Εξόρυξη δεδομένων με χρήση τεχνικών μηχανικής μάθησης

Τμ. Μηχανικών Πληροφορικής και Υπολογιστών

Παραδείγματα Προετοιμασίας Δεδομένων

Περιγραφή

- Γίνεται αναφορά σε σημαντικές έννοιες εξόρυξης δεδομένων (data mining) που αφορούν την **προετοιμασία δεδομένων (Data Preparation)**. Η διεκπεραίωση των θεμάτων γίνεται κυρίως με χρήση παραδειγμάτων.
- Επιπλέον, γίνεται αναφορά στο εργαλείο Rapid Miner

Καθαρισμός δεδομένων (data scrubbing)

Σε ένα dataset συχνά απαιτείται να διαχειριστούμε κάποιες μη αναμενόμενες τιμές/καταστάσεις («ανωμαλίες») στα δεδομένα και να τις διορθώσουμε σύμφωνα με τις ανάγκες μας.

Ο όρος στην Αγγλική είναι καθαρισμός δεδομένων “(process of) data scrubbing”.

παράδειγμα

Row ...	Gender	Race	Birth_Year	Marital...	Years_on_Internet	Hours_Per_Day	Preferred_Browser	Preferred_Search_Engine
1	M	White	1972	M	8	1	Firefox	Google
2	M	Hispanic	1981	S	14	2	Chrome	Google
3	F	African American	1977	S	6	2	Firefox	Yahoo
4	F	White	1961	D	8	6	Firefox	Google
5	M	White	1954	M	2	3	Internet Explorer	Bing
6	M	African American	1982	D	15	4	Internet Explorer	Google
7	M	African American	1981	D	11	2	Firefox	Google
8	M	White	1977	S	3	3	Internet Explorer	Yahoo
9	F	African American	1969	M	6	2	Firefox	Google
10	M	White	1987	S	12	1	Safari	Yahoo
11	F	Hispanic	1959	D	12	5	Chrome	Google

Row No.	Preferred_E...	Read_News	Online_Sho...	Online_Gam...	Facebook	Twitter	Other_Socia...
1	Yahoo	Y	N	N	Y	N	?
2	Hotmail	Y	N	N	Y	N	?
3	Yahoo	Y	Y	?	Y	N	?
4	Hotmail	N	Y	N	N	Y	?
5	Hotmail	Y	Y	N	Y	N	?
6	Yahoo	Y	N	Y	N	N	?
7	Yahoo	?	Y	?	Y	Y	LinkedIn
8	Yahoo	Y	?	?	Y	99	LinkedIn
9	Gmail	N	Y	N	N	N	?
10	Yahoo	Y	?	Y	Y	N	MySpace
11	Gmail	Y	N	N	Y	N	Google+

Περιπτώσεις που απαιτούν καθαρισμό ή γενικότερα διαχείριση δεδομένων (data scrubbing):

(1) Missing Data (Ελλιπή δεδομένα),

(2) Reducing Data (Observations) (μείωση των δεδομένων (παρατηρήσεων)),

(3) Inconsistent Data (ασυνεπή δεδομένα),

(4) Reducing Attributes (μείωση χαρακτηριστικών)

Περιπτώσεις που απαιτούν καθαρισμό ή γενικότερα διαχείριση δεδομένων (data scrubbing):

Για τις ελλιπείς τιμές ενός χαρακτηριστικού μπορούμε να χρησιμοποιήσουμε μέτρα θέσης (mean, median, mode). Για τη διαχείριση ακραίων τιμών (outliers) μπορούμε να χρησιμοποιήσουμε μέτρα θέσης και μέτρα διακύμανσης (range, standard deviation)

Mean (μέση τιμή), median (διάμεσος), mode (κορυφή ή επικρατούσα τιμή), range (εύρος)

1) Έστω σύνολο τιμών χαρακτηριστικού:

13, 13, 13, 13, 14, 14, 16, 18, 21

Mean=(13 + 13 + 13 + 13 + 14 + 14 + 16 + 18 + 21) ÷ 9 = 15

13, 13, 13, 13, **14**, 14, 16, 18, 21

Median=14, Mode (δηλαδή, πιο συχνή τιμή)=13, Range=21–13 = 8.

2) Έστω σύνολο τιμών χαρακτηριστικού:

1, 2, 4, 7

mean=(1 + 2 + 4 + 7) ÷ 4 = 14 ÷ 4 = 3.5, median=(2 + 4) ÷ 2 = 6 ÷ 2 = 3

mode: δεν υπάρχει. Range=7-1=6.

standard deviation (τυπική απόκλιση)

Standard Deviation	$s_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$

Outliers (τιμές εκτός ορίων)

Ένας εμπειρικός κανόνας:

Αν η τιμή ενός χαρακτηριστικού δεν ανήκει στο διάστημα

[mean - 2*Standard Deviation, mean + 2*Standard Deviation] **ΤΌΤΕ**

ίσως πρέπει να εξαιρείται.

Ενδεικτική Βιβλιογραφία

Διασκευή παραδείγματος κεφαλαίου 3 του βιβλίου

M. North, Data Mining for the Masses, 2012,

ISBN: 978-0615684376

CRISP-DM, the CRoss-Industry Standard Process for Data Mining.

CRISP-DM Step 1:
Business (Organizational) Understanding

- Πώς μπορούμε να αυξήσουμε το περιθώριο κέρδους ανά μονάδα προϊόντος; Πώς μπορούμε να προβλέψουμε και να διορθώσουμε ατέλειες κατασκευής έτσι ώστε να αποφύγουμε την αποστολή ενός ελαττωματικού προϊόντος;
- Από εκεί, μπορείτε να αρχίσετε και να αναπτύξετε πιο συγκεκριμένες ερωτήσεις που θέλετε να απαντήσετε, και αυτό θα σας δώσει τη δυνατότητα να προχωρήσετε σε ...

CRISP-DM Step 2: Data Understanding

- Από πού προέρχονται τα δεδομένα; Από ποιόν συλλέγονται; Χρησιμοποιήθηκε μια τυποποιημένη μέθοδος συλλογής (a standard method of collection); Τι σημαίνουν οι διάφορες στήλες και οι γραμμές των δεδομένων; Υπάρχουν ακρωνύμια ή συντομογραφίες που είναι άγνωστα ή ασαφή;

CRISP-DM Step 3: Data Preparation

(Data Mining for the Masses)

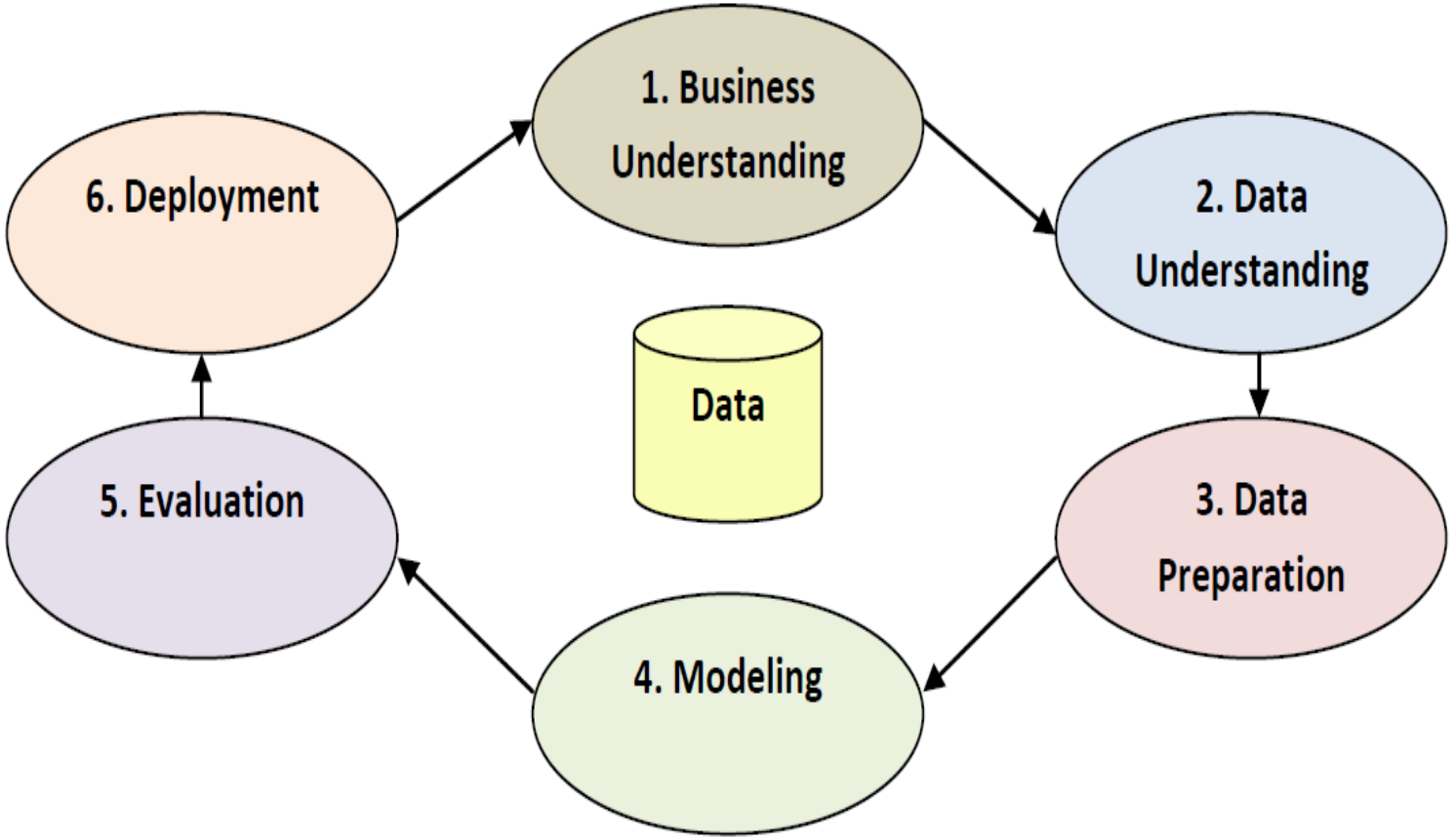
- Η **Προετοιμασία των δεδομένων (Data Preparation)** περιλαμβάνει μια σειρά από δραστηριότητες.
- Μπορεί να ενώνει δύο ή περισσότερα σύνολα δεδομένων, να περιορίζει σύνολα δεδομένων μόνον σε εκείνες τις μεταβλητές που έχουν ενδιαφέρον σε μια συγκεκριμένη περίπτωση εξόρυξης δεδομένων, να καθαρίζει δεδομένα από «ακραίες» παρατηρήσεις, να συμπληρώνει – διαχειρίζεται ελλείποντα δεδομένα, να μορφοποιεί εκ νέου δεδομένα για λόγους συνέπειας κ.λπ.

CRISP-DM Step 4: Modeling

(Data Mining for the Masses)

- Απλουστεύοντας, ένα **μοντέλο**, στην εξόρυξη δεδομένων, είναι μια ηλεκτρονική αναπαράσταση παρατηρήσεων – μετρήσεων (observations) του πραγματικού κόσμου. Τα μοντέλα προκύπτουν από την εφαρμογή αλγορίθμων που «αναλαμβάνουν» την αναζήτηση, τον εντοπισμό, και την εμφάνιση προτύπων ή μηνυμάτων στα δεδομένα.
- Υπάρχουν δύο βασικά είδη μοντέλων εξόρυξης: εκείνα που ταξινομούν **(classify)** και εκείνα που προβλέπουν (**predict)**.

CRISP-DM Conceptual Model



παράδειγμα: Data Preparation (με χρήση RapidMiner)

Ο διευθυντής μάρκετινγκ μιας μικρής εταιρείας σχεδιασμού εφαρμογών Διαδικτύου και διαφήμισης θέλει να αναπτύξει ένα σύνολο δεδομένων που θα περιέχει πληροφορίες σχετικά με τους χρήστες του Διαδικτύου. Η εταιρεία θα χρησιμοποιήσει αυτά τα στοιχεία για να καθορίσει τι είδους άνθρωποι χρησιμοποιούν το Διαδίκτυο και πώς η επιχείρηση θα είναι σε θέση να εμπορευτεί υπηρεσίες σε αυτή την ομάδα χρηστών.
Δημιουργεί μια online έρευνα και τοποθετεί συνδέσεις (links) σχετικές με διάφορες δημοφιλείς ιστοσελίδες. Μέσα σε δύο εβδομάδες, ο διευθυντής έχει συλλέξει αρκετά δεδομένα για να ξεκινήσει την ανάλυση, αλλά ο ίδιος θεωρεί ότι τα στοιχεία του πρέπει να κανονικοποιηθούν. Συνειδητοποιεί ότι κάποιες πρόσθετες εργασίες σχετικά με τα δεδομένα πρέπει να λάβουν χώρα πριν από την έναρξη της ανάλυσης.

(M. North, Data Mining for the Masses, 2012

Διασκευή παραδείγματος κεφαλαίου 3)

Survey & RapidMiner Attributes

Gender (Φύλο) M/F

Race (Φυλή), American, Hispanic, White, …

Birth_Year (Έτος γέννησης) 1981

Marital_Status (Οικογενειακή κατάσταση) M/S/D

Years_on_Internet (Χρόνια στο Διαδίκτυο) 14

Hours_Per_Day (Ώρες ανά ημέρα στο Διαδίκτυο) 2

Preferred_Browser (Προτιμώμενο πρόγραμμα περιήγησης) Chrome, Firefox, Internet Explorer, Safari

Preferred_Search_Engine: Προτιμώμενη μηχανή αναζήτησης Google, Yahoo, Bing

Preferred_Email: Προτιμώμενο ηλεκτρονικό ταχυδρομείο Hotmail, Yahoo, Gmail

Read_News Y/N

Survey & RapidMiner Attributes

Online_Shopping (Online αγορές) Y/N

Online_Gaming (Διαδικτυακό παιχνίδι) Y/N

Facebook Y/N

Twitter Y/N

Other_Social_Network (Άλλο κοινωνικό δίκτυο) African, LinkedIn

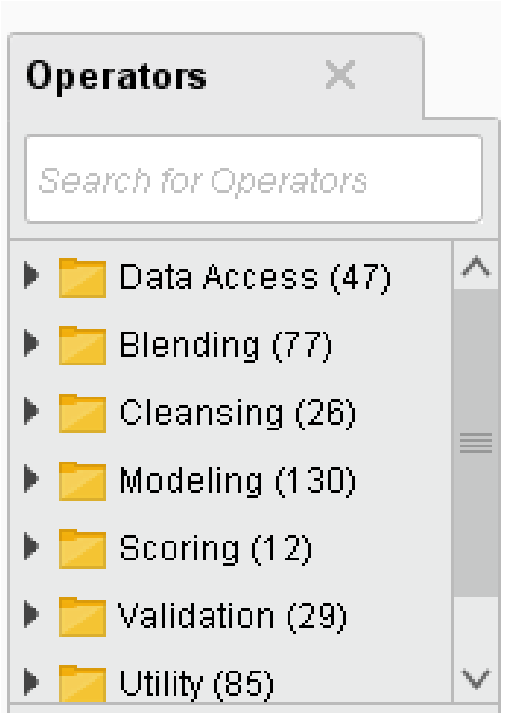
https://sites.google.com/site/dataminingforthemasses/

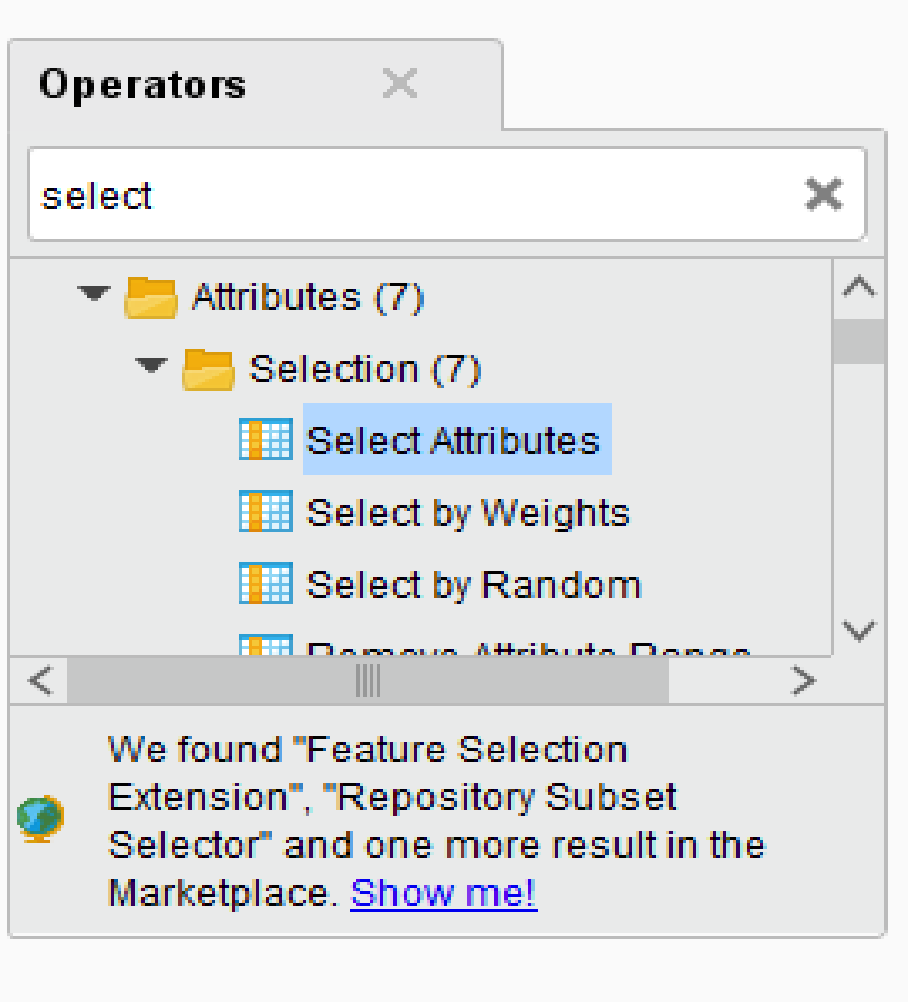
The process of data scrubbing
ή πως θα διαχειριστούμε με RapidMiner ανωμαλίες στα δεδομένα

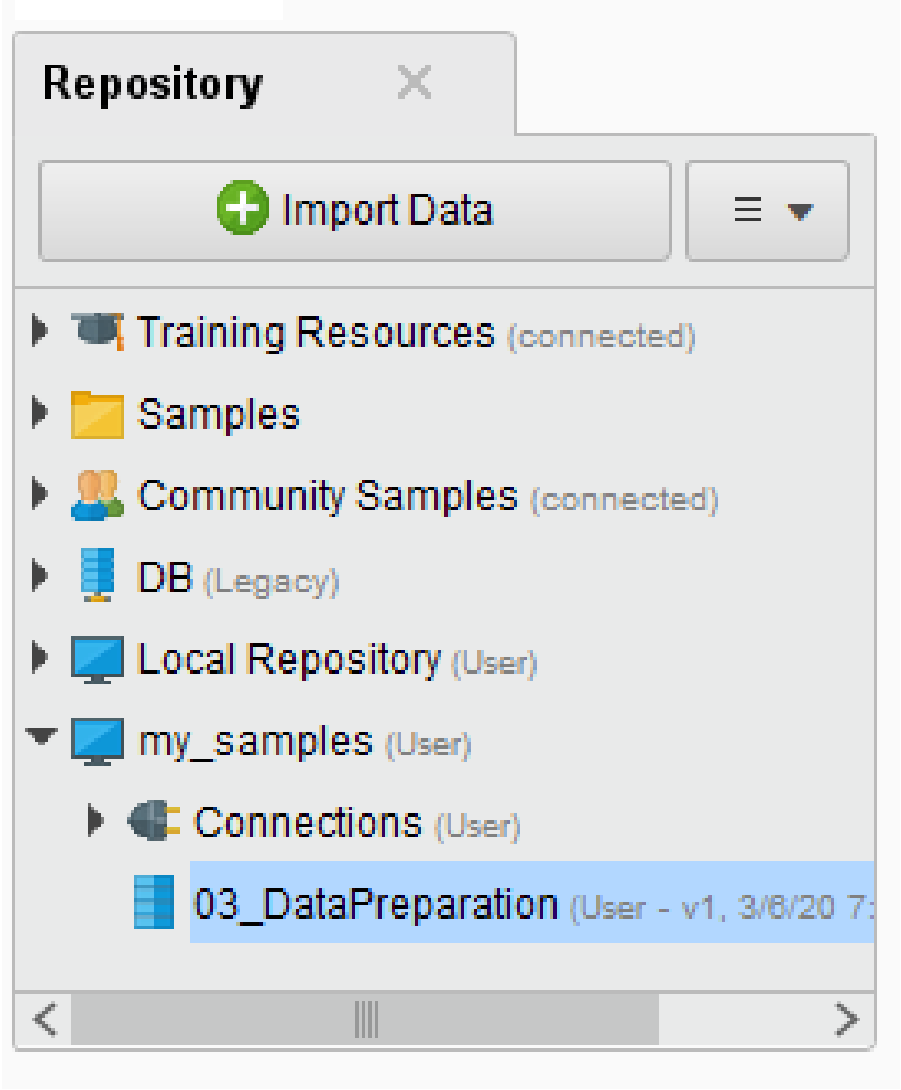
Τρόποι διαχείρισης data scrubbing:

1. handling missing data
2. reducing data (observations)
3. handling inconsistent data
4. reducing attributes.

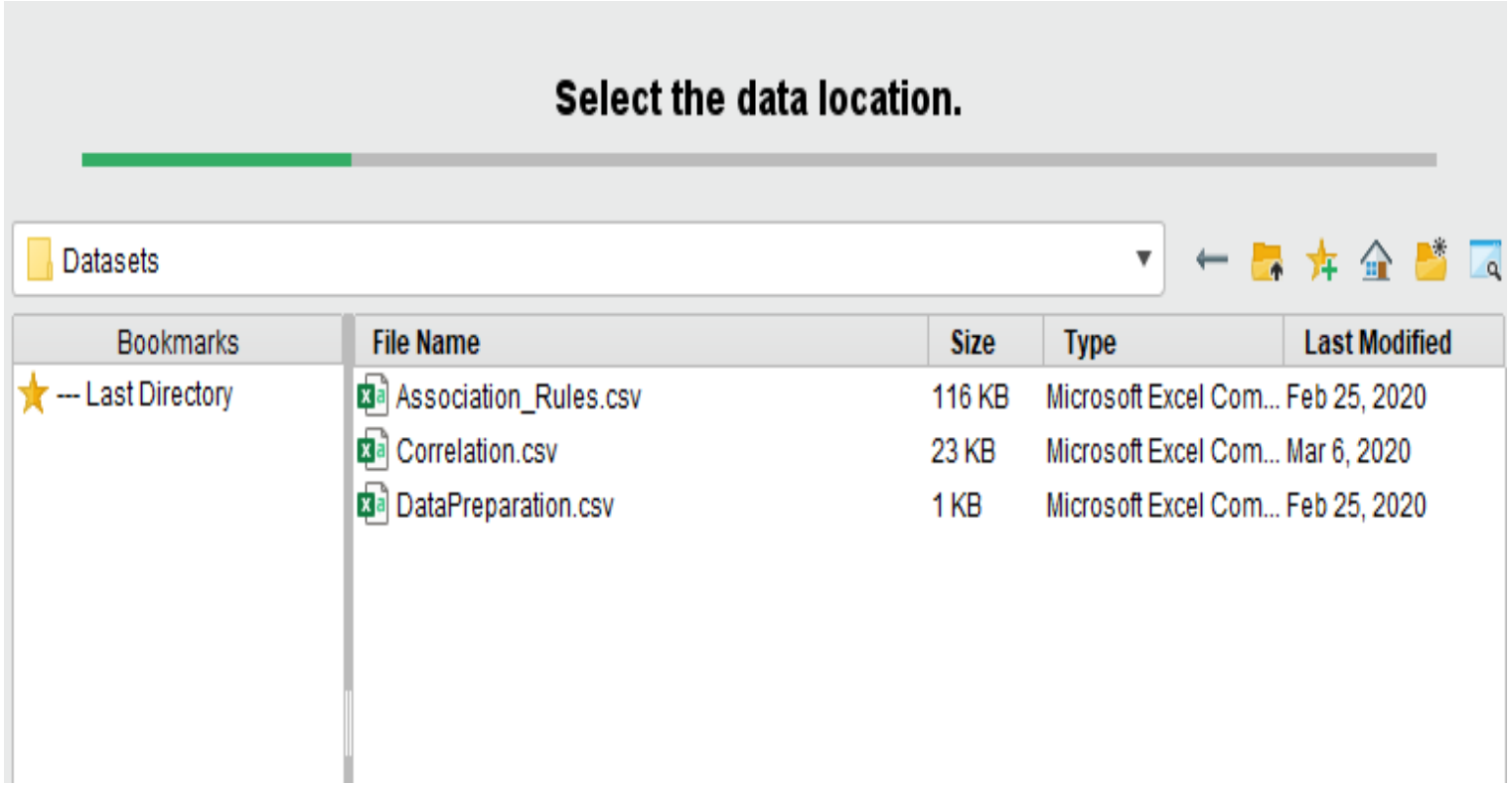
εργαλείο Rapid Miner



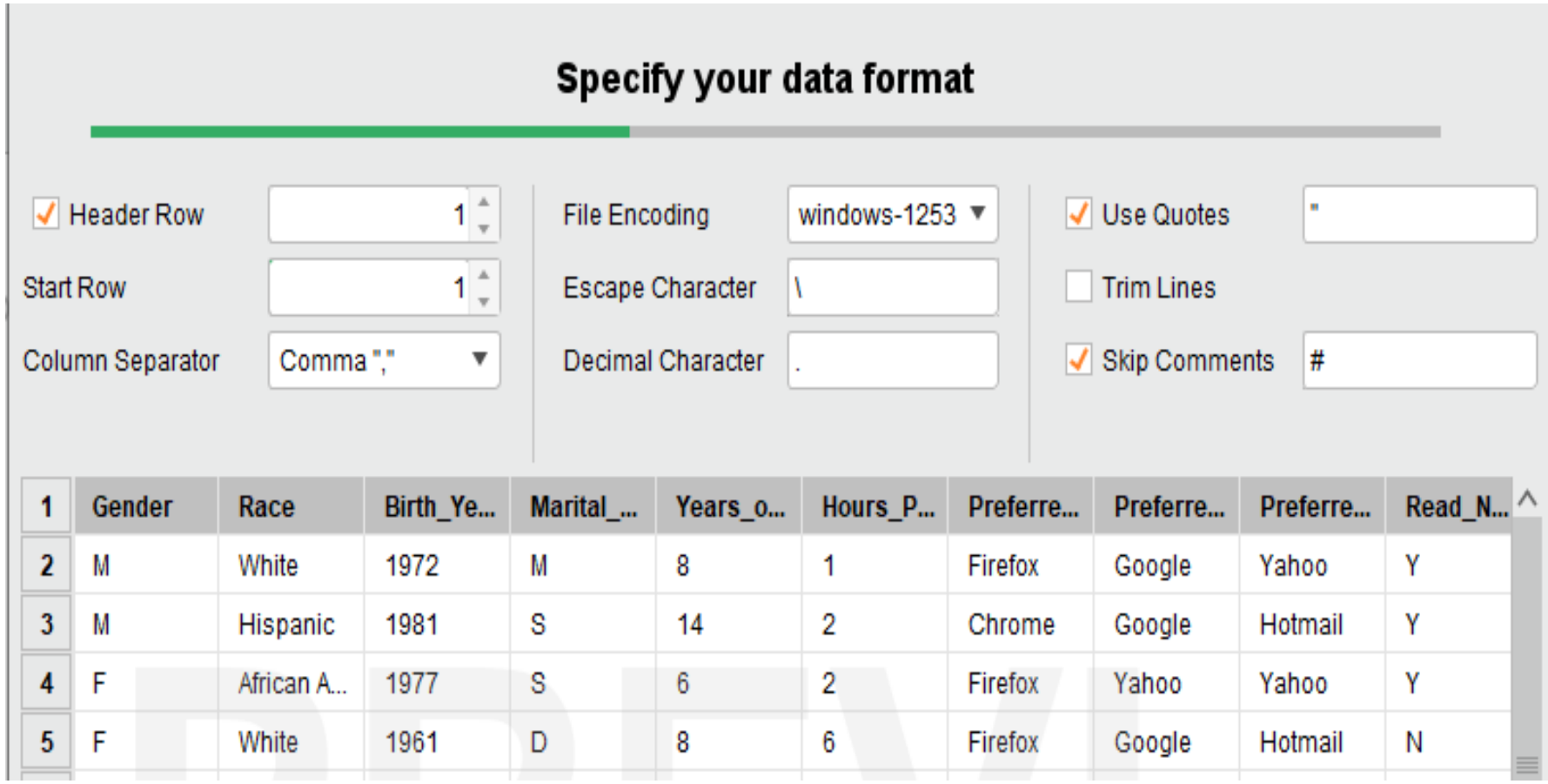



Repository
Import Data
Training Resources (connected)
Samples
Community Samples (connected)
DB (Legacy)
Local Repository (User)
my_samples (User)
Connections (User)
03_DataPreparation (User - v1, 3/6/20 7:

Φορτώνουμε το σύνολο δεδομένων DataPreparation.csv (Import)



Μπορούμε να δούμε τα δεδομένα (Data View)



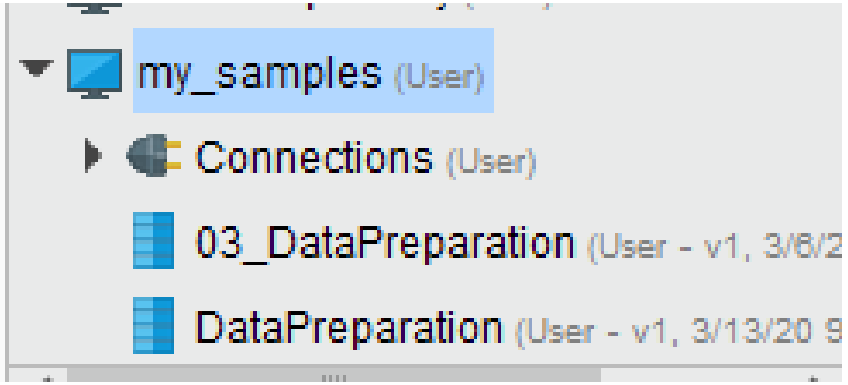
1	Gender	Race	Birth_Ye...	Marital_...	Years_o...	Hours_P...	Preferre...	Preferre...	Preferre...	Read_N...
2	M	White	1972	M	8	1	Firefox	Google	Yahoo	Y
3	M	Hispanic	1981	S	14	2	Chrome	Google	Hotmail	Y
4	F	African A...	1977	S	6	2	Firefox	Yahoo	Yahoo	Y
5	F	White	1961	D	8	6	Firefox	Google	Hotmail	N

Format your columns.

Date format Enter value...

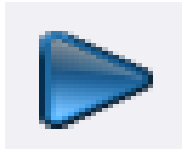
Replace errors with missing values

	Gender *polynominal*	Race *polynominal*	Birth_Year *integer*	Marital_Sta... *polynominal*	Years_on_I... *integer*	Hours_Per_... *integer*
1	M	White	1972	M	8	1
2	M	Hispanic	1981	S	14	2
3	F	African American	1977	S	6	2
4	F	White	1961	D	8	6



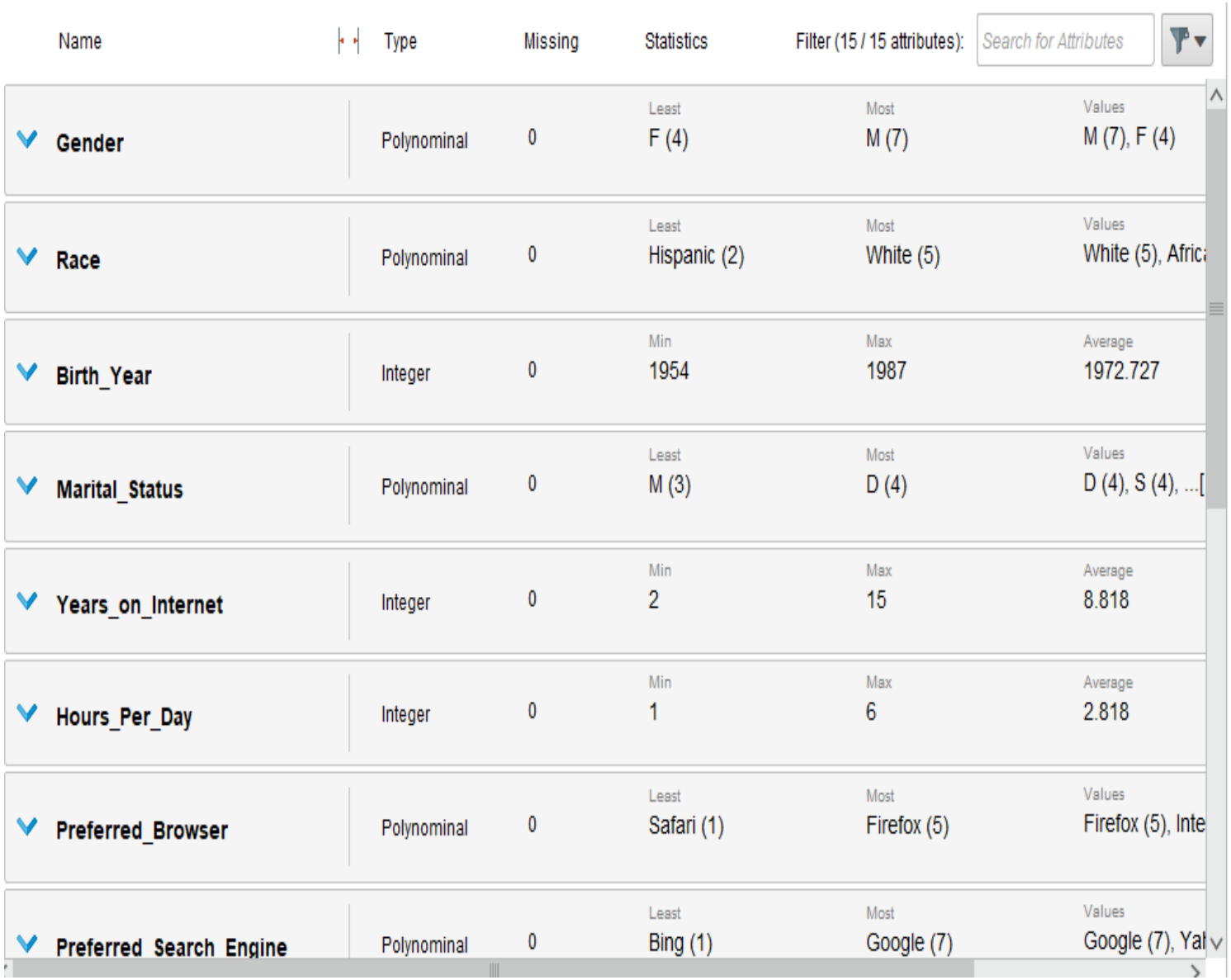
Design Perspective

εκτέλεση (Results Perspective)



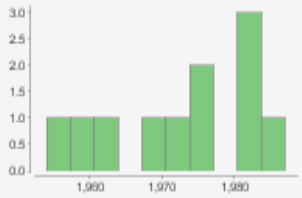
Row No.	Gender	Race	Birth_Year	Marital_Stat...	Years_on_In...	Hours_Per_...	Preferred_B...	Preferred_S...	Preferr
1	M	White	1972	M	8	1	Firefox	Google	Yahoo
2	M	Hispanic	1981	S	14	2	Chrome	Google	Hotmai
3	F	African Ameri...	1977	S	6	2	Firefox	Yahoo	Yahoo
4	F	White	1961	D	8	6	Firefox	Google	Hotmai

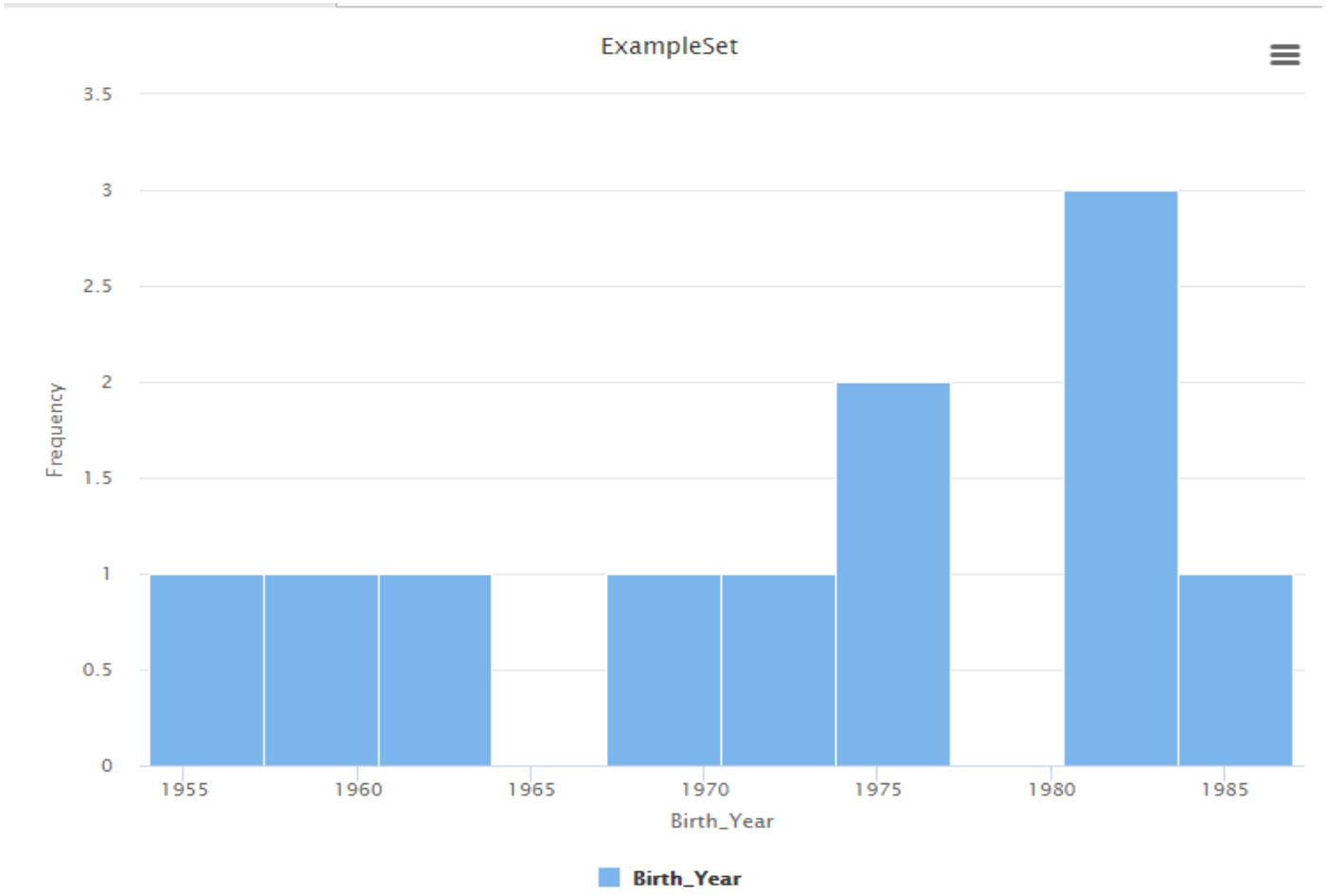
Statistics

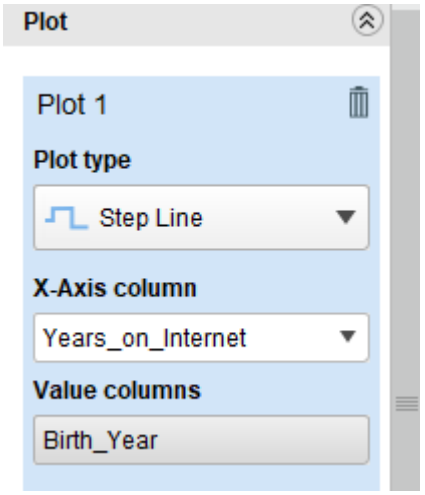


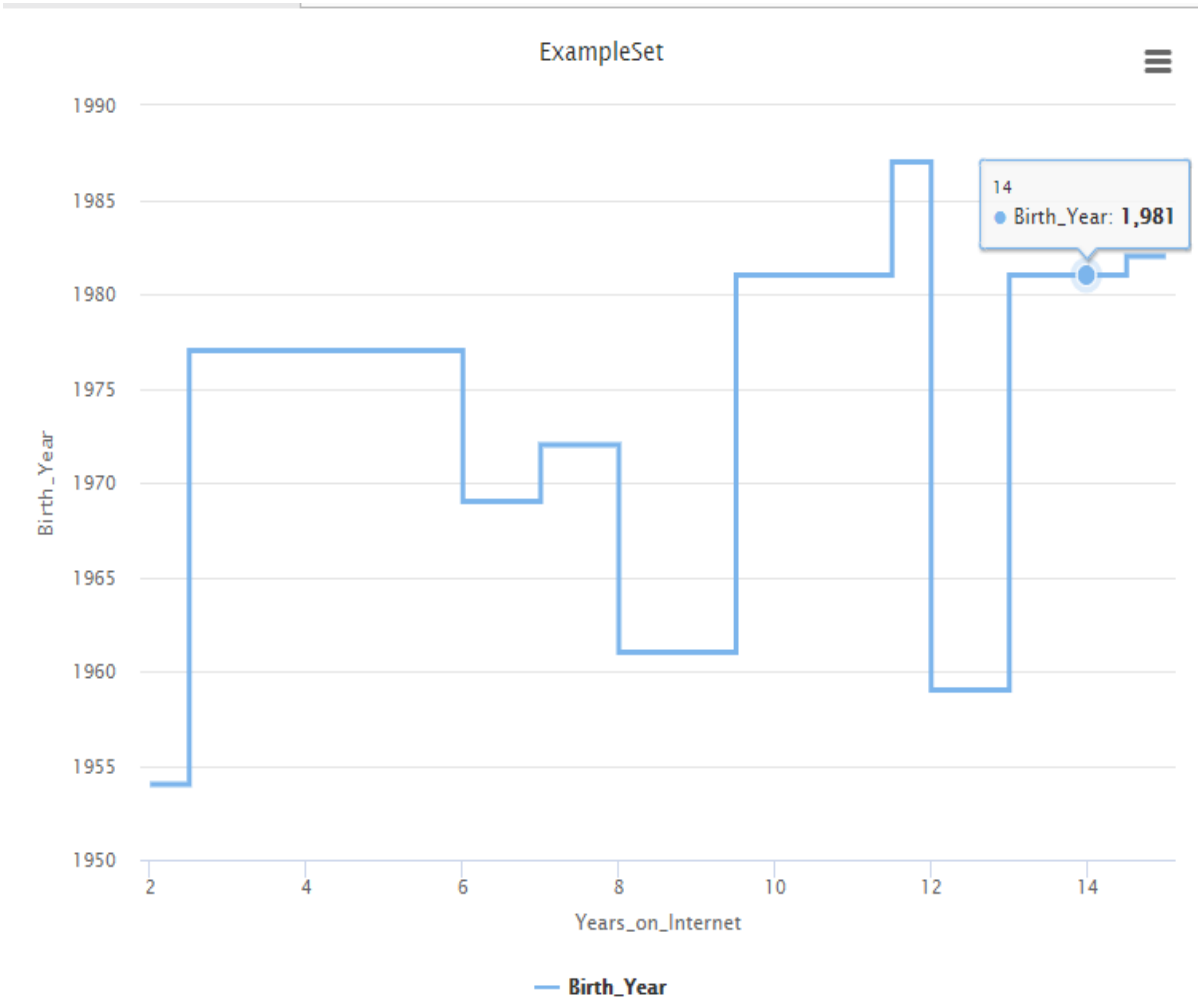
Name	Type	Missing	Statistics		Filter (15 / 15 attributes): Search for Attributes
Gender	Polynominal	0	Least: F (4)	Most: M (7)	Values: M (7), F (4)
Race	Polynominal	0	Least: Hispanic (2)	Most: White (5)	Values: White (5), Africa
Birth_Year	Integer	0	Min: 1954	Max: 1987	Average: 1972.727
Marital_Status	Polynominal	0	Least: M (3)	Most: D (4)	Values: D (4), S (4), ...
Years_on_Internet	Integer	0	Min: 2	Max: 15	Average: 8.818
Hours_Per_Day	Integer	0	Min: 1	Max: 6	Average: 2.818
Preferred_Browser	Polynominal	0	Least: Safari (1)	Most: Firefox (5)	Values: Firefox (5), Inte
Preferred_Search_Engine	Polynominal	0	Least: Bing (1)	Most: Google (7)	Values: Google (7), Yal

			Min	Max	Average
Birth_Year	Integer	0	1954	1987	1972.727

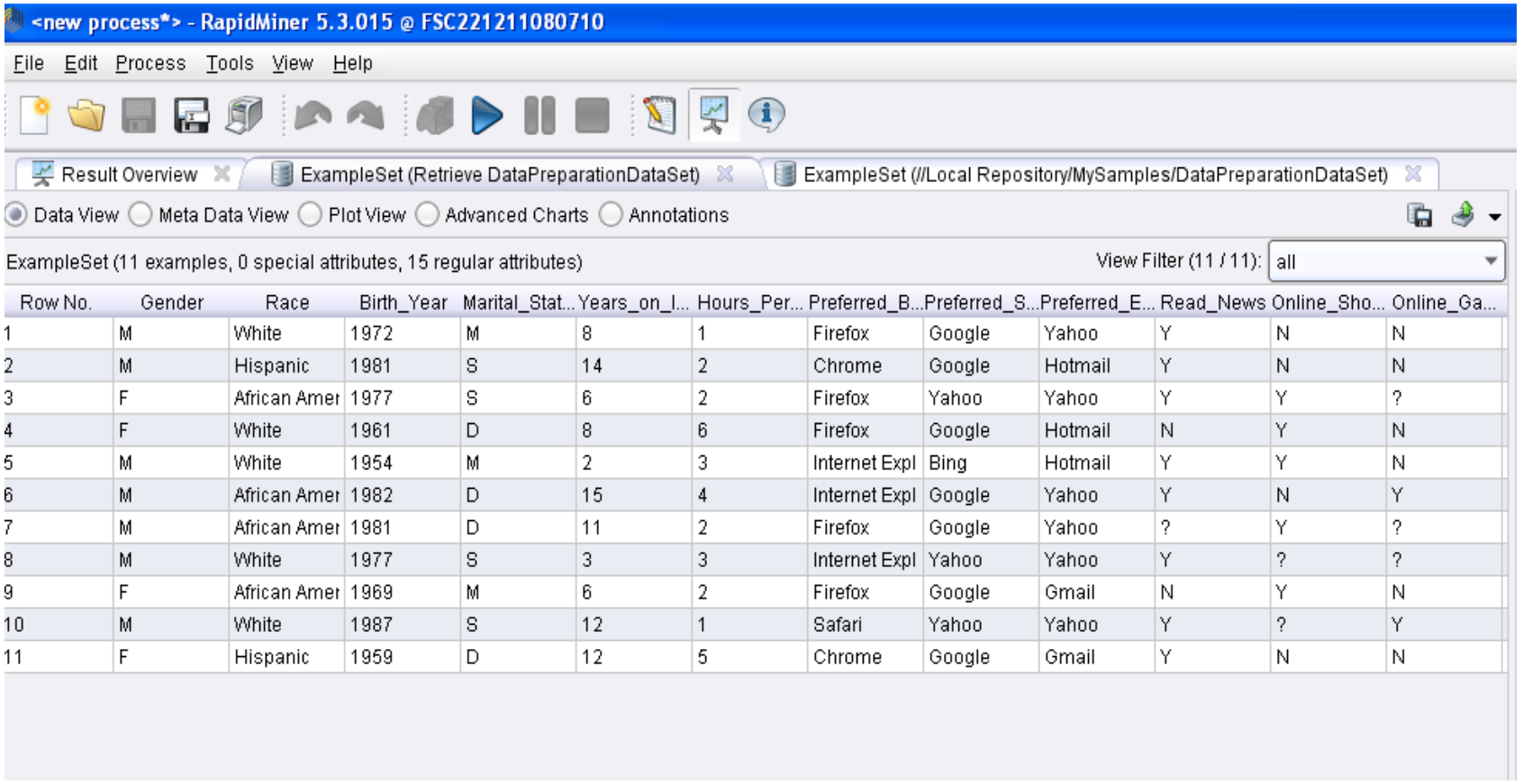
				Min	Max
Birth_Year	Integer	0	 Open visualizations	1954	1987


ExampleSet
3.5
3
2.5
2
1.5
1
0.5
0
Frequency
1955
1960
1965
1970
1975
1980
1985
Birth_Year
Birth_Year


Plot
Plot 1
Plot type
Step Line
X-Axis column
Years_on_Internet
Value columns
Birth_Year

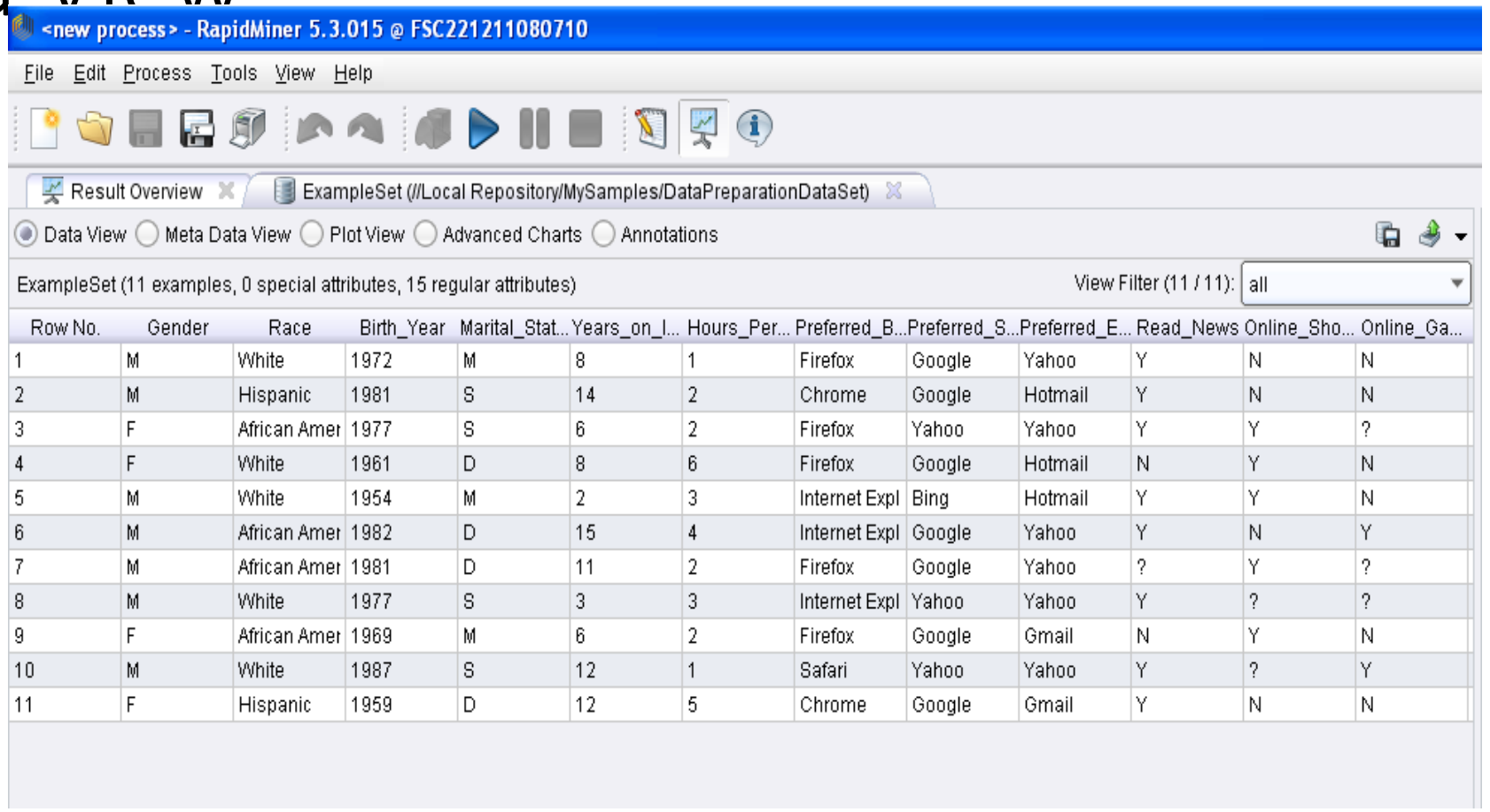

ExampleSet
1990
1985
1980
1975
1970
1965
1960
1955
1950
Birth_Year
2
4
6
8
10
12
14
Years_on_Internet
14
Birth_Year: 1,981
Birth_Year

Res Perspective



Row No.	Gender	Race	Birth_Year	Marital_Stat...	Years_on_I...	Hours_Per...	Preferred_B...	Preferred_S...	Preferred_E...	Read_News	Online_Sho...	Online_Ga...
1	M	White	1972	M	8	1	Firefox	Google	Yahoo	Y	N	N
2	M	Hispanic	1981	S	14	2	Chrome	Google	Hotmail	Y	N	N
3	F	African Amer	1977	S	6	2	Firefox	Yahoo	Yahoo	Y	Y	?
4	F	White	1961	D	8	6	Firefox	Google	Hotmail	N	Y	N
5	M	White	1954	M	2	3	Internet Expl	Bing	Hotmail	Y	Y	N
6	M	African Amer	1982	D	15	4	Internet Expl	Google	Yahoo	Y	N	Y
7	M	African Amer	1981	D	11	2	Firefox	Google	Yahoo	?	Y	?
8	M	White	1977	S	3	3	Internet Expl	Yahoo	Yahoo	Y	?	?
9	F	African Amer	1969	M	6	2	Firefox	Google	Gmail	N	Y	N
10	M	White	1987	S	12	1	Safari	Yahoo	Yahoo	Y	?	Y
11	F	Hispanic	1959	D	12	5	Chrome	Google	Gmail	Y	N	N

Data View



Row No.	Gender	Race	Birth_Year	Marital_Stat...	Years_on_I...	Hours_Per...	Preferred_B...	Preferred_S...	Preferred_E...	Read_News	Online_Sho...	Online_Ga...
1	M	White	1972	M	8	1	Firefox	Google	Yahoo	Y	N	N
2	M	Hispanic	1981	S	14	2	Chrome	Google	Hotmail	Y	N	N
3	F	African Amer	1977	S	6	2	Firefox	Yahoo	Yahoo	Y	Y	?
4	F	White	1961	D	8	6	Firefox	Google	Hotmail	N	Y	N
5	M	White	1954	M	2	3	Internet Expl	Bing	Hotmail	Y	Y	N
6	M	African Amer	1982	D	15	4	Internet Expl	Google	Yahoo	Y	N	Y
7	M	African Amer	1981	D	11	2	Firefox	Google	Yahoo	?	Y	?
8	M	White	1977	S	3	3	Internet Expl	Yahoo	Yahoo	Y	?	?
9	F	African Amer	1969	M	6	2	Firefox	Google	Gmail	N	Y	N
10	M	White	1987	S	12	1	Safari	Yahoo	Yahoo	Y	?	Y
11	F	Hispanic	1959	D	12	5	Chrome	Google	Gmail	Y	N	N

Meta Data View (basic descriptive statistics)

Result Overview | ExampleSet (Retrieve DataPreparationDataSet) | ExampleSet (//Local Repository/MySamples/DataPreparationDataSet)

Data View | Meta Data View | Plot View | Advanced Charts | Annotations

ExampleSet (11 examples, 0 special attributes, 15 regular attributes)

Role	Name	Type	Statistics	Range	Missings
regular	Gender	binominal	mode = M (7), least = F (4)	M (7), F (4)	0
regular	Race	polynominal	mode = White (5), least = His	White (5), Hispanic (2), Africa	0
regular	Birth_Year	integer	avg = 1972.727 +/- 10.743	[1954.000 ; 1987.000]	0
regular	Marital_Status	polynominal	mode = S (4), least = M (3)	M (3), S (4), D (4)	0
regular	Years_on_Internet	integer	avg = 8.818 +/- 4.332	[2.000 ; 15.000]	0
regular	Hours_Per_Day	integer	avg = 2.818 +/- 1.601	[1.000 ; 6.000]	0
regular	Preferred_Browser	polynominal	mode = Firefox (5), least = Sa	Firefox (5), Chrome (2), Intern	0
regular	Preferred_Search_Engine	polynominal	mode = Google (7), least = B	Google (7), Yahoo (3), Bing (	0
regular	Preferred_Email	polynominal	mode = Yahoo (6), least = Gr	Yahoo (6), Hotmail (3), Gmai	0
regular	Read_News	binominal	mode = Y (8), least = N (2)	Y (8), N (2)	1
regular	Online_Shopping	binominal	mode = Y (5), least = N (4)	N (4), Y (5)	2
regular	Online_Gaming	binominal	mode = N (6), least = Y (2)	N (6), Y (2)	3
regular	Facebook	binominal	mode = Y (8), least = N (3)	Y (8), N (3)	0
regular	Twitter	polynominal	mode = N (8), least = 99 (1)	N (8), Y (2), 99 (1)	0
regular	Other_Social_Network	polynominal	mode = LinkedIn (2), least =	LinkedIn (2), MySpace (1), Gr	7

Data Type

- Σε κάθε χαρακτηριστικό του συνόλου δεδομένων (data set) έχει εκχωρηθεί από το εργαλείο ένας τύπος δεδομένων με βάση το είδος των δεδομένων που είναι αποθηκευμένα στο χαρακτηριστικό.
- Οι τύποι δεδομένων ανήκουν σε τρεις κατηγορίες: Character (Text), Numeric, Date/Time. Στην κατηγορία Character (Text) το RapidMiner έχει τύπους δεδομένων Polynominal, Binominal, κλπ., στην κατηγορία Numeric έχει τύπους Real, Integer, κλπ. Ο Τύπος δεδομένων Binomial είναι Numeric και σημαίνει έναν από τους δύο αριθμούς (συνήθως 0 και 1).
- Ο Τύπος **Binominal** σημαίνει μία από τις δύο τιμές και μπορεί να είναι Character ή Numeric.

(M. North, Data Mining for the Masses, 2012

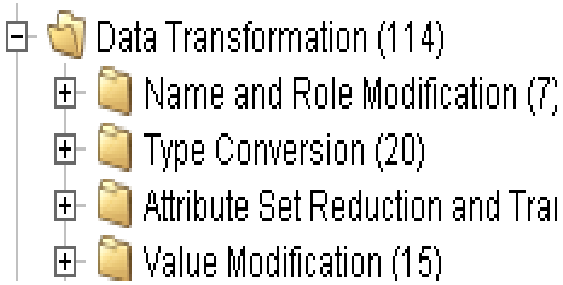
Data Type

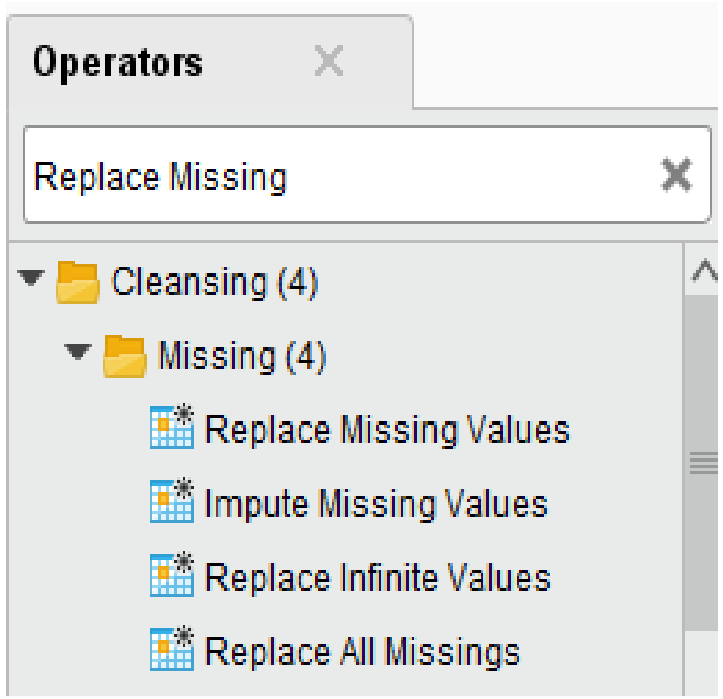
Παρατηρήστε ότι κάποια χαρακτηριστικά έχουν ελλιπείς τιμές (missing values). Για παράδειγμα το χαρακτηριστικό Online_Gaming attribute έχει 3 ελλιπείς τιμές.

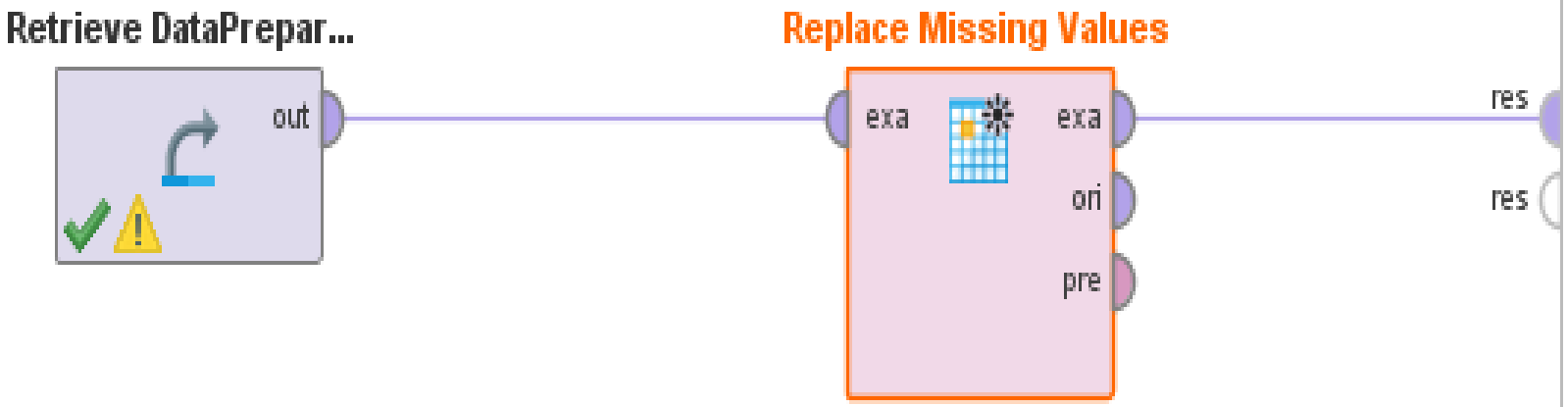
Online_Gaming attribute (3 missing values)

Preferred_Email	Polynominal	0
Read_News	Polynominal	1
Online_Shopping	Polynominal	2
Online_Gaming	Polynominal	3
Facebook	Polynominal	0
Twitter	Polynominal	0
Other_Social_Network	Polynominal	7

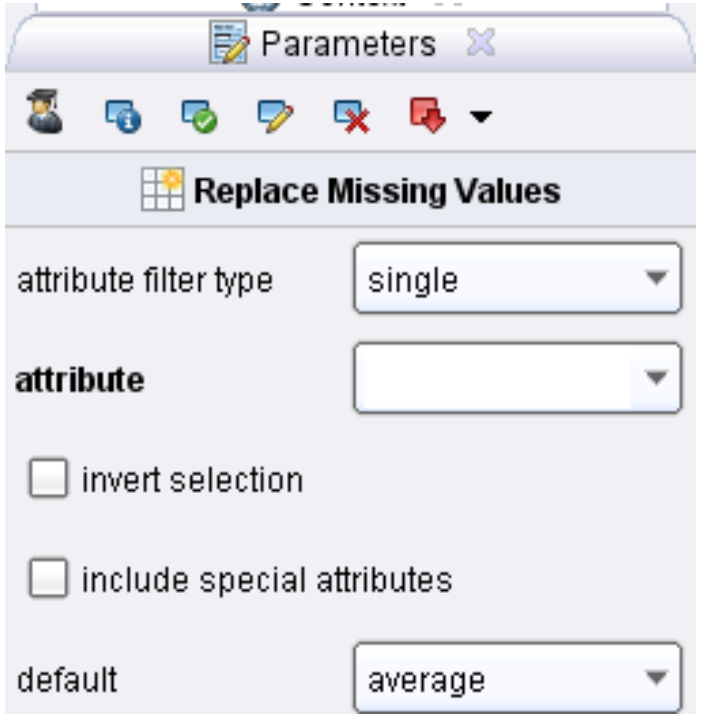
Θα διορθώσουμε αυτό το πρόβλημα χρησιμοποιώντας τον τελεστή Replace Missing Values

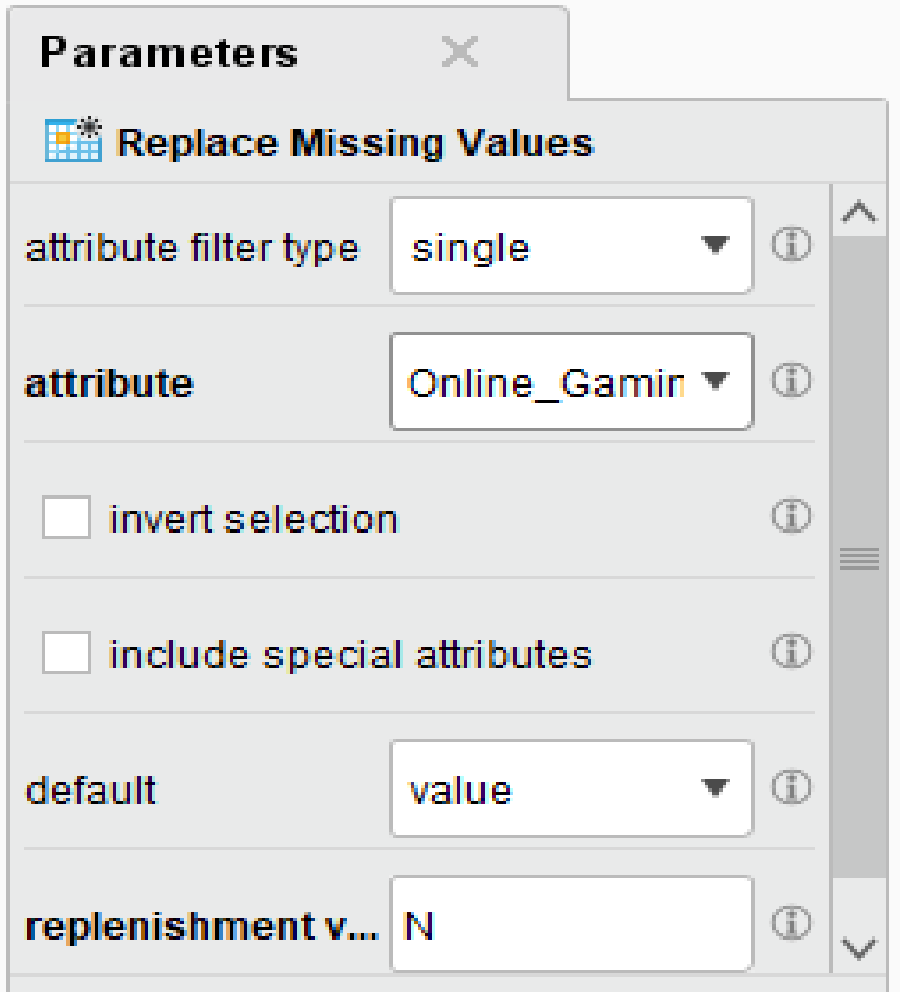






Η αλλαγή των ελλιπών τιμών (missing values) και η εκχώρηση συγκεκριμένων τιμών γίνεται με χρήση μεταβλητών στο παράθυρο παραμέτρων (parameter pane). Στην περίπτωσή μας κάνουμε κλικ στον τελεστή οπότε βλέπουμε το αντίστοιχο παράθυρο.





Εκτελούμε  και βλέπουμε ότι το πρόβλημα διορθώθηκε

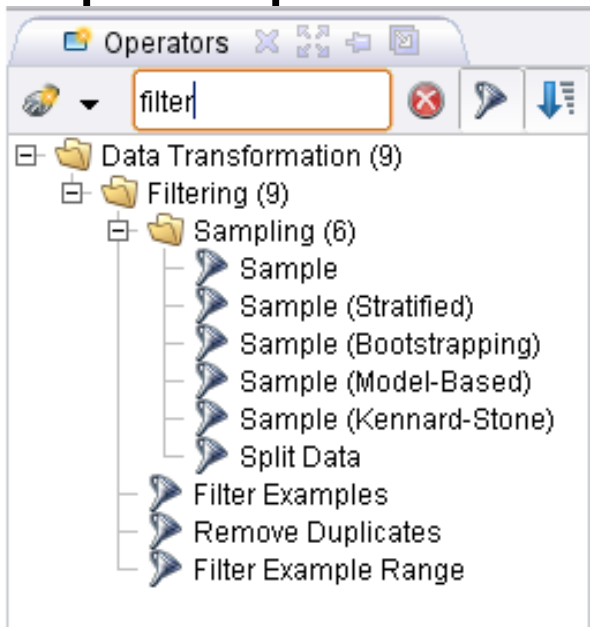
Data

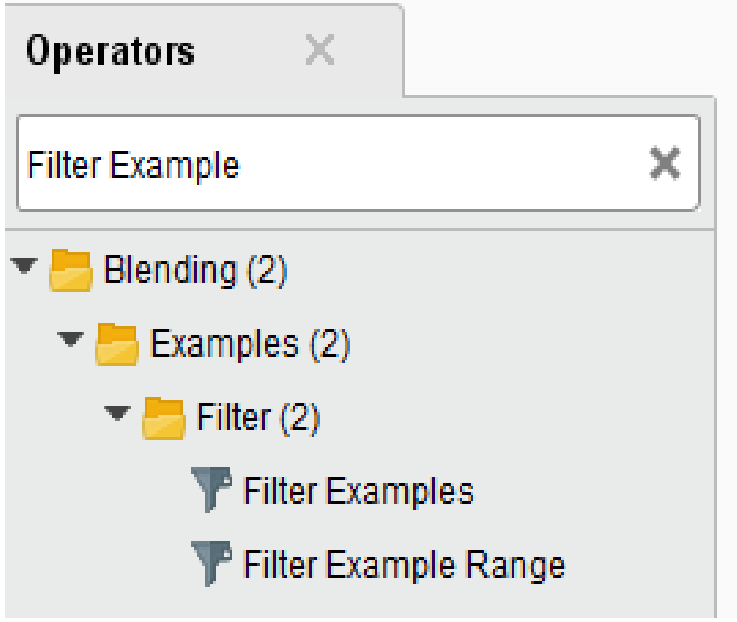
Row No.	Online_Gam...
1	N
2	N
3	N
4	N
5	N
6	Y
7	N
8	N
9	N
10	Y
11	N

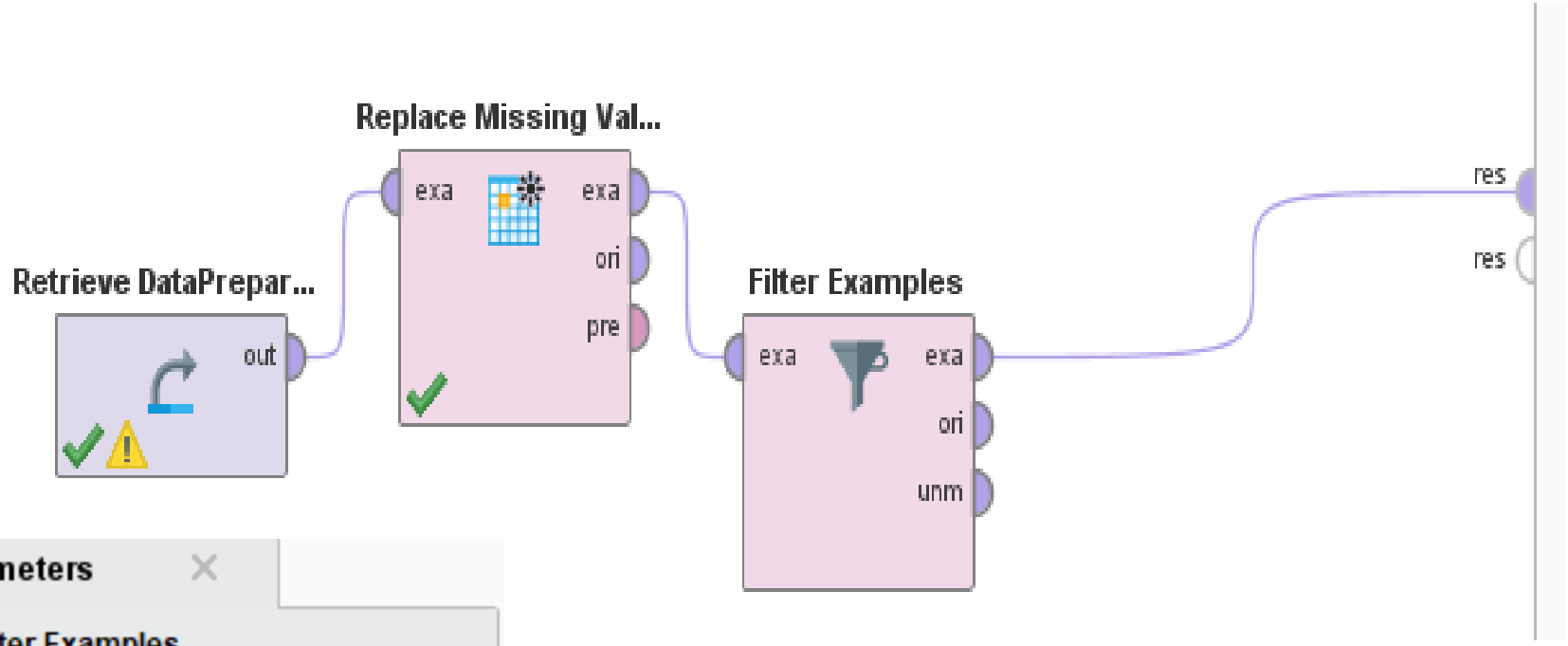
Statistics

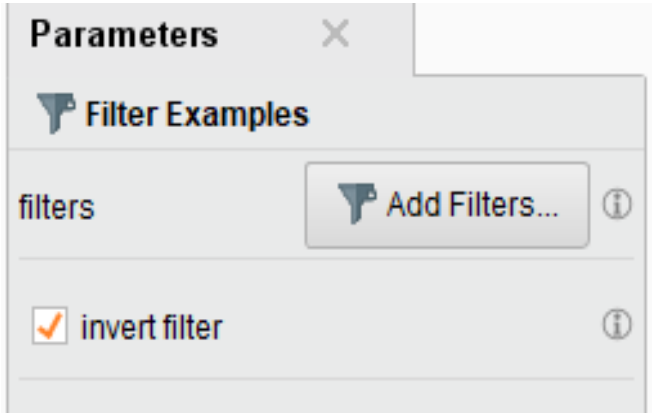
Name	Type	Missing	Statistics	Filter (15 / 15 attributes): Search for Attributes	
			Least	Most	Values
Online_Gaming	Polynominal	0	Y (2)	N (6)	N (6), N (3), ...[

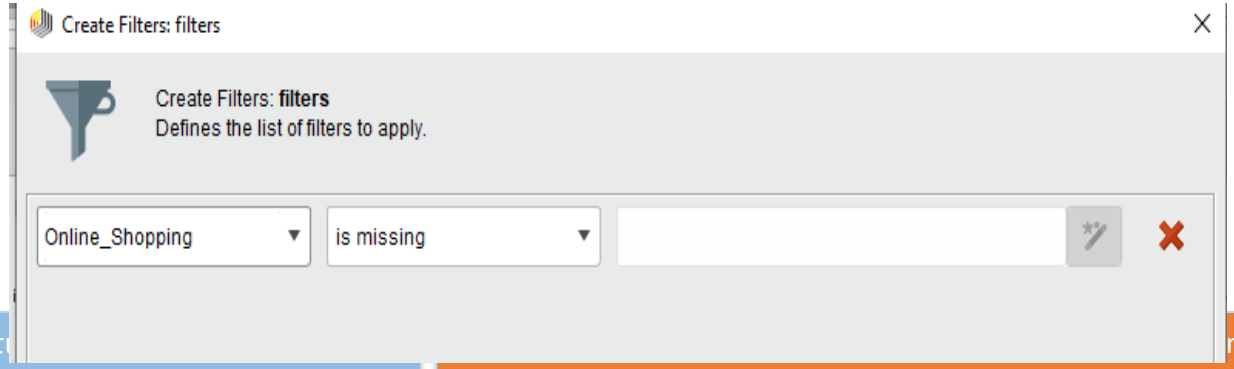
Διαχείριση Data Reduction






Replace Missing Val...
Retrieve DataPrepar...
Filter Examples
exa
out
exa
ori
pre
exa
exa
ori
unm
res
res

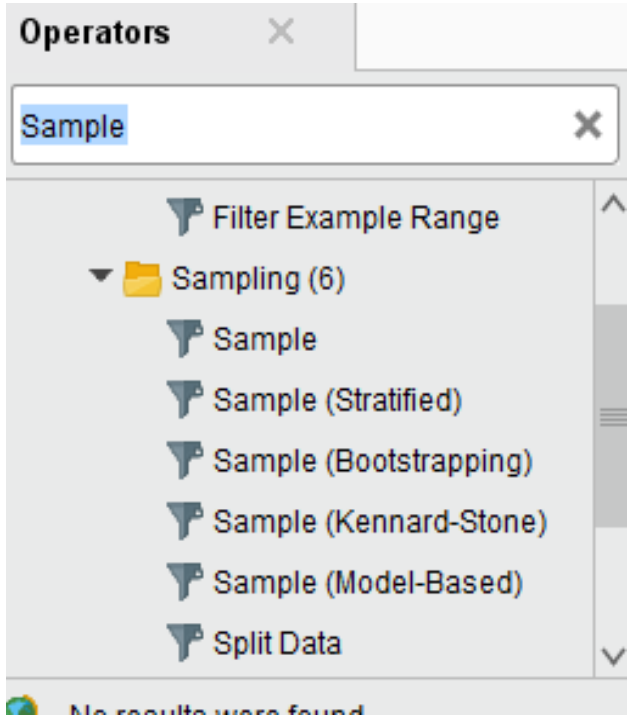

Parameters
Filter Examples
filters
Add Filters...
invert filter

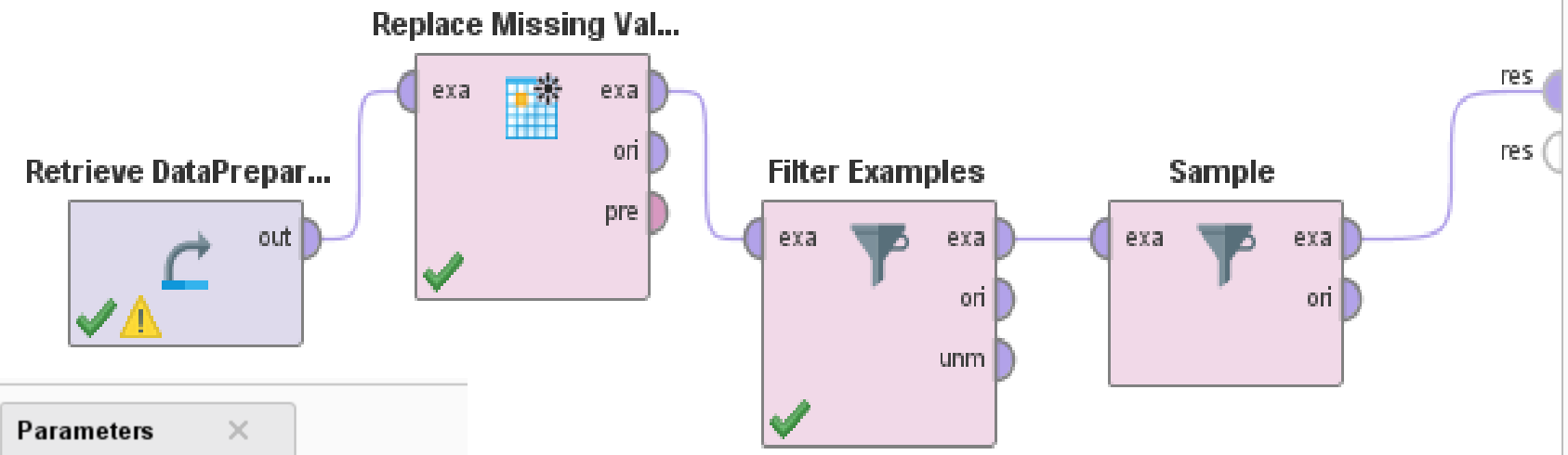

Create Filters: filters
Create Filters: filters
Defines the list of filters to apply.
Online_Shopping
is missing

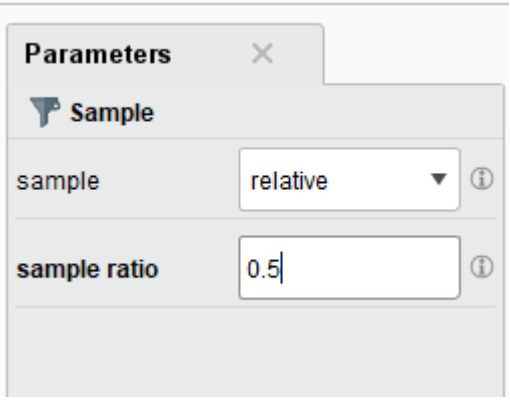
Φιλτράραμε τις τιμές και τώρα βλέπουμε εννέα παρατηρήσεις αντί για 11.

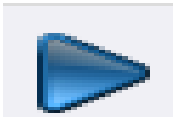
Row No.	Online_Gam...	Gender	Race	Birth_Year	Marital_Stat...	Years_on_In...	Hours_Per_...	Preferred_B...	Preferr
1	N	M	White	1972	M	8	1	Firefox	Google
2	N	M	Hispanic	1981	S	14	2	Chrome	Google
3	N	F	African Ameri...	1977	S	6	2	Firefox	Yahoo
4	N	F	White	1961	D	8	6	Firefox	Google
5	N	M	White	1954	M	2	3	Internet Explo...	Bing
6	Y	M	African Ameri...	1982	D	15	4	Internet Explo...	Google
7	N	M	African Ameri...	1981	D	11	2	Firefox	Google
8	N	F	African Ameri...	1969	M	6	2	Firefox	Google
9	N	F	Hispanic	1959	D	12	5	Chrome	Google

Διαχείριση δείγματος (Sample)







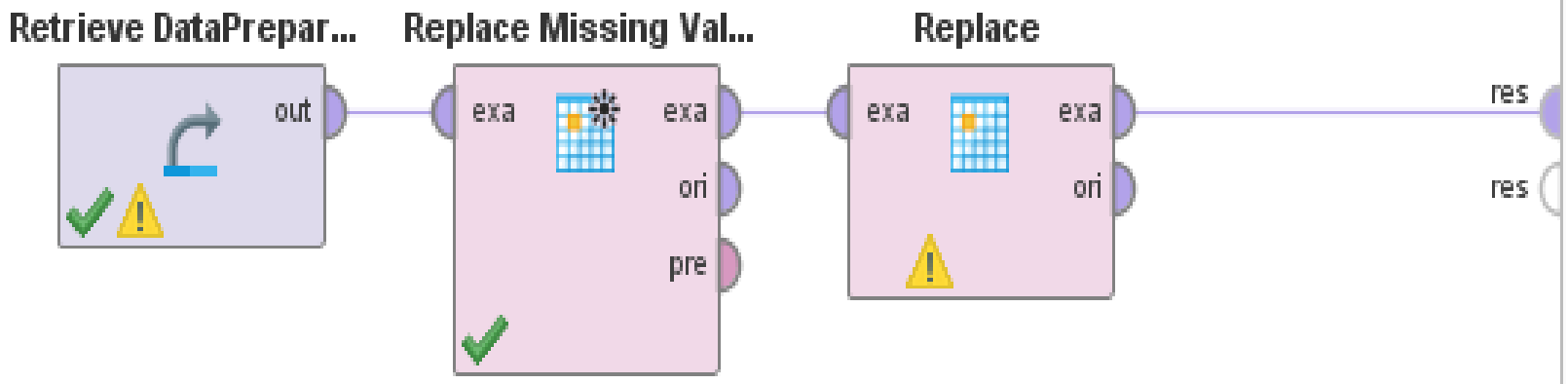


Row No.	Online_Gam...	Gender	Race	Birth_Year	Marital_Stat...	Years_on_In...	Hours_Per_...	Preferred_B...	Preferr
1	N	M	White	1972	M	8	1	Firefox	Google
2	N	F	White	1961	D	8	6	Firefox	Google
3	N	M	White	1954	M	2	3	Internet Explo...	Bing
4	N	F	Hispanic	1959	D	12	5	Chrome	Google

Διαχείριση ασυνεπών δεδομένων (Handling Inconsistent Data)

Το χαρακτηριστικό twitter έχει μία τιμή 99 αντί για τιμή Y/N.

Twitter	Other_Socia...
N	?
N	?
N	?
Y	?
N	?
N	?
Y	LinkedIn
99	LinkedIn
N	?
N	MySpace
N	Google+



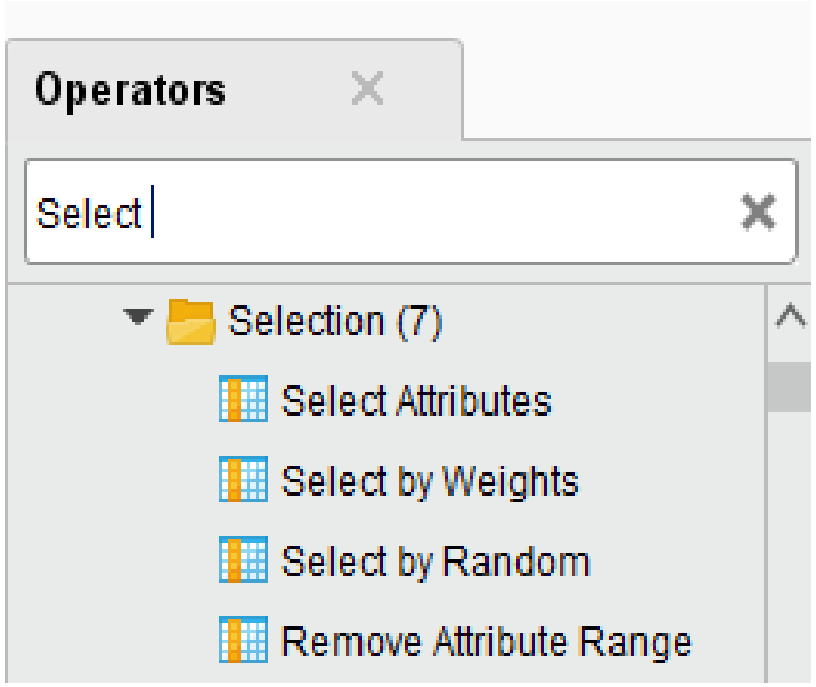
Parameters

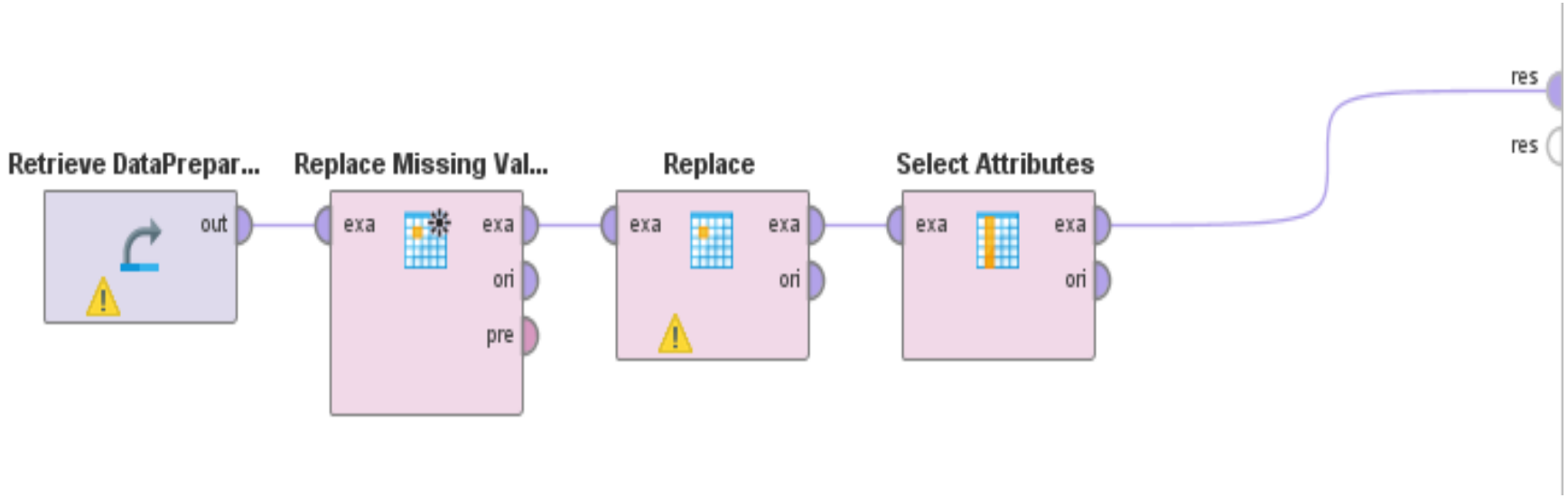
Replace

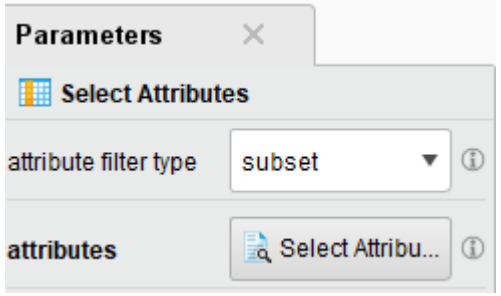
attribute filter type	single
attribute	Twitter
invert selection	
include special attributes	
replace what	99
replace by	N

Twitter
N
N
N
Y
N
N
Y
N
N
N
N

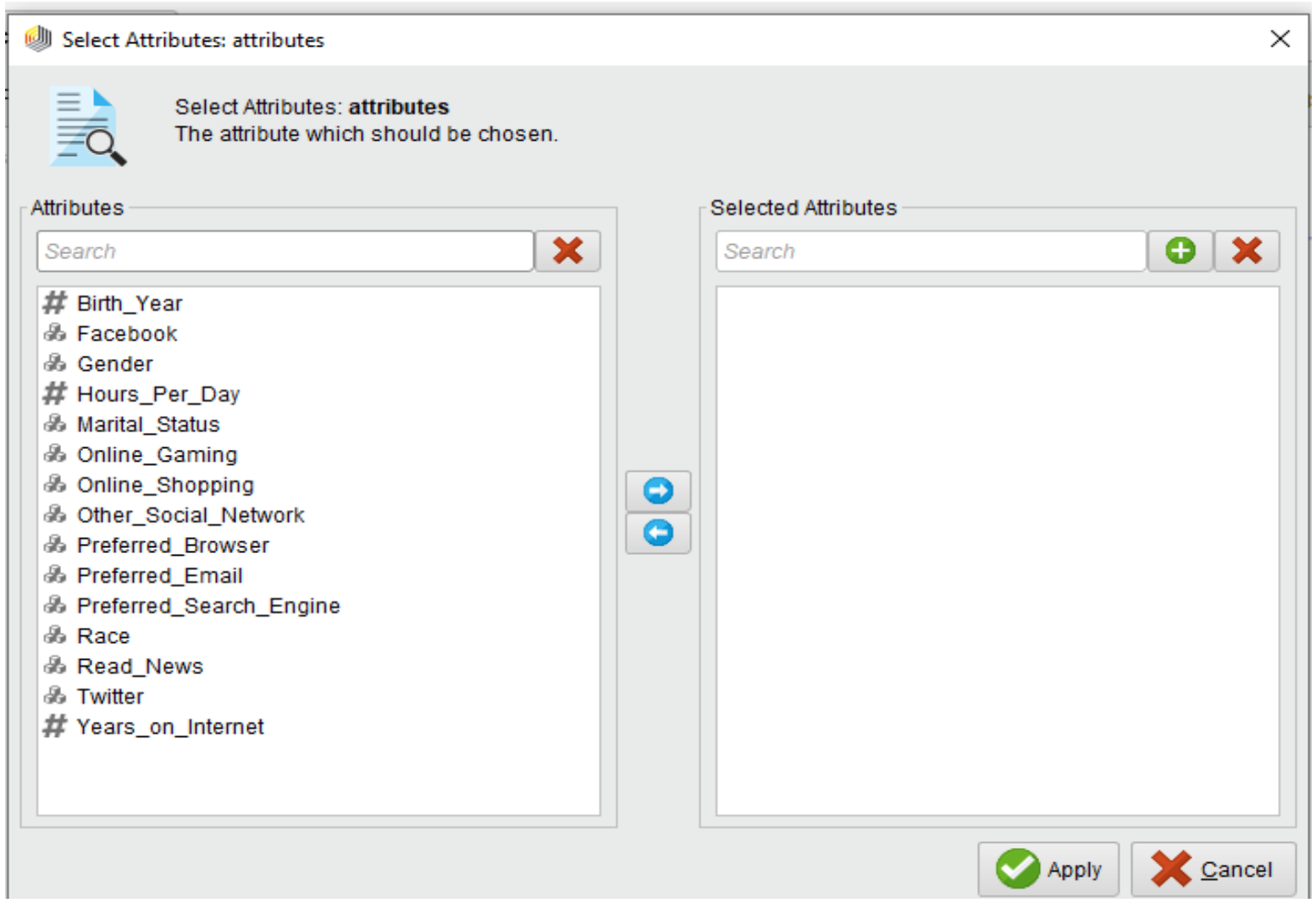
Διαχείριση Attribute Reduction




Retrieve DataPrepar...
out
Replace Missing Val...
exa
exa
ori
pre
Replace
exa
exa
ori
Select Attributes
exa
exa
ori
res
res


Parameters
Select Attributes
attribute filter type
subset
attributes
Select Attribu...


Select Attribute...


Select Attributes: attributes
Select Attributes: attributes
The attribute which should be chosen.
Attributes
Search
Birth_Year
Facebook
Gender
Hours_Per_Day
Marital_Status
Online_Gaming
Online_Shopping
Other_Social_Network
Preferred_Browser
Preferred_Email
Preferred_Search_Engine
Race
Read_News
Twitter
Years_on_Internet
Selected Attributes
Search
Apply
Cancel

ΤΕΛΟΣ ΕΝΟΤΗΤΑΣ