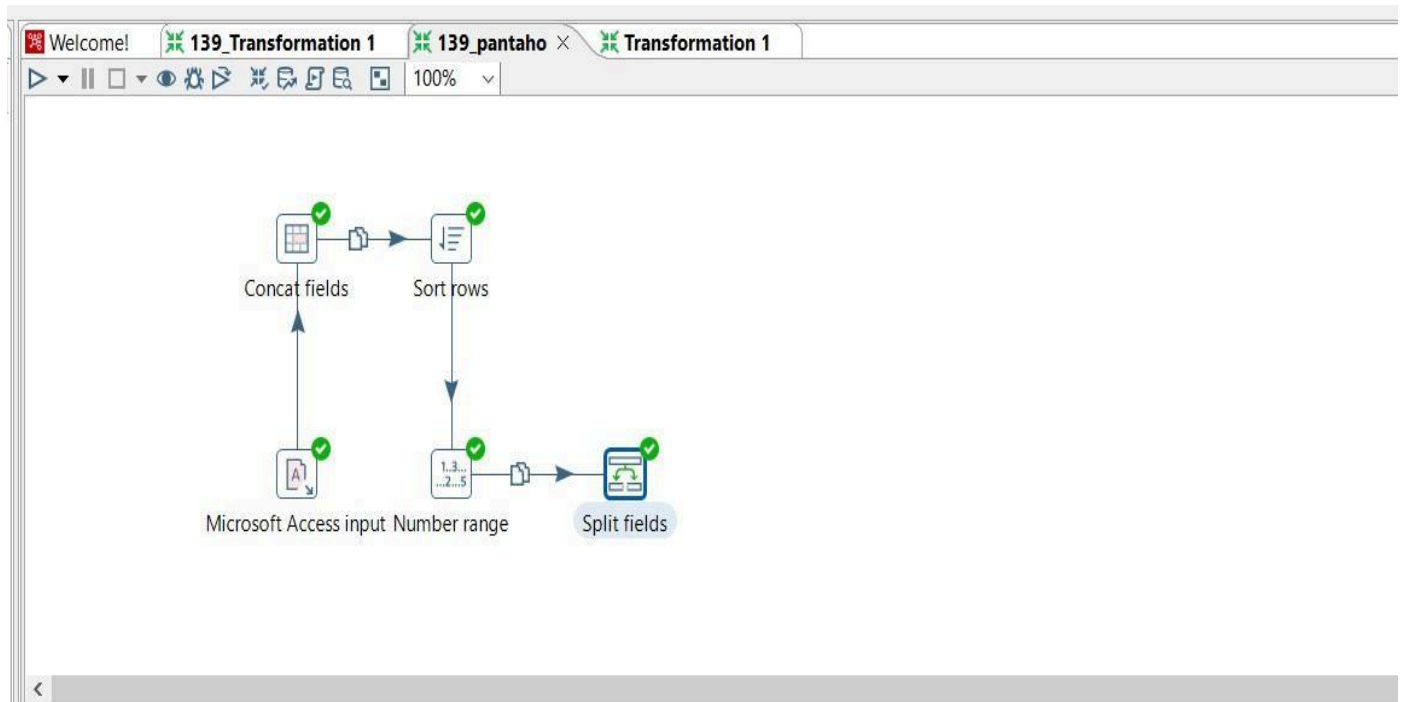
Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	400203	11	56000	2400	Viraj_Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	400102	18	34000	1200	Samay_Khurana	1-5	2
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	400206	15	76000	1300	Ram_Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	400101	5	32000	1200	Mary_Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	Lucknow	UP	400207	25	67000	3400	Akshay_Kumar	Less than 1	5

5. Split fullname column into two columns Fname and Lname using staff table



Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Fname	Lname	range
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	400203	11	56000.0	2400.0	Viraj	Gupta	Less than 1
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	400102	18	34000.0	1200.0	Samay	Khurana	1-5
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	400206	15	76000.0	1300.0	Ram	Singh	Less than 1
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	400101	5	32000.0	1200.0	Mary	Joe	Less than 1
5	3	Akshay	Kumar	SV Road	Lucknow	UP	400207	25	67000.0	3400.0	Akshay	Kumar	Less than 1

6. Retrieve data from file staff.accdb and set value of zip column of staff table in staffid Column

Microsoft Access input

Concat fields

Sort rows

Number range

Add sequence

Set field value

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	4	11	56000	2400	Viraj_Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	5	18	34000	1200	Samay_Khurana	1-5	2
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	2	15	76000	1300	Ram_Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	1	5	32000	1200	Mary_Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	Lucknow	UP	3	25	67000	3400	Akshay_Kumar	Less than 1	5

7. Load data transformed in step 1-6 in another MS – Access database.

139_Transformation 1 139_pantaho Transformation 1

100%

```

graph TD
    Input[Microsoft Access input] --> Concat[Concat fields]
    Concat --> Sort[Sort rows]
    Sort --> AddSeq[Add sequence]
    AddSeq --> Split[Split fields]
    Split --> SetVal[Set field value]
    SetVal --> Output[Microsoft Access output]
  
```

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

☒ First rows ☐ Last rows ☐ Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Fname	Lname	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	Jaipur	Rajasthan	4	11	56000.0	2400.0	Viraj	Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	Nagpur	Maharastra	5	18	34000.0	1200.0	Samay	Khurana	1-5	2
3	2	Ram	Singh	LT Marg	Bardoa	Gujrat	2	15	76000.0	1300.0	Ram	Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	Mumbai	Maharastra	1	5	32000.0	1200.0	Mary	Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	Lucknow	UP	3	25	67000.0	3400.0	Akshay	Kumar	Less than 1	5

8. Retrieve data from file staff.accdb and apply following string operation

a. Convert city column in upper case

Microsoft Access input → Concat fields → Sort rows → Number range → Add sequence → Set field value → String operations → AR

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range
1	4	Viraj	Gupta	Linking Road	JAIPUR	Rajasthan	4	11	56000	2400	Viraj_Gupta	Less than
2	5	Samay	Khurana	WE highway	NAGPUR	Maharastra	5	18	34000	1200	Samay_Khurana	1-5
3	2	Ram	Singh	LT Marg	BARDOA	Gujrat	2	15	76000	1300	Ram_Singh	Less than
4	1	Mary	Joe	Thakur Marg	MUMBAI	Maharastra	1	5	32000	1200	Mary_Joe	Less than
5	3	Akshay	Kumar	SV Road	LUCKNOW	UP	3	25	67000	3400	Akshay_Kumar	Less than

b. Convert state column in lower case

Microsoft Access input → Concat fields → Sort rows → Number range → Add sequence → Set field value → String operations → AR

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	JAIPUR	rajasthan	4	11	56000	2400	Viraj_Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	NAGPUR	maharastra	5	18	34000	1200	Samay_Khurana	1-5	2
3	2	Ram	Singh	LT Marg	BARDOA	gujrat	2	15	76000	1300	Ram_Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	MUMBAI	maharastra	1	5	32000	1200	Mary_Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	LUCKNOW	up	3	25	67000	3400	Akshay_Kumar	Less than 1	5

c. Apply left padding in city and fix width 20 and padding character should be "**"

139.pantaho

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	*****JAIPUR	rajasthan	4	11	56000	2400	Viraj_Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	*****NAGPUR	maharastra	5	18	34000	1200	Samay_Khurana	1-5	2
3	2	Ram	Singh	LT Marg	*****BARDOA	gujrat	2	15	76000	1300	Ram_Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	*****MUMBAI	maharastra	1	5	32000	1200	Mary_Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	*****LUCKNOW	up	3	25	67000	3400	Akshay_Kumar	Less than 1	5

d. Apply left padding in state and fix width 20 and padding character should be “#”

139.pantaho

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number
1	4	Viraj	Gupta	Linking Road	*****JAIPUR	#####rajasthan	4	11	56000	2400	Viraj_Gupta	Less than 1	1
2	5	Samay	Khurana	WE highway	*****NAGPUR	#####maharastra	5	18	34000	1200	Samay_Khurana	1-5	2
3	2	Ram	Singh	LT Marg	*****BARDOA	#####gujrat	2	15	76000	1300	Ram_Singh	Less than 1	3
4	1	Mary	Joe	Thakur Marg	*****MUMBAI	#####maharastra	1	5	32000	1200	Mary_Joe	Less than 1	4
5	3	Akshay	Kumar	SV Road	*****LUCKNOW	#####up	3	25	67000	3400	Akshay_Kumar	Less than 1	5

9. Retrieve data from file staff.accdb and calculate total salary using which is sum of salary

and bonus column

Microsoft Access input Concat fields Sort rows Number range Add sequence Set field value String operations Calculator

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name	range	Sequence Number	Total Salary
1	4	Viraj	Gupta	Linking Road	*****JAIPUR	*****rajasthan	4	11	56000.0	2400.0	Viraj_Gupta	Less than 1	1	58400.0
2	5	Samay	Khurana	WE highway	*****NAGPUR	*****maharashtra	5	18	34000.0	1200.0	Samay_Khurana	1-5	2	35200.0
3	2	Ram	Singh	LT Marg	*****BARDOA	*****gujrat	2	15	76000.0	1300.0	Ram_Singh	Less than 1	3	77300.0
4	1	Mary	Joe	Thakur Marg	*****MUMBAI	*****maharashtra	1	5	32000.0	1200.0	Mary_Joe	Less than 1	4	33200.0
5	3	Akshay	Kumar	SV Road	*****LUCKNOW	*****up	3	25	67000.0	3400.0	Akshay_Kumar	Less than 1	5	70400.0

10. Retrieve data from file staff.accdb and validate Age field where age should be greater or equal to 18

Microsoft Access input Concat fields Sort rows Number range Add sequence Set field value String operations Calculator Data validator Dummy (do nothing) 2

Execution Results

Logging Execution History Step Metrics Performance Graph Metrics Preview data

First rows Last rows Off

#	ID	First Name	Last Name	Street	City	State	ZipCode	Age	Sal	Bonus	Full Name
1	4	Viraj	Gupta	Linking Road	*****JAIPUR	*****rajasthan	4	11	56000.0	2400.0	Viraj_Gupta
2	5	Samay	Khurana	WE highway	*****NAGPUR	*****maharashtra	5	18	34000.0	1200.0	Samay_Khurana
3	2	Ram	Singh	LT Marg	*****BARDOA	*****gujrat	2	15	76000.0	1300.0	Ram_Singh