

Procedimientos de campo no convencionales.

La implementación en el caso de Respondent Driven Sampling (RDS)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

SIESTA - Abril 2019

Planteamiento del problema

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Naturaleza de las Poblaciones Ocultas

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

- ▶ No existe marco muestral
- ▶ Privacidad entre los miembros de la población (estigma social, cultural, comportamiento ilegal)

Ejemplos: consumidores de sustancias, inmigrantes, identidades de género/sexuales atípicas.

Técnicas de muestreo no probabilístico

- Muestreo por bola de nieve (Snow-ball)
- Muestreo basado en establecimientos (Facility-based)
- Muestreo basado en informantes calificados (Key informant)
- *Muestreo por lugar y tiempo (Time Location Sampling TLS)*

Inconvenientes:

- ▶ Conjunto de partida (semillas) no es aleatorio
- ▶ Probabilidades de inclusión desconocidas
- ▶ Sesgo por homofilia
- ▶ Sesgo por reclutamiento diferencial (voluntarismo, sobremuestreo de personas bien conectadas)

Planteamiento del problema
Historia del RDS
Descripción del RDS
Estimación en RDS
Aplicaciones

Técnicas de muestreo no probabilístico

- Muestreo por bola de nieve (Snow-ball)
- Muestreo basado en establecimientos (Facility-based)
- Muestreo basado en informantes calificados (Key informant)
- *Muestreo por lugar y tiempo (Time Location Sampling TLS)*

Inconvenientes:

- ▶ Conjunto de partida (semillas) no es aleatorio
- ▶ Probabilidades de inclusión desconocidas
- ▶ Sesgo por homofilia
- ▶ Sesgo por reclutamiento diferencial (voluntarismo, sobremuestreo de personas bien conectadas)

Planteamiento del problema
Historia del RDS
Descripción del RDS
Estimación en RDS
Aplicaciones

Dilema-Solución

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Dilema

- ▶ Muestra estadísticamente válida de una parte de la población, o
- ▶ Muestra estadísticamente no válida de cobertura más amplia?

Dilema-Solución

Solución

RDS Es capaz de generar una muestra representativa de la población objetivo y basado en la teoría de cadenas de Markov y en la teoría de redes logra:

- ▶ una muestra final independiente de las características de la inicial (semillas)
- ▶ el cálculo de probabilidades de inclusión a través de la construcción de la matriz de reclutamiento
- ▶ reducir el sesgo producido por el encubrimiento de pares y por los distintos tamaños de las redes sociales de los participantes.

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Historia del RDS

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Historia RDS



Douglas Heckathorn.
Dpto de Sociología.
Universidad de Cornell.

- 1997: primera aproximación como parte de un proyecto de investigación para la prevención de HIV
- 2002: Salganik desarrollo con más detalle, incluye IC y ponderadores para controlar los diferentes tamaños de red
- 2004: publican un artículo en el que muestran analíticamente y mediante simulaciones q los estimadores son asintóticamente insesgados (teoría de Markov y teoría de redes sociales)
- 2008: Volz publican en la Journal of Official Statistics un artículo que introduce el estimador RDS II (agregan la teoría de muestreo con probabilidades desiguales, el estimador Hansen Hurwitz)
- 2012: RDSAT (RDS Analysis Tool)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Descripción del RDS

Planteamiento del problema

[Historia del RDS](#)

Descripción del RDS

Estimación en RDS

Aplicaciones

Recolección de la muestra

- ▶ Se selecciona un grupo inicial (“Grupo semilla”) a través de un contacto preexistente con la población objetivo.
- ▶ A cada semilla se le provee de una cantidad c de cupones únicos que deben entregar a otros miembros de la población objetivo.
- ▶ A los reclutados por las semillas que participan en el estudio también se les proporciona una cantidad c de cupones únicos, y el proceso continúa hasta que se alcance el nivel de muestra que se requiere.
- ▶ Cada grupo de reclutados constituyen lo que se denominan *olas de reclutamiento*.

Al finalizar se obtiene una muestra compuesta por el grupo semilla y el conjunto de todos los reclutados.

Planteamiento del problema
Historia del RDS
Descripción del RDS
Estimación en RDS
Aplicaciones

Sistema dual de incentivos

Cada participante recibirá un pago por la participación en el estudio (**Incentivo Primario**) y otro por las personas reclutadas que efectivamente participen en la encuesta (**Incentivo Secundario**)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Requisitos de la población objeto de estudio

- 1 Los informantes deben reconocerse los unos a los otros como miembros de la población objetivo (deben saber a quien seleccionar como nuevo informante)
- 2 Las redes sociales deben ser lo suficientemente densas (para garantizar cierta profundidad sociométrica)
- 3 La población no debe estar muy segmentada en grupos (para no encapsular los subgrupos)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Fase Formativa

Utilización de técnicas rápidas de evaluación para explorar factores claves en la implementación del RDS

Resultados de la fase formativa

- ▶ Diversidad y los tamaños de la comunidad y las redes
- ▶ Accesibilidad a la población y posibles barreras operacionales
- ▶ Voluntad para participar
- ▶ Identificación de los tipos y las potenciales semillas
- ▶ Incentivos apropiados
- ▶ Determinación de las horas óptimas de operación
- ▶ Diseminación y conciencia del estudio
- ▶ Conocer necesidades de la población

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Selección de Semillas

Características de las semillas

- ▶ Diversidad con respecto a aspectos demográficos de la población (edad, educación, ubicación geográfica)
- ▶ Deben estar bien conectadas socialmente: buena actividad social, muchos amigos en la comunidad
- ▶ Articuladas: capaces de explicar a otros la propuesta de estudio y qué se espera de los demás
- ▶ Motivadas: que estén de acuerdo con el proyecto y con voluntad de promoverlo
- ▶ Claramente las semillas deberán pasar todos los criterios de elegibilidad

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Entrevistas

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Sistema y “sitio” de agendas de entrevistas

- ▶ Lugar fijo y/o coordinación previa del lugar y horario
- ▶ Cantidad de entrevistas que se pueden realizar por día

Incentivos

Los participantes serán recompensados por su participación en el estudio y por los reclutamientos exitosos de sus pares

- ▶ Monto (ni muy alto ni muy bajo)
- ▶ Momento y lugar de pago (efectivo, ticket, cupones para...)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cupones

- ▶ Los cupones son un elemento crucial en el estudio de RDS ya que linkea a cada uno de los reclutados con su reclutador.
- ▶ Deben estar bien diseñados (fondo artístico q haga referencia a la población de interés), con información clara y deben ser durables.
- ▶ Se debe diseñar un **sistema para codificarlos**

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Diseño de cupones

Los cupones deben tener la siguiente información

- ▶ El número de cupón (o lugar donde escribirlo)
- ▶ Nombre del proyecto
- ▶ Número de teléfono para agendar entrevista
- ▶ Horario de atención
- ▶ Dirección y mapa del lugar donde se realizarán las entrevistas
- ▶ Espacio para escribir el tiempo de vigencia del cupón (se recomienda 4 semanas, pero deberá ajustarse durante el reclutamiento)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cupón estudio Población Trans 2011-2012



Planteamiento del problema

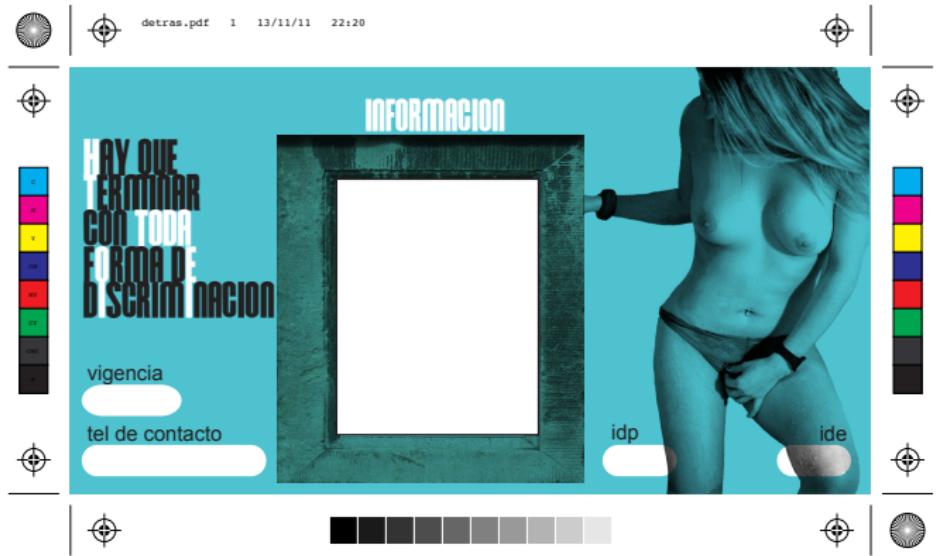
Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cupón estudio Población Trans 2011-2012



Planteamiento del problema

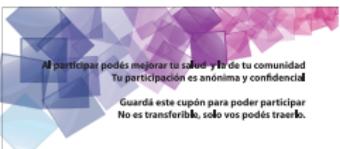
Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cupón estudio Población Consumidores de PBC



Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cupón estudio Inmigrantes



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

Para:

Nº de cupón:

Sumate a la Encuesta de Inmigración

¡Tu participación es muy valiosa!

Más info en <http://goo.gl/EZrrq1>
o llamando al 2413 6400 int 411

Planteamiento del
problema

Historia del RDS

Descripción del
RDS

Estimación en RDS

Aplicaciones

Manejo de Cupones

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Portapapeles	Fuente	Alineación	Número	Estilos	Celdas	Modificar	
P14	Vino otra persona con el cupon y los tres de richard a cobrar y trajo dos mas. Puede ser que venga el verdadero richard.						
A	B	C	ENTREGADOS			P	
			Cupón 1	Cupón 2	Cupón 3		
1	Fecha	Nro. Cupón	Control entrega	Cupón 1	Cupón 2	Cupón 3	Observaciones
2	14/may	101	0	1001	1002	1003	OK
3	15/may	102	0	1004	1005	1006	OK
4	17/may	103	0	1057	1054	1055	OK
5	16/may	104	0	1017	1018	1019	OK
6	17/may	105	0	1026	1049	1048	OK
7		1001	1				OK
8		1002	1	1050	1051	1052	OK
9		1003	1	1039	1040	1041	OK
10		1004	1	1010	1011	1012	OK
11		1005	1	1007	1008	1009	OK
12		1006	1	1013	1014	1015	OK
13		1007	1	1030	1031	1032	OK
14							Vino otra persona con el cupon

Manejo de Cupones - Pagos

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Nro. Cupón	VOLVIERON			PAGAR			Incentivo primario	Fecha Entrega			Pagar	Resta Pagar
	Cupón 1	Cupón 2	Cupón 3	Cupón 1	Cupón 2	Cupón 3		Cupón 1	Cupón 2	Cupón 3		
66	1558	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
67	1559	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
68	1560	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
69	1561	No vino	Sin entregar	Sin entregar	400	0	0	0	11/jul		0	600
70	1562	No vino	Sin entregar	Sin entregar	400	0	0	0	11/jul		0	600
71	1563	No vino	Sin entregar	Sin entregar	400	0	0	0	11/jul		0	600
72	1564	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
73	1565	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
74	1566	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
75	1567	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
76	1568	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000
77	1569	Sin entregar	Sin entregar	Sin entregar	400	0	0	0			0	1000

Entrenamiento de encuestadores

Determinación de elegibilidad

- ▶ Cupón
- ▶ Pertenercer a la población objetivo
- ▶ Especial atención en la posible duplicación

Al final de la entrevista

- ▶ Explicación clara sobre como aproximarse a sus contactos (conocerlo, elegible, no haber participado, presentar el cupón)
- ▶ Registro de la actividad de los cupones (recibidos, entregados, canjeados, cancelados)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cuestionario

El cuestionario debe incluir:

- ▶ Las variables necesarias para monitorear el proceso de reclutamiento y la verificación del equilibrio de la muestra.
- ▶ Requerimiento del RDS: **tamaño de la red**
- ▶ Preguntas que permitan estimar el tamaño de la población (Captura-Recaptura)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cómo preguntar el tamaño de la red

- ▶ En total, ¿cuántos consumidores de pasta base conoces personalmente? (que sean amigos, pareja, hermanos, primos, vecinos, conocidos, etc.)
- ▶ Entre los consumidores de pasta base que conoces personalmente (que sean amigos, pareja, hermanos, primos, vecinos, conocidos, etc.), ¿cuántos tienen entre 18 y 65 años?
- ▶ Entre los consumidores de pasta base que conoces personalmente (que sean amigos, pareja, hermanos, primos, vecinos, conocidos, etc.), y tienen entre 18 y 65 años,
 - ▶ ¿cuántos viven en Montevideo o el área metropolitana?
 - ▶ ¿cuántos son varones y cuántas son mujeres?
 - ▶ ¿a cuántos los conoces por su nombre o su apodo?
 - ▶ ¿cuántos confirmarían que te conocen?

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Cómo preguntar el tamaño de la red

- ▶ Entre los consumidores de pasta base que conoces personalmente (que sean amigos, pareja, hermanos, primos, vecinos, conocidos, etc.), tienen entre 18 y 65 años, viven en Montevideo o el área metropolitana y confirmarían que te conocen,
 - ▶ ¿con cuántos podrías comunicarte en las próximas 24 horas? (puede ser personalmente, por teléfono o por correo electrónico.)
 - ▶ ¿con cuántos te has comunicado durante los últimos 6 meses? (puede ser en vivo, por teléfono, o por correo electrónico)
 - ▶ ¿con cuántos te has comunicado durante los últimos 10 días? (puede ser en vivo, por teléfono, o por correo electrónico)
 - ▶ ¿cuántos estarían dispuestos a participar en este estudio? (si fueron invitados por ti)

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Estimación en RDS

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Supuestos

- Grado.
- El reclutamiento es al azar.
- Reciprocidad $A \rightleftharpoons B$ ($T_{AB} = T_{BA}$).
- Convergencia: reclutamiento modelado como un Proceso de Markov irreducible, cada estado tiene un tiempo de retorno finito y existe un único punto de equilibrio (red social bien conectada).

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Deducción de los estimadores de prevalencia

$$R_A = \sum_{i \in A} d_i = N_A \cdot D_A$$

$$C_{AB} = \frac{T_{AB}}{R_A}$$

$$\Rightarrow \begin{cases} T_{AB} = N_A \cdot D_A \cdot C_{AB} \\ T_{BA} = N_B \cdot D_B \cdot C_{BA} \end{cases}$$

Utilizando la hipótesis de Reciprocidad:

$$N_A \cdot D_A \cdot C_{AB} = N_B \cdot D_B \cdot C_{BA}$$

A, B : conjuntos disjuntos de individuos

R_A : n° total lazos o flechas del grupo A

d_i : grado del individuo i

N_A : n° de elementos en la población del grupo A

D_A : grado promedio de los miembros del grupo

C_{AB} : probabilidad de que alguien del grupo A elija a alguien del grupo B

T_{AB} : n° de lazos de amistad (o flechas) del grupo A al grupo B

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Para obtener P_A y P_B se divide ambos lados entre N

$$P_A \cdot D_A \cdot C_{AB} = P_B \cdot D_B \cdot C_{BA}$$

Y al incluir la restricción $\sum P_g = 1$

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

$$P_A = \frac{D_B \cdot C_{BA}}{D_A \cdot C_{AB} + D_B \cdot C_{BA}}$$

$$P_B = \frac{D_A \cdot C_{AB}}{D_A \cdot C_{AB} + D_B \cdot C_{BA}}$$

Estimación de la probabilidad de que alguien del grupo A elija a alguien del grupo B : \hat{C}_{AB} y \hat{C}_{BA}

Surge de la Matriz de Reclutamiento

	A	B
A	r_{AA}	r_{AB}
B	r_{BA}	r_{BB}

r_{AB} : elementos de la matriz de reclutamiento. N° de reclutados (en la muestra) del grupo B por el grupo A

$$\hat{C}_{AB} = \frac{r_{AB}}{r_{AA} + r_{AB}}$$

$$\hat{C}_{BA} = \frac{r_{BA}}{r_{BB} + r_{BA}}$$

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Estimación del grado promedio de los miembros del grupo: \hat{D}_A y \hat{D}_B

A este estimador se llega con dos enfoques diferentes:

- ▶ basado en las distribuciones del grado
- ▶ basado en los estimadores Hansen Hurwitz

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

$$\hat{D}_A = \frac{n_A}{\sum_{i=1}^{n_A} \frac{1}{d_i}}$$

$$\hat{D}_B = \frac{n_B}{\sum_{i=1}^{n_B} \frac{1}{d_i}}$$

n_A : nº de unidades en la muestra del grupo A

Estimadores de Prevalencia

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

$$\hat{P}_A = \frac{\hat{D}_B \cdot \hat{C}_{BA}}{\hat{D}_A \cdot \hat{C}_{AB} + \hat{D}_B \cdot \hat{C}_{BA}}$$

$$\hat{P}_B = \frac{\hat{D}_A \cdot \hat{C}_{AB}}{\hat{D}_A \cdot \hat{C}_{AB} + \hat{D}_B \cdot \hat{C}_{BA}}$$

Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Aplicaciones

Aplicaciones

Planteamiento del problema

Historia del RDS

Descripción del RDS

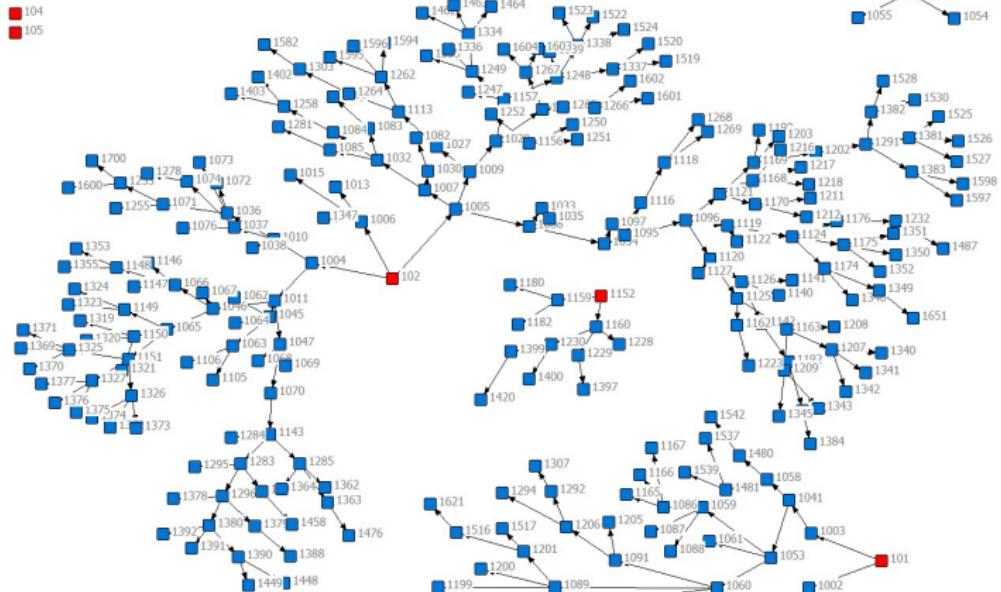
Estimación en RDS

Aplicaciones

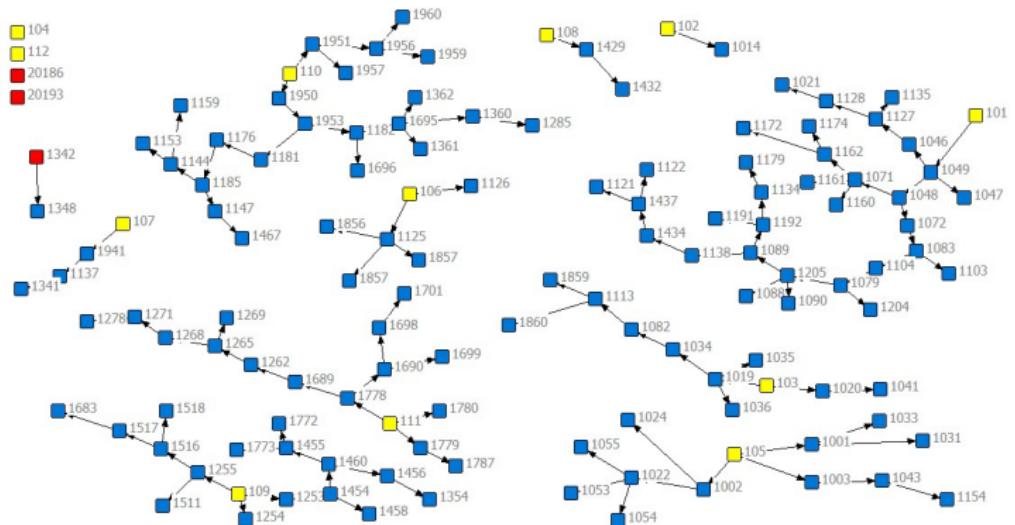
→ Estudio sobre los efectos, práctica y representaciones vinculadas con el consumo problemático de pasta base de cocaína (PBC) en usuarios problemáticos de esta sustancia. Montevideo 2018.

→ Etnoencuesta de Inmigración reciente en Montevideo, 2018

Usuarios problemáticos de PBC



Inmigrantes: Cuba



Planteamiento del problema

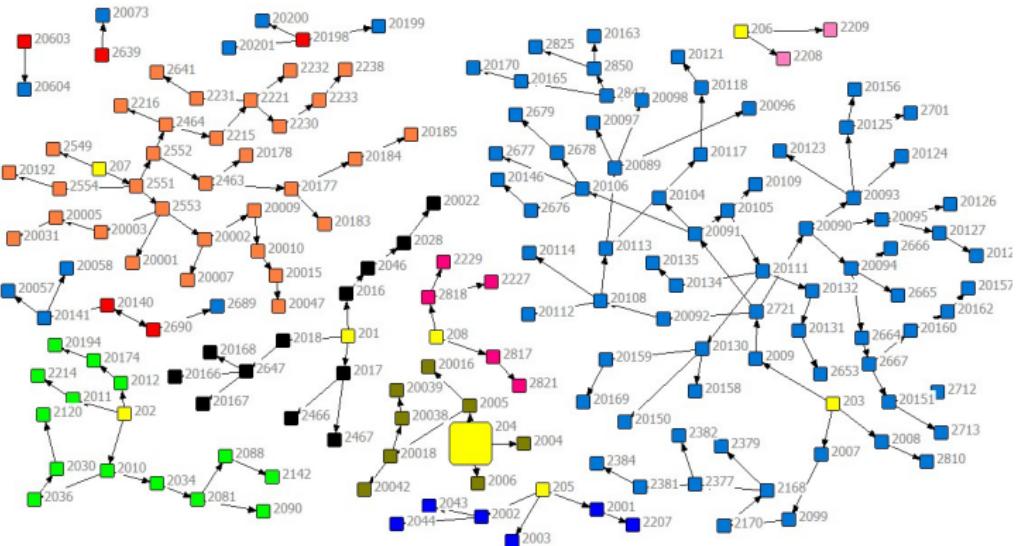
Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Inmigrantes: República Dominicana



Planteamiento del problema

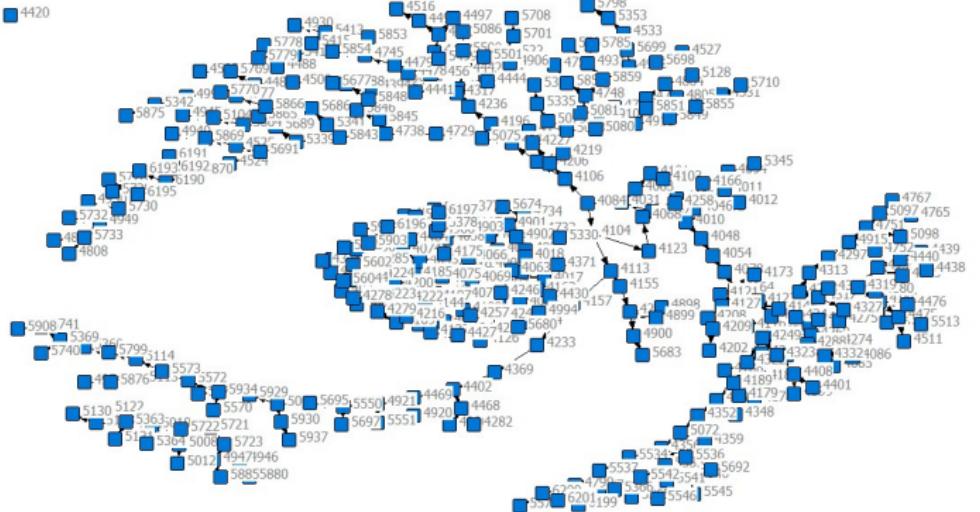
Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones

Inmigrantes: Venezuela



Planteamiento del problema

Historia del RDS

Descripción del RDS

Estimación en RDS

Aplicaciones