


```
In [1]: import pandas as pd
from pandas import ExcelWriter
from pandas import ExcelFile

grupos_totales=3

#Se importa el listado de La población ( muestra con numeros aleatorios en Excel)
#respetando las primisas del problema planteado

df1 = pd.read_excel('Listado_Muestras.xlsx', sheet_name='Hoja1')
df1
```

Out[1]:

	ID	PESO	SEXO
0	1	37.186414	M
1	2	35.337085	N
2	3	37.315491	M
3	4	36.346271	M
4	5	35.540932	M
5	6	36.414193	M
6	7	38.239603	M
7	8	35.291130	M
8	9	35.617051	N
9	10	37.917627	N
10	11	36.048995	N
11	12	37.818631	N
12	13	37.402008	M
13	14	39.071792	M
14	15	36.592887	M
15	16	36.001324	M
16	17	36.646902	M
17	18	36.649568	M
18	19	38.433277	N
19	20	38.675614	M
20	21	36.137090	N
21	22	35.326259	M
22	23	38.053866	M
23	24	36.721128	N
24	25	38.597233	N
25	26	38.357930	N
26	27	34.955637	M
27	28	36.588311	M
28	29	35.873402	M
29	30	35.937008	M
...
170	171	36.540325	M
171	172	35.659285	M
172	173	37.057584	N
173	174	37.933576	M

	ID	PESO	SEXO
174	175	37.273530	N
175	176	36.967561	N
176	177	38.102654	N
177	178	35.936788	M
178	179	35.357114	M
179	180	37.074293	M
180	181	36.626433	N
181	182	37.682006	N
182	183	38.297468	M
183	184	35.679164	M
184	185	37.509240	M
185	186	38.730123	N
186	187	38.809690	M
187	188	38.621036	M
188	189	35.443673	M
189	190	37.144577	M
190	191	36.519466	N
191	192	35.935682	N
192	193	35.455943	M
193	194	35.541238	M
194	195	37.623136	N
195	196	35.224890	N
196	197	37.605130	N
197	198	36.481435	M
198	199	38.306237	N
199	200	36.692669	M

200 rows × 3 columns

```
In [2]: # Acomodo de forma ascendente Los valores de la columna 'Peso':  
df1 = df1.sort_values(by='PESO', ascending=True)  
df1
```

Out[2]:

	ID	PESO	SEXO
167	168	34.903393	N
63	64	34.937366	M
26	27	34.955637	M
124	125	34.957056	N
162	163	34.966741	N
41	42	34.992446	M
101	102	35.015142	M
71	72	35.019647	M
76	77	35.028918	M
130	131	35.068189	N
86	87	35.076313	N
154	155	35.111852	N
128	129	35.126644	N
195	196	35.224890	N
157	158	35.228785	M
7	8	35.291130	M
112	113	35.308637	M
93	94	35.311613	N
21	22	35.326259	M
121	122	35.333560	N
1	2	35.337085	N
118	119	35.342822	M
178	179	35.357114	M
57	58	35.361154	N
168	169	35.388043	M
45	46	35.423296	N
188	189	35.443673	M
192	193	35.455943	M
55	56	35.461512	N
52	53	35.472442	M
...
143	144	38.405617	M
18	19	38.433277	N
51	52	38.510411	N
74	75	38.514911	N

	ID	PESO	SEXO
54	55	38.518286	M
32	33	38.547323	M
140	141	38.590762	N
161	162	38.593280	M
24	25	38.597233	N
60	61	38.619586	N
187	188	38.621036	M
114	115	38.648887	M
19	20	38.675614	M
94	95	38.678882	N
153	154	38.681074	N
166	167	38.705926	N
185	186	38.730123	N
47	48	38.757116	N
186	187	38.809690	M
135	136	38.814334	N
123	124	38.843512	M
117	118	38.864751	M
79	80	38.889760	M
106	107	38.947146	N
69	70	38.987573	M
102	103	39.047712	N
120	121	39.048672	N
13	14	39.071792	M
39	40	39.086421	M
80	81	39.087500	M

200 rows × 3 columns

```
In [3]: #Realizo el acomodo del ID para mantener de menor a mayor la tabla  
df1.ID = range(df1.shape[0])  
df1
```


Out[3]:

	ID	PESO	SEXO
167	0	34.903393	N
63	1	34.937366	M
26	2	34.955637	M
124	3	34.957056	N
162	4	34.966741	N
41	5	34.992446	M
101	6	35.015142	M
71	7	35.019647	M
76	8	35.028918	M
130	9	35.068189	N
86	10	35.076313	N
154	11	35.111852	N
128	12	35.126644	N
195	13	35.224890	N
157	14	35.228785	M
7	15	35.291130	M
112	16	35.308637	M
93	17	35.311613	N
21	18	35.326259	M
121	19	35.333560	N
1	20	35.337085	N
118	21	35.342822	M
178	22	35.357114	M
57	23	35.361154	N
168	24	35.388043	M
45	25	35.423296	N
188	26	35.443673	M
192	27	35.455943	M
55	28	35.461512	N
52	29	35.472442	M
...
143	170	38.405617	M
18	171	38.433277	N
51	172	38.510411	N
74	173	38.514911	N

	ID	PESO	SEXO
54	174	38.518286	M
32	175	38.547323	M
140	176	38.590762	N
161	177	38.593280	M
24	178	38.597233	N
60	179	38.619586	N
187	180	38.621036	M
114	181	38.648887	M
19	182	38.675614	M
94	183	38.678882	N
153	184	38.681074	N
166	185	38.705926	N
185	186	38.730123	N
47	187	38.757116	N
186	188	38.809690	M
135	189	38.814334	N
123	190	38.843512	M
117	191	38.864751	M
79	192	38.889760	M
106	193	38.947146	N
69	194	38.987573	M
102	195	39.047712	N
120	196	39.048672	N
13	197	39.071792	M
39	198	39.086421	M
80	199	39.087500	M

200 rows × 3 columns

```
In [4]: # Ordenar por los valores de la columna 'Peso' para la tabla:
tabla = df1.sort_values(by='PESO', ascending=True)
```

```
In [5]: # Ordenar por los valores de la columna 'Peso' para la tabla2:
tabla2 = df1.sort_values(by='PESO', ascending=True)
```

```
In [6]: # Ordenar por los valores de la columna 'Peso' para la tabla3:
tabla3 = df1.sort_values(by='PESO', ascending=True)
```

In [7]: *#Realizo la cuenta de cuantos especimenes hay en el archivo para saber de cuantos especimenes van a ser mis grupos*

```
cuenta = 0
for i in df1.index:
    cuenta = cuenta + 1
print(cuenta)
```

200

In [8]: *# Realizo la división de Los especimenes totales y Los grupos totales que de acuerdo al problema son 3*

```
muestras = cuenta // grupos_totales
print(muestras)
```

66

In [9]: *#Calculo el residuo de la division de Los especimenes vs grupos totales*

```
sobranter = cuenta % 3
print(sobranter)
```

2

```
In [10]: #Realizo una comparación de los especímenes para tomar los primeros elementos
          de la variable "muestras"
          #y agrego una columna para el grupo 1

          tabla['Grupo_1'] = tabla['ID'].apply(lambda x: 'True' if x <= muestras else 'False')

          # Ordeno los valores de la columna 'Grupo 1' y elimino, los que estaría fuera
          del rango de acuerdo a la variable muestras

          grupo1 = tabla.sort_values(by='Grupo_1', ascending=False)
          grupo1 = grupo1.drop(grupo1[grupo1.ID >= muestras].index)

          #Exporto el listado
          grupo1.to_excel("Grupo_1.xlsx", sheet_name='Grupo_1')

          #Imprimo la tabla
          grupo1
```

Out[10]:

	ID	PESO	SEXO	Grupo_1
167	0	34.903393	N	True
65	43	35.758964	N	True
146	49	35.877169	N	True
28	48	35.873402	M	True
49	47	35.829574	N	True
58	46	35.807852	M	True
152	45	35.785460	N	True
92	44	35.768222	N	True
111	42	35.746335	M	True
89	34	35.612335	N	True
56	41	35.696900	M	True
163	40	35.683993	N	True
183	39	35.679164	M	True
67	38	35.667361	M	True
171	37	35.659285	M	True
155	36	35.649107	M	True
30	50	35.901564	N	True
141	51	35.910504	N	True
191	52	35.935682	N	True
177	53	35.936788	M	True
29	54	35.937008	M	True
64	55	35.937347	N	True
35	56	35.965202	N	True
15	57	36.001324	M	True
31	58	36.034972	N	True
10	59	36.048995	N	True
38	60	36.057183	N	True
110	61	36.066070	N	True
33	62	36.133709	M	True
20	63	36.137090	N	True
...
4	32	35.540932	M	True
124	3	34.957056	N	True
162	4	34.966741	N	True
41	5	34.992446	M	True

	ID	PESO	SEXO	Grupo_1
101	6	35.015142	M	True
71	7	35.019647	M	True
76	8	35.028918	M	True
130	9	35.068189	N	True
86	10	35.076313	N	True
154	11	35.111852	N	True
128	12	35.126644	N	True
195	13	35.224890	N	True
157	14	35.228785	M	True
7	15	35.291130	M	True
26	2	34.955637	M	True
93	17	35.311613	N	True
45	25	35.423296	N	True
21	18	35.326259	M	True
72	31	35.537894	N	True
164	30	35.487381	N	True
52	29	35.472442	M	True
55	28	35.461512	N	True
188	26	35.443673	M	True
192	27	35.455943	M	True
168	24	35.388043	M	True
57	23	35.361154	N	True
178	22	35.357114	M	True
118	21	35.342822	M	True
1	20	35.337085	N	True
121	19	35.333560	N	True

66 rows × 4 columns

```
In [11]: #Realizo una comparación de los especímenes para tomar los primeros elementos
         de la variable "muestras"
         #y agrego una columna para el grupo 2

         tabla2['Grupo_2'] = tabla2['ID'].apply(lambda x: 'True' if ((x >= muestras) &
         (x <=(muestras*2+sobranante))) else 'False')

         # Ordeno los valores de la columna 'Grupo 2' y elimino los que estaría fuera d
         el rango de acuerdo a la variable muestras

         grupo2 = tabla2.sort_values(by='Grupo_2', ascending=False)
         grupo2 = grupo2.drop(grupo2[grupo2.ID < (muestras)].index)
         grupo2 = grupo2.drop(grupo2[grupo2.ID >= (muestras*2+sobranante)].index)

         #Exporto el listado
         grupo2.to_excel("Grupo_2.xlsx", sheet_name='Grupo_2')

         #Imprimo la tabla
         grupo2
```

Out[11]:

	ID	PESO	SEXO	Grupo_2
165	100	36.859276	N	True
73	84	36.528349	N	True
23	98	36.721128	N	True
148	97	36.717552	M	True
199	96	36.692669	M	True
17	95	36.649568	M	True
16	94	36.646902	M	True
107	93	36.643048	M	True
180	92	36.626433	N	True
14	91	36.592887	M	True
27	90	36.588311	M	True
125	89	36.576477	N	True
59	88	36.570500	M	True
105	87	36.552773	N	True
170	86	36.540325	M	True
99	85	36.536691	M	True
190	83	36.519466	N	True
126	67	36.264250	M	True
119	82	36.515153	N	True
197	81	36.481435	M	True
104	80	36.447956	N	True
5	79	36.414193	M	True
83	78	36.412277	N	True
149	77	36.400193	M	True
159	76	36.394301	M	True
109	75	36.371164	N	True
134	74	36.358653	N	True
91	73	36.353964	N	True
3	72	36.346271	M	True
75	71	36.344855	N	True
...
184	128	37.509240	M	True
66	127	37.487900	M	True
40	126	37.469712	M	True
84	125	37.407041	M	True

	ID	PESO	SEXO	Grupo_2
12	124	37.402008	M	True
68	123	37.393224	M	True
81	122	37.370770	M	True
147	121	37.345504	N	True
2	120	37.315491	M	True
174	119	37.273530	N	True
0	118	37.186414	M	True
160	117	37.174508	M	True
189	116	37.144577	M	True
131	115	37.142539	N	True
150	114	37.127286	M	True
82	113	37.078844	M	True
179	112	37.074293	M	True
70	111	37.062603	N	True
116	110	37.060822	N	True
172	109	37.057584	N	True
90	108	37.049608	N	True
77	107	37.047977	M	True
85	106	37.044217	N	True
115	105	37.030902	N	True
142	104	37.030273	N	True
144	103	36.990943	M	True
108	68	36.284411	N	True
61	66	36.250120	M	True
194	133	37.623136	N	True
196	132	37.605130	N	True

68 rows × 4 columns

```
In [12]: #Realizo una comparación de los especímenes para tomar los primeros elementos
          de la variable "muestras"
          #y agrego una columna para el grupo 3

          tabla3['Grupo_3'] = tabla3['ID'].apply(lambda x: 'True' if ((x >= muestras*2+s
          obrante)) else 'False')

          # Ordeno los valores de la columna 'Grupo 3' y elimino los que estaría fuera d
          el rango de acuerdo a la variable muestras

          grupo3 = tabla3.sort_values(by='Grupo_3', ascending=False)
          grupo3 = grupo3.drop(grupo3[grupo3.ID < (muestras*2+sobrante)].index)

          #Exporto el listado
          grupo3.to_excel("Grupo_3.xlsx", sheet_name='Grupo_3')

          #Imprimo la tabla
          grupo3
```

Out[12]:

	ID	PESO	SEXO	Grupo_3
80	199	39.087500	M	True
173	149	37.933576	M	True
182	164	38.297468	M	True
139	163	38.273645	N	True
6	162	38.239603	M	True
46	161	38.230807	M	True
95	160	38.148819	M	True
88	159	38.106902	M	True
176	158	38.102654	N	True
145	157	38.088024	M	True
22	156	38.053866	M	True
87	155	38.039846	M	True
132	154	38.029919	M	True
158	153	38.023170	N	True
50	152	37.975459	M	True
169	151	37.961702	N	True
34	148	37.931067	M	True
25	166	38.357930	N	True
9	147	37.917627	N	True
43	146	37.914324	M	True
138	145	37.833964	M	True
62	144	37.832361	M	True
11	143	37.818631	N	True
78	142	37.808048	N	True
133	141	37.767197	M	True
53	140	37.733145	M	True
103	139	37.726313	N	True
136	138	37.708629	M	True
181	137	37.682006	N	True
100	136	37.681808	M	True
...
166	185	38.705926	N	True
185	186	38.730123	N	True
47	187	38.757116	N	True
186	188	38.809690	M	True

	ID	PESO	SEXO	Grupo_3
135	189	38.814334	N	True
117	191	38.864751	M	True
106	193	38.947146	N	True
94	183	38.678882	N	True
69	194	38.987573	M	True
102	195	39.047712	N	True
120	196	39.048672	N	True
13	197	39.071792	M	True
39	198	39.086421	M	True
127	168	38.387935	M	True
153	184	38.681074	N	True
123	190	38.843512	M	True
42	169	38.389656	N	True
140	176	38.590762	N	True
19	182	38.675614	M	True
143	170	38.405617	M	True
18	171	38.433277	N	True
51	172	38.510411	N	True
74	173	38.514911	N	True
32	175	38.547323	M	True
54	174	38.518286	M	True
161	177	38.593280	M	True
24	178	38.597233	N	True
60	179	38.619586	N	True
187	180	38.621036	M	True
114	181	38.648887	M	True

66 rows × 4 columns

De esa manera se realizan los 3 grupos indicados en la actividad con los pesos lo más parecidos.