

Probabilidad y estadística

Docente a Cargo:

Ing. Pablo Vega

Mail: pablorgega@gmail.com

Material extracto del apunte teórico del Ing. Fugiglando

- Materiales de estudio: diapositivas de clases y cualquier texto de Estadística básica.

Walpole, R. E., Myers, R. H., Myers, S. L. y Ye, K. (2012) Probabilidad y Estadística para Ingeniería y Ciencias. Novena edición. Pearson Educación, México.

- Software estadístico: MINITAB 18 o Excel
- Evaluación: examen vía Moddle

UNIDADES TEMATICAS

1. Metodología estadística
2. Medias de posición y dispersión
3. Algebra de probabilidades
4. Variable aleatoria
5. Modelos especiales de probabilidad
6. Elementos de muestreo
7. Teoría de la estimación estadística
8. Contraste, prueba o docimasia de hipótesis

UNIDADES TEMATICAS

1. Metodología estadística
2. Medias de posición y dispersión
3. Algebra de probabilidades
4. Variable aleatoria
5. Modelos especiales de probabilidad
6. Elementos de muestreo
7. Teoría de la estimación estadística
8. Contraste, prueba o docimasia de hipótesis



Cual es la diferencia entre la probabilidad y la estadística?

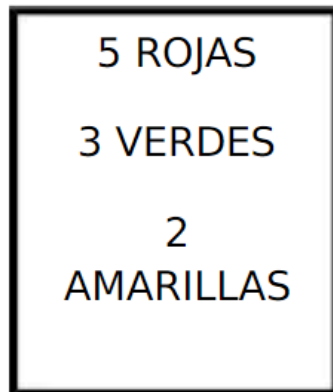
- La probabilidad propone modelos para los fenómenos aleatorios, es decir, los que se pueden predecir con certeza y estudia sus consecuencias lógicas.
- La Estadística ofrece métodos y técnicas que permiten entender los datos a partir de modelos.

De esta manera el calculo de las Probabilidades es una teoría matemática y la Estadística es una ciencia aplicada donde hay que dar un contenido concreto a la noción de probabilidad.

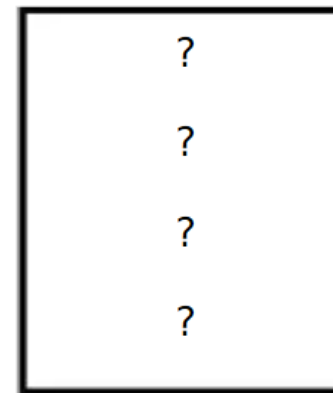
La probabilidad y la Estadística son dos campos distintos aunque relacionados entre si, en las matemáticas, se dice a veces que la probabilidad es el vehículo de la estadística es decir, de no ser por las Leyes de la Probabilidad, no existirían las leyes estadísticas.

En la caja de probabilidad hay 10 fichas (5 rojas, 3 verdes y 2 amarillas) en el terreno de la probabilidad se intenta responder preguntas "por ejemplo, cual es la probabilidad de que si se saca una ficha de la caja, sea roja? Cual es la probabilidad de ocurrencia de un evento? Por otra parte, dentro de la estadística necesitaremos extraer una muestra de ella para hacernos conjeturas sobre lo que en ella existe. Por lo que se podría decir como diferencia bien marcada entre una y otra, lo siguiente: "La Probabilidad estudia la oportunidad de que algo ocurra, cuando se conocen las posibilidades; mientras que la estadística pregunta, culés son esas posibilidades a partir de los resultados de una muestra"

PROBABILIDAD



ESTADISTICA



Qué se entiende por Estadística?

Podemos señalar al menos dos significados:

- La Estadística se refiere a información numérica. Ejemplos: el salario inicial promedio de los egresados de la carrera de Ingeniería electrónica, el porcentaje de estudiantes de la UTN FRC que trabajan en relación de dependencia, etc.
- La Estadística es la ciencia de recolectar, organizar, presentar, analizar e interpretar datos para ayudar en una toma de decisiones más efectiva.

¿Por qué estudiar Estadística?

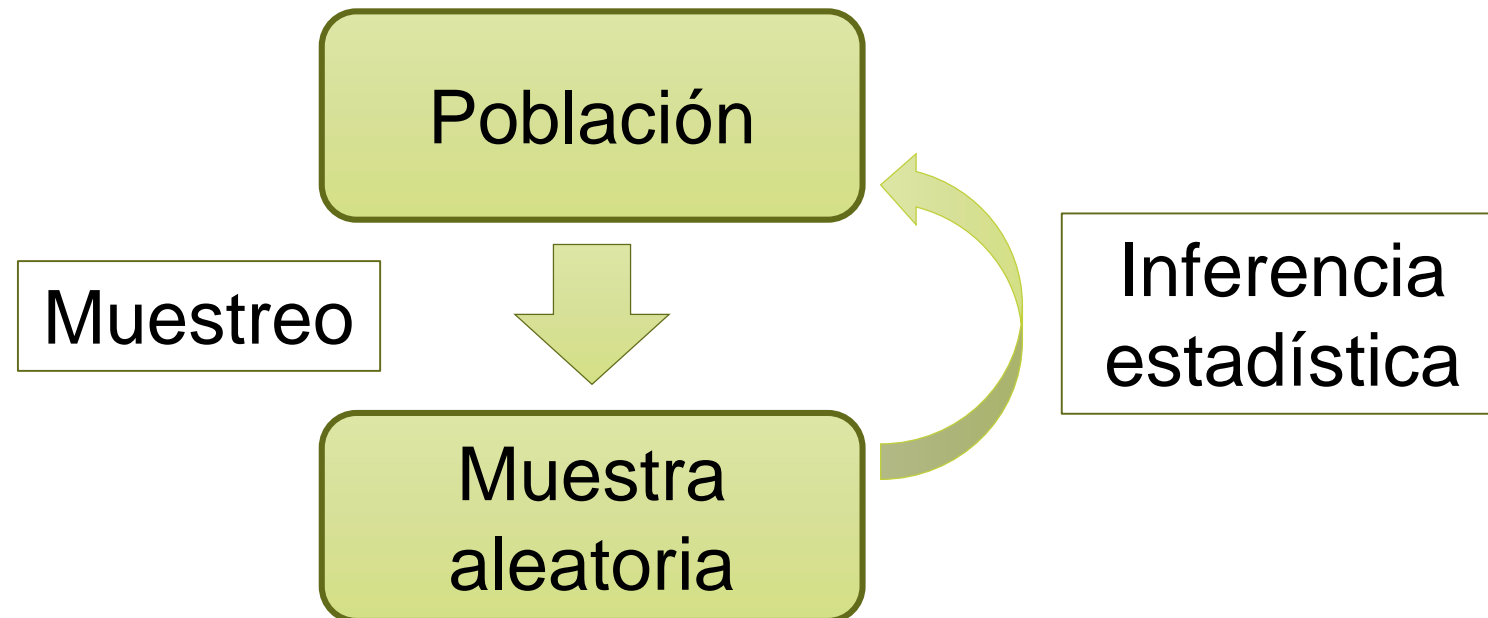
- En todos lados encontramos información numérica. (periódicos, revistas especializadas, informes internos de la empresa).
- Las técnicas estadísticas se utilizan para tomar decisiones que afectan nuestra vida diaria. (Ej. Las compañías de seguro utilizan análisis estadísticos para fijar las primas de los seguros).
- El conocimiento de las técnicas estadísticas ayudarán a entender por qué se toman ciertas decisiones con la ayuda del análisis de datos.

¿Cómo organizar un estudio?

Etapas: 1- definir el problema 2- planificar la recolección de datos 3- recolectar y analizar los datos 4- extraer conclusiones (toma de decisiones)

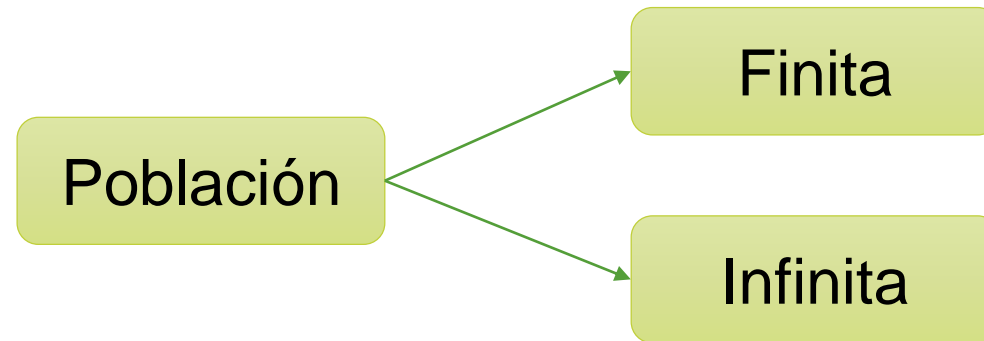
Se clasifica en:

- **Estadística Descriptiva:** se corresponde con la metodología utilizada para describir un conjunto de datos a través de métodos numéricos o gráficos. Por ejemplo, la evolución de las ventas en una empresa en un cierto periodo de tiempo
- **Inferencia Estadística:** incluye métodos que permiten hacer generalizaciones con respecto a la población en base a la información proporcionada por una muestra aleatoria, con un grado de incertidumbre cuantificable.



Precisando algunos conceptos

- **Población:** es un conjunto completo de elementos o individuos de interés sobre los cuales se quiere realizar un estudio. Por ejemplo: el total de empresas de una cierta rama de actividad o el agua contenida en un cierto dique.
















Precisando algunos conceptos

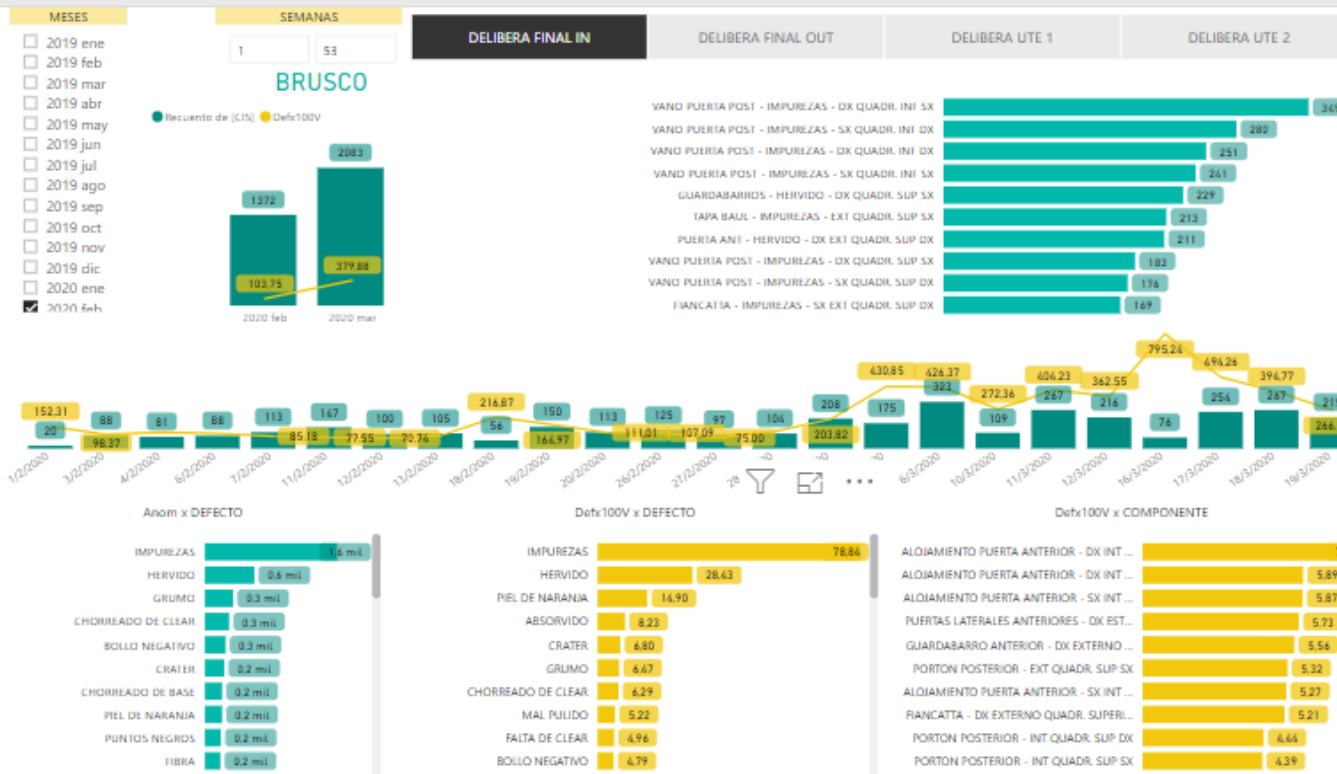
- **Datos:** es toda información susceptible de ser sometida al análisis estadístico, es decir, que debe poder ser comparada, analizada e interpretada, razón por la cual se dice que los datos estadísticos presentan relaciones significativas.
- Fuentes de datos:
 1. **Datos primarios:** son los que recopila el propio investigador para resolver un problema particular. Métodos más conocidos para recolectar datos primarios: investigación mediante encuesta, experimental y observacional.
 2. **Datos secundarios:** son aquellos que han sido reunidos y publicados por otras instituciones y se encuentran disponibles. En general provenientes de organismos públicos tales como INDEC, BCRA, AFIP.

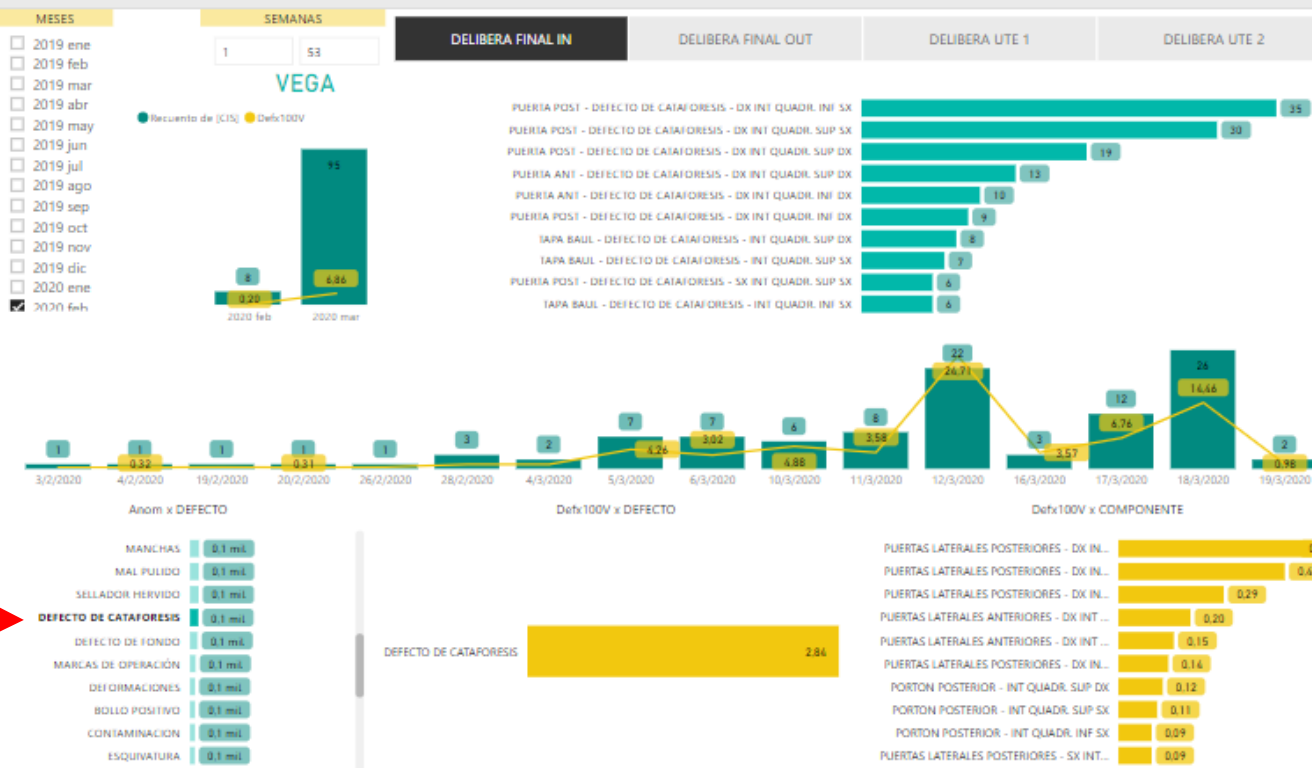
FAASA

Inicio > FAASA

CARPETAS (13)

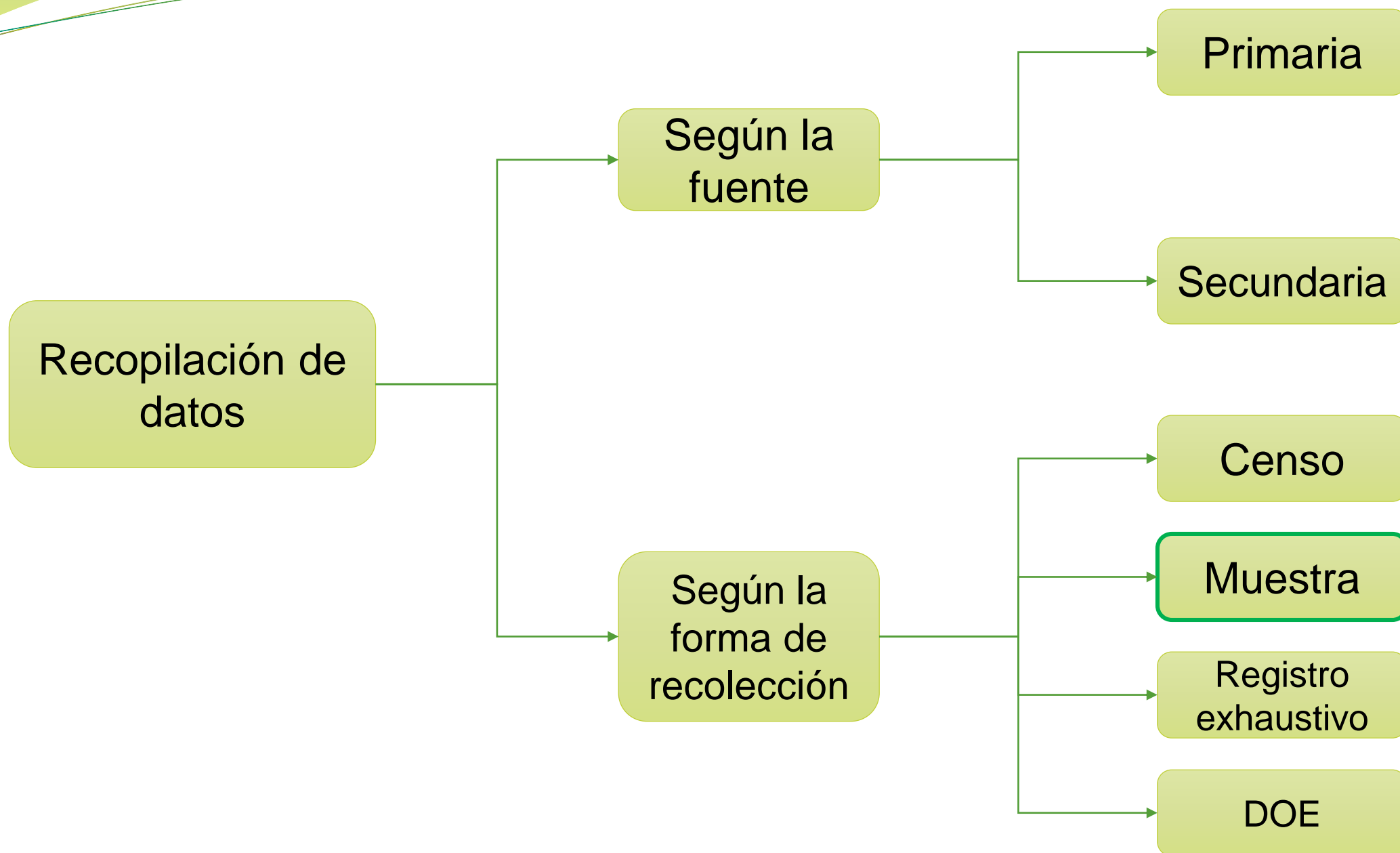
 CALIDAD	 COMMERCIAL	 FiatPlan	 FINANCE	 GARANTIA
 GOVERNANCE	 INDUSTRIAL	 LIVEON	 MARKETING	 PATENTAMIENTO
 PRODUCT_DEVELOPMENT	 PURCHASING	 RRHH		





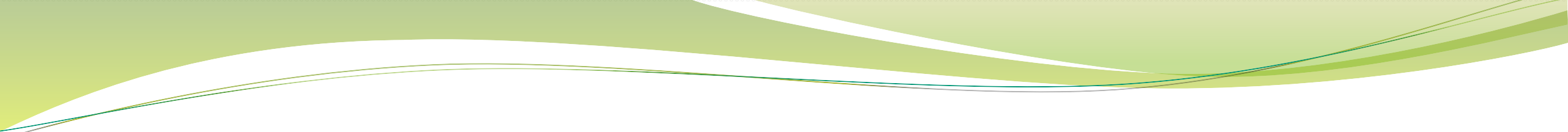
Precisando algunos conceptos

- Proceso de generación de datos puede llevarse a cabo de distintas formas:
1. **Censo:** es el proceso consistente en estudiar características de cada uno de *todos* los elementos de una cierta población estadística, debiendo ser esta *finita*, siendo un relevamiento *estadístico*.
 2. **Muestra:** consiste en estudiar características de una cierta población estadística tomando en consideración solo una *parte representativa* de la misma, a través de las denominadas *técnicas de muestreo*, de tal modo que las conclusiones obtenidas de esa parte puedan ser generalizadas al resto de los elementos de la población, a través de la metodología de la *inferencia estadística*. Resulta aplicable tanto a poblaciones infinitas o ensayos destructivos.
 3. **Registro exhaustivo:** la información estadística se obtiene por un acto administrativo. Por ejemplo cuando se quiere estudiar la procedencia de los alumnos de una Facultad, o el número de nacimientos inscriptos en un registro civil en un período de tiempo.
 4. **Diseño de experimentos:** es el proceso que se presenta cuando el investigador genera sus propios datos, diseñando un experimento a la medida de sus necesidades. Por ejemplo combinar determinados valores de los parámetros de cataforesis para lograr un determinado valor de espesor de la misma al menor costo posible.



Precisando algunos conceptos

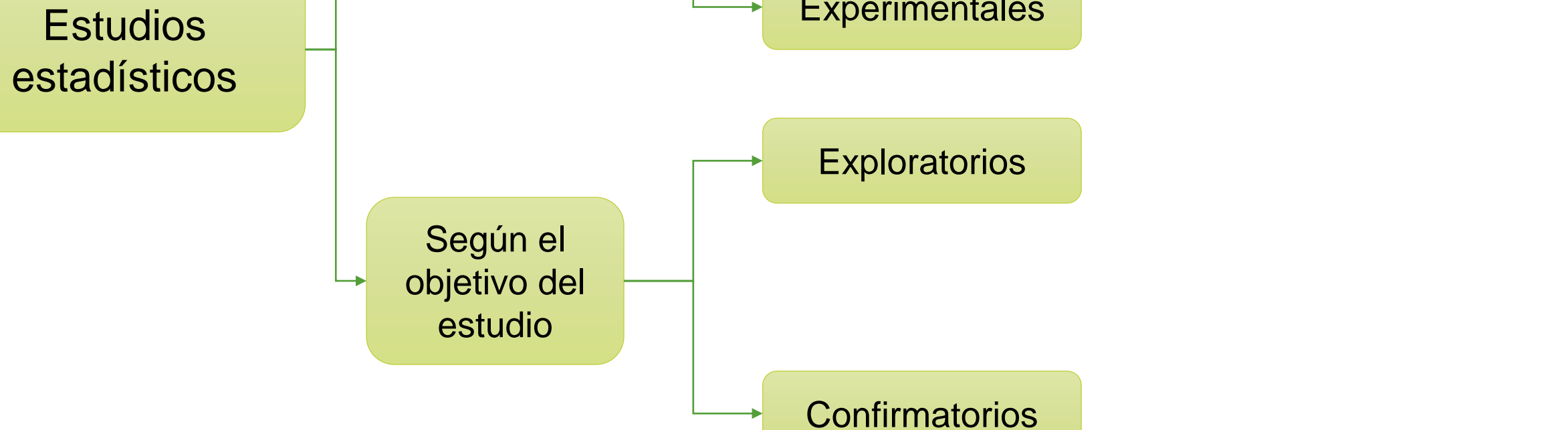
- **Unidad estadística o unidad elemental:** cada componente de una población estadística, por ejemplo, si se está estudiando el modelo de auto de una cierta marca más vendido, cada auto de esta marca será la unidad estadística.
- **Unidad de relevamiento:** es donde la información estadística va a ser recopilada, o por quien proporciona la misma. Por ejemplo, si queremos saber de la producción diaria de una empresa cuántos artículos fallados fueron producidos y existe en la misma un área de control de calidad, este sector sería la unidad de relevamiento. Por ejemplo, en un censo de población la unidad de relevamiento es el hogar.



Considere los siguientes casos e identifique: población, muestra y unidad estadística.

- Caso 1. La Unión Industrial de Córdoba, durante el año 2017, realizó un estudio para analizar los distintos aspectos de la actividad industrial. Para ello, de la población de 2000 industrias radicadas en la provincia, se tomó una muestra aleatoria de 150 empresas.
- Caso 2. Suponga que un ingeniero se encuentra con datos de un proceso de producción en el cual se muestrean 100 artículos y se obtienen 10 defectuosos.

Estudios estadísticos



```
graph LR; A[Estudios estadísticos] --> B[Según la intervención del investigador]; A --> C[Según el objetivo del estudio]; B --> D[Observacionales]; B --> E[Experimentales]; D --> F[Retrospectivos]; D --> G[Prospectivos]; C --> H[Exploratorios]; C --> I[Confirmatorios];
```

The diagram is a hierarchical flowchart classifying statistical studies. It begins with a root node 'Estudios estadísticos' on the left. Two arrows branch out to 'Según la intervención del investigador' and 'Según el objetivo del estudio'. From the first node, two arrows lead to 'Observacionales' and 'Experimentales'. From 'Observacionales', two arrows lead to 'Retrospectivos' and 'Prospectivos'. From the second node, two arrows lead to 'Exploratorios' and 'Confirmatorios'. All nodes are contained within light green rounded rectangular boxes, and the flow is indicated by green arrows.

Según la
intervención
del
investigador

Observacionales

Retrospectivos

Prospectivos

Experimentales

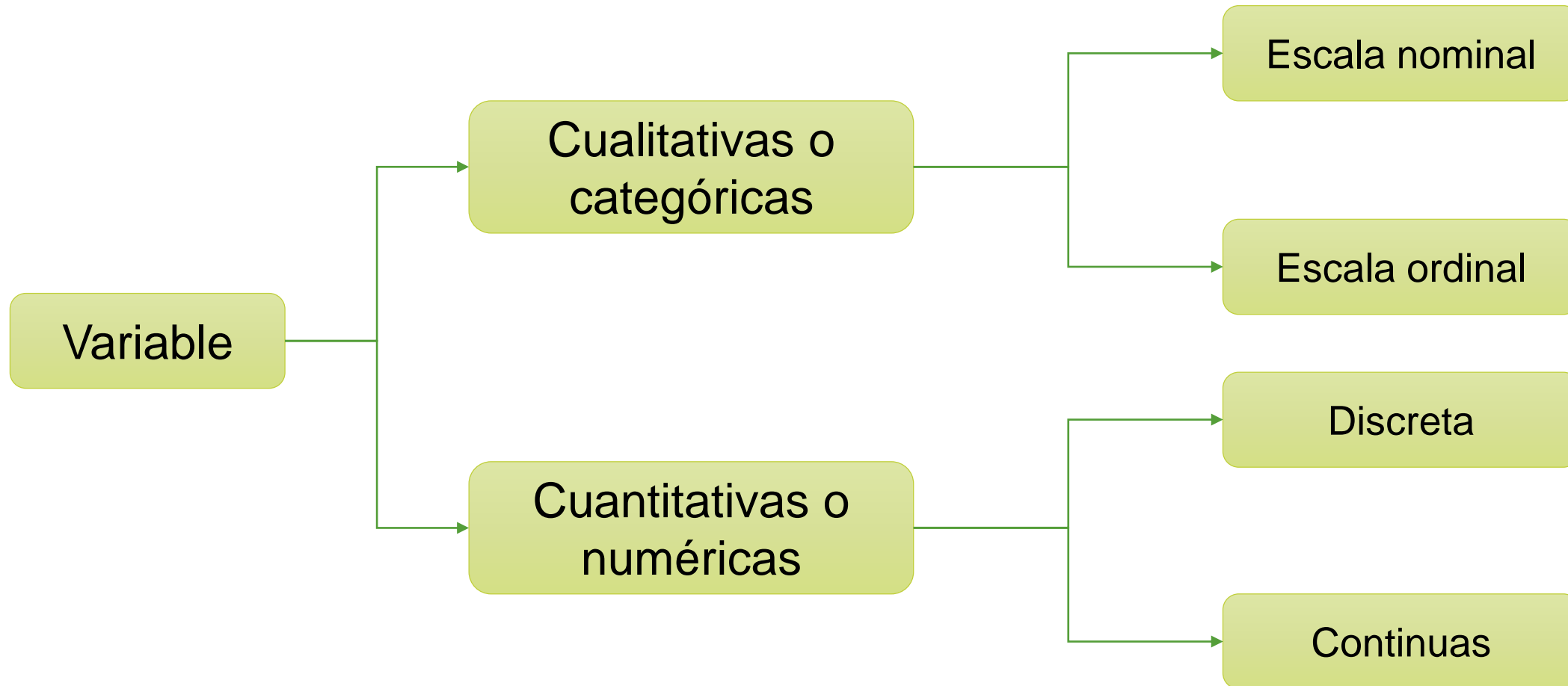
Según el
objetivo del
estudio

Exploratorios

Confirmatorios

Precisando algunos conceptos

- **Variable:** de las unidades estadísticas de una determinada población podemos estudiar una o mas características dadas por la o las *variables*, es decir es la característica de interés a estudiar.



Precisando algunos conceptos

- **Cualitativas:** aquellas en donde la característica bajo estudio es un *atributo o categoría*. Por ejemplo: la profesión de los empleados de una empresa. De acuerdo a la forma de medición las variables cualitativas pueden ser clasificadas en escalas dadas por:
 1. **Escala Nominal:** es aquella en la que no es posible establecer una relación de magnitud entre los distintos atributos en que puede ser expresada la variable. Por ejemplo: nacionalidad de las personas.
 2. **Escala Ordinal:** es el caso en donde es posible establecer un orden de importancia entre los distintos atributos en que puede ser expresada la variable. Por ejemplo: la opinión de las personas acerca del servicio brindado por una empresa, siendo las posibles respuestas: Excelente, Muy Bueno, Bueno, Regular, Malo. Se observa que no es posible decir que los que responden Bueno están doblemente satisfechos en relación a los que respondieron Malo.

Precisando algunos conceptos

- **Cuantitativas:** aquellas en donde la característica bajo estudio *numérica*. Según sea la forma de obtener el valor numérico, se las clasifica en:
 1. **Discretas:** cuando el valor surge de un *proceso de conteo*, asumiendo por lo tanto valores enteros. Por ejemplo: el numero de empleados de una empresa, cantidad de autos que llegan a un peaje en una cantidad de tiempo determinada.
 2. **Continuas:** cuando el valor surge de un *proceso de medición*, siendo la diferencia entre un valor y otro un infinitésimo, asumiendo por lo tanto cualquier valor real. Por ejemplo: la altura de los empleados de una empresa.

Precisando algunos conceptos

- **Cuantitativas:** aquellas en donde la característica bajo estudio *numérica*. Según sea la forma de obtener el valor numérico, se las clasifica en:
 1. **Discretas:** cuando el valor surge de un *proceso de conteo*, asumiendo por lo tanto valores enteros. Por ejemplo: el numero de empleados de una empresa, cantidad de autos que llegan a un peaje en una cantidad de tiempo determinada.
 2. **Continuas:** cuando el valor surge de un *proceso de medición*, siendo la diferencia entre un valor y otro un infinitésimo, asumiendo por lo tanto cualquier valor real. Por ejemplo: la altura de los empleados de una empresa.



Clasifique las siguientes variables

1. Número de empleados de una empresa
2. Peso de un cargamento de cereal
3. Estado civil de una persona
4. Vida útil de una lámpara LED
5. Cantidad de habitaciones de un hotel
6. Ventas mensuales de una empresa
7. Sexo de una persona
8. Cantidad de artículos defectuosos
9. Número de buques que llegan diariamente a un puerto
10. Distancia recorrida diariamente por un taxi

Precisando algunos conceptos

- **Parámetro:** es toda medida estadística calculada considerando a *todos* los elementos de una población bajo estudio. Por ejemplo calcular las ventas promedio de las empresas de una determinada rama de actividad, considerando las ventas de cada una de *todas* esas empresas.
- **Estadígrafo, estadístico o estimador:** a toda medida estadística calculada considerando una *muestra* de los elementos de una determinada población.

Esta medida servirá de **estimador** del verdadero valor del **parámetro** de esa población (inferencia estadística)

- **Actividad 1:** determine si las siguientes variables son cualitativas o cuantitativas, en este caso, clasificarlas

1. La vida útil de una bombita de luz.
2. La marca de una bombita de luz.
3. El numero de alumnos inscriptos en una Facultad.
4. La cantidad de lluvia caída en una localidad en un determinado periodo.
5. El numero de accidentes por semana en una fabrica.
6. Los tipos de accidentes que ocurren en una fabrica.
7. El numero de empleados que no asisten a su lugar de trabajo en una empresa.

- **Actividad 2:** la siguiente información se incluye en el registro personal de una empresa: determine si las siguientes variables son cualitativas o cuantitativas

1. Edad: 37 años
2. Sexo: masculino
3. Estado civil: casado
4. Estudios: universitarios completo
5. Altura: 1,76 m
6. Peso: 78 kg
7. Antigüedad en la empresa: 8 años
8. Salario mensual: U\$S 2500

- **Actividad 3:** la dirección de turismo de Marpla recopila datos sobre los turistas que llegan allá. Entre las 16 preguntas de un cuestionario que se repartió a los pasajeros en los vuelos de llegada durante una cierta temporada de verano, estaban las siguientes:

1. Cuantas veces vino a Marpla?
2. Cual es el motivo principal de este viaje?
3. Cual será su alojamiento en este viaje?
4. Cuantos días tiene previsto quedarse?
5. Cual es el numero de integrantes de su grupo familiar?
6. Cual es el sexo de los integrantes del mismo?

a) Cual es la población que se estudia? b) Cual es la muestra en estudio? c) Porque se necesita usar una muestra en este caso? d) Las respuestas a los interrogantes anteriores son cualitativas o cuantitativas? En este ultimo caso, clasificarlas? e) La información recopilada es representativa de todos los turistas que llegan a Marpla? porque?

- **Actividad 4:** se consulto a cada estudiante de los 32 de un curso acerca de su color favorito, obteniéndose que 12 preferían el rojo, 8 el verde, 8 el azul y 4 el amarillo.

1. Cual es la población estadística?
2. Cual es la unidad estadística?
3. Cual es la unidad de relevamiento?
4. Cual es la variable en estudio?
5. El estudio es censal o muestral?
6. Contribuir una distribución de frecuencias de variable discreta considerando la variable color.
7. Representar gráficamente las frecuencias absolutas y relativas, simples y acumuladas.

ORGANIZACION Y PRESENTACION DE DATOS ESTADISTICOS

- **Serie simple:** se corresponde con la metodología utilizada para describir un conjunto de datos a través de métodos numéricos o gráficos. Por ejemplo, la evolución de las ventas en una empresa en un cierto periodo de tiempo