GRUPO 3

UNIVERSIDAD CÉSAR VALLEJO

EVALUACIÓN ESTADÍSTICA DEL PESO Y TAMAÑO DE HUEVOS EN UNA GRANJA AVÍCOLA

ESTADÍSTICA APLICADA-A1T1

UNIVERSIDAD CÉSAR VALLEJO

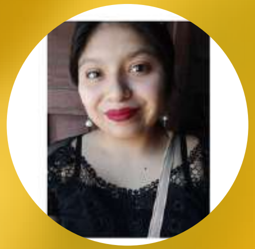
DOCENTE:

OBREGÓN DOMINGUEZ JESÚS

INTEGRANTES:

- ✓ Abanto Alvarez, Sarai Indira

- ✓ Bardales Bautista, Noelia Lizbet



- ✓ Gallo Senmache, Natalia

- ✓ Rojas Abanto, Jenifer Lizeth

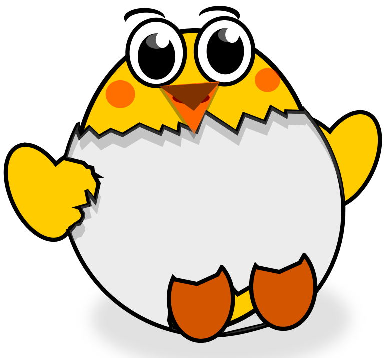
- ✓ Guillermo Ramirez, Tracy

UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

INTRODUCCIÓN:

EN ESTE INFORME, SE PRESENTA UN ESTUDIO QUE TIENE COMO OBJETIVO EVALUAR EL PORCENTAJE DE ÉXITO EN LA PUESTA DE HUEVOS Y EN LA TOMA DE DATOS LA CALIDAD DE LOS HUEVOS EN UNA GRANJA AVÍCOLA. LA EVALUACIÓN SE CENTRARA EN LA PRIMERA ETAPA EN LA PUESTA DE HUEVOS, Y SE CONTINUARÁ CON EL PESO Y EL TAMAÑO DE LOS HUEVOS, Y SE HA REALIZADO UTILIZANDO UN MUESTREO ALEATORIO SIMPLE, EL ESTUDIO SE APLICARÁ EN VARIOS GALPONES DE LA GRANJA.



UNIVERSIDAD CÉSAR VALLEJO

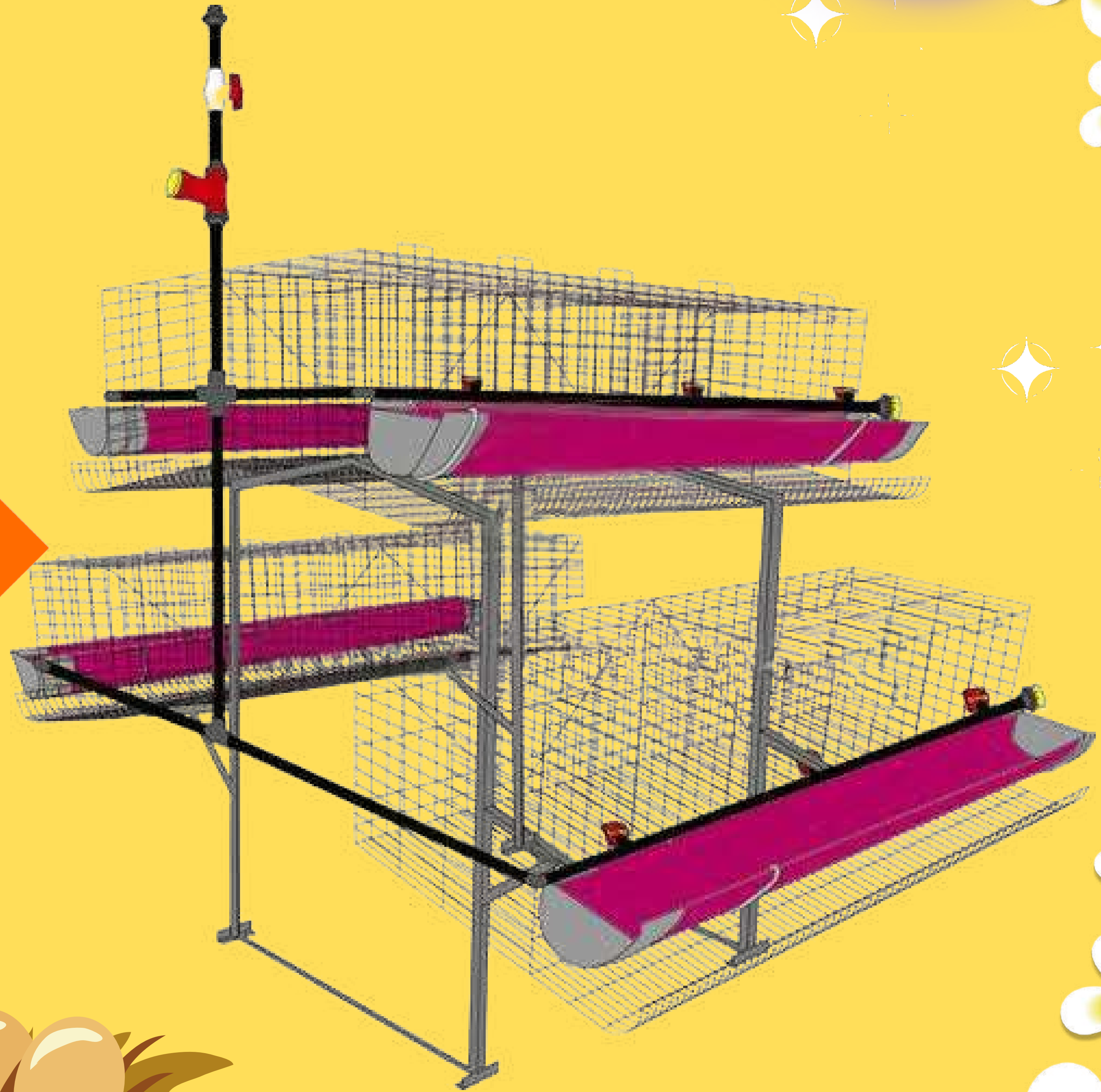
GRUPO 3

3 GALPONES

UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

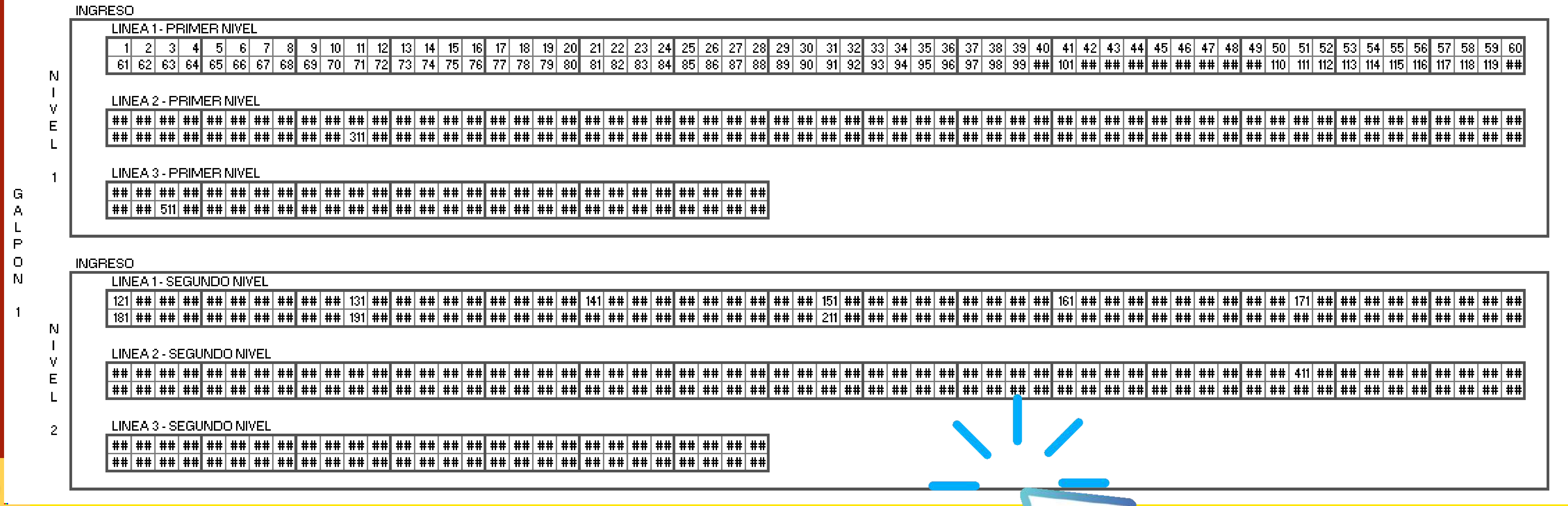
FORMA DE LOS GALPONES



UNIVERSIDAD CÉSAR VALLEJO

DISTRIBUCIÓN

GALPÓN 01



GALPON 1

INGRESO

NIVEL 1

LINEA 1 - PRIMER NIVEL

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	##	101	##	##	##	##	##	##	##	##	110	111	112	113	114	115	116	117	118	119	##

LINEA 2 - PRIMER NIVEL

##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	##	311	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##

LINEA 3 - PRIMER NIVEL

##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##
##	##	511	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##

INGRESO

NIVEL 2

LINEA 1 - SEGUNDO NIVEL

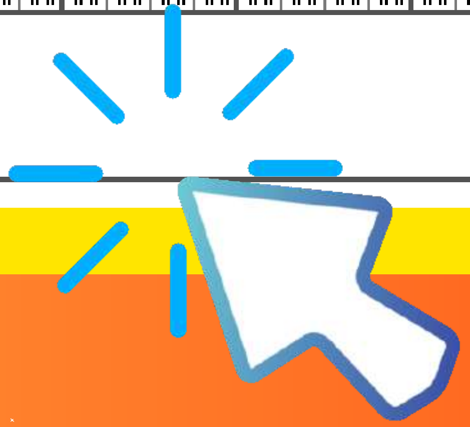
121	##	##	##	##	##	##	##	##	##	131	##	##	##	##	##	##	##	##	##	141	##	##	##	##	##	##	##	##	##	151	##	##	##	##	##	##	##	##	##	161	##	##	##	##	##	##	##	##	##	171	##	##	##	##	##	##	##	##	##
181	##	##	##	##	##	##	##	##	##	191	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	211	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##

LINEA 2 - SEGUNDO NIVEL

##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	411	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##

LINEA 3 - SEGUNDO NIVEL

##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##	##

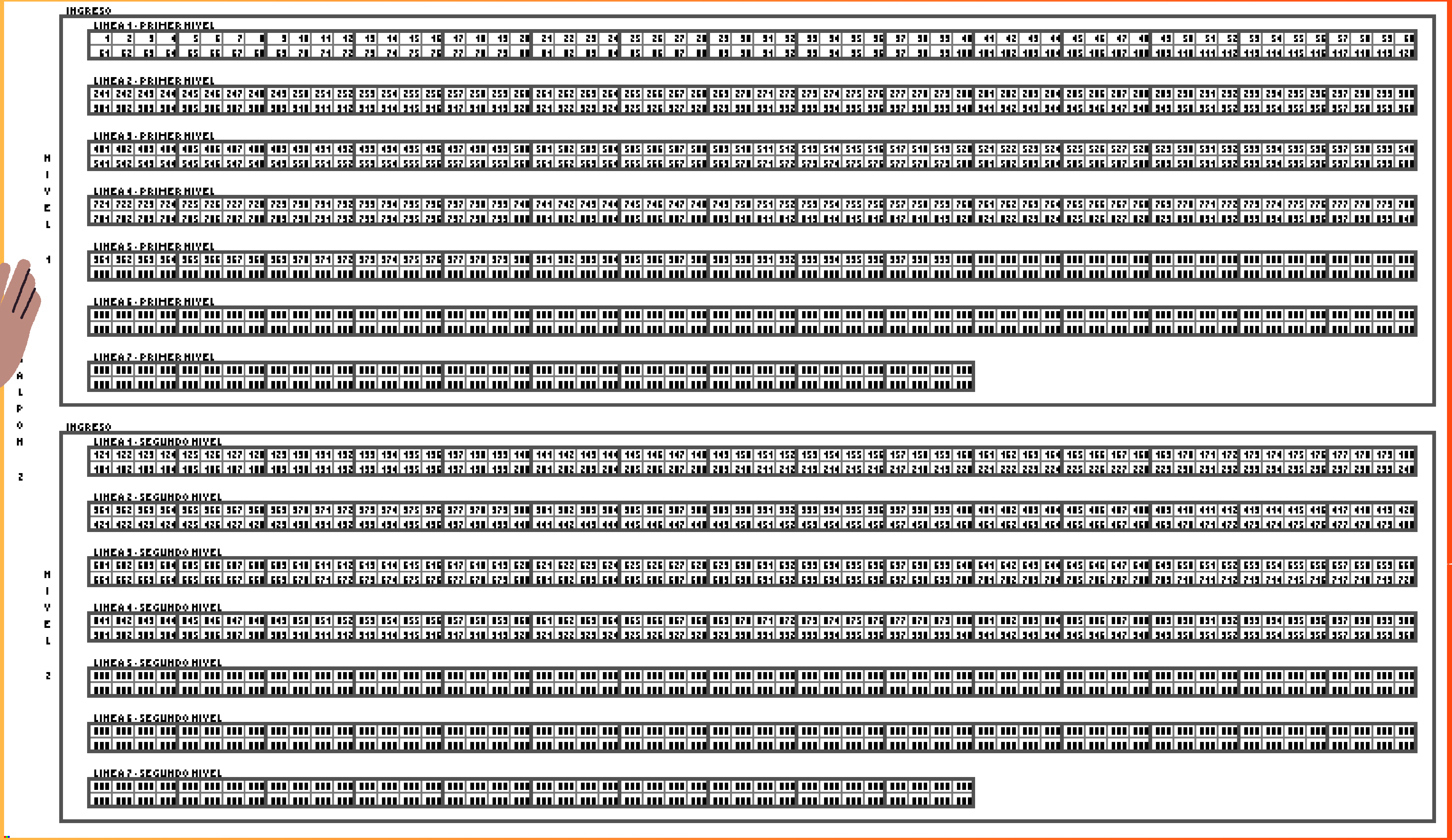


UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

DISTRIBUCIÓN

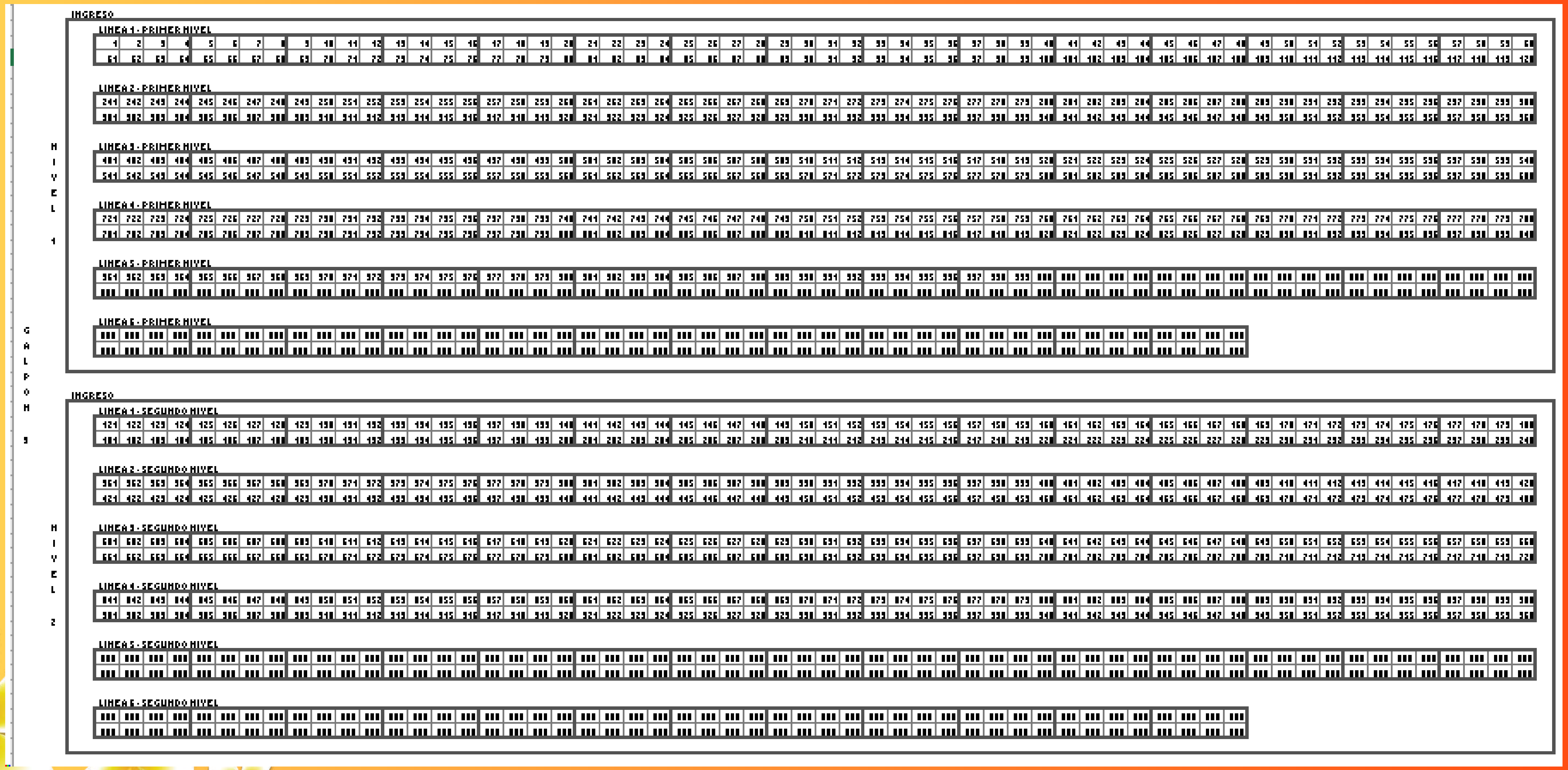
GALPÓN
02


INGRESO
LINEA 1 - PRIMER NIVEL
LINEA 2 - PRIMER NIVEL
LINEA 3 - PRIMER NIVEL
LINEA 4 - PRIMER NIVEL
LINEA 5 - PRIMER NIVEL
LINEA 6 - PRIMER NIVEL
LINEA 7 - PRIMER NIVEL
NIVEL 1
GALPON 2
INGRESO
LINEA 1 - SEGUNDO NIVEL
LINEA 2 - SEGUNDO NIVEL
LINEA 3 - SEGUNDO NIVEL
LINEA 4 - SEGUNDO NIVEL
LINEA 5 - SEGUNDO NIVEL
LINEA 6 - SEGUNDO NIVEL
LINEA 7 - SEGUNDO NIVEL
NIVEL 2

UNIVERSIDAD CÉSAR VALLEJO

DISTRIBUCIÓN

GALPÓN 03



INGRESO

LINEA 1 - PRIMER NIVEL

LINEA 2 - PRIMER NIVEL

LINEA 3 - PRIMER NIVEL

LINEA 4 - PRIMER NIVEL

LINEA 5 - PRIMER NIVEL

LINEA 6 - PRIMER NIVEL

NIVEL 1

GALPON 3

INGRESO

LINEA 1 - SEGUNDO NIVEL

LINEA 2 - SEGUNDO NIVEL

LINEA 3 - SEGUNDO NIVEL

LINEA 4 - SEGUNDO NIVEL

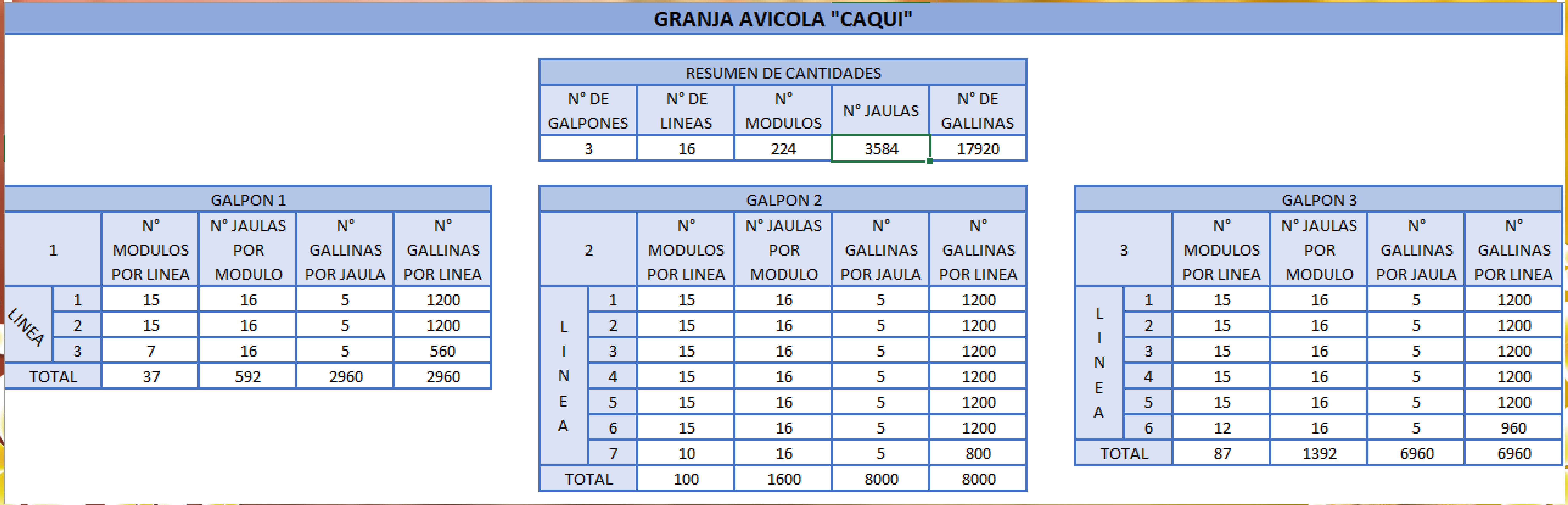
LINEA 5 - SEGUNDO NIVEL

LINEA 6 - SEGUNDO NIVEL

NIVEL 2

UNIVERSIDAD CÉSAR VALLEJO

GRANJA AVICOLA "CAQUI"



GRANJA AVICOLA "CAQUI"

RESUMEN DE CANTIDADES				
N° DE GALPONES	N° DE LINEAS	N° MODULOS	N° JAULAS	N° DE GALLINAS
3	16	224	3584	17920

GALPON 1					
1		N° MODULOS POR LINEA	N° JAULAS POR MODULO	N° GALLINAS POR JAULA	N° GALLINAS POR LINEA
LINEA	1	15	16	5	1200
	2	15	16	5	1200
	3	7	16	5	560
TOTAL		37	592	2960	2960

GALPON 2					
2		N° MODULOS POR LINEA	N° JAULAS POR MODULO	N° GALLINAS POR JAULA	N° GALLINAS POR LINEA
LINEA	1	15	16	5	1200
	2	15	16	5	1200
	3	15	16	5	1200
	4	15	16	5	1200
	5	15	16	5	1200
	6	15	16	5	1200
	7	10	16	5	800
TOTAL		100	1600	8000	8000

GALPON 3					
3		N° MODULOS POR LINEA	N° JAULAS POR MODULO	N° GALLINAS POR JAULA	N° GALLINAS POR LINEA
LINEA	1	15	16	5	1200
	2	15	16	5	1200
	3	15	16	5	1200
	4	15	16	5	1200
	5	15	16	5	1200
	6	12	16	5	960
TOTAL		87	1392	6960	6960

UNIVERSIDAD CÉSAR VALLEJO

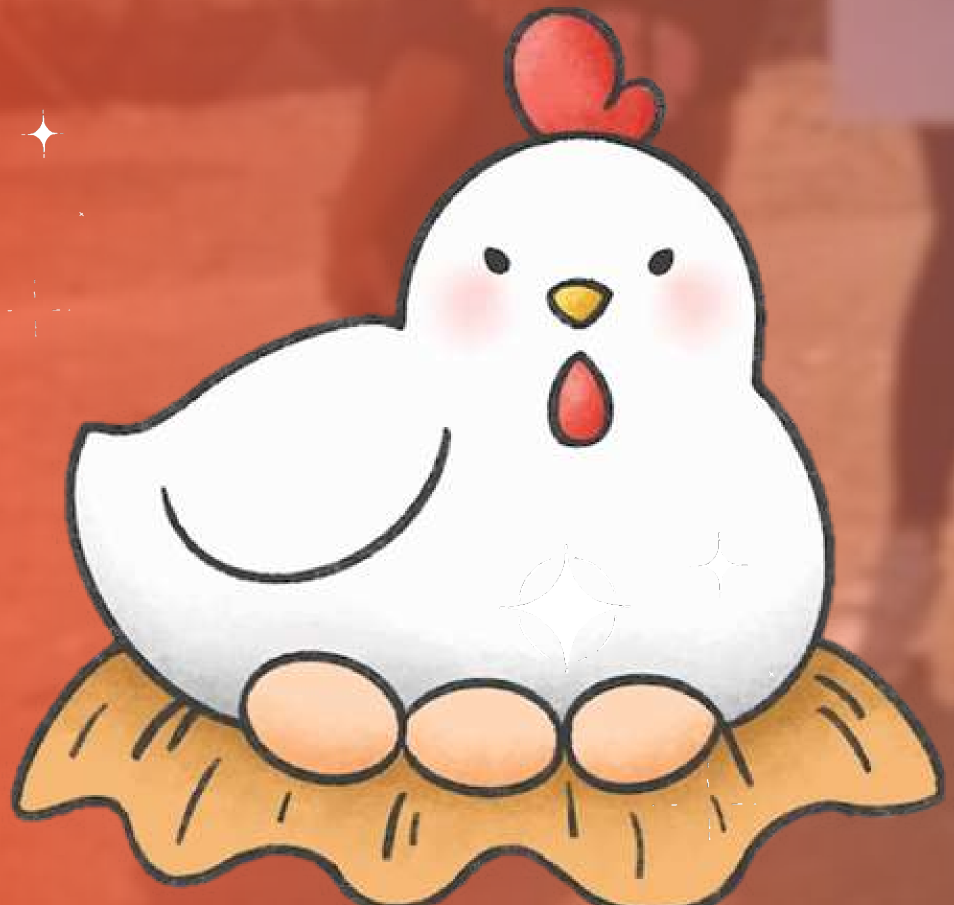
PRUEBA PILOTO

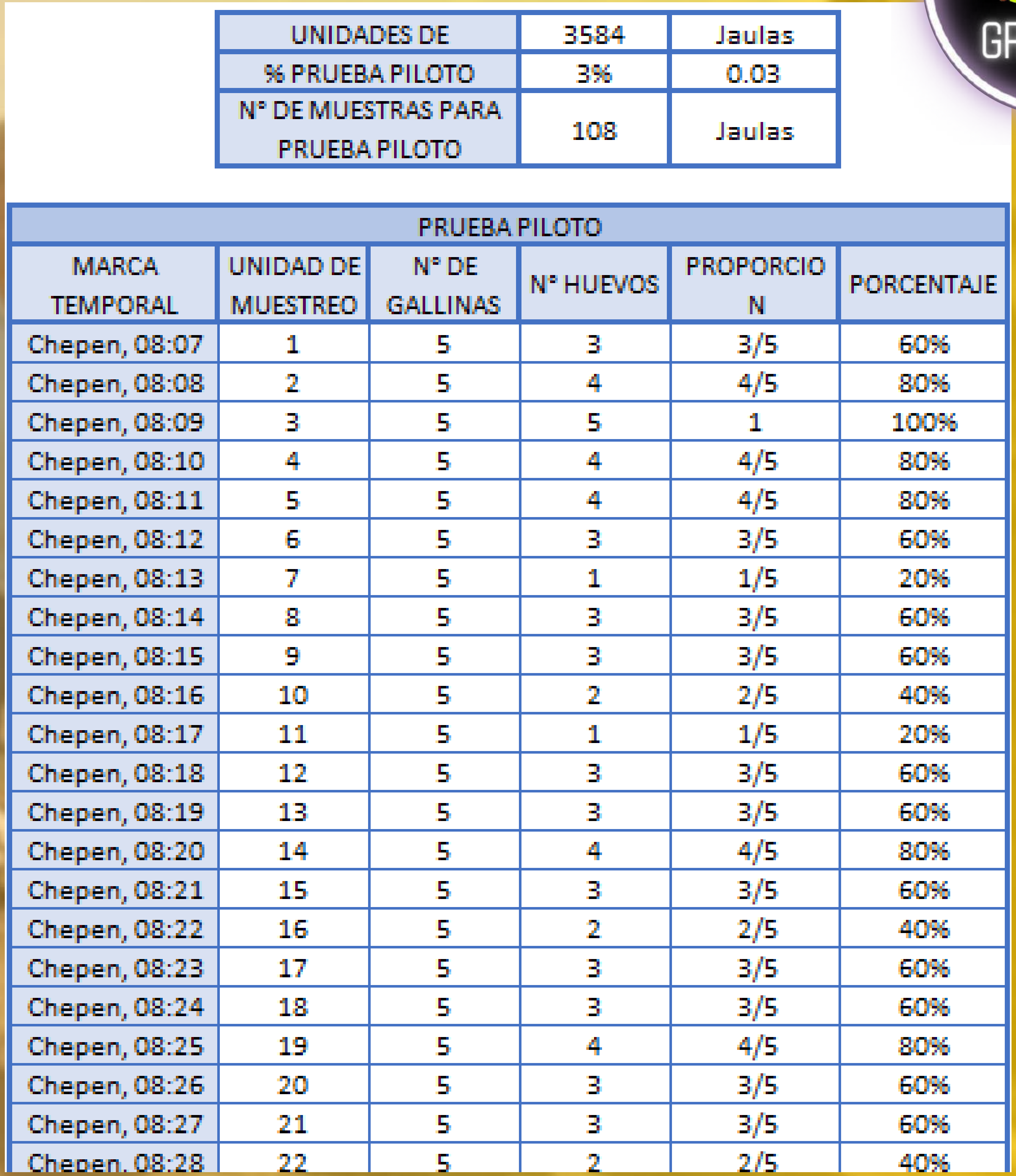
GRUPO 3

Durante la prueba piloto, estaremos recopilando datos de una muestra reducida de jaulas para evaluar la eficacia de nuestra metodología de recolección de datos. Queremos asegurarnos de que nuestros procedimientos sean precisos y eficientes, y esta fase piloto nos proporcionará información valiosa para lograrlo.

GRUPO 3

PRUEBA PILOTO:





UNIDADES DE	3584	Jaulas
% PRUEBA PILOTO	3%	0.03
N° DE MUESTRAS PARA PRUEBA PILOTO	108	Jaulas

PRUEBA PILOTO

MARCA TEMPORAL	UNIDAD DE MUESTREO	N° DE GALLINAS	N° HUEVOS	PROPORCION	PORCENTAJE
Chepen, 08:07	1	5	3	3/5	60%
Chepen, 08:08	2	5	4	4/5	80%
Chepen, 08:09	3	5	5	1	100%
Chepen, 08:10	4	5	4	4/5	80%
Chepen, 08:11	5	5	4	4/5	80%
Chepen, 08:12	6	5	3	3/5	60%
Chepen, 08:13	7	5	1	1/5	20%
Chepen, 08:14	8	5	3	3/5	60%
Chepen, 08:15	9	5	3	3/5	60%
Chepen, 08:16	10	5	2	2/5	40%
Chepen, 08:17	11	5	1	1/5	20%
Chepen, 08:18	12	5	3	3/5	60%
Chepen, 08:19	13	5	3	3/5	60%
Chepen, 08:20	14	5	4	4/5	80%
Chepen, 08:21	15	5	3	3/5	60%
Chepen, 08:22	16	5	2	2/5	40%
Chepen, 08:23	17	5	3	3/5	60%
Chepen, 08:24	18	5	3	3/5	60%
Chepen, 08:25	19	5	4	4/5	80%
Chepen, 08:26	20	5	3	3/5	60%
Chepen, 08:27	21	5	3	3/5	60%
Chepen, 08:28	22	5	2	2/5	40%

La prueba piloto, una etapa crucial de nuestro estudio de evaluación de puesta de huevos en la granja avícola, se llevó a cabo con éxito para afinar nuestra metodología y determinar el tamaño de muestra óptimo necesario para el estudio completo.

EL OBJETIVO PRINCIPAL

El objetivo principal de esta fase era garantizar que nuestros procedimientos fueran precisos y eficientes antes de avanzar con la evaluación completa.

UNIVERSIDAD CÉSAR VALLEJO

TAMAÑO DE LA MUESTRA

VARIABLE	Cuantitativa
POBLACIÓN	Finita
UNIDAD DE MUESTRA	Jaula de gallinas
UNIDAD DE ANALISIS	Gallinas

N	POBLACION		3584
Z	NIVEL DE CONFIANZA	97%	2.17009038
E	ERROR	4%	0.04
P	% EXITO	67%	0.66851852
Q	% FRACASO	33%	0.33148148
S^2	DESVIACION ESTANDAR		0.92483148

NIVEL DE CONFIANZA		NIVEL DE SIGNIFICANCIA	Z
%	DECIMAL		
90%	0.9	0.1	1.644853627
91%	0.91	0.09	1.69539771
92%	0.92	0.08	1.750686071
93%	0.93	0.07	1.811910673
94%	0.94	0.06	1.880793608
95%	0.95	0.05	1.959963985
96%	0.96	0.04	2.053748911
97%	0.97	0.03	2.170090378
98%	0.98	0.02	2.326347874
99%	0.99	0.01	2.575829304

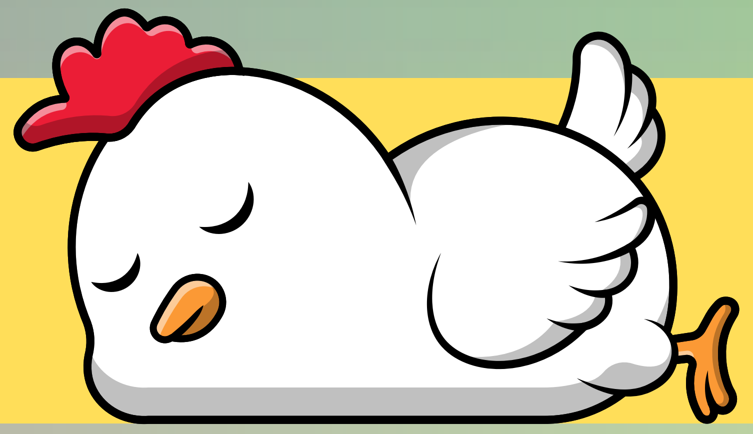
GRUPO 3

$$n = \frac{Z^2 S^2 N}{E^2(N-1) + Z^2 S^2}$$

FÓRMULA

n	1478.995962

MUESTREO ALEATORIO SIMPLE



TAMAÑO DE MUESTRA

El tamaño de muestra óptimo se establece en 1479, y este valor se basa en una combinación de factores, incluyendo el nivel de confianza requerida y considerando un porcentaje de error bajo, también las variabilidades observadas en la prueba piloto. Este número nos proporcionará una base sólida para evaluar la calidad de nuestros huevos en la granja avícola de manera confiable.

UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

TOMA DE DATOS

COMO EQUIPO, NUESTRA LABOR EN LA TOMA DE DATOS IMPLICA RECOLECTAR INFORMACIÓN PRECISA SOBRE EL LARGO, ANCHO Y PESO DE LOS HUEVOS. ESTA RECOPILACIÓN METICULOSA NOS PERMITE OBTENER DATOS PRECISOS Y RELEVANTES PARA NUESTRO ANÁLISIS Y TRABAJO CONJUNTO.

MATERIALES:

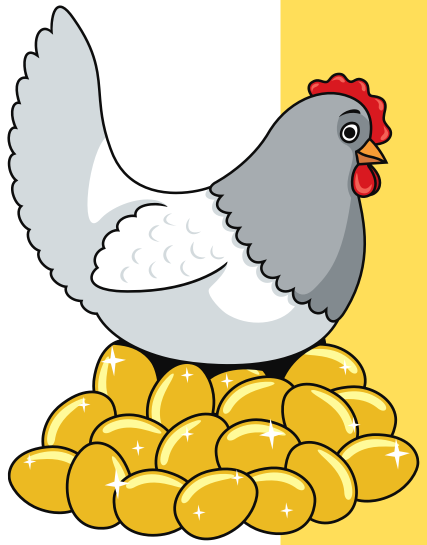
GRUPO 3

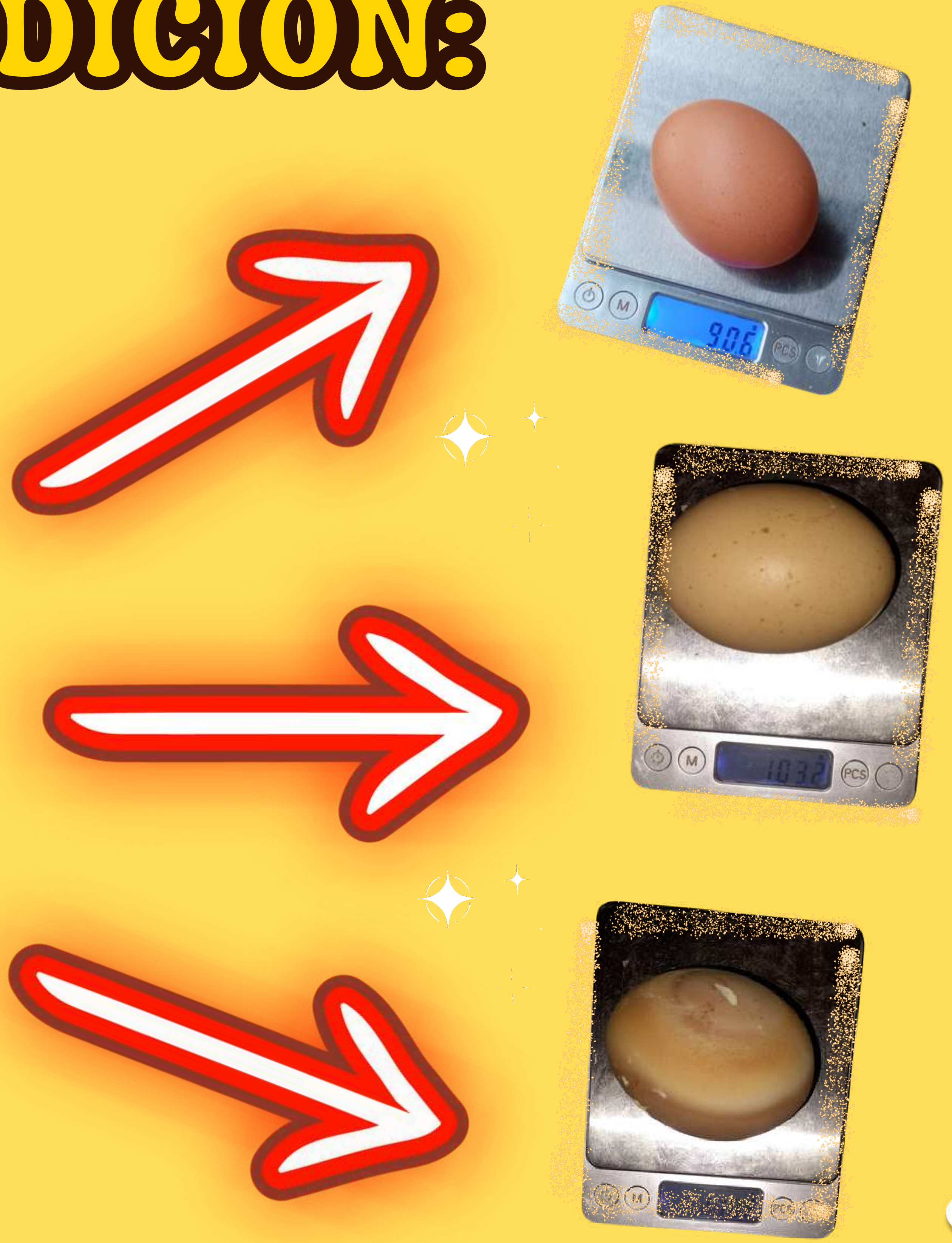
Balanza Gramera

Palitos de Chupete

Cinta Metrica

MODO DE MEDICIÓN:

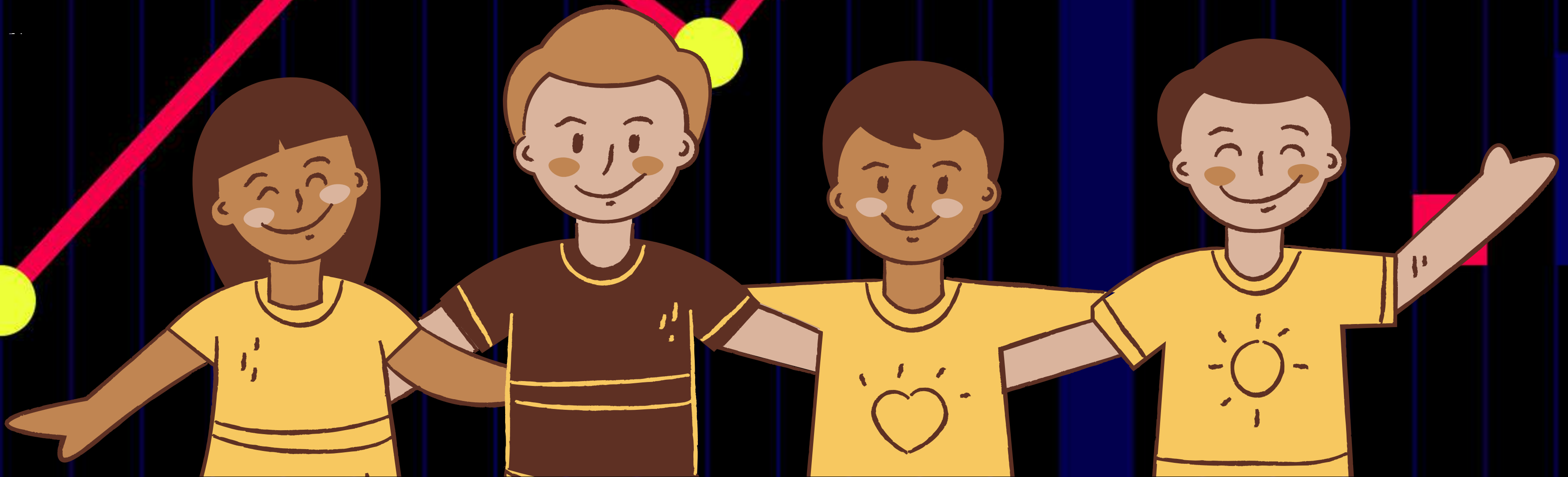




GRUPO 3

DATA

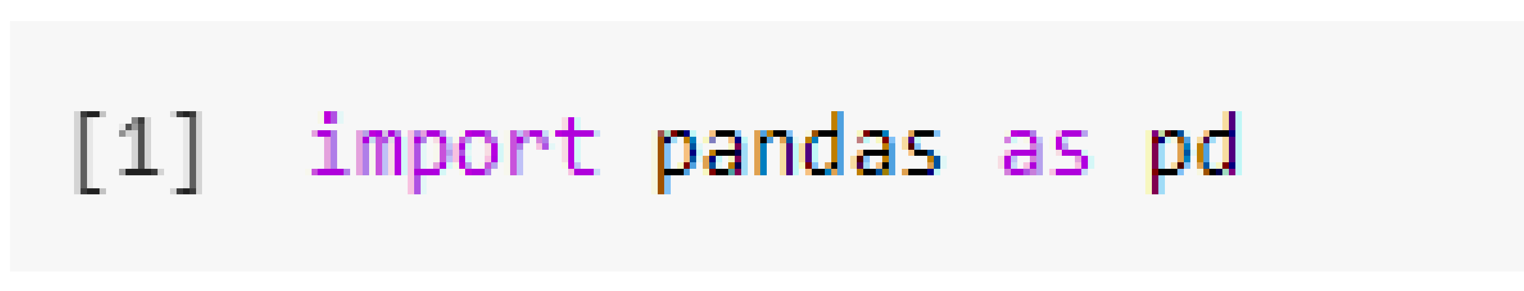
STATISTICS



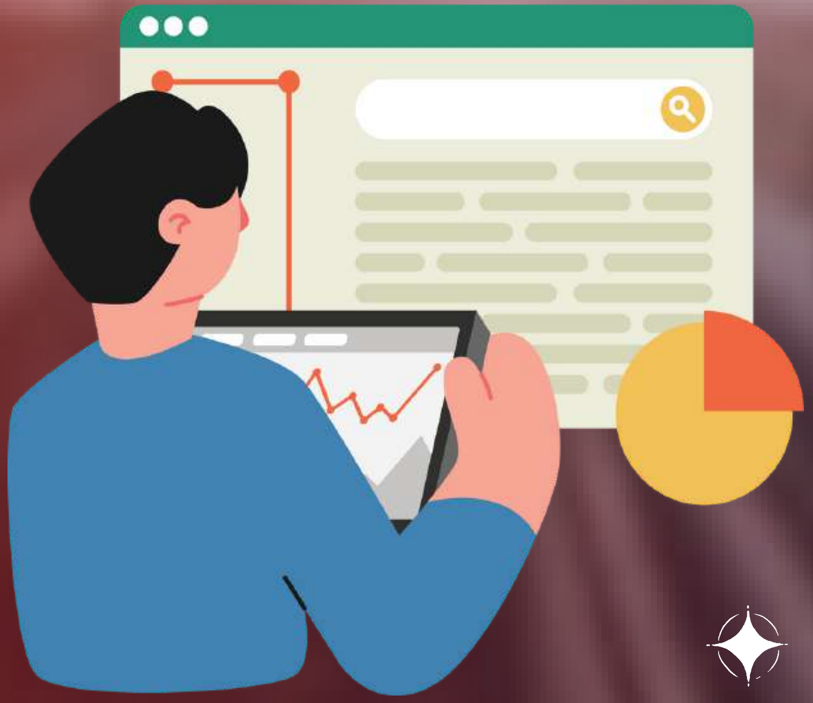
UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

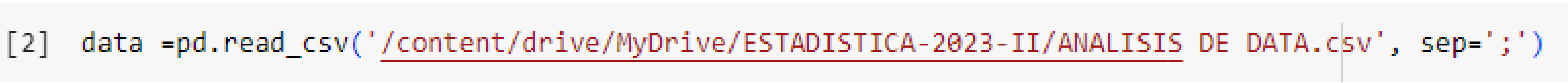
Importación de librerias



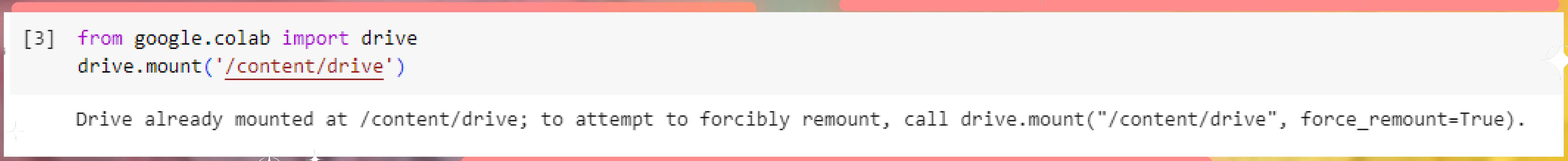
```
[1]  import pandas as pd
```



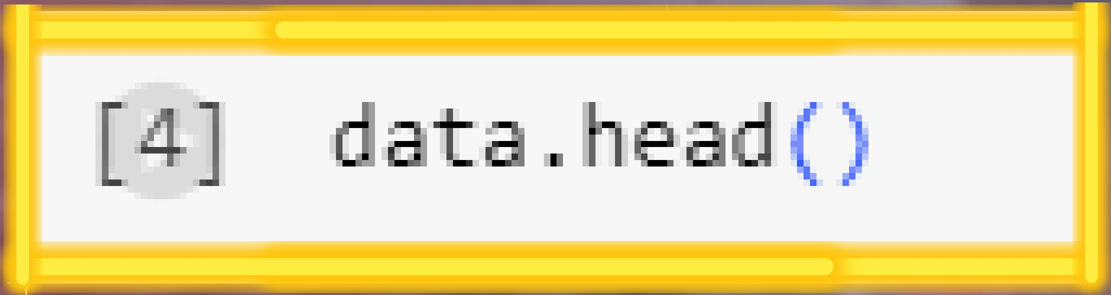
Importación de datos



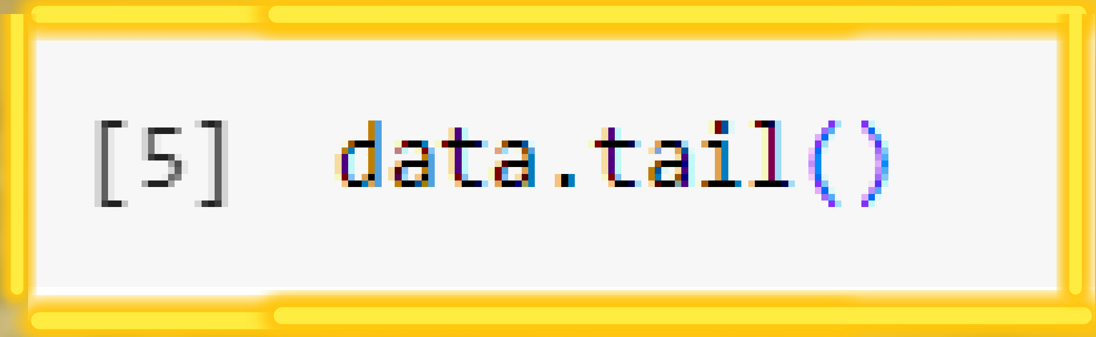
```
[2]  data =pd.read_csv('/content/drive/MyDrive/ESTADISTICA-2023-II/ANALISIS DE DATA.csv', sep=';')
```



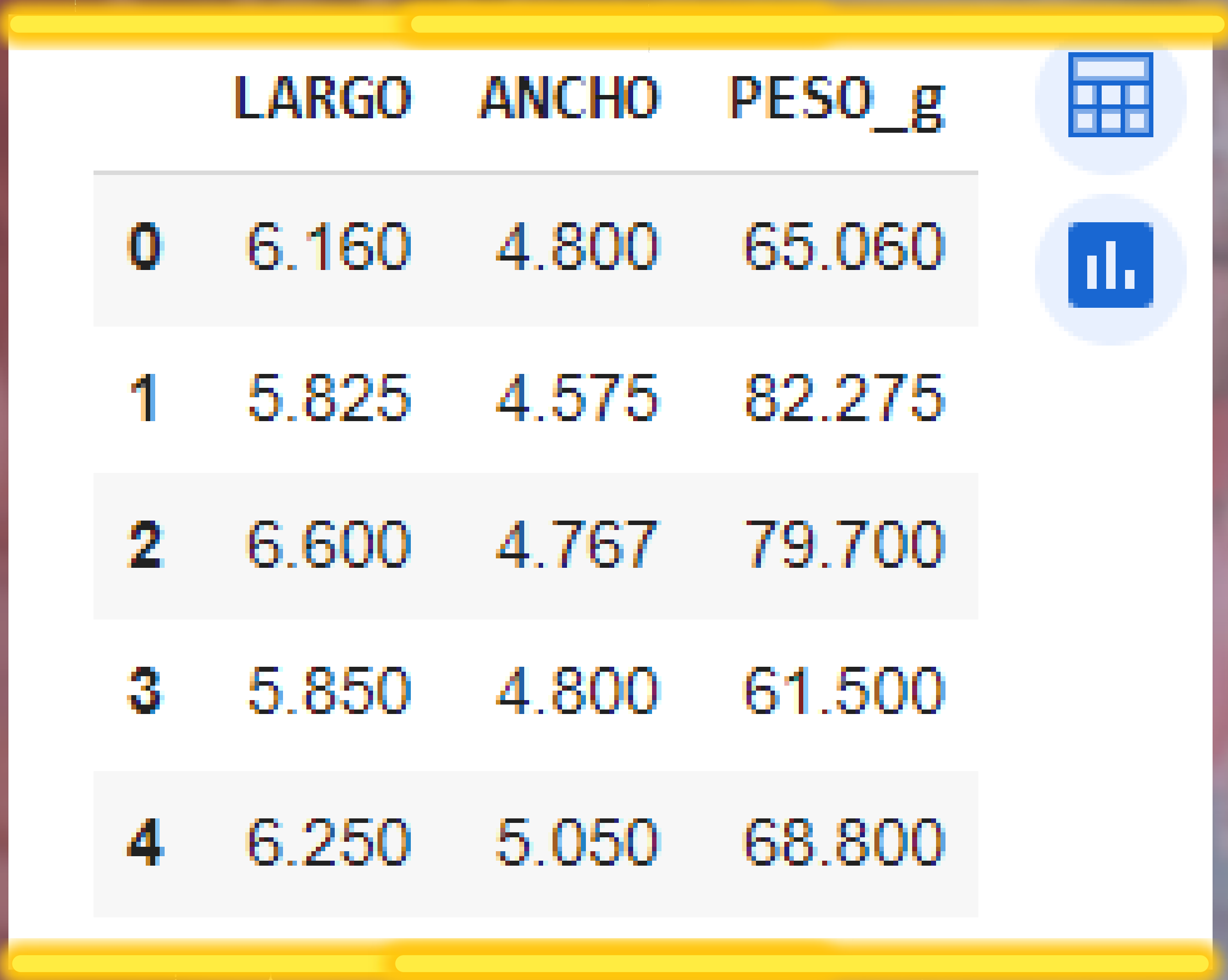
```
[3] from google.colab import drive
    drive.mount('/content/drive')

    Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```



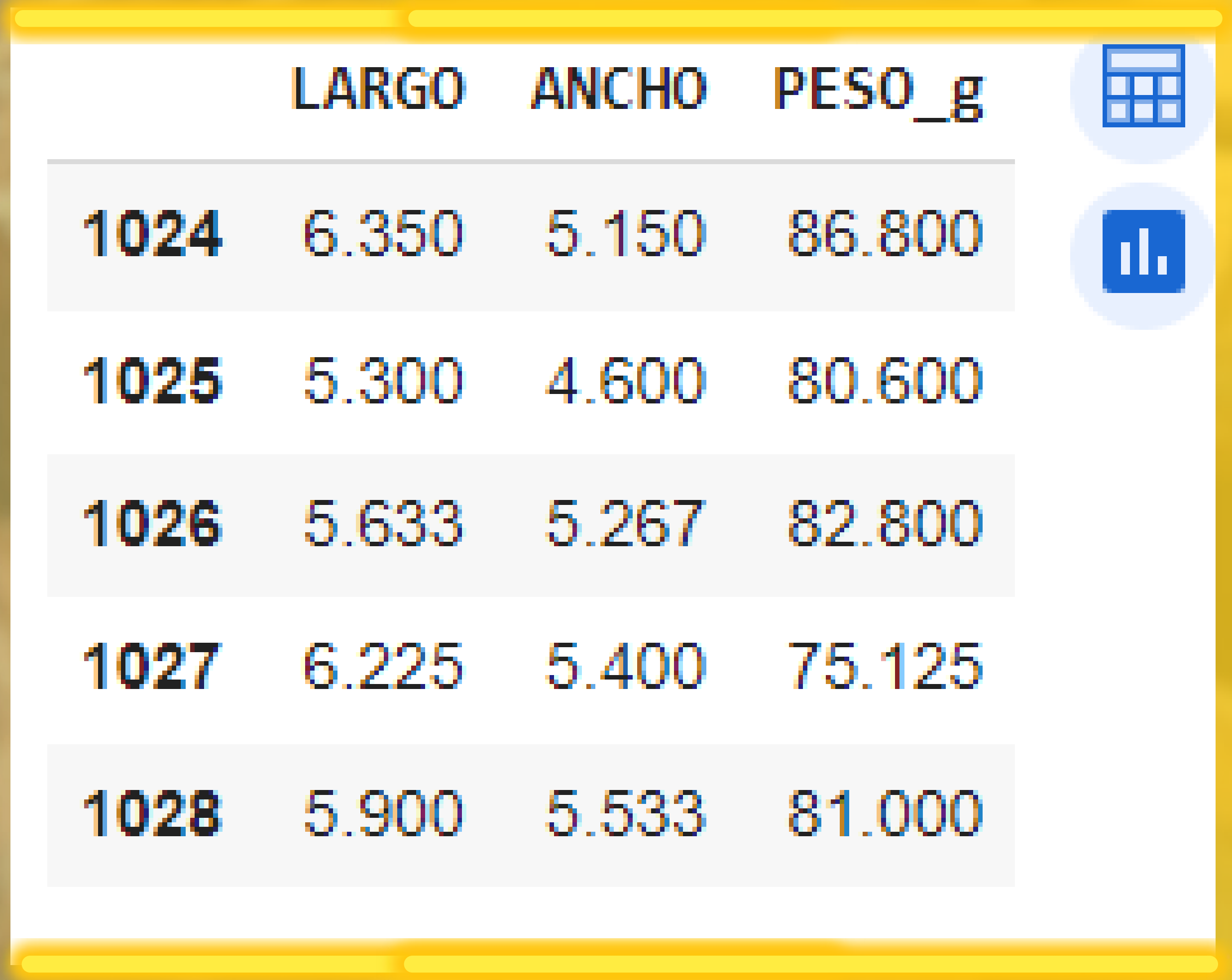
```
[4] data.head()
```



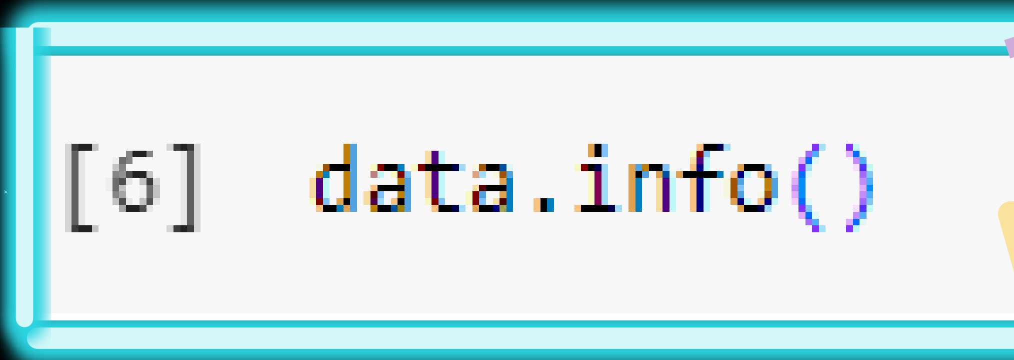
```
[5] data.tail()
```



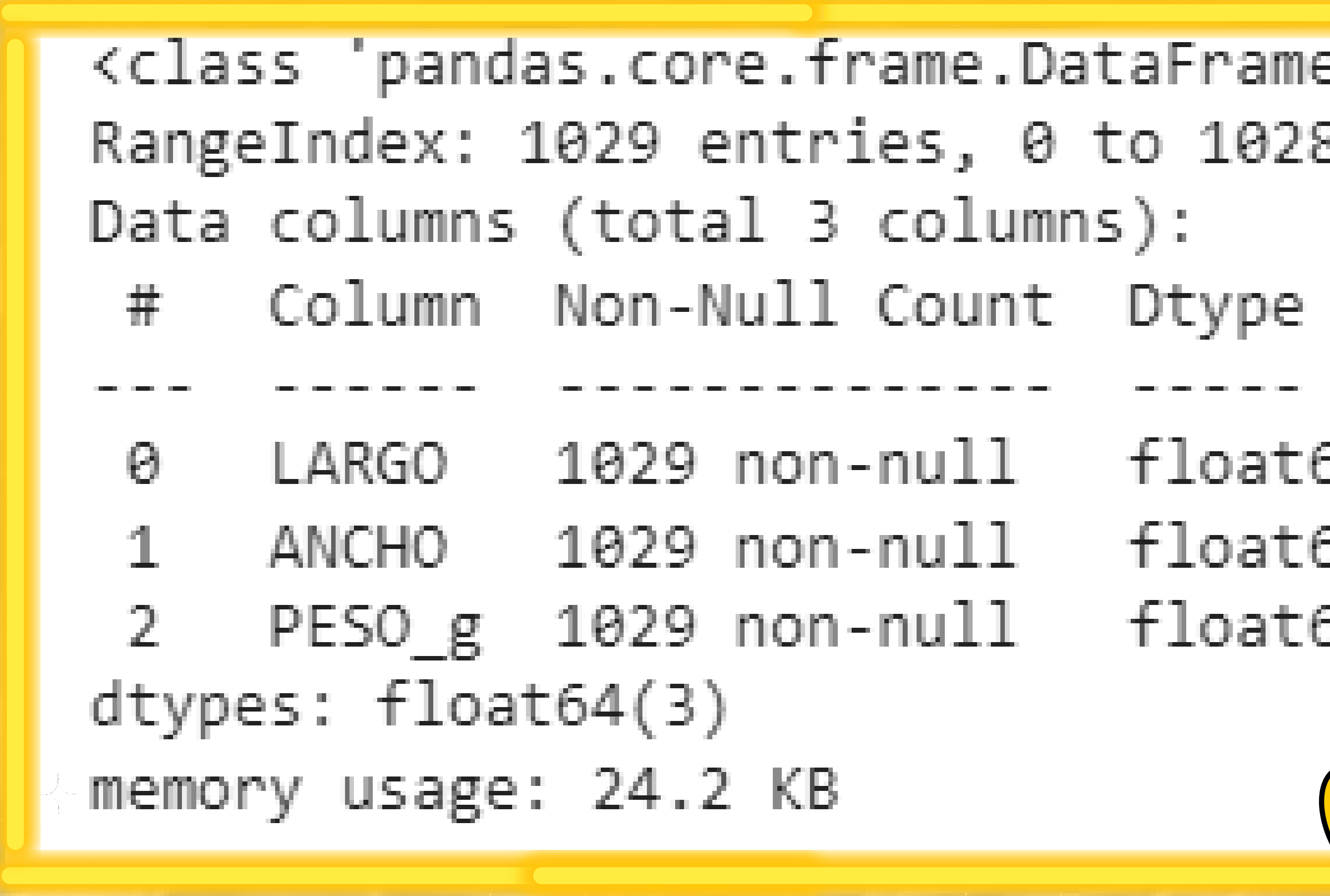
	LARGO	ANCHO	PESO_g
0	6.160	4.800	65.060
1	5.825	4.575	82.275
2	6.600	4.767	79.700
3	5.850	4.800	61.500
4	6.250	5.050	68.800



	LARGO	ANCHO	PESO_g
1024	6.350	5.150	86.800
1025	5.300	4.600	80.600
1026	5.633	5.267	82.800
1027	6.225	5.400	75.125
1028	5.900	5.533	81.000

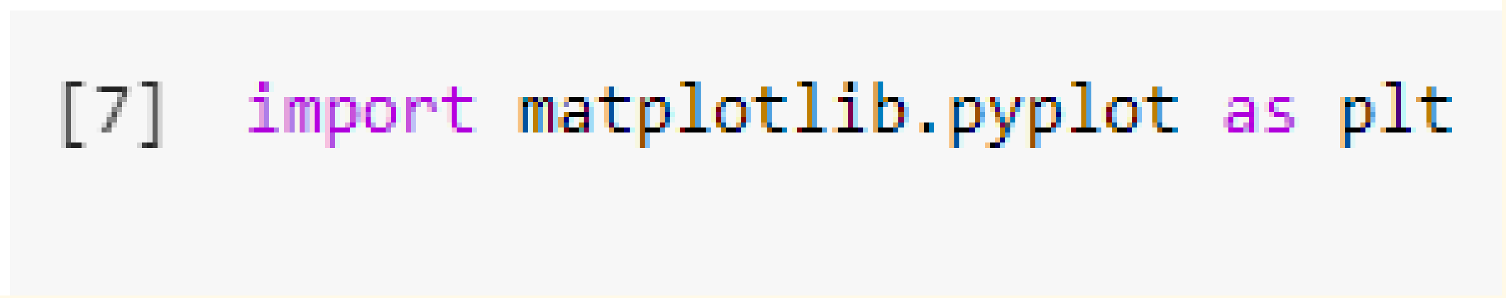


```
[6]  data.info()
```

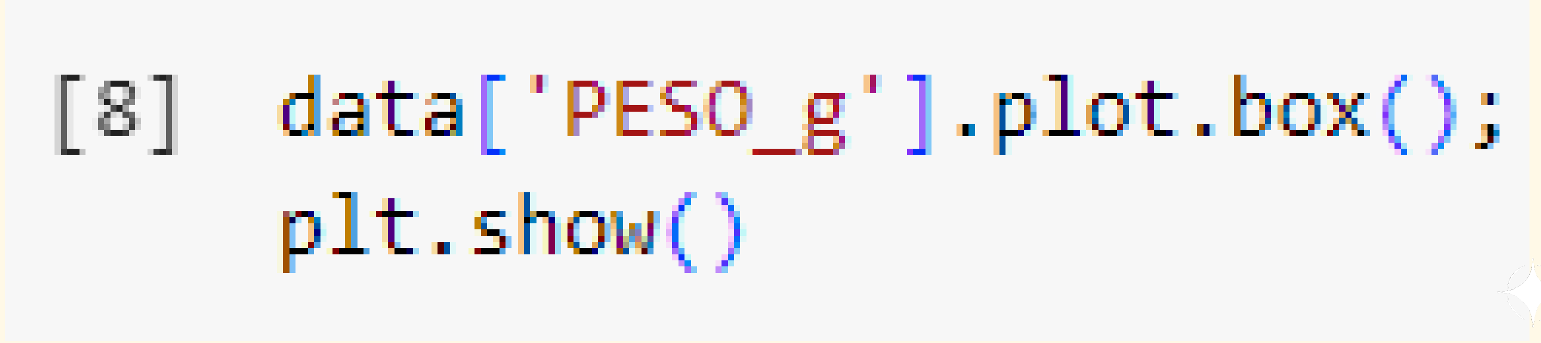


```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1029 entries, 0 to 1028
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   LARGO   1029 non-null   float64
 1   ANCHO   1029 non-null   float64
 2   PESO_g  1029 non-null   float64
dtypes: float64(3)
memory usage: 24.2 KB
```

VISUALIZACION DE DATOS

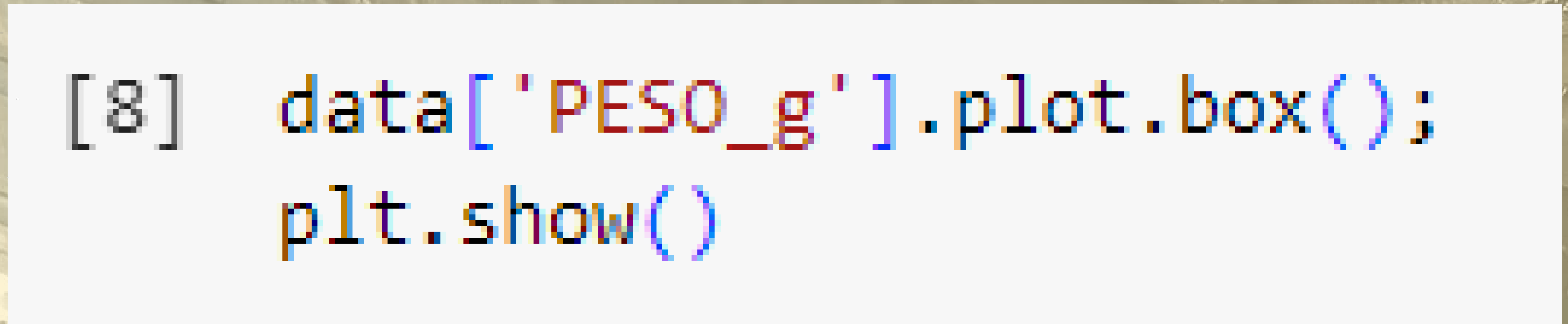


```
[7]  import matplotlib.pyplot as plt
```

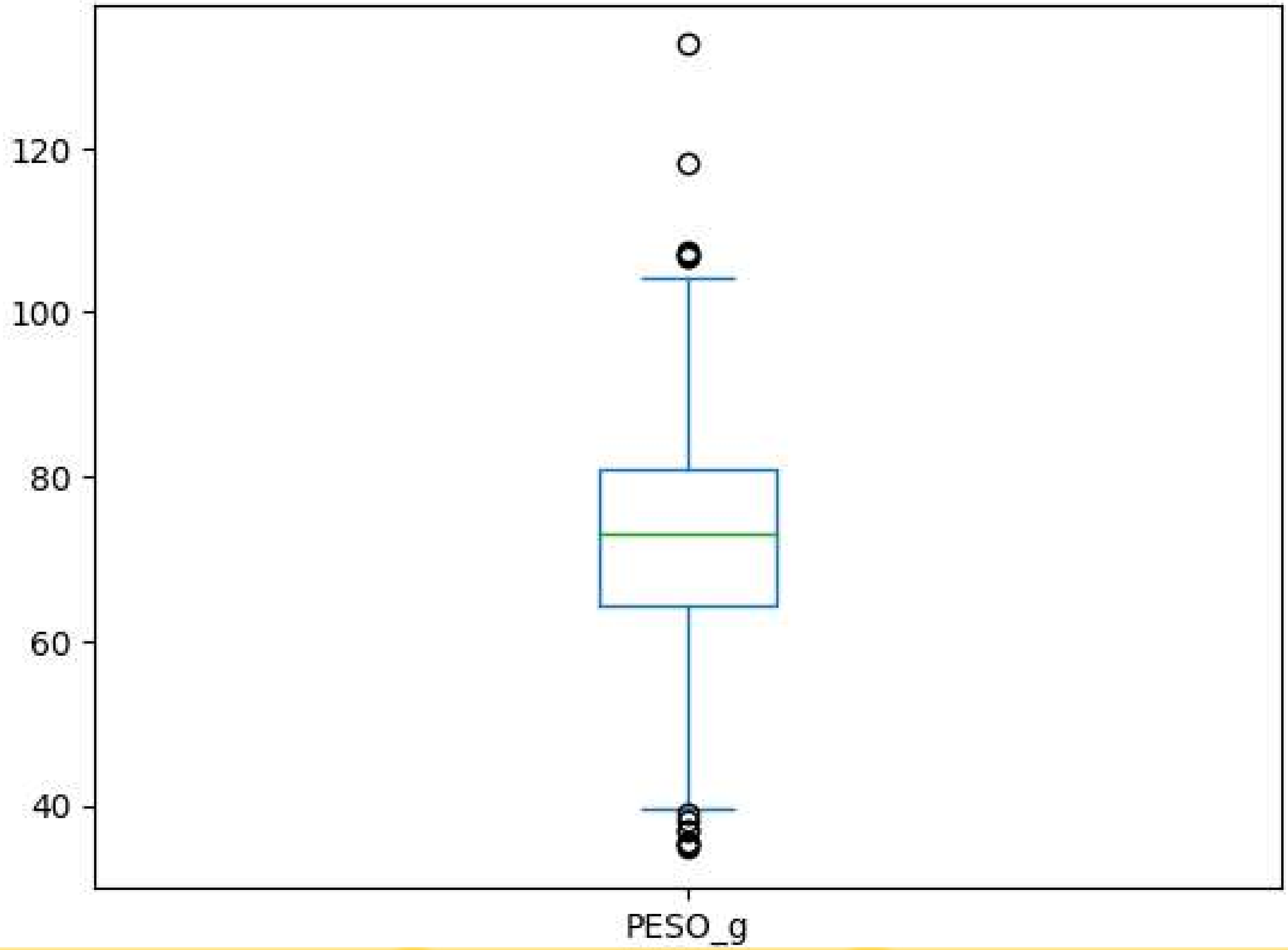


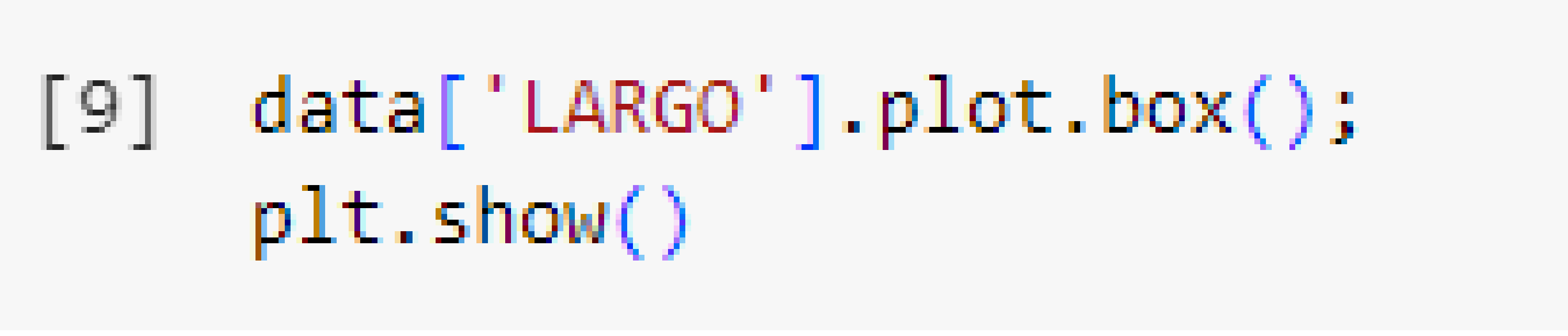
```
[8]  data['PESO_g'].plot.box();
     plt.show()
```

UNIVERSIDAD CÉSAR VALLEJO

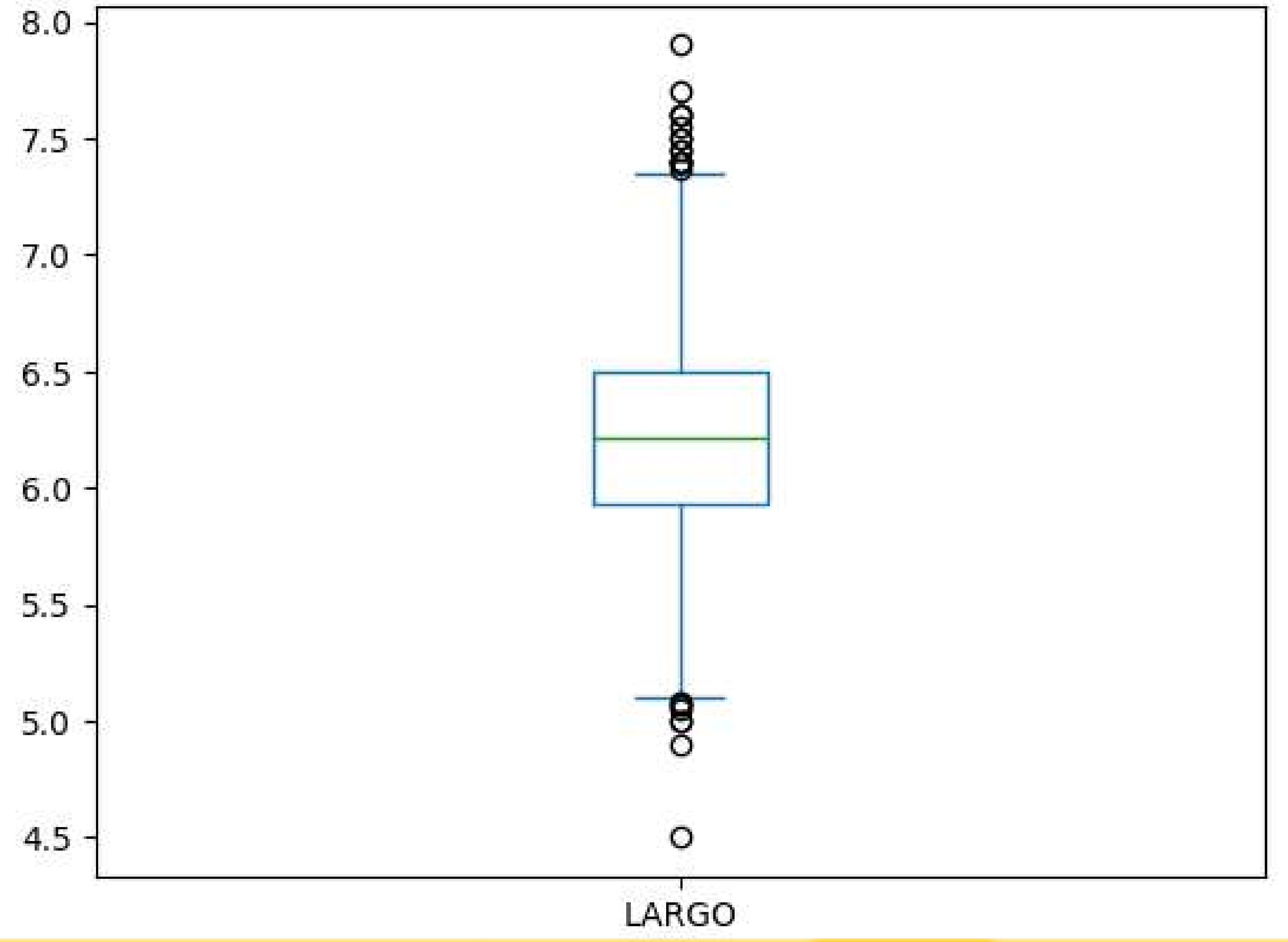


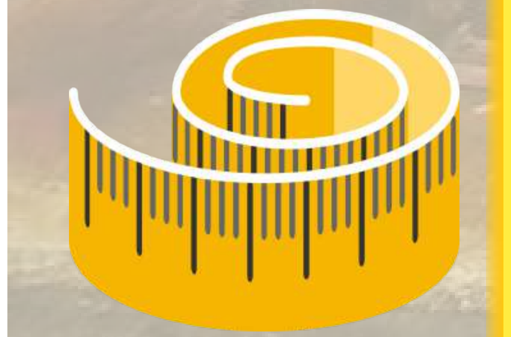
```
[8]  data['PESO_g'].plot.box();
     plt.show()
```

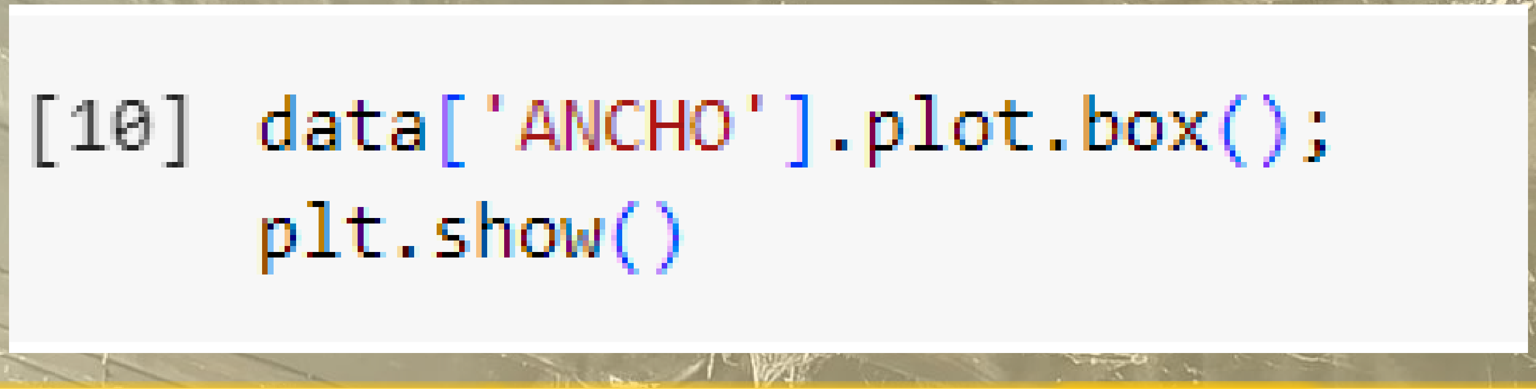




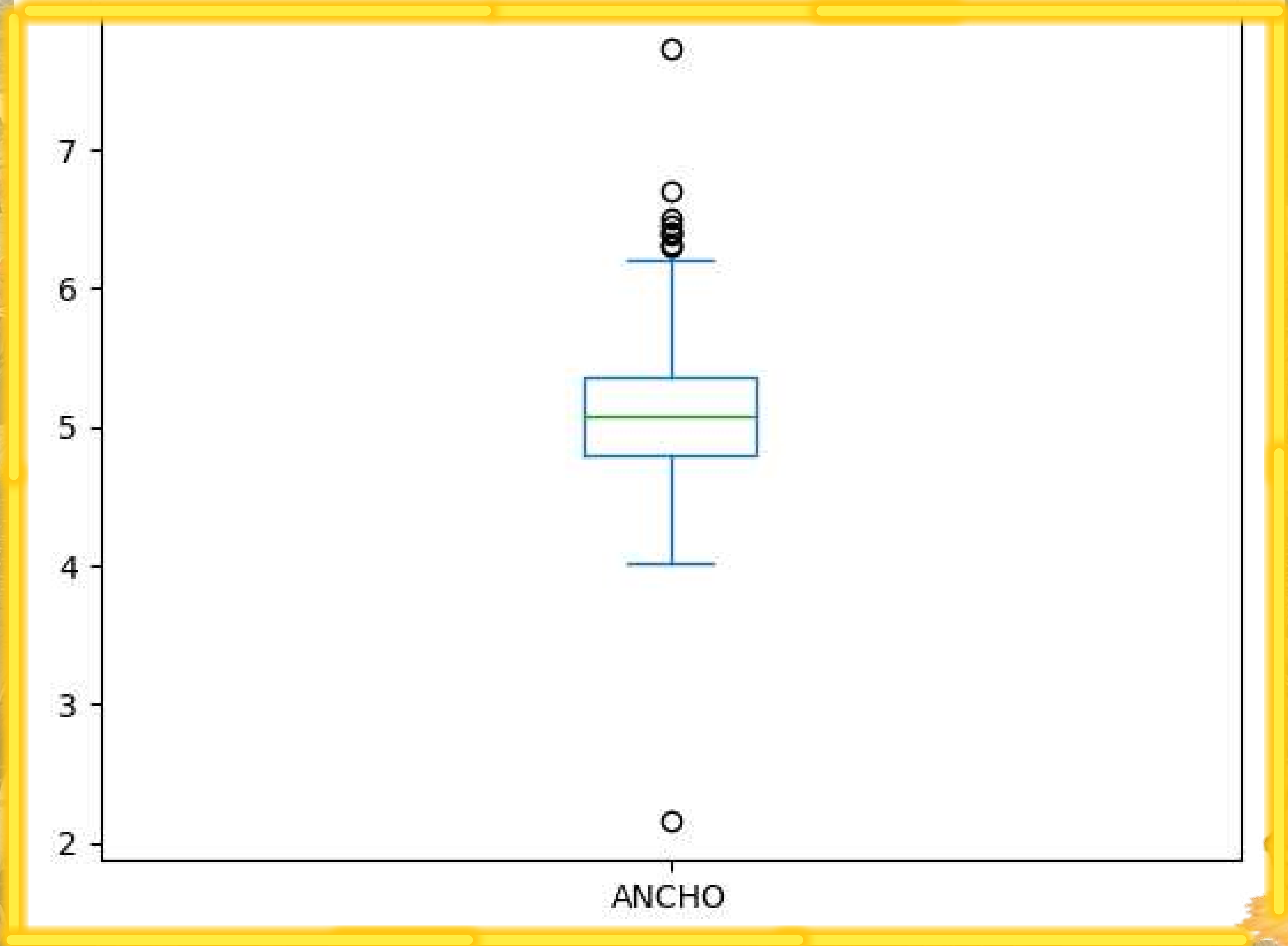
```
[9]  data['LARGO'].plot.box();
     plt.show()
```

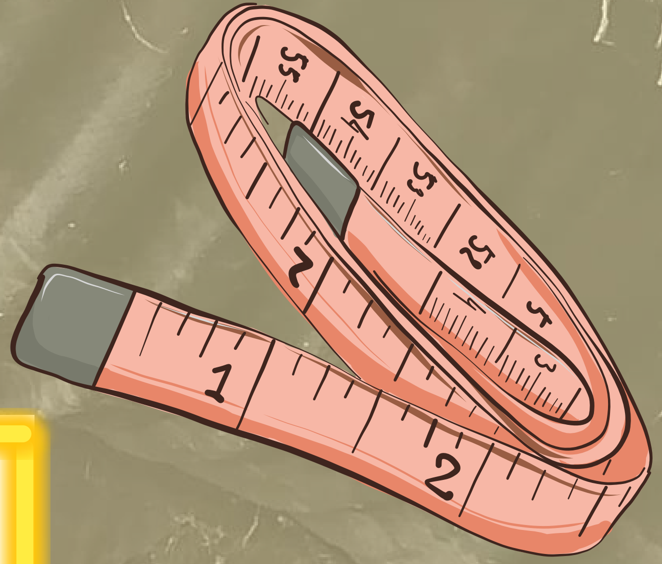


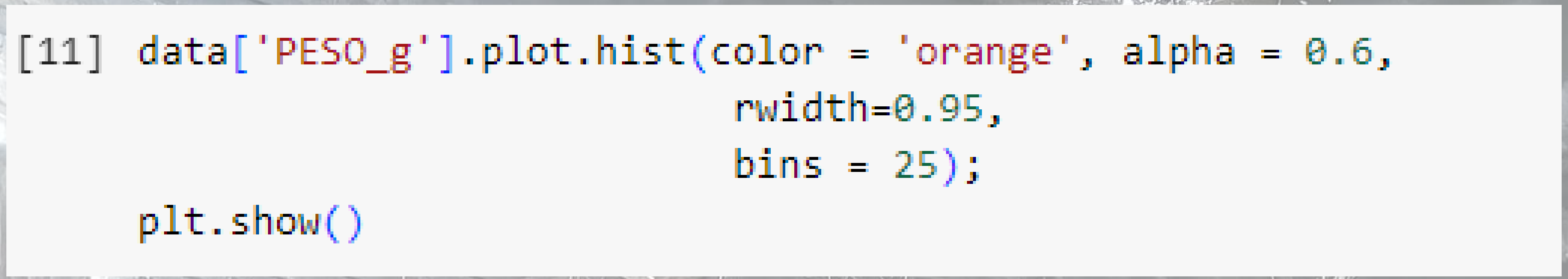




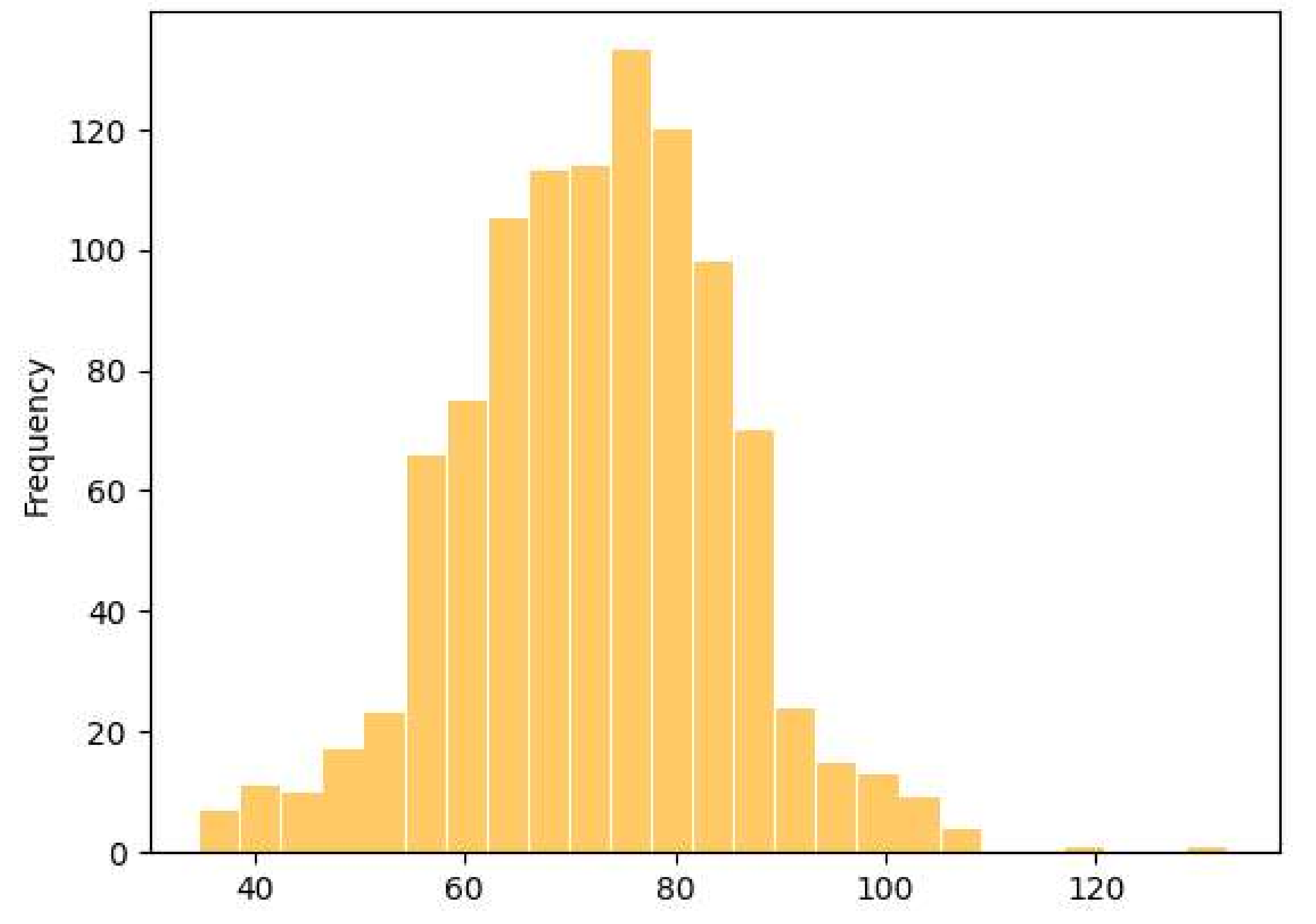
```
[10] data['ANCHO'].plot.box();
     plt.show()
```



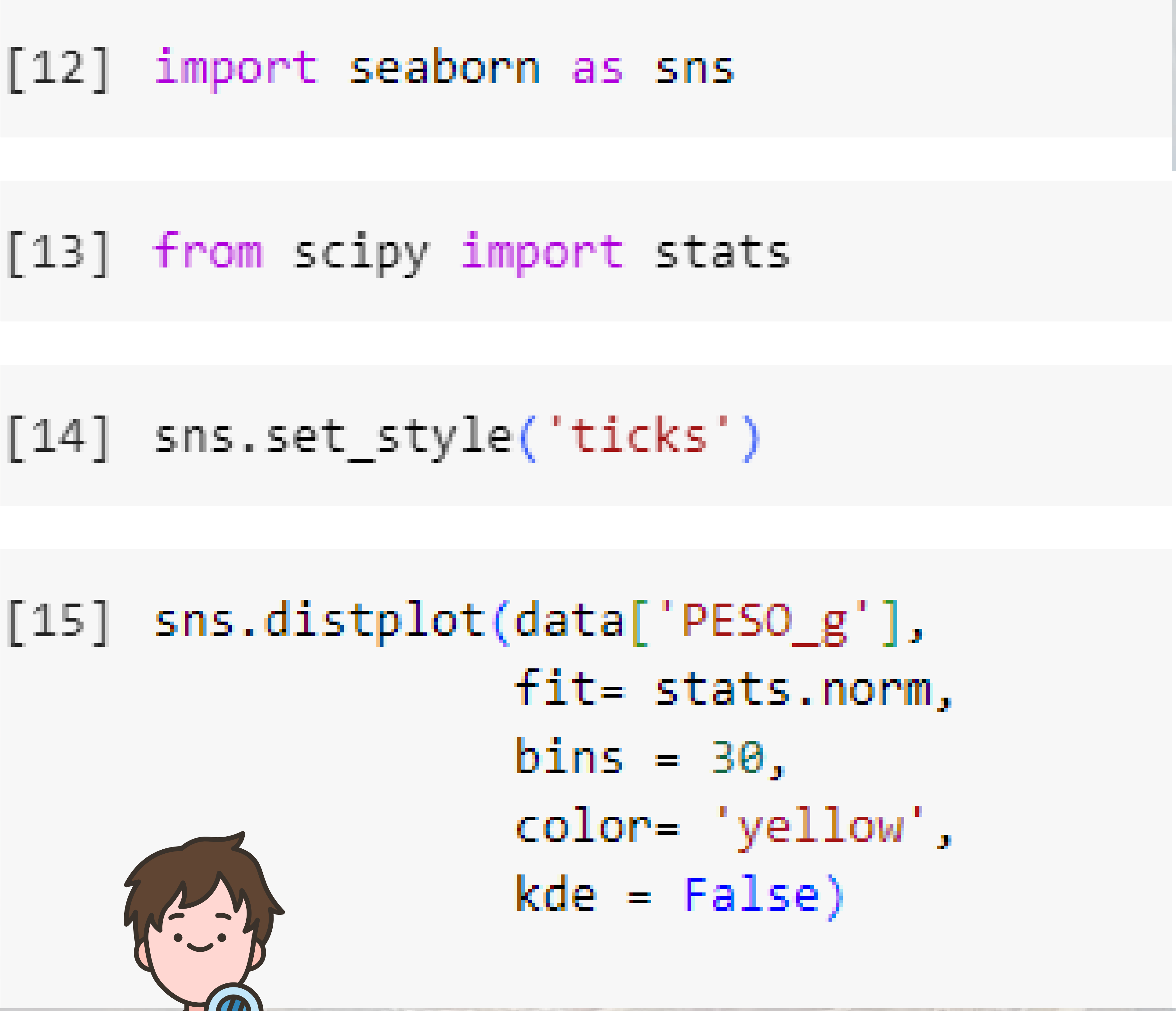




```
[11]  data['PESO_g'].plot.hist(color = 'orange', alpha = 0.6,
                                rwidth=0.95,
                                bins = 25);

      plt.show()
```



GRUPO 3

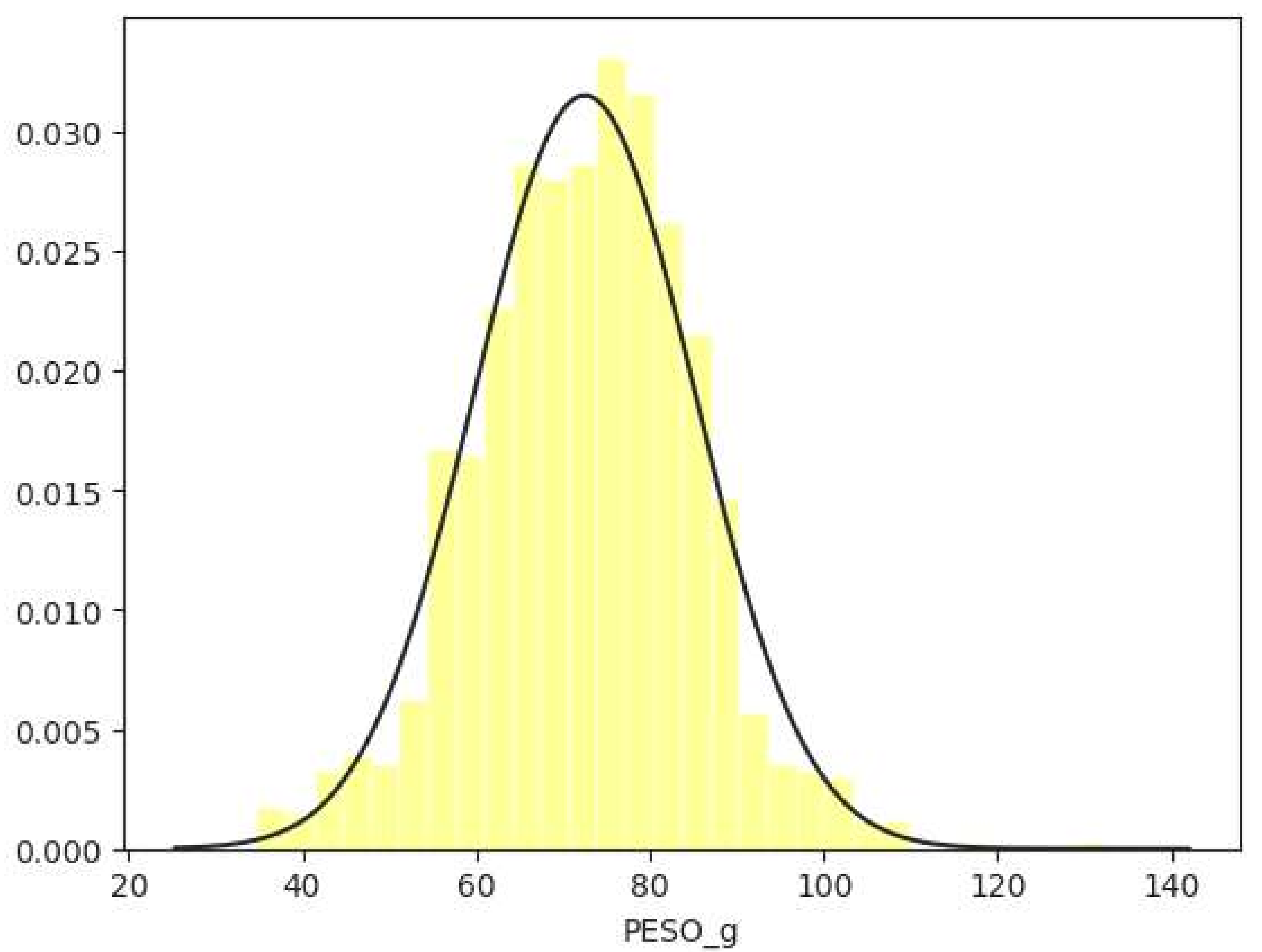


```
[12]  import seaborn as sns
```

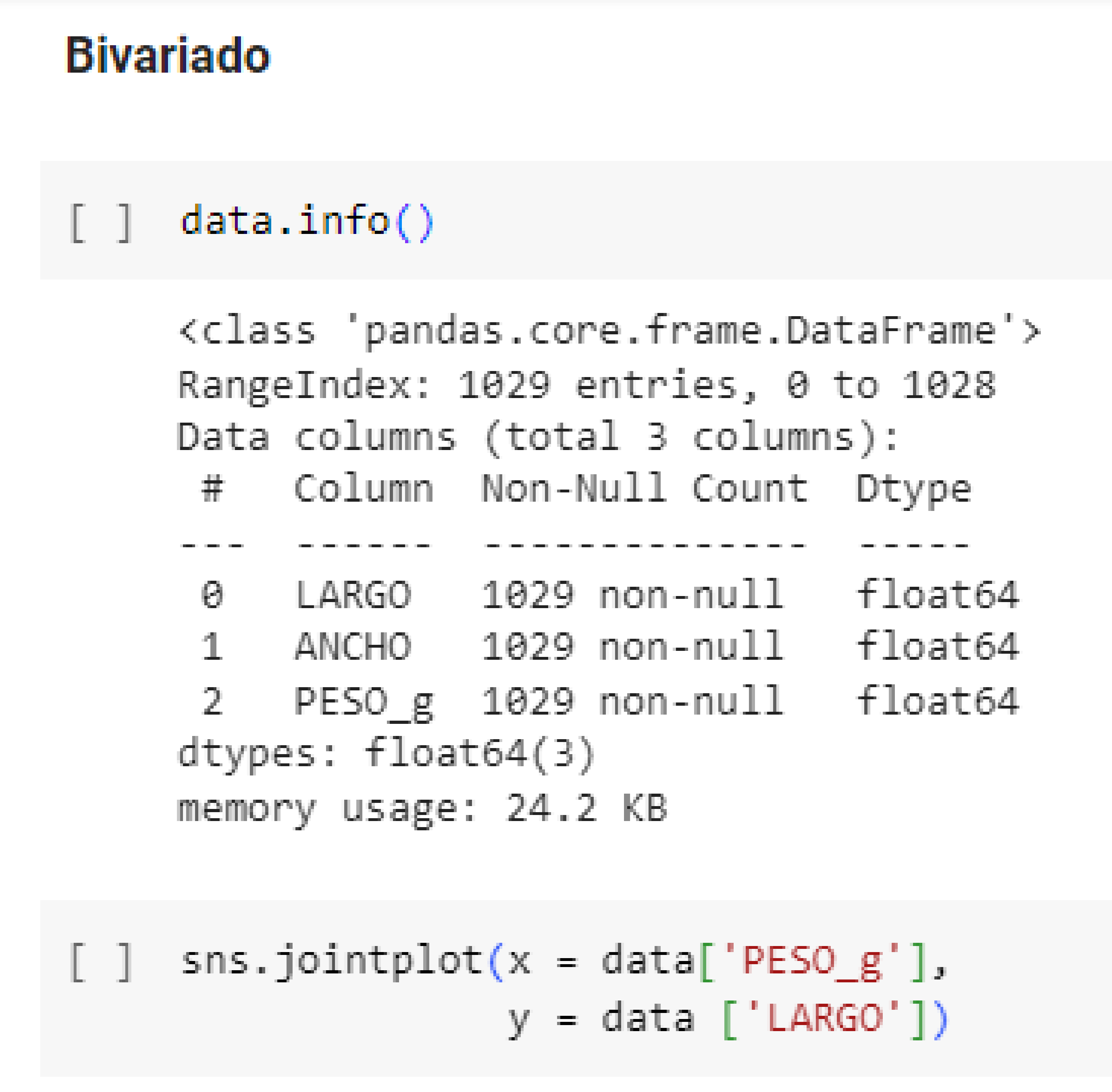
```
[13]  from scipy import stats
```

```
[14]  sns.set_style('ticks')
```

```
[15]  sns.distplot(data['PESO_g'],
                   fit= stats.norm,
                   bins = 30,
                   color= 'yellow',
                   kde = False)
```



UNIVERSIDAD CÉSAR VALLEJO

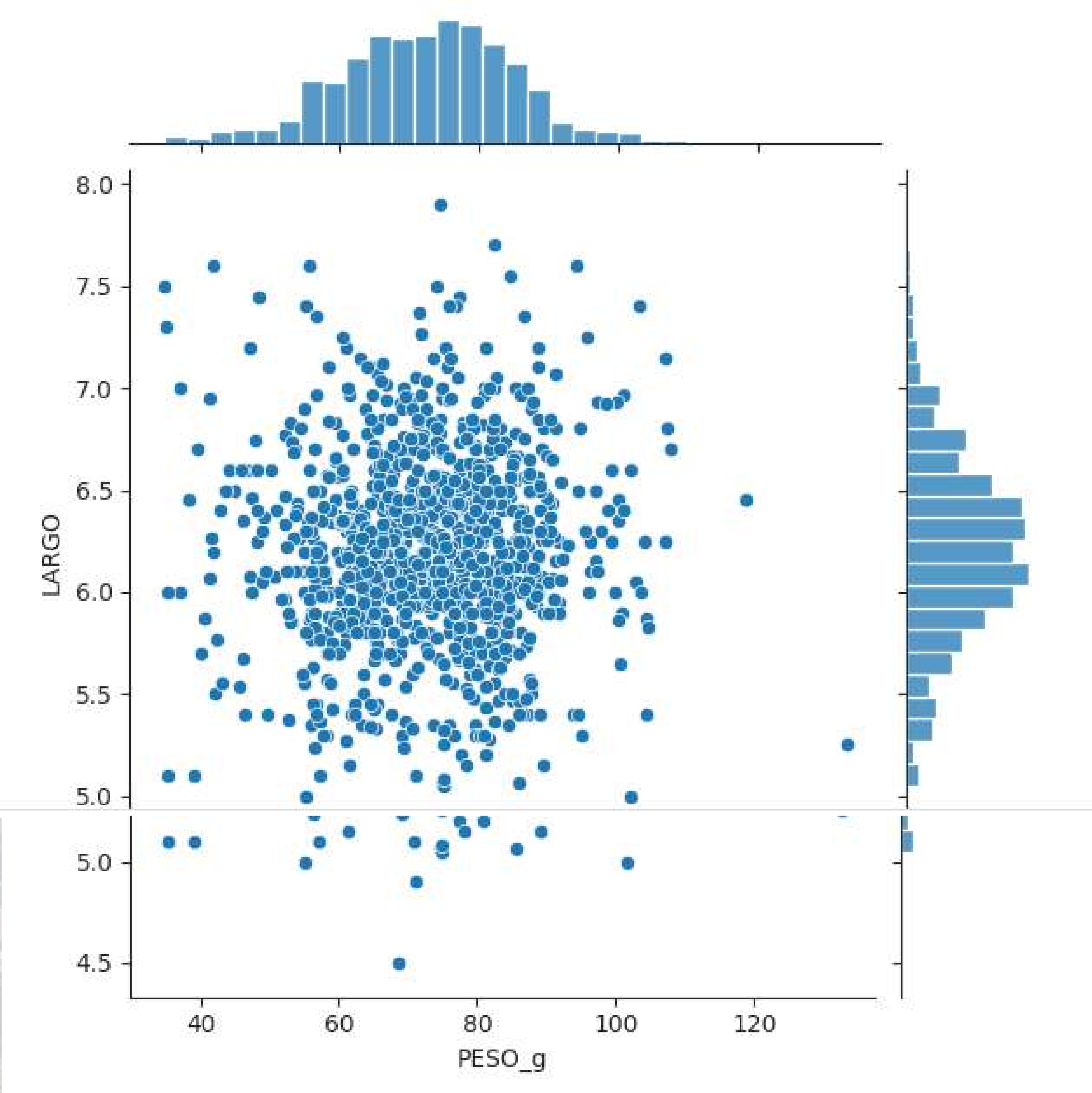


Bivariado

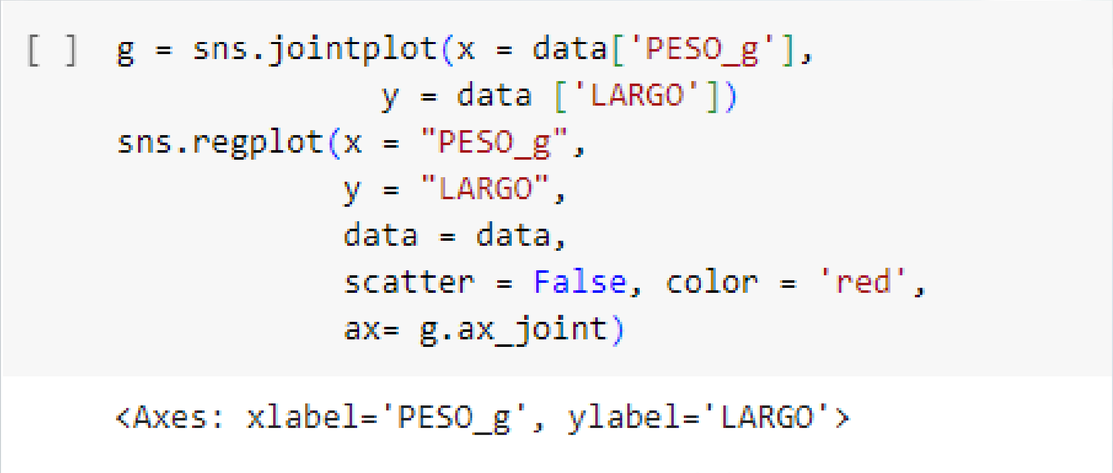
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1029 entries, 0 to 1028
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   LARGO   1029 non-null   float64
 1   ANCHO   1029 non-null   float64
 2   PESO_g  1029 non-null   float64
dtypes: float64(3)
memory usage: 24.2 KB
```

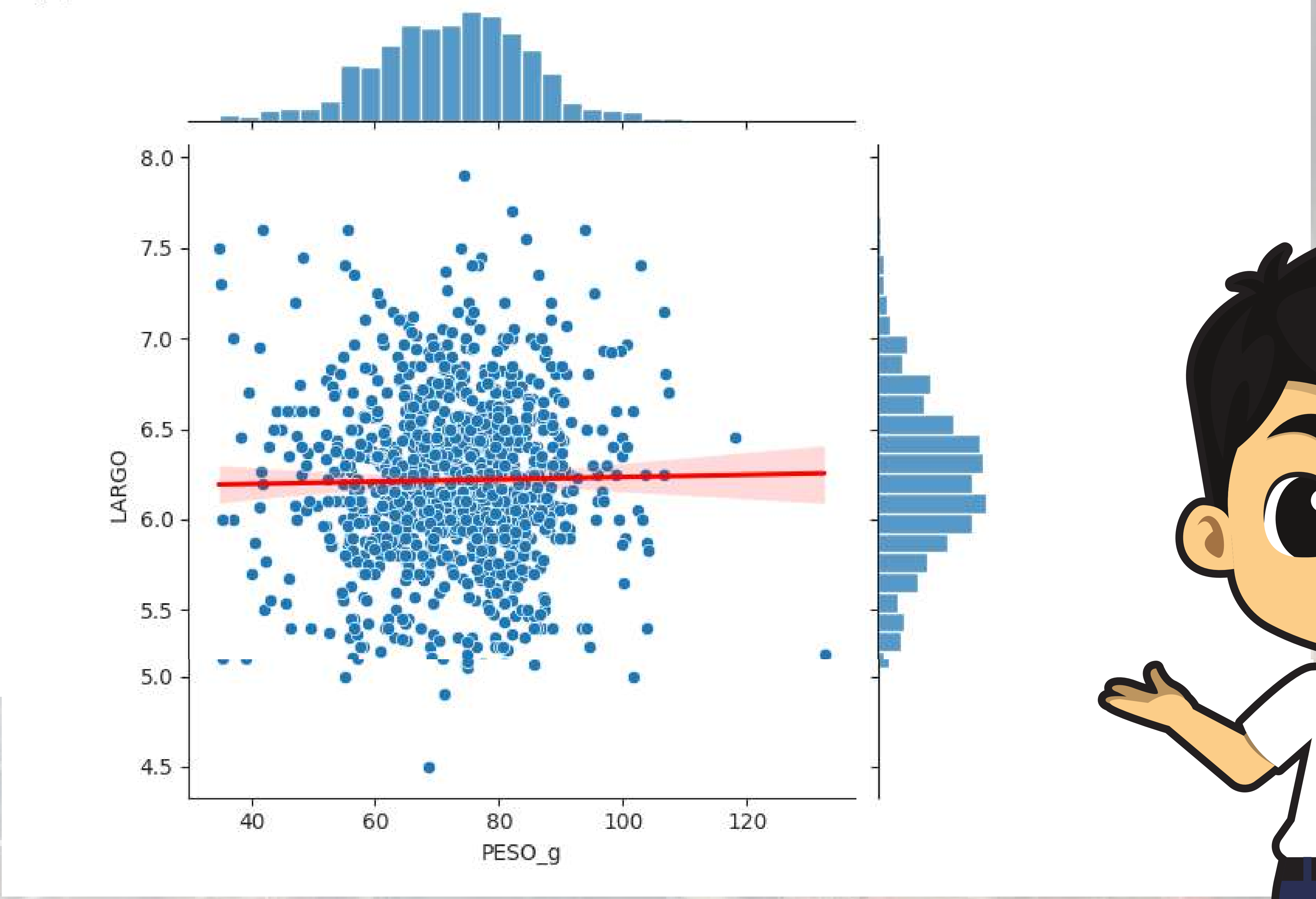
```
sns.jointplot(x = data['PESO_g'],
              y = data ['LARGO'])
```

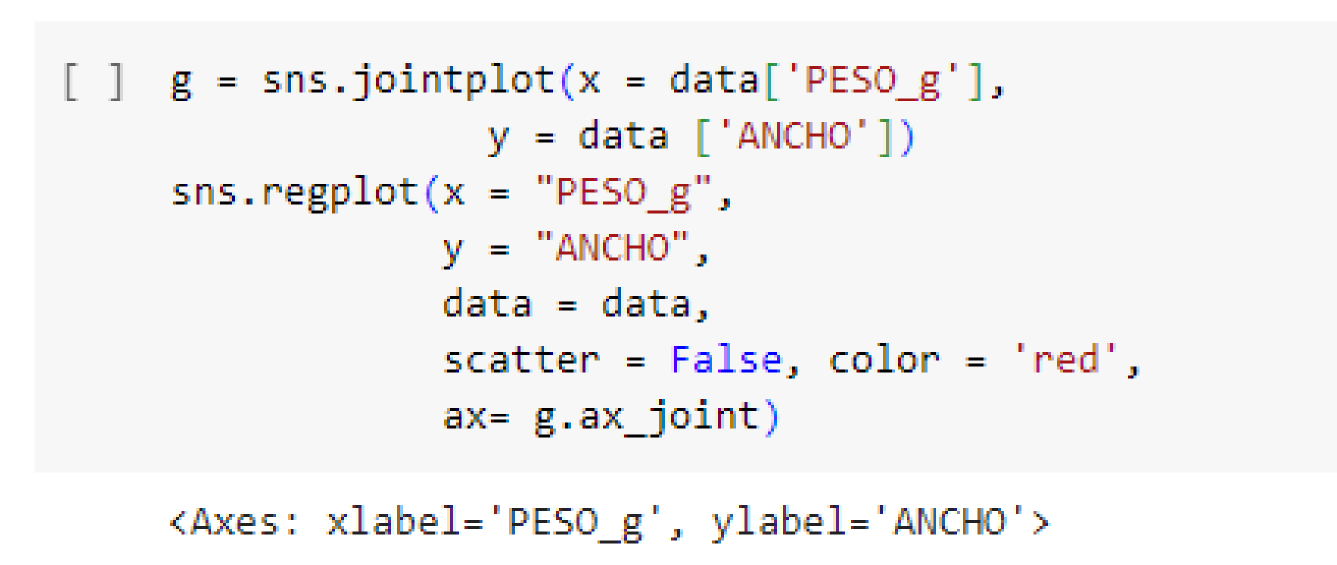


GRUPO 3

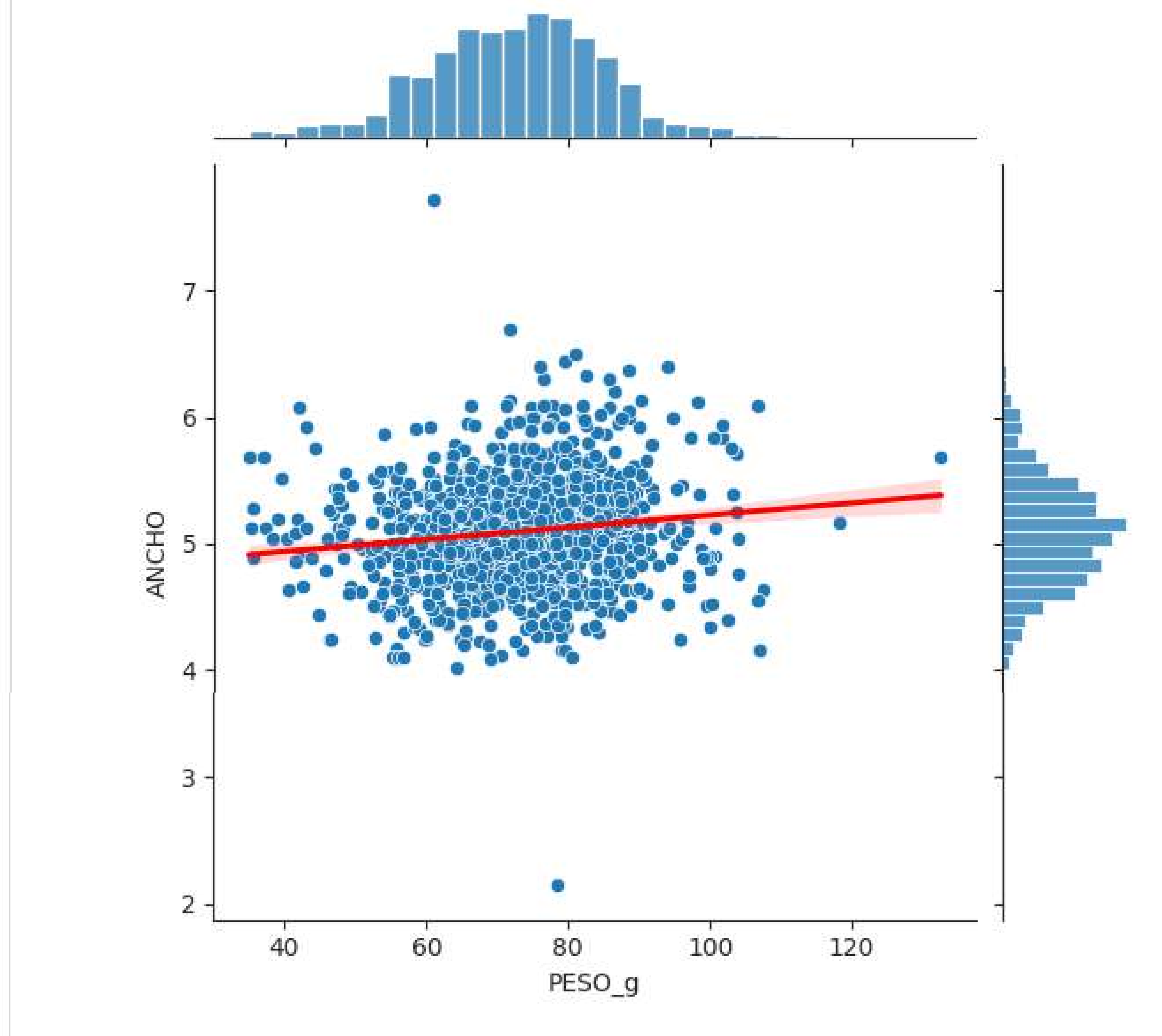


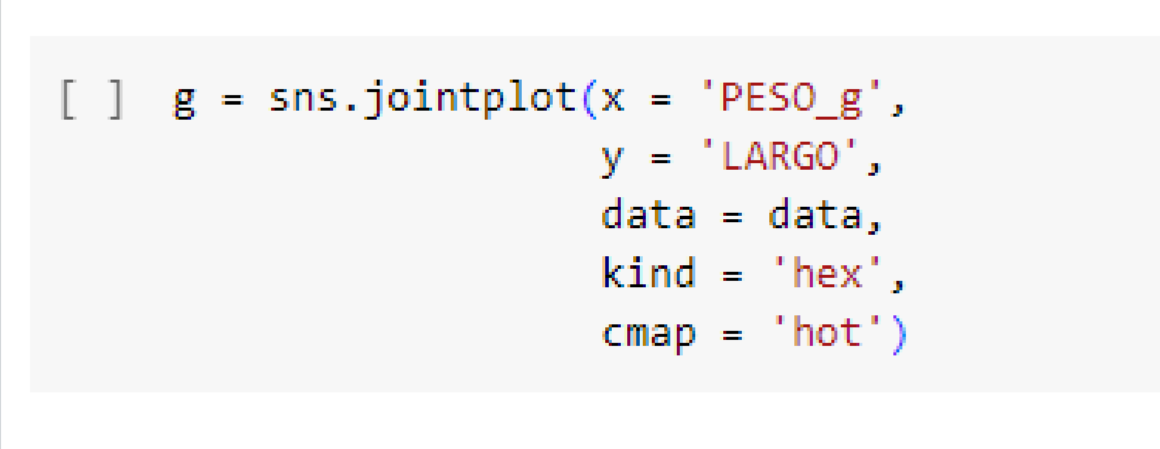
```
[ ]  g = sns.jointplot(x = data['PESO_g'],
                       y = data ['LARGO'])
     sns.regplot(x = "PESO_g",
                 y = "LARGO",
                 data = data,
                 scatter = False, color = 'red',
                 ax= g.ax_joint)

     <Axes: xlabel='PESO_g', ylabel='LARGO'>
```



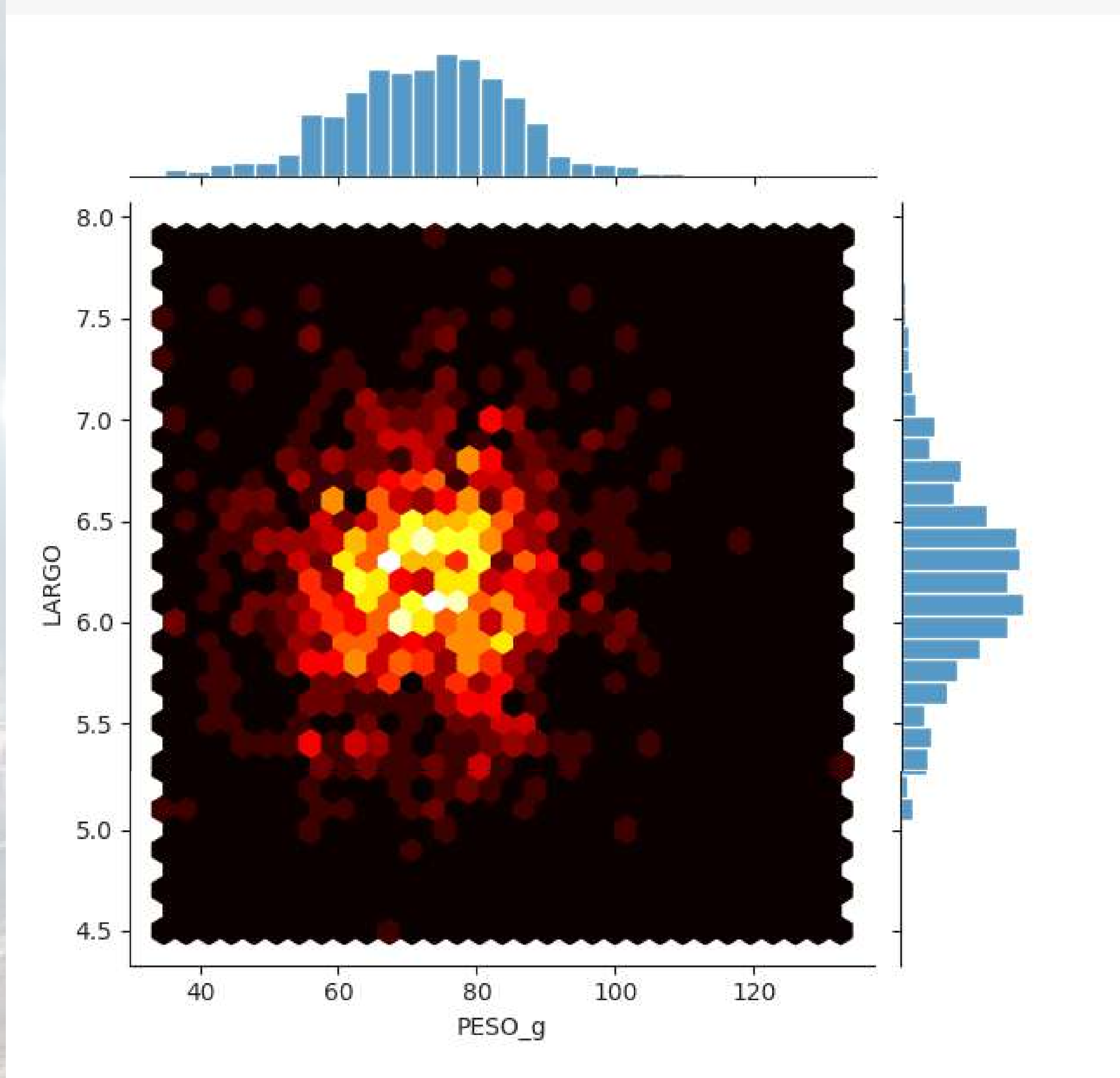


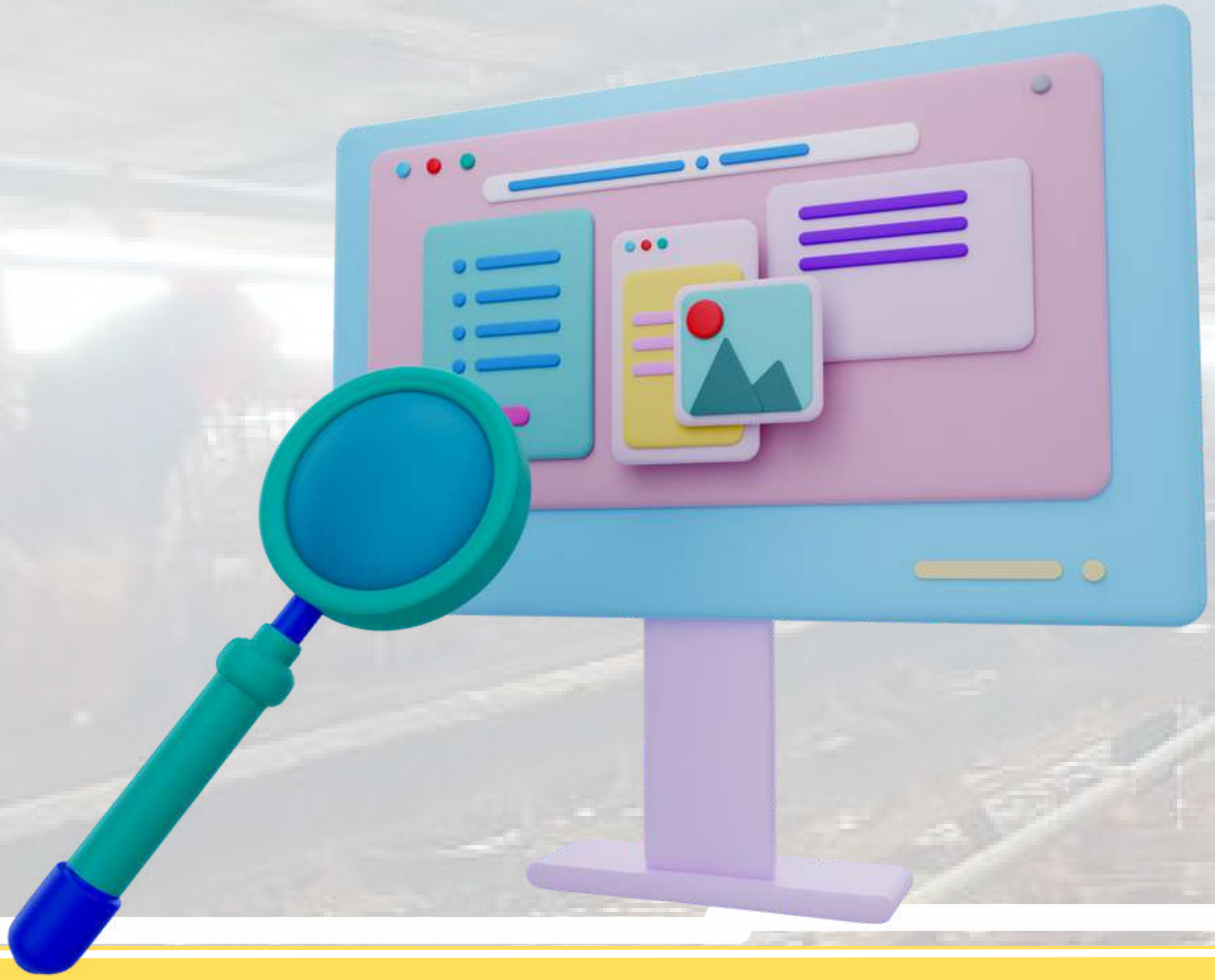
```
[ ]  g = sns.jointplot(x = data['PESO_g'],
                       y = data ['ANCHO'])
     sns.regplot(x = "PESO_g",
                 y = "ANCHO",
                 data = data,
                 scatter = False, color = 'red',
                 ax= g.ax_joint)

     <Axes: xlabel='PESO_g', ylabel='ANCHO'>
```

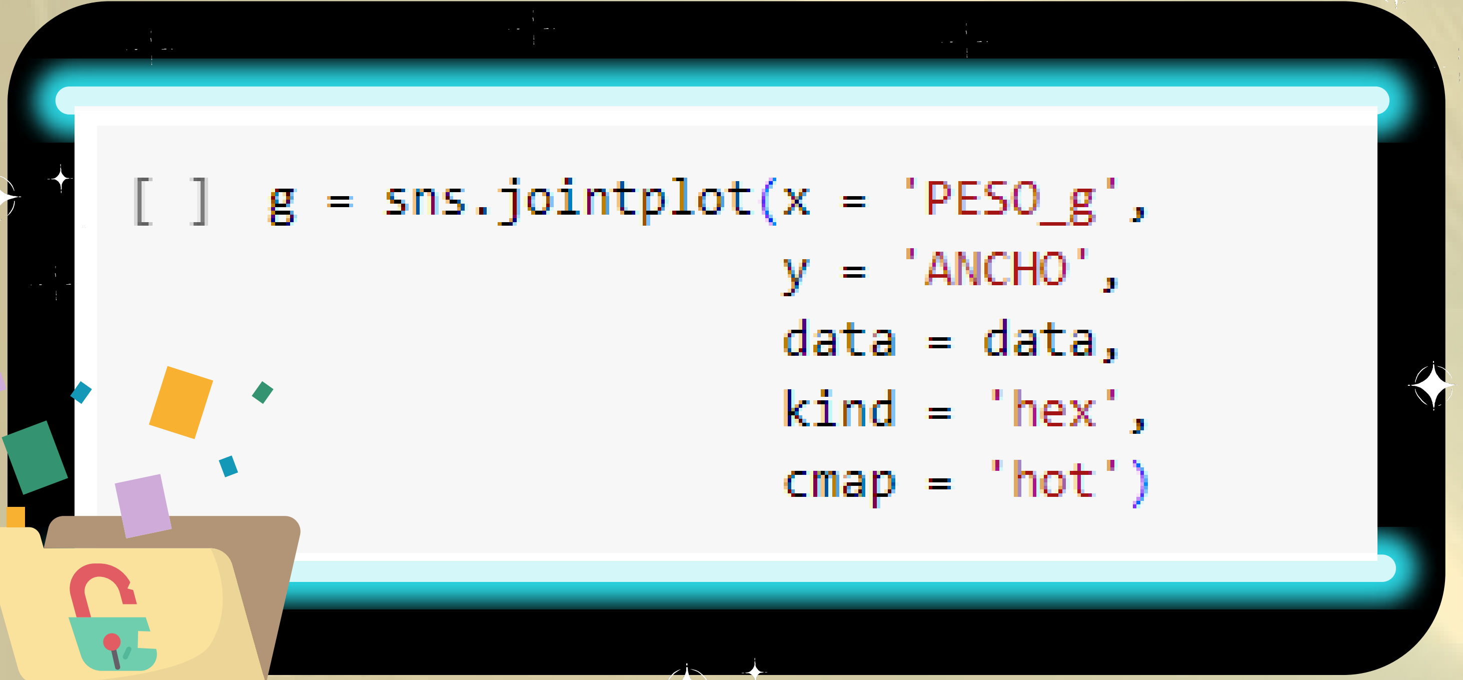




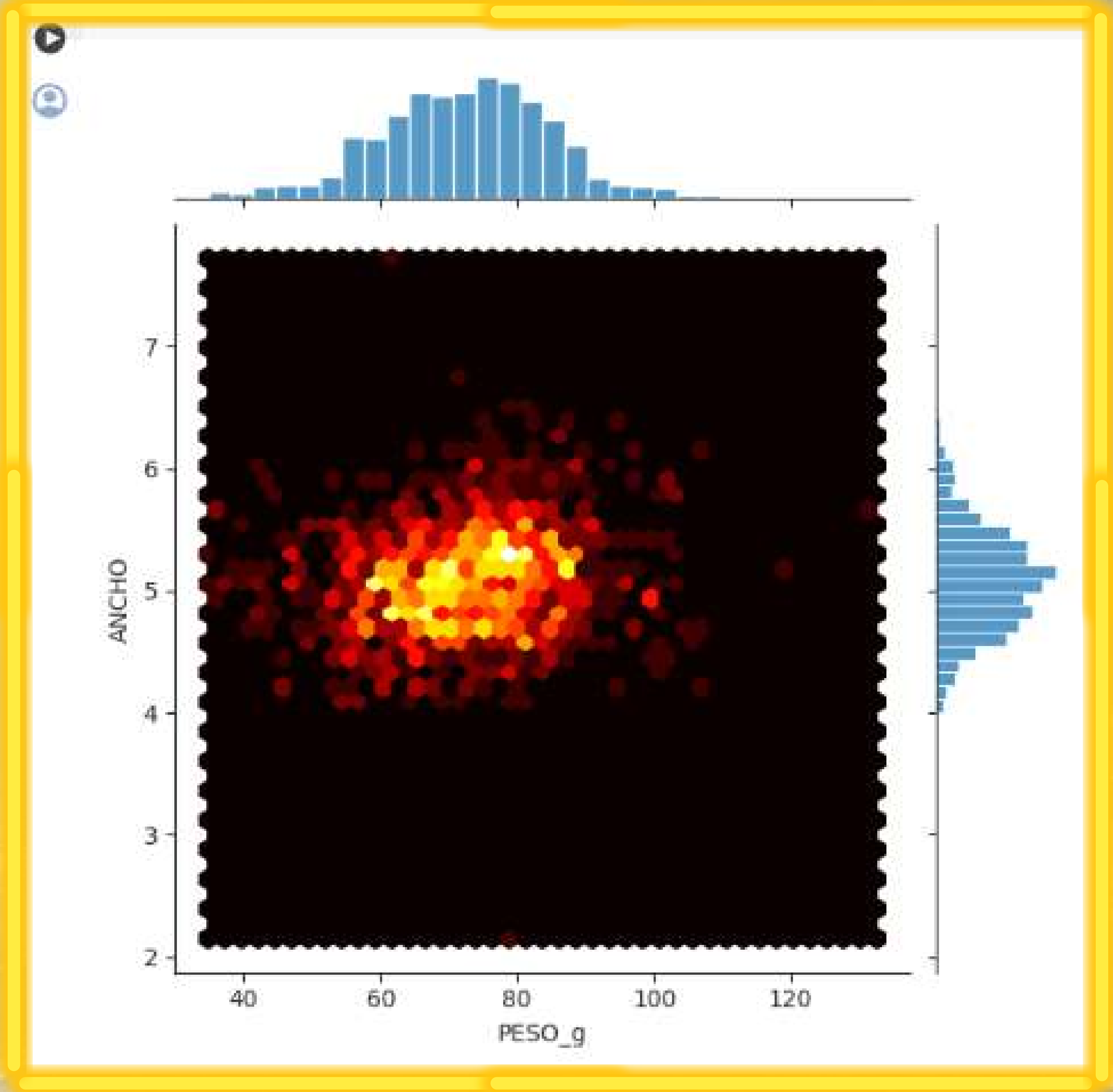
```
[ ]  g = sns.jointplot(x = 'PESO_g',
                       y = 'LARGO',
                       data = data,
                       kind = 'hex',
                       cmap = 'hot')
```





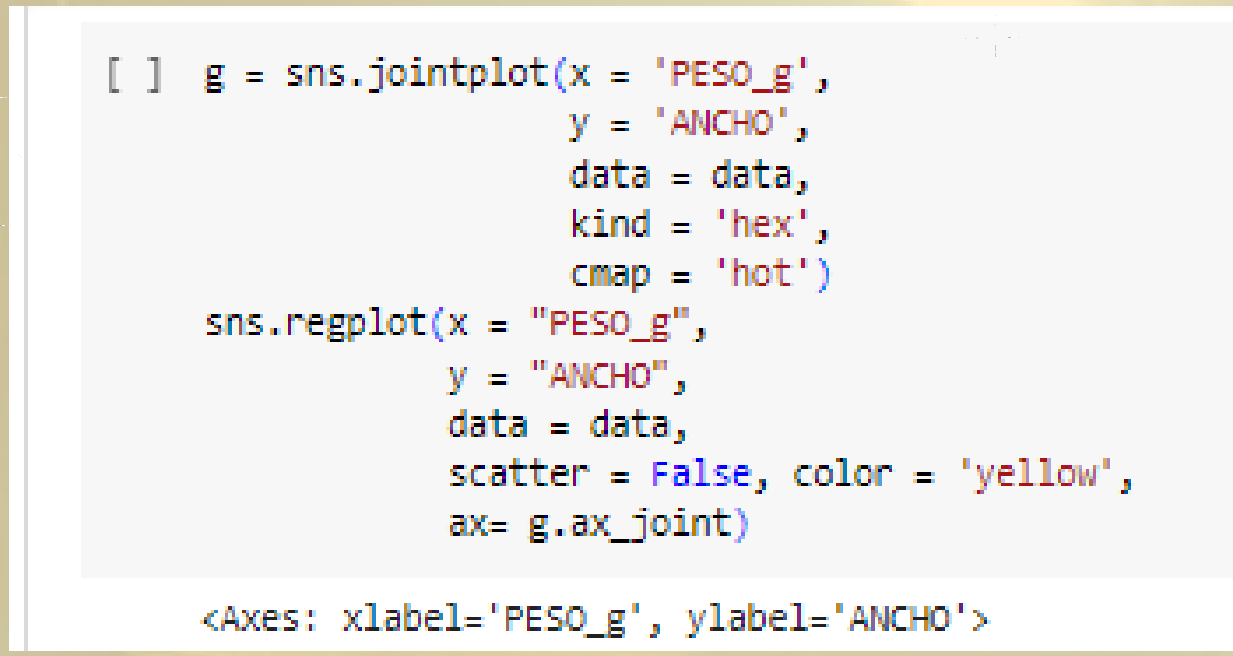


```
[ ]  g = sns.jointplot(x = 'PESO_g',
                       y = 'ANCHO',
                       data = data,
                       kind = 'hex',
                       cmap = 'hot')
```

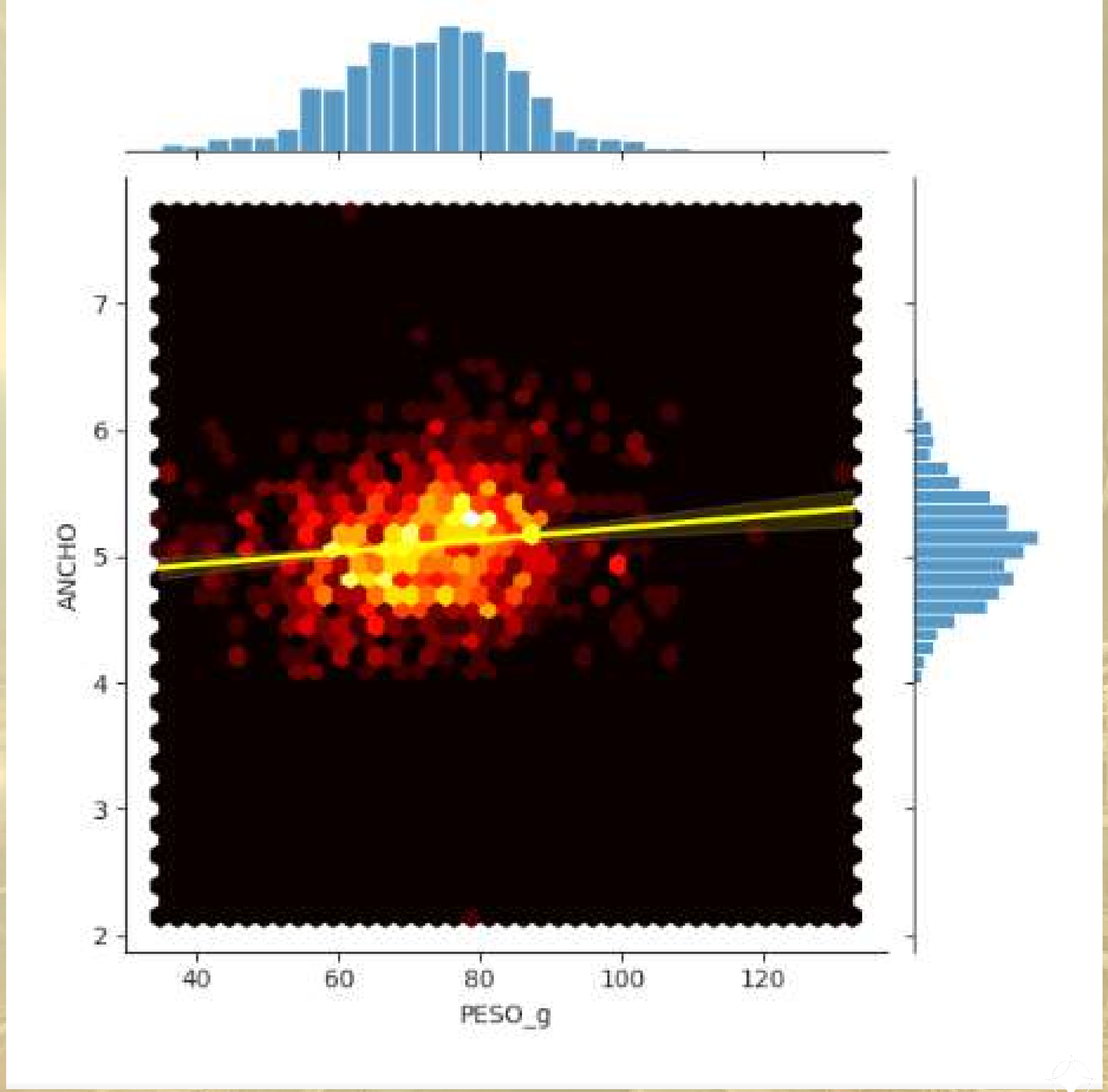


UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

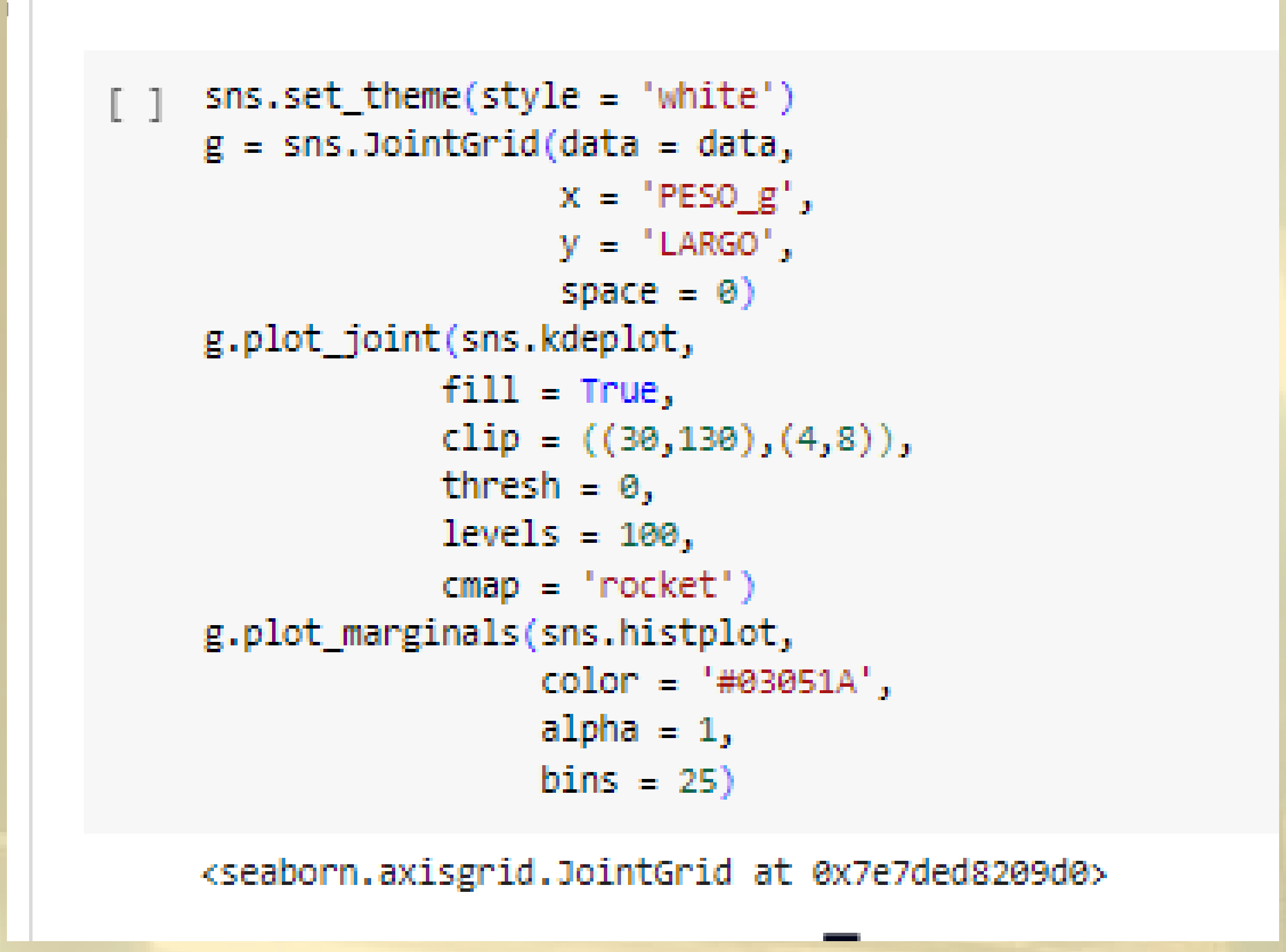


```
[ ]  g = sns.jointplot(x = 'PESO_g',
                       y = 'ANCHO',
                       data = data,
                       kind = 'hex',
                       cmap = 'hot')
     sns.regplot(x = "PESO_g",
                 y = "ANCHO",
                 data = data,
                 scatter = False, color = 'yellow',
                 ax= g.ax_joint)

     <Axes: xlabel='PESO_g', ylabel='ANCHO'>
```

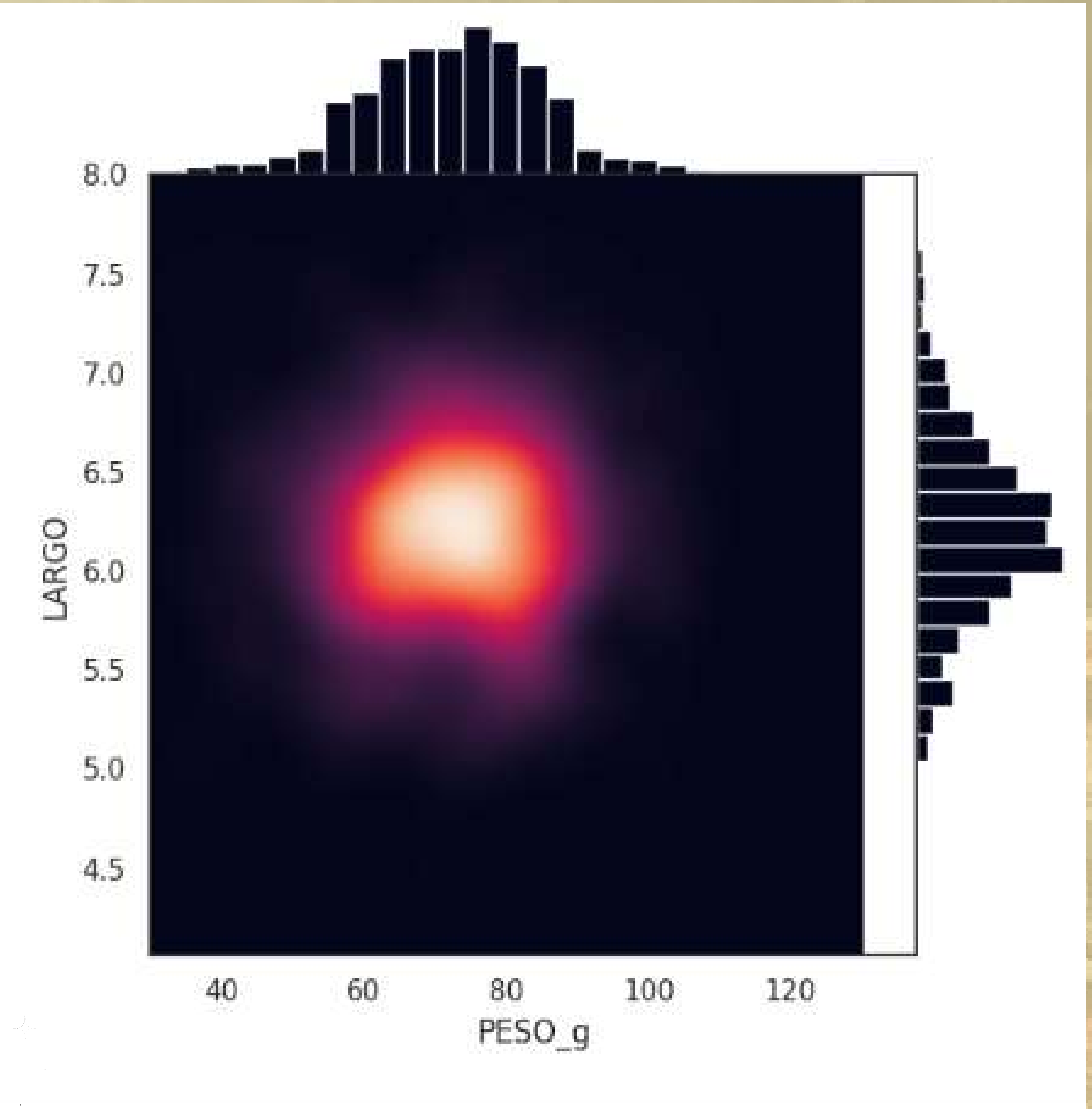


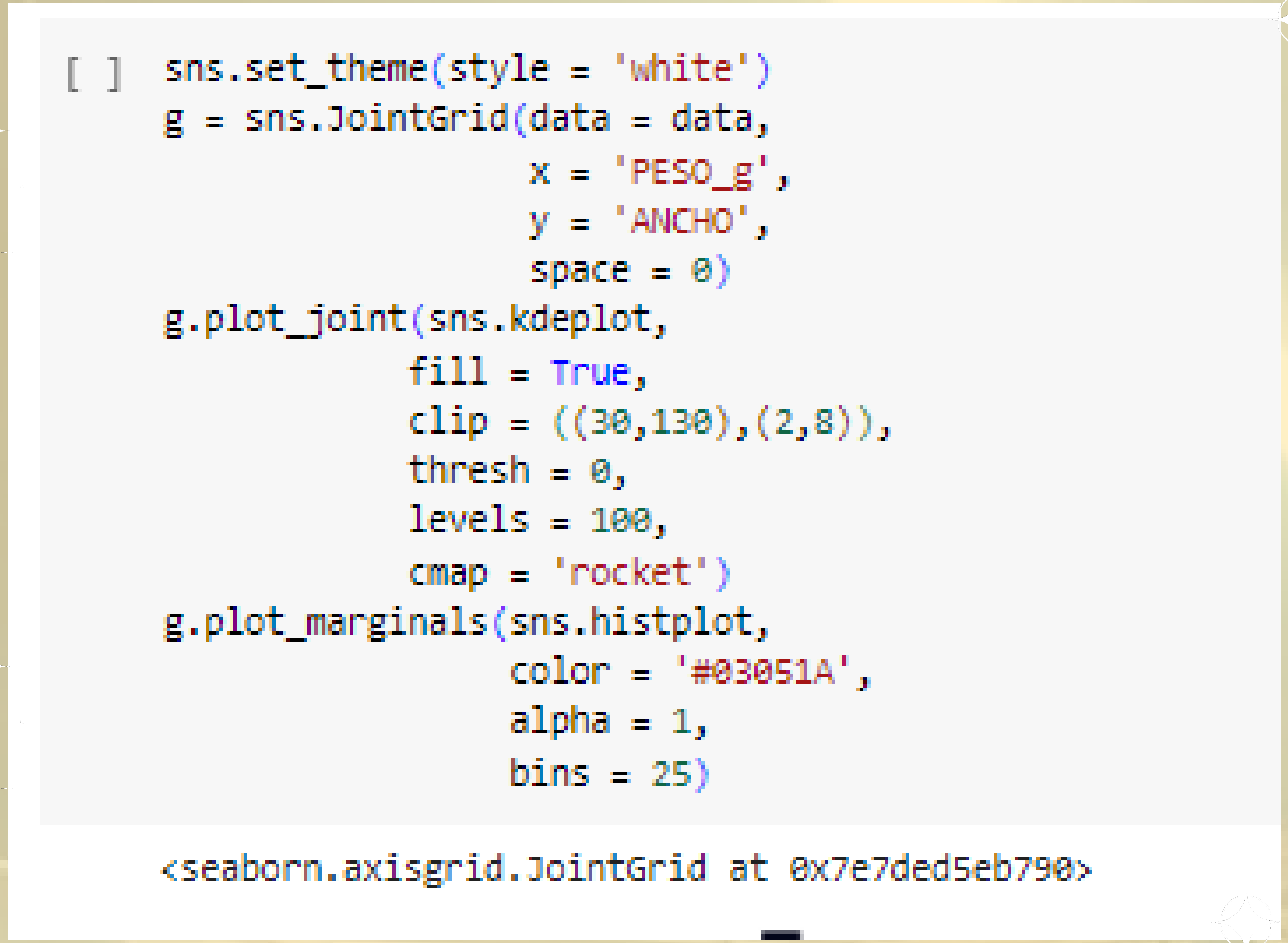
UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3

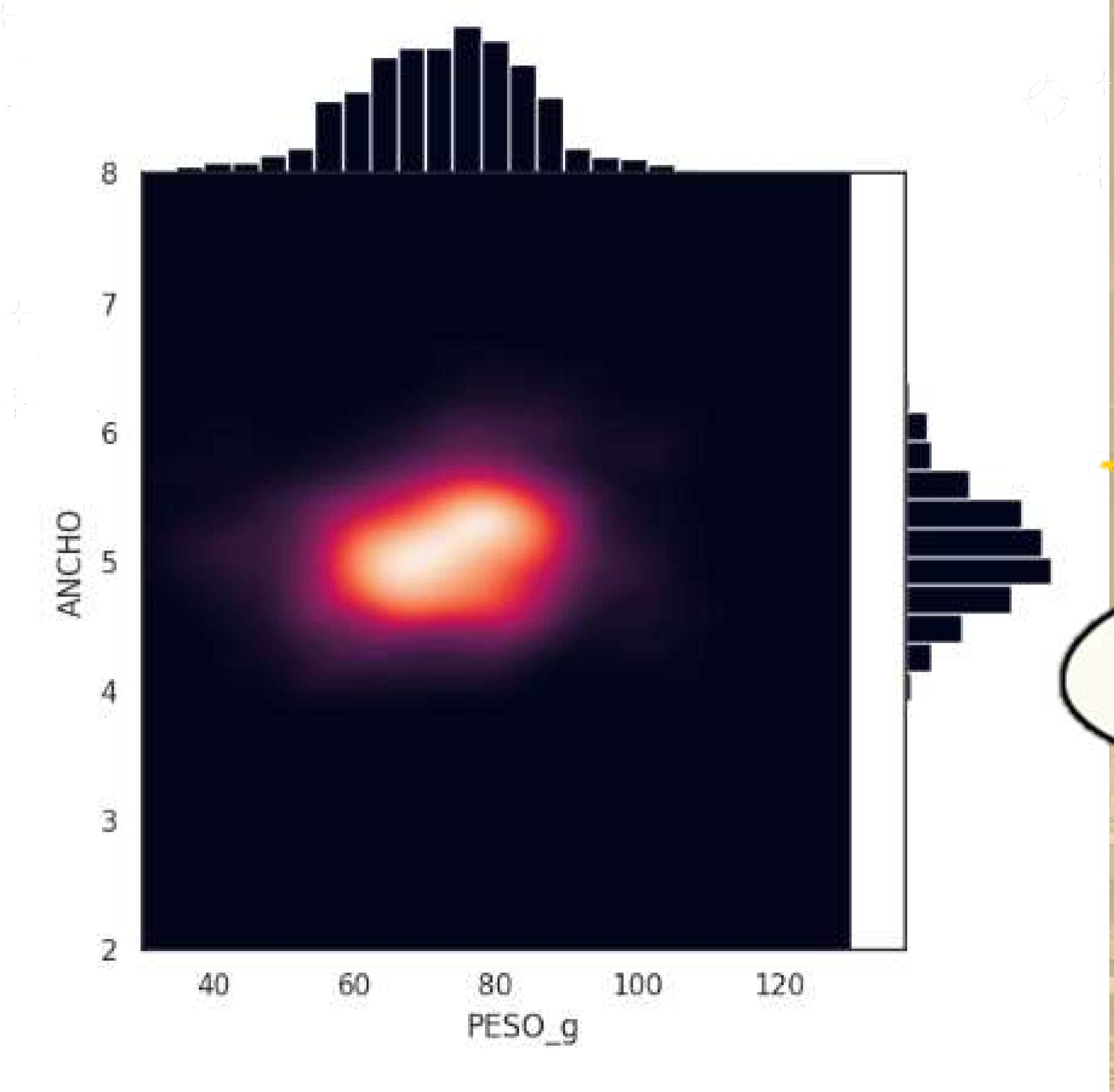


```
[ ]  sns.set_theme(style = 'white')
     g = sns.JointGrid(data = data,
                       x = 'PESO_g',
                       y = 'LARGO',
                       space = 0)
     g.plot_joint(sns.kdeplot,
                  fill = True,
                  clip = ((30,130),(4,8)),
                  thresh = 0,
                  levels = 100,
                  cmap = 'rocket')
     g.plot_marginals(sns.histplot,
                      color = '#03051A',
                      alpha = 1,
                      bins = 25)

<seaborn.axisgrid.JointGrid at 0x7e7ded8209d0>
```



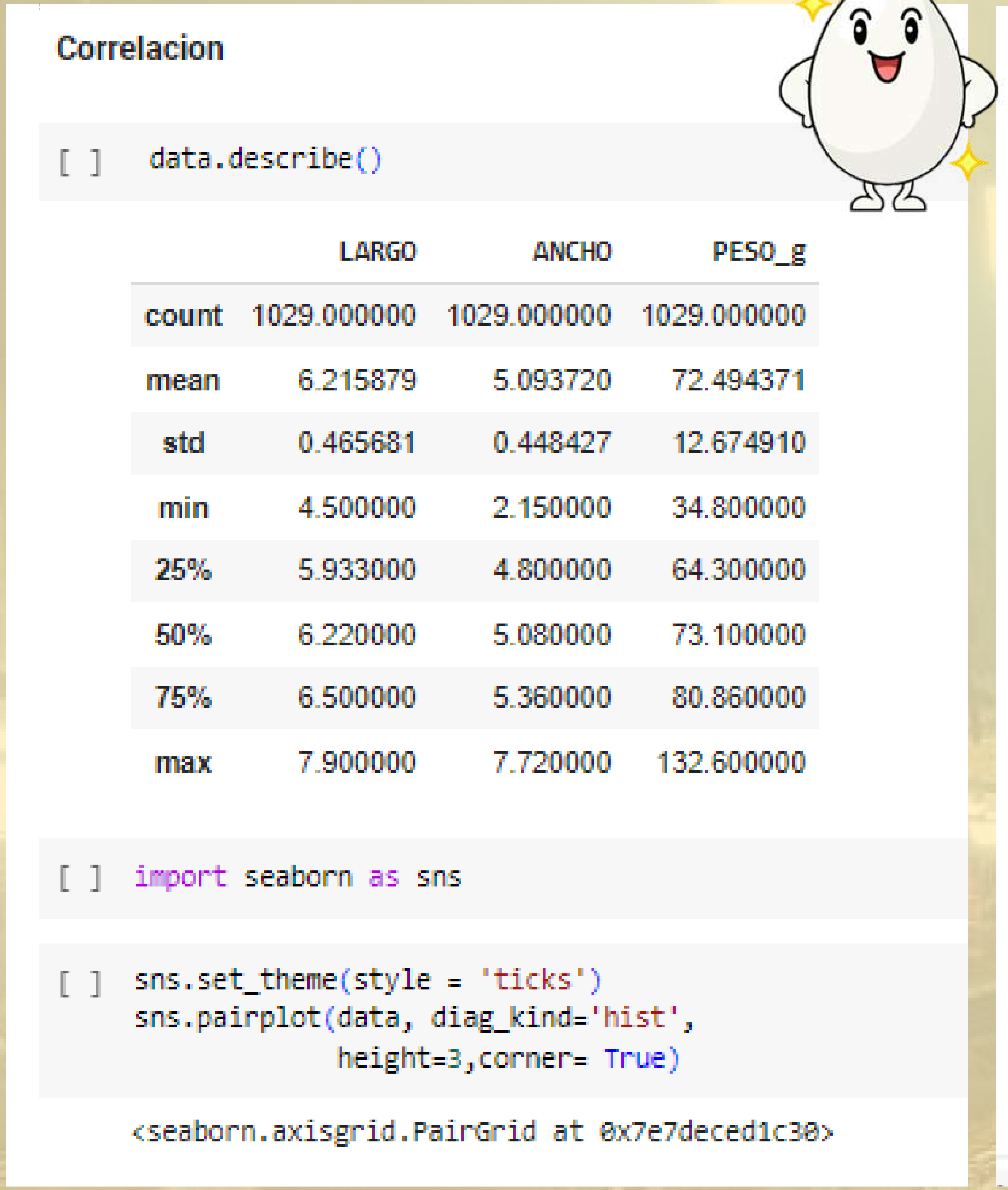


```
[ ]  sns.set_theme(style = 'white')
     g = sns.JointGrid(data = data,
                       x = 'PESO_g',
                       y = 'ANCHO',
                       space = 0)
     g.plot_joint(sns.kdeplot,
                  fill = True,
                  clip = ((30,130),(2,8)),
                  thresh = 0,
                  levels = 100,
                  cmap = 'rocket')
     g.plot_marginals(sns.histplot,
                      color = '#03051A',
                      alpha = 1,
                      bins = 25)

     <seaborn.axisgrid.JointGrid at 0x7e7ded5eb790>
```



UNIVERSIDAD CÉSAR VALLEJO

GRUPO 3



Correlacion

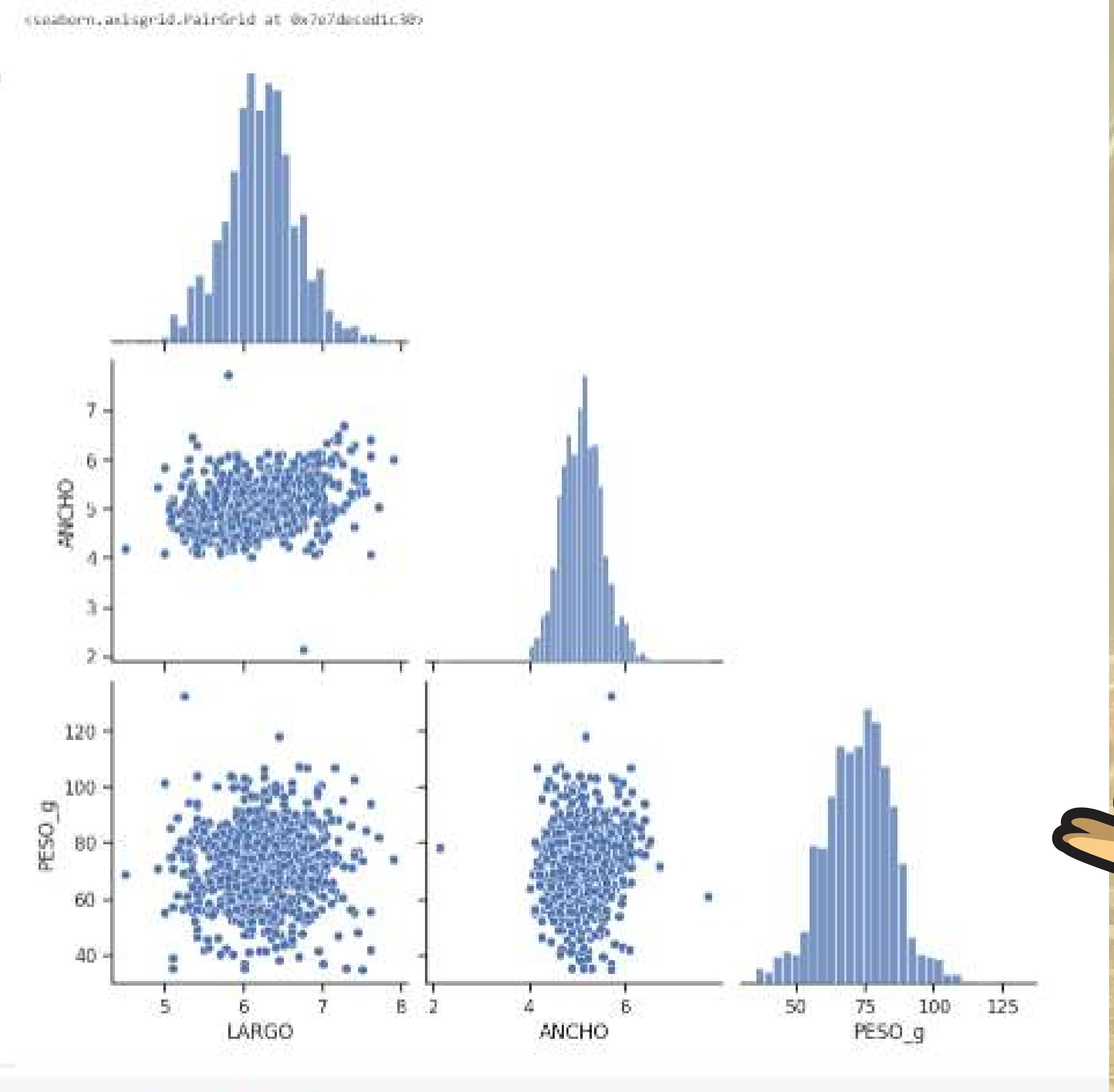
```
[ ]  data.describe()
```

	LARGO	ANCHO	PESO_g
count	1029.000000	1029.000000	1029.000000
mean	6.215879	5.093720	72.494371
std	0.465681	0.448427	12.674910
min	4.500000	2.150000	34.800000
25%	5.933000	4.800000	64.300000
50%	6.220000	5.080000	73.100000
75%	6.500000	5.360000	80.860000
max	7.900000	7.720000	132.600000

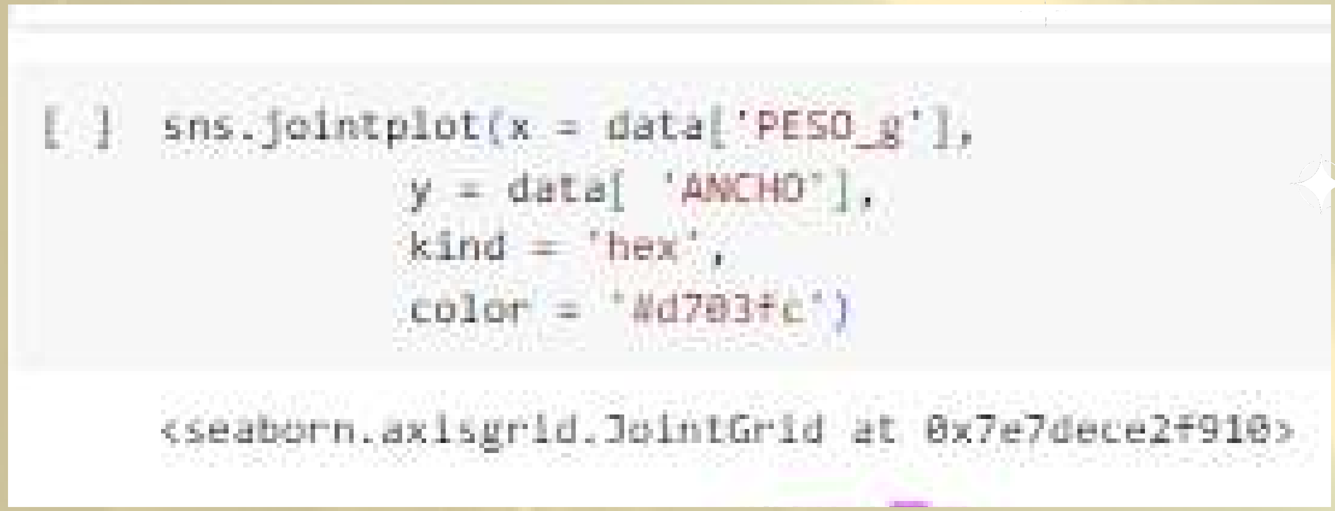
```
[ ]  import seaborn as sns
```

```
[ ]  sns.set_theme(style = 'ticks')
     sns.pairplot(data, diag_kind='hist',
                  height=3,corner= True)
```

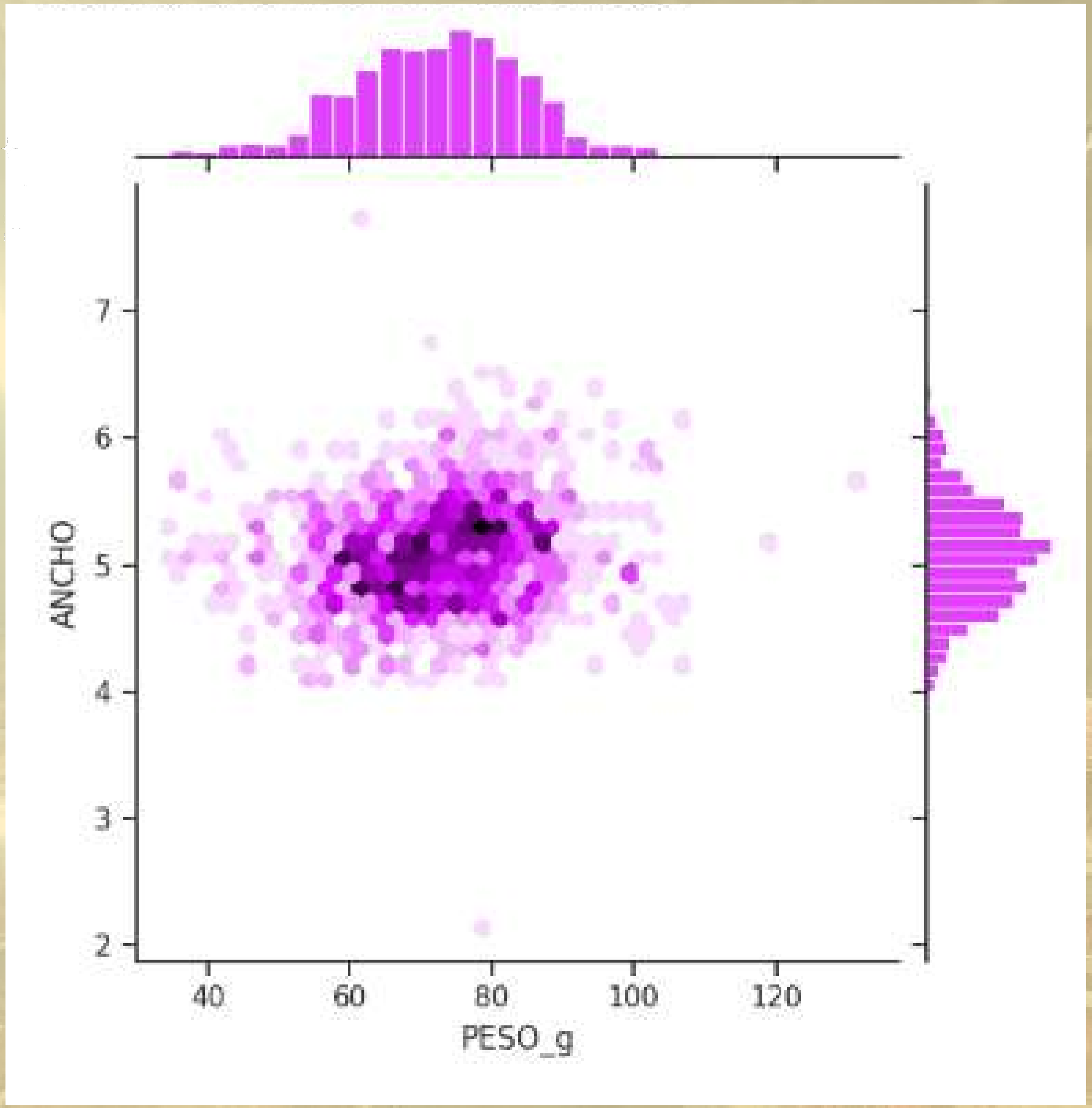
```
<seaborn.axisgrid.PairGrid at 0x7e7deced1c30>
```

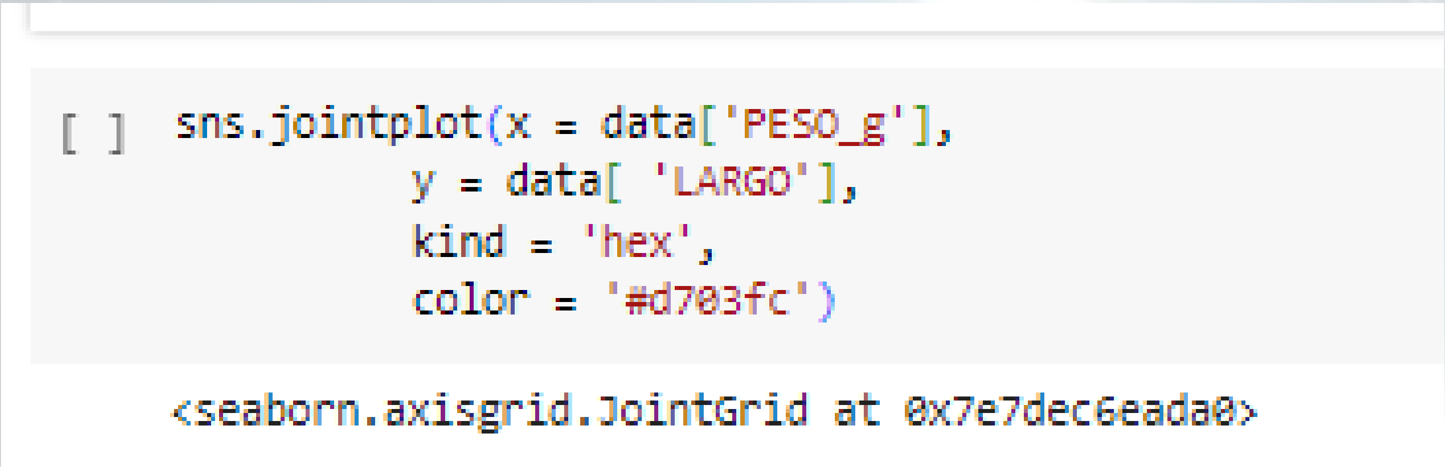


```
<seaborn.axisgrid.PairGrid at 0x7e7deced1c30>
```

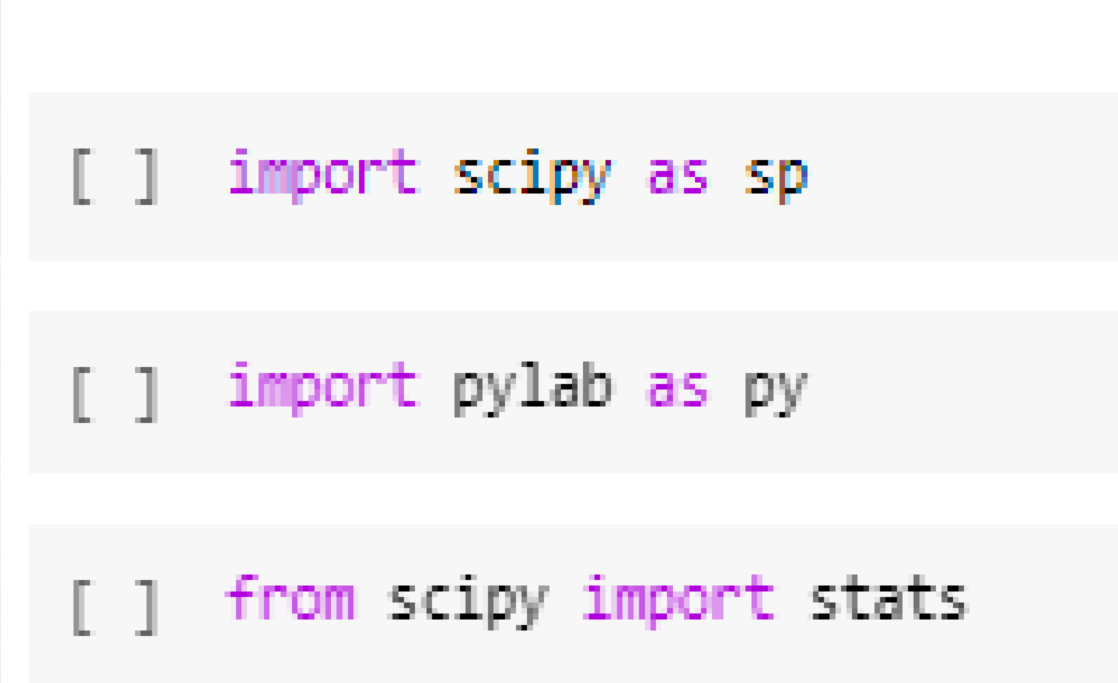


```
[ ]  sns.jointplot(x = data['PESO_g'],
                   y = data[ 'ANCHO'],
                   kind = 'hex',
                   color = '#d703fc')

     <seaborn.axisgrid.JointGrid at 0x7e7dece2f910>
```





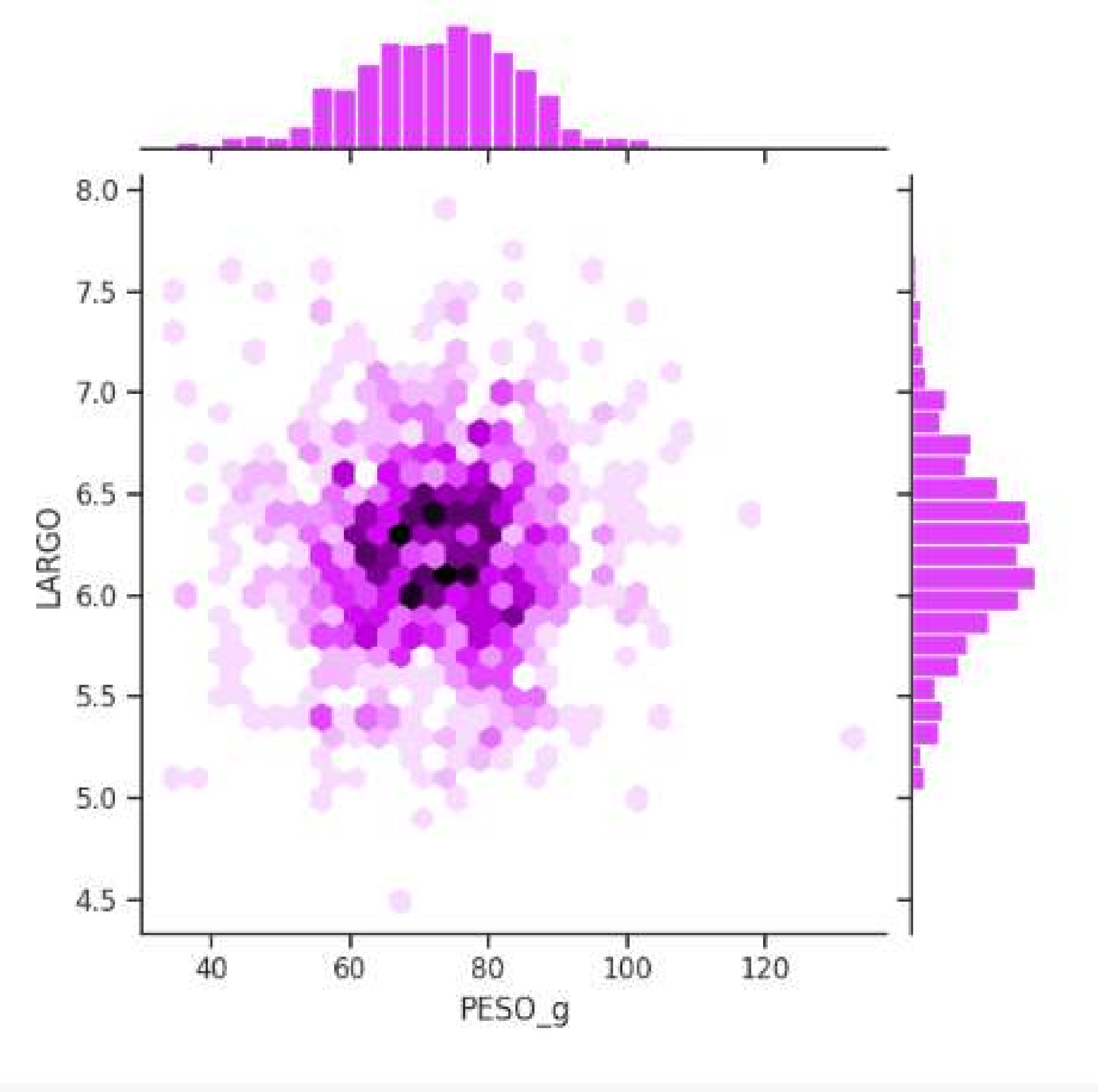
```
[ ]  sns.jointplot(x = data['PESO_g'],
                   y = data[ 'LARGO'],
                   kind = 'hex',
                   color = '#d703fc')

     <seaborn.axisgrid.JointGrid at 0x7e7dec6eada0>
```

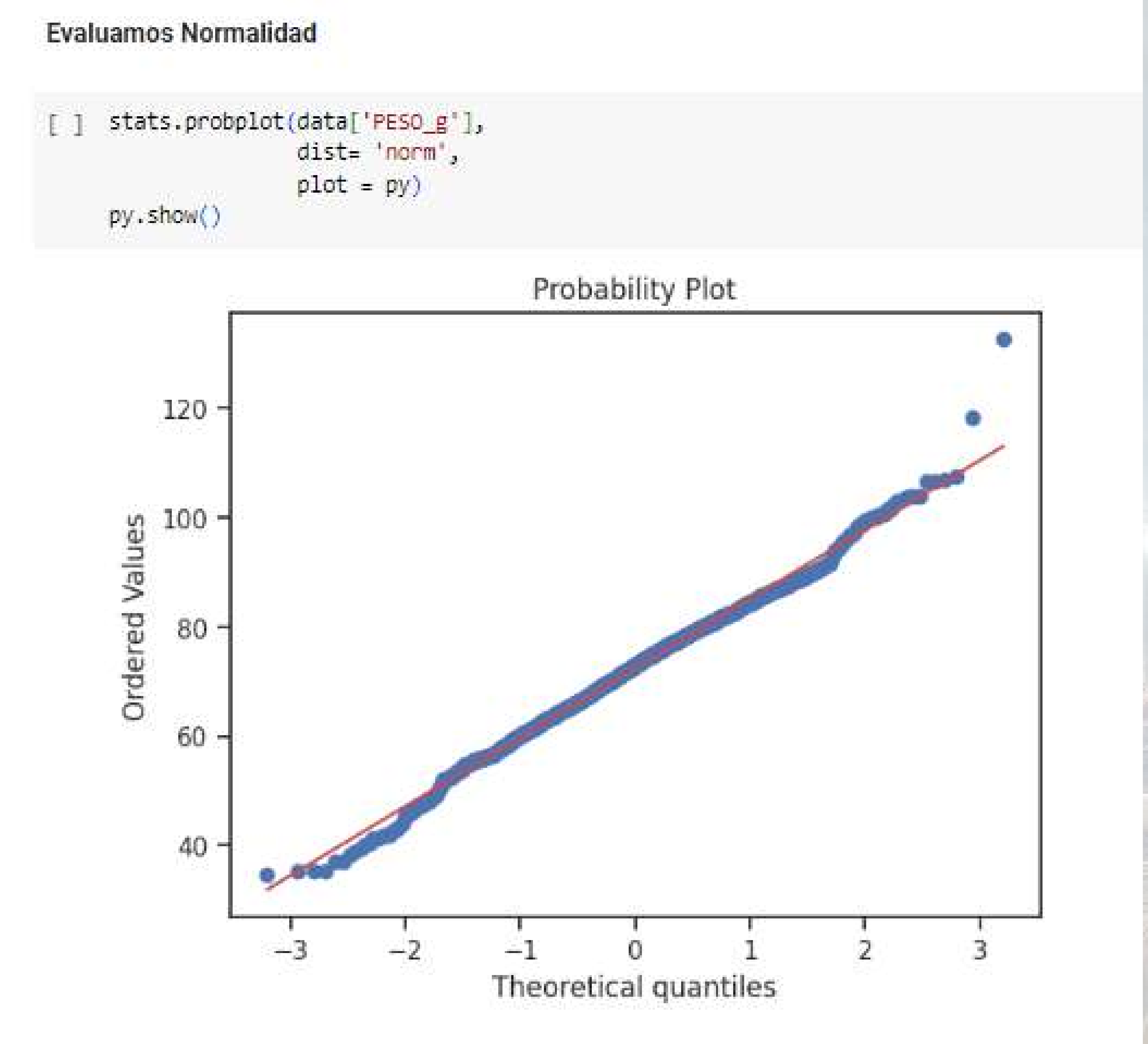


```
[ ]  import scipy as sp
```

```
[ ]  import pylab as py
```

```
[ ]  from scipy import stats
```

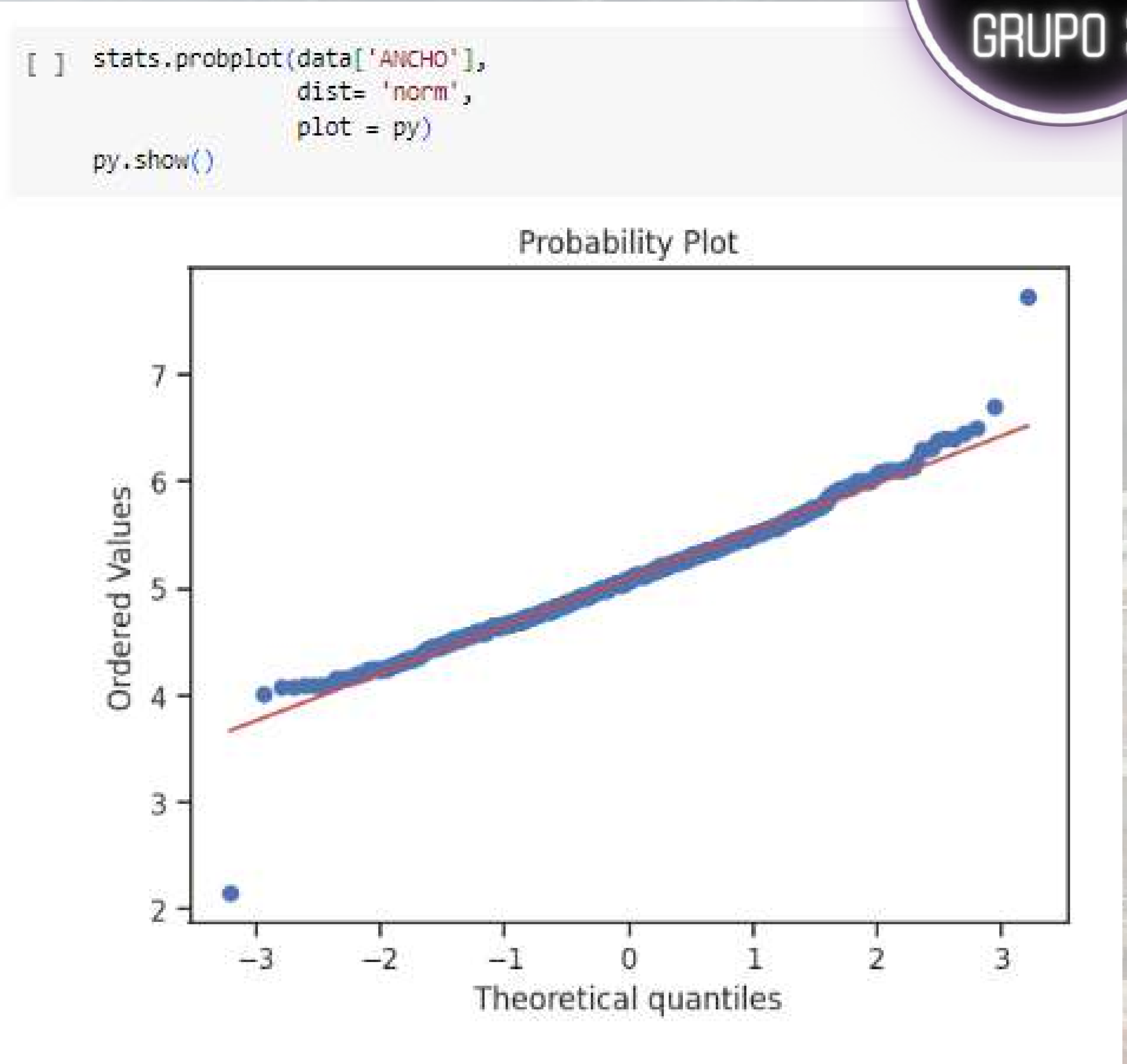




Evaluamos Normalidad

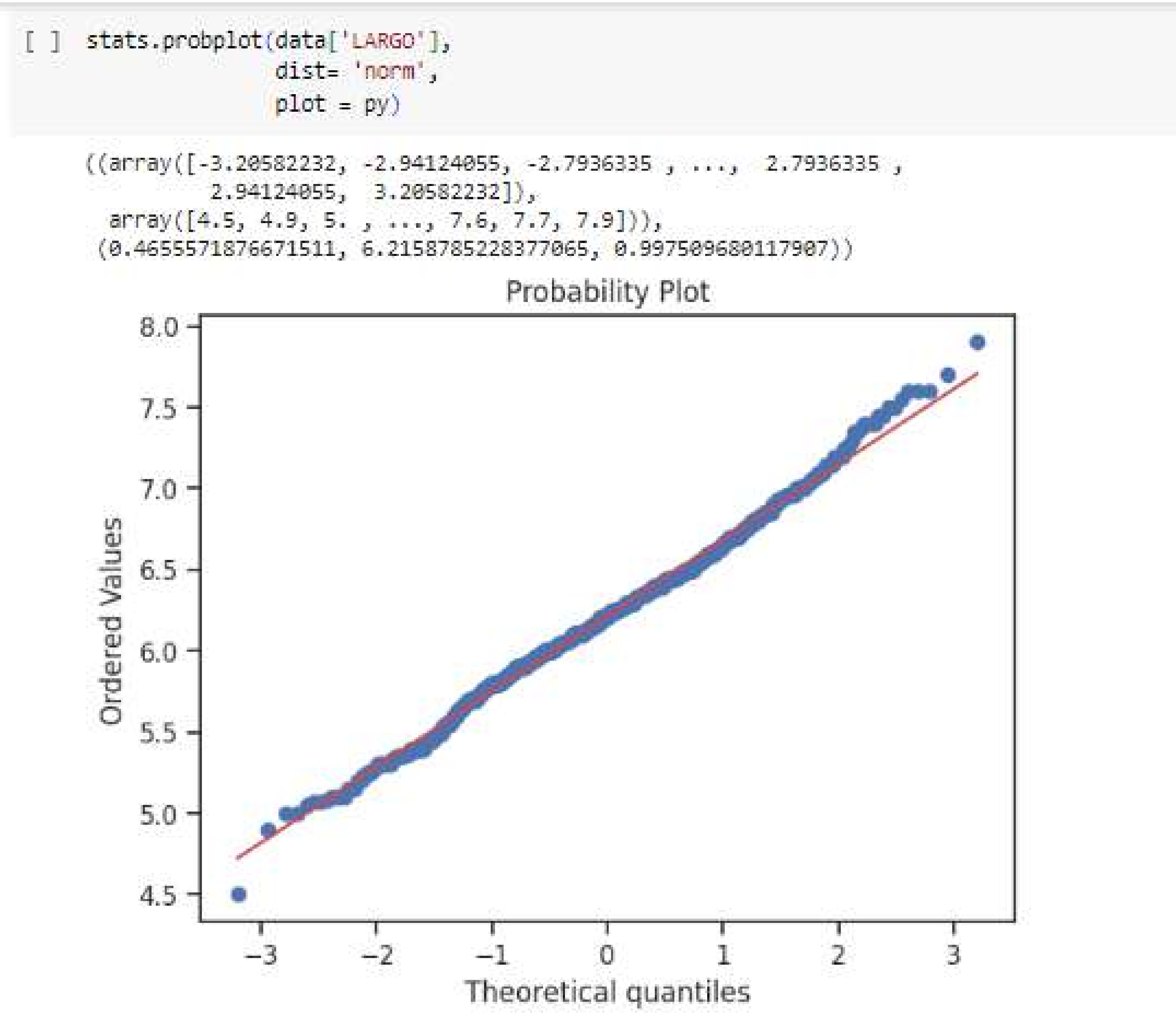
```
[ ]  stats.probplot(data['PESO_g'],
                    dist= 'norm',
                    plot = py)
     py.show()
```

Probability Plot
Ordered Values
40
60
80
100
120
Theoretical quantiles
−3
−2
−1
0
1
2
3



```
[ ]  stats.probplot(data['ANCHO'],
                    dist= 'norm',
                    plot = py)
     py.show()
```

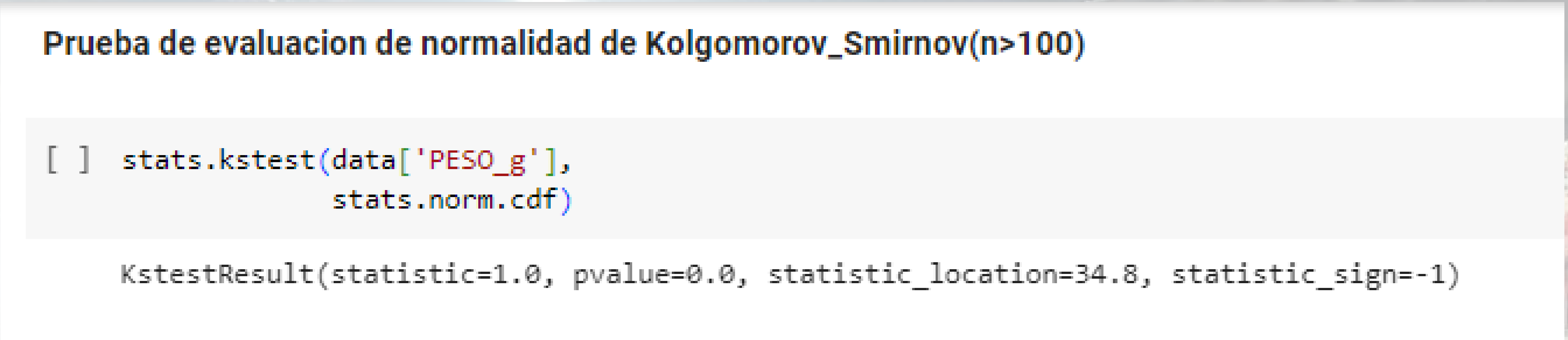
Probability Plot
Ordered Values
2
3
4
5
6
7
Theoretical quantiles
−3
−2
−1
0
1
2
3



```
[ ]  stats.probplot(data['LARGO'],
                   dist= 'norm',
                   plot = py)
```

```
((array([-3.20582232, -2.94124055, -2.7936335 , ...,  2.7936335 ,
          2.94124055,  3.20582232]),
  array([4.5, 4.9, 5. , ..., 7.6, 7.7, 7.9])),
 (0.4655571876671511, 6.2158785228377065, 0.997509680117907))
```

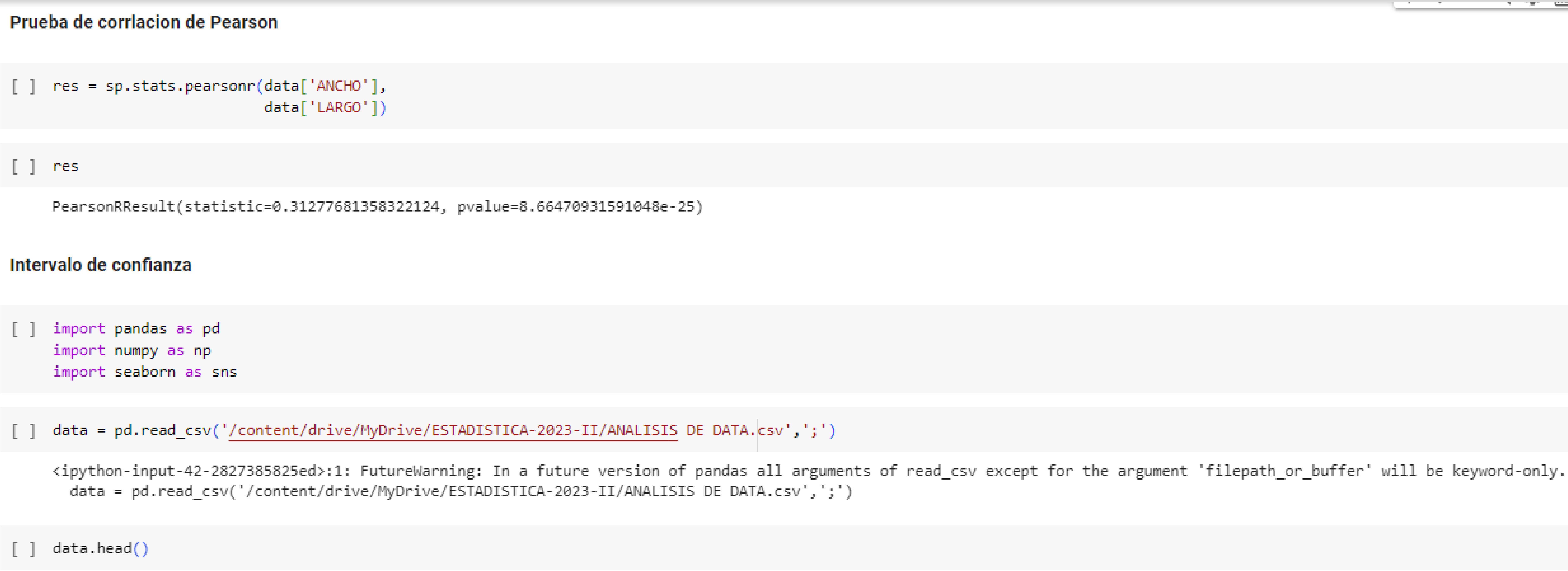
Probability Plot
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
Ordered Values
−3
−2
−1
0
1
2
3
Theoretical quantiles



Prueba de evaluacion de normalidad de Kolgomorov_Smirnov(n>100)

```
[ ]  stats.kstest(data['PESO_g'],
                  stats.norm.cdf)
```

```
KstestResult(statistic=1.0, pvalue=0.0, statistic_location=34.8, statistic_sign=-1)
```



Prueba de corrlacion de Pearson

```
[ ] res = sp.stats.pearsonr(data['ANCHO'],
                            data['LARGO'])
```

```
[ ] res
```

```
PearsonRResult(statistic=0.31277681358322124, pvalue=8.66470931591048e-25)
```

Intervalo de confianza

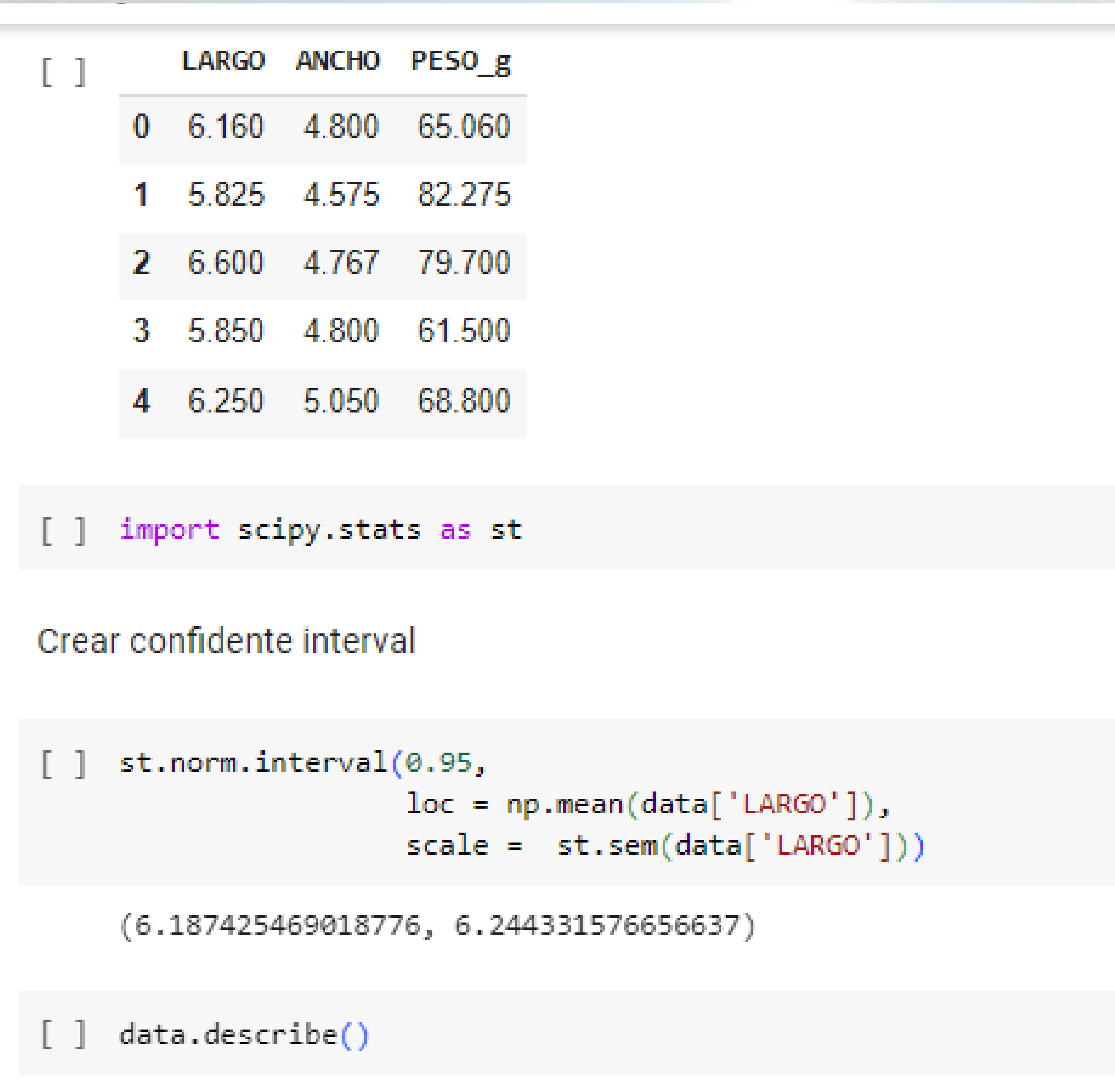
```
[ ] import pandas as pd
    import numpy as np
    import seaborn as sns
```

```
[ ] data = pd.read_csv('/content/drive/MyDrive/ESTADISTICA-2023-II/ANALISIS DE DATA.csv',';')
```

```
<ipython-input-42-2827385825ed>:1: FutureWarning: In a future version of pandas all arguments of read_csv except for the argument 'filepath_or_buffer' will be keyword-only.
  data = pd.read_csv('/content/drive/MyDrive/ESTADISTICA-2023-II/ANALISIS DE DATA.csv',';')
```

```
[ ] data.head()
```

GRUPO 3



	LARGO	ANCHO	PESO_g
0	6.160	4.800	65.060
1	5.825	4.575	82.275
2	6.600	4.767	79.700
3	5.850	4.800	61.500
4	6.250	5.050	68.800

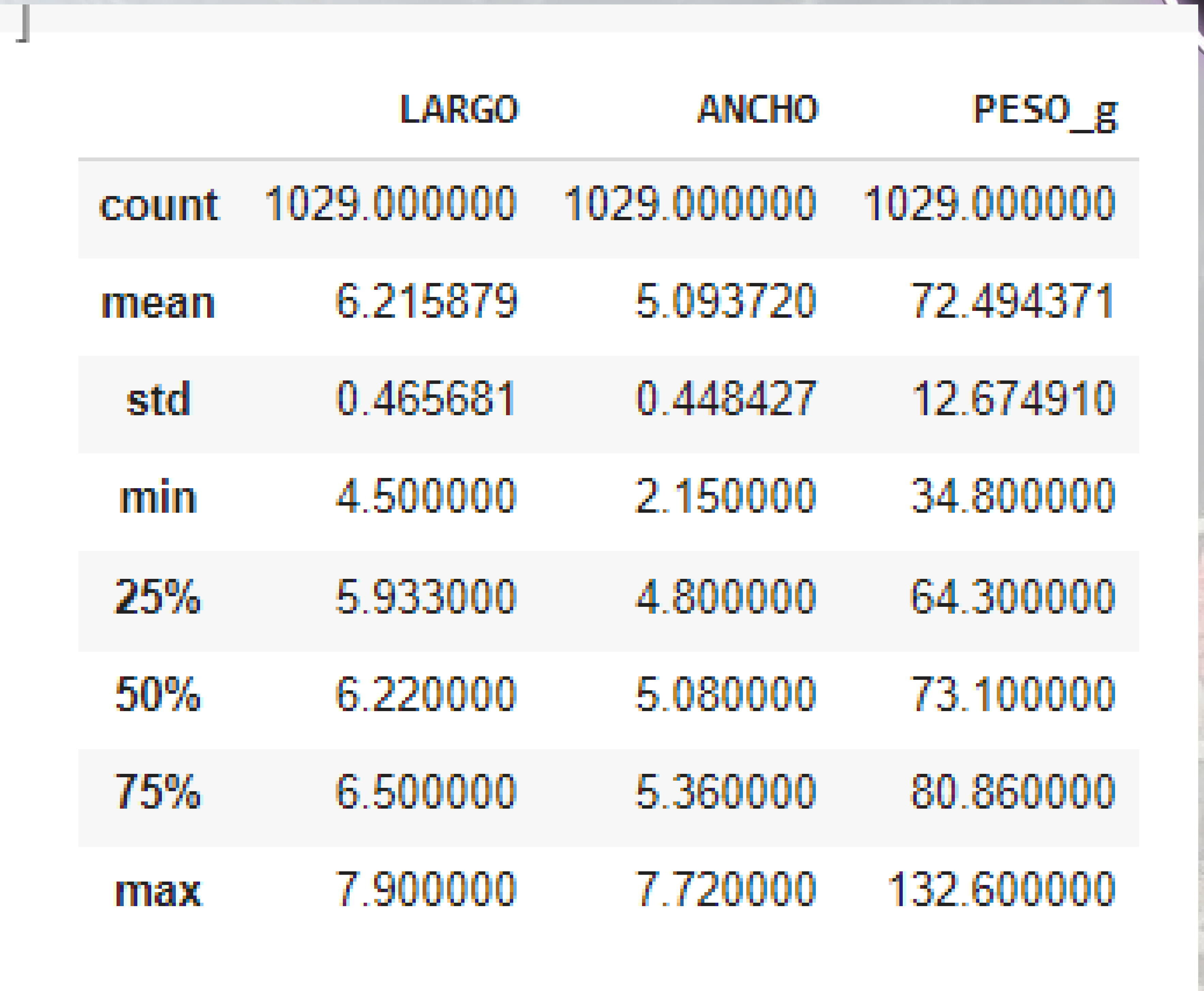
```
import scipy.stats as st
```

Crear confidente interval

```
st.norm.interval(0.95,
                 loc = np.mean(data['LARGO']),
                 scale =  st.sem(data['LARGO']))
```

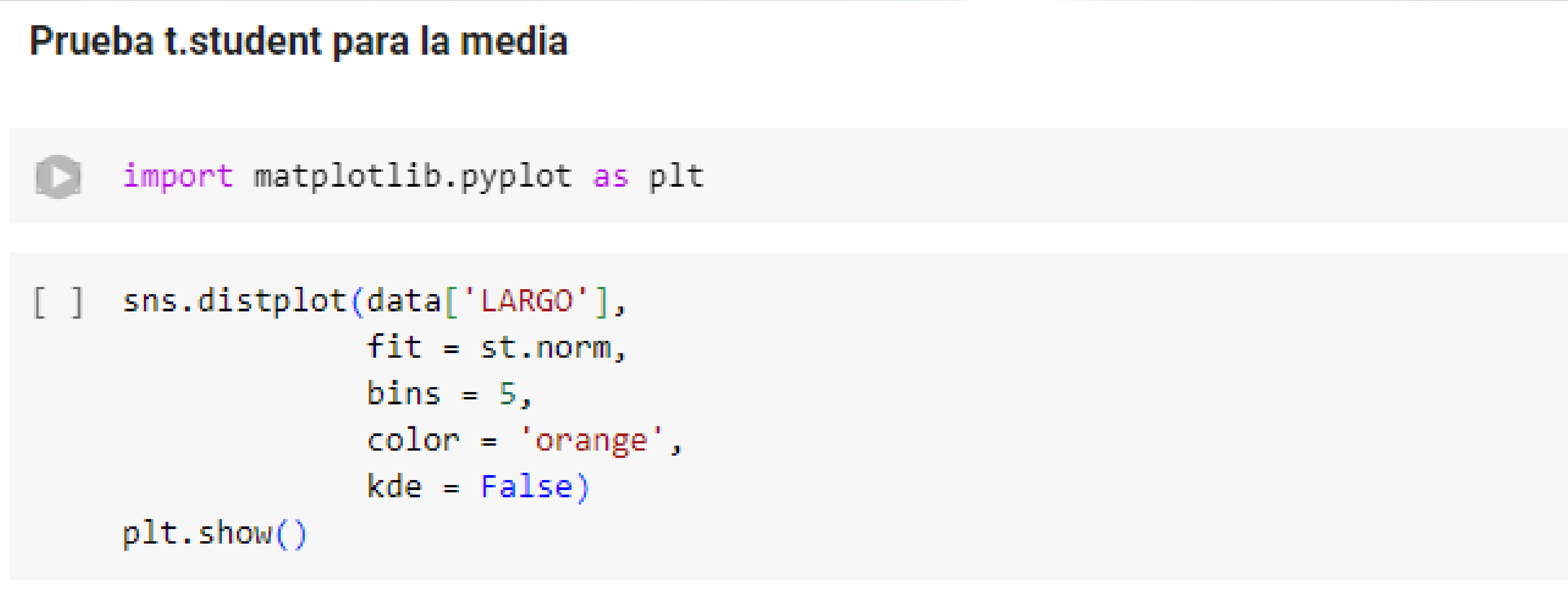
```
(6.187425469018776, 6.244331576656637)
```

```
data.describe()
```



	LARGO	ANCHO	PESO_g
count	1029.000000	1029.000000	1029.000000
mean	6.215879	5.093720	72.494371
std	0.465681	0.448427	12.674910
min	4.500000	2.150000	34.800000
25%	5.933000	4.800000	64.300000
50%	6.220000	5.080000	73.100000
75%	6.500000	5.360000	80.860000
max	7.900000	7.720000	132.600000

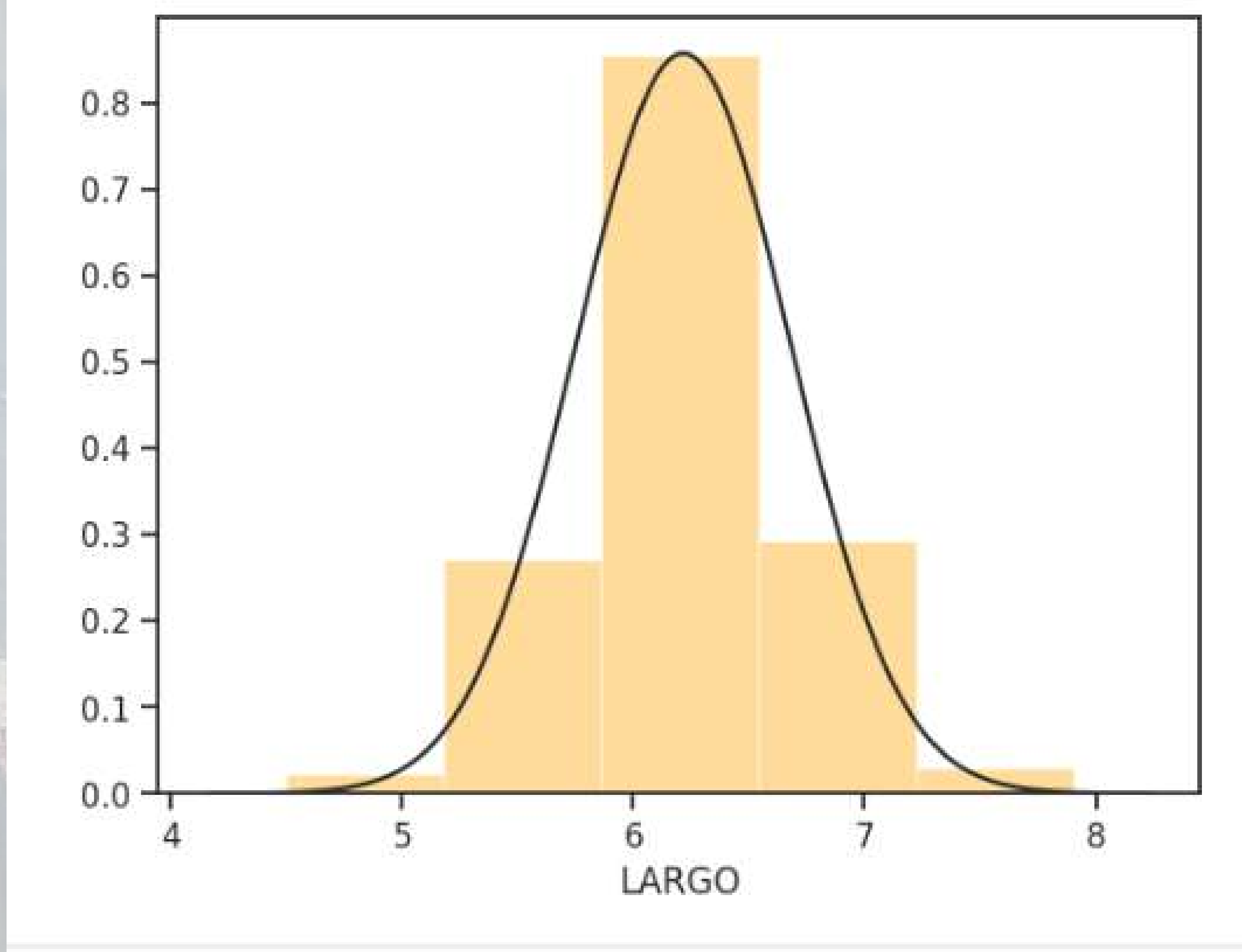
GRUPO 3

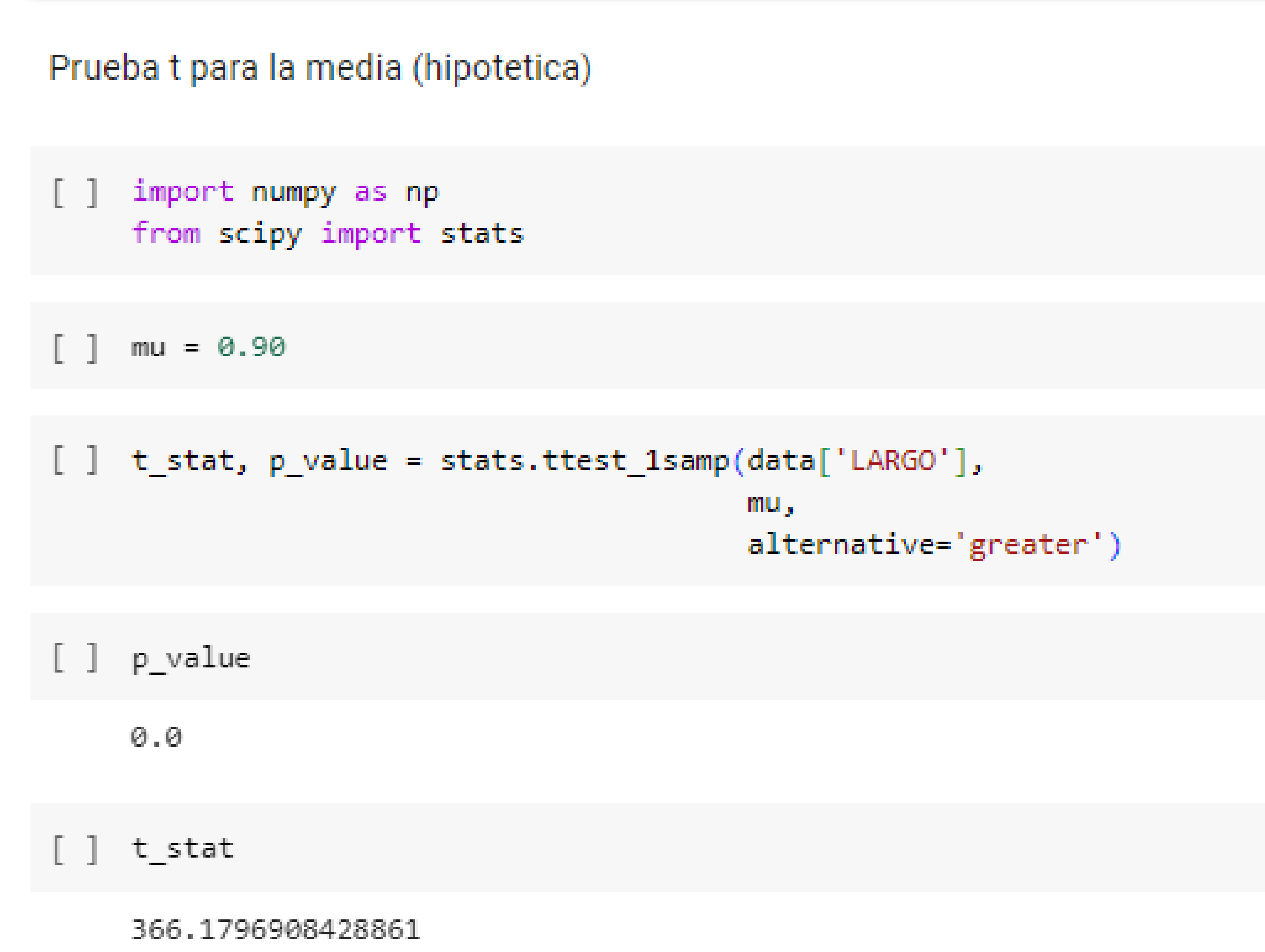


Prueba t.student para la media

```
import matplotlib.pyplot as plt
```

```
[ ]  sns.distplot(data['LARGO'],
                  fit = st.norm,
                  bins = 5,
                  color = 'orange',
                  kde = False)
     plt.show()
```





Prueba t para la media (hipotetica)

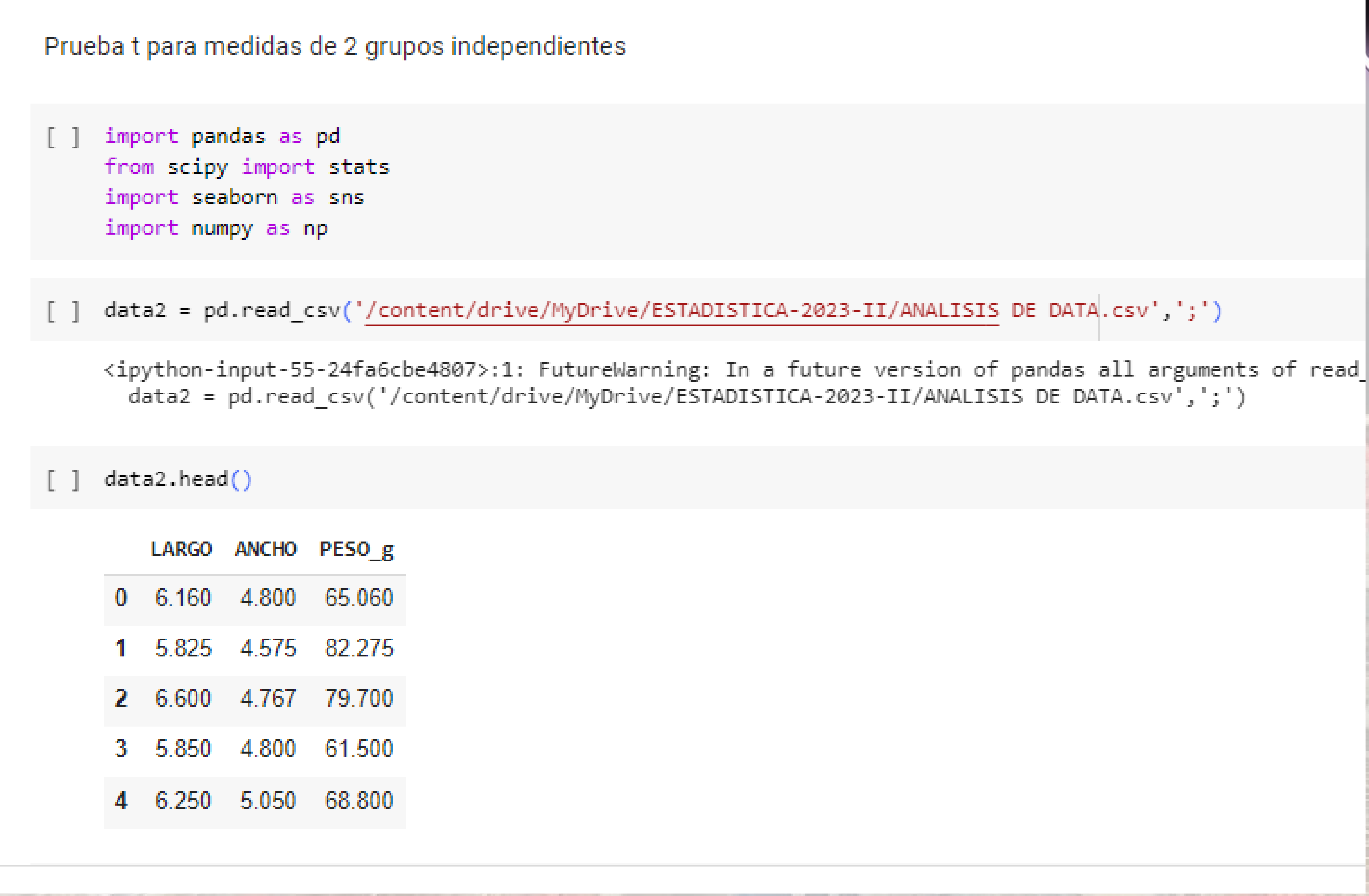
```
[ ]  import numpy as np
     from scipy import stats
```

```
[ ]  mu = 0.90
```

```
[ ]  t_stat, p_value = stats.ttest_1samp(data['LARGO'],
                                        mu,
                                        alternative='greater')
```

```
[ ]  p_value

     0.0
```

```
[ ]  t_stat

     366.1796908428861
```



Prueba t para medidas de 2 grupos independientes

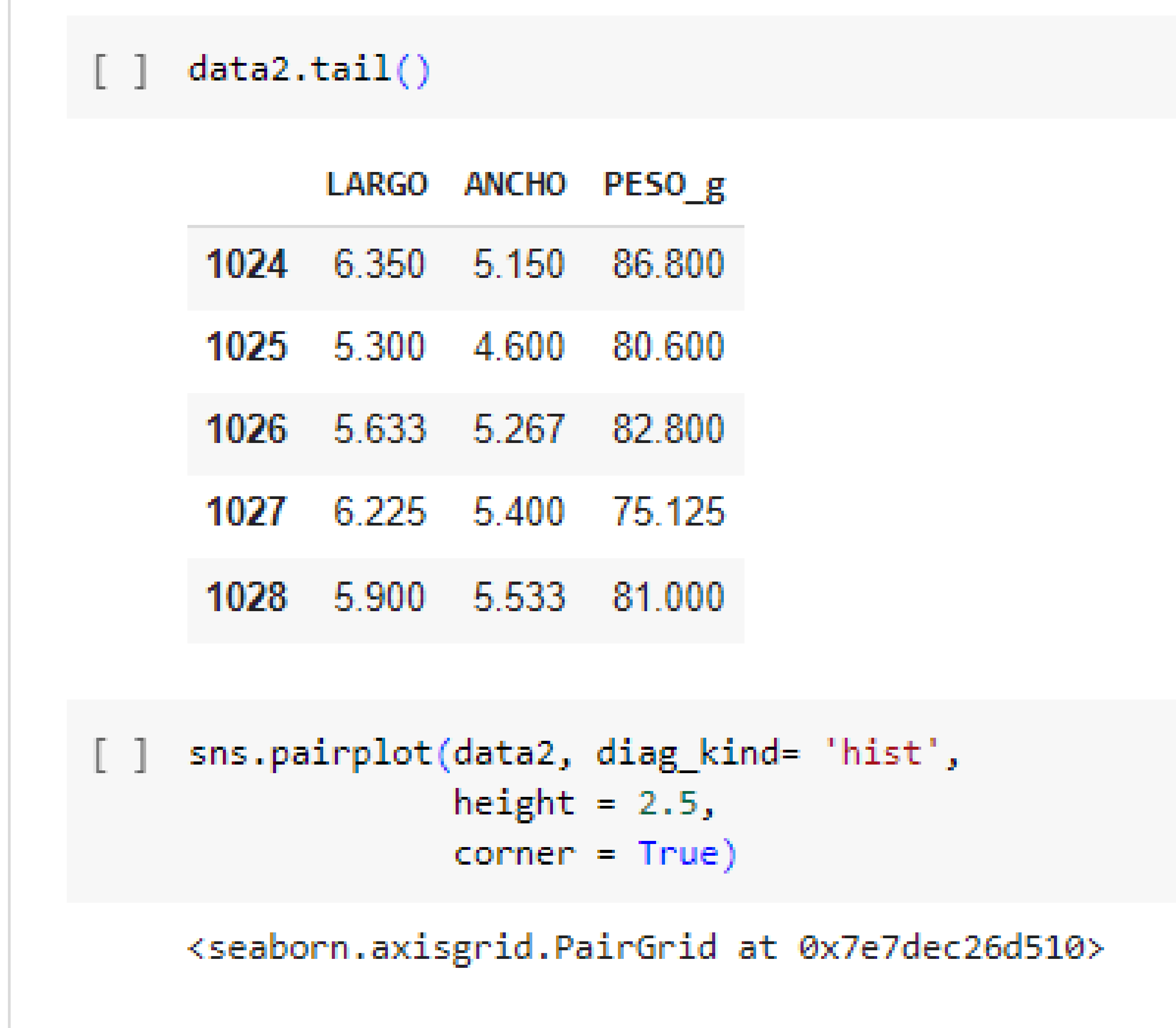
```
[ ] import pandas as pd
    from scipy import stats
    import seaborn as sns
    import numpy as np
```

```
[ ] data2 = pd.read_csv('/content/drive/MyDrive/ESTADISTICA-2023-II/ANALISIS DE DATA.csv',';')
```

```
<ipython-input-55-24fa6cbe4807>:1: FutureWarning: In a future version of pandas all arguments of read_
  data2 = pd.read_csv('/content/drive/MyDrive/ESTADISTICA-2023-II/ANALISIS DE DATA.csv',';')
```

```
[ ] data2.head()
```

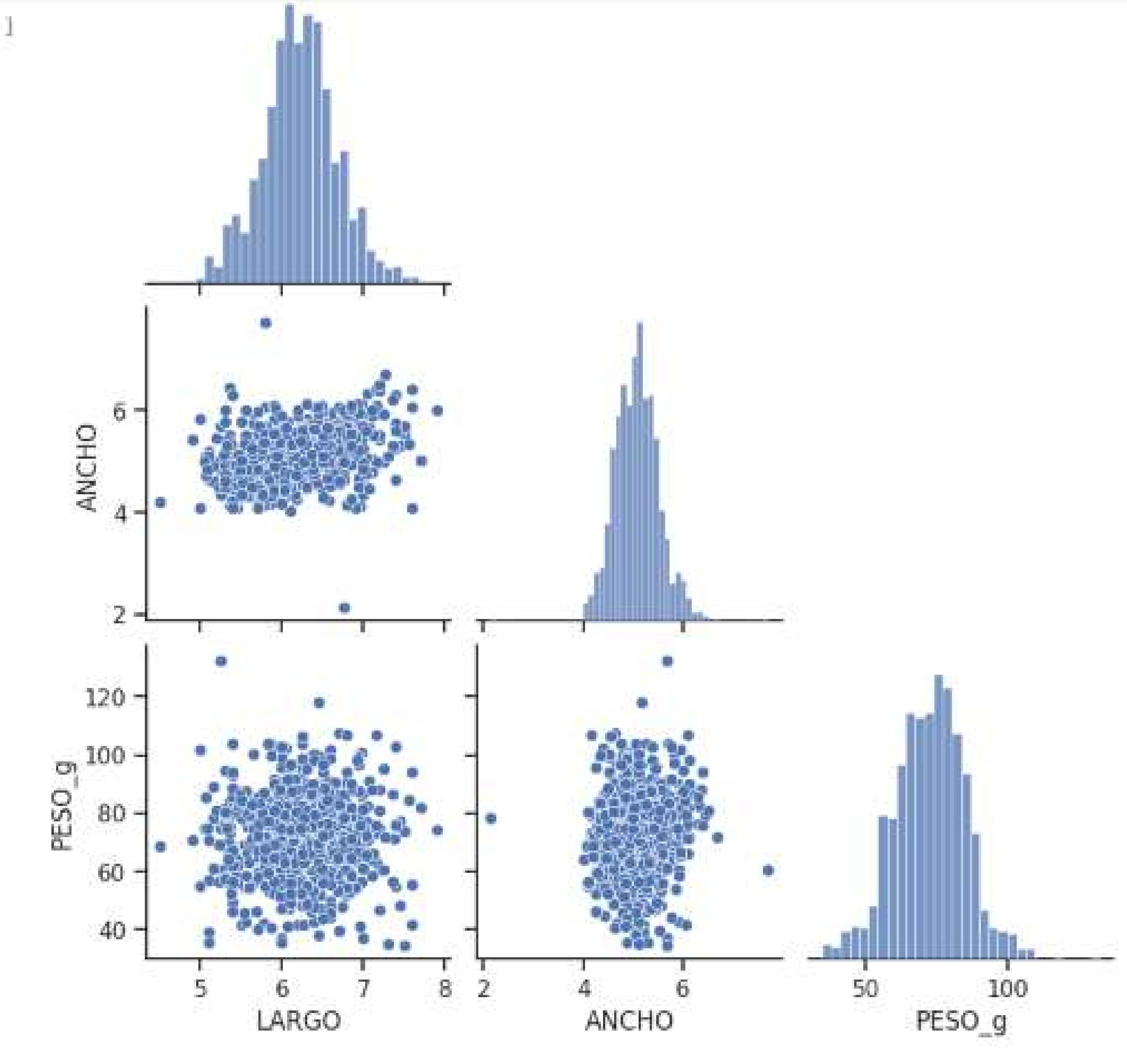
	LARGO	ANCHO	PESO_g
0	6.160	4.800	65.060
1	5.825	4.575	82.275
2	6.600	4.767	79.700
3	5.850	4.800	61.500
4	6.250	5.050	68.800

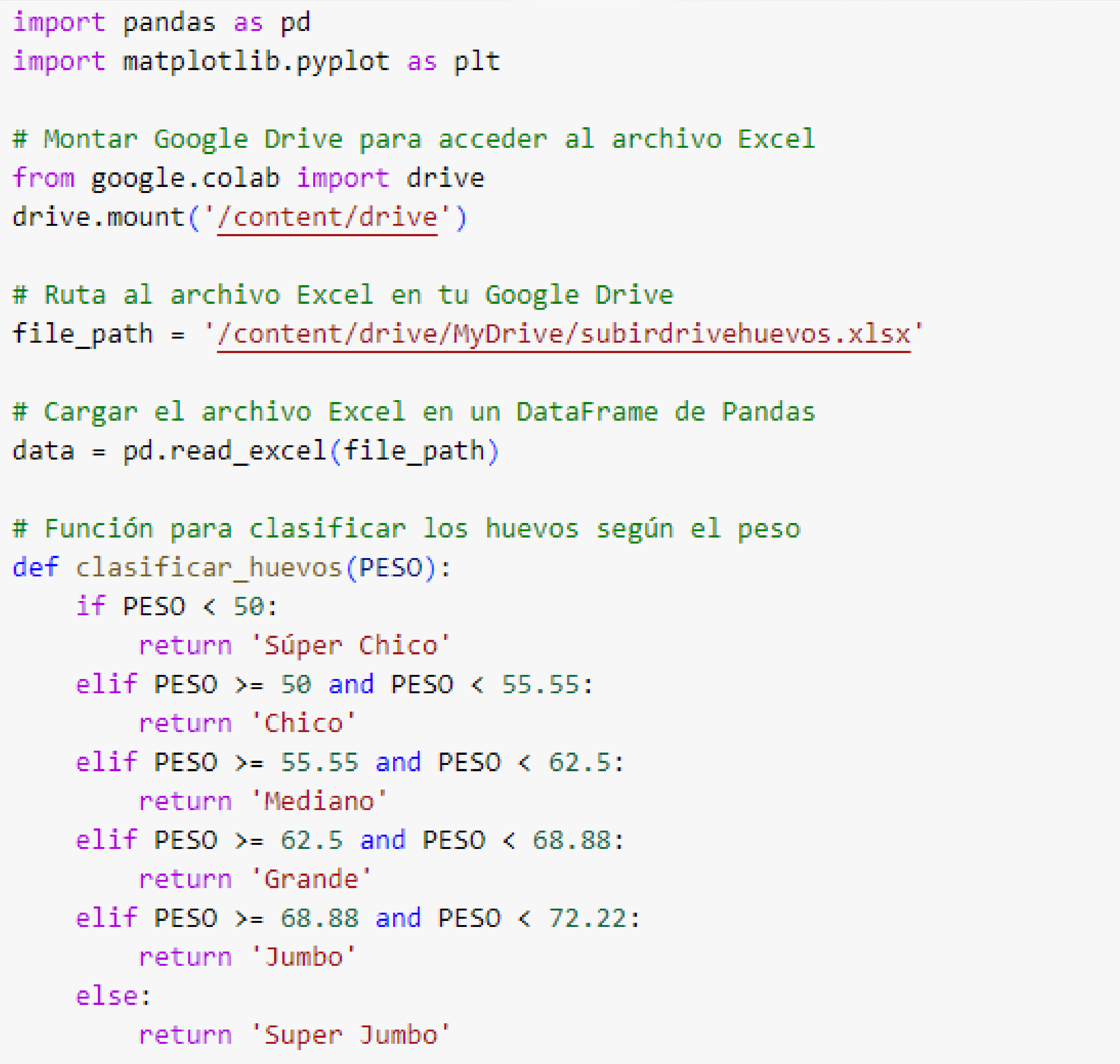


```
[ ]  data2.tail()
```

	LARGO	ANCHO	PESO_g
1024	6.350	5.150	86.800
1025	5.300	4.600	80.600
1026	5.633	5.267	82.800
1027	6.225	5.400	75.125
1028	5.900	5.533	81.000

```
[ ]  sns.pairplot(data2, diag_kind= 'hist',
                  height = 2.5,
                  corner = True)

<seaborn.axisgrid.PairGrid at 0x7e7dec26d510>
```





```
import pandas as pd
import matplotlib.pyplot as plt

# Montar Google Drive para acceder al archivo Excel
from google.colab import drive
drive.mount('/content/drive')

# Ruta al archivo Excel en tu Google Drive
file_path = '/content/drive/MyDrive/subirdrivehuevos.xlsx'

# Cargar el archivo Excel en un DataFrame de Pandas
data = pd.read_excel(file_path)

# Función para clasificar los huevos según el peso
def clasificar_huevos(PESO):
    if PESO < 50:
        return 'Súper Chico'
    elif PESO >= 50 and PESO < 55.55:
        return 'Chico'
    elif PESO >= 55.55 and PESO < 62.5:
        return 'Mediano'
    elif PESO >= 62.5 and PESO < 68.88:
        return 'Grande'
    elif PESO >= 68.88 and PESO < 72.22:
        return 'Jumbo'
    else:
        return 'Super Jumbo'
```

NTP 011.2019 HUEVO DE GALLINA

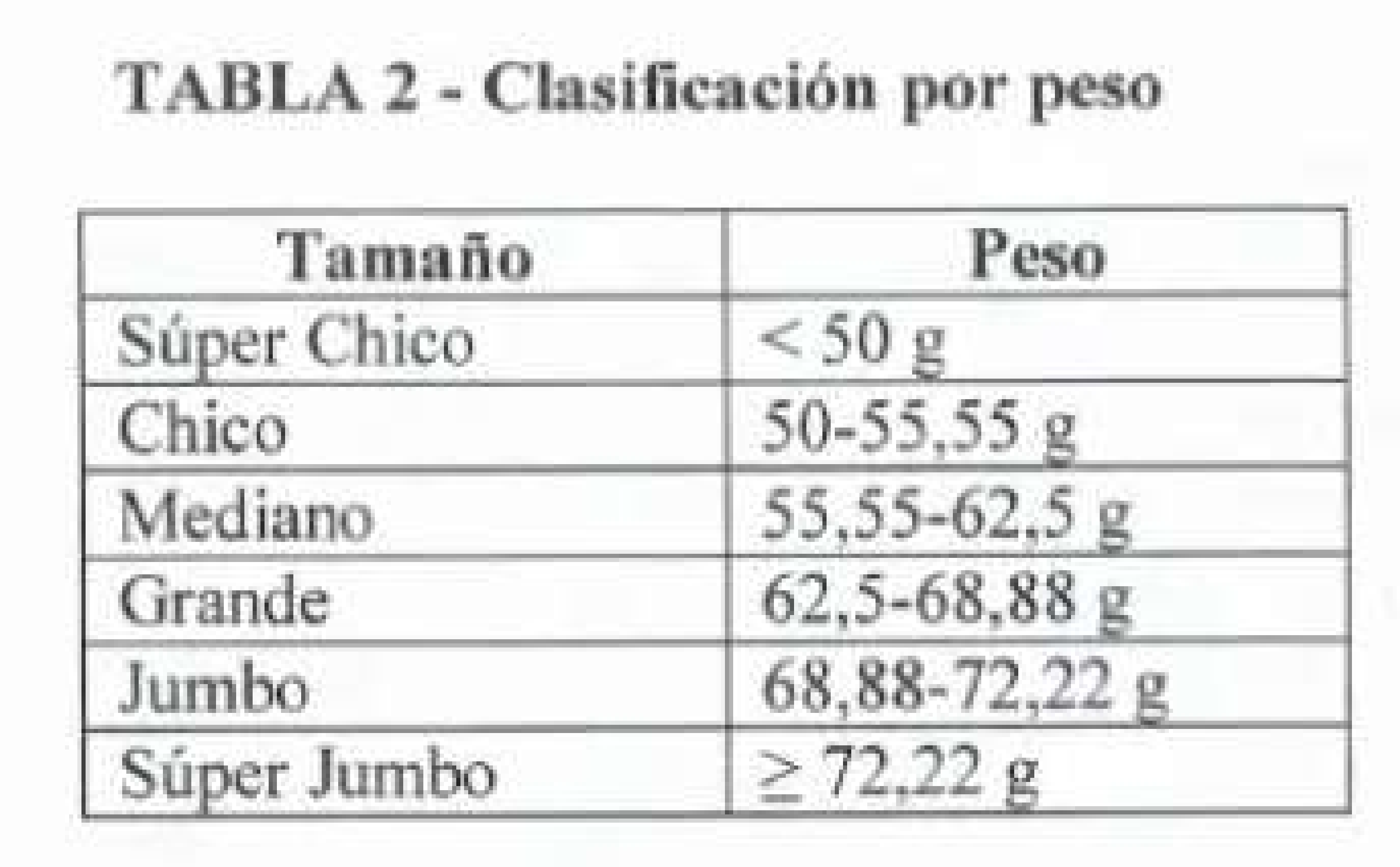
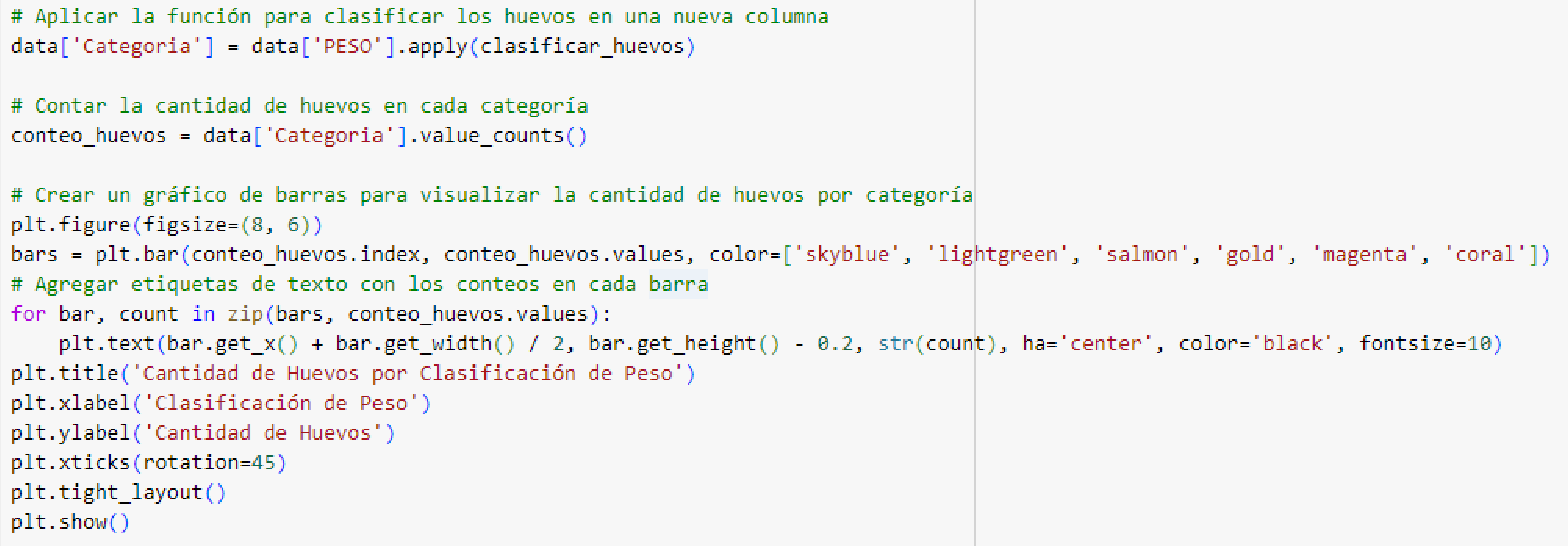
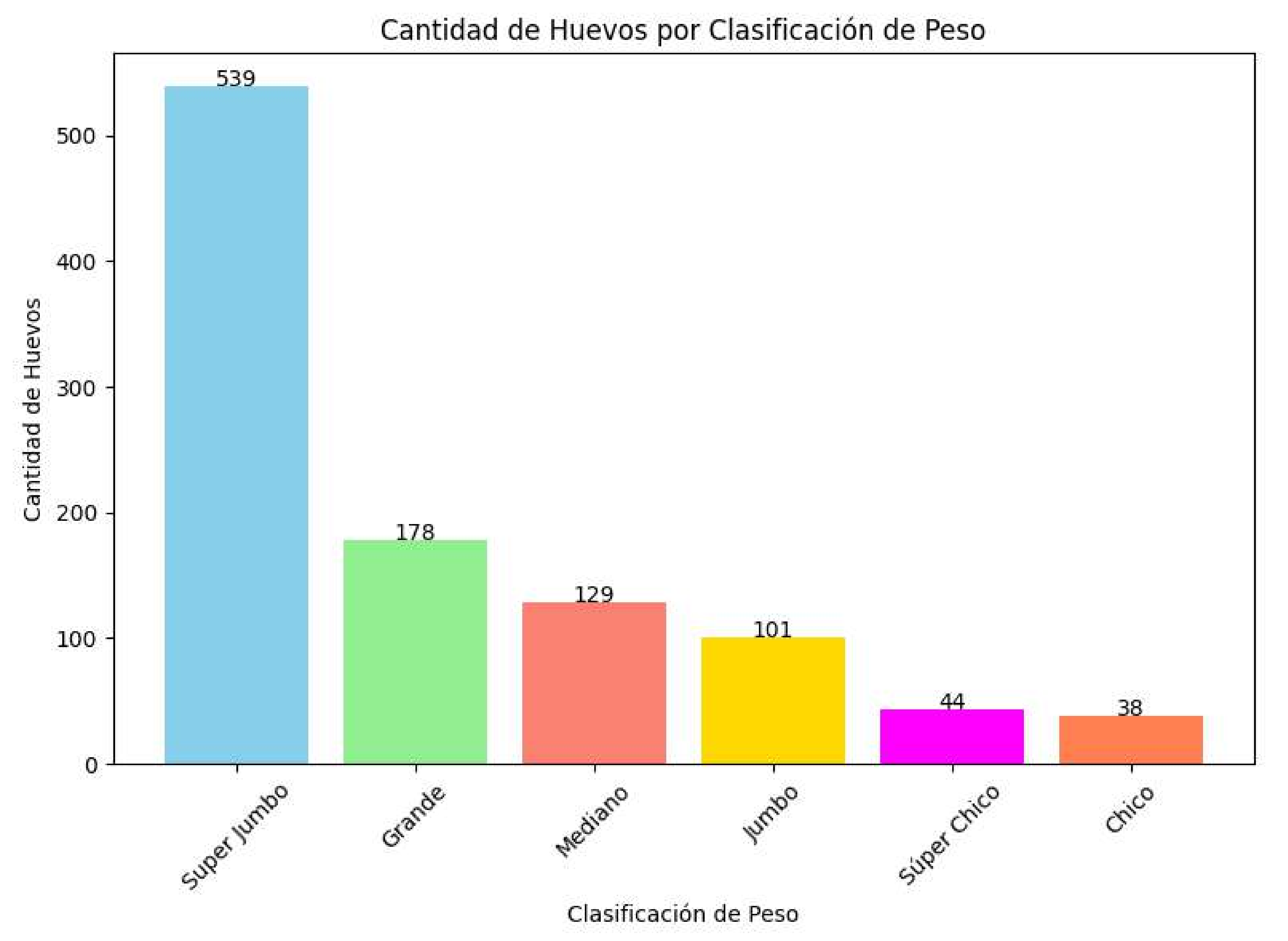


TABLA 2 - Clasificación por peso

Tamaño	Peso
Súper Chico	< 50 g
Chico	50-55,55 g
Mediano	55,55-62,5 g
Grande	62,5-68,88 g
Jumbo	68,88-72,22 g
Súper Jumbo	≥ 72,22 g



```
# Aplicar la función para clasificar los huevos en una nueva columna
data['Categoria'] = data['PESO'].apply(clasificar_huevos)

# Contar la cantidad de huevos en cada categoría
conteo_huevos = data['Categoria'].value_counts()

# Crear un gráfico de barras para visualizar la cantidad de huevos por categoría
plt.figure(figsize=(8, 6))
bars = plt.bar(conteo_huevos.index, conteo_huevos.values, color=['skyblue', 'lightgreen', 'salmon', 'gold', 'magenta', 'coral'])
# Agregar etiquetas de texto con los conteos en cada barra
for bar, count in zip(bars, conteo_huevos.values):
    plt.text(bar.get_x() + bar.get_width() / 2, bar.get_height() - 0.2, str(count), ha='center', color='black', fontsize=10)
plt.title('Cantidad de Huevos por Clasificación de Peso')
plt.xlabel('Clasificación de Peso')
plt.ylabel('Cantidad de Huevos')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```


Cantidad de Huevos por Clasificación de Peso
Cantidad de Huevos
0
100
200
300
400
500
539
178
129
101
44
38
Super Jumbo
Grande
Mediano
Jumbo
Súper Chico
Chico
Clasificación de Peso

UNIVERSIDAD CÉSAR VALLEJO

La estadística es la gramática de la ciencia

-Karl Pearson

GRUPO 3

¡MUCHAS GRACIAS!