

Name - Harshit Parsai
Registration no. - 11702448
Roll no. - 06
Section - KM054
Question Number - 06
Subject - INT305

Informatica Data Integration (INT325)

Academic Task-2 - Report

Q6)

A) Create a mapping that performs the following operation using multiple transformations:

- 1.Display the countries whose sum of gold medals is more than 20
- 2.Display the countries who have won less than 10 medals till now
- 3.how many gold medal in swimming country wise

B)What is the difference between domain and node in informatica power center architecture?

Solution - >

For this solution in the coming pages i am attaching the snapshots and screenshots of the procedure and the practicals. I am explaining the procedure that we need to follow to do this question step wise step and then i will be putting the screenshots of the practical and the result of that as answer .

To do this question we must have our source file and the output files that we can create as per the requirement.

To import file in Informatica developer tool.

- Writing the steps here.

→ select the physical data object → select the flat files → choose the file location & file → click next → next & finish. Same steps for importing target files.

→ After importing we need to change the source file directory and target file directory,
C:\Informatica\10.1.1\server\infra-shared\srcfiles\

→ After this we have to create a mapping to create a mapping right click on the folder in project then select new → mapping → give name to mapping then click finish and then open the mapping and.

→ drag the source file in mapping in read mode.

- 1/ We need to display the country names whose sum of gold medals is more than 20.
- select a aggregator connect the input from the source file's country & Gold as input and output, add a port in aggregator as Total-gold and add a group as country we can do this by simply navigating to properties of aggregator and selecting port and group by option.
 - Select a filter and connect the output of aggregator country and Total-gold to filter select the filter and go to properties and add a filter as $\text{Total-gold} > 20$ in condition
 - Drag the output file in writing mode and connect the output of filter to the output file; only country field.
 - Run the mapping by selecting the mapping and right click and run mapping.
 - To see the output run the output file in Data viewer.
 - After doing this we can clearly see the result.

2) We have to display countries who have won less than 10 medals till now.

→ As we already have our source file imported.

→ Select the aggregator and connect country & total medal from the source to aggregator.

then add a port in aggregator following the same steps as above. give that port a name Totalmedal-country and add groupby-country.

→ Select the filter and add the output country and Totalmedal-country port from aggregator to filter add filter condition as Totalmedal-country < 10. Following the same steps as above question

→ Add the output file in write mode and add the country port from filter to output file.

→ Run the mapping by right-clicking on mapping and then run-mapping.

→ To view the result run the output file in data viewer you can see the output. (result).

3) how many gold medals in swimming country wise

→ select the aggregator and connect the port country, category, gold from source to aggregator and add a port Total-G-S and give definition $\text{Sum}(\text{Gold})$ to it. add the groupby two fields Country & Category.

→ select the filter transformation add the filter condition as category: 'Swimming' by following the same steps as mentioned in 1st part.

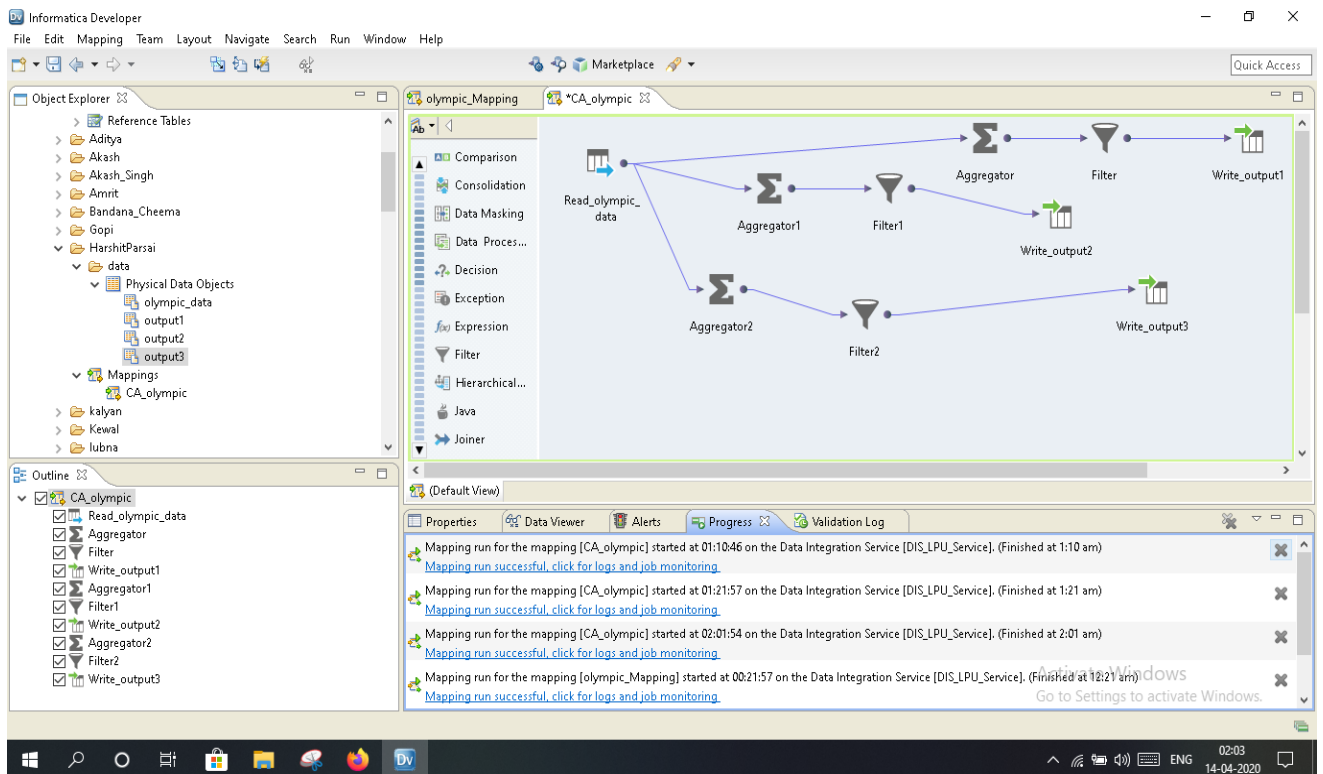
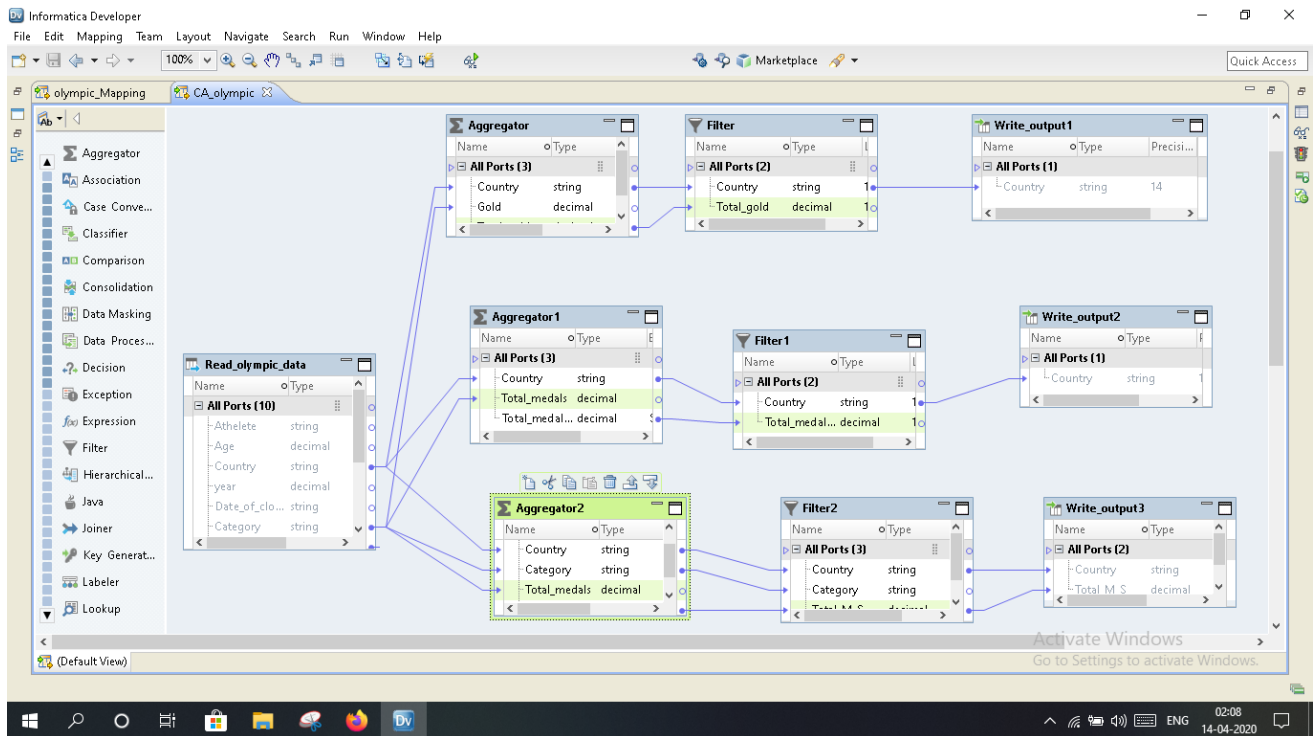
→ Import, drag the output file that will keep the output as country name and Total-gold-medal in swimming. Connect the port Country & total-G-S. from filter to output.

→ Run mapping and select the output file and run data viewer to see the result.

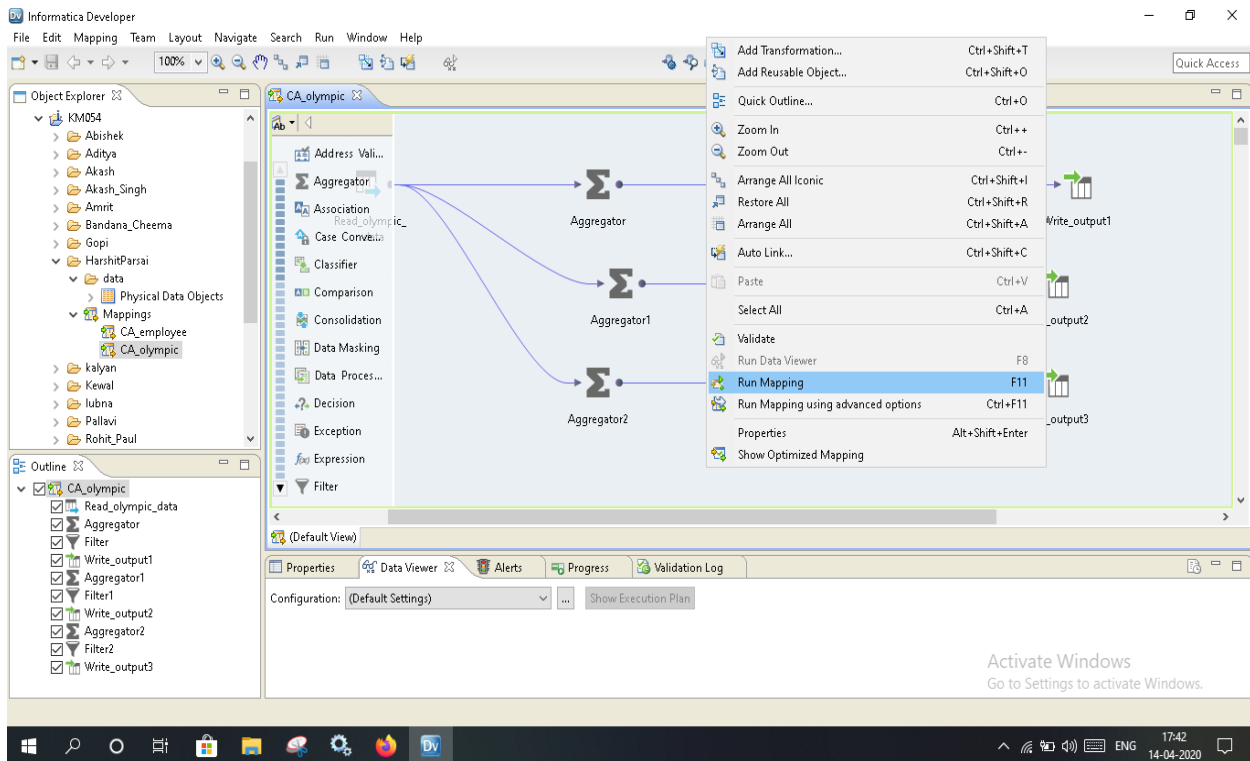
→ As we can see 1st file output contains name who won more than 20 country (Gold-medals)

2nd file who won less than 10 medal and 3rd file countrywise - gold medal in Swimming.

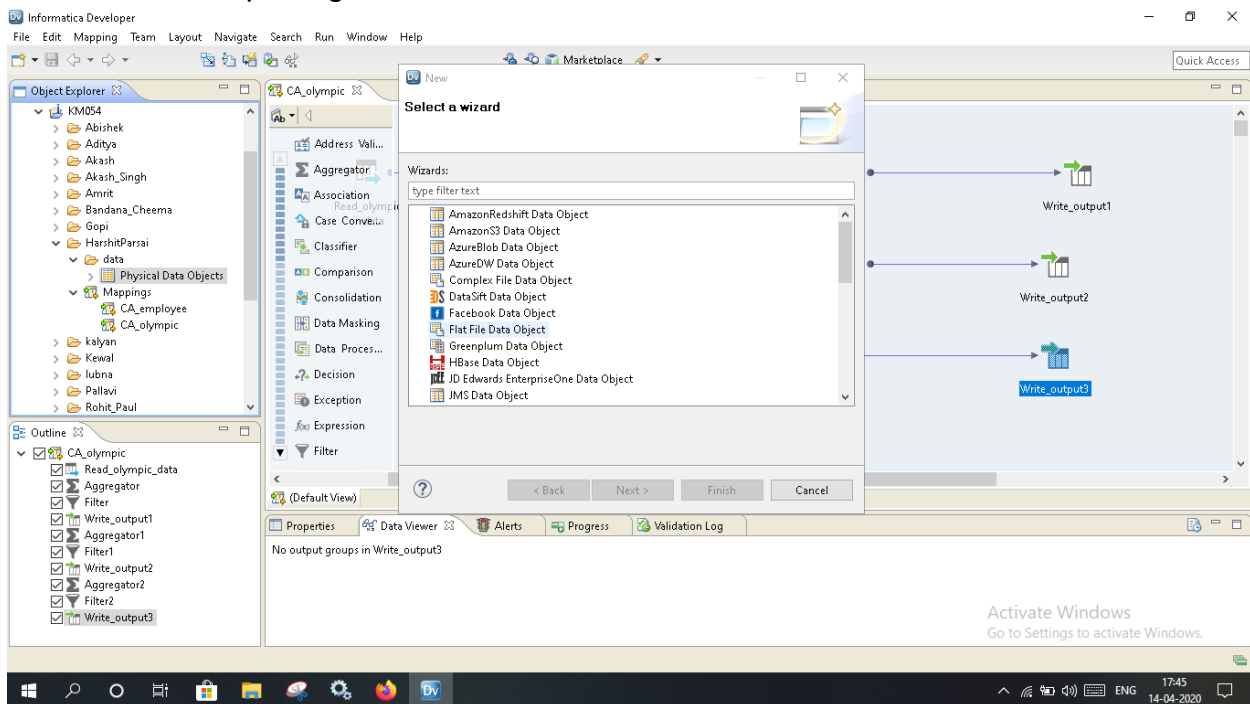
Screenshot of the final mapping using the multiple transformation and filtration



Screenshot of running the complete mapping that is made.



Screenshot of importing the files



Screenshot of the Data in the source file

Informatica Developer

File Edit Mapping Team Layout Navigate Search Run Window Help

Properties Data Viewer Alerts Progress Validation Log

Configuration: (Default Settings) Run Show: (All Outputs) Choose...

Output

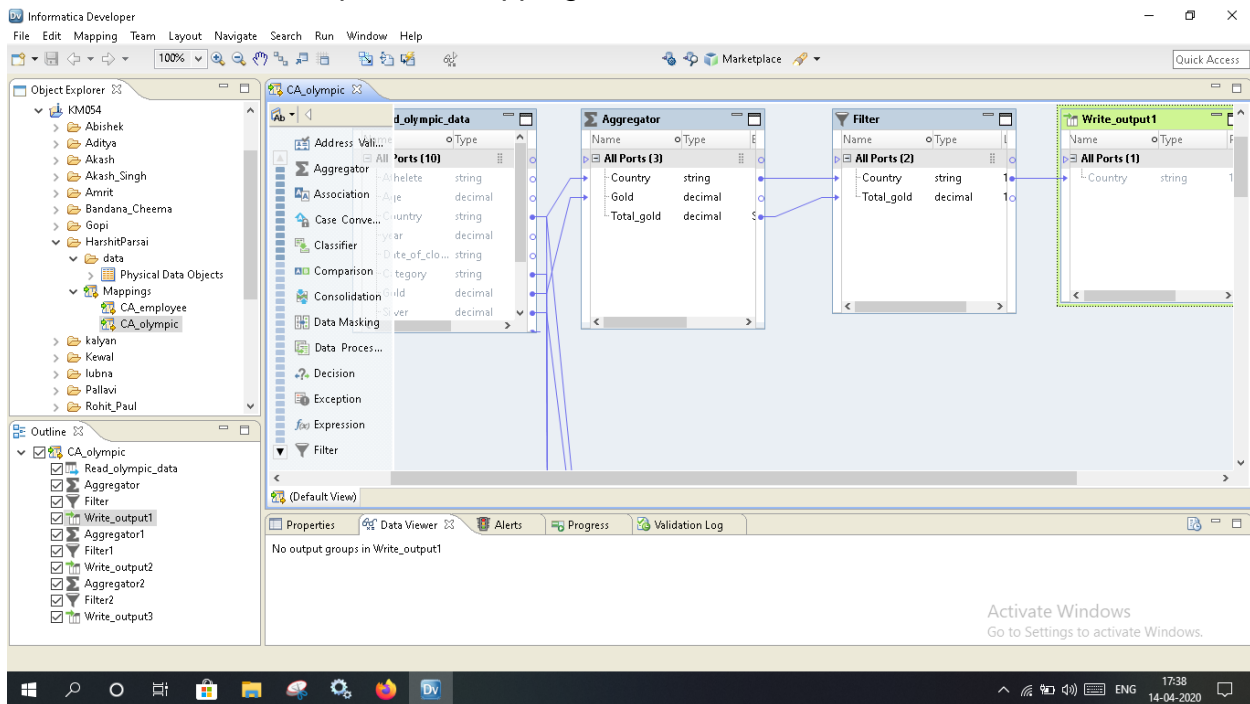
Name: Read_olympic_data

	Athlete	Age	Country	Year	Date_of_closing	Category	Gold	Silver	Bronze	Total_medals
1	Michael Phelps	23	United States	2008	8/24/2008	Swimming	8	0	0	8
2	Michael Phelps	19	United States	2004	8/29/2004	Swimming	6	0	2	8
3	Michael Phelps	27	United States	2012	8/12/2012	Swimming	4	2	0	6
4	Natalie Coughlin	25	United States	2008	8/24/2008	Swimming	1	2	3	6
5	Aleksey Nemov	24	Russia	2000	10/1/2000	Gymnastics	2	1	3	6
6	Alicia Coutts	24	Australia	2012	8/12/2012	Swimming	1	3	1	5
7	Missy Franklin	17	United States	2012	8/12/2012	Swimming	4	0	1	5
8	Ryan Lochte	27	United States	2012	8/12/2012	Swimming	2	2	1	5
9	Allison Schmitt	22	United States	2012	8/12/2012	Swimming	3	1	1	5
10	Natalie Coughlin	21	United States	2004	8/29/2004	Swimming	2	2	1	5
11	Ian Thorpe	17	Australia	2000	10/1/2000	Swimming	3	2	0	5
12	Dara Torres	33	United States	2000	10/1/2000	Swimming	2	0	3	5
13	Cindy Klassen	26	Canada	2006	2/26/2006	Speed Skating	1	2	2	5
14	Nastia Liukin	18	United States	2008	8/24/2008	Gymnastics	1	3	1	5
15	Marit Bjørgen	29	Norway	2010	2/28/2010	Cross Country...	3	1	1	5
16	Sun Yang	20	China	2012	8/12/2012	Swimming	2	1	1	4
17	Kirsty Coventry	24	Zimbabwe	2008	8/24/2008	Swimming	1	3	0	4
18	Libby Lenton	23	Australia	2008	8/24/2008	Swimming	2	1	1	4
19	Ryan Lochte	24	United States	2008	8/24/2008	Swimming	2	0	2	4
20	Inge de Bruijn	30	Netherlands	2004	8/29/2004	Swimming	1	1	2	4
21	Petria Thomas	28	Australia	2004	8/29/2004	Swimming	3	1	0	4
22	Ian Thorpe	21	Australia	2004	8/29/2004	Swimming	2	1	1	4
23	Inge de Bruijn	27	Netherlands	2000	10/1/2000	Swimming	3	1	0	4
24	Gary Hall Jr.	25	United States	2000	10/1/2000	Swimming	2	1	1	4

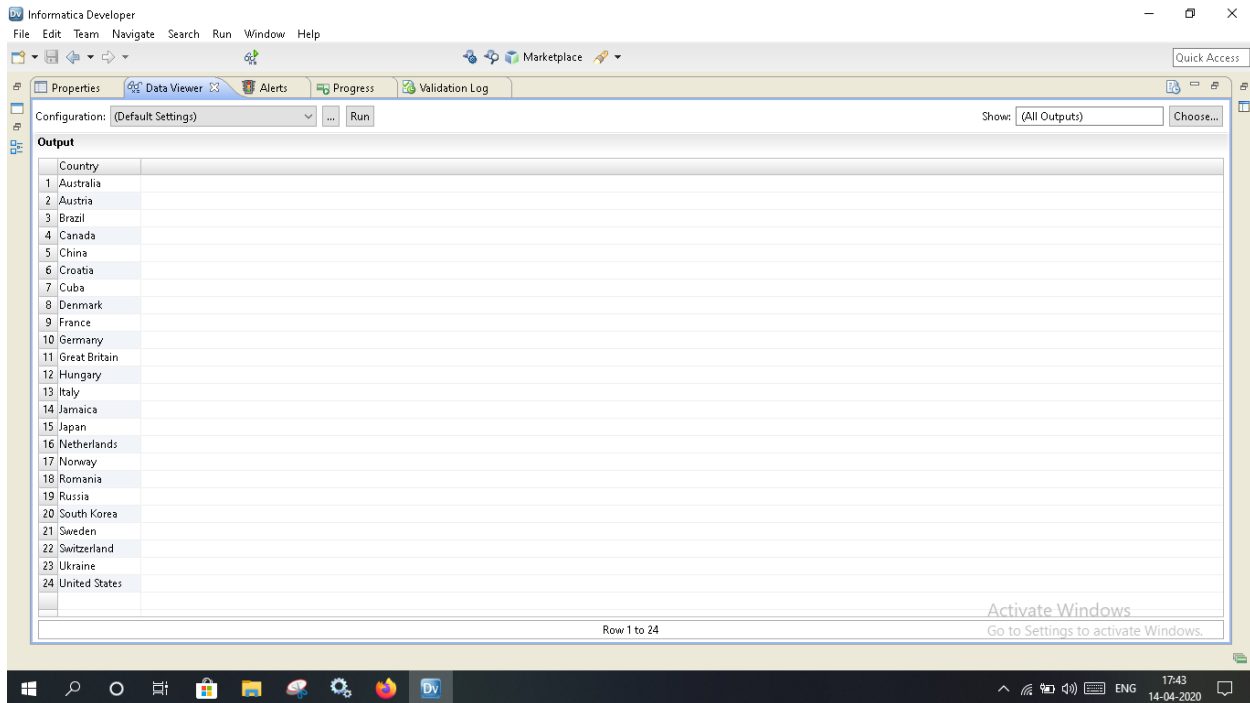
Row 1 to 1,000

Activate Windows
Go to Settings to activate Windows.

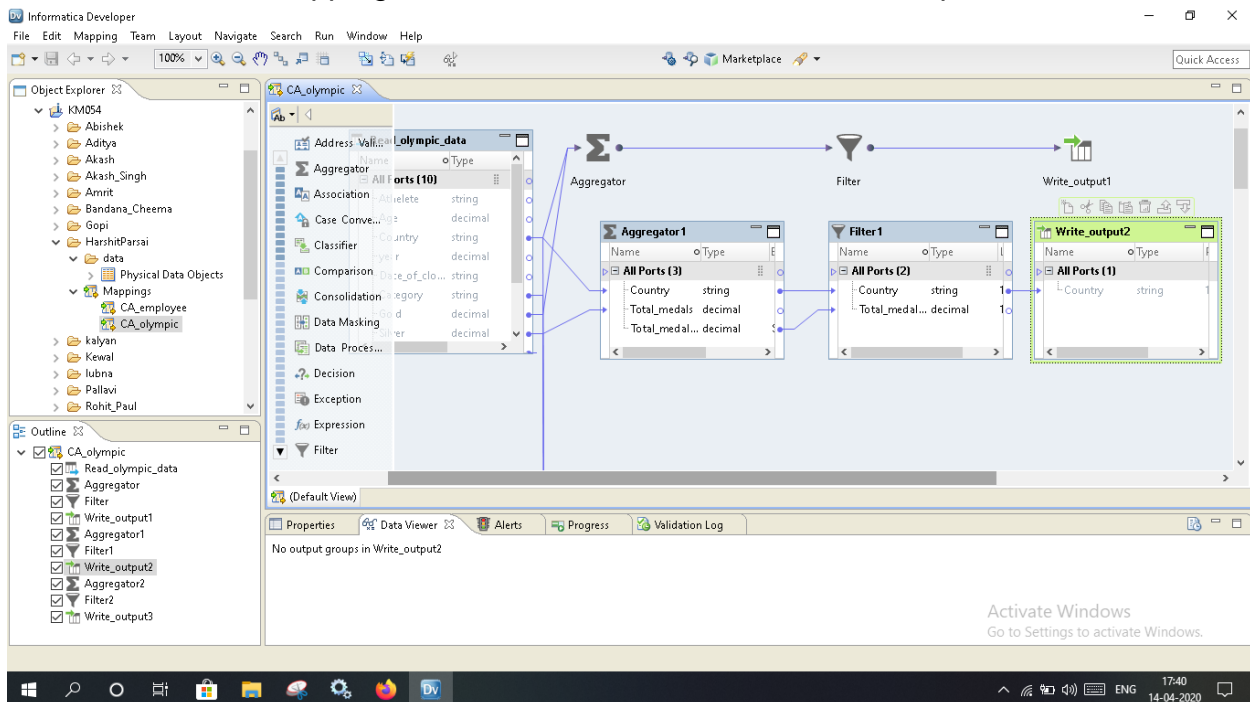
Screenshot of the first question mapping and transformation



Screenshot of the result that we got after running the first question mapping and transformation : showing the countries who more than 10 gold medals



Screenshot of the mapping and the transformation of the second Ports question



Screenshot of the result of the 2nd option in 1st question : showing the countries won less than 10 medals till now

Informatica Developer

File Edit Team Navigate Search Run Window Help

Marketplace Quick Access

Properties Data Viewer Alerts Progress Validation Log

Configuration: (Default Settings) Run Show: (All Outputs) Choose...

Output

Country
1 Algeria
2 Bahrain
3 Barbados
4 Botswana
5 Costa Rica
6 Cyprus
7 Ecuador
8 Egypt
9 Eritrea
10 Gabon
11 Grenada
12 Guatemala
13 Hong Kong
14 Ireland
15 Israel
16 Kuwait
17 Kyrgyzstan
18 Macedonia
19 Malaysia
20 Mauritius
21 Moldova
22 Mozambique
23 Panama
24 Portugal
25 Puerto Rico

Row 1 to 41

Activate Windows
Go to Settings to activate Windows.

Windows taskbar: 17:44 14-04-2020

Screenshot of the result of the 3rd option of this question : showing the country wise gold medal in the swimming

Informatica Developer

File Edit Team Navigate Search Run Window Help

Marketplace Quick Access

Properties Data Viewer Alerts Progress Validation Log

Configuration: (Default Settings) Run Show: (All Outputs) Choose...

Output

Country	Total_M_S
1 Australia	163
2 Austria	3
3 Belarus	2
4 Brazil	8
5 Canada	5
6 China	35
7 Costa Rica	2
8 Croatia	1
9 Denmark	1
10 France	39
11 Germany	32
12 Great Britain	11
13 Hungary	9
14 Italy	16
15 Japan	43
16 Lithuania	1
17 Netherlands	46
18 Norway	2
19 Poland	3
20 Romania	6
21 Russia	20
22 Serbia	1
23 Slovakia	2
24 Slovenia	1
25 South Africa	11

Row 1 to 33

Activate Windows
Go to Settings to activate Windows.

Windows taskbar: 17:45 14-04-2020

Solution B -)

→ Difference between the Nodes and Domain in informatica power center architecture.

Node

- Installation of informatica server on our machine then our machine acts as a node.
- Nodes can be of two types Gateway node or Worker node.
- Nodes are used to perform the operation and processes on domain
- Node is the machine where the stuff runs

Domain

- Domain is a group of one or more informatica Nodes.
- Only the node which is running as master gateway node is considered to be domain.
- Domain forms the environment upon which the informatica processes run.
- Domain is where the node keeps stuff.

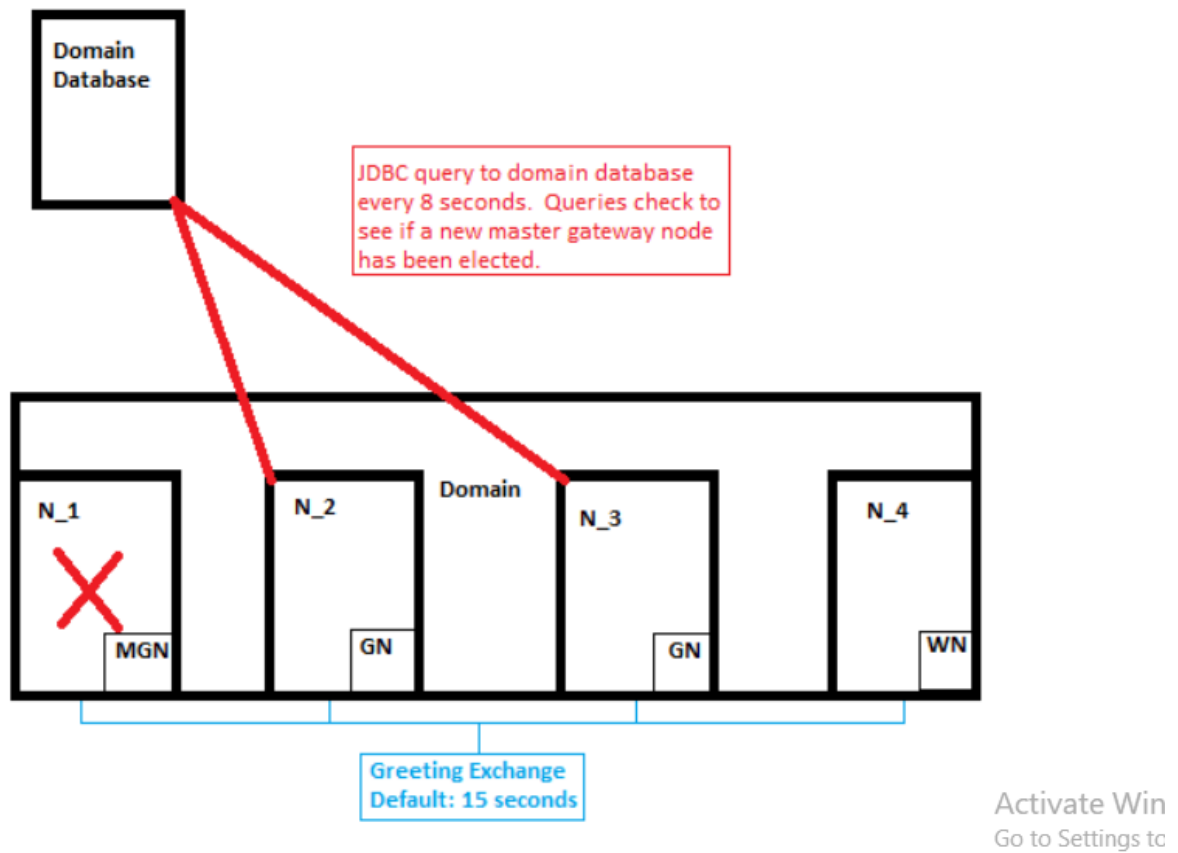
Node

- Node connects to the domain database using JDBC. It does not use ODBC.
- If a node goes down in a domain with multiple node domain with more than two gateway nodes where the master gateway dies, the remaining node becomes master gateway node.
- We can clearly say that Nodes are logical representation of machine in the domain.

Domain

- Domain Database is a standard relational database. The domain database is "backbone" that supports movement in domain and stores records and services.
- Domain remains active till it gets any issue with the Database, but if there is an issue in the Domain then it affects the entire connection and nodes connected to it.
- A domain is collection of node, services and databases.

Best Picture demonstrating the domain and the node difference :-



THANKYOU