

Big Data & Data Analytics

Prodi Psikologi Kelas III & IV

PERTEMUAN 4

Data Cleansing

- Diskritisasi Dengan Binning
- Menghilangkan Missing Value
- Mengidentifikasi Outlier

Diskritisasi Dengan Binning

Menggunakan Data “Titanic” Yang ada pada Repository di folder Sampel.

- ❑ Retrieve data “Titanic” lalu drag ke Panel Process
- ❑ Definisikan tipe attribute & Role :
 - Pilih Operator “Discretize by Binning” – Hubungkan dengan data Titanic
 - Klik operator “Discretize by Binning” lalu isikan pada panel parameter
 - a. Attribute filter Type : Pilih Single (Karena hanya 1 yang akan kita kelompokkan)
 - b. Attribute to : Pilih Age
 - c. Number of Bins : Isilah dengan 3

Result History

ExampleSet (Discretize)

ExampleSet (Filter Examples (2))



Data



Statistics



Visualizations



Annotations

Open in



Turbo Prep



Auto Model

Filter (1,309 / 1,309 examples):

all

Row No.	Age	Passenger ...	Name	Sex	No of Sibling...	No of Parent...	Ticket Num
1	range2 [26.7...	First	Allen, Miss. E...	Female	0	0	24160
2	range1 [-∞ - 2...	First	Allison, Mast...	Male	1	2	113781
3	range1 [-∞ - 2...	First	Allison, Miss. ...	Female	1	2	113781
4	range2 [26.7...	First	Allison, Mr. H...	Male	1	2	113781
5	range1 [-∞ - 2...	First	Allison, Mrs. ...	Female	1	2	113781
6	range2 [26.7...	First	Anderson, Mr....	Male	0	0	19952
7	range3 [53.3...	First	Andrews, Mis...	Female	1	0	13502
8	range2 [26.7...	First	Andrews, Mr. ...	Male	0	0	112050
9	range2 [26.7...	First	Appleton, Mrs...	Female	2	0	11769
10	range3 [53.3...	First	Artagaveytia, ...	Male	0	0	PC 17609
11	range2 [26.7...	First	Astor, Col. Jo...	Male	1	0	PC 17757
12	range1 [-∞ - 2...	First	Astor, Mrs. Jo...	Female	1	0	PC 17757
13	range1 [-∞ - 2...	First	Aubart, Mme. ...	Female	0	0	PC 17477
14	range1 [-∞ - 2...	First	Barber, Miss. ...	Female	0	0	19877

ExampleSet (1,309 examples, 0 special attributes, 12 regular attributes)

Repository

+ Import Data

- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting

- processes
- Templates
- Time Series
- Tutorials

- Local Repository (Local)
- DB (Legacy)

❑ Definisikan attribute yang digunakan untuk memprediksi :

- Pilih Operator “Set Role” – hubungkan dengan operator “Discretize by Binning”
- Klik operator “Set Role ” lalu isikan pada panel parameter
 - a. Attribute Name : Pilih Survived
 - b. Role target : Ubah menjadi “label”

❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio ▾

Result History

ExampleSet (Set Role) ×

ExampleSet (Filter Examples (2)) ×



Data



Statistics



Visualizations



Annotations

Open in



Turbo Prep



Auto Model

Filter (1,309 / 1,309 examples):

all ▾

Row No.	Survived	Age	Passenger ...	Name	Sex	No of Sibling...	No of Pare
1	Yes	range2 [26.7...	First	Allen, Miss. E...	Female	0	0
2	Yes	range1 [-∞ - 2...	First	Allison, Mast...	Male	1	2
3	No	range1 [-∞ - 2...	First	Allison, Miss. ...	Female	1	2
4	No	range2 [26.7...	First	Allison, Mr. H...	Male	1	2
5	No	range1 [-∞ - 2...	First	Allison, Mrs. ...	Female	1	2
6	Yes	range2 [26.7...	First	Anderson, Mr....	Male	0	0
7	Yes	range3 [53.3...	First	Andrews, Mis...	Female	1	0
8	No	range2 [26.7...	First	Andrews, Mr. ...	Male	0	0
9	Yes	range2 [26.7...	First	Appleton, Mrs...	Female	2	0
10	No	range3 [53.3...	First	Artagaveytia, ...	Male	0	0
11	No	range2 [26.7...	First	Astor, Col. Jo...	Male	1	0
12	Yes	range1 [-∞ - 2...	First	Astor, Mrs. Jo...	Female	1	0
13	Yes	range1 [-∞ - 2...	First	Aubart, Mme. ...	Female	0	0
14	Yes	range1 [-∞ - 2...	First	Barber, Miss. ...	Female	0	0

ExampleSet (1,309 examples, 1 special attribute, 11 regular attributes)

Repository ×

+ Import Data

- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting

▶ processes

▶ Templates

▶ Time Series

▶ Tutorials

▶ Local Repository (Local)

▶ DB (Legacy)

Menghilangkan Attribute/Kolom Dengan Banyak Missing Value

Menggunakan Dataset “Titanic” yang ada pada Repository Sampel.

- ❑ Retrieve Data Titanic ke Panel Process
- ❑ Letakkan Pointer ke Data Titanic lalu perhatikan bahwa ada 5 attribute yang memuat Missing Value yaitu Age, Passenger Fare, Cabin, Port of Embarkation dan Life Boat.
- ❑ Pilih operator “Select Attributes” lalu Hubungkan dengan data Titanic
- ❑ Klik Operator Select Attributes, lalu pada panel Parameter :
 - Attribute Filter Type : Pilih Subset (karena akan dihilangkan beberapa attribute)
 - Pada Attributes – klik Select Attribute – pindahkan semua attribute kecuali Cabin & Life Boat
 - Klik Apply
- ❑ Hubungkan ke Port Result – Klik ➤

	Name	Type	Missing	Filter (10 / 10 attributes): <input type="text" value="Search for Attributes"/>	
Data	Age	Real	263	Min 0.167	Max 80
Statistics	No of Siblings or Spouses on B...	Integer	0	Min 0	Max 8
Visualizations	No of Parents or Children on B...	Integer	0	Min 0	Max 9
Annotations	Ticket Number	Polynomial	0	Least W/C 14208 (1)	Most CA. 2343 (11)
	Passenger Fare	Numeric	1	Min 0	Max 512.329
	Port of Embarkation	Polynomial	2	Least Queenstown (123)	Most Southampton (9)
	Survived	Binominal	0	Negative Yes	Positive No

Pada Statistics terlihat masih ada Attribute yang mengandung Missing Value yaitu Age, Passenger Fare & Port of Embarkation

Me-Replace Missing Value.

- ❑ Pilih Operator “Replace Missing Values” – Hubungkan dengan Operator “Select Attributes”
- ❑ Klik Operator Replace Missing Values, lalu pada panel Parameter :
 - Attribute Filter Type : Pilih Single (karena akan dihilangkan 1 attribute)
 - Pada Attributes – pilih Age
- ❑ Hubungkan ke Port Result – Klik ►

Pada Statistics terlihat bahwa pada Attribute Age sudah tidak memuat Missing Value. Tetapi masih ada attribute Passenger Fare & Port of Embarkation yang memuat missing value dan harus dihilangkan.

Me-Remove Examples/Baris Yang Memuat Missing Value.

- ❑ Pilih Operator “Filter Examples” – Hubungkan dengan Operator “Replace Missing Values”
- ❑ Klik Operator Filter Examples, lalu pada panel Parameter :
 - Condition class : pilih “no_missing_attributes”
- ❑ Hubungkan ke Port Result – Klik ▷

Pada Statistics terlihat bahwa seluruh attribute sudah tidak memuat missing value.

Mengidentifikasi Outlier

➤ **Tujuan:** Mengidentifikasi kasus yang tak biasa (adanya outlier,dsb) dan membuangnya dari data set.

➤ **Menggunakan Dataset “Titanic” yang sudah dilakukan cleansing.**

❑ Pilih operator “Select Attributes” - Hubungkan dengan operator “Replace Missing Values”

❑ Klik Operator Select Attributes, lalu pada panel Parameter :

- Attribute Filter Type : Pilih Subset (karena akan dihilangkan beberapa attribute)
- Pada Attributes – klik Select Attribute – pindahkan semua attribute kecuali Name & Ticket Number
- Klik Apply

❑ Hubungkan ke Port Result – Klik ▷

Diperoleh Dataset yang hanya memuat attribute-attribute yang berkontribusi dalam deteksi outlier.

Melakukan Normalisasi Data

➤ **Normalisasi Data** adalah proses mengubah attribute menggunakan rentang nilai yang sama sedemikian sehingga data tersebut akan menyebar normal.

❑ Pilih Operator “Normalized” – Hubungkan dengan Operator “Select Attributes”

❑ Pilih Operator “Detect Outlier (Distances)” - Isikan Number of Neighbour =10, Number of outliers =10, distance function= Euclidian distance – Hubungkan dengan Operator “Normalized”

❑ Pilih Operator “Filter Examples” – Hubungkan dengan Operator “Detect Outlier (Distances)”

❑ Klik Operator Filter Examples, lalu pada panel Parameter :

- Filter : Klik Add a New Filter – Pilih Outlier (Kiri) , Equals (tengah), False (Kanan)

❑ Hubungkan ke Port Result – Klik ▷



Result History ExampleSet (Filter Examples (2))

Open in Turbo Prep Auto Model

Data

Statistics

Pada Outlier, False berarti pada baris tersebut tidak memuat outlier

Data di Attribute Age, No of Sibling, No of Parents, Passenger Fare Hasil Normalisasi sehingga menyebar Normal standar di sekitar mean = 0 dengan standar deviasi 1

Filter (1,296 / 1,296 examples): all

Row No.	outlier	Age	No of Sibling...	No of Parent...	Passenger F...	Passenger ...	Sex
528	false	0.109	-0.480	-0.445	-0.391	Second	Female
529	false	0.325	-0.480	-0.445	-0.391	Second	Male
530	false	0.013	-0.480	-0.445	-0.396	Second	Male
	false	-0.220	-0.480	-0.445	-0.351	Second	Male
	false	-2.168	0.480	0.709	-0.140	Second	Female
	false	-1.700	0.480	0.709	-0.140	Second	Female
	false	0.247	-0.480	1.863	-0.140	Second	Female
	false	0.481	-0.480	-0.445	-0.439	Second	Male
	false	0.325	0.480	-0.445	-0.236	Second	Male
	false	0.013	2.398	-0.445	-0.236	Second	Female
	false	-0.142	-0.480	-0.445	-0.391	Second	Female
	false	-0.532	-0.480	-0.445	-0.351	Second	Male
	false	-2.259	0.480	0.709	-0.280	Second	Male
	false	-2.090	0.480	0.709	-0.280	Second	Male
	false	0.454	1.430	2.017	0.280	Second	Female

ExampleSet (1,296 examples, 1 special attribute, 8 regular attributes)

- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting
- processes
- Templates
- Time Series
- Tutorials
- Local Repository (Local)
- DB (Legacy)

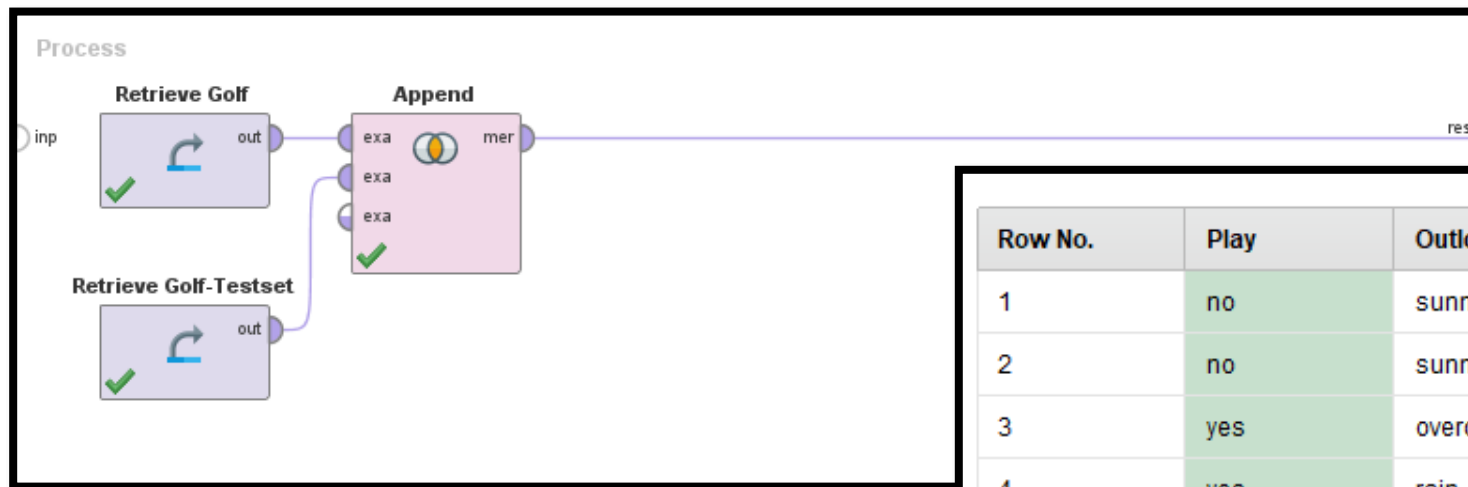
Data Blending

Menggabungkan Data Set dengan Join, Append, Set Minus, Intersect, Union, Cartesian Product

Menggabungkan Data Set Dengan Appends

- **Operator Appends** digunakan untuk menggabungkan beberapa data set yang memiliki attribute yang sama sehingga barisnya (Example) bertambah.
- ❑ Retrieve data “Golf” (14 baris) dan “Golf Testset” (14 baris) ke panel process
- ❑ Pilih operator “Appends” – Hubungkan dengan kedua data tersebut
- ❑ Klik ▷

Hasil : dataset dengan 28 baris



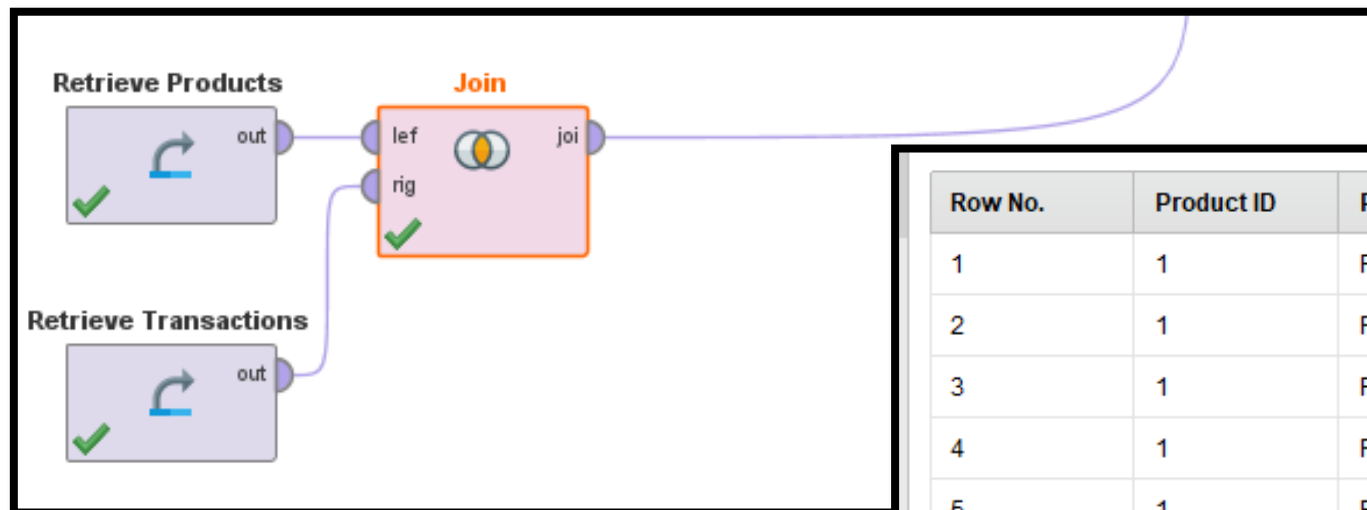
Hasil →

Row No.	Play	Outlook	Temperature	Humidity	Wind
1	no	sunny	85	85	false
2	no	sunny	80	90	true
3	yes	overcast	83	78	false
4	yes	rain	70	96	false
5	yes	rain	68	80	false
6	no	rain	65	70	true
7	yes	overcast	64	65	true
8	no	sunny	72	95	false
9	yes	sunny	69	70	false
10	yes	rain	75	80	false
11	yes	sunny	75	70	true
12	yes	overcast	72	90	true
13	yes	overcast	81	75	false
14	no	rain	71	80	true
15	yes	sunny	85	85	false

ExampleSet (28 examples, 1 special attribute, 4 regular attributes)

Menggabungkan Data Set Dengan Join

- **Operator Join** digunakan untuk menggabungkan beberapa data set yang memiliki key attribute (seperti primary key) yang sama sehingga attributenya bertambah.
- ❑ Retrieve data “Transactions” dan “Products” (14 baris) yang keduanya memiliki key attribute yang sama yaitu product id.
- ❑ Pilih operator “Joins” – Hubungkan dengan kedua data tersebut
- ❑ Pada parameter – Key Attributes : Klik Edit List – Isikan Left Key & Right Key dengan Product ID
- ❑ Klik ▷



Hasil

Row No.	Product ID	Product Na...	Product Cat...	Price	Customer ID	Amount
1	1	Repressitol	9	35.540	220	4
2	1	Repressitol	9	35.540	125	3
3	1	Repressitol	9	35.540	352	2
4	1	Repressitol	9	35.540	408	3
5	1	Repressitol	9	35.540	334	1
6	1	Repressitol	9	35.540	31	2
7	1	Repressitol	9	35.540	62	3
8	1	Repressitol	9	35.540	79	3
9	1	Repressitol	9	35.540	64	1
10	1	Repressitol	9	35.540	191	3
11	1	Repressitol	9	35.540	365	4
12	1	Repressitol	9	35.540	231	4
13	1	Repressitol	9	35.540	566	4
14	1	Repressitol	9	35.540	99	2
15	1	Repressitol	9	35.540	447	5

ExampleSet (2,328 examples, 0 special attributes, 6 regular attributes)

Menggabungkan Data Set Dengan Union

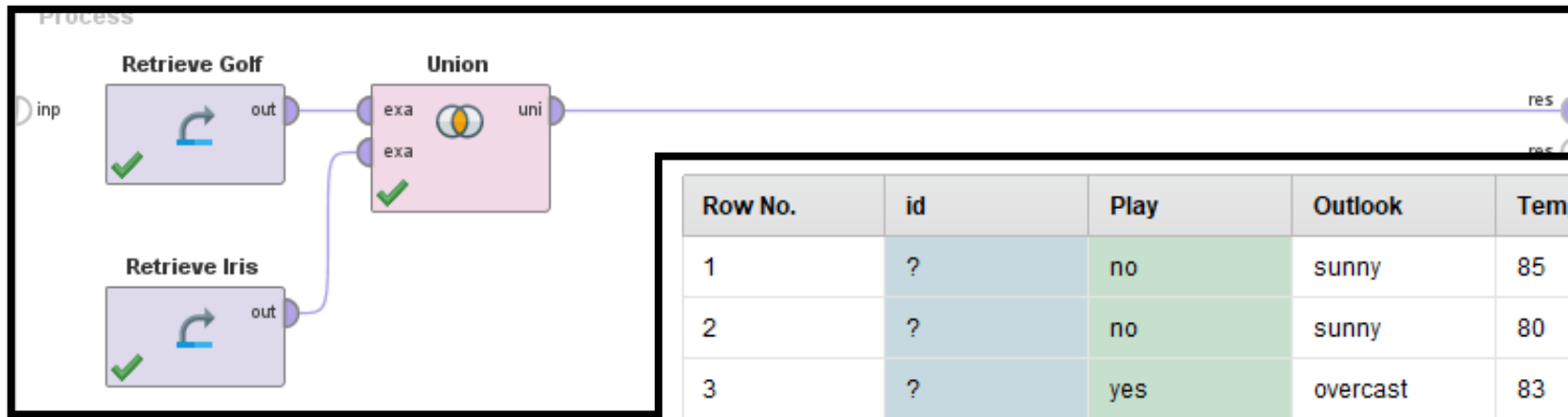
➤ **Operator Union** digunakan untuk menggabungkan beberapa data set yang memiliki attribute yang sama sehingga baris (Example) & Kolom (Attribute) bertambah.

- ❑ Retrieve data “Golf” dan “Iris” ke panel process

- ❑ Pilih operator “Union” – Hubungkan dengan kedua data tersebut

- ❑ Klik ▷

Hasil : dataset dengan 28 baris



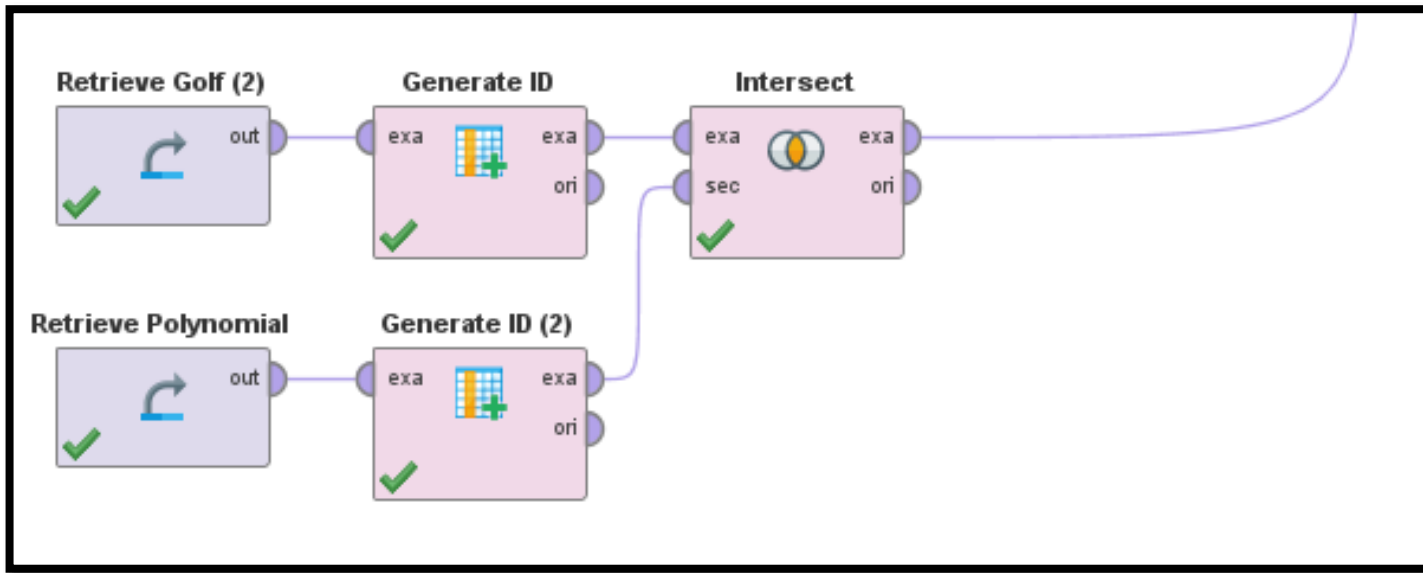
Hasil

Row No.	id	Play	Outlook	Temperature	Humidity	Wind	a1
1	?	no	sunny	85	85	false	?
2	?	no	sunny	80	90	true	?
3	?	yes	overcast	83	78	false	?
4	?	yes	rain	70	96	false	?
5	?	yes	rain	68	80	false	?
6	?	no	rain	65	70	true	?
7	?	yes	overcast	64	65	true	?
8	?	no	sunny	72	95	false	?
9	?	yes	sunny	69	70	false	?
10	?	yes	rain	75	80	false	?
11	?	yes	sunny	75	70	true	?
12	?	yes	overcast	72	90	true	?
13	?	yes	overcast	81	75	false	?
14	?	no	rain	71	80	true	?

ExampleSet (164 examples, 2 special attributes, 8 regular attributes)

Mengkombinasikan Data Set Dengan Intersect (Irisan)

- **Operator Intersect** digunakan untuk melihat irisan atau data yang sama dari 2 data set.
- ❑ Retrieve data “Golf” dan “Polynomial” ke panel process
- ❑ Pilih operator generate ID – Hubungkan dengan data Golf – pada parameter isikan offset = 0 (artinya urutan dari data golf adalah 1 – 14)
- ❑ Pilih operator generate ID – Hubungkan dengan data Polynomial – pada parameter isikan offset = 10 (artinya urutan dari data polynomial dimulai dari 11 – 210)
- ❑ Pilih operator “Intersect” – Hubungkan dengan kedua Generate ID data tersebut
- ❑ Klik ▷



Hasil

Row No.	id	Play	Outlook	Temperature	Humidity	Wind
1	11	yes	sunny	75	70	true
2	12	yes	overcast	72	90	true
3	13	yes	overcast	81	75	false
4	14	no	rain	71	80	true

Mengkombinasikan Data Set Dengan Set Minus

➤ **Operator Set Minus** digunakan untuk melihat isi dari dataset pertama tetapi yang tidak termuat di dataset yang kedua.

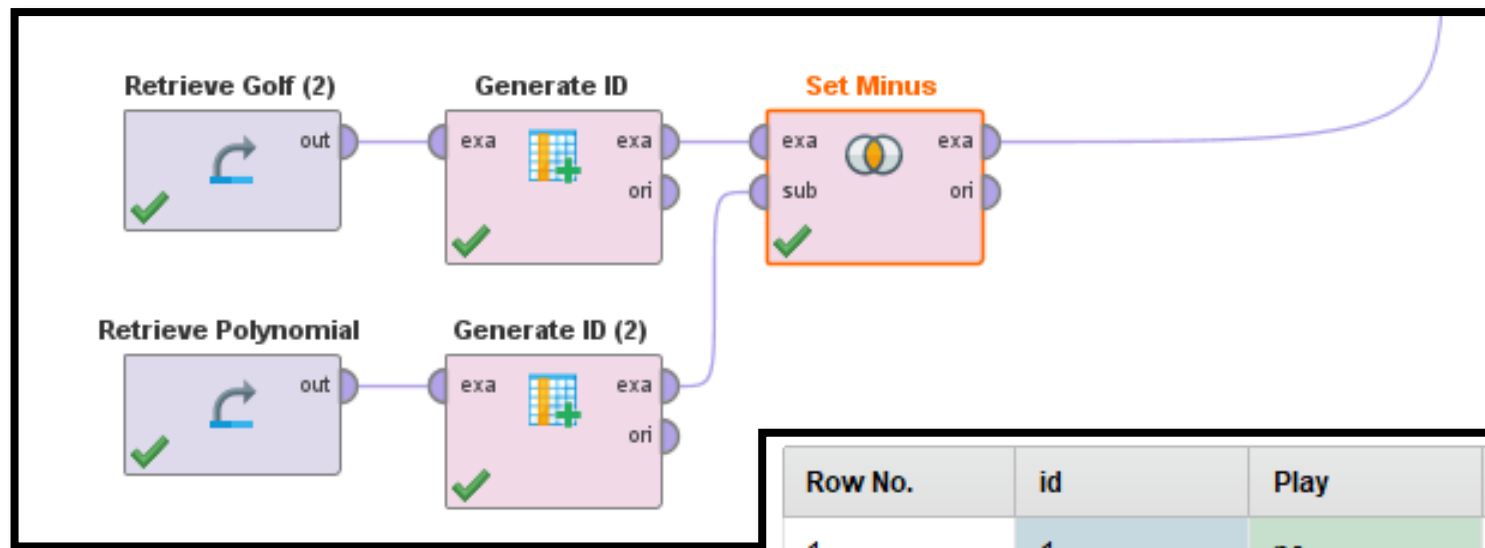
- ❑ Retrieve data “Golf” dan “Polynomial” ke panel process

- ❑ Pilih operator generate ID – Hubungkan dengan data Golf – pada parameter isikan offset = 0 (artinya urutan dari data golf adalah 1 – 14)

- ❑ Pilih operator generate ID – Hubungkan dengan data Polynomial – pada parameter isikan offset = 10 (artinya urutan dari data polynomial dimulai dari 11 – 210)

- ❑ Pilih operator “Set Minus” – Hubungkan dengan kedua Generate ID data tersebut

- ❑ Klik ▷



Hasil →

Row No.	id	Play	Outlook	Temperature	Humidity	Wind
1	1	no	sunny	85	85	false
2	2	no	sunny	80	90	true
3	3	yes	overcast	83	78	false
4	4	yes	rain	70	96	false
5	5	yes	rain	68	80	false
6	6	no	rain	65	70	true
7	7	yes	overcast	64	65	true
8	8	no	sunny	72	95	false
9	9	yes	sunny	69	70	false
10	10	yes	rain	75	80	false

Mengkombinasikan Data Set Dengan Cartesian Product

➤ **Operator Cartesian Product** digunakan untuk membuat kombinasi / pasangan berurutan dari seluruh data di dataset 1 dengan seluruh data di dataset 2.

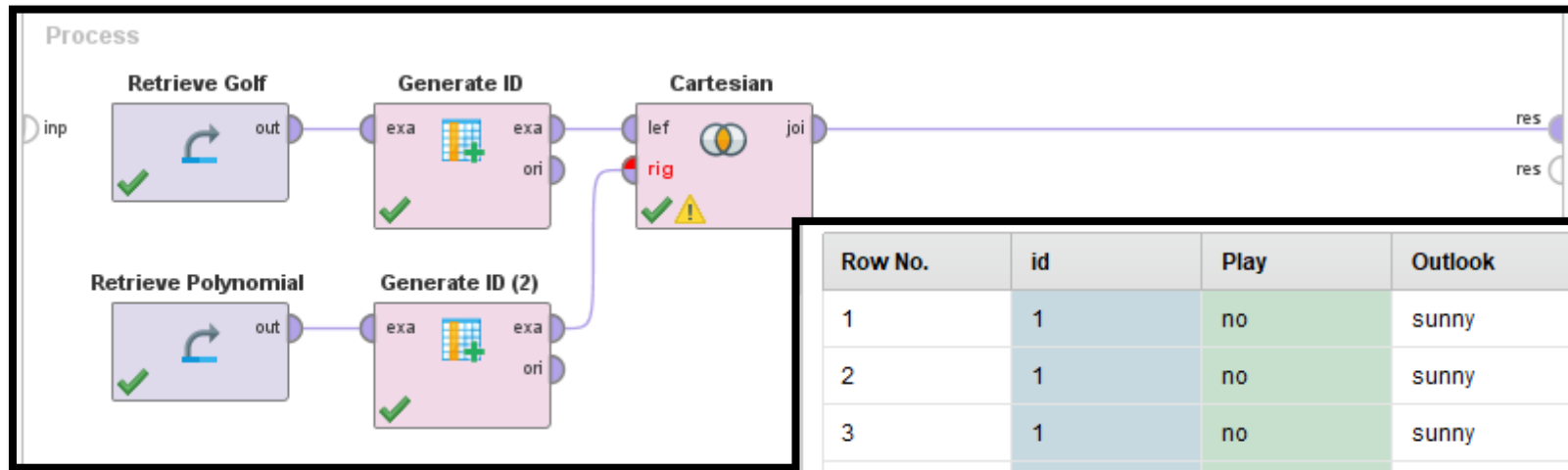
- ❑ Retrieve data “Golf” dan “Polynomial” ke panel process

- ❑ Pilih operator generate ID – Hubungkan dengan data Golf – pada parameter isikan offset = 0 (artinya urutan dari data golf adalah 1 – 14)

- ❑ Pilih operator generate ID – Hubungkan dengan data Polynomial – pada parameter isikan offset = 10 (artinya urutan dari data polynomial dimulai dari 11 – 210)

- ❑ Pilih operator “Cartesian Product” – Hubungkan dengan kedua Generate ID data tersebut

- ❑ Klik ▷



Hasil →

Row No.	id	Play	Outlook	Temperature	Humidity	Wind	a1
1	1	no	sunny	85	85	false	0.637
2	1	no	sunny	85	85	false	4.595
3	1	no	sunny	85	85	false	4.292
4	1	no	sunny	85	85	false	8.560
5	1	no	sunny	85	85	false	2.272
6	1	no	sunny	85	85	false	3.956
7	1	no	sunny	85	85	false	2.220
8	1	no	sunny	85	85	false	6.105
9	1	no	sunny	85	85	false	2.691
10	1	no	sunny	85	85	false	4.263
11	1	no	sunny	85	85	false	1.364
12	1	no	sunny	85	85	false	6.226
13	1	no	sunny	85	85	false	9.812
14	1	no	sunny	85	85	false	1.383

ExampleSet (2,800 examples, 2 special attributes, 9 regular attributes)

Big Data & Data Analytics

Prodi Psikologi Kelas III & IV

PERTEMUAN 5

Rekayasa Attribute

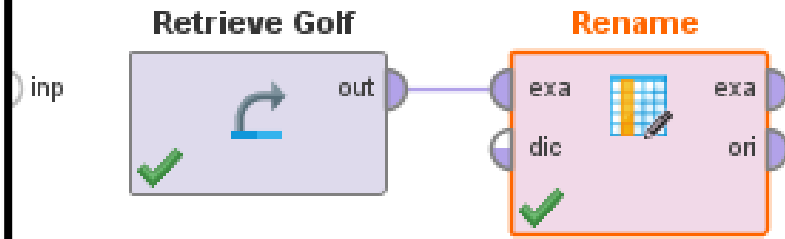
- Me-Rename attribute
- Filter Data
- Mengubah Tipe Data

Mengubah Nama Attribute

Akan dilakukan perubahan nama attribute pada Data “Golf” menggunakan bahasa Indonesia

- ☐ Retrieve Data “Golf” ke Panel Process
- ☐ Pilih Operator “Rename” – Hubungkan dengan data Golf
- ☐ Pada Parameter – Rename Attributes : Klik Edit List
 - a. Old Name : Pilih “Play” – New Name : Ketikkan “Bermain”
 - b. Old Name : Pilih “Humidity” – New Name : Ketikkan “Kelembaban”
 - c. Old Name : Pilih “ Outlook” – New Name : Ketikkan “Cuaca”
 - d. Old Name : Pilih “Temperature” – New Name : Ketikkan “Suhu”
 - e. Old Name : Pilih “Wind” – New Name : Ketikkan “Angin”
- ☐ Klik apply
- ☐ Klik ►

Process



Hasil

Row No.	Bermain	Cuaca	Suhu	Kelembaban	Angin
1	no	sunny	85	85	false
2	no	sunny	80	90	true
3	yes	overcast	83	78	false
4	yes	rain	70	96	false
5	yes	rain	68	80	false
6	no	rain	65	70	true
7	yes	overcast	64	65	true
8	no	sunny	72	95	false
9	yes	sunny	69	70	false
10	yes	rain	75	80	false
11	yes	sunny	75	70	true
12	yes	overcast	72	90	true
13	yes	overcast	81	75	false
14	no	rain	71	80	true

Akan dilakukan perubahan nama attribute pada Data “Sonar” dengan menghilangkan tanda “_”

- ❑ Retrieve Data “Sonar” ke Panel Process

- ❑ Pilih Operator “Rename by Replacing” – Hubungkan dengan data Sonar

- ❑ Pada Parameter :

- a. Attribute filter type : Pilih all (Karena akan diaplikasikan untuk semua atribut)

- b. Replace what : Ketikkan dengan tanda underscore (_) (karena yang akan kita hilangkan tanda ini)

- ❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training

Operators

rename

Attributes (5)

Names & Roles (5)

- Rename
- Rename by Replacing
- Rename by Example Va
- Rename by Generic Nar
- Rename by Construction

[Get more operators from the Marketplace](#)

Process

Process

inp

Retrieve Sonar

out

Rename by Replacing

exa

exa

ori

res

res

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Rename by Replacing

attribute filter type: all

☐ invert selection

☐ include special attributes

replace what: -

replace by:

Help

Rename by Replacing



RapidMiner Studio Core

Tags: [Names](#), [Regex](#), [Regular expressions](#), [Names & Roles](#)

Synopsis

This operator can be used to rename a set of attributes by replacing parts of the attribute names by a specified replacement.

Result History ExampleSet (Rename by Replacing) X

Open in  Turbo Prep  Auto Model Filter (208 / 208 examples): all

Row No.	class	attribute1	attribute2	attribute3	attribute4	attribute5	attribute6
1	Rock	0.020	0.037	0.043	0.021	0.095	0.099
2	Rock	0.045	0.052	0.084	0.069	0.118	0.258
3	Rock	0.026	0.058	0.110	0.108	0.097	0.228
4	Rock	0.010	0.017	0.062	0.021	0.021	0.037
5	Rock	0.076	0.067	0.048	0.039	0.059	0.065
6	Rock	0.029	0.045	0.028	0.017	0.038	0.099
7	Rock	0.032	0.096	0.132	0.141	0.167	0.171
8	Rock	0.052	0.055	0.084	0.032	0.116	0.092
9	Rock	0.022	0.037	0.048	0.048	0.065	0.059
10	Rock	0.016	0.017	0.035	0.007	0.019	0.067
11	Rock	0.004	0.006	0.015	0.034	0.031	0.028
12	Rock	0.012	0.031	0.017	0.031	0.036	0.010
13	Rock	0.008	0.009	0.005	0.025	0.034	0.055
14	Rock	0.009	0.006	0.025	0.049	0.120	0.159

ExampleSet (208 examples, 1 special attribute, 60 regular attributes)

Repository

- Dead Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting
- processes
- Templates
- Time Series
- Tutorials
- Local Repository (Local)
- DB (Legacy)

Hasil

Akan dilakukan perubahan nama attribute pada Data “Sonar” dengan nama inisial “att1, att2,dst”

- ❑ Retrieve Data “Sonar” ke Panel Process

- ❑ Pilih Operator “Rename by Replacing” – Hubungkan dengan data Sonar

- ❑ Pada Parameter :

- a. Attribute filter type : Pilih all (Karena akan diaplikasikan untuk semua atribut)

- b. Replace what : Ketikkan dengan (att)ribute_ (karena yang kita ambil hanya att)

- c. Replace by : \$1

- ❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training

Operators

rename

Attributes (5)

Names & Roles (5)

- Rename
- Rename by Replacing
- Rename by Example Va
- Rename by Generic Nar
- Rename by Construction

[Get more operators from the Marketplace](#)

Process

Process

inp

Retrieve Sonar

out

Rename by Replacing

exa

ori

res

res

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Rename by Replacing

attribute filter type: all

☐ invert selection

☐ include special attributes

replace what: (att)ribute_

replace by: \$1

Help

Rename by Replacing

RapidMiner Studio Core



Tags: Names, Regex, Regular expressions, Names & Roles

Synopsis

This operator can be used to rename a set of attributes by replacing parts of the attribute names by a specified replacement.

HASIL

Result History ExampleSet (Rename by Replacing) X


Open in  Turbo Prep  Auto Model

Filter (208 / 208 examples): all

Row No.	class	att1	att2	att3	att4	att5	att6
1	Rock	0.020	0.037	0.043	0.021	0.095	0.099
2	Rock	0.045	0.052	0.084	0.069	0.118	0.258
3	Rock	0.026	0.058	0.110	0.108	0.097	0.228
4	Rock	0.010	0.017	0.062	0.021	0.021	0.037
5	Rock	0.076	0.067	0.048	0.039	0.059	0.065
6	Rock	0.029	0.045	0.028	0.017	0.038	0.099
7	Rock	0.032	0.096	0.132	0.141	0.167	0.171
8	Rock	0.052	0.055	0.084	0.032	0.116	0.092
9	Rock	0.022	0.037	0.048	0.048	0.065	0.059
10	Rock	0.016	0.017	0.035	0.007	0.019	0.067
11	Rock	0.004	0.006	0.015	0.034	0.031	0.028
12	Rock	0.012	0.031	0.017	0.031	0.036	0.010
13	Rock	0.008	0.009	0.005	0.025	0.034	0.055
14	Rock	0.009	0.006	0.025	0.049	0.120	0.159

ExampleSet (208 examples, 1 special attribute, 60 regular attributes)

Repository

 Import Data

- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting
- processes
- Templates
- Time Series
- Tutorials
- Local Repository (Local)
- DB (Legacy)

Akan dilakukan perubahan nama attribute pada Data “Titanic” dengan Baris Tertentu Pada Data Tersebut

- ☐ Retrieve Data “Titanic” ke Panel Process
- ☐ Pilih Operator “Rename by Examples Value” – Hubungkan dengan data Titanic
- ☐ Pada Parameter : Row Number diisi dengan 3 (Misal dinamai dengan baris ketiga)
- ☐ Klik ▷

Repository

+ Import Data

- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting

processes

Operators

Rename

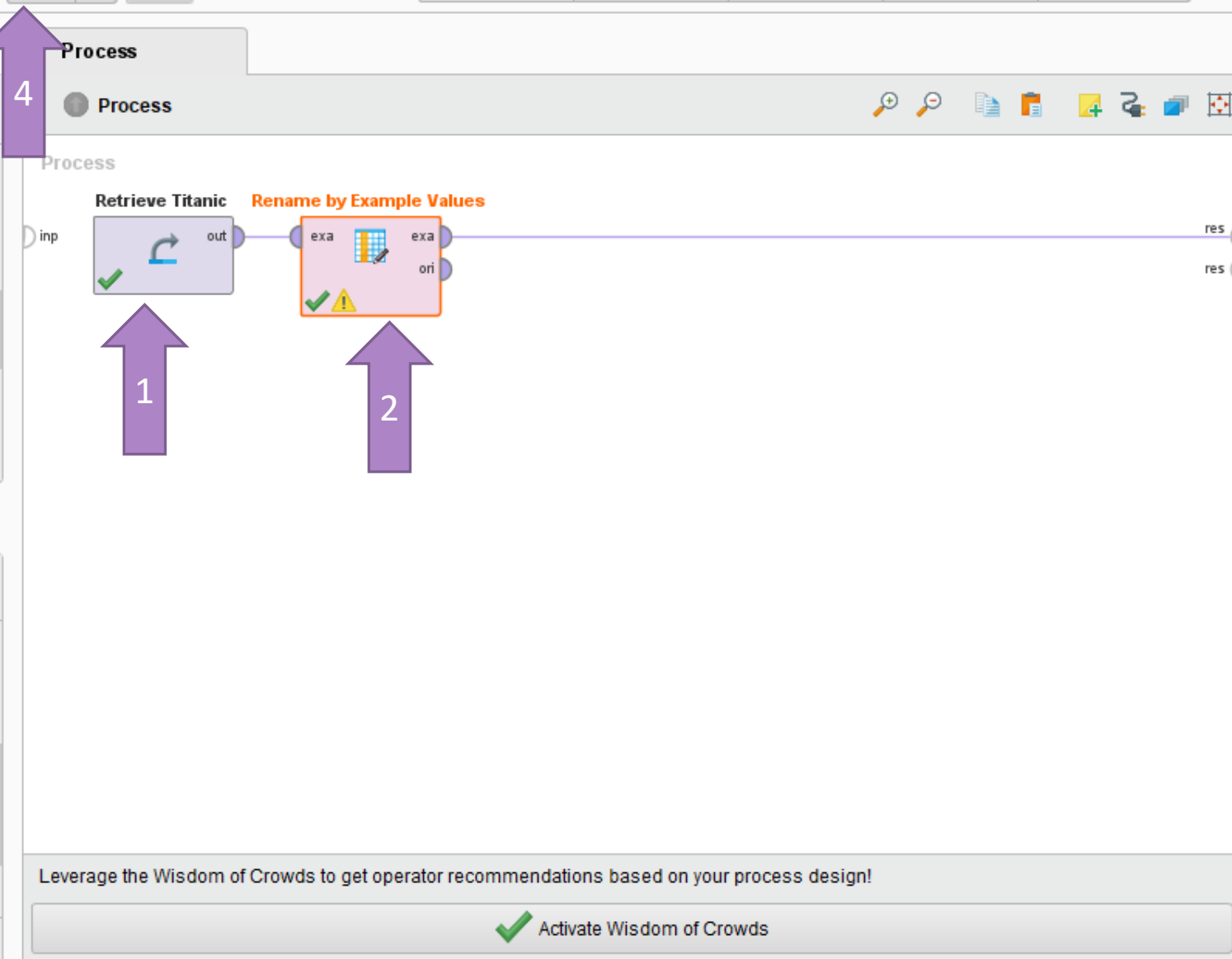
- Rename by Re
- Rename by Exa
- Rename by Ge
- Rename by Co

Utility (1)

Files (1)

- Rename File

No results were found.



Parameters

Rename by Example Values

row number 2

3

Help

Rename by Example Values

Blending

Tags: Names, Columns, Attributes, Features, Variables, Names & Roles

Synopsis

This operator renames the attributes of an ExampleSet by assigning the values of a specified example as attribute names and deleting that

HASIL

Result History ExampleSet (Rename by Example Values)

- Data
- Statistics
- Visualizations
- Annotations

Open in Turbo Prep Auto Model Filter (1,308 / 1,308 examples): all

Row No.	First	Allison, Mas...	Male	0.9167	1	2	113781
1	First	Allen, Miss. E...	Female	29	0	0	24160
2	First	Allison, Miss. ...	Female	2	1	2	113781
3	First	Allison, Mr. H...	Male	30	1	2	113781
4	First	Allison, Mrs. ...	Female	25	1	2	113781
5	First	Anderson, Mr....	Male	48	0	0	19952
6	First	Andrews, Mis...	Female	63	1	0	13502
7	First	Andrews, Mr. ...	Male	39	0	0	112050
8	First	Appleton, Mrs...	Female	53	2	0	11769
9	First	Artagaveytia, ...	Male	71	0	0	PC 17609
10	First	Astor, Col. Jo...	Male	47	1	0	PC 17757
11	First	Astor, Mrs. Jo...	Female	18	1	0	PC 17757
12	First	Aubart, Mme. ...	Female	24	0	0	PC 17477
13	First	Barber, Miss. ...	Female	26	0	0	19877
14	First	Barkworth, Mr...	Male	80	0	0	27042

ExampleSet (1,308 examples, 0 special attributes, 12 regular attributes)

- Repository
- + Import Data
- Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
 - processes
 - Templates
 - Time Series
 - Tutorials
 - Community Samples (connected)
 - Local Repository (Local)
 - DB (Legacy)

Akan dilakukan perubahan nama attribute pada Data “Titanic” dengan Generic Name

- ❑ Retrieve Data “Titanic” ke Panel Process
- ❑ Pilih Operator “Rename by Generic Names” – Hubungkan dengan data Titanic
- ❑ Pada Parameter : Generic Name Stem diisi dengan Titanic (Misal)
- ❑ Klik ▷

Repository

+ Import Data

- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting

processes

Operators

Rename

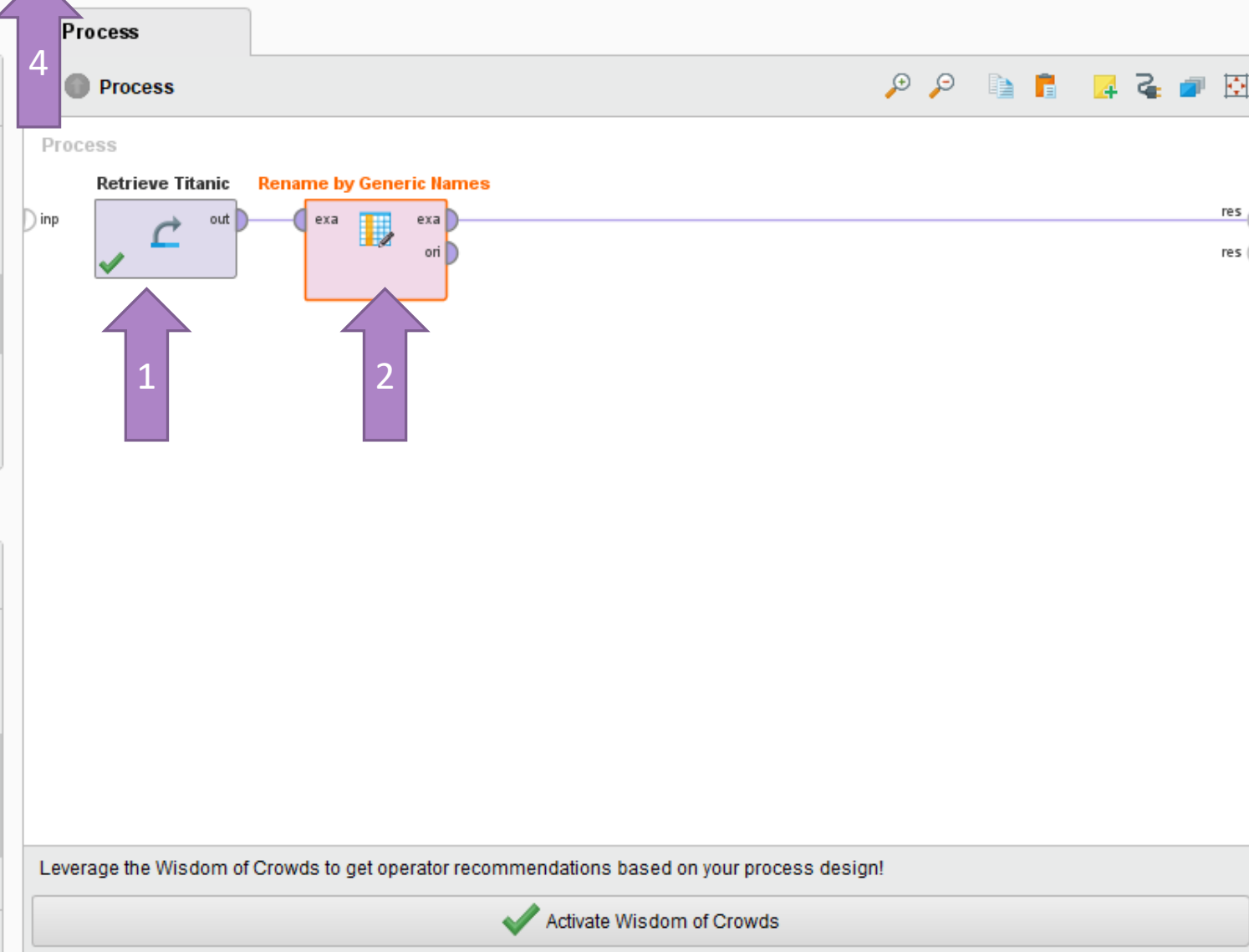
- Rename by Reg
- Rename by Exa
- Rename by Ge
- Rename by Co

Utility (1)

Files (1)

- Rename File

No results were found.



Parameters

Rename by Generic Names

attribute filter type all

invert selection

include special attributes

generic name stem Titanic

3

Help

Rename by Generic Names

RapidMiner Studio Core

Tags: Names & Roles

Synopsis

This operator renames the selected attributes of the given ExampleSet to a set of generic names like att1, att2, att3 etc.

[Jump to Tutorial Process](#)

HASIL

Result History ExampleSet (Rename by Generic Names) X

- Data
- Statistics
- Visualizations
- Annotations

Open in Turbo Prep Auto Model Filter (1,309 / 1,309 examples): all

Row No.	Titanic1	Titanic2	Titanic3	Titanic4	Titanic5	Titanic6	Titanic7
1	First	Allen, Miss. E...	Female	29	0	0	24160
2	First	Allison, Mast...	Male	0.917	1	2	113781
3	First	Allison, Miss. ...	Female	2	1	2	113781
4	First	Allison, Mr. H...	Male	30	1	2	113781
5	First	Allison, Mrs. ...	Female	25	1	2	113781
6	First	Anderson, Mr....	Male	48	0	0	19952
7	First	Andrews, Mis...	Female	63	1	0	13502
8	First	Andrews, Mr. ...	Male	39	0	0	112050
9	First	Appleton, Mrs...	Female	53	2	0	11769
10	First	Artagaveytia, ...	Male	71	0	0	PC 17609
11	First	Astor, Col. Jo...	Male	47	1	0	PC 17757
12	First	Astor, Mrs. Jo...	Female	18	1	0	PC 17757
13	First	Aubart, Mme. ...	Female	24	0	0	PC 17477
14	First	Barber, Miss. ...	Female	26	0	0	19877

ExampleSet (1,309 examples, 0 special attributes, 12 regular attributes)

- Repository
- + Import Data
- Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
 - processes
 - Templates
 - Time Series
 - Tutorials
 - Community Samples (connected)
 - Local Repository (Local)
 - DB (Legacy)

Melakukan Filter By Ranges

Untuk menampilkan data set sesuai dengan Range/Rentang yang diinginkan

❑ Retrieve Data “Titanic” Pada Panel Proses

❑ Pilih Operator “Filter by Ranges” – Klik

❑ Pada Parameter :

- First Example : 3
- Last Example : 10

❑ Klik ►

Menghilangkan Simbol Tidak Berguna Pada Attribute

- ❑ Menggunakan dataset yang sudah di download, akan dibaca dataset kz_users.csv

menggunakan operator Read CSV



- ❑ Double klik pada operator Read CSV tersebut
- ❑ Pilih folder dimana dataset bertipe CSV yang akan dibaca diletakkan
- ❑ Pilih file kz_users.csv – Klik Next

Specify your data format

<input checked="" type="checkbox"/> Header Row	<input type="text" value="1"/>	File Encoding	<input type="text" value="windows-12..."/>	<input checked="" type="checkbox"/> Use Quotes	<input type="text" value=""/>
Start Row	<input type="text" value="1"/>	Escape Character	<input type="text" value="\"/>	<input type="checkbox"/> Trim Lines	
Column Separator	<input type="text" value="Semicolon ';'"/>	Decimal Character	<input type="text" value="."/>	<input checked="" type="checkbox"/> Skip Comments	<input type="text" value="#"/>

1	user_id	first_name	last_name	city
2	1515915625441993984	Marlon	Callum	Pueblo
3	1515915625441993984	Johnathan	Gordon	Oceanside
4	1515915625441993984	Markus		Lubbock
5	1515915625441993984	Brenna		Westminster
6	1515915625448766480	Braden	Armani	Louisville
7	1515915625450561165	Simeon		Quincy
8	1515915625446798439	Maxwell	Zackery	McKinney
9	1515915625450899340	Andy	Levi	Sunnyvale
10	1515915625451131565	Franklin	Eduardo	Wilmington
11	1515915625451212869	Izayah	Arjun	Pomona

Tanda yang akan
dihapus

✓ no problems.

← Previous → Next ✕ Cancel

Klik

Ubah Tipe Data
Menjadi Polynominal

Import Data - Format your columns. ✕

Format your columns.

Date format ▾ ☐ Replace errors with missing values ⓘ

	user_id <i>real</i>	first_name	last_name <i>polynominal</i>	city <i>polynominal</i>
1	1515915625441993980		Callum	Pueblo
2	1515915625447879420		Gordon	Oceanside
3	1515915625443148030		?	Lubbock
4	1515915625450382850.000	Brenna	?	Westminster
5	1515915625448766460.000	Braden	Armani	Louisville
6	1515915625450561280.000	Simeon	?	Quincy
7	1515915625446798340.000	Maxwell	Zackery	McKinney
8	1515915625450899460.000	Andy	Levi	Sunnyvale
9	1515915625451131650.000	Franklin	Eduardo	Wilmington
10	1515915625451212800.000	Izayah	Arjun	Pompano Beach
11	1515915625443158780.000	Peyton	Jamie	Long Beach
12	1515915625447780100.000	Frankie	Rohan	Memphis

no problems.

← Previous Finish Cancel

Klik

❑ Pilih operator “Rename by Replacing” – Hubungkan dengan Read CSV

❑ Double klik pada operator Rename by Replacing tersebut

❑ Pada Parameter :

- Replace What : diisi dengan [\W _] (untuk menghilangkan symbol yang bukan alphanumeric termasuk symbol _)

❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Purchases
- Ripley-Set
- Sonar
- Titanic
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting

processes

Operators

Rename

- Blending (5)
 - Attributes (5)
 - Names & Roles (5)
 - Rename
 - Rename by Repl
 - Rename by Exa
 - Rename by Ge

No results were found.

Process

Process

Read CSV

Rename by Replacing

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Rename by Replacing

attribute filter type: all

☐ invert selection

☐ include special attributes

replace what: [W _]

replace by:

Help

Rename by Replacing

RapidMiner Studio Core



Tags: Names, Regex, Regular expressions, Names & Roles

Synopsis

This operator can be used to rename a set of attributes by replacing parts of the attribute names by a specified replacement.

HASIL

Result History ExampleSet (Rename by Replacing) X


Open in  Turbo Prep  Auto Model

Filter (24,688 / 24,688 examples): all

Row No.	userid	firstname	lastname	city
1	1515915625...	Marlon	Callum	Pueblo
2	1515915625...	Johnathan	Gordon	Oceanside
3	1515915625...	Markus	?	Lubbock
4	1515915625...	Brenna	?	Westminster
5	1515915625...	Braden	Armani	Louisville
6	1515915625...	Simeon	?	Quincy
7	1515915625...	Maxwell	Zackery	McKinney
8	1515915625...	Andy	Levi	Sunnyvale
9	1515915625...	Franklin	Eduardo	Wilmington
10	1515915625...	Izayah	Arjun	Pompano Be...
11	1515915625...	Peyton	Jamie	Long Beach
12	1515915625...	Frankie	Rohan	Memphis
13	1515915625...	Isaiah	Dexter	Westminster
14	1515915625...	Adrien	Ali	Jackson
15	1515915625...	Jayce	Brycen	Greensboro

ExampleSet (24,688 examples, 0 special attributes, 4 regular attributes)

Repository

 Import Data

- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products
- Purchases
- Ripley-Set
- Sonar
- Titanic**
- Titanic Training
- Titanic Unlabeled
- Transactions
- Weighting
- processes
- Templates
- Time Series
- Tutorials
- Community Samples (connected)
- Local Repository (Local)
- DB (Legacy)

Mengubah Tipe Data Real (Pecahan) Menjadi Integer (Bulat)

Menggunakan data “Polynomial”, attribute 1 dan 2 akan diubah menjadi bilangan bulat.

- ❑ Retrieve Data “Polynomial”

- ❑ Pilih Operator “Real to Integer” – klik

- ❑ Pada Parameter :

- Attribute filter type : pilih subset

- Attribute : klik Select Attributes – pindahkan attributes “a1” dan “a2” ke sebelah kanan – klik apply

- ❑ Klik ➤

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products

Operators

Type

Types (18)

- Set Positive Value
- Numerical to Binomina
- Numerical to Polynomi
- Numerical to Real
- Numerical to Date
- Real to Integer

No results were found.

Process

Process

Process

Retrieve Polynomial

Real to Integer

inp

out

exa

exa

ori

res

res

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Real to Integer

attribute filter type: subset

attributes: Select Attributes...

☐ invert selection

☐ include special attributes

☐ round values

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Real to Integer

RapidMiner Studio Core

Tags: Types

Synopsis

This operator changes the type of the selected real attributes to integer type. It also maps all values of these attributes to integer values.



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products

Operators

Type

Types (18)

- Set Positive Value
- Numerical to Binomina
- Numerical to Polynomi
- Numerical to Real
- Numerical to Date
- Real to Integer

No results were found.

Process

Process

Retrieve Polyno

inp

Select Attributes: attributes

Select Attributes: attributes
The attribute which should be chosen.

Attributes

Search

- # a3
- # a4
- # a5
- # label

Selected Attributes

Search

- # a1
- # a2

Klik

Apply Cancel

Parameters

Real to Integer

Attribute filter type subset

Select Attributes...

Invert selection

Include special attributes

Round values

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Real to Integer

RapidMiner Studio Core

Types

Overview

This operator changes the type of the selected real attributes to integer type. It also maps all values of these attributes to integer values.

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!



Activate Wisdom of Crowds

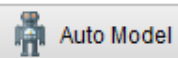
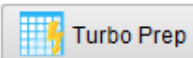


Views:

Hasilnya : data pada attributes “a1” dan “a2” berubah menjadi nilai pembulatan

Result History ExampleSet (Real to Integer) x

Open in



Filter (200 / 200 examples): all

Row No.	label	a1	a2	a3	a4	a5
1	88.016	0	8	2.116	2.772	8.448
2	138.755	4	4	4.926	2.682	8.618
3	38.284	4	0	9.158	6.607	4.085
4	82.148	8	3	1.038	1.688	2.893
5	33.790	2	1	6.301	1.938	2.070
6	7.389	3	0	4.337	7.563	8.591
7	191.453	2	7	6.463	8.732	9.282
8	15.349	6	1	1.163	9.241	7.239
9	97.405	2	3	7.806	6.843	3.054
10	111.497	4	7	0.821	6.688	3.976
11	203.574	1	9	8.860	2.573	1.538
12	140.767	6	8	0.351	3.639	6.234
13	259.381	9	3	5.305	6.661	7.118
14	20.330	1	1	7.947	2.043	0.030
15	105.813	0	9	0.001	4.557	3.258

ExampleSet (200 examples, 1 special attribute, 5 regular attributes)

- data
 - Deals
 - Deals-Testset
 - Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
- processes
- Templates
- Time Series
- Tutorials

Mengubah Tipe Data

Menggunakan data “electronics-b-100k.csv” yang sudah didownload.

- ❑ Pilih Operator “Read CSV” – pilih file “electronics-b-100k.csv” – klik next – klik next
- ❑ Ubah Tipe Data :
 1. Order_id, product_id, user_id diubah dari Real menjadi Polynominal
 2. Klik Finish
- ❑ Tambahkan Operator “Date to Nominal” untuk mengubah tipe data Tanggal (Date) menjadi Nominal – klik

❑ Pada parameter :

- a. Attribute name : pilih attribute “Event Time”
- b. Isikan date format : pilih dd/MM/YYYY (misal dalam format Indonesia)
- c. Isikan locale : pilih Indonesian (Indonesia)

❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

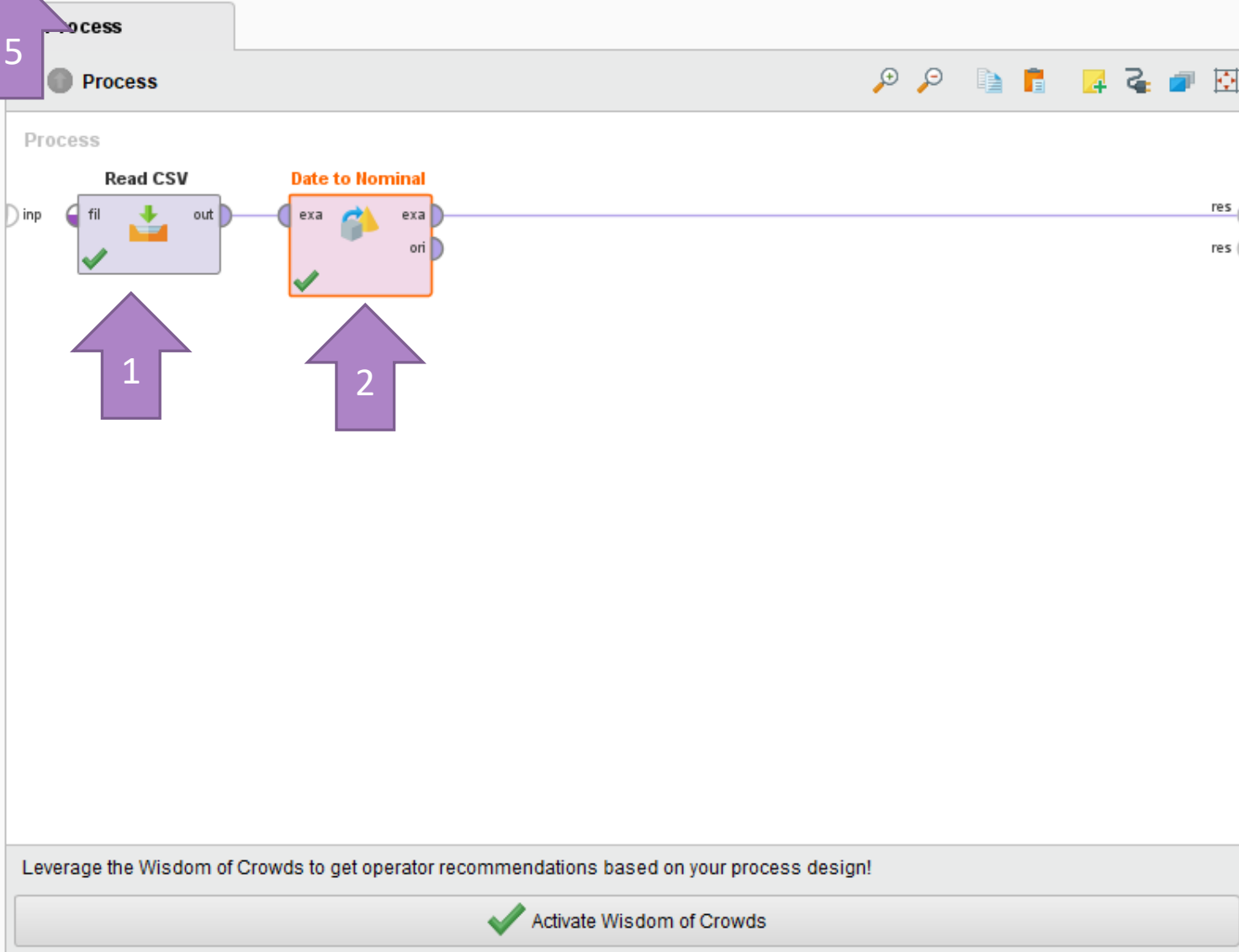
- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products

Operators

type

- Date to Numerical
- Date to Nominal
- Parse Numbers
- Format Numbers
- Guess Types
- One-Hot Encoding
- Target Encoding

No results were found.



Parameters

Date to Nominal

attribute name: event_time

date format: dd/MM/yyyy

time zone: SM

locale: Indonesian (Indonesia)

☐ keep old attribute

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Date to Nominal

RapidMiner Studio Core

Tags: [Types](#)

Synopsis

This operator parses the date values of the specified date attribute with respect to the given date format string and transforms the values into nominal values.

Result History ExampleSet (Date to Nominal)

	Name	Type	Missing	Filter (8 / 8 attributes): Search for Attributes	
	product_id	Polynomial	0	Least 2388434452476051570...	Most 15159662 [...]
	category_id	Polynomial	10437	Least 2374498914001945685...	Most 22681054 [...]
	category_code	Polynomial	23103	Least sport.ski (1)	Most electron [...] e (
	brand	Polynomial	5718	Least xigmatek (1)	Most samsung (1524
	price	Real	10437	Min 0	Max 8101.830
	user_id	Polynomial	14468	Least 1515915625514859200...	Most 15159156 [...]
	event_time	Nominal	0	Least 25/10/2020 (1211)	Most 22/10/2020 (10

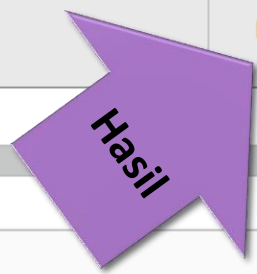
Showing attributes 1 - 8

Examples: 100,000 Special Attributes: 0 Regular Attributes: 8

Repository

+ Import Data

- data
 - Deals
 - Deals-Testset
 - Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
- processes
- Templates
- Time Series
- Tutorials



Karena untuk Event Time apabila bertipe Nominal akan berpengaruh pada hasil Analisa data, maka Event Time harus diubah bertipe Date.

- ❑ Tambahkan operator “Nominal to Date”

- ❑ Pada parameter :

- a. Attribute name : pilih attribute “Event Time”

- b. Isikan date format : pilih dd/MM/YYYY (misal dalam format Indonesia)

- c. Isikan locale : pilih Indonesian (Indonesia)

- ❑ Klik ►



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiations
- Market-Data
- Polynomial
- Products

Operators

type

- Nominal to Date
- Real to Integer
- Nominal to Binomial
- Nominal to Text
- Nominal to Numerical
- Nominal to Date
- Text to Nominal
- Date to Numerical

No results were found.

Process

Process

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

✓ Activate Wisdom of Crowds

Parameters

Nominal to Date

attribute name: event_time

date type: date

date format: dd/MM/yyyy

time zone: SYSTEM

locale: Indonesian (Indonesia)

☐ keep old attribute

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Nominal to Date

RapidMiner Studio Core

Tags: Dates, Times, Datetimes, Formats, Nominal, Categorical, Polynominal, Weeks, Days, Months, Years, Quarters, Hours, Minutes, Seconds, Types

Synopsis

This operator converts the selected nominal attribute into the selected date time type. The nominal values

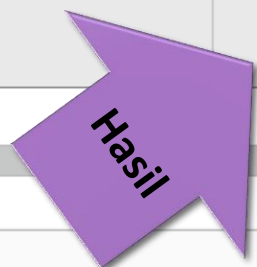
Result History ExampleSet (Nominal to Date)

	Name	Type	Missing	Filter (8 / 8 attributes): Search for Attributes	
▼	product_id	Polynomial	0	Least 2388434452476051570...	Most 15159662 [...]
▼	category_id	Polynomial	10437	Least 2374498914001945685...	Most 22681054 [...]
▼	category_code	Polynomial	23103	Least sport.ski (1)	Most electron [...] e (
▼	brand	Polynomial	5718	Least xigmatek (1)	Most samsung (1524
▼	price	Real	10437	Min 0	Max 8101.830
▼	user_id	Polynomial	14468	Least 1515915625514859200...	Most 15159156 [...]
▼	event_time	Date	0	Earliest date Sep 30, 2020	Latest date Oct 25, 2020

Showing attributes 1 - 8 Examples: 100,000 Special Attributes: 0 Regular Attributes: 8

Repository

- + Import Data
- data
 - Deals
 - Deals-Testset
 - Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
 - processes
 - Templates
 - Time Series
 - Tutorials



Pada attribute Price akan diubah menjadi bilangan bulat (Integer) dan dimunculkan mata uang Rupiah.

- ❑ Tambahkan operator “Real to Integer” – klik

- ❑ Pada parameter :

- a. Attribute filter type : pilih “single”

- b. Attribute : pilih “Price”

❑ Tambahkan operator “Format Number” – klik

❑ Pada parameter :

a. Attribute filter type : pilih “single”

c. Format type : pilih currency

b. Attribute : pilih “Price”

d. Locale : pilih Indonesian

(Indonesia)

❑ Klik ▷



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiation
- Market-Data
- Polynomial

Operators

Real t

- Blending (4)
 - Attributes (4)
 - Types (4)
 - Numerical to Real
 - Real to Integer
 - Parse Numbers
 - Guess Types

No results were found.

Process

Process

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

✓ Activate Wisdom of Crowds

Parameters

Real to Integer

attribute filter type: single

attribute: price

☐ invert selection

☐ include special attributes

☐ round values

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Real to Integer

RapidMiner Studio Core

Tags: [Types](#)

Synopsis

This operator changes the type of the selected real attributes to integer type. It also maps all values of these attributes to integer values.

[Jump to Tutorial Process](#)



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiation
- Market-Data
- Polynomial

Operators

format n

- Nominal to Dat
- Parse Number
- Format Numbe

Cleansing (1)

Normalization (1)

We found "Information Extraction", "Information Selection" and 4 more results in the Marketplace. [Show me!](#)

Process

Process

inp

res

res

Read CSV

Real to Integer

Format Numbers

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

✓ Activate Wisdom of Crowds

Parameters

Format Numbers

attribute filter type: single

attribute: price

☐ invert selection

☐ include special attributes

format type: currency

locale: Indonesian (Indonesia)

[Hide advanced parameters](#)

[Change compatibility \(9.10.000\)](#)

Help

Format Numbers

RapidMiner Studio Core



Tags: [Types](#)

Synopsis

This operator reformats the selected numerical attributes according to the specified format and changes the attributes to nominal.

[Jump to Tutorial Process](#)

Result History ExampleSet (Format N)

Open in  Turbo Prep  Model

Filter (100,000 / 100,000 examples): all

Row No.	price	time	order_id	product_id	category_id	category_co...	brand
1	Rp115,00	Oct 25, 2020 ...	2388440981...	1515966223...	2268105409...	computers.p...	samsung
2	Rp648,00	Oct 25, 2020 ...	2388440981...	2388434452...	2268105407...	computers.n...	lenovo
3	Rp97,00	Oct 25, 2020 ...	2388440981...	2273948227...	2268105392...	appliances.e...	samsung
4	Rp37,00	Oct 25, 2020 ...	2388440981...	2388434452...	2268105442...	furniture.kitch...	tw
5	Rp37,00	Oct 25, 2020 ...	2388440981...	1515966223...	2268105442...	furniture.kitch...	?
6	Rp31,00	Oct 25, 2020 ...	2388440981...	2273948186...	2268105406...	?	svc
7	?	Oct 25, 2020 ...	2388440981...	1515966223...	?	17.36	15159156
8	Rp289,00	Oct 25, 2020 ...	2388440981...	1515966223...	2268105389...	appliances.ki...	samsung
9	Rp358,00	Oct 25, 2020 ...	2388440981...	1515966223...	2374498914...	electronics.vi...	lg
10	?	Oct 25, 2020 ...	2388440981...	1515966223...	?	8.33	15159156
11	Rp32,00	Oct 25, 2020 ...	2388440981...	2273948247...	2268105423...	?	deluxe
12	?	Oct 25, 2020 ...	2388440981...	1515966223...	?	9.26	15159156
13	Rp43,00	Oct 25, 2020 ...	2388440981...	1515966223...	2268105430...	electronics.a...	hyperx
14	Rp543,00	Oct 25, 2020 ...	2388440981...	1515966223...	2268105389...	appliances.ki...	bosch

ExampleSet (100,000 examples, 0 special attributes, 8 regular attributes)

Repository

+ Import Data

- data
 - Deals
 - Deals-Testset
 - Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
- processes
- Templates
- Time Series
- Tutorials

Susunan attribute pada hasil running tersebut masih belum rapi sehingga perlu dirapikan susunannya.

- ❑ Pilih Operator “Reorder attributes” – klik

- ❑ Pada parameter , Isikan :

- Attribute ordering : klik define order
- Tentukan urutan attribute yang dikehendaki
- Klik apply

- ❑ Klik ▷

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiation
- Market-Data
- Polynomial

Operators

reorder

- Blending (1)
 - Attributes (1)
 - Reorder Attributes

No results were found.

Process

Process

inp

Read CSV

Real to Integer

Format Numbers

Reorder Attributes

res

res

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Parameters

Reorder Attributes

sort mode: user specified

attribute ordering: Define order...

☐ use regular expressions

handle unmatched: append

sort direction: ascending

[Hide advanced parameters](#)

Help

Reorder Attributes

RapidMiner Studio Core

Tags: Sort, Arrange, Ascending, Descending, Attributes

Synopsis

This operator allows to reorder regular Attributes of an ExampleSet. Reordering can be done alphabetically, by user specification (including Regular Expressions) or with a reference ExampleSet.



Views:

Design

Results

Turbo Prep

Auto Model

Deployments

Find data, operators...etc



All Studio

Repository

+ Import Data

- Deals
- Deals-Testset
- Golf
- Golf-Testset
- Iris
- Labor-Negotiation
- Market-Data
- Polynomial

Operators

reorder

- Blending (1)
 - Attributes (1)
 - Reorder Attributes

No results were found.

Process

Process

inp

Read CSV

fil

Select Ordering Rules: attribute ordering

Select Ordering Rules: attribute ordering

Rules to order attributes.

Attributes

Search

Attribute Ordering

- user_id
- order_id
- category_id
- category_code
- product_id
- brand
- price
- event_time

Order Cancel

Parameters

Reorder Attributes

mode user specified

Attribute ordering Define order...

use regular expressions

Handle unmatched append

direction ascending

[Hide advanced parameters](#)

Reorder Attributes

RapidMiner Studio Core

Sort, Arrange, Ascending, Descending, Attributes

Synopsis

This operator allows to reorder regular Attributes of an ExampleSet. Reordering can be done alphabetically, by user specification (including Regular Expressions) or with a reference ExampleSet.

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!



Activate Wisdom of Crowds

Klik

Result History

ExampleSet (Reorder Attributes)



Data



Statistics



Visualizations



Annotations

Open in



Turbo Prep



Auto Model

Filter (100,000 / 100,000 examples):

all

Row No.	user_id	order_id	category_id	category_co...	product_id	brand	price
1	1515915625...	2388440981...	2268105409...	computers.p...	1515966223...	samsung	Rp115,00
2	1515915625...	2388440981...	2268105407...	computers.n...	2388434452...	lenovo	Rp648,00
3	1515915625...	2388440981...	2268105392...	appliances.e...	2273948227...	samsung	Rp97,00
4	1515915625...	2388440981...	2268105442...	furniture.kitch...	2388434452...	tv	Rp37,00
5	1515915625...	2388440981...	2268105442...	furniture.kitch...	1515966223...	?	Rp37,00
6	1515915625...	2388440981...	2268105406...	?	2273948186...	svc	Rp31,00
7	?	2388440981...	?	17.36	1515966223...	1515915625...	?
8	1515915625...	2388440981...	2268105389...	appliances.ki...	1515966223...	samsung	Rp289,00
9	1515915625...	2388440981...	2374498914...	electronics.vi...	1515966223...	lg	Rp358,00
10	?	2388440981...	?	8.33	1515966223...	1515915625...	?
11	1515915625...	2388440981...	2268105423...	?	2273948247...	deluxe	Rp32,00
12	?	2388440981...	?	9.26	1515966223...	1515915625...	?
13	1515915625...	2388440981...	2268105430...	electronics.a...	1515966223...	hyperx	Rp43,00
14	1515915625...	2388440981...	2268105389...	appliances.ki...	1515966223...	bosch	Rp543,00

ExampleSet (100,000 examples, 0 special attributes, 8 regular attributes)

Repository

+ Import Data

- data
 - Deals
 - Deals-Testset
 - Golf
 - Golf-Testset
 - Iris
 - Labor-Negotiations
 - Market-Data
 - Polynomial
 - Products
 - Purchases
 - Ripley-Set
 - Sonar
 - Titanic
 - Titanic Training
 - Titanic Unlabeled
 - Transactions
 - Weighting
- processes
- Templates
- Time Series
- Tutorials