

# Aplicada Proyecto Final

Rodrigo Zepeda

November 2021

El proyecto puede ser en equipos de 1, 2 ó 3. **Algunos proyectos pueden ser seleccionados al azar para responder dudas del mismo uno de los dos días después de su entrega.**

## 1 Introducción

El proyecto consiste en diseñar (para vender a sus clientes o sea yo) una encuesta. Deben decirme cómo la van a hacer, cuánto va a costar. Así mismo deben simular un ejemplo de cómo va a quedar la base de datos resultante de dicha encuesta y código de cómo deberá analizarse para que (si fuera la vida real) en cuanto tuvieran los datos ya estuviera listo el código que analiza. El proyecto consiste en las siguientes etapas. Supongan que tienen infinitos recursos.

1. **Establecer un diseño de encuesta** para el problema que se va a resolver. Esto involucra diseñar el cómo se van a medir las cosas usando un instrumento que ya exista para ese problema o un problema similar. Este instrumento no es que se les ocurra a ustedes sino que investiguen cómo lo hace la gente.
2. **El diseño debe ser multietápico y estratificado** Es decir una vez que decidieron qué van a medir y cómo lo van a medir un punto importante es que el diseño sea complejo (múltiples etapas tipo seleccionar primero la casa luego la persona) y a la vez tiene que estar estratificado por una otra variable (digamos estado). Es importante que se distinga cuál variable es una etapa cuál es un estado. Es importante señalar quiénes son los  $\pi_k$ , las unidades primarias y secundarias de muestreo y el estimador.
3. **Se debe obtener un tamaño de muestra** para una confianza del 90% y una precisión de  $\pm 100$  a menos que el enunciado específico diga lo contrario. El cálculo debe ser claro cómo se hizo (pueden hacer código pero necesito las mates).
4. **El diseño debe considerar valores faltantes** Es decir suponer que una proporción  $p$  de valores serán no registrados. Esta  $p$  es una constante ya sea obtenida de algún estudio previo o bien supongan  $p = 0.2$ . Debe incluirse  $p$  en el tamaño de muestra.

5. **El proyecto debe incluir un costo asociado a la encuesta** Se sugiere un costo lineal como el que vimos de estratificado. El costo debe tener lógica de de dónde sale. No se preocupen porque sea el más barato o que sea exacto sólo denme un estimado (ejemplo: googleen cuánto cuesta un encuestador en México por hora y eso agregan al costo + el costo de medir + el costo de analizar aunque no sea exacto)
6. **Simular** Una base de datos que me dé idea de cómo esperan que queden los resultados. Lo que estoy buscando es algo del estilo “la base va a tener 4 columnas y 50 renglones y bueno por ahora le pusimos estos numeritos a los renglones pero obvio si tomamos la muestra va a cambiar”. **Ojo** No olviden que una proporción  $p$  deben quedar como faltantes.
7. **Crear el código** Tener el código (en el lenguaje de su preferencia) que a partir de la base obtenga el estimador. Es decir tomen la base simulada como si fuera cierta y hagan el código que estimaría los valores. Yo sé que los valores que salgan no van a tener sentido pero en cuanto vayan y tomen la muestra lo único que harían sería tener otra base pero igualita (“4 columnas y 50 renglones” en el ejemplo anterior) y bastaría con correr el código para tener el resultado.
8. **Entrega** Día del final a cualquier hora yo empiezo la revisión al día siguiente a las 8am.

## 2 Elegir uno de los siguientes problemas a resolver y sólo uno

1. Estimar el **total** de perritos callejeros en Guadalajara y su zona conurbada.
2. Estimar el total de personas con **VIH** en México.
3. Determinar la cantidad de petróleo (en barriles) en el golfo de México.
4. Establecer qué porcentaje de trabajadores de tiendas de abarrotes en Toluca están contratados por honorarios (precisión de  $\pm 0.1$ ).
5. Determinar si la media de la diferencia de ingreso (en pesos) de hombres y mujeres en Nayarit es cero (y no hay diferencia) o es distinta de cero.
6. Establecer qué porcentaje de ardillas en Ciudad de México tiene rabia (precisión de  $\pm 0.2$ ).
7. Estimar el tiempo promedio (en minutos, precisión de  $\pm 15$  minutos) de traslado de una persona de su trabajo hacia su casa en Monterrey.
8. Determinar cuál es la proporción de personas asintomáticas de COVID-19 en el mundo (precisión de  $\pm 0.15$ ).

9. Estimar es el número promedio de asaltos por día en el transporte CDMX-EdoMex que no se denuncian (precisión de  $\pm 5$ ).

### 3 Entrega

Se debe entregar un archivo pdf (así como el código que puede ser en el pdf o aparte y el archivo simulado que es aparte) con la propuesta de proyecto. Debe tener objetivo, referencias de gente que haya hecho lo mismo y las matemáticas deben estar a computadora (no escaneos o fotos de lo que escribieron a mano). La idea es que sea parecido a los pdf metodológicos del INEGI.

### 4 Registro

Por favor regístrense en el forms <https://forms.gle/7C24dHawbkuTofMX9>. Sólo se pueden a lo más dos registros por tipo de proyecto. Pueden ver qué proyecto ya está registrado acá: [https://docs.google.com/spreadsheets/d/14iTl6jbnS7bhyqgjgNvjS\\_539byyo5da5nPeud8HQbM/edit?usp=sharing](https://docs.google.com/spreadsheets/d/14iTl6jbnS7bhyqgjgNvjS_539byyo5da5nPeud8HQbM/edit?usp=sharing)