

Assignment 1

Course Code - CAP447

Course Teacher – Dr. Punam Rattan

Name – Aniket Kumar

Registration Number - 12108348

Section - D2112

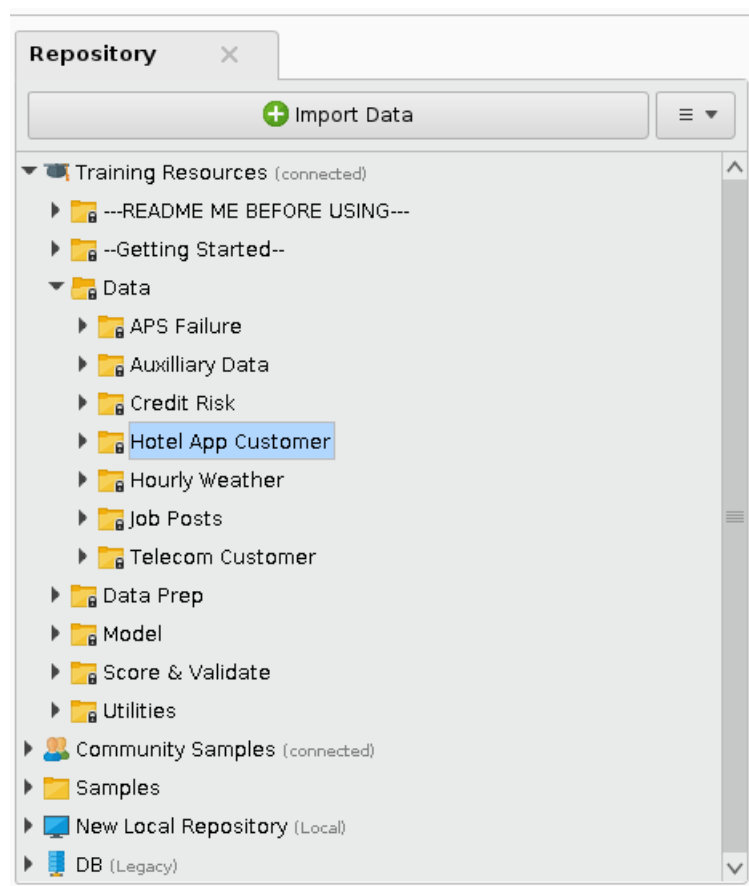
Roll Number - RD2112B81

Q1. Explain briefly the main components of RAPID MINER Studio.

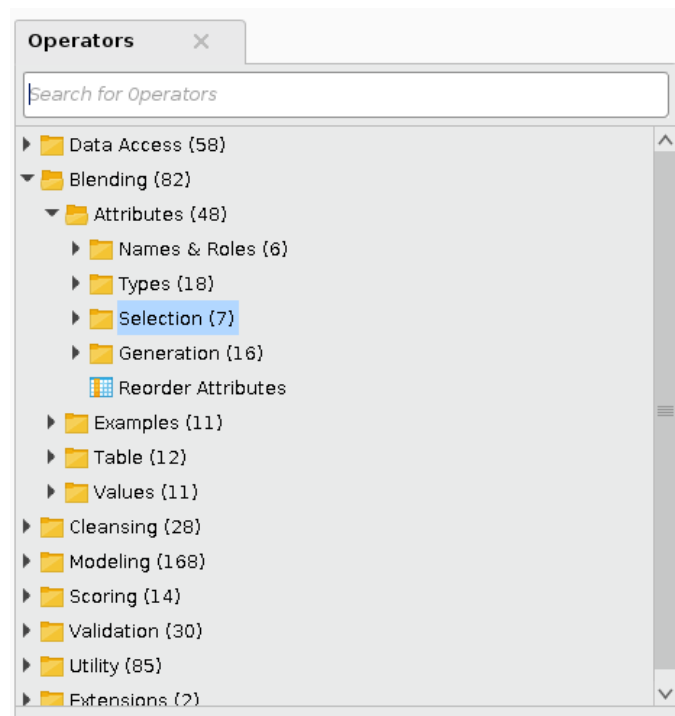
Rapid Miner studio is a platform to analyze and perform various operations on a data set to obtain the useful informations, graphical representations of data like line chart, bar chart, pie chart, etc.

Following are some main components of RapidMiner studio:

- **Repository**
 - It is the place where the data, processes and results are stored, either locally or remotely.
 - It is the collection of datasets, also known as folder, workspace or project.
 - We can import external datasets too and add to the repository.
 - The Repository can be used to store:
 - Data
 - Processes
 - Results
 - Operators



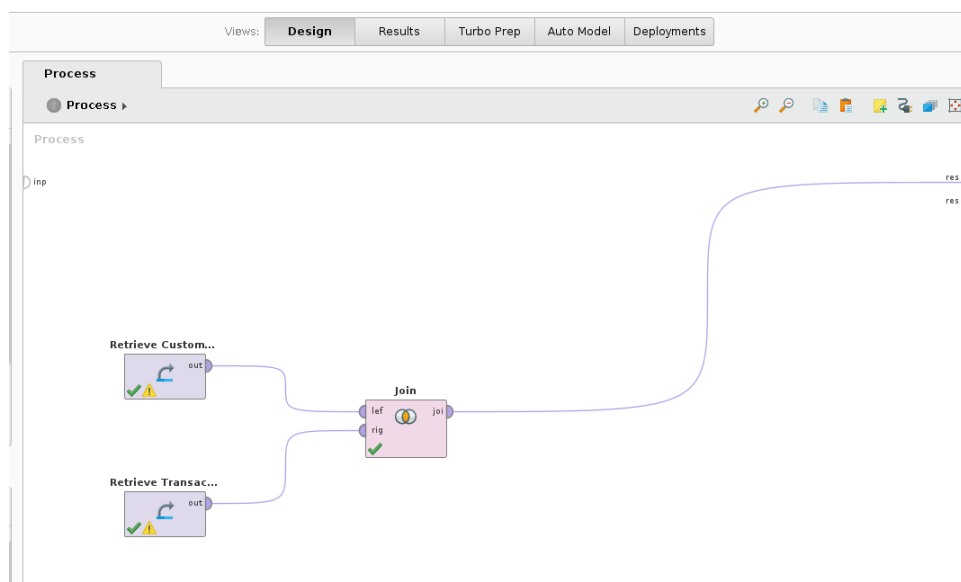
- **Operators**
 - This section contains various kinds of operators that are used to perform different operations.
 - Some of the operators are sort, join, aggregate, select attribute, etc.



- **Views**

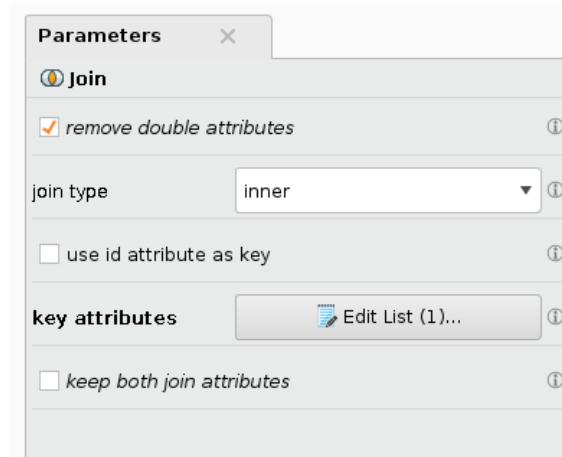
There are five kinds of view in Rapidminer:

- **Design**
 - It is the default view, which contains the **Process**.
 - In Process, we drag different datasets and operators and connect them.
- **Results**
 - After Running the process, we get the results in Results section, in which there are Data, Statistics, Visualization and annotation.
- **Turbo Prep**
 - It is used for blending and cleaning the data as well as to build data preparation process automatically.
- **Auto Model**
 - It is used to create models.
- **Deployments**
 - It is used to manage projects and deployed models.



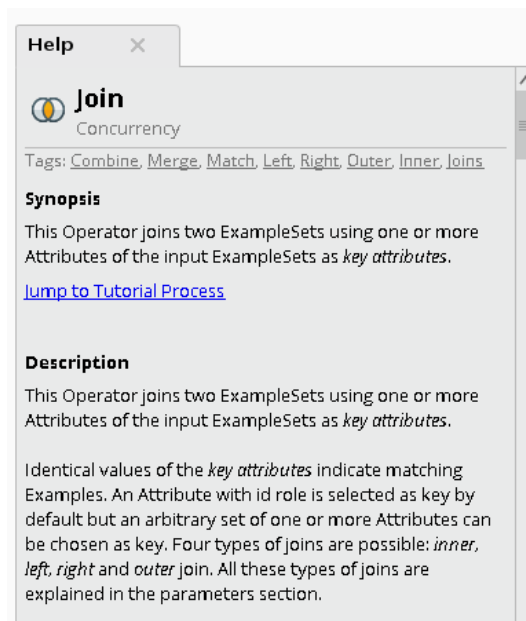
- **Parameters**

- This section is used to configure the behavior of different operators.
- It's content is displayed when we select a particular operator.
- By default, RapidMiner Studio shows only the more commonly used parameters.
- To see all of the available parameters, we have to click Show advanced parameters.



- **Help**

- It displays a help text for the current Operator.
- The content of the Help Panel is also context-dependent.
- If we select any Operator that is displayed in the Process Panel, the Help Panel displays a help text for that Operator.
- The Help Panel provides useful background informations.



These were the components of the RapidMiner that are used together to perform various operations on the datasets and get the meaningful result.

2.

a. Explain any 5 operators along with its usage and snapshots in rapidminer.

Data sets used : Customer Data and Transaction Data

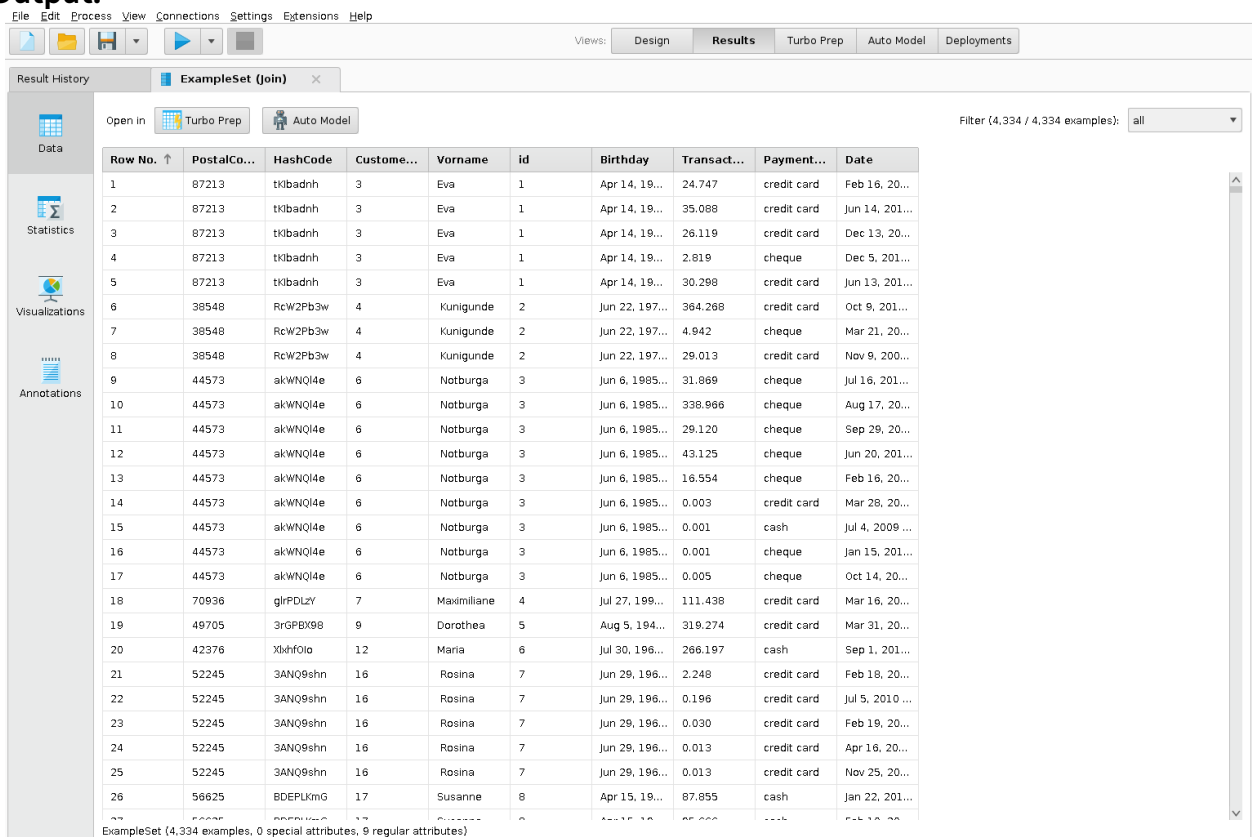
Path: Repository → Training Resources → Data → Hotel App Customer

(I). Join – This operator joins two Data Sets using one or more attributes that are key attributes in both the Data sets.

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Now, connect the output port of join operator to the result port.
- Press Run to execute the process.

Output:



Row No. ↑	PostalCo...	HashCode	Custome...	Vorname	id	Birthday	Transact...	Payment...	Date
1	87213	tkibadnh	3	Eva	1	Apr 14, 19...	24.747	credit card	Feb 16, 20...
2	87213	tkibadnh	3	Eva	1	Apr 14, 19...	35.088	credit card	Jun 14, 201...
3	87213	tkibadnh	3	Eva	1	Apr 14, 19...	26.119	credit card	Dec 13, 20...
4	87213	tkibadnh	3	Eva	1	Apr 14, 19...	2.819	cheque	Dec 5, 201...
5	87213	tkibadnh	3	Eva	1	Apr 14, 19...	30.298	credit card	Jun 13, 201...
6	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	364.268	credit card	Oct 9, 201...
7	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	4.942	cheque	Mar 21, 20...
8	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	29.013	credit card	Nov 9, 200...
9	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	31.869	cheque	Jul 16, 201...
10	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	338.966	cheque	Aug 17, 20...
11	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	29.120	cheque	Sep 29, 20...
12	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	43.125	cheque	Jun 20, 201...
13	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	16.554	cheque	Feb 16, 20...
14	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	0.003	credit card	Mar 28, 20...
15	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	0.001	cash	Jul 4, 2009 ...
16	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	0.001	cheque	Jan 15, 201...
17	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	0.005	cheque	Oct 14, 20...
18	70936	glrFDLzY	7	Maximiliane	4	Jul 27, 199...	111.438	credit card	Mar 16, 20...
19	49705	3rGPBX9B	9	Dorothea	5	Aug 5, 194...	319.274	credit card	Mar 31, 20...
20	42376	XlxfFOlo	12	Maria	6	Jul 30, 196...	266.197	cash	Sep 1, 201...
21	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	2.248	credit card	Feb 18, 20...
22	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	0.196	credit card	Jul 5, 2010 ...
23	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	0.030	credit card	Feb 19, 20...
24	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	0.013	credit card	Apr 16, 20...
25	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	0.013	credit card	Nov 25, 20...
26	56625	BDEPLKmG	17	Susanne	8	Apr 15, 19...	87.895	cash	Jan 22, 201...

(ii). Aggregate - The Aggregate operator creates a new ExampleSet from the input ExampleSet showing the results of the selected aggregation functions. Many aggregation functions are supported including SUM, COUNT, MIN, MAX, AVERAGE and many other similar functions known from SQL.

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Search for Aggregate operator and drag it to the process.
- Now, connect the output port of join operator to the input port of Aggregate operator.
- In the Parameters of Aggregate, go to edit list of group by attributes and select the postal code.
- Now, connect the output port of Aggregate to the result port.
- Press Run to execute the process.

Output:

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Result History ExampleSet (Aggregate) x

Open in Turbo Prep Auto Model

Filter (999 / 999 examples): all

Row No.	PostalCo...
1	1
2	9012
3	10807
4	10869
5	11286
6	13014
7	13223
8	13641
9	13826
10	15002
11	15024
12	15607
13	16287
14	16423
15	17193
16	17429
17	17715
18	18006
19	18418
20	18610
21	18888
22	19111
23	19635
24	19881
25	20041
26	20069

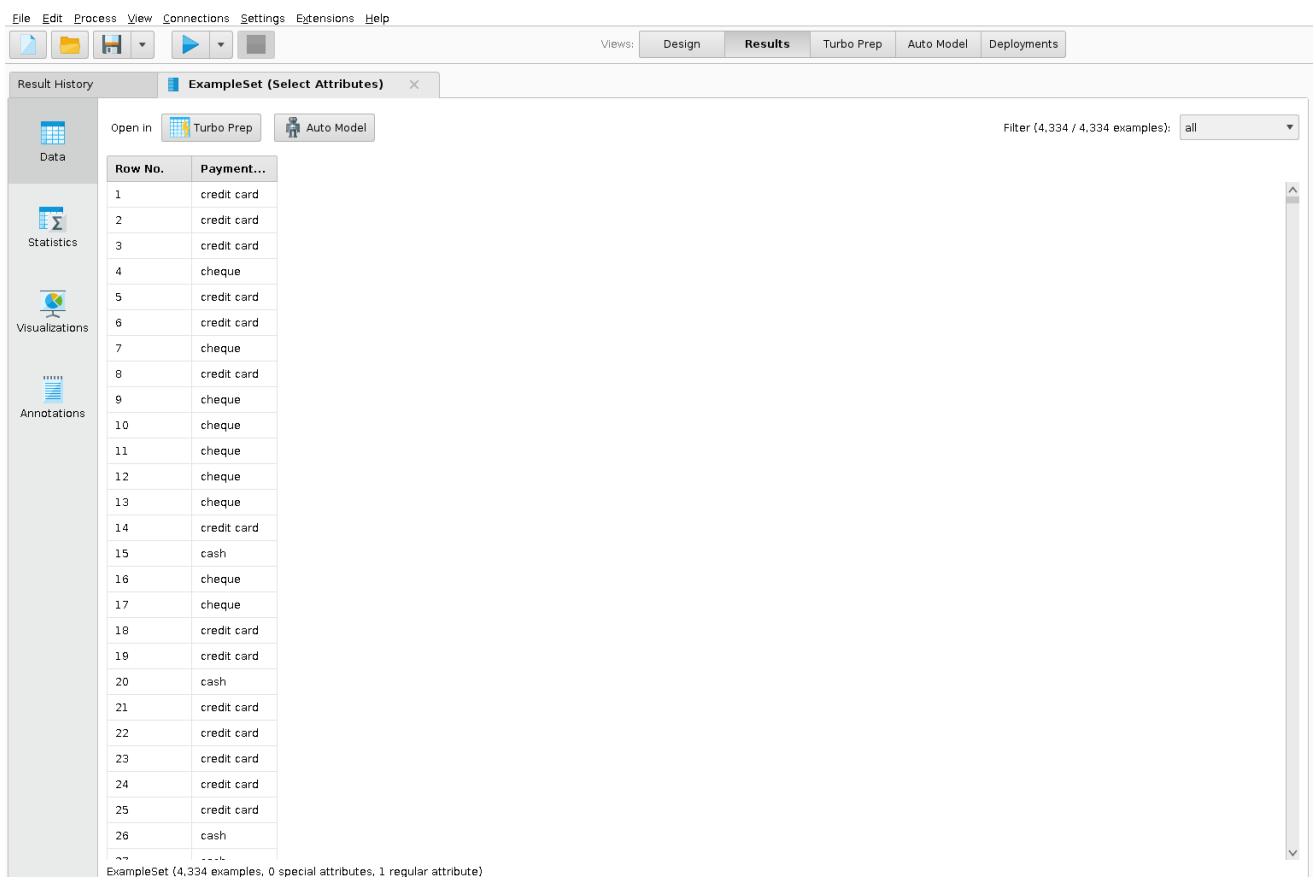
ExampleSet (999 examples, 0 special attributes, 1 regular attribute)

(iii). Select Attributes - This Operator selects a subset of Attributes of an ExampleSet and removes the other Attributes.

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Search for Select Attribute operator and drag it to the process.
- Now, connect the output port of join operator to the input port of Select Attribute operator.
- In the Parameters of Select Attribute, set attribute filter type to single and attribute to Payment Method.
- Now, connect the output port of Select Attribute to the result port.
- Press Run to execute the process.

Output:



File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Result History ExampleSet (Select Attributes) x

Open in Turbo Prep Auto Model

Filter (4,334 / 4,334 examples): all

Row No.	Payment...
1	credit card
2	credit card
3	credit card
4	cheque
5	credit card
6	credit card
7	cheque
8	credit card
9	cheque
10	cheque
11	cheque
12	cheque
13	cheque
14	credit card
15	cash
16	cheque
17	cheque
18	credit card
19	credit card
20	cash
21	credit card
22	credit card
23	credit card
24	credit card
25	credit card
26	cash

ExampleSet (4,334 examples, 0 special attributes, 1 regular attribute)

(iv). Sort - This operator sorts the input data set in ascending or descending order according to several attributes.

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Search for Sort operator and drag it to the process.
- Now, connect the output port of join operator to the input port of sort operator.
- In the Parameters of sort, go to the edit list of sort by and select one attribute and ascending or descending.
- Now, connect the output port of sort to the result port.
- Press Run to execute the process.

Output:

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Result History ExampleSet (Sort) x

Open in Turbo Prep Auto Model

Filter (4,334 / 4,334 examples): all

Row No.	PostalCo...	HashCode	Custome...	Vorname	id	Birthday	Transact...	Payment...	Date
1	80619	dVdjH2Uj	270	Scholastika	133	Dec 18, 19...	2.614	credit card	Mar 12, 20...
2	35583	MemBg2vN	47	Stefanie	27	Nov 9, 196...	43.247	credit card	Mar 13, 20...
3	48578	QZNLpGkp	189	Stephan	94	Mar 10, 19...	19.236	cash	Mar 17, 20...
4	80619	dVdjH2Uj	270	Scholastika	133	Dec 18, 19...	133.977	credit card	Mar 18, 20...
5	29907	LB3fgpFh	59	Notburga	29	Apr 10, 19...	0.000	cheque	Mar 27, 20...
6	55349	jPuStBk5	182	Babette	92	Apr 27, 19...	0.001	cash	Apr 20, 20...
7	62427	tpxwJRRU	765	Friedrich	414	Dec 4, 193...	19.130	cash	May 1, 200...
8	32949	s4mqiUx	344	Georgine	169	Jun 5, 1938...	153.954	cash	May 16, 20...
9	25257	VTRHXs1F	67	Leonhard	35	Nov 23, 19...	220.896	credit card	Jun 11, 200...
10	31280	80PGkjHR	507	Katharina	233	Apr 12, 19...	1.386	cash	Jun 21, 200...
11	35583	MemBg2vN	47	Stefanie	27	Nov 9, 196...	0.016	credit card	Jun 29, 200...
12	68734	THcgfMaD	239	Annemarie	119	Jul 10, 196...	6.664	cheque	Jul 4, 2008 ...
13	52623	Cu7KhuFs	202	Josefa	104	Dec 1, 196...	185.672	credit card	Jul 12, 200...
14	40973	DDEoEUqq	217	Scholastika	109	Mar 9, 198...	0.767	cash	Jul 18, 200...
15	40973	DDEoEUqq	217	Scholastika	109	Mar 9, 198...	0.000	credit card	Jul 22, 200...
16	39804	Hv5jreXj	139	Fridolin	72	Jun 15, 194...	131.101	credit card	Jul 22, 200...
17	59736	UVy9PnW	26	Waltraud	14	Sep 21, 19...	58.693	credit card	Jul 27, 200...
18	69484	yk7RmqC3	488	Cäcilia	221	Mar 5, 196...	127.414	credit card	Aug 1, 200...
19	35583	MemBg2vN	47	Stefanie	27	Nov 9, 196...	20.668	credit card	Aug 9, 200...
20	89598	3UkCKVJl	334	Rupert	170	Jun 24, 196...	1.562	cheque	Aug 18, 20...
21	69484	yk7RmqC3	488	Cäcilia	221	Mar 5, 196...	33.744	credit card	Sep 4, 200...
22	67034	7GDdhe6R	176	Margarete	88	Jun 1, 1962...	13.144	cash	Sep 7, 200...
23	48578	QZNLpGkp	189	Stephan	94	Mar 10, 19...	41.176	cash	Sep 17, 20...
24	29907	LB3fgpFh	59	Notburga	29	Apr 10, 19...	0.008	cheque	Sep 23, 20...
25	35583	MemBg2vN	47	Stefanie	27	Nov 9, 196...	4.128	credit card	Sep 25, 20...
26	25161	gFkFRdal	192	Melchior	95	May 15, 19...	10.785	cash	Sep 25, 20...
27	61072	81V8WU	288	Teodora	128	Dec 28, 19...	8.878	cheque	Sep 28, 20...

ExampleSet (4,334 examples, 0 special attributes, 9 regular attributes)

(v). Discretize by Binning - This operator discretizes the selected numerical attributes into user-specified number of bins. Bins of equal range are automatically generated, the number of the values in different bins may vary.

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Search for Discretize by Binning operator and drag it to the process.
- Now, connect the output port of join operator to the input port of sort operator.
- In the Parameters of Discretize by Binning, select attribute type to single, attribute to TransactionValue and number of bins to 3.
- Now, connect the output port of Discretize by Binning to the result port.
- Press Run to execute the process.

Output:

File Edit Process View Connections Settings Extensions Help

Views: Design Results Turbo Prep Auto Model Deployments

Result History ExampleSet (Discretize) X

Open in Turbo Prep Auto Model

Filter (4,334 / 4,334 examples): all

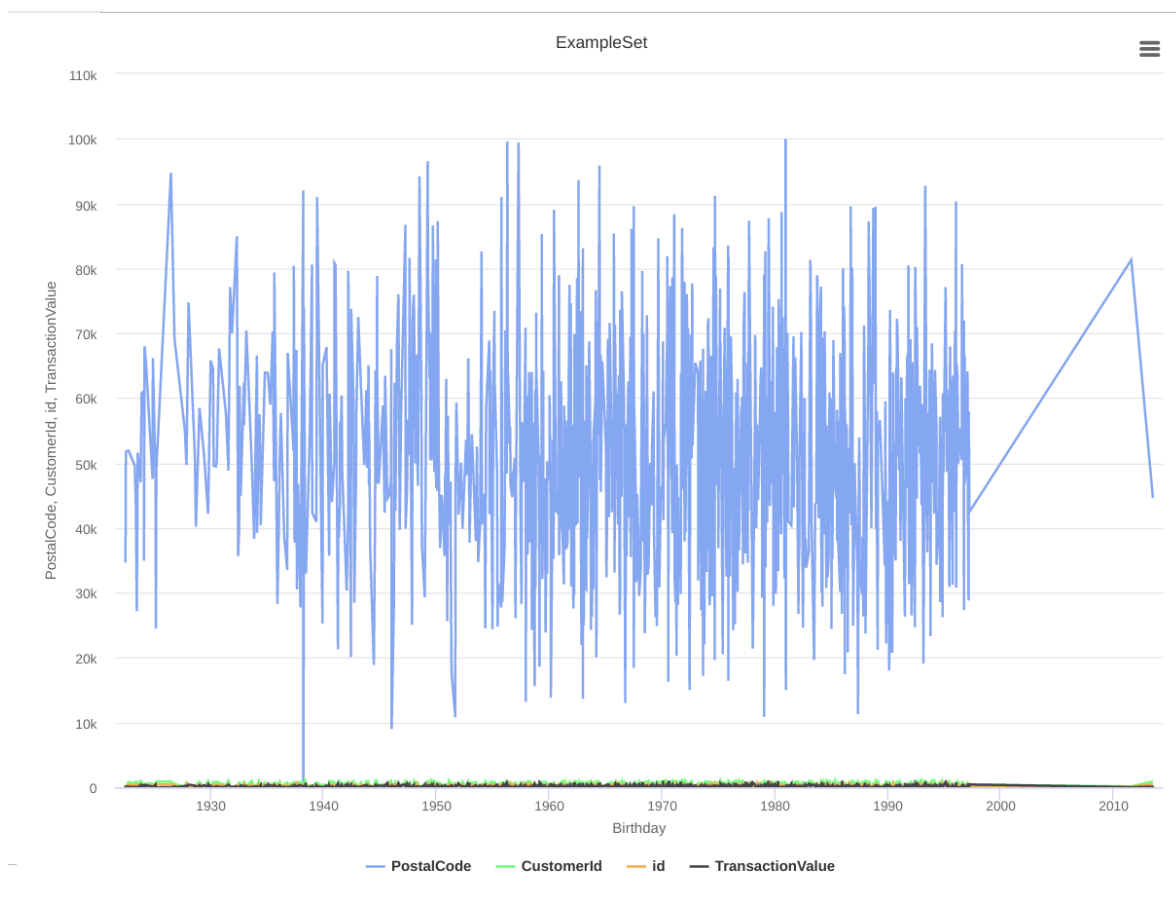
Row No.	Transact...	PostalCo...	HashCode	Custome...	Vorname	id	Birthday	Payment...	Date
1	range1 [-∞ ...	87213	tklbadnh	3	Eva	1	Apr 14, 19...	credit card	Feb 16, 20...
2	range1 [-∞ ...	87213	tklbadnh	3	Eva	1	Apr 14, 19...	credit card	Jun 14, 201...
3	range1 [-∞ ...	87213	tklbadnh	3	Eva	1	Apr 14, 19...	credit card	Dec 13, 20...
4	range1 [-∞ ...	87213	tklbadnh	3	Eva	1	Apr 14, 19...	cheque	Dec 5, 201...
5	range1 [-∞ ...	87213	tklbadnh	3	Eva	1	Apr 14, 19...	credit card	Jun 13, 201...
6	range2 [35...	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	credit card	Oct 9, 201...
7	range1 [-∞ ...	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	cheque	Mar 21, 20...
8	range1 [-∞ ...	38548	RcW2Pb3w	4	Kunigunde	2	Jun 22, 197...	credit card	Nov 9, 200...
9	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Jul 16, 201...
10	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Aug 17, 20...
11	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Sep 29, 20...
12	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Jun 20, 201...
13	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Feb 16, 20...
14	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	credit card	Mar 28, 20...
15	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cash	Jul 4, 2009 ...
16	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Jan 15, 201...
17	range1 [-∞ ...	44573	akWNQl4e	6	Notburga	3	Jun 6, 1985...	cheque	Oct 14, 20...
18	range1 [-∞ ...	70936	glrPDZy	7	Maximiliane	4	Jul 27, 199...	credit card	Mar 16, 20...
19	range1 [-∞ ...	49705	3rGPBX98	9	Dorothea	5	Aug 5, 194...	credit card	Mar 31, 20...
20	range1 [-∞ ...	42376	Xbhf0to	12	Maria	6	Jul 30, 196...	cash	Sep 1, 201...
21	range1 [-∞ ...	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	credit card	Feb 18, 20...
22	range1 [-∞ ...	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	credit card	Jul 5, 2010 ...
23	range1 [-∞ ...	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	credit card	Feb 19, 20...
24	range1 [-∞ ...	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	credit card	Apr 16, 20...
25	range1 [-∞ ...	52245	3ANQ9shn	16	Rosina	7	Jun 29, 196...	credit card	Nov 25, 20...
26	range1 [-∞ ...	56625	BDEPLKmG	17	Susanne	8	Apr 15, 19...	cash	Jan 22, 201...
27	range1 [-∞ ...	56625	BDEPLKmG	17	Susanne	8	Apr 15, 19...	cash	Feb 16, 20...

ExampleSet (4,334 examples, 0 special attributes, 9 regular attributes)

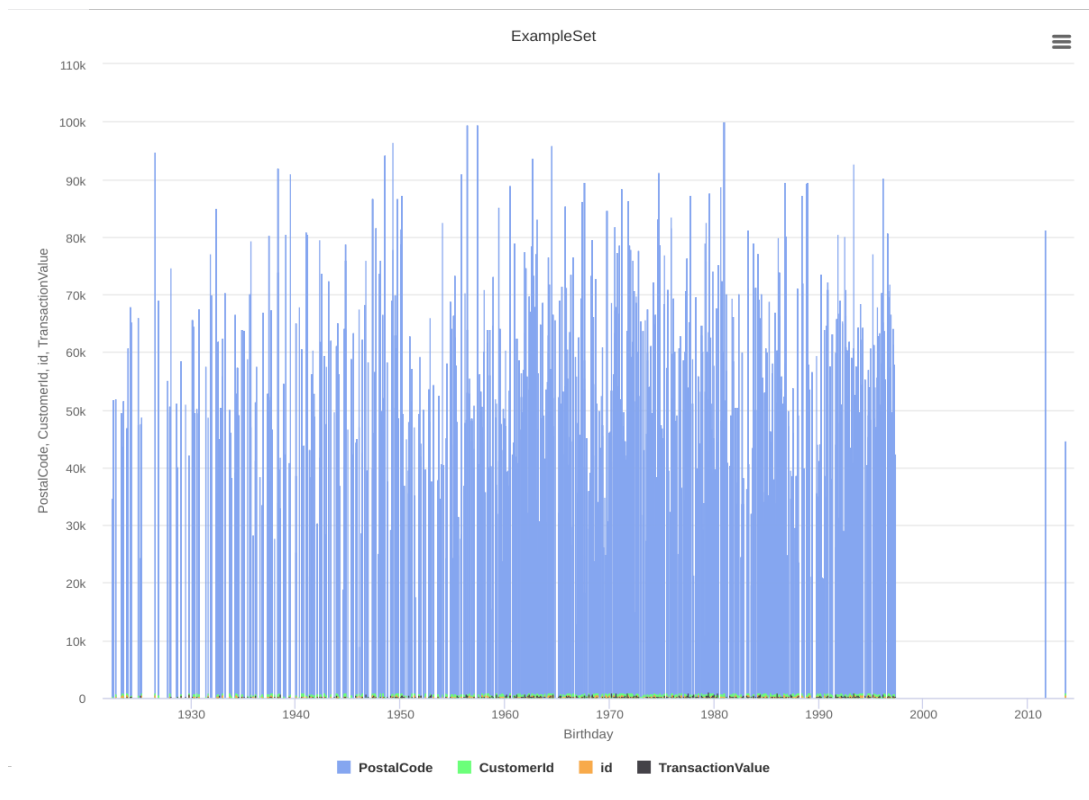
2b. What do you understand by graphical representation and statistics in rapidminer. Attach any 5 different types of graph and their interpretation.

Graphical Representation:

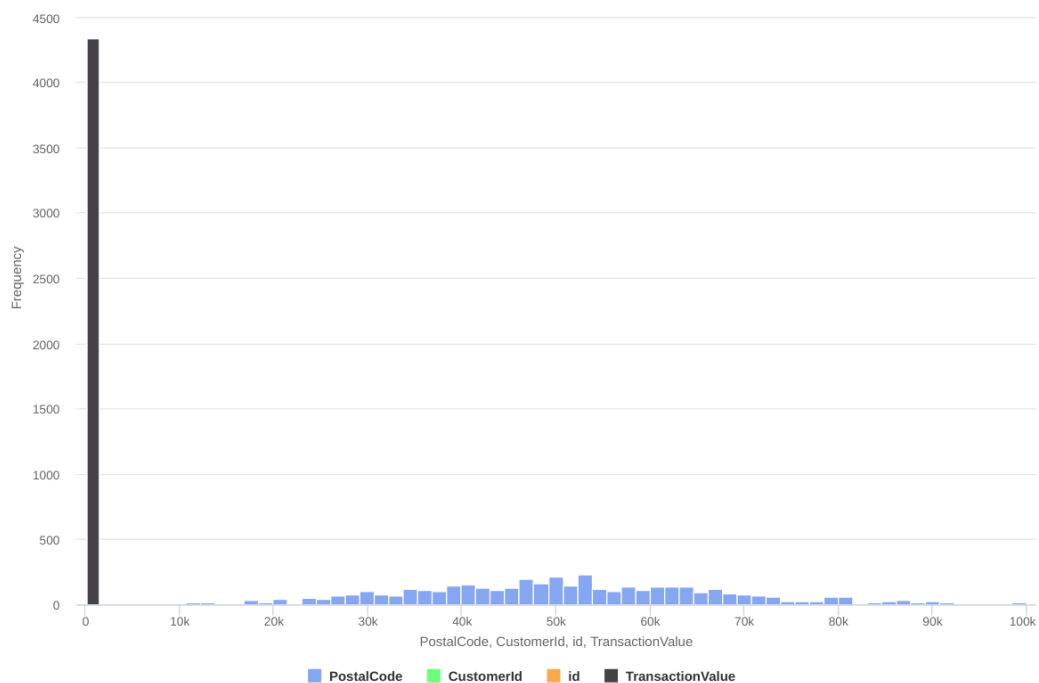
- It is a way of analysing numerical data.
- It exhibits the relation between data, ideas, information and concepts in a diagram.
- It is easy to understand and it is one of the most important learning strategies.
- It always depends on the type of information in a particular domain.
- There are different types of graphical representation. Some of them are as follows:
- **Data sets used :** Customer Data and Transaction Data
- **Path:** Repository → Training Resources → Data → Hotel App Customer
- **Line Graphs** – It is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. Line graph or the linear graph is used to display the continuous data and it is useful for predicting future events over time.



- **Bar Graphs** – Bar Graph is used to display the category of data and it compares the data using solid bars to represent the quantities.

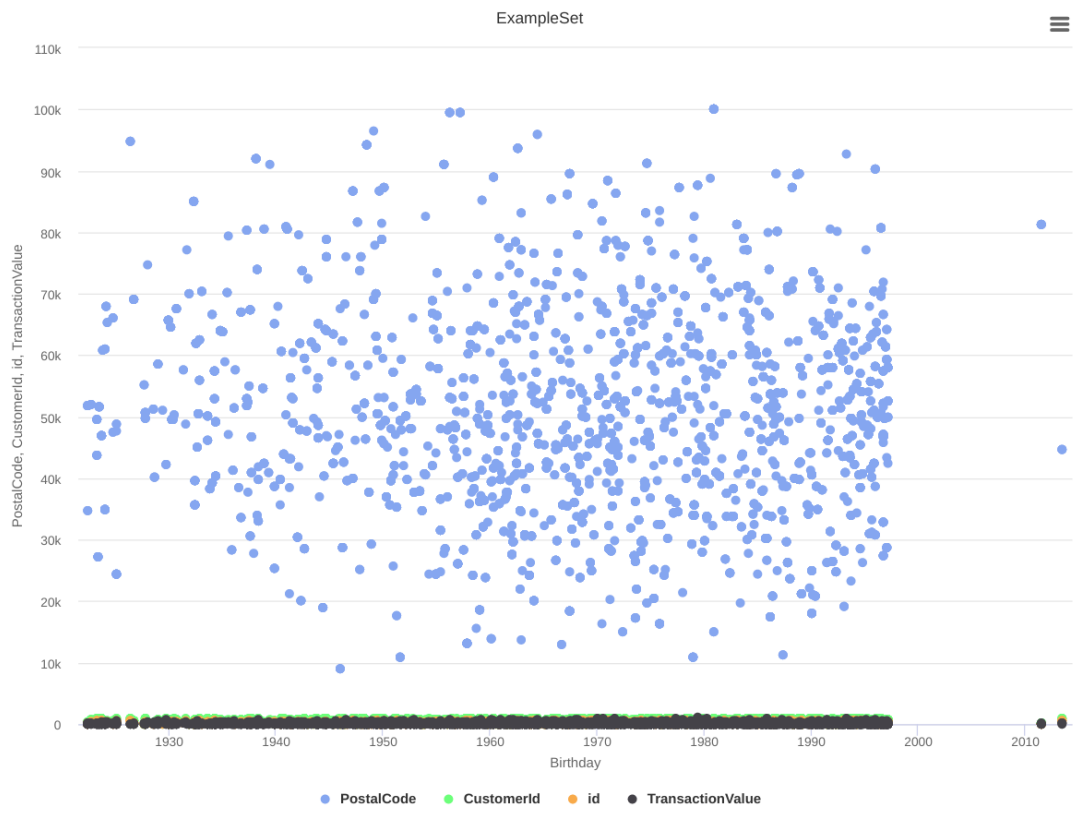


- **Histograms** – The graph that uses bars to represent the frequency of numerical data that are organised into intervals. Since all the intervals are equal and continuous, all the bars have the same width.

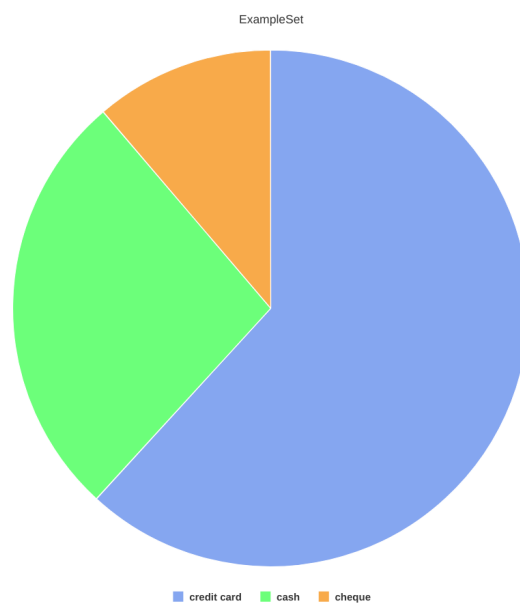


- **Scatter plot** - A scatter plot uses dots to represent values for two different numeric variables. The position of each dot on the horizontal and vertical axis indicates values for an individual data point. Scatter plots are used to observe

relationships between variables.



- **Pie Chart** – It shows the relationships of the parts of the whole. The circle is considered with 100% and the categories occupied is represented with that specific percentage like 15%, 56%, etc.



Statistics in RapidMiner:

Statistics in Rapidminer gives the output after analyzing the data set, which contains class labels, no of missing values, minimum, maximum and average of a particular attribute.

Name	Type	Missing	Statistics			Filter (9 / 9 attributes): <input type="text" value="Search for Attribute"/>
▼ PostalCode	Integer	0	Min 1	Max 99999	Average 51297.960	
▼ HashCode	Nominal	0	Least zPLl3DcR (1)	Most cujNb6GQ (20)	Values cujNb6GQ (20), wHh275mF (20), ...[869 more]	
▼ CustomerId	Integer	0	Min 1	Max 1000	Average 489.938	
▼ Vorname	Nominal	0	Least Benedikt (3)	Most Georg (80)	Values Georg (80), Leopold (64), ...[166 more]	
▼ id	Integer	0	Min 1	Max 551	Average 249.098	
▼ ⚠ Birthday	Date-time	1	Earliest date Jun 8, 1922 5:55 AM	Latest date Jun 24, 2013 4:55 AM	Duration 33253d 23h 0m 0s	
▼ TransactionValue	Real	0	Min 0.000	Max 1051.191	Average 77.944	
▼ PaymentMethod	Nominal	0	Least cheque (487)	Most credit card (2679)	Values credit card (2679), cash (1168), ...[1 more]	
▼ Date	Date-time	0	Earliest date Mar 12, 2008 5:46 PM	Latest date Feb 27, 2014 7:52 AM	Duration 2177d 14h 5m 34s	

3. What do you mean by preprocessing in Data Mining. Explain any 5 operations to handle missing values and attach screenshots of same.

Preprocessing in Data Mining:

- Data preprocessing is the process of transforming raw data into an understandable format. It is also an important step in data mining as we cannot work with raw data.
- The quality of the data should be checked before applying machine learning or data mining algorithms.
- Preprocessing of data is mainly to check the data quality. The quality can be checked by the following
 - **Accuracy:** To check whether the data entered is correct or not.
 - **Completeness:** To check whether the data is available or not recorded.
 - **Consistency:** To check whether the same data is kept in all the places that do or do not match.
 - **Timeliness:** The data should be updated correctly.
 - **Believability:** The data should be trustable.
 - **Interpretability:** The understandability of the data.
- Following are some major tasks in pre-processing:
 - **Data Cleaning** – It is the process to remove incorrect data, incomplete data and inaccurate data from the datasets, and it also replaces the missing values.
 - **Data Integration** – It is the process of combining multiple sources into a single dataset.
 - **Data reduction** - This process helps in the reduction of the volume of the data which makes the analysis easier yet produces the same or almost the same result.
 - **Data transformation** - The change made in the format or the structure of the data is called data transformation.

Handling missing values

There are several ways to handle missing values, some of them are:

- **Ignore the data row.**
 - If there are lots of missing values in a particular row, we just ignore that row.
- **Use a global constant to fill in for missing values.**
 - We just put a common value to all those missing values.
- **Use attribute mean.**
 - Replacing missing values of an attribute with the mean (or median if its discrete) value for that attribute in the database.

- **Use attribute mean for all samples belonging to the same class.**
 - Instead of using the mean (or median) of a certain attribute calculated by looking at all the rows in a database, we can limit the calculations to the relevant class to make the value more relevant to the row we're looking at.
- **Use a data mining algorithm to predict the most probable value.**
 - The value can be determined using regression, inference based tools using Bayesian formalism, decision trees, clustering algorithms (K-Mean\Median etc.).

We can use **Replace Missing values** operator to perform these operations to handle missing values in the datasets.

- **Data sets used :** Customer Data and Transaction Data
- **Path:** Repository → Training Resources → Data → Hotel App Customer

Steps:

- Drag the Customer Data and Transaction Data into the process.
- Search for sort operator and drag it into the process.
- Connect the out ports of Customer Data and Transaction Data into the input port of Join operator.
- Click on join to select it.
- In the Parameters of join, go to the edit list of key attributes.
- Select the key attributes in both side, here CustomerId.
- Hover your mouse pointer over the output port and wait for the tooltip to show the meta data.
- Press F3 while the tooltip is shown. It is now transformed into a window and you can scroll down to see the information about all the columns.
- Check for columns with missing values.
- Add a new operator **Select Attributes**.
- Connect the new operator to the **Retrieve** operator and the output to the result port "res" on the right.
- In Parameters, change the **attribute filter type** to *Subset* and select all attributes but *Birthday*. This means that this will be removed by the operator.
- Run the process.
- Click on the **Statistics** tab and check which attributes with missing values are still left.
- Here, in this dataset, there's no missing value left, but to understand this, unselect the Birthday and perform the below operations.
- Search for the operator **Replace Missing Values** and add it to the process. You can drop it on the connection line between **Select Attributes** and the result port (move the mouse pointer until the connection is highlighted before you drop the new operator). This way you do not need to manually reconnect the operators.
- In the Parameters of this operator, use *single* for the **attribute filter type** and select *Birthday* for **attribute**.
- Run the process again and interpret results

Name	Type	Missing	Statistics			Filter (9 / 9 attributes): <input type="text" value="Search for Attribute"/>
▼ Birthday	Date time	1	Earliest date Jun 8, 1922 5:55 AM	Latest date Jun 24, 2013 4:55 AM	Duration 33253d 23h 0m 0s	
▼ PostalCode	Integer	0	Min 1	Max 99999	Average 51297.960	
▼ HashCode	Polynomial	0	Least zPLl3DcR (1)	Most cuJNb6GQ (20)	Values cuJNb6GQ (20), wHh275mF (20), ...[869 more]	
▼ CustomerId	Integer	0	Min 1	Max 1000	Average 489.938	
▼ Vorname	Polynomial	0	Least Benedikt (3)	Most Georg (80)	Values Georg (80), Leopold (64), ...[166 more]	
▼ id	Integer	0	Min 1	Max 551	Average 249.098	
▼ TransactionValue	Real	0	Min 0.000	Max 1051.191	Average 77.944	
▼ PaymentMethod	Polynomial	0	Least cheque (487)	Most credit card (2679)	Values credit card (2679), cash (1168), ...[1 more]	
▼ Date	Date time	0	Earliest date Mar 12, 2008 5:46 PM	Latest date Feb 27, 2014 7:52 AM	Duration 2177d 14h 5m 34s	

- Search for **Filter Examples** and drop it again on the connection line to the result port. In case you miss the connection, you can of course always manually connect the operator instead.
- Note the **link** at the bottom of the **Parameters** panel which shows/hides advanced parameters Click on **Show advanced parameters** to display all operator parameters.
- New parameters should appear. Set the **condition class** to *no_missing_attributes*.
- Run the process again.
- Now there's no missing value left in the dataset.

Name	Type	Missing	Statistics			Filter (9 / 9 attributes): <input type="text" value="Search for Attribute"/>
▼ Birthday	Date time	0	Earliest date Jun 8, 1922 5:55 AM	Latest date Jun 24, 2013 4:55 AM	Duration 33253d 23h 0m 0s	
▼ PostalCode	Integer	0	Min 1	Max 99999	Average 51291.322	
▼ HashCode	Polynomial	0	Least EOGj6tDH (0)	Most cuJNb6GQ (20)	Values cuJNb6GQ (20), wHh275mF (20), ...[869 more]	
▼ CustomerId	Integer	0	Min 1	Max 1000	Average 490.048	
▼ Vorname	Polynomial	0	Least Benedikt (3)	Most Georg (80)	Values Georg (80), Leopold (64), ...[166 more]	
▼ id	Integer	0	Min 1	Max 551	Average 249.153	
▼ TransactionValue	Real	0	Min 0.000	Max 1051.191	Average 77.879	
▼ PaymentMethod	Polynomial	0	Least cheque (487)	Most credit card (2678)	Values credit card (2678), cash (1168), ...[1 more]	
▼ Date	Date time	0	Earliest date Mar 12, 2008 5:46 PM	Latest date Feb 25, 2014 8:47 AM	Duration 2175d 15h 0m 36s	