

Análisis descriptivo, tablas y gráficos dinámicos utilizando Excel



Andrés M. Castaño

Instituto de Economía Aplicada Regional (IDEAR)

2 de junio de 2016

1 Introducción

2 Tipos de variables y estadísticas descriptivas

3 Aplicación a datos reales de Chile a nivel comunal, año 2011

- ¿ A qué problemas de interés social y económico nos podemos aproximar con esta base de datos?

4 Análisis descriptivo en Excel: paso a paso

5 Tablas dinámicas e informes de tabla dinámica

Introducción

- Durante su curso de Estadística van a discutir la importancia de convertir bases de datos complejas en información y en muchas ocasiones, conocimiento. Esta información debe tener algún sentido: toma de decisiones, entendimiento del problema, formulación de política pública, etc.
- El propósito del análisis descriptivo es describir situaciones o eventos. Esto es, decir **cómo** es, y **cómo** se manifiesta cierto fenómeno, ya sea del ambiente económico, social, político o cultural.
- Los estudios descriptivos buscan especificar las propiedades importantes de personas, grupos, comunidades o cualquier otro fenómeno que sea sometido a análisis (Dankhe, 1986).
- El **análisis descriptivo** sirve como una fase previa a los **estudios correlacionales y explicativos**.
- **Ejemplo 1:** piense que desea estudiar las características que tiene el emprendimiento en las regiones de Chile (estudio descriptivo).

Ejemplo 1: emprendimiento en Chile

- piense que desea estudiar las características que tiene el emprendimiento en las regiones de Chile (estudio descriptivo)

Consideraciones iniciales

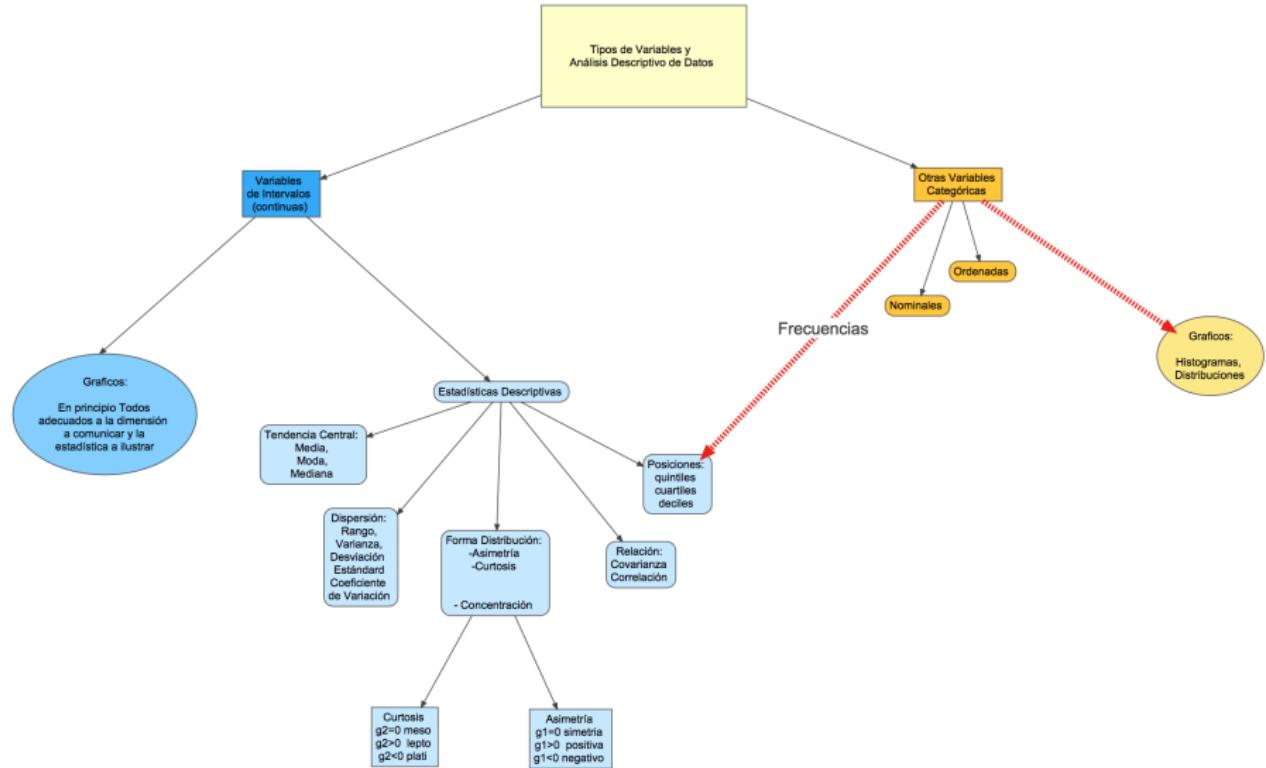
- Es probable que muchos estén más interesados en saber qué factores afectan el emprendimiento en las regiones de Chile? ⇒ **estudio explicativo**.
- O en determinar: ¿Cómo es la relación entre emprendimiento y escolaridad? ¿El emprendimiento es más grande en las comunas con mayor desempleo? ¿Cómo es la relación entre emprendimiento en las comunas y el salario que perciben los habitantes de dicha comuna? ⇒ **estudio correlacional**.

Ejemplo 1: emprendimiento en Chile

Consideraciones iniciales

- No obstante, no podemos entrar a explicar causas del emprendimiento, cuando ni siquiera conocemos:
 - ¿Qué tipo de variable de emprendimiento tengo? ⇒ **¿Continuas o categóricas?**
 - ¿Cuáles son sus propiedades estadísticas?
 - ¿Cómo es a nivel nacional y regional?
 - ¿Es variable a lo largo de las regiones? ⇒ Esto permitirá identificar potenciales datos atípicos (outliers) que pueden afectar un estudio de carácter explicativo.
- Dichas preguntas **no las responde** un estudio explicativo, sino un estudio descriptivo que incluya: medidas de **tendencia central, dispersión y posicionamiento**, y estudiar la forma de la **distribución**.

Tipos de variables y estadísticas descriptivas



Nota sobre las escalas de medición

- Escalas de medición de las variables (criterios: distancia, proporción y orden natural)
- Escala de proporción (ejemplo: $\frac{PIB_2}{PIB_1}$) \implies cumple todos los criterios.
- Escala de intervalo \implies sólo es significativa la distancia pero no la proporción, ni el orden (ejemplo: tiempo).
- Escala ordinal \implies sólo cumplen el orden natural (ejemplo: sistema de calificación, tipos de ingresos).
- Escala nominal \implies no cumple ninguno de los criterios (ejemplo: sexo, estado civil).

- Con fines de ilustrar como realizar análisis estadístico, tablas y gráficos dinámicos en Microsoft Excel 2013, he construido una base de datos para Chile con variables de distinta índole (nominales, intervalo) que están relacionados con problemas económicos y sociales.

Aplicación a datos reales de Chile a nivel comunal, año 2011

base_comunal-2 - Excel

Iniciar sesión

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA

Cortar Copiar Copiar formato Portapapeles Fuente Alineación Número Estilos Celdas Rellenar Ordenar Buscar y filtrar Modificar

A B C D E F G H I J K L M N O P Q

macro	region	idregion	comuna	idcomun	tipo	tamaño	inicios	esc	educ1%	educ2%	educ3%	desempleo%	robos	homicidios	pob_tot	extens	km2	p
2	Norte	Tarapaca	Iquique	1101	urbano	5	1551	11.5	34.5	54.4	11.0	6.6	2347.0	10.0	186.0	2262.0		
3	Norte	Tarapaca	Alto Hospicio	1107	rural	5	564	10.2	48.9	47.8	3.3	7.8	1130.0	7.0	93.8	573.0		
4	Norte	Tarapaca	Pozo Almonte	1401	rural	3	71	9.7	53.7	43.7	2.6	4.5	41.0	0.0	17.3	13766.0		
5	Norte	Tarapaca	Camiña	1402	rural	1	8	5.2	82.6	16.7	0.8	0.0	0.0	0.0	1.0	2200.0		
6	Norte	Tarapaca	Huara	1404	rural	1	25	8.9	54.9	41.7	3.4	4.7	1.0	0.0	3.4	10475.0		
7	Norte	Tarapaca	Pica	1405	rural	3	26	9.2	51.5	40.7	7.8	7.7	10.0	0.0	18.6	8934.0		
8	Norte	Antofagasta	Antofagasta	2101	urbano	5	1896	11.6	35.2	52.7	12.1	10.9	4012.0	19.0	373.0	30718.0		
9	Norte	Antofagasta	Mejillones	2102	rural	2	72	10.0	50.0	45.2	4.8	17.2	67.0	0.0	11.1	3804.0		
10	Norte	Antofagasta	Sierra Gorda	2103	rural	1	19	9.4	47.1	50.0	2.9	13.3	2.0	0.0	3.7	12886.0		
11	Norte	Antofagasta	Taltal	2104	rural	2	62	10.3	43.0	47.0	10.0	11.1	20.0	2.0	10.6	20405.0		
12	Norte	Antofagasta	Calama	2201	urbano	5	740	10.9	40.6	49.3	10.1	9.5	1297.0	10.0	148.8	15597.0		
13	Norte	Antofagasta	San Pedro de Atacama	2203	rural	2	59	9.4	55.2	36.0	8.8	11.8	30.0	0.0	9.3	23439.0		
14	Norte	Antofagasta	Tocopilla	2301	rural	3	135	10.6	41.7	51.7	6.6	6.5	100.0	1.0	21.3	4039.0		
15	Norte	Antofagasta	Maria Elena	2302	rural	1	13	11.2	34.1	59.6	6.3	12.9	13.0	0.0	3.7	12197.0		
16	Norte	Atacama	Copiapo	3101	urbano	5	657	10.6	41.7	49.5	8.8	9.1	1379.0	5.0	163.9	16681.0		
17	Norte	Atacama	Caldera	3102	rural	2	96	10.5	38.5	53.6	7.8	8.6	132.0	0.0	15.2	4667.0		
18	Norte	Atacama	Tierra Amarilla	3103	rural	2	70	9.5	52.0	44.8	3.2	12.3	36.0	1.0	13.9	11191.0		
19	Norte	Atacama	Chafaral	3201	rural	2	66	10.1	45.5	49.9	4.6	6.9	64.0	1.0	12.7	5772.0		
20	Norte	Atacama	Diego de Almagro	3202	rural	2	57	10.5	46.3	45.5	8.2	8.4	20.0	0.0	11.7	18664.0		
21	Norte	Atacama	Vallenar	3301	rural	4	224	9.7	45.6	49.7	4.8	8.6	188.0	0.0	46.5	7084.0		
22	Norte	Atacama	Alto Del Carmen	3302	rural	1	22	8.8	61.1	33.2	5.6	10.0	1.0	0.0	4.8	5939.0		
23	Norte	Atacama	Freirina	3303	rural	1	24	9.7	47.7	44.4	8.0	6.2	4.0	0.0	5.9	3578.0		
24	Norte	Atacama	Huasco	3304	rural	1	27	10.2	44.7	47.9	7.3	11.6	22.0	0.0	8.0	1601.0		
25	Norte	Coquimbo	La Serena	4101	urbano	5	1179	11.1	37.3	50.4	12.4	9.0	1265.0	4.0	214.7	1893.0		
26	Norte	Coquimbo	Coquimbo	4102	rural	5	972	10.5	42.4	49.7	7.9	8.8	1391.0	5.0	215.1	1429.0		
27	Norte	Coquimbo	Andacollo	4103	rural	1	67	9.1	51.2	43.4	5.4	7.2	32.0	0.0	8.5	310.0		
28	Norte	Coquimbo	La Higuera	4104	rural	1	28	7.8	64.1	34.1	1.8	10.1	9.0	0.0	3.9	4158.0		

Fuente: elaboración propia

Descripción de variables

base_comunal-2 - Excel

Iniciar sesión

A	B	C	D
1 macrozona	Ubicación de las comunas en grandes zonas (norte, centro, sur, austral)		
2 region	Nombre de la región		
3 idregion	Identificador para las regiones de chile (1=primera region, 2=segunda region, etc)		
4 comuna	Nombre de la comuna		
5 idcomuna	Identificador de comuna de acuerdo a la Encuesta de Caracterización Socioeconómica Nacional del año 2011 (CASEN)		
6 tipo	comuna rural o urbana		
7 tamaño	Tamaño de la comuna de acuerdo a población en orden ascendente (1=las comunas más pequeñas, 5=las comunas más grandes)		
8 inicios	Total de empresas de primera categoría que inician actividades en el Servicio de Impuestos Internos (SII), puede ser considerada una buena aproximación al emprendimiento		
9 esc	Años promedio de escolaridad de la población encuestada en la CASEN 2011		
10 educ1%	Porcentaje de personas que completaron estudios básicos		
11 educ2%	Porcentaje de personas que completaron estudios de enseñanza media		
12 educ3%	Porcentaje de personas que completaron la educación superior		
13 desempleo%	Tasa de desempleo comunal (CASEN 2011)		
14 robos	Número de robos con violencia e intimidación		
15 homicidios	Número de homicidios		
16 pob_tot	Población comunal total (en miles de habitantes)		
17 exten_km2	Extensión en kilómetros cuadrados		
18 pobre	Porcentaje de población pobre (%)		
19 ne	Total de empresas existentes en la comuna (en miles de empresas)		
20 ventas	Total de ventas de las empresas existentes en la comuna (en millones de UF)		
21 ntrab	Total de trabajadores de las empresas existentes (en miles de trabajadores)		
22 remu	Total de remuneraciones de las empresas existentes (en millones de UF)		
23 docentes_em	Total de docentes de aula en la enseñanza media		
24 matl_em	Total de matriculados en la enseñanza media		
25 colegios_em	Total de colegios de enseñanza media		
26 salario	Salario promedio a partir de CASEN 2011 (en miles de pesos chilenos)		
27 pea	Población Económicamente Activa de la región (en miles de trabajadores)		
28 emprendimiento	Indicador de emprendimiento normalizado (inicios de empresas por cada 10000 miembros de la PEA)		

Fuente: elaboración propia

Contenido

1 Introducción

2 Tipos de variables y estadísticas descriptivas

3 Aplicación a datos reales de Chile a nivel comunal, año 2011

- ¿ A qué problemas de interés social y económico nos podemos aproximar con esta base de datos?

4 Análisis descriptivo en Excel: paso a paso

5 Tablas dinámicas e informes de tabla dinámica

¿Potenciales preguntas de investigación?

- ¿Cómo es el emprendimiento a lo largo de las comunas de Chile?
- ¿Qué relación tiene el emprendimiento con la escolaridad, y el desempleo?
- ¿Cómo ha sido el comportamiento de las ventas de las empresas de la Región de Antofagasta en comparación al resto de regiones?
- ¿Cómo se comportan los robos en la región de Antofagasta respecto al resto de regiones de Chile?
- ¿Cómo es la relación entre los robos cometidos en una comuna y su nivel de desempleo?
- Con el ¿cómo? Nos referimos a problemas dentro de las ciencias sociales y empresariales que hacen referencia a la descripción estadística (media, máximo, mínimo, varianza, etc) de la variable objeto de estudio.
- Con el ¿qué? Nos referimos a problemas dentro de las ciencias sociales y empresariales que hacen referencia a una relación estadística entre las variables de estudio (covarianza, correlación).

Análisis descriptivo en Excel: paso a paso

- Ahora pasaremos a realizar un análisis estadístico descriptivo de algunas variables en la base de datos y aprenderemos a interpretar los resultados.
- Aunque existen muchos programas especializados para elaborar informes estadísticos (SPSS, Stata, entre otros), Microsoft Excel 2013 para Windows cuenta con complementos adicionales que permiten aprender de manera rápida a obtener las principales estadísticas descriptivas.
- Queremos realizar un reporte estadístico rápido de la variable **"inicios"** (como una aproximación al emprendimiento) e interpretarlo apropiadamente.

- Seleccionar la opción **Archivo**.
- Aparecerá un menú, escoger la alternativa **opciones**.
- Se desplegará un conjunto de Opciones, escoger **Complementos**.
- En la parte derecha inferior aparecerá una opción que indica administrar, a su lado presionar la opción **Ir**.
- Se desplegará un asistente de complementos, con varias opciones, marcar por ahora la opción que dice **Herramientas para análisis**.
- con estos pasos hemos activado el complemento para análisis estadístico.

Paso 1: activar el análisis estadístico en Microsoft Excel 2010

Screenshot of Microsoft Excel 2010 showing the 'ANÁLISIS DE DATOS' ribbon tab selected. A context menu is open over cell B6, listing various statistical tools like 'Herramientas para análisis', 'Herramientas para análisis - VBA', 'Herramientas para el euro', and 'Solver'. The 'Herramientas para análisis' option is checked.

Variable	Descripción
idregion	Identificador para las regiones de chile (1=primera region, 2=segunda region, etc)
region	Nombre de la región
id_comuna	Identificador de comuna de acuerdo a la Encuesta de Caracterización Socioeconómica Nacional del año 2011 (CASEN)
comuna	Nombre de la comuna
tipo-1	identificador de tipo de comuna (1=urbana, 0=rural)
tipo	comuna rural o urbana
inicios	Total de empresas de primera categoría que iniciaron actividades en el Servicio de Impuestos Internos (SII), puede ser considerada una buena aproximación al emprendimiento
esc	Año promedio de escolaridad de la población encuestada
educ1%	Porcentaje de personas que completaron estudios básicos
educ2%	Porcentaje de personas que completaron estudios de educación media
educ3%	Porcentaje de personas que completaron la educación superior
desempleo%	Tasa de desempleo comunal (CASEN 2011)
robos	Número de robos con violencia e intimidación
homicidios	Número de homicidios
pob_tot	Población comunal total (en miles de habitantes)
pobre%	Porcentaje de población pobre (%)
ne	Total de empresas existentes en la comuna (en miles de pesos)
ventas	Total de ventas de las empresas existentes en la comuna (en miles de pesos)
ntrab	Total de trabajadores de las empresas existentes (en miles de pesos)
remu	Total de remuneraciones de las empresas existentes (en miles de pesos)
docentes_em	Total de docentes de aula en la enseñanza media
matl_em	Total de matriculados en la enseñanza media
colegios_em	Total de colegios de enseñanza media
salario	Salario promedio a partir de CASEN 2011 (en miles de pesos)
pea	Populación Económicamente Activa de la región (en miles de pesos)
emprendimiento	Indicador de emprendimiento normalizado (inicios de empresas)
robos-2	Indicador de robos normalizado (robos por cada 100 mil habitantes)

Paso 1: activar el análisis estadístico en Microsoft Excel 2010

Screenshot of Microsoft Excel 2010 showing the ribbon and a data analysis example.

The ribbon tabs shown are: ARCHIVO, INICIO, INSERTAR, DISEÑO DE PÁGINA, FÓRMULAS, DATOS (highlighted with a red arrow), REVISAR, VISTA.

The DATOS tab has the following options: Desde Access, Desde web, De otras fuentes..., Conexiones existentes, Actualizar todo, Propiedades, Conexiones (highlighted with a red arrow), Borrar, Ordenar, Filtro, Avanzadas, Relleno rápido, Quitar duplicados, Análisis de hipótesis, Texto en columnas, Validación de datos, Relaciones, Agrupar, Desagrupar, Subtotal, Esquema, and Análisis.

The data analysis example is titled "base_comunal-2 - Excel". It contains a table with columns: id_comuna, comuna, tipo-1, tipo, inicios, esc, educ1%, educ2%, educ3%, desempleo%, robos, homicidios, pob_tot, pobre%, ne, ventas, ntrab, remu, docentes, em_m.

The data shows various statistics for different communes in Chile, such as Iquique, Alto Hospicio, Pozo Almonte, etc., comparing urban and rural populations across various socio-economic indicators.

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	id_comuna	comuna	tipo-1	tipo	inicios	esc	educ1%	educ2%	educ3%	desempleo%	robos	homicidios	pob_tot	pobre%	ne	ventas	ntrab	remu	docentes	em_m
2	1101	Iquique	1 urbano		1551	11,5	34,5	54,4	11,0	6,6	2347,0	10,0	186,0	14,0	13,2	175,7	82,5	8,7	2190,0	1
3	1107	Alto Hospicio	0 rural		564	10,2	48,9	47,8	3,3	7,8	1130,0	7,0	93,8	20,0	2,9	8,6	8,4	0,7	980,0	
4	1401	Pozo Almonte	0 rural		71	9,7	53,7	43,7	2,6	4,5	41,0	0,0	17,3	13,0	0,6	32,9	4,5	1,2	157,0	
5	1402	Camiña	0 rural		8	5,2	82,6	16,7	0,8	0,0	0,0	0,0	1,0	15,0	0,1	0,1	0,0	0,0	21,0	
6	1404	Huara	0 rural		25	8,9	54,9	41,7	3,4	4,7	1,0	0,0	3,4	10,0	0,1	0,8	0,4	0,0	37,0	
7	1405	Pica	0 rural		26	9,2	51,5	40,7	7,8	7,7	10,0	0,0	18,6	14,0	0,3	0,3	0,6	0,1	72,0	
8	2101	Antofagasta	1 urbano		1896	11,6	35,2	52,7	12,1	10,9	4012,0	19,0	373,0	7,0	16,9	292,5	121,3	23,5	3214,0	2
9	2102	Mejillones	0 rural		72	10,0	50,0	45,2	4,8	17,2	67,0	0,0	11,1	6,0	0,5	11,5	2,1	0,4	107,0	
10	2103	Sierra Gorda	0 rural		19	9,4	47,1	50,0	2,9	13,3	2,0	0,0	3,7	2,0	0,1	5,8	0,9	0,5	12,0	
11	2104	Tal Tal	0 rural		62	10,3	43,0	47,0	10,0	11,1	20,0	2,0	10,6	11,0	0,7	161,9	2,4	0,5	120,0	
12	2201	Calama	1 urbano		740	10,9	40,6	49,3	10,1	9,5	1297,0	10,0	148,8	9,0	7,2	60,2	41,4	7,0	1506,0	1
13	2203	San Pedro de Atacama	0 rural		59	9,4	55,2	36,0	8,8	11,8	30,0	0,0	9,3	6,0	0,5	0,9	1,3	0,1	55,0	
14	2301	Tocopilla	0 rural		135	10,6	41,7	51,7	6,6	6,5	100,0	1,0	21,3	13,0	1,2	24,3	4,3	0,9	238,0	
15	2302	María Elena	0 rural		13	11,2	34,1	59,6	6,3	12,9	13,0	0,0	3,7	7,0	0,2	0,7	1,4	0,2	53,0	
16	3101	Copiapó	1 urbano		657	10,6	41,7	49,5	8,8	9,1	1379,0	5,0	163,9	18,0	7,5	129,4	59,3	10,4	1659,0	1
17	3102	Caldera	0 rural		96	10,5	38,5	53,6	7,8	8,6	132,0	0,0	15,2	16,0	0,9	3,6	3,4	0,6	193,0	
18	3103	Tierra Amarilla	0 rural		70	9,5	52,0	44,8	3,2	12,3	36,0	1,0	13,9	15,0	0,6	5,5	5,3	0,7	142,0	
19	3201	Chañaral	0 rural		66	10,1	45,5	49,9	4,6	6,9	64,0	1,0	12,7	22,0	0,9	6,5	5,1	0,9	168,0	
20	3202	Diego de Almagro	0 rural		57	10,5	46,3	45,5	8,2	8,4	20,0	0,0	11,7	10,0	0,7	5,5	4,4	1,0	229,0	
21	3301	Vallenar	0 rural		224	9,7	45,6	49,7	4,8	8,6	188,0	0,0	46,5	19,0	3,0	9,1	11,6	1,5	579,0	
22	3302	Alto Del Carmen	0 rural		22	8,8	61,1	33,2	5,6	10,0	1,0	0,0	4,8	6,0	0,4	0,3	0,9	0,1	35,0	
23	3303	Freirina	0 rural		24	9,7	47,7	44,4	8,0	6,2	4,0	0,0	5,9	16,0	0,3	0,4	0,7	0,1	61,0	
24	3304	Huasco	0 rural		27	10,2	44,7	47,9	7,3	11,6	22,0	0,0	8,0	9,0	0,4	0,9	2,1	0,2	108,0	
25	4101	La Serena	1 urbano		1179	11,1	37,3	50,4	12,4	9,0	1265,0	4,0	214,7	14,0	10,2	103,1	63,3	8,9	2631,0	1
26	4102	Coquimbo	1 urbano		972	10,5	42,4	49,7	7,9	8,8	1391,0	5,0	215,1	15,0	9,0	45,9	57,6	7,5	2177,0	1
27	4103	Andacollo	0 rural		67	9,1	51,2	43,4	5,4	7,2	32,0	0,0	8,5	19,0	0,5	19,5	2,5	0,7	94,0	
28	4104	La Higuera	0 rural		28	7,8	64,1	34,1	1,8	10,1	9,0	0,0	3,9	14,0	0,2	0,4	0,3	0,0	50,0	

Bottom status bar: datoscomu, descripción de variables, 4, 100%.

Paso 2: análisis descriptivo para la variable "inicios"

- Seleccionar la opción **Datos**.
- Aparecerá debajo un conjunto de opciones, en la parte derecha escoger la alternativa **Análisis de datos**.
- Se desplegará un conjunto de opciones, escoger **Estadística descriptiva** y marcar **Aceptar**.
- Aparecerá un menú con varios opciones, en la opción entrada, subopción **Rango de entrada** se debe seleccionar el rango de celdas que se quiere utilizar para efectuar el análisis estadístico, seleccionaremos las celdas de la columna G ("inicios"), desde G1 hasta G325 que pertenecen a la variable **inicios**.

Paso 2: análisis de estadística descriptiva para la variable "inicios"

The screenshot shows a Microsoft Excel spreadsheet titled "base_comunal-2 - Excel". The ribbon menu is visible at the top, with the "DATOS" tab selected. A context menu is open over a range of cells in the sheet, specifically over the header row of a data table. This menu includes options like "Entrada", "Rango de entrada: \$G\$1:\$G\$325", "Agrupado por: Columnas", "Bóltulos en la primera fila" (which is checked), and "Opciones de salida". The "Opciones de salida" section contains radio buttons for "Rango de salida:" (unchecked), "En una hoja nueva:" (checked), and "En un libro nuevo". Below these are several check boxes: "Resumen de estadísticas" (checked), "Nivel de confianza para la media: 95 %", "K-ésimo mayor: 1", and "K-ésimo menor: 1". The main data table below the header row contains 32 rows of data, starting with "13 Metropolitano" and ending with "15 Arica". The columns represent various variables such as "Comuna", "ID", "Nombre", "Tipos", "Población", and "Porcentaje".

- En el la opción **Agrupado por** marcar **Columnas**.
- Marcar la opción **Rótulos en la primera fila** esto le indica a Excel que la primera fila contiene los nombres de las variables.
- En el sub menú **Opciones de salida** marcar **En una hoja nueva**, luego se deben marcar las opciones **Resumen de estadísticas** y **Nivel de estadístico para la media**.
- Luego marcar la opción **Aceptar** y obtendremos nuestra resumen de estadística descriptiva en una nueva hoja de Excel.

Paso 2: análisis de estadística descriptiva para la variable "inicios"

base_comunal-2 - Excel

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA

Desde Access Desde web Desde texto De otras fuentes Conexiones existentes Actualizar todo Conexiones Propiedades Editar vínculos

Obtener datos externos

Ordenar Filtro Avanzadas Borrar Volver a aplicar Texto en columnas Quitar Valida

C18 A B C D E F G

	A	B	C	D	E	F	G
1		<i>inicios</i>					
2							
3	Media	282.6882716					
4	Error típico	30.95952084					
5	Mediana	100					
6	Moda	25					
7	Desviación estándar	557.2713751					
8	Varianza de la muestra	310551.3855					
9	Curtosis	38.69262467					
10	Coeficiente de asimetría	5.404011949					
11	Rango	5249					
12	Mínimo	8					
13	Máximo	5257					
14	Suma	91591					
15	Cuenta	324					
16	Nivel de confianza(095%)	60.90776773					
17							
18							
19							
20							
21							
22							

Paso 3: interpretar los resultados

- **Media (tendencia central):** en promedio en el año 2011, se inician casi 283 nuevas empresas por comuna.
- **Error típico (dispersión):** este estadístico corresponde al error típico de la media. Se calcula como $ET = \frac{\sigma}{\sqrt{n}}$, siendo σ el error o desviación estándar de la muestra y n el tamaño de la muestra. Nos permite saber entre que rango está la media de la población. Esto quiere decir que si sacamos muchas muestras aleatorias de tamaño $n=324$, hay una probabilidad del 95 % que el emprendimiento promedio en las comunas de Chile se va a encontrar entre $(283 - (30,95 * 1,96))$ y $(283 + (30,95 * 1,96)) \Rightarrow (222,3; 343,6)$.
- **Mediana y moda (tendencia central):** para el caso de la mediana, nos quiere decir que el 50 % de las comunas de Chile contaron con 100 o más inicios de empresas en el año 2011. La moda carece de sentido práctico para el caso de variables continuas en una escala de intervalo.

Paso 3: interpretar los resultados

- **Desviación estándar, varianza y rango (dispersión):** en general la desviación estándar tiende a interpretarse como una medida de incertidumbre, en el año 2011 el promedio de la variación del emprendimiento respecto a la media es 557. Esto quiere decir que si el emprendimiento siguiera una distribución normal cerca del 68 % de los datos se encuentran entre $(283 - 557,27)$ y $(283 + 557)$. Como se puede ver esto nos puede dar un indicio de que existen muchos valores atípicos en la distribución de la variable inicios, por lo cual es poco confiable (ya volveremos a este problema).

Paso 3: interpretar los resultados

- **Curtosis y asimetría (forma de la distribución):** estos estadísticos nos van a confirmar nuestras sospechas respecto a que la variable *inicios* probablemente no se distribuye de manera normal. Por ejemplo el coeficiente de simetría positivo (5.4) nos indica que la mayoría de los datos se encuentran por debajo del valor de la media (282.6). En una variable que se distribuye normal el coeficiente de asimetría se encuentra entre -0.5 y +0.5. El coeficiente de curtosis (apuntamiento), nos indica que la distribución de la variable *inicios* es leptocurtica, es decir existe una gran concentración de datos en la región central de la media. Para una distribución normal el coeficiente de curtosis debe estar cercano a 3.
- **Qué problemas puede tener representar el emprendimiento con esta variable?**
- **Qué se le ocurriría como un mejor indicador de emprendimiento?**

Práctica

- Formar grupos de 5.
- Tomar dos variables de la base de datos que representen problemas (historias) que usted quiera describir.
- Calcule los principales estadísticos como se ha explicado.
- Prepare una presentación corta discutiendo su historia y analizando los estadísticos.
- Tiempo para tener listo el ejercicio 10 minutos. Tiempo para presentar sus resultados 5 minutos.

Tablas dinámicas e informes de tabla dinámica

- Las tablas dinámicas te brindarán la posibilidad de resumir, analizar, explorar y presentar datos de resumen.
- Un informe de tabla dinámica es una forma interactiva de resumir rápidamente grandes volúmenes de datos. Utiliza un informe de tabla dinámica para analizar datos numéricos en profundidad y para responder preguntas no anticipadas sobre los datos.
- A través de los informes de gráfico dinámico podrás ver los datos de resumen contenidos en un informe de tabla dinámica para realizar comparaciones, patrones y tendencias, a través de los cuales podrás optimizar tus tareas repetitivas.

Tablas dinámicas e informes de tabla dinámica

- Las tablas dinámicas te brindarán la posibilidad de resumir, analizar, explorar y presentar datos de resumen.
- Un informe de tabla dinámica es una forma interactiva de resumir rápidamente grandes volúmenes de datos. Utiliza un informe de tabla dinámica para analizar datos numéricos en profundidad y para responder preguntas no anticipadas sobre los datos..
- A través de los informes de gráfico dinámico podrás ver los datos de resumen contenidos en un informe de tabla dinámica para realizar comparaciones, patrones y tendencias, a través de los cuales podrás optimizar tus tareas repetitivas.

Sugerencias para trabajar con Tablas Dinámicas

- Se sugiere que todas las columnas estén rotuladas. Pues los títulos se convertirán en los campos de la tabla dinámica.
- No deben existir columnas vacías.
- Se sugiere que el archivo esté en su formato plano más sencillo.
- Los nombres de las columnas deben estar relacionados con la información que contiene cada columna.

- Con la base de datos construida se pueden describir algunos fenómenos de interés económico y social.
- **Ejemplo 1:** la TED intenta responder a la pregunta respecto de las razones que llevan a la concentración desigual del delito y de la violencia en el espacio urbano. Esta teoría predice que los lugares con una alta concentración espacial de las viviendas baratas, con niveles de pobreza altos, con alta densidad de población y con prevalencia de la pasividad social los delitos son mayores. Para una revisión conceptual véase:

http://www.seguridadenbarrios.cl/docs/marco_conceptual.pdf

Ejercicio práctico propuesto para la TED

- Juntarse en grupos de 3 estudiantes.
- Realizar el análisis descriptivo de las variables que involucran la TED (hoja 1), de acuerdo a la base de datos propuesta.
- A través de gráficos muestra el tipo de relación que existe entre las variables de la TED (hoja 2).
- A través de tablas y gráficos dinámicos muestre el comportamiento en simultaneo de las variables que involucran la TED por región, macrozona, utilizando los estadísticos que crea convenientes, por ejemplo, la media, desviación estándar, etc (hoja 3).
- Prepare una presentación donde muestra hasta que punto la TED puede explicar el delito en Chile.
- El profesor pasará a resolver dudas grupo a grupo.
- El tiempo para preparar el material son 30 minutos. Tiempo para presentar y discutir los resultados: 10 minutos.