

## **Actividad 2**

### Obtención de estadísticas descriptivas

Andrea Valeria Pintor Valencia A01067424

Victoria Ramírez Castro A01640824

Carlos Bernabé Rojas Medina A01641668

Juan Diego Salcedo García A01368540

octubre 2023

Herramientas computacionales: el arte de la analítica

Gpo 201

[https://github.com/apintorv/Semana\\_Tec.git](https://github.com/apintorv/Semana_Tec.git)

[https://colab.research.google.com/drive/1DcAQ\\_BcnDeAvKkK2trEbT3KIm6ZdkgXb?usp=sharing](https://colab.research.google.com/drive/1DcAQ_BcnDeAvKkK2trEbT3KIm6ZdkgXb?usp=sharing)

### **Análisis de variables**

Los datos que elegimos para el análisis, son sobre la inversión extranjera directa por sector de actividad económica, subsector y rama, período del primer trimestre de 2005 al primer trimestre de 2019. Presenta cifras expresadas en millones de dólares. Serie histórica de 16 años.

<b>Variable</b>	<b>Tipo</b>	<b>Rango</b>
<b>Año</b>	Entero	2005 - 2019
<b>Trimestre</b>	String	Primero - Cuarto
<b>Agricultura, Ganadería, Silvicultura, Caza y Pesca</b>	Entero	237,000 - 366,000
<b>Industria Extractiva, Electricidad</b>	Entero	7,000 - 21,000
<b>Industria manufacturera</b>	Entero	588,00 - 689,000
<b>Construcción</b>	Entero	220,000 - 316,000
<b>Comercio</b>	Entero	587,000 - 799,000
<b>Restaurantes y alojamiento</b>	Entero	181,000 - 392,000
<b>Transportes, comunicaciones, correo y almacenamiento</b>	Entero	105,000 - 175,000
<b>Servicios profesionales, financieros y corporativos</b>	Entero	141,000 - 320,000
<b>Servicios sociales</b>	Entero	205,000 - 311,000
<b>Servicios diversos</b>	Entero	302,000 - 451,000
<b>Gobierno y organismos internacionales</b>	Entero	104,000 - 125,000
<b>No especificado</b>	Entero	6,000 - 25,000

## Conclusiones de los datos

Para analizar la inversión extranjera directa (IED) a lo largo de los años, se pueden calcular la media, mediana y desviación estándar de la inversión. La media representa el valor promedio de la IED, la mediana muestra el valor central de la inversión y la desviación estándar mide la dispersión de la inversión a lo largo de los años. Estas estadísticas pueden ayudar a identificar la tendencia general de la inversión a lo largo de los años.

Para cada sector, se pueden calcular la media, mediana y desviación estándar de la inversión. La media indica el valor promedio de la inversión en el sector a lo largo de los años, la mediana puede mostrar si hay valores atípicos en la inversión en el sector y la desviación estándar mide cuánto varía la inversión de año en año en el sector. Estas estadísticas pueden ayudar a determinar si la inversión en el sector ha sido consistente o ha experimentado fluctuaciones significativas.

MEDIA:

```
#estadisticas
media = dt.mean()
print(media)
```

Año	2011.631579
Agricultura, Ganadería, Silvicultura, Caza y Pesca	306105.561404
Industria Extractiva, Electricidad	12224.456140
Industria manufacturera	612987.210526
Construcción	258467.543860
Comercio	701738.964912
Restaurantes y alojamiento	259692.368421
Transportes, comunicaciones, correo y almacenamiento	131751.052632
Servicios profesionales, financieros y corporativos	232135.105263
Servicios sociales	251327.122807
Servicios diversos	367316.824561
Gobierno y organismos internacionales	117399.543860
No especificado	13852.473684

dtype: float64  
<ipython-input-11-2b481d23818f>:2: FutureWarning: The default value of numer  
media = dt.mean()

Como se explicó anteriormente la media representa el valor promedio de la IED, por lo que de esta información se puede analizar y llegar la siguiente conclusión en base a la media hablando puramente de ingresos como datos más relevantes:

**Más alta producción económica:** El sector de Comercio es el que tiene la producción económica más alta, con 713,142 unidades.

**Más baja producción económica:** El sector no especificado tiene la producción económica más baja, con 12,889 unidades, se puede deber a diversos factores de inversión, pero sin información extra se toma como ingresos que no se pueden clasificar por lo que terminan siendo la menor fuente de ingresos.

**Sectores con alta producción económica:** Los sectores con una producción económica a considerar son Servicios diversos (370,538 unidades), Agricultura, Ganadería, Silvicultura, Caza y Pesca (307,985 unidades) y Servicios profesionales, financieros y corporativos (235,673 unidades).

**Sectores con producción económica más baja:** Sin tomar en cuenta el “no especificado”, la Industria Extractiva, Electricidad (12,045 unidades) y Gobierno y organismos internacionales (117,243 unidades) también tienen una producción económica relativamente baja en comparación con otros sectores.

MEDIANA:

```
1 mediana = dt.median()
2 print(mediana)
```

```
Año                2012.0
Agricultura, Ganadería, Silvicultura, Caza y Pesca  307985.0
Industria Extractiva, Electricidad  12045.0
Industria manufacturera  604364.0
Construcción  256376.0
Comercio  713142.0
Restaurantes y alojamiento  255826.0
Transportes, comunicaciones, correo y almacenamiento  128793.0
Servicios profesionales, financieros y corporativos  235673.0
Servicios sociales  248205.0
Servicios diversos  370538.0
Gobierno y organismos internacionales  117243.0
No especificado  12889.0
dtype: float64
<ipython-input-12-58855bfd55dd>:1: FutureWarning: The default value of numeric_only in
mediana = dt.median()
```

A partir de los datos obtenidos se puede determinar que la mediana no tiene relevancia en el análisis de esta base de datos debido a que los datos corresponden a la información recolectada durante distintos años.

#### DESVIACIÓN:

```
1 desviacion = dt.std()
2 print(desviacion)
```

```
Año                4.151728
Agricultura, Ganadería, Silvicultura, Caza y Pesca  27208.877127
Industria Extractiva, Electricidad  3362.255700
Industria manufacturera  43526.369520
Construcción  26712.068479
Comercio  49599.149009
Restaurantes y alojamiento  44750.399604
Transportes, comunicaciones, correo y almacenamiento  17816.135220
Servicios profesionales, financieros y corporativos  51188.326529
Servicios sociales  30953.747353
Servicios diversos  31671.676042
Gobierno y organismos internacionales  8382.351925
No especificado  4871.590145
dtype: float64
<ipython-input-13-9d875b4d858a>:1: FutureWarning: The default value of n
desviacion = dt.std()
```

La desviación estándar es utilizada para la cuantificación de la dispersión de los datos por cada uno de los sectores que se vieron influenciados por la IED a lo largo del tiempo establecido.