

INSTITUTO SUPERIOR DEL PROFESORADO DE SALTA Nro. 6005

PLAN PEDAGÓGICO

- Llenar la propuesta en forma concreta, clara y útil para los estudiantes, las actividades deben estar acompañadas por los links en PDF de la bibliografía solicitada además de video, blogs, sitio web o mail para la realización y entrega de actividades.
- Una vez completo se debe enviar por mail al Coordinador de Carrera a la cual pertenezca.

INSTITUTO SUPERIOR DEL PROFESORADO DE SALTA Nro. 6005

PLAN PEDAGÓGICO – N° 2

**CARRERA: TECNICATURA SUPERIOR EN ADMINISTRACIÓN CON
ORIENTACIÓN EN GESTIÓN DE PROYECTOS**

(DESDE EL 21 DE SEPTIEMBRE AL 03 OCTUBRE DEL 2020)

ASIGNATURA: ESTADÍSTICA

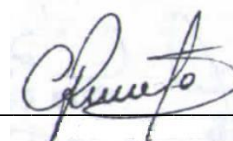
APELLIDO Y NOMBRE DEL DOCENTE: PINTO, CRISTIAN VICTOR

DIA: JUEVES Y VIERNES

HORARIO: 19:00 HASTA 20:20

CONTENIDO O TEMA A DESARROLLAR
Unidad II: técnicas de obtención de datos. Recomendación en la construcción de los instrumentos. Dificultades en su aplicación. Trabajo con fuentes: primarias y secundarias.
GUÍA O ACTIVIDADES
Lectura n. ° 4: Recolección de datos y fuentes de datos. Trabajo Práctico n. ° 2. Debe entregarse por medio de la opción “Realizar actividad” del Campus Virtual.
BIBLIOGRAFÍA
Está indicada al final de cada ficha de cátedra.
CANALES DE COMUNICACIÓN
Dudas y consultas a través de mensajería interna dentro del Campus Virtual (Educativa) de la Institución. Grupo de trabajo en Telegram: https://t.me/joinchat/UDvFVxxs7O2q2QHAHUAoAA Casilla de correo electrónico: iesestadistica@gmail.com

Se adjunta al presente material de estudio para el estudiante (de ser necesario).



FIRMA DEL DOCENTE

Lectura N° 4: Recolección de datos y fuentes de datos

Palabras iniciales

¿Cómo van llevando la materia? Recuerden no quedarse con dudas. Esta es la última lectura teórica, el resto del contenido es práctico. Ahora vamos a desarrollar un apartado que les será muy útil para cuando deban hacer captura de datos.

Continuamos con el proceso de investigación

Si seguimos la secuencia de pasos en la investigación estadística presentada en la lectura n. ° 2, luego del planteamiento del problema tenemos la recolección de datos. La recolección de datos tiene 2 etapas: la primera es diseñar un plan para recolectar los datos y la segunda es la recolección de datos.

Diseñar instrumentos de recolección de datos no es tarea sencilla. Si queremos obtener datos de calidad y que aporten a lo que queremos dilucidar es crucial. Es crucial porque si los datos que obtenemos son de buena calidad los resultados que obtendremos de todo el proceso estadístico serán de calidad (Torres et al., s. f.)

Los instrumentos de recolección de datos

Cualquier medio para recolectar datos se lo puede denominar instrumento de recolección de datos. Entre los más comunes están la observación, las encuestas, los cuestionarios, las cédulas censales y las entrevistas.

Vamos a tener en cuenta algunos pasos al momento de diseñar un instrumento:

1. Definir el objeto de la recolección de datos, formulando con precisión los objetivos a conseguir, desmenuzando el problema, eliminando lo superfluo y centrando el contenido de la encuesta.
2. Formulación adecuada del cuestionario/encuesta/entrevista.
3. Trabajo de campo (trabajo en terreno), consistente en la obtención de los datos. Para ello será preciso seleccionar a los entrevistadores, formarlos y distribuirles el trabajo a realizar de forma homogénea.
4. Procesar, codificar y tabular los resultados de la encuesta, que serán presentados en el informe y para posteriores análisis.

Características de un buen instrumento de recolección

Los instrumentos deben reunir las siguientes características:

1. Operativos: fáciles de manejar. Fáciles de procesar y tabular.
2. Fidedignos: que sea confiable, que permita la recolección real de los objetivos y que sean fáciles de ser depurados.
3. Válidos: que sea conciso, claro, firme, consistente, que no se preste a ambigüedades. Preguntas claras, breves, concretas y lógicas.

Recomendaciones fundamentales

1. Las preguntas han de ser pocas (no más de 30) y cortas.
2. Las preguntas deben tener una secuencia lógica y ordenada en función del tema.
3. Las preguntas preferentemente cerradas y numéricas.
4. Redactar las preguntas con lenguaje sencillo.
5. Formular las preguntas de forma concreta y precisa.
6. Evitar usar palabras abstractas y ambiguas.
7. Las preguntas formularlas de manera neutral.
8. En las preguntas abiertas no dar ninguna alternativa.
9. Que no levanten prejuicios en los encuestados.
10. Redactar las preguntas limitadas a una sola idea o referencia.
11. Evitar preguntas condicionantes con palabras con una carga emocional grande.
12. Evitar estimular una respuesta condicionada. Es el caso se preguntas que presentan varias respuestas alternativas y una de ellas va unida a un objetivo Tan altruista que difícilmente puede uno negarse.

Tipo de preguntas

Vamos a usar 3 criterios de clasificación.

En función a la contestación del encuestado.	Abierta: aquellas en la que no se le presentan las respuestas al encuestado dejando a éste libertad para que conteste según su criterio. A veces, suele encontrarse un listado de opciones y una que dice "otra, ¿cuál?". Cerrada: aquellas que contienen la respuesta, pudiendo decidirse entre aquellas que tienen sólo dos posibles o de elección múltiple. Estos son los valores de la variable.
En función del contenido de la pregunta.	De identificación: para conocer al sujeto de estudio. Ejemplo: preguntas sobre la edad, sexo, profesión, nacionalidad, etc.

	<p>De hechos: referidas a acontecimientos concretos. Ejemplo: ¿tiene automóvil?</p> <p>De acción: referida a las actividades que realiza el encuestado. Ejemplo: ¿viaja en transporte urbano?</p> <p>De información: para conocer los conocimientos del encuestado. Ejemplo: ¿conoce las funciones de la Autoridad Metropolitana de Transporte?</p> <p>De intencionalidad: para conocer la intención del encuestado. Ejemplo: ¿piensa comprar un automóvil para el 2021?</p> <p>De opinión: para conocer las apreciaciones sobre determinada cuestión. Ejemplo: ¿considera mejor viajar en transporte público que manejar?</p>
<p>En función a la ubicación dentro del cuestionario. Una pregunta puede cumplir más de una función simultáneamente.</p>	<p>Filtro: son aquellas que se realizan previamente a otras para eliminar a los que no les afecte. Ejemplo: Pregunta 1: ¿usted fuma? Pregunta 2: ¿qué tipo? La pregunta 1 es filtro de la 2. Si la persona me responde que no fuma, no le preguntaré la 2.</p> <p>Trampa o de control: para descubrir la intención con que se responde. Ejemplo: Pregunta 1: ¿es graduado universitario? Pregunta 2: ¿de qué universidad? Pregunta 3: ¿en qué año se graduó? Las preguntas 2 y 3 son de control para verificar si la persona me responde que es graduada universitaria.</p> <p>En batería: conjunto de preguntas que están encadenadas y se complementan entre sí. Ejemplo: Pregunta 1: ¿cuántas horas estudia Estadística Laboral? Pregunta 2: ¿prefiere estudiar con música? Pregunta 3: ¿prefiere estudiar en grupo o individual?</p> <p>Embudo: se empieza de lo general, hasta llegar a lo esencial. Pregunta 1: ¿le gusta la pastelería? Pregunta 2: ¿qué estilo de pastelería prefiere?</p>

Algunos problemas que se pueden enfrentar durante la recolección de datos

Problemas causados por los encuestados:

- No dicen la verdad.
- No comprenden las preguntas.
- Dan distintas respuestas a distintos entrevistadores.
- Se olvidan hechos.
- No dicen lo que piensan para impresionar al encuestador.

Problemas causados por los entrevistadores:

- Lee mal pregunta.
- Registra mal la respuesta.
- Utiliza palabras ambiguas.
- Formula preguntas en un orden incorrecto.

Manejo de errores: datos perdidos y datos anómalos

Dado que es un tema bastante complejo, se presentan muy someramente algunas definiciones.

- Datos perdidos: podemos decir que son datos nulos. Ocurre cuando se administra una encuesta/cuestionario y el encuestado no responde o dejan el espacio en blanco. También se puede dar cuando se empieza a construir el lote de datos en base a las encuestas/cuestionarios. Quien tiene la responsabilidad de organizar el lote de datos, no procesa todos los datos disponibles.
- Datos anómalos: podemos decir también datos atípicos. Son datos que están fuera del conjunto de los valores, o son valores que no se corresponden con la variable.
- Datos vacíos: en realidad no es un dato propiamente dicho. Existe cuando luego de una pregunta filtro ya no es necesaria la respuesta de las próximas preguntas.

El tratamiento de errores depende de las decisiones que tome quien tiene la responsabilidad de construir el lote de datos.

Para el caso de datos perdidos, se puede reducir el tamaño del lote y trabajar sólo en base a los datos disponible. Se puede volver a recolectar datos hasta llegar al tamaño deseado. Otra decisión es reponer los datos faltantes mediante alguna técnica estadística, para lo cual se requiere de mayor formación en el área de muestreo.

Para el caso de datos anómalos, a veces, se puede deducir y reponer los datos. Por ejemplo, consulta por el “año” de nacimiento y le responden con la “edad”, por lo cual se puede construir el dato. Cuando es imposible su reposición, se puede descartar esa cadena de datos. Cuando el dato pertenece a un conjunto de valores que puede estar contenido dentro de otro conjunto de valores más grandes, se puede redefinir la variable objeto de estudio para que comprenda dichos datos.

Más vocabulario

Ahora vamos a definir 3 conceptos.

Cadena de datos: secuencia de datos obtenidos de un elemento.

Lote de datos: conjunto de datos obtenidos en un proceso de recolección. Se suele agrupar por variable.

Matriz de datos: conjunto de datos obtenidos que están ordenados en un arreglo de rectangular. Se indica en cada columna el nombre de la variable.

Simbología por usar

Llaves y, punto y coma: cuando hablo de conjuntos de valores y cadenas de datos. { ; ; }

Paréntesis y, punto y coma: cuando hablo de datos y lote de datos. (; ;)

Problema muy frecuente en los estudiantes

Antes de avanzar quiero retomar el tema de variable. Quiero mostrar un error muy frecuente detectado en los trabajos prácticos y en los exámenes finales. Cuando estudio una variable es muy difícil captarla de manera directa, pero cuando es una cuestión subjetiva. Una variable se constituye por otras variables, y está muy vinculado al instrumento de recolección de datos.

Si quiero medir la satisfacción por un servicio prestado y pregunto ¿está satisfecho con el producto?, por cómo está formulada la pregunta admitirá un {sí; no} y tal vez no aporte a un proceso de mejora continua propuesto. Si pregunto con más detalle ¿le gustó la presentación?, ¿cómo le resultó el servicio de entrega? ¿recomendaría el producto? Etc. Todas estas preguntas captan una variable y todas apuntan medir el nivel de satisfacción del producto.

Vamos con otro ejemplo: en un estudio sobre inserción laboral de las mujeres se interesó conocer la variable “Acceso a puestos directivos”. No hay una pregunta que capte esa información de manera directa. Si hago la pregunta “¿Accedió a un puesto directivo?” estoy como en la situación anterior, {sí; no}. No hay un aporte rico para la investigación. Entonces, se formulan una serie de preguntas, donde cada pregunta capta una variable, y entre todas las variables se tiene una idea de “Acceso a puestos directivos”.

- ¿La organización tiene mayor confianza en personas de género masculino para puestos de liderazgo?
- ¿Considera que la mujer no es tenida en cuenta para puestos directivos por pertenecer al género femenino?
- ¿La mujer se considera más incompetente que el hombre para ocupar un cargo directivo?
- ¿Considera que la mujer tiene mayores dificultades en alcanzar puestos directivos?

En el ejemplo vimos como cada pregunta capta una idea sobre el “Acceso a puestos directivos”.

Repito, cada pregunta capta una variable porque las respuestas serán diferentes según a la persona a la que se pregunte. Pero entre todas las respuestas se puede tener una idea acerca de la variable de interés.

Fuentes de datos

Las fuentes de datos son todos los hechos o documentos a los que acude el investigador y que le permite tener información. Las fuentes de datos las podemos clasificar de varias formas. Un criterio es acorde con la disponibilidad de los datos. Bajo ese criterio, tenemos fuentes de datos primarias y fuentes de datos secundarias. ¿En qué se diferencian ambas fuentes? Las fuentes de datos primarias son construidas por el propio investigador. Son datos de “prima mano” que se obtienen de manera directa a través de algún instrumento de recolección de datos (encuestas, entrevistas, cuestionarios, etc.). Las fuentes de datos secundarias no son construidas por el investigador. Son datos que están disponibles para consulta, ya sea en instituciones públicas como privadas u otras investigaciones previas.

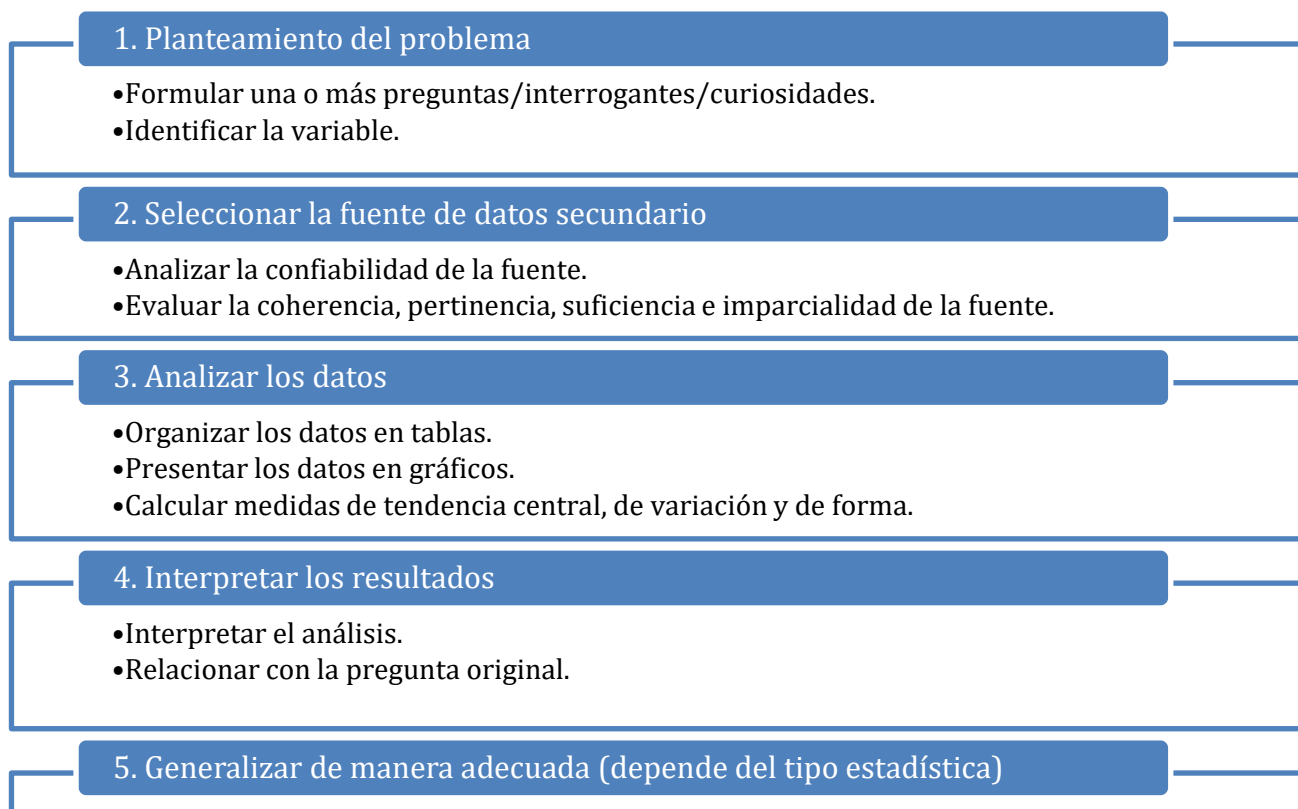
Trabajo con fuentes secundarias

¿Recuerdan la secuenciación básica para una investigación de la lectura n.º 2? Esa secuenciación es válida únicamente para el trabajo con fuentes de datos primarias. Cuando se decide hacer una investigación con fuentes secundarias los pasos van a diferir un poco.

Los datos ya están disponibles para ser analizados cuando trabajo con fuentes secundarias, por lo que no se debe pensar en hacer muestreos ni en los diseños de instrumentos de recolección de datos. Es lógico, porque los datos ya están disponibles para ser analizados.

Presentar atención porque este es un error común. Si ustedes definen que van a trabajar con fuentes de datos secundarias igualmente proponen o diseñan instrumentos de recolección. Esto es ilógico porque los datos ya los tienen. Lo importante en el trabajo con fuentes de datos secundarios es que la fuente sea confiable y cumpla con los requisitos de coherencia, pertinencia, suficiencia e imparcialidad. Vamos a suponer que las fuentes de datos públicas (INDEC [Instituto Nacional de

Estadísticas y Censos] y todos los organismos públicos] cumplen con todas las características y por lo tanto son confiables. Vamos a redefinir los pasos de la investigación cuando se trabaja con fuente de datos secundario (Juárez Duarte et al., 2012).



Una cuestión más que agregar que tiene que ver con los errores frecuentes que comenten en los exámenes finales. Las fuentes de datos pueden provenir de muestras y de censos. Por lo que es posible hacer inferencia cuando se trabaja con fuentes de datos secundarias. La única precaución que deben procurar tener es que esa muestra sea representativa, obvio además de los recaudos propios de la fuente de datos secundarios. Es muy común que concluyan que no es posible hacer inferencias cuando se trabaja con fuentes de datos secundarias.

Otro error común es que confunden el autor de una publicación de una noticia o artículo de investigación con el propietario de los datos. Por ejemplo, en la revista de la Facultad de Ciencias Económicas de la Universidad Buenos Aires, se publica que a partir de la información del censo nacional se tiene que Salta tiene una problemática de falta de oferta educativa de secundaria en las zonas rurales. El error es que asignen a la Facultad de Ciencias Económicas como fuente de datos, cuando en realidad es el INDEC. El INDEC es quien recolecta y almacena los datos de los censos de hogares, viviendas y personas. Recuerden que siempre deben estar atentos y pensar quién hizo la recolección de los datos.

Con esta lectura concluimos los aspectos teóricos y en la próxima empezamos con la práctica.

No se queden con dudas.

Bibliografía

Juárez Duarte, J. A., Ylé Martínez, A., Flórez Arco, A., & Inzunsa Cázares, S. (2012). *Estadística. Exploración de datos*. (3a Edición). Servicios Editoriales Once Ríos.

Torres, P. I. M., Paz, I. K., & Salazar, I. F. G. (s. f.). *Métodos de recolección de datos para una investigación*. 03, 21.

Trabajo Práctico N. ° 2

Estudiante	Escriba el nombre	Clase	Probabilidad y Estadística Estadística
Profesor	Pinto, Cristian Victor	Período	2020
Disponible desde	17/09/2020	Fecha de entrega	23/09/2020
Indique si presentó el TP N. ° 1 y el puntaje			Ejemplo: no presenté Ejemplo: sí presenté. 70 puntos.

Lea atentamente las consignas y responda luego de la lectura n. ° 2 y 3. Responda preferentemente a continuación de cada actividad.

Para cada situación completar el cuadro. La primera situación es a modo de ejemplo. Si hay más de una variable agregue las filas necesarias.

- Un estudio realizado en 2004 por el Ministerio de Educación, a 5000 estudiantes de los diferentes niveles educativos en el Reinado de España, concluyen que el género es una variable significativa en cuanto a la edad de inicio en el uso de videojuegos. Los varones se iniciaron, en general, 3 años antes que las mujeres.

Variable	Elemento	Criterio de variación		
		Temporal	Espacial	Elemento a elemento
Género	Estudiante			X
Edad	Estudiante			X

- Las ciencias físicas aseguran que la altura tiene incidencia en la temperatura en la que hierve el agua. Así, por cada 300 m. de altura, el punto de ebullición del agua se reduce en 1° C. Por tanto, en la cima del Mont Blanc (4500 m), el agua hierve a 84,4°C y en la del Everest a 70°C. ¿A cuántos grados en la ciudad de Salta?

Variable	Elemento	Criterio de variación		
		Temporal	Espacial	Elemento a elemento

- c. Un estudio sobre el consumo de chocolate en Chile concluye que la edad no es significativa. Por tanto, los niños y jóvenes consumen tanto que los adultos y ancianos.

Variable	Elemento	Criterio de variación		
		Temporal	Espacial	Elemento a elemento

En función de los casos del punto 1, justifiquen porque marcaron dicho criterio de variación. A modo de ejemplo, se continua con la primera situación.

- a. Elemento a elemento: porque cada estudiante tiene un género y una edad diferente. Los valores fueron cambiando en función de cada estudiante.
- b.
- c.

Para las siguientes situaciones indique si la/s variable/s están correctamente enunciadas. Para cada variable proponga el conjunto de valores, clasifíquela e indique el nivel de medición. Si los valores son infinitos utilice los puntos suspensivos. A modo de ejemplo tiene el primer caso.

- a. La canasta básica de alimentos se incrementó en 4%, de febrero 2019 a marzo 2019. La variable es mes.
- Es incorrecto pensar que la variable es mes. La variable es el precio y el elemento la canasta básica de alimentos.
 - Valores = {... 10000... 20000... 30000...}
 - Tipo: cuantitativa.
 - Nivel de medición: continua.
- b. En la revista *Reporter Consumer* [junio, 2004] mostraron que el contenido graso en los productos de *Starkbucks* son menores a los de *Donkin´ Donuts*.
La variable es tipo de local de café.
- c. En un estudio realizado en 4 colegios de élite en la ciudad de Cali, Colombia sobre la elección de carrera en estudiantes del último año de secundaria. Los hombres eligieron las ingenierías (33,3%), mientras que las mujeres prefirieron las ciencias de la salud (29,7%).
La variable es año que cursa en el secundario.

Complete el cuadro y concluya si se trata de un estudio por experimentación o por observación. A modo de ejemplo, tiene la primera situación.

- a. En el artículo “El impacto de la educación sobre la fecundidad adolescente: evidencia de la Ley Federal de Educación en Argentina” [Universidad Nacional de La Plata, 2015], analiza desde 1990 a 2011, las jóvenes y mujeres de 15 a 40 años, y entre otras cosas se concluye que la fecundidad se reduce por cada año de educación completada.

- b. En el artículo “Expectativas de los *millennians* en el contexto laboral” [Instituto Tecnológico de Celaya y Universidad de Guanajuato, mayo 2019] se expone que en base a 422 encuestas a jóvenes menores de 30 años de la ciudad de Guanajuato que las mujeres piden de sus empleadores mayor compromiso con la responsabilidad social. En general [varones y mujeres], se tiene una preferencia por ambientes de trabajo diversos e igualitarios.
- c. La Universidad de Oxford está en búsqueda de una vacuna contra uno de los tipos de coronavirus. La investigación de vacunas, en general constan de 3 fases, la exploratoria, la preclínica y la clínica. Dentro de la fase clínica, se organizan en 3 etapas: en la primera se prueba la vacuna en un grupo pequeño, en la segunda se prueba en cientos de personas y en la tercera se prueba en miles de personas. Para asegurar la calidad es necesario probar la vacuna en grandes cantidades de personas porque en pequeños grupos muchas veces no se detectan ciertos efectos secundarios. [Revisar la lectura n.º 1 sobre frecuencias escasas]

Caso	Cantidad de variables	Efecto tratamiento	Cantidad de mediciones	Estudio	
				Experimental	Observacional
a	2 [fecundidad y año de educación completada]	Ninguno	Por año debe realizar al menos 1 medición	No	Sí
b					
c					