

Taller 1:

Diferenciar métodos para recolectar información de redes

Apellidos	Nombre	Código	Login

Esta actividad va a estar basada en el estudio realizado por Mendoza-Palechor y de la Hoz-Manotas (2019). La base de datos suministrada recopila información de 2.111 niños provenientes de Colombia, Perú y México. Para estos niños se realizó una estimación de los niveles de obesidad basado en hábitos de comida y sus condiciones de actividad física.

En este ejercicio vamos a comparar la diferencia entre dos metodologías de muestreo. La primera es el muestreo aleatorio y la segunda es el muestreo por bola de nieve en redes. El objetivo de esta actividad es entender la importancia del muestreo en redes para capturar las relaciones y la estructura de la red y la diferencia existente con las metodologías de muestreo estadístico que lo que buscan es representatividad muestral.

Suponga que un grupo de la OMS le ha encargado analizar desde la perspectiva de las redes sociales la muestra del estudio de Mendoza-Palechor y de la Hoz-Manotas. Este grupo quiere que usted realice un estudio que permita determinar cuáles son los niños más influenciadores que puedan ser utilizados para compartir un programa de actividad física con sus compañeros. El problema radica en que no hay suficiente presupuesto para preguntarle a cada uno de los niños cuáles son sus amigos. Por lo tanto, le han pedido que cree una muestra a partir de la población de niños y con base en esa muestra realice el análisis de redes para determinar a los niños más influenciadores. Para esto, usted se ha propuesto seguir los siguientes pasos que le permiten cumplir el requerimiento.

1. Dada la restricción presupuestal, usted sabe que lo primero que hay que hacer es determinar el tamaño de la muestra. Un aspecto relevante que le han solicitado es que la muestra que usted seleccione mantenga las proporciones de sexo de la población total. Calcule el tamaño de la muestra utilizando como referencia un nivel de confianza del 5% y un margen de error del 5%.
2. El grupo de la OMS no está seguro sobre la relevancia de utilizar metodologías de muestreo en redes y cree que la mejor estrategia es realizar un muestreo estratificado de acuerdo con el sexo. Por lo cual le han solicitado que diseñe una muestra que cumpla las especificaciones del tamaño muestral. ¿Cuál es la edad promedio de los niños y las niñas en la muestra?
3. Dado que usted ya seleccionó su muestra, ahora va a reconstruir la red de estos niños. Para esto, usted le ha preguntado a cada uno de los niños en la muestra que le

mencione a lo sumo a 20 de sus amigos con los que pasa la mayoría del tiempo. Esa información ya fue recopilada y se encuentra en el archivo de arcos (por simplicidad usted tiene la información de los arcos de todos los niños). De acuerdo con el grado total, ¿quiénes son los 5 niños con mayor número de conexiones?

4. Usted recuerda que en el curso de Estructuración de Redes Sociales le enseñaron que el muestro estadístico tiene problemas a la hora de capturar las redes de los niños. Por lo tanto, usted le ha propuesto realizar de manera paralela al grupo de la OMS un muestreo por bola de nieve. El equipo aceptó su propuesta, sin embargo, le solicitaron que el número de niños que capturara el muestro no fuera superior al tamaño de la muestra determinado previamente. ¿Cuál es la edad promedio de los niños y las niñas en la nueva muestra?
5. Usted va a replicar el ejercicio de reconstruir la red de los niños con esta nueva muestra. Grafique la red. De acuerdo con el grado total, ¿quiénes son los 5 niños con mayor número de conexiones?
6. El grupo de la OMS está complacido con los resultados, sin embargo, le han hecho la siguiente pregunta: ¿cuáles de los niños más influyentes debería elegir: los del muestreo estadístico o los del muestreo en redes? Justifique su respuesta argumentada con base en algunas de las métricas de cada una de las redes.

Realice esta actividad enteramente en **Python**. Adicionalmente, en el notebook adjunto se describen una serie de pasos que deberán desarrollar en la herramienta Gephi. Una vez terminen el estudio deberán presentar sus resultados en un informe autocontenido mostrando su desarrollo y análisis (máximo 2 páginas). Concluida la actividad, compriman en un zip su documento y el Notebook de soporte, colocando su código en el archivo.