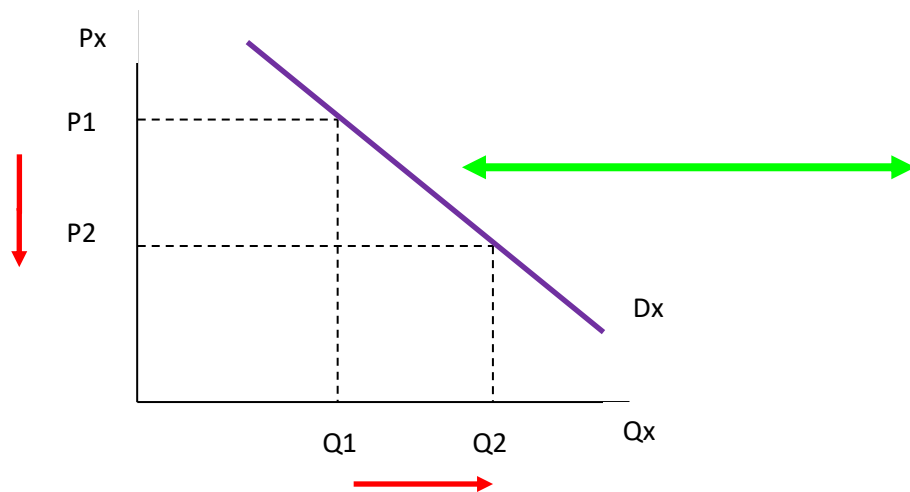


ESTRUCTURA DE LOS DATOS

Se llegó a la conclusión de que la econometría es *“una fusión de teoría económica, economía matemática, estadística económica y estadística matemática que debe estudiarse de manera independiente para estimar las relaciones económicas, y así probar o refutar la validez de las teorías.”*

¿QUÉ HACEMOS CON ESO?
R: VAMOS POR PASOS

La teoría de la demanda (microeconomía) establece:



La ley de la demanda establece que: *“a medida que aumenta el precio de un bien, la cantidad demandada del mismo disminuye, en proporciones diferentes dependiendo del tipo de bien”.*

La función matemática de esta relación es:

$$D = a - bP$$

Cantidad demandada (Variable dependiente) → D

← Precio (Variable independiente) P

↑ a Punto de corte

↑ b Pendiente

↖ $\frac{\Delta D}{\Delta P}$

Función econométrica:

$$D = \beta_1 + \beta_2 P + \varepsilon$$

Cantidad demandada (Variable dependiente) → D

↑ Precio (Variable independiente) P

↓ β_1 ↓ β_2 Parámetros de estimación

→ ε Término de error ε ?

¿Cómo se estiman β_1 y β_2 ?

R: A medida que avancemos en el curso daremos la respuesta

La econometría emplea la siguiente metodología según Hendry:

- 1) Planteamiento de la teoría o hipótesis. ➔ FORMULACIÓN DE LA PREGUNTA DE INTERÉS
- 2) Formulación del modelo matemático.
- 3) Formulación del modelo econométrico.
- 4) Obtención de los datos.
- 5) Estimación de los parámetros.
- 6) Prueba de hipótesis.
- 7) Proyección o predicción.
- 8) Uso del modelo para fines de control o política.

Ejemplo: Estimación de la demanda de pollo en EEUU 1960-1982.

1) Se buscan los datos:

	YEAR	Y	X2	X3	X4	X5	X6								
	1	1960	27.8	397.5	42.2	50.7	78.3	65.8	1970	40.4	768.2	38.6	73.2	106.1	93.3
	2	1961	29.9	413.3	38.1	52.0	79.2	66.9	1971	40.3	843.3	39.8	67.8	104.8	89.7
	3	1962	29.8	439.2	40.3	54.0	79.2	67.8	1972	41.8	911.6	39.7	79.1	114.0	100.7
	4	1963	30.8	459.7	39.5	55.3	79.2	69.6	1973	40.4	931.1	52.1	95.4	124.1	113.5
	5	1964	31.2	492.9	37.3	54.7	77.4	68.7	1974	40.7	1021.5	48.9	94.2	127.6	115.3
	6	1965	33.3	528.6	38.1	63.7	80.2	73.6	1975	40.1	1165.9	58.3	123.5	142.9	136.7
	7	1966	35.6	560.3	39.3	69.8	80.4	76.3	1976	42.7	1349.6	57.9	129.9	143.6	139.2
	8	1967	36.4	624.6	37.8	65.9	83.9	77.2	1977	44.1	1449.4	56.5	117.6	139.2	132.0
	9	1968	36.7	666.4	38.4	64.5	85.5	78.1	1978	46.7	1575.5	63.7	130.9	165.5	132.1
	10	1969	38.4	717.8	40.1	70.0	93.7	84.7	1979	50.6	1759.1	61.6	129.8	203.3	154.4
									1980	50.1	1994.2	58.9	128.0	219.6	174.9
									1981	51.7	2258.1	66.4	141.0	221.6	180.8
									1982	52.9	2478.7	70.4	168.2	232.6	189.4

2) Se estiman el modelo

$$Dp = 34,51 - 0,21Pp + 0,01Y$$

Consumo de pollo que no depende ni del precio (Pp) ni del ingreso (Y)

$$\frac{\Delta Dp}{\Delta Pp}$$

$$\frac{\Delta Dp}{\Delta Y}$$

¿Qué hacemos con eso?

- Determinar la dirección de la relación.
- Cuantificar el impacto y la relación entre las variables.
- Determinar la significancia de la relación.

De estas cuestiones es precisamente de lo que se encarga la econometría.

¿De dónde salieron esos números?

S
A
L
I
D
A

E
V
I
E
W
S

Dependent Variable: Y
Method: Least Squares
Sample: 1960 1982
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	34.51561	3.855779	8.951655	0.0000
X2	0.014884	0.002193	6.785304	0.0000
X3	-0.213592	0.121905	-1.752115	0.0951
R-squared	0.910821	Mean dependent var	39.66957	
Adjusted R-squared	0.901903	S.D. dependent var	7.372950	
S.E. of regression	2.309239	Akaike info criterion	4.632821	
Sum squared resid	106.6517	Schwarz criterion	4.780929	
Log likelihood	-50.27744	Hannan-Quinn criter.	4.670070	
F-statistic	102.1340	Durbin-Watson stat	0.432741	
Prob(F-statistic)	0.000000			

Salida R:

```
> fix(Datos)
> Datos <- read.table("C:/Users/Laura/Desktop/TABLA.txt", header=TRUE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
> RegModel.6 <- lm(Y~X2+X3, data=Datos)
> summary(RegModel.6)
```

Call:

```
lm(formula = Y ~ X2 + X3, data = Datos)
```

Residuals:

```
   Min     1Q  Median     3Q    Max
-3.6182 -2.1812  0.4679  1.7500  3.1545
```

Coefficients:

```
              Estimate Std. Error  t value Pr(>|t|)
(Intercept) 34.515607   3.855779   8.952 1.97e-08 ***
X2           0.014884   0.002193   6.785 1.34e-06 ***
X3          -0.213592   0.121905  -1.752  0.0951 .
---

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.309 on 20 degrees of freedom
Multiple R-squared: 0.9108, Adjusted R-squared: 0.9019
F-statistic: 102.1 on 2 and 20 DF, p-value: 3.181e-11

Salida STATA 11:

```
. regress Y X2 X3
```

Source	SS	df	MS
Model	1089.277	2	544.638499
Residual	106.651738	20	5.33258691
Total	1195.92874	22	54.3603971

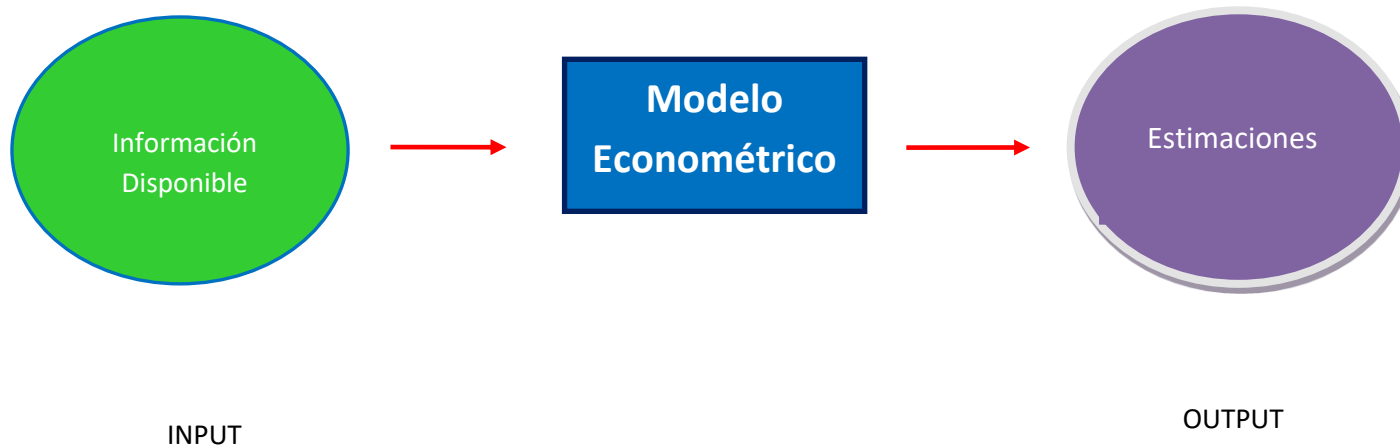
```
Number of obs = 23  
F( 2, 20) = 102.13  
Prob > F = 0.0000  
R-squared = 0.9108  
Adj R-squared = 0.9019  
Root MSE = 2.3092
```

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
X2	.0148836	.0021935	6.79	0.000	.010308	.0194591
X3	-.2135923	.1219054	-1.75	0.095	-.4678825	.0406979
_cons	34.51561	3.85578	8.95	0.000	26.47259	42.55862

TAREA: revise las salidas del Eviews, del R y del STATA, marque las similitudes y las diferencias.

Recuerde que según Goldberger (1960):

“la econometría puede ser definida como la ciencia social en la cual las herramientas de la teoría económica, las matemáticas y la inferencia estadística son aplicadas al análisis de los fenómenos económicos”.



TIPOS DE MODELOS:

DETERMINÍSTICOS O MATEMÁTICOS	$C(Y) = \beta_1 + \beta_2 Y$	Abstraen los fenómenos reales para aislar los aspectos y relaciones relevantes
ESTOCÁSTICOS	$C(Y) = \beta_1 + \beta_2 Y + \varepsilon$	Incluye elementos estocásticos en los modelos matemáticos con la finalidad de generar modelos más parsimoniosos y abstracciones de la realidad más precisas que la de los modelos determinísticos.
ESTADÍSTICOS O ECONOMÉTRICOS	Son un caso particular de los modelos estocásticos, en donde los parámetros son desconocidos y se estiman a partir de datos empíricos haciendo uso de procedimientos estadísticos.	

CLASIFICACIÓN DE LOS MODELOS ECONOMÉTRICOS:

SEGÚN TIPO	➤ DE REGRESIÓN CLÁSICOS ➤ DE SERIES DE TIEMPO
SEGÚN LA FORMA FUNCIONAL	➤ LINEALES ➤ NO LINEALES
SEGÚN EL NÚMERO DE ECUACIONES	➤ UNIECUACIONALES ➤ MULTIECUACIONALES
SEGÚN LA NATURALEZA DE LOS DATOS	➤ SERIES DE TIEMPO ➤ CORTE TRANSVERSAL ➤ DATOS COMBINADOS (MIXTOS)
SEGÚN LAS CARACTERÍSTICAS DINÁMICAS	➤ ESTÁTICOS ➤ DINÁMICOS

TIPOS DE VARIABLES:

- **DEPENDIENTES (VD):** variable explicada, predicha, regresada, respuesta, endógena, resultado, variable controlada. Es la variable que interesa estudiar.
- **INDEPENDIENTE (VI):** variable independiente, predictora, regresora, estímulo, exógena, covariante, variable de control.
- **CUALITATIVAS:** también llamadas DUMMY o CATEGÓRICAS representan cualidades de las unidades muestrales. Región donde habita, sexo, religión, entre otras.
- **CUANTITATIVAS:** representan cantidades de las unidades muestrales. Tasa de desempleo, inflación, PIB. Estas serán las variables dependientes usadas durante el curso.

TIPOS DE DATOS:

Se denomina **DATOS** a las observaciones que suministran información sobre un evento o un individuo.

El econometrista, a diferencia del estadístico económico, busca utilizar las cifras recopiladas por el último para probar o refutar la validez de una teoría económica. El hacedor de modelos se enfrenta al uso de datos *no observados* (NO EXPERIMENTALES) que se contraponen a los experimentales por ende:

- Quien hace modelos econométricos debe dominar muy bien distintas habilidades para el manejo de datos no experimentales.
- El que elabora modelos se familiariza por completo con la naturaleza y estructura de los datos.

A) **DATOS SERIES DE TIEMPO:** es un conjunto de observaciones sobre los valores que toma la variable en diferentes momentos del tiempo. Tal información se recopila a intervalos regulares: diaria, semanal, quincenal, mensual, semestral, entre otros.

- Temporales.
- Orden cronológico.
- Aleatorios.
- Periódicos.

Ejemplo: Datos Tasa de cambio para varios países

obs	SUIZA	SUECIA	JAPON	GRAN_BRE TANA	FRANCIA	CANADA
1977	2.406500	4.480200	268.6200	1.744900	4.916100	1.063300
1978	1.790700	4.520700	210.3900	1.918400	4.509100	1.140500
1979	1.664400	4.289300	219.0200	2.122400	4.256700	1.171300
1980	1.677200	4.231000	226.6300	2.324600	4.225100	1.169300
1981	1.967500	5.066000	220.6300	2.024300	5.439700	1.199000
1982	2.032700	6.283900	249.0600	1.748000	6.579400	1.234400
1983	2.100700	7.671800	237.5500	1.515900	7.620400	1.232500
1984	2.350000	8.270800	237.4600	1.336800	8.735600	1.295200
1985	2.455200	8.603200	238.4700	1.297400	8.980000	1.365900
1986	1.797900	7.127300	168.3500	1.467700	6.925700	1.389600
1987	1.491800	6.346900	144.6000	1.639800	6.012200	1.325900
1988	1.464300	6.137000	128.1700	1.781300	5.959500	1.230600

B) DATOS DE CORTE TRANSVERSAL (CROSS SECTION): consiste en datos de una o más variables sobre distintos individuos recogidos en un mismo momento del tiempo. Un ejemplo tradicional son los conocidos censo poblacionales.

- Atemporales.
- Proviene de un muestreo aleatorio
- Las observaciones no siempre pertenecen al mismo periodo de tiempo con precisión.
- No existe un orden natural.
- Son más comunes en microeconomía.

Ejemplo: Datos de automóviles de pasajeros. Tabla 11.7 Gujarati 4ta Edición.

obs	Velocidad Máxima	Espacio Cabina en Pies cúbicos	Peso del vehículo Libras	Millas por medio galón	Caballos de fuerza
1	96.00000	89.00000	17.50000	65.40000	49.00000
2	97.00000	92.00000	20.00000	56.00000	55.00000
3	97.00000	92.00000	20.00000	55.90000	55.00000
4	105.0000	92.00000	20.00000	49.00000	70.00000
5	96.00000	92.00000	20.00000	46.50000	53.00000
6	105.0000	89.00000	20.00000	46.20000	70.00000
7	97.00000	92.00000	20.00000	45.40000	55.00000
8	98.00000	50.00000	22.50000	59.20000	62.00000
9	98.00000	50.00000	22.50000	53.30000	62.00000
10	107.0000	94.00000	22.50000	43.40000	80.00000
11	103.0000	89.00000	22.50000	41.10000	73.00000
12	113.0000	50.00000	22.50000	40.90000	92.00000
13	113.0000	99.00000	22.50000	40.90000	92.00000

C) **INFORMACIÓN COMBINADA:** llamados también datos agrupados, tienen elementos de series de tiempo y corte transversal. Un tipo especial de estos datos son los de *información de panel o longitudinal*, también llamados *micropanel*, en los cuales la misma unidad de corte transversal (individuo, empresa, o familia) es encuestada a través del tiempo.

Sector	Salario	Sexo	Individuo	Tiempo
1	500	M	1	1
1	450	F	2	2
2	750	M	3	3
1	800	M	4	1
2	450	M	5	2

FUENTES DE INFORMACIÓN:

- **PRIMARIAS:** son aquellas derivadas de la recolección directa de los datos a través de distintos procedimientos de muestreo.
- **SECUNDARIAS:** datos recolectados por otros investigadores previamente, son los más utilizados en economía.

PRECISIÓN DE LA INFORMACIÓN:

A pesar de la existencia de cuantiosa información económica para su investigación, la calidad de ésta no siempre es buena por ello el investigador debe tener presente que:

- La mayor parte de la información económica es de carácter no experimental por ende, hay mayor probabilidad de incurrir en errores de observación, bien sea por acción u omisión.
- Errores de medida.
- Sesgo de selectividad muestral.
- Cambios en los métodos de muestreo.
- Información agregada.

RECUERDE: “EL INVESTIGADOR DEBE TENER SIEMPRE EN MENTE QUE EL RESULTADO DE LA INVESTIGACIÓN SOLAMENTE SERÁ TAN BUENO COMO LO SEA LA CALIDAD DE LOS DATOS”. D. Gujarati (p.29)

UNA OBSERVACIÓN SOBRE LAS ESCALAS DE MEDICIÓN DE LAS VARIABLES:

- **ESCALA DE PROPORCION:** para una variable Y , al ésta tomar dos valores (Y_1 y Y_2) la proporción (Y_1/Y_2) y la distancia (Y_1-Y_2) son cantidades que tienen sentido. Los valores que toma la variable presentan un ordenamiento natural. Ejemplo: PIB, Inflación, Exportaciones, Reservas internacionales, entre otras.
- **ESCALA DE INTERVALO:** satisface las dos últimas propiedades de la variable de escala de proporción pero no la primera.
- **ESCALA ORDINAL:** pertenecen a esta categoría aquellas variables que satisfacen la tercera propiedad de las variables de escala de proporción. Ejemplo: estrato social, grado académico, calificaciones, entre otras.
- **ESCALA NOMINAL:** pertenecen a este grupo aquellas variables que representan cualidades, género, estado civil, estudios, entre otros; es decir, denotan categorías.