

# Informe preprocesamiento de datos de Becas Crema

**Eric Bellet. Cédula: 24.463.483.**

## **Tidy Data:**

Tidy Data hace que sea fácil de llevar a cabo el análisis de datos, ya que estos se encuentran ordenados. Un conjunto de datos se dice que es ordenado, siempre que cumplan las siguientes condiciones:

- Las observaciones están en filas
- Las variables son las columnas
- Están contenidas en un solo conjunto de datos.

Hay varias características de datos desordenados que se pueden observar en la práctica. Éstos son algunos de los patrones más comúnmente observados:

- Los encabezados de columna son los valores, no los nombres de variables.
- Múltiples variables se almacenan en una columna.
- Las variables se almacenan en ambas filas y columnas.
- Múltiples tipos de unidad experimental almacenados en la misma mesa.
- Un tipo de unidad experimental almacenada en múltiples tablas.

## **Análisis exploratorio de los datos de las características (EDA)**

El análisis exploratorio de datos se lleva a cabo después de la recolección y limpieza de datos, y antes de cualquier modelado y visualización/presentación de los resultados. Sin embargo, es parte de un proceso iterativo. Después de hacer algunas EDA podemos tratar de construir algunos modelos o presentar algunas visualizaciones. Al mismo tiempo, en base a los resultados de la tarde podemos realizar algunos más EDA y así sucesivamente. Es todas las pistas para encontrar acerca rápidamente y no tanto sobre los detalles o la estética. Entre los objetivos principales de este tipo de análisis son, por supuesto, conocer a nuestros datos, sus tendencias y su calidad, y también para comprobar o incluso empezar a formular nuestra hipótesis.

# Explicación de la decisiones tomadas para realizar la vista minable y análisis exploratorio de los datos:

La primera decisión fue cambiar el nombre de todas las columnas a nombres mas entendibles y que puedan ser utilizados sin problema en la herramienta Spyder con el lenguaje de programación Python. En segundo lugar, ordenar todas las columnas por su Cédula de identidad y estandarizar a un mismo formato y tipo de datos, para poder aplicar métodos estadísticos sobre estas. A continuación explicaré detalladamente que se realizó en cada columna junto con un análisis exploratorio de los datos:

1. Columna **ID** (En realidad no tiene nombre pero asumo que son los ID específicos y únicos de cada estudiante) e **Indique.el.Período.Académico.a.renovar**:

La columna ID fue transformada a valores numéricos y ordenada.

La columna **Indique.el.Período.Académico.a.renovar** contiene 2 variables, el semestre del año y el año en el cual estudiante quiere renovar su beca, por lo tanto no cumple con la definición de Tidy Data, y separe estos valores en 2 columnas: **SemestreAcademicoarenovar** y **AnoAcademicoarenovar**.

	ID	SemestreAcademicoarenovar	AnoAcademicoarenovar
count	190.000000	190.000000	190.000000
mean	99.700000	1.584211	2014.568421
std	58.152413	0.494160	0.496605
min	1.000000	1.000000	2014.000000
25%	49.250000	1.000000	2014.000000
50%	98.500000	2.000000	2015.000000
75%	148.750000	2.000000	2015.000000
max	200.000000	2.000000	2015.000000

Existen 190 estudiantes, donde más de la mitad de ellos desean renovar el período II-2015.

La columna ID fue eliminada ya que me parece innecesaria, ya se tiene un identificador único por estudiante que es la cedula.

En las columnas **SemestreAcademicoarenovar** y **AnoAcademicoarenovar** se uso imputación de datos utilizando la moda, donde en el caso de no tener el **AnoAcademicoarenovar** se imputa la moda de los años, y en el caso de no tener el semestre se imputa la moda de los **SemestreAcademicoarenovar**, donde si el año es 2014 se imputa el semestre en moda de ese año, igualmente para el 2015.

2. Columna **Cédula.de.Identidad** y **Fecha.de.Nacimiento..colocar.sólo.datos.numéricos**:

La columna **Cédula.de.Identidad** fue cambiada de nombre a **CedulaDeIdentidad** y sus tipos de datos a numéricos.

La columna **Fecha.de.Nacimiento..colocar.sólo.datos.numéricos** contiene 3 variables, el día, mes y el año de nacimiento del estudiante, por lo tanto no cumple con la definición de Tidy Data, y separe estos valores en 3 columnas: **DiadeNacimiento**, **MesdeNacimiento** y **AnodeNacimiento**.

	CedulaDeIdentidad	DiadeNacimiento	MesdeNacimiento	AnodeNacimiento \
count	190.000000	190.000000	190.000000	190.000000
mean	95.500000	16.268421	6.615789	1991.668421
std	54.992424	8.423107	3.360959	4.056552

min	1.000000	1.000000	1.000000	1951.000000
25%	48.250000	9.000000	4.000000	1991.000000
50%	95.500000	18.000000	7.000000	1992.000000
75%	142.750000	23.000000	9.000000	1994.000000
max	190.000000	31.000000	12.000000	1996.000000

Podemos observar que los estudiantes nacieron entre el año 1951 y 1996 , donde la mayoría de estos nacieron en el año 1991,1992, 1993 y 1994.

### 3. Columna **Edad**, **Estado.Civil** y **Sexo**:

La columna **Edad** fue transformada a valores numéricos.

La columna **Estado.Civil** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Soltero (a)': 0, 'Casado (a)': 1, 'Viudo (a)': 2, 'Unido (a)':3.

La columna **Sexo** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Femenino': 0, 'Masculino': 1.

	Edad	EstadoCivil	Sexo
count	190.000000	190.000000	190.000000
mean	22.626316	0.036842	0.142105
std	4.013514	0.279286	0.350081
min	18.000000	0.000000	0.000000
25%	21.000000	0.000000	0.000000
50%	22.000000	0.000000	0.000000
75%	24.000000	0.000000	0.000000
max	63.000000	3.000000	1.000000

```
mode(df['EstadoCivil'] )
ModeResult(mode=array([0]), count=array([186]))
```

```
mode(df['Sexo'])
ModeResult(mode=array([0]), count=array([163]))
```

Podemos observar que la edad de los estudiantes se encuentran entre 18 años y 63 años, donde la media es de 22 años, y la moda es soltero (a) y de sexo femenino. En total hay 183 estudiantes solteros (as), 163 mujeres.

La columna Edad fue eliminada, ya que tenemos la fecha de nacimiento y la edad se puede calcular, por lo tanto esta columna es redundante.

### 4. Columna **Escuela**, **Año.de.Ingreso.a.la.UCV** y **Modalidad.de.Ingreso.a.la.UCV**:

La columna **Escuela** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Bioanálisis': 0, 'Enfermería': 1.

La columna **Año.de.Ingreso.a.la.UCV** fue transformada a valores numéricos.

La columna **Modalidad.de.Ingreso.a.la.UCV** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Convenios Interinstitucionales (nacionales e internacionales)': 'Prueba Interna y/o propedéutico': 1, 'Convenios Internos (Deportistas, artistas, hijos empleados docente y obreros, Samuel Robinson)': 2, 'Asignado OPSU': 3.

	Escuela	AnodeIngresoalaUCV	ModalidaddeIngresoalaUCV \
count	190.000000	190.000000	190.000000
mean	0.321053	2010.294737	1.784211
std	0.468114	2.007255	0.965414
min	0.000000	2005.000000	0.000000
25%	0.000000	2009.000000	1.000000
50%	0.000000	2010.000000	1.000000
75%	1.000000	2012.000000	3.000000
max	1.000000	2014.000000	3.000000

```
mode(df['Escuela'])
ModeResult(mode=array([0]), count=array([129]))
```

```
mode(df['ModalidaddeIngresoalaUCV'])
ModeResult(mode=array([1]), count=array([110]))
```

Podemos observar la moda de estudiantes son de la escuela de Bioanálisis (129 estudiantes). La media de año de ingreso fue en el 2010, y la moda de modalidad de ingreso a la UCV es por Prueba Interna y/o propedéutico, donde 110 estudiantes ingresaron mediante esta modalidad.

#### 5. Columna **Semestre.que.cursa** y **Ha.cambiado.usted.de.dirección.**:

La columna **Semestre.que.cursa** fue transformada a valores numéricos.

La columna **Ha.cambiado.usted.de.dirección.** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'No': 0, 'Si': 1.

	Semestrequecursa	Hacambiadousteddedireccion
count	190.000000	190.000000
mean	6.236842	0.047368
std	2.223423	0.212987
min	1.000000	0.000000
25%	5.000000	0.000000
50%	6.000000	0.000000
75%	8.000000	0.000000
max	10.000000	1.000000

```
mode(df['Hacambiadousteddedireccion'])
ModeResult(mode=array([0]), count=array([181]))
```

Podemos observar que la media de estudiantes cursan el 6to semestre, y que la moda es que no han cambiado de dirección (181 estudiantes).

#### 6. Columna **De.ser.afirmativo..indique.el.motivo:**

La columna **De.ser.afirmativo..indique.el.motivo** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'nan': 0, 'Situacion de salud de un familiar': 1, 'No me llegan correos de ustedes': 3, 'mudanza': 2, 'olvido de clave': 4.

```
Deserafirmativoindiqueelmotivo \
count      190.000000
mean       0.236842
std        1.204801
min        0.000000
25%        0.000000
50%        0.000000
75%        0.000000
max        9.000000
```

```
mode(respondieron['Deserafirmativoindiqueelmotivo'])
ModeResult(mode=array([2]), count=array([5]))
```

Podemos observar que la moda de las personas que cambiaron de dirección es por mudanza.

Esta columna la unifique con **Ha.cambiado.usted.de.dirección**, ya que no hace falta tener 2 columnas, en una sola se puede tener toda la información.

#### 7. Columna **Número.de.materias.inscritas.en.el.semestre.o.año.anterior:**

La columna **Número.de.materias.inscritas.en.el.semestre.o.año.anterior** fue transformada a valores numéricos.

```
Numerodemateriasinscritasenelsemestreoanoanterior \
count      190.000000
mean       3.889474
std        1.826725
min        0.000000
25%        3.000000
50%        4.000000
75%        5.000000
max        15.000000
```

Podemos observar que la media de materias inscritas por los estudiantes es entre 3 y 4 materias. Esta columna fue eliminada porque es redundante, ya que la suma de las materias aprobadas, reprobadas y aprobadas, es igual a los valores de esta columna.

#### 8. Columna **Número.de.materias.aprobadas.en.el.semestre.o.año.anterior:**

La columna **Número.de.materias.aprobadas.en.el.semestre.o.año.anterior** fue transformada a valores numéricos.

```
Numerodemateriaaprobadasenelsemestreoanoanterior \
count      190.000000
mean       3.310526
std        1.537329
min        0.000000
25%        2.000000
```

50%	4.000000
75%	4.000000
max	10.000000

Podemos observar que la media de los estudiantes es aprobar 3 materias.

#### 9. Columna **Número.de.materias.retiradas.en.el.semestre.o.año.anterior:**

La columna **Número.de.materias.retiradas.en.el.semestre.o.año.anterior** fue transformada a valores numéricos.

Numerodemateriasretiradasenelsemestreoanoanterior \	
count	190.000000
mean	0.215789
std	0.903118
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	10.000000

Podemos observar que la media de los estudiantes es no retirar materias.

#### 10. Columna **Número.de.materias.reprobadas.en.el.semestre.o.año.anterior:**

La columna **Número.de.materias.reprobadas.en.el.semestre.o.año.anterior** fue transformada a valores numéricos.

Numerodemateriasreprobadasenelsemestreoanoanterior \	
count	190.000000
mean	0.363158
std	0.809801
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	5.000000

Podemos observar que la media de los estudiantes es reprobado 0 materias.

#### 11. Columna **Promedio.ponderado.aprobado** y **Eficiencia:**

Las columnas **Promedio.ponderado.aprobado** y **Eficiencia** fueron transformadas a valores flotantes.

Promedioponderadoaprobado Eficiencia \		
count	190.000000	190.000000
mean	13.580000	0.852737
std	1.759475	0.228496
min	0.600000	0.100000

25%	12.670000	0.800000
50%	13.580000	0.960000
75%	14.587500	1.000000
max	17.220000	1.000000

Podemos observar que la media de promedio de los estudiantes es 13,5 y la eficiencia 0,85.

## 12. Columna **Si.reprobó.una.o.más.materias.indique.el.motivo:**

Las columnas **Si.reprobó.una.o.más.materias.indique.el.motivo** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: (1, 'Elevada carga horaria'), (2, 'Falta de tiempo para asistir a clases'), (3, 'Falta de tiempo para estudiar, Elevada carga horaria, Mala distribuci\xc3\xbn del tiempo'), (4, 'Falta de tiempo para estudiar, Mala distribuci\xc3\xbn del tiempo, Lejan\xc3\xada entre la residencia y la universidad'), (5, 'Lejan\xc3\xada entre la residencia y la universidad'), (6, 'Mala distribuci\xc3\xbn del tiempo'), (7, 'Mala distribuci\xc3\xbn del tiempo, Lejan\xc3\xada entre la residencia y la universidad'), (8, 'Mala distribuci\xc3\xbn del tiempo, me costaron esas materias'), (9, 'Problemas con el/la docente, Lejan\xc3\xada entre la residencia y la universidad'), (10, 'Problemas de adaptaci\xc3\xbn de la UCV'), (11, 'Problemas de salud'), (12, 'Problemas de salud, Deficientes h\xc3\xa1bitos de estudios, Elevada carga horaria'), (13, 'Problemas de salud, enfermedad'), (14, 'Problemas econ\xc3\xbmicos'), (15, 'Problemas econ\xc3\xbmicos, Problemas de salud'), (16, 'Problemas familiares'), (17, 'Problemas familiares, Deficiente formaci\xc3\xbn del bachillerato'), (18, 'Problemas familiares, Mala distribuci\xc3\xbn del tiempo'), (19, 'Problemas familiares, Problemas econ\xc3\xbmicos, Falta de tiempo para estudiar, Deficientes h\xc3\xa1bitos de estudios'), (20, 'Problemas familiares, Problemas econ\xc3\xbmicos, Falta de tiempo para estudiar, Elevada carga horaria, Mala distribuci\xc3\xbn del tiempo, Lejan\xc3\xada entre la residencia y la universidad'), (21, 'Problemas familiares, Problemas econ\xc3\xbmicos, Problemas de salud, Problemas vocacionales, Falta de tiempo para asistir a clases, Falta de tiempo para estudiar, Deficientes h\xc3\xa1bitos de estudios'), (22, 'Problemas personales'), (23, 'Problemas personales, Deficientes h\xc3\xa1bitos de estudios'), (24, 'Problemas personales, Deficientes h\xc3\xa1bitos de estudios, Deficiente formaci\xc3\xbn del bachillerato, Mala distribuci\xc3\xbn del tiempo'), (25, 'Problemas personales, Falta de tiempo para estudiar'), (26, 'Problemas personales, Falta de tiempo para estudiar, Elevada carga horaria, Mala distribuci\xc3\xbn del tiempo'), (27, 'Problemas personales, Mala distribuci\xc3\xbn del tiempo'), (28, 'Problemas personales, Problemas con el/la docente'), (29, 'Problemas personales, Problemas econ\xc3\xbmicos'), (30, 'Problemas personales, Problemas econ\xc3\xbmicos, Problemas de salud, Mala distribuci\xc3\xbn del tiempo, Problemas con el/la docente'), (31, 'Problemas personales, Problemas familiares, Falta de tiempo para estudiar'), (32, 'Problemas personales, Problemas familiares, Problemas econ\xc3\xbmicos, Problemas de salud, Falta de tiempo para asistir a clases'), (33, 'Problemas vocacionales, Deficiente formaci\xc3\xbn del bachillerato, Mala distribuci\xc3\xbn del tiempo'), (34, 'no eh reprobado materias'). El valor 0 indica que no ha reprobado materias durante la carrera.

Sireprobounaomasmateriasindiqueelmotivo \	
count	190.000000
mean	3.536842
std	8.094430
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000

max 34.000000

```
mode(df['Sireprobounaomasmateriasindiqueelmotivo'])
ModeResult(mode=array([0]), count=array([148]))
```

Podemos observar que la moda por la cual los estudiantes reprobaron materias, fue por problemas económicos.

### 13. Columna **Número.de.materias.inscritas.en.el.semestre.en.curso:**

La columna **Número.de.materias.inscritas.en.el.semestre.en.curso** fue transformada a valores numéricos.

```
Numerodemateriasinscritasenelsemestreencurso \
count      190.000000
mean        3.510526
std         1.612581
min         1.000000
25%         2.000000
50%         4.000000
75%         5.000000
max         7.000000
```

Podemos observar que la media de materias inscritas es 3 y 4 materias,

### 14. Columna **X.Estás.realizado.tesis...trabajo.de.grado.o.pasantías.de.grado.:**

La columna **X.Estás.realizado.tesis...trabajo.de.grado.o.pasantías.de.grado.** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'No': 0, 'Si': 1.

```
Seencuentrarealizandotesisopasantiasdegrado \
count      190.000000
mean        0.252632
std         0.435669
min         0.000000
25%         0.000000
50%         0.000000
75%         0.750000
max         1.000000
```

### 15. Columna **Tesis...trabajo.de.grado.o.pasantías.de.grado:**

La columna **Tesis...trabajo.de.grado.o.pasantías.de.grado** fue transformada a valores numéricos.

```
Cantidaddevecesqueharealizadotesisopasantiasdegrado \
count      190.000000
mean        0.368421
std         0.735389
```



min	0.000000
25%	0.000000
50%	0.000000
75%	1.000000
max	3.000000

Donde podemos observar que la media de los estudiantes no ha realizado ninguna de estas actividades. Esta columna fue unificada con **X.Estás.realizado.tesis...trabajo.de.grado.o.pasantías.de.grado.**, ya que toda esta información puede estar en una sola columna.

#### 16. Columna **Procedencia y Lugar.donde.reside.mientras.estudia.en.el.Universidad**:

La columna **Procedencia** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Municipio Sucre':0,'Guarenas - Guatire': 1, 'Municipio Libertador Caracas': 2, 'Aragua': 3,'Municipio Baruta': 4, 'Valles del Tuy': 5, 'Altos Mirandinos': 6,'Apure': 7, 'Municipio El Hatillo': 8, 'Municipio Chacao': 9,'Táchira': 10, 'Vargas':11, 'Monagas': 12, 'Portuguesa':13, 'Nueva Esparta': 14, 'Trujillo':15, 'Bolívar': 16, 'Barinas':17, 'Sucre': 18, 'Barlovento': 19, 'Anzoategui':20, 'Mérida': 21, 'Delta Amacuro': 22, 'Lara':23, 'Yaracuy': 24, 'Guárico': 25.

La columna **Lugar.donde.reside.mientras.estudia.en.el.Universidad** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Municipio Sucre':0,'Guarenas - Guatire': 1, 'Municipio Libertador Caracas': 2, 'Aragua': 3,'Municipio Baruta': 4, 'Valles del Tuy': 5, 'Altos Mirandinos': 6,'Apure': 7, 'Municipio El Hatillo': 8, 'Municipio Chacao': 9,'Táchira': 10, 'Vargas':11, 'Monagas': 12, 'Portuguesa':13, 'Nueva Esparta': 14, 'Trujillo':15, 'Bolívar': 16, 'Barinas':17, 'Sucre': 18, 'Barlovento': 19, 'Anzoategui':20, 'Mérida': 21, 'Delta Amacuro': 22, 'Lara':23, 'Yaracuy': 24, 'Guárico': 25.

En esta columna fueron imputados en los datos vacíos, la Procedencia, ya que es probable que el estudiante resida allí mientras estudia.

	Procedencia	LugardonderesidemientrasestudiaenlaUniversidad \
count	190.000000	190.000000
mean	4.926316	2.926316
std	5.809574	2.804917
min	0.000000	0.000000
25%	2.000000	2.000000
50%	2.000000	2.000000
75%	6.000000	4.000000
max	25.000000	19.000000

```
mode(df['Procedencia'])
ModeResult(mode=array([2]), count=array([73]))
```

```
mode(df['LugardonderesidemientrasestudiaenlaUniversidad'])
ModeResult(mode=array([2]), count=array([90]))
```

Podemos observar que la moda, es que los estudiantes proceden del Municipio Libertador Caracas y el lugar donde residen mientras estudian en la universidad es el Municipio Libertador Caracas.

### 17. Columna **Personas.con.las.cuales.usted.vive..mientras.estudia.en.la.universidad.:**

La columna **Personas.con.las.cuales.usted.vive..mientras.estudia.en.la.universidad.** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: residencia estudiantil':0, 'residencia':0, 'residencia':0,'Residencia':0,'Esposo (a) Hijos (as) ': 1,'Esposo (a) Hijos (as)': 1, 'Familiares paternos': 2, 'Madre': 3,'Familiares maternos': 4, 'Ambos padres': 5, 'Mi Mamá y mi hijo ': 6,'Padre': 7, 'Amigos': 8, 'madre y su esposo,abuela,y mi esposo': 9,'hermana': 10, 'Hermana': 10, 'dos hermanos':11, 'OTROS INQUILINOS': 12, 'hermanas':13, 'madre y hermana': 14, 'madre y hermanos':15, 'Madre y Hermanos':15, 'prima': 16, 'madrina':17, 'sola': 18, 'Mamá y Abuela': 19, 'madre,hermano e hijo':20, 'Madre, Hermano y Sobrina': 21, 'hermano, hermana y mi hijo': 22, 'compañeros de habitacion alquilada':23, 'Padres, hermana y abuelos maternos': 24, 'Madre, Hermana, Abuela': 25, 'abuela': 26, 'Dueños del apartamento donde alquilo la habitacion': 27, 'ambos padres y dos hermanis':28, 'Solo': 29, 'hermano':30, 'dueña del apartamento': 31, 'Madre y hermano': 32.

```
Personasconlascualesustedvivemientrasestudiaenlauniversidad \
count                190.000000
mean                 6.284211
std                  6.365687
min                  0.000000
25%                  3.000000
50%                  5.000000
75%                  5.000000
max                  32.000000
```

```
mode(df['Personasconlascualesustedvivemientrasestudiaenlauniversidad'])
ModeResult(mode=array([3]), count=array([60]))
```

Podemos observar que la moda es que los estudiantes viven con su madre mientras estudian en la universidad.

### 18. Columna **Tipo.de.vivienda.donde.reside.mientras.estudia.en.la.universidad:**

La columna **Tipo.de.vivienda.donde.reside.mientras.estudia.en.la.universidad** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Quinta o casa quinta':0, 'Apartamento en edificio': 1, 'Casa en barrio urbano': 2, 'Habitación alquilada': 3,'Casa de vecindad': 4, 'Residencia estudiantil': 5, 'Apartamento en quinta - casa quinta o casa': 6,'conserjería ': 7, 'Casa en barrio rural': 8, 'Casa de vecindad': 9,'casa': 10.

```
Tipodeviviendadonderesidemientrasestudiaenlauniversidad \
count                190.000000
mean                 2.205263
std                  2.183385
min                  0.000000
25%                  1.000000
50%                  1.000000
75%                  2.000000
max                  10.000000
```

```
mode(df["Tipodeviviendadonderesidemientrasestudiaenlauniversidad"])
ModeResult(mode=array([1]), count=array([85]))
```

Podemos observar que la moda es que los estudiantes vivan en apartamentos en edificios.

#### 19. Columna

**En.caso.de.vivir.en.habitación.alquilada.o.residencia.estudiantil..indique.el.monto.mensual:**

```
Encasodevivirenhabitacionalquiladaoresidenciaestudiantilindiqueelmontomensual \
count      190.000000
mean       384.973684
std        1164.331279
min         0.000000
25%         0.000000
50%         0.000000
75%         0.000000
max        6500.000000
```

Solamente los que viven:

```
count    23.000000
mean    3180.217391
std    1533.905057
min     125.000000
25%    1950.000000
50%    3000.000000
75%    4210.000000
max    6500.000000
3180.217391304348
```

Podemos observar que la media de los que respondieron, es de 3180,21. En esta columna los valores en cero significan que esos estudiantes no viven en habitación alquilada o en residencia estudiantil,

#### 20. Columna **Dirección.donde.se.encuentra.ubicada.la.residencia.o.habitación.alquilada:**

**La columna Dirección.donde.se.encuentra.ubicada.la.residencia.o.habitación.alquilada** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 1, 'Plaza La concordia. Parroquia Santa Teresa. Centro'), (2, '(Vivienda propia), urbanizaci\xc3\xb3n Ra\xc3\xbal Leoni Casalta 3, residencias Salto \xc3\x81ngel bloque 8 piso 11, apartamento 06 - Propatria'), (3, '23 de Enero'), (4, 'ANTIMANO BARIO NUEVO CASA52'), (5, 'AV. LEONARDO RUIZ PINEDA RES. HORNOS DE CAL TORRE C PISO 5 ,APTO. 58C, SAN AGUSTIN DEL SUR.'), (6, 'AVENIDA SANZ. CALLE SECTOR CONVENTO 1. EL MARQUES - LA CALIFORNIA NORTE. EDIFICIO CONTINENTAL. PISO 10. APTO 103. '), (7, 'Altagracia norte 4, esq. De salas.'), (8, 'Av Sanz calle convento II El Marques Edo. Miranda'), (9, 'Av. Alma Mater Qta, Carmen Los Chaguaramos'), (10, 'Av. Alma Mater Qta. Carmen Los Chaguaramos '), (11, 'Av. Alma Mater Qta. Carmen Los chaguaramos'), (12, 'Av. Almamater. Edificio Romeo. Piso 6. Apartamento 136'), (13, 'Av. Guzm\xc3\xa1n Blanco, cota 905. Res. La Veguita. Edif Los Monjes. Piso 9 '), (14, 'Av. Principal San Souci con Av. Francisco Solano, Municipio Cacao, Edf. Mijao, PB, Apart. #3.'), (15, 'Avenida Luis Roche de Altamira, municipio Chacao, edf. Sassola, piso 2 apto. 23. '), (16, 'CALLE EL DESVIO, DIAGONAL A LA AV. NUEVA GRANADA. URB. LA BANDERA. PARROQUIA SANTA

ROSALIA. MUNICIPIO LIBERTADOR. '), (17, 'Calle la providencia Edf Mary Piso 2 Apto 4. El cementerio'), (18, 'Caracas. Municipio Libertador. Jardines del Valle'), (19, 'Distrito Libertador. Municipio Libertador URB. La campi\xc3\xbb1a Calle los mangos. Residencia Villa Magna piso 02'), (20, 'EL PARAISO '), (21, 'Edificio Doralta. Avenida Las Acacias.Caracas. '), (22, 'Entre cuarta y quinta transversal monte cristo los ruices '), (23, 'Esquina de Amadores. Edif. Jose Gregorio Apart 23 piso 2 parroquia altagracia- caracas'), (24, 'Final Pasaje 11 Calle Carbonell Casa N 10 El Amparo Catia '), (25, 'Guarenas, Edo. Miranda Urb. Oropeza Castillo, zona 1, vereda 14, casa1'), (26, 'LOS ROSALES. AV EL PASEO. QUINTA VILLA COROMOTO'), (27, 'Municipio Libertador, Parroquia San Pedro. Urb. Valle Abajo, Residencias Valle Abajo, Torre F Piso 4, Apto 4-1. '), (28, 'Palo verde. Calle Jose Feliz Ribas. Residencias Jose Gregorio Hernandez. Torre A. Piso 3. Apartamento 34'), (29, 'Plaza Venezuela'), (30, 'Sta. Eduvigis, Municipio Sucre, Caracas. '), (31, 'Urb. La Rosaleda Sur, Edf. Atabapo'), (32, 'Urbanizaci\xc3\xbb3n Cumbres de Curumo, Avenida Principal, Edificio Araucaria'), (33, 'av principal francisco de miranda, boleita sur, edificio sarina piso 4 apartamento 4-1. '), (34, 'calle miranda, chacao'), (35, 'calle soublette casa numero 19 sector borde sequia cerca del jardin numero 1 guarenas pueblo '), (36, 'caracas san agustin del norte esquina arismendi'), (37, 'colinas de Bello Monte av-Caurimare edf-san antonio de padua piso uno apartamento uno'), (38, 'el cementerio Av. Soosevelt, sector los jabillos, , quinta floriver'), (39, 'el junquito, km 12, urbanizacion luis hurtado, calle el campesino, casa sin numero'), (40, 'los Teques, municipio Guaicaipuro, urbanizaci\xc3\xbb3n El Paso bloque 2 edificio 2 piso 3 apartamento 04'), (41, 'los valles del tuy. Edif virgen del valle. sabana de la cruz. apartamento 3'), (42, 'parroquia petare,urbanizacion el llanito, edf.orinoco, piso 1 apto 2'), (43, 'urb horizonte, av petare, qta merpe, el marques'), (44, 'urbanizacion el marques av sanz edificio bremen piso 4 apartamento 44'). En esta columna los valores en cero significan que esos estudiantes no viven en habitación alquilada o en residencia estudiantil.

Direcciondondeseencuentraubicadalar residenciao habitacional alquilada \

count	190.000000
mean	5.336842
std	11.393501
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	44.000000

## 21. Columna **X.Contrajo.Matrimonio.:**

La columna **X.Contrajo.Matrimonio.** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'No': 0, 'Si': 1.

Contrajomatrimonio \

count	190.000000
mean	0.005263
std	0.072548
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1.000000

```
mode(df['Contrajomatrimonio'])
ModeResult(mode=array(['0'], dtype=object), count=array([189]))
```

Podemos observar que la moda es que los estudiantes no se han casado.  
Esta columna fue eliminada, ya que la columna EstadoCivil indica lo mismo con mas detalles.

## 22. Columna **X.Ha.solicitado.algún.otro.beneficio.a.la.Universidad.u.otra.Institución.:**

La columna **X.Ha.solicitado.algún.otro.beneficio.a.la.Universidad.u.otra.Institución** fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'No': 0, 'Si': 1.

```
HasolicitadoalgunotrobeneficioalaUniversidaduoInstitucion \
count          190.000000
mean           0.068421
std            0.253134
min            0.000000
25%            0.000000
50%            0.000000
75%            0.000000
max            1.000000
```

```
mode(df['HasolicitadoalgunotrobeneficioalaUniversidaduoInstitucion'] )
ModeResult(mode=array([0]), count=array([177]))
```

Podemos observar que la moda es que los estudiantes no hayan solicitado beneficios a otras instituciones.

## 23. Columna **En.caso.afirmativo.señale..año.de.la.solicitud..institución.y.motivo**

La columna fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: '2012, OBE. AYUDA ECONÓMICA. SOLVENTAR GASTOS DE ESTUDIOS Y TRANSPORTE.':13, 'OBE 2011 UCV . Motivo: ayuda económica para mis estudios ': 1, 'AÑO 2011 AYUDA ECONÓMICA POR OBE PARA CUBRIR GASTOS DE MATERIAL DE ESTUDIO.': 2, 'Beca ayudantía en el año 2011 Universidad Central de Venezuela (OBE) apoyo económico para material de estudio': 3,'solicitud en 2013. \nayuda economica \n': 4, 'Año solicitud 2013\nInstitución: Universidad Central de Venezuela (OBE)\nMotivo: Transporte, Materiales Estudios, Comida': 5, 'fames ayuda para maternidad': 6,'2015, ucv, falta de ingresos, ayuda economica': 7, 'Año: 2013. Institucion: OBE. Motivo: ayuda económica para rehabilitación (fisioterapia), ya que me atropello una moto.': 8, 'Año:2014 \nInstitucion: OBE\nMotivo: Compra de materiales de estudio. ': 9,'fecha:2014\ninstituto:OBE\ncompras de material de estudio': 10,'solicitud 2015 en el instituto OBE , motivo ayuda economica para gastos de la universidad (libro, pasaje, equipo de enfermeria, etc)': 11,'año 2014, OBE, solicitada para cubrir parte de los gastos mensuales estudiantiles': 12.

```
Encasoaafirmativosenaleelanodelasolicitudinstitucionymotivo \
count          190.000000
mean           0.478947
```

std	2.025519
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	13.000000

```
mode(respondieron['Encasoafirmativosenaleelanodelasolicitudinstitucionymotivo'])
ModeResult(mode=array([1]), count=array([1]))
```

Podemos observar que la moda por la cual los estudiantes solicitaron la ayuda y la institución, fue por OBE 2011 UCV . Motivo: ayuda económica para mis estudios. Esta columna fue unificada con **X.Ha.solicitado.algún.otro.beneficio.a.la.Universidad.u.otra.Institución.** Ya que ambas informaciones pueden estar en una sola columna.

#### 24. Columna **X.Se.encuentra.usted..realizando.alguna.actividad.que.le.genere.ingresos.**

La siguiente columna fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'No': 0, 'Si': 1.

Seencuentraustedrealizandoalgunaactividadquelegenereingresos \	
count	190.000000
mean	0.021053
std	0.143939
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	1.000000

```
mode(df['Seencuentraustedrealizandoalgunaactividadquelegenereingresos'])
ModeResult(mode=array([0]), count=array([186]))
```

Podemos observar que la moda es que los estudiantes no se encuentran realizando actividades que le generen ingresos.

#### 25. Columna **En.caso.de.ser.afirmativo..indique.tipo.de.actividad.y.su.frecuencia**

La siguiente columna fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'cuido pacientes parapoder obtener un ingreso economico': 1, 'STAFF DE EMPRESA DEPORTAIVA. UNO O DOS FINES DE SEMANA AL MES APROXIMADAMENTE.': 2, 'Recreación, algunos fines de semana (poco frecuente por falta de tiempo).': 3, 'Secretaria, sólo los sábados medio turno.': 4.

Encasodeserafirmativoindiquetipodeactividadysufrecuencia \

count	190.000000
mean	0.052632
std	0.394899
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	4.000000

mode(respondieron['Encasodeserafirmativoindicquetipodeactividadysufrecuencia'])

ModeResult(mode=array([1]), count=array([1]))

Esta columna fue unificada con

**X.Se.encuentra.usted..realizando.alguna.actividad.que.le.genere.ingresos.** Ya que ambas informaciones pueden estar en una misma columna.

## 26. Columna **Monto.mensual.de.la.beca** y **Aporte.mensual.que.le.brinda.su.responsable.económico**

Las columnas fueron transformadas a valores numéricos.

	Montomensualdelabeca	Aportemensualquelebrindasuresponsableeconomico \
count	190.000000	190.000000
mean	1566.315789	1618.947368
std	987.861302	1414.871485
min	200.000000	0.000000
25%	1500.000000	600.000000
50%	1500.000000	1000.000000
75%	1500.000000	2000.000000
max	15000.000000	9000.000000

Podemos observar que la media del monto mensual de la beca es de 1566,32 y la media del aporte que le brinda mensualmente los responsables económicos a los estudiantes es de 1618,95.

## 27. Columna **Aporte.mensual.que.recibe.de.familiares.o.amigos**

La siguiente columna fue transformada a valores numéricos.

	Aportemensualquerecibedefamiliaresoamigos \
count	190.000000
mean	331.763158
std	718.473302
min	0.000000
25%	0.000000
50%	0.000000
75%	500.000000
max	5000.000000

Podemos observar que la media es de 331,76.

## 28. Columna **Ingreso.mensual.que.recibe.por.actividades.a.destajo.o.por.horas**

La siguiente columna fue transformada a valores numéricos.

	Ingresomensualquerecibeporactividadesadestajooporhoras \
count	190.000000
mean	107.894737
std	419.360548
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	3000.000000

## 29. Columna **Ingreso.mensual.total** y **Alimentación**

Las siguientes columnas fueron transformadas a valores numéricos.

	Ingresomensualtotal	Gastosenalimentacionpersonal \
count	190.000000	190.000000
mean	3624.921053	1232.368421
std	2042.712202	2319.777441
min	1500.000000	0.000000
25%	2500.000000	412.500000
50%	3000.000000	1000.000000
75%	4275.000000	1500.000000
max	17000.000000	30000.000000

La columna total de ingresos fue eliminada, ya que con la suma de todos los ingresos se puede obtener estos valores.

## 30. Columna **Transporte.público** y **Gastos.médicos**

Las siguientes columnas fueron transformadas a valores numéricos.

	Gastosentransportepersonal	Gastosmedicospersonal \
count	190.000000	190.000000
mean	608.084211	345.000000
std	588.287366	869.69176
min	0.000000	0.000000
25%	250.000000	0.000000
50%	500.000000	0.000000
75%	750.000000	500.000000
max	3500.000000	10000.000000



### 31. Columna **Gastos.odontológicos** y **Gastos.personales**

Las siguientes columnas fueron transformadas a valores numéricos.

	Gastosodontologicospersonal	Gastospersonales \
count	190.000000	190.000000
mean	194.736842	569.000000
std	443.529237	557.436071
min	0.000000	0.000000
25%	0.000000	200.000000
50%	0.000000	500.000000
75%	0.000000	800.000000
max	2500.000000	3000.000000

### 32. Columna **Residencia.o.habitación.alquilada**

La siguiente columna fue transformada a valores numéricos.

	Gastosenresidenciaohabitacionalquiladapersonal \
count	190.000000
mean	300.894737
std	1039.751568
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	6500.000000

### 33. Columna **Materiales.de.estudio**

La siguiente columna fue transformada a valores numéricos.

	GastosenMaterialesdeestudiopersonal
count	190.000000
mean	1040.894737
std	702.272004
min	100.000000
25%	500.000000
50%	1000.000000
75%	1500.000000
max	4000.000000

### 34. Columna **Recreación, Otros.gastos** y **Total.egresos**

Las siguientes columnas fueron transformadas a valores numéricos.

	Gastosenrecreacionpersonal	Otrosgastospersonal	Totalegresospersonal \
count	190.000000	190.000000	190.000000

mean	274.315789	326.500000	4891.794737
std	446.952414	861.794601	4431.362572
min	0.000000	0.000000	1300.000000
25%	0.000000	0.000000	2512.500000
50%	0.000000	200.000000	3500.000000
75%	500.000000	300.000000	6000.000000
max	2500.000000	10900.000000	47500.000000

La columna total de egresos fue eliminada, ya que con la suma de todos los egresos se puede obtener estos valores.

### 35. Columna **Indique.quién.es.su.responsable.económico** y **Carga.familiar**

La columna Indique.quién.es.su.responsable.económico fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número: 'Usted mismo':0,'esposo': 1, 'ninguno': 2, 'Madre': 3,'Familiares': 4, 'Padre': 5, 'Ambos padres': 6,'Cónyugue': 7, 'tia': 8, 'Hermano': 9,'hermana': 10,'MI HERMANA': 10, 'abuela': 10, 'dos hermanos':11, 'OTROS INQUILINOS': 12, 'hermanas':13, 'madre y hermana': 14, 'madre y hermanos':15, 'Madre y Hermanos':15, 'prima': 16, 'madrina':17, 'sola': 18, 'Mamá y Abuela': 19, 'madre,hermano e hijo':20, 'Madre, Hermano y Sobrina': 21, 'hermano, hermana y mi hijo': 22, 'compañeros de habitacion alquilada':23, 'Padres, hermana y abuelos maternos': 24, 'Madre, Hermana, Abuela': 25, 'abuela': 26, 'Dueños del apartamento donde alquilo la habitacion': 27, 'ambos padres y dos hermanis':28, 'Solo': 29, 'hermano':30, 'dueña del apartamento': 31, 'Madre y hermano': 32.

La columna Carga.familiar fue transformada a valores numéricos.

	Indiquequienessuresponsableeconomico	Cargafamiliar \
count	190.000000	190.000000
mean	4.121053	3.121053
std	2.662910	1.264526
min	0.000000	1.000000
25%	3.000000	2.000000
50%	3.000000	3.000000
75%	5.000000	4.000000
max	26.000000	8.000000

```
mode(df['Indiquequienessuresponsableeconomico'])
ModeResult(mode=array([3]), count=array([107]))
```

### 36. Columna **Ingreso.mensual.de.su.responsable.económico**, **Otros.ingresos** y **Total.de.ingresos**

Las siguientes columnas fueron transformadas a valores numéricos.

	Ingresomensualdesuresponsableeconomico	Otrosingresos	Totaldeingresos \
count	190.000000	190.000000	190.000000
mean	109635.059526	2250.997263	111886.056789
std	765604.138161	3273.834547	765568.465509
min	0.000000	0.000000	0.000000
25%	6275.000000	0.000000	7500.000000

50%	8450.000000	0.000000	10411.500000
75%	12347.735000	4000.000000	15787.250000
max	8773759.000000	17879.000000	8775759.000000

La columna total de ingresos fue eliminada, ya que con la suma de todos los ingresos se puede obtener estos valores.

### 37. Columna **Vivienda**

La siguiente columna fue transformada a valores numéricos.

Gastosenviviendadesusresponsableseconomicos \	
count	190.000000
mean	3161.957895
std	38617.901371
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	532498.000000

### 38. Columna **Alimentación.1**

La siguiente columna fue transformada a valores numéricos.

Gastosenalimentaciondesusresponsableseconomicos	
count	190.000000
mean	6113.068421
std	4402.714218
min	0.000000
25%	3500.000000
50%	5000.000000
75%	8000.000000
max	36000.000000

### 39. Columna **Transporte**

La siguiente columna fue transformada a valores numéricos.

Gastosentransportedesusresponsableseconomicos \	
count	190.000000
mean	1326.778947
std	1526.241052
min	0.000000
25%	500.000000
50%	1000.000000
75%	2000.000000
max	12960.000000

#### 40. Columna **Gastos.médicos.1**

La siguiente columna fue transformada a valores numéricos.

```
Gastosmedicosdesusresponsableseconomicos \
count          190.000000
mean          1364.497474
std           1787.681958
min            0.000000
25%            0.000000
50%          1000.000000
75%          2000.000000
max          12000.000000
```

#### 41. Columna **Gastos.odontológicos.**

La siguiente columna fue transformada a valores numéricos.

```
Gastosodontologicosdesusresponsableseconomicos \
count          190.000000
mean           357.473684
std            829.479302
min            0.000000
25%            0.000000
50%            0.000000
75%            0.000000
max           5000.000000
```

#### 42. Columna **Gastos.educativos**

La siguiente columna fue transformada a valores numéricos.

```
Gastoseducativosdesusresponsableseconomicos \
count          190.000000
mean           954.600000
std           1640.818394
min            0.000000
25%            0.000000
50%            0.000000
75%           1500.000000
max           10000.000000
```

#### 43. Columna **Servicios.públicos.de.agua..luz..teléfono.y.gas**

La siguiente columna fue transformada a valores numéricos.

```
Gastosenserviciospublicosdeagualuztelefonoygasdesusresponsableseconomicos \
count          190.000000
mean           770.700000
```

std	768.695686
min	0.000000
25%	300.000000
50%	600.000000
75%	1000.000000
max	5650.000000

#### 44. Columna **Condominio**

La siguiente columna fue transformada a valores numéricos.

```
Condominiodesusresponsableseconomicos \
```

count	190.000000
mean	420.686579
std	711.307975
min	0.000000
25%	0.000000
50%	0.000000
75%	500.000000
max	3500.000000

Podemos observar que la media es de 420.686579.

#### 45. Columna **Otros.gastos.1**

La siguiente columna fue transformada a valores numéricos.

```
Otrosgastosdesusresponsableseconomicos \
```

count	190.000000
mean	3432.277684
std	32199.561638
min	0.000000
25%	157.500000
50%	700.000000
75%	1500.000000
max	444482.000000

Podemos observar que la media es de 3432.277684.

#### 46. Columna **Total.de.egresos**

La siguiente columna fue transformada a valores numéricos.

```
Totaldeegresosdesusresponsableseconomicos \
```

count	190.000000
mean	17902.040684
std	52665.489274
min	0.000000
25%	7300.000000
50%	10700.000000
75%	16000.000000

max 581628.450000

Podemos observar que la media es de 17902.040684.

La columna total de egresos fue eliminada, ya que con la suma de todos los egresos se puede obtener estos valores.

#### 47. Columna

**Deseamos.conocer.la.opinión.de.nuestros.usuarios..para.mejorar.la.calidad.de.los.servicios.ofrecidos.por.el.Dpto..de.Trabajo.Social.OBE**

La siguiente columna fue transformada a valores numéricos.

DeseamosconocerlaopiniondenuestrosusuariosparamejorarlacalidaddelosserviciosofrecidosporelDptode TrabajoSocialOBE \

count	190.000000
mean	4.547368
std	0.701750
min	1.000000
25%	4.000000
50%	5.000000
75%	5.000000
max	5.000000

mode(df.Deseamosconocerlaopiniondenuestrosusuariosparamejorarlacalidaddelosserviciosofrecidospor elDptodeTrabajoSocialOBE)

ModeResult(mode=array([5], dtype=int64), count=array([121]))

Podemos ver que la moda de la opinion de los estudiantes esta evaluada con un 5.

#### 48. Columna **Sugerencias.y.recomendaciones.para.mejorar.nuestra.atención**

La siguiente columna fue categorizada y transformada a valores numéricos, donde cada uno corresponde al siguiente número:

(0, ' muy buena la atencion agradecida con la ayuda que mensualmente me brindan '), (1, '..'), (2, '...'), (3, 'Agradecida, declaro la ayuda econ\x3\xb3mica muy buena y provechosa. '), (4, 'Agradezco toda la ayuda prestada por parte del servicio de trabajo social y ser comprensivos con respecto a las situaciones de cada estudiante'), (5, 'Alargar el plazo de entrega de recaudos para la renovaci\x3\xb3n de la beca a 1 semana. Avisar con m\x3\xa1s tiempo de anticipaci\x3\xb3n la fecha de la renovaci\x3\xb3n de la beca para as\x3\xad recaudar los documentos necesarios con tiempo.'), (6, 'Algunas personas no poseen Internet, por ello en la cartelera de las Escuelas deber\x3\xadan colocar mas informaci\x3\xb3n de las actividades o alguna jornada que este realizando OBE'), (7, 'Buen Servicio'), (8, 'Buena'), (9, 'Buena atenci\x3\xb3n y proceso mas r\x3\xa1pido al estar en la escuela de bioanálisis para la recepci\x3\xb3n de documentos.'), (10, 'Buena atenci\x3\xb3n y puntualidad en los pagos... Gracias'), (11, 'Colocar un amplio horario de atenci\x3\xb3n, debido a que chocan con los horario acad\x3\xa9micos'), (12, 'Colocar una ofinica de obe en la escuela de enfermeria '), (13, 'Como recomendaci\x3\xb3n ofrecer becas a los estudiantes de licenciaturas muchos no tienen la posibilidad ni el tiempo de trabajar por horarios de estudios y trabajos en cl\x3\xadnicas, mucha veces se ve

forzado para otros no, lo ideal, es que se estudie cada caso seg\xfc\xban lo amerite el estudiante, en mi caso lo veo forzado salir de clases a las 12 m para ir corriendo a la cl\xfc\xadnica y salir a las 7 pm en que tiempo se descansa y se estudia o otros casos q trabajan de noche salir amanecidos para recibir clases por tal motivo no trabajo '), (14, 'Concidero que es apta ya que por ser la \xü\xbanica escuela que se encuentra lejos de la ciudad universitaria nos facilitan aceptar los datos requeridos en dicha escuela'), (15, 'Consideracion para los estudiantes , ya que somos una de las victimas mas afectadas de esta horrible inflaci\xü\xbn , los materiales de estudio se han vuelto cada vez mas costosos y pues estudiar es ya casi una mision dificil por no decir imposible'), (16, 'Considero que la atenci\xü\xbn es acorde, con un trato grato y satisfactorio. '), (17, 'Continuar ofreciendo la ayuda y orientacion a los estudiantes. '), (18, 'DEBE HABER MAS DIAS PARA LA RECEPCION DE LOS RECAUDOS, ASI COMO EL HORARIO PARA LA ENTREGA DE LOS RECAUDOS DEBE SER MAS LARGO'), (19, 'Dar mayor informacion cuando no se cancela el pago de la beca a tiempo y el por que...y los retroactivos!'), (20, 'Deber\xü\xadan estar mas tiempo en cada escuela, espec\xü\xadficamente las sedes que quedan retiradas de la oficina principal de la OBE. '), (21, 'Debido a lo apretado del horario de clases, muchas veces se nos hace dif\xü\xadcil reunir ciertos requisitos con prontitud o incluso entregarlos . Excelente ser\xü\xada una extensi\xü\xbn de los horarios de renovaci\xü\xbn '), (22, 'ES IMPORTANTE QUE EL HORARIO DE ATENCIO SEA MAS FLEXIBLE'), (23, 'ESTA MUY BIEN\n'), (24, 'ESTOY SATISFECHA CON SU ATENCION'), (25, 'EVALUAR LA POSIBILIDAD DE SUBIR EL MONTO DE LA BECA YA QUE 1500BSF NO SE AJUSTA A LA REALIDAD DE GASTOS QUE POSEE UN ESTUDIANTE'), (26, 'EXCELENTE '), (27, 'EXCELENTEATENCI\xü\x93N. SIGAN AS\xü\x8d!!'), (28, 'El servicio brindado a trav\xü\xa9s de la trabajadora social es bueno,con grata atenci\xü\xbn, responsabilidad y puntualidad. Los pagos se realizan siempre durante la fecha estipulada, aunque aun no se ha recibido el aumento becario. '), (29, 'El servicios es bueno, lo \xü\xbanico que tiene que mejor es que a todos le lleguen la informaci\xü\xbn de las renovaciones al correo. '), (30, 'En general la atenci\xü\xbn prestada es buena. '), (31, 'Enviar un mensaje un d\xü\xada para recordar si se tiene un cita con el trabajador social'), (32, 'Es mi primera renovaci\xü\xbn. La beca me a ayudado a cubrir parte de mis gastos. '), (33, 'Espero que los problemas en cuanto al aumento de la beca se solventen pronto. En lo dem\xü\xa1s no tengo quejas. '), (34, 'Establecer horarios m\xü\xa1s flexibles de atenci\xü\xbn al estudiante de Bioanálisis, al menos para la recepci\xü\xbn de las planillas. '), (35, 'Este a\xü\xbo no recibí el correo electronico para la renovacion como en los a\xü\xbos pasados, me entere por el instagram de mi centro de estudiantes , me parecia mas efectiva la via de e mail. '), (36, 'Estoy a gusto con su atenci\xü\xbn. '), (37, 'Estoy de acuerdo con la atenci\xü\xbn prestada'), (38, 'Excelente que exista un punto de asistencia para renovaci\xü\xbn en la escuela de enfermer\xü\xada. '), (39, 'Excelente atenci\xü\xbn'), (40, 'Excelente atenci\xü\xbn '), (41, 'Excelente atenci\xü\xbn y organizaci\xü\xbn. Hubo un poco de problema para cargar la p\xü\xa1g. al principio hasta lograrlo despu\xü\xa9s de varios intentos. '), (42, 'Excelente forma de renovar. '), (43, 'Excelente la idea de dirigirse a la escuela de Bioanálisis ya que por las clases nos queda poco tiempo para dirigirnos a la oficina y realizar los tr\xü\xa1mites. Gracias. '), (44, 'Excelente trabajo, agradecida por la ayuda aportada todo este tiempo. '), (45, 'Excelente trabajo, seria fantastico que presionaran a los organismos competentes para que llegue el dinero del aumento a la universidad '), (46, 'Expandir la hora para renovar las becas hasta las 3/4pm porque habemos muchos que no podemos ir en el horario entre las 8.30am y la 1.30pm debido a los laboratorios pr\xü\xa1cticos de la carrera, de los cuales no podemos salir, para dirigirnos a hacer la renovaci\xü\xbn. Si se expande la hora podemos ir a las 2pm cosa de no perder clases y cumplir con este requisito de renovaci\xü\xbn. '), (47, 'FLEXIBILIZAR LOS HORARIOS DE ATENCION PARA LOS QUE ESTAMOS DE PASANTIAS\n\nMANTENERNOS INFORMADOS'), (48, 'Felicidades por hacer esta parte del proceso v\xü\xada internet'), (49, 'Gracias por su apoyo, hacen un excelente trabajo, estoy muy agradecida'), (50, 'Hasta ahora me ha gustado la facilidad del proceso, espero se mantenga as\xü\xad. '), (51, 'Hasta ahora, todo excelente... Gente

comprometida a prestar un servicio de calidad, se les agradece. '), (52, 'Informar con anticipaci\u00f3n el proceso de renovaci\u00f3n, ya que la constancia de trabajo demora dos semanas en ser emitida por la secretaria del dador de trabajo. '), (53, 'La atenci\u00f3n me parece totalmente agradable y satisfactoria, por lo que no es necesario una mejora. '), (54, 'La atenci\u00f3n recibida fue bastante buena, como recomendaci\u00f3n que sigan trabajando por tener un sistema cada vez mas beneficioso. '), (55, 'La labor realizada por la trabajadora social de la escuela de Enfermer\u00eda es muy eficiente y responsable. sin embargo, se sugiere establecer una fecha fija para el dep\u00f3sito de la beca e incluir los pagos atrasados'), (56, 'La presente organizaci\u00f3n presta un buen servicio, la cual agradezco. En este momento no tengo alguna sugerencia o recomendaci\u00f3n que pueda aportar, en el momento de tener alguna se los har\u00e9 saber, muchas gracias por su atenci\u00f3n. '), (57, 'Las veces que me he dirigido hasta all\u00ed he recibido una atenci\u00f3n de calidad. '), (58, 'Luchar por un aumento de beca ya que la situaci\u00f3n econ\u00f3mica es realmente fuerte, aun m\u00e1s para aquellos que presentan bajos recursos monetarios al mes'), (59, 'MUY BUEN SERVICIO'), (60, 'Mandar correos con informaci\u00f3n acerca de los distintos servicios y beneficios ofrecidos por la universidad. Y sobre los cambios en la beca. '), (61, 'Mayor seguridad para la oficina de OBE '), (62, 'Me parece han realizado un excelente trabajo ya que proporcionan la informaci\u00f3n precisa y con antelaci\u00f3n. '), (63, 'Me parece muy eficiente este m\u00e9todo de internet y no es como otros q se caen o tienen errores. '), (64, 'Me parece prestan buena atenci\u00f3n al p\u00fablico '), (65, 'Me parece que el servicio es muy bueno no tengo ninguna recomendaci\u00f3n, me parece que la trabajadora actual es excelente. '), (66, 'Me parece que estan haciendo un buen trabajo, con excelente comunicaci\u00f3n hacia los estudiantes. '), (67, 'Me parece que la atenci\u00f3n es muy buena, fueron muy cordiales. '), (68, 'Mejor puntualidad'), (69, 'Mejorar el cumplimiento del horario de trabajo para la atenci\u00f3n de los estudiantes. '), (70, 'Muchas gracias por su labor. Contin\u00faen as\u00ed. '), (71, 'Muy buen servicio'), (72, 'Muy buen trabajo. Continuar con la buen desempe\u00f1o en el Servicio ofrecido a toda la comunidad estudiantil. '), (73, 'Muy buena atenci\u00f3n por parte de la OBE '), (74, 'Muy buena atenci\u00f3n por parte de nuestra trabajadora social. sin inconvenientes'), (75, 'Muy buena, la renovacion online ahorra mucho tiempo y hace el proceso de renovacion mas rapido y eficaz'), (76, 'Muy bueno el trato siempre'), (77, 'NINGUNA, COMPLETAMENTE SATISFECHO. '), (78, 'Nada que mejorar. '), (79, 'Ninguna'), (80, 'Ninguna Sugerencia o Recomendaci\u00f3n, sus atenciones han sido Excelentes '), (81, 'Ninguna esto nos ayuda a agilizar el proceso '), (82, 'No cambiar tanto a la trabajadora social. '), (83, 'No presento quejas al respecto, de verdad que la beca es una gran ayuda para los estudiantes y m\u00e1s ahora en la situaci\u00f3n en la que vivimos'), (84, 'No tengo ninguna sugerencia'), (85, 'No tengo ninguna sugerencia ni recomendacion, es un buen servicio para la situacion economica que se vive en el pais, estoy agradecida con mi beca y con el servicio que me presta la universidad'), (86, 'No tengo ninguna sugerencia y recomendaci\u00f3n '), (87, 'Promover e informar mejor los servicios que OBE presta. '), (88, 'Que envíen la fecha de renovaci\u00f3n de beca con anticipaci\u00f3n por correo'), (89, 'Que la trabajadora social envíe correos con la actualizaci\u00f3n de la beca'), (90, 'Que se contin\u00fae realizando la recepci\u00f3n de documentos en nuestra escuela ya que resulta mas f\u00c1cil por horario que podamos realizar el tramite de la renovaci\u00f3n. Tambi\u00e9n que la trabajadora social tenga un espacio en la Escuela de Enfermer\u00eda destinado a la atenci\u00f3n de becarios y de mas informaci\u00f3n referente a OBE. '), (91, 'Que sigan as\u00ed! '), (92, 'Realizar jornadas de las q ofrece OBE en la Escuela de Enfermer\u00eda'), (93, 'Realizar todos los procesos via internet para ahorrar tiempo y que sea mas comodo. '), (94, 'Realmente la atenci\u00f3n prestada es bastante satisfactoria, el personal es sumamente amable y cordial. \n\nLa \u00fanica sugerencia seria ver si se podr\u00eda hacer un comunicado para aumentar un poco m\u00e1s la beca estudio ya que hoy en d\u00eda todo se ha encarecido bastante. \n\nMuchas Gracias por la ayuda prestada. '), (95, 'Realmente, el proceso es muy satisfactorio y



mucho mas al agregar la deficiencia presupuestaria del pais. La calidad de servicio es de las mas alta y profesional. '), (96, 'Resguardar informaci3n importante en lugares seguros..'), (97, 'SER UN POCO M3S EFICIENTE'), (98, 'SIGAN TRABAJANDO EN EL PROCESO DE AUTOMATIZACI3N '), (99, 'SU ATENCI3N ES MUY BUENA'), (100, 'SU ATENCI3N ES MUY BUENA'), (101, 'Seguir atendiendo con calidad humana '), (102, 'Seguir con el buen servicio prestado. '), (103, 'Seguir con la misma t3cnica de trabajo que tienen hasta ahora. '), (104, 'Siempre muy atentos y con mucha consideraci3n con respecto a los problemas presentados. '), (105, 'Siempre son amables y responsables. Ojal3 pudieran tener mas apoyo para poder hacer mas!'), (106, 'Sigan as3, claro y entendible. '), (107, 'Sigan prestando su servicio como lo han hecho hasta ahora, ya que su trabajo ayuda a muchos estudiantes que lo necesitamos. '), (108, 'Sugiero que el monto de la beca llegue completo al momento de que el Presidente aumente el monto de la beca. '), (109, 'Tal vez extender un poco el horario de atencion. '), (110, 'Tener m3s amabilidad con los estudiantes'), (111, 'Todo bien'), (112, 'Todo me parece bien, c3modo y sencillo, bien explicado y resumido. '), (113, 'Tomar en cuenta que continuamos cursando estudios en la universidad y debido a la situaci3n del pa3s se ha hecho dif3cil conseguir un empleo en la rama de enfermer3da, que facilite el horario y colabora al crecimiento profesional educativo. '), (114, 'Tomar siempre en cuenta la situaci3n econ3mica del pa3s y el beneficio de esta beca estudio y la importancia que tiene para los beneficiarios puesto es una de las 3banicas entradas que gozan algunos, Gracias. '), (115, 'Un horario un poco m3s amplio de atenci3n ya que el horario de clases siempre choca'), (116, 'Yo particularmente nunca he tenido problemas con este sistema...siempre est3n atentos a cualquier inconveniente y siempre lo solventan lo mas r3pido posible. '), (117, 'agradecida, por las atenciones recibidas'), (118, 'alargar las fecgas para la entrega de documentos para la renovaci3n de becas, ya que muchos alumnos por el estudio no tiene tiempo de entregar los recaudos a tiempo'), (119, 'aumentar la beca. De resto, es un buen servicio. '), (120, 'aumente la becas por favor por que la situaci3n econ3mica esta grave no alcanza'), (121, 'colocar las secciones de la distribuci3n de gastos un poco m3s claras, por ejemplo a que se refieren con gastos personales y especificar si en el gasto familiar se incluya el gasto personal, ya que con el monto de la beca apor to a los gatos de alimentaci3n, y al momento de llenar esas secciones, estuve en confusi3n'), (122, 'colocar un lapso mas largo para la renovaci3n de beca (semana) '), (123, 'cumplir con el pago a tiempo. '), (124, 'dar mas d3adas para entregar los recaudos de renovacion. '), (125, 'de verdad que la he utilizado muy poco pero ,cuando la he necesitado he tenido una respuesta muy satisfactoria '), (126, 'deben ser preguntas mas claras y no pedir tantos monto, ya que con la situaci3n del pa3s se hace dif3cil discernir cuanto se gasta. '), (127, 'deber3dan colocar una oficina de OBE en la escuela de enfermer3da :)'), (128, 'efectiva'), (129, 'en lugar de tener un solo trabajador social para gestionar la beca, asignar un segundo para cubrir la gran demanda '), (130, 'es excelente la atenci3n recibida'), (131, 'es muy buena su atenci3n. han pagado los d3adas correspondientes. que era la 3banica queja que se tenia. '), (132, 'es un buen servicio'), (133, 'excelente'), (134, 'excelente atencion'), (135, 'excelente el servicio. '), (136, 'excelente servicio, se recomienda que sea mas automatico '), (137, 'excelente trato'), (138, 'excelente!'), (139, 'extender el horario de atencion a los estudiantes'), (140, 'extender el horario de atenci3n'), (141, 'extender por unos dias mas la renovacion de la beca ya que por la complejidad de nuestros horarios se nos hace dif3cil asistir'), (142, 'ha sido una excelente atenci3n!! '), (143, 'hasta los momentos todo marcha excelente'), (144, 'informar a los estudiantes acerca de la ubicaci3n de las oficinas, debido a que muchos no saben. '), (145, 'la atencion brindada es satisfactoria'), (146, 'la atencion es muy buena'), (147, 'la atencion siempre a sido excelente y tiempo. y lo mejor es que ahora la renovaci3n la podemos hacer en la escuela'), (148, 'la atenci3n brindada esta acorde con el tramite realizado, lo puedo catalogar como buena. '), (149, 'la gesti3n es mucho mas r3pida por este medio, no tengo inconvenientes de ning3ban tipo con el servicio de becas estudantil. '), (150, 'lo estan

haciendo muy bien. felicidades'), (151, 'manejar el sistema de msj de texto para los becarios de esta manera ayuda a estar mas informado'), (152, 'mayor flexibilidad en el horario de renovacion de beca.'), (153, 'me gustaria que la difusion de la de avisos tales como la renovacion de becas u alguna otra informacion sea con mayor tiempo de antelacion para que el estudiante tome las medidas necesarias'), (154, 'me parece muy buena su atenci\xc3\xbb3n prestada '), (155, 'me parece que realizan un trabajo excelente, muy eficientes, ya que son puntuales dan respuesta oportuna a nuestras dudas u interrogantes. \n\nexcelente que la Lic. Analeth se dirija a nuestra escuela, ya que en muchas cosas esta escuela no la toman mucho en cuenta.'), (156, 'mejorar el servicio en OBE'), (157, 'mejorar la comunicacion con los becarios, ya que me entere el ultimo dia que habia renovacion de la beca'), (158, 'mejorar la comunicaci\xc3\xbb2n de las fechas de renovaci\xc3\xbb2n.'), (159, 'mejorar un poco los horarios de atenci\xc3\xbb3n.'), (160, 'muy buen servicio'), (161, 'muy buen trabajo, graciaas'), (162, 'muy buena la atenci\xc3\xbb3n '), (163, 'muy larga'), (164, 'ninguna'), (165, 'ninguna me parece muy buena su atencion'), (166, 'ninguna por el momento, ya que que no tengo queja sobre la calidad de los servicios ofrecidos.'), (167, 'ninguna. el metodo actual es realmente eficaz.'), (168, 'no hay sugerencia, todo bien'), (169, 'no tengo nada que acotar debido a que desde que obtuve este beneficio la atenci\xc3\xbb3n a sido eficiente'), (170, 'no tengo sugerencias'), (171, 'ojala pudieran tener presupuesto para que todos sus servicios funcionaran sin problemas, por ejemplo odontologia'), (172, 'para la renovacion de becas extender el horario de atencion .'), (173, 'prolongar horarios de atencion'), (174, 'que extiendan un poco mas el horario de la entrega de los requisitos para la renovacion ya que muchos tenemos horarios de clases con poca flexibilidad como por ejemplo laboratorios todo el dia con un tiempo limitado de almuerzo y descanso'), (175, 'que la trabajadora social trabaje todo el dia cuando este recopilando los documentos, ya que estoy cursando el internado rotatorio y no me da chance llegar antes de las 1 y 30. gracias.'), (176, 'que las jornadas duren mas tiempo'), (177, 'que todo el proceso se pueda llevar en base electronica, mediante el scan de documentos'), (178, 'realizar mas jornadas en la universidad '), (179, 'se recomienda mas difusion acerca de los servicios que ofrece obe la cual podriaser otorgada por correo electronico '), (180, 'seguir la renovaci\xc3\xbb3n por este medio, me parece excelente esta v\xc3\xada ya que se evita el uso de hojas y se contribuye con el reciclaje y por lo tanto con el medio ambiente.'), (181, 'seguir luchando para una beca mas considerada por los estudiantes ya que no nos alcanza para cubrir nuestros gasto de estudio un libro ya hoy en dia estan minimo 5000 bueno se que no es su culpa del moto que estamos recibiendo horita '), (182, 'seguir manteniendo el buen trato con los alumnos, y mantenernos informados de lo que ocurre en la univerisidad con respecto a la beca'), (183, 'ser un poco mas amables con las personas.'), (184, 'siempre son muy receptivos. Hasta los momentos no tengo ninguna sugerencia.'), (185, 'sigan asi, atendiendo al estudiante con amor.. prestando el servicio de orientaci\xc3\xbb3n estudiantil como hasta ahora se han venido destacando. cada dia mejoraremos todos como universidad y sociedad.'), (186, 'todo est\xc3\xa1 bastante bien...!')

#### Sugerenciasyrecomendacionesparamejorar nuestraatencion

count	190.000000
mean	92.468421
std	54.664407
min	0.000000
25%	45.250000
50%	92.500000
75%	139.750000
max	186.000000

```

mode(df.Sugerenciasyrecomendacionesparamejorar nuestraatencion)
ModeResult(mode=array([1]), count=array([2]))
  
```

## Estandarización y normalización

No aplique ninguno de estos métodos ya que me pareció que no hacen falta, ya que se pueden visualizar los valores de forma adecuada ya que no hay muchas filas y todos tienen valores muy cercanos. Y el otro motivo es que no se sabe cual es el uso que se le va dar a la vista minable, y si aplico estas técnicas puedo alterar el set de datos de forma no adecuada, por ejemplo si se quiere usar el dataset para saber a quien dar becas, y he aplicado estas técnicas es posible que los estudiantes que más se merecen una beca, tengan menos prioridad.

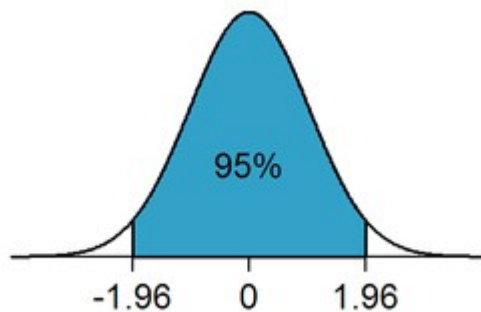
## Detección de Outliers

Para la detección de outliers o valores anómalos, utilice el siguiente criterio:

Media:

```
dataframe['x-Mean'] = abs(dataframe['Columna'] - dataframe['Columna'].mean())
```

Desviación: Utilizo la medida de centralidad de 1,96, es decir que si el valor no se encuentra en el 95% del conjunto de datos entonces es un outlier, como podemos observar en la siguiente figura.



```
dataframe['1.96*std'] = 1.96*dataframe['Columna'].std()
```

```
dataframe['Outlier'] = abs(dataframe['Columna'] - dataframe['Columna'].mean()) >  
1.96*dataframe['Columna'].std()
```

En la carpeta Outlier, señalo todas las columnas que según este criterio poseen outliers.

