Data Import

RDS ANALYST

Lots of data formats supported

Valid File Formats:

- RDS Data (does not need conversion):

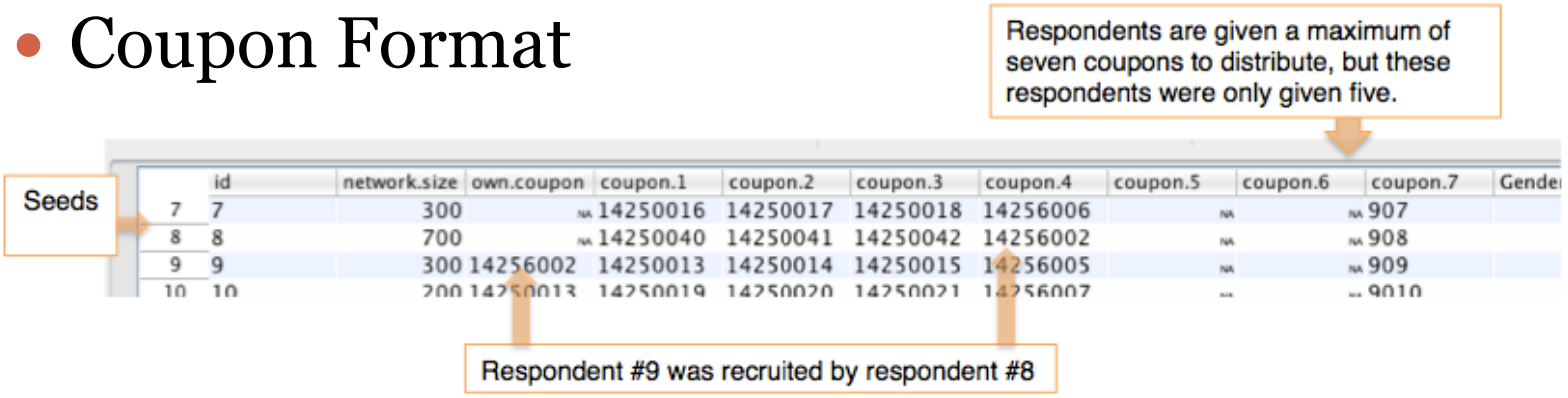
File Type	Extension
RDS Object	*.rdsobj, *.rdsat
R Object	*.robj

- Accepted Data Types (must be converted to RDS format for RDS analysis):

Comma separated	*.csv
Text file	*.txt
SPSS	*.sav
SAS export	*.xpt
DBase	*.dbf
Stata	*.dta
Systat	*.sys and *.syd
ARFF	*.arff
Epiinfo	*.rec
Minitab	*.mtp
S data dump	*.s3

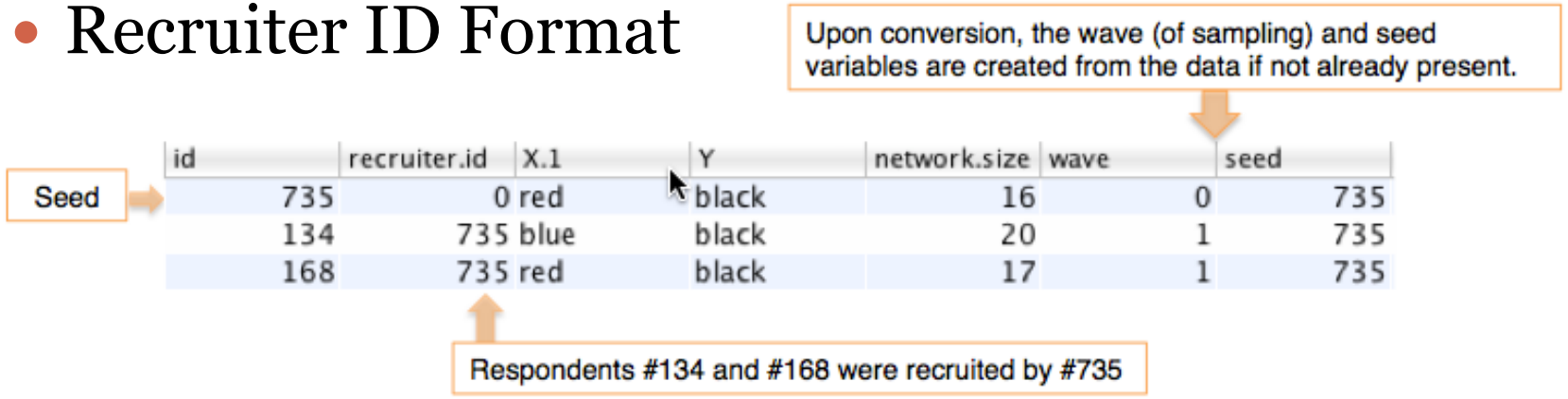
Two File Structures Supported

- Coupon Format



	id	network.size	own.coupon	coupon.1	coupon.2	coupon.3	coupon.4	coupon.5	coupon.6	coupon.7	Gende
7	7	300	NA	14250016	14250017	14250018	14256006	NA	NA	907	
8	8	700	NA	14250040	14250041	14250042	14256002	NA	NA	908	
9	9	300	14256002	14250013	14250014	14250015	14256005	NA	NA	909	
10	10	200	14250013	14250019	14250020	14250021	14256007	...	...	9010	

- Recruiter ID Format



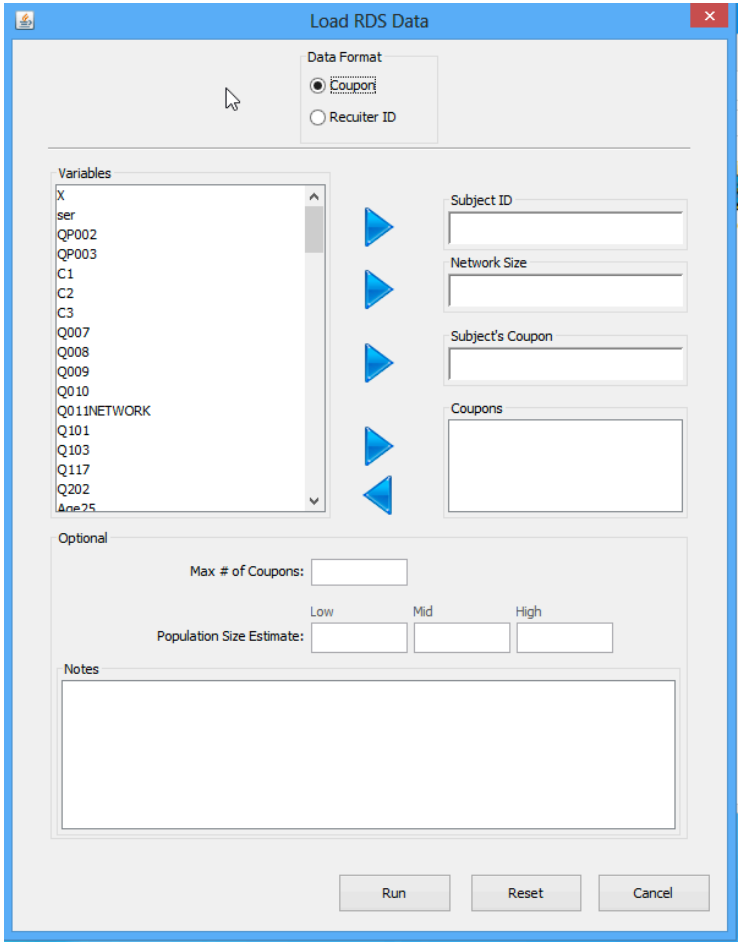
id	recruiter.id	X.1	Y	network.size	wave	seed
735	0	red	black	16	0	735
134	735	blue	black	20	1	735
168	735	red	black	17	1	735

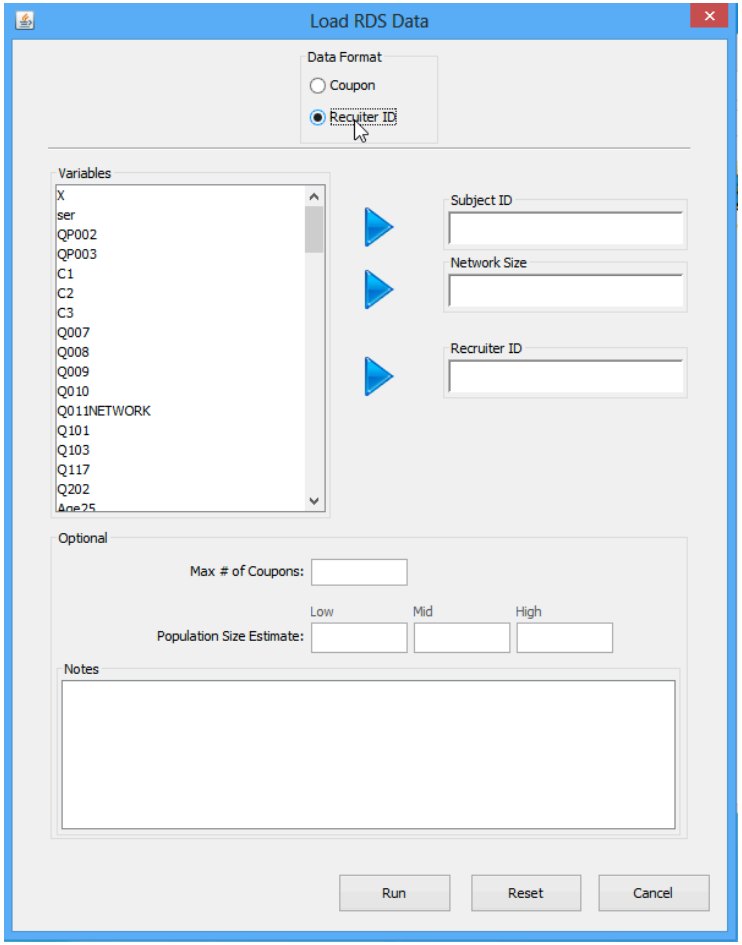
Steps to Import Data

1. Read data file into RDS Analyst.

2. Parse recruitment structure.

Let's do it!


Load RDS Data
Data Format
Coupon
Recuiter ID
Variables
X
ser
QP002
QP003
C1
C2
C3
Q007
Q008
Q009
Q010
Q011NETWORK
Q101
Q103
Q117
Q202
Age25
Subject ID
Network Size
Subject's Coupon
Coupons
Optional
Max # of Coupons:
Low
Mid
High
Population Size Estimate:
Notes
Run
Reset
Cancel


Load RDS Data
Data Format
Coupon
Recuiter ID
Variables
X
ser
QP002
QP003
C1
C2
C3
Q007
Q008
Q009
Q010
Q011NETWORK
Q101
Q103
Q117
Q202
Age25
Subject ID
Network Size
Recruiter ID
Optional
Max # of Coupons:
Low
Mid
High
Population Size Estimate:
Notes
Run
Reset
Cancel

Data Problems: Data Import

Problem

Each subject must have a unique non-missing ID

Symptom

Error message in console when importing

Solution

1. Edit data to insert missing ids
2. Remove extraneous rows

Data Problems: Importing Coupon Format

Problem

The subject's coupon must match exactly one of the coupon variables in the recruiter's row.

Symptom

Incorrect recruitment tree plots

Solution

1. Fix data entry errors in coupon variables.
2. Look for trailing and/or leading white space (spaces) in the coupon variables

Data Problems: Importing Recruiter ID Format

Problem

The subject's coupon must match exactly one of the coupon variables in the recruiter's row.

Symptom

Error shown in console, or incorrect recruitment tree plots

Solution

1. Fix data entry errors in the subject ID and recruiter ID variables.
2. Look for trailing and/or leading white space (spaces).

Data Problems: Importing Recruiter ID Format

Problem

All seeds must have the same recruiter ID which can not be present in the Subject ID

Symptom

Error shown in console or incorrect recruitment tree plots

Solution

1. Manually edit the recruiter id's of the seeds.
2. Look to make sure the seeds' recruiter id is not present in the subject id.

Data Problems: Rounding error in IDS

Problem

Coupon or ID variables are numeric and some have 15 or more digits.

Symptom

Recruitment structure incorrectly read in except when using .rdsat and .rdsobj formats.

Solution

1. Edit variables to have non-numeric type (e.g. add a “c” to the beginning of all coupon variables).
2. Put data file in .rdsat format.