



**INSTITUTO POLITÉCNICO NACIONAL**



**ESCUELA SUPERIOR DE CÓMPUTO**

Número de Proyecto: 1

Nombre de la Actividad: **Análisis descriptivo de Iris Dataset**

Materia: **Analítica y Visualización de Datos**

Alumnos:

**Juárez Sánchez Pablo Darío**

**Loeza Saldaña Sebastián**

Grupo: **5AM1**

Fecha: **22/04/2024**

1. Calcular el área del sépalos y pétalo para cada especie de Iris.

petal area	sepal area	species
0.28	17.85	Iris-setosa
0.28	14.7	Iris-setosa
0.26	15.04	Iris-setosa
0.3	14.26	Iris-setosa
0.28	18	Iris-setosa
0.68	21.06	Iris-setosa
0.42	15.64	Iris-setosa
0.3	17	Iris-setosa
0.28	12.76	Iris-setosa
0.15	15.19	Iris-setosa
0.3	19.98	Iris-setosa
0.32	16.32	Iris-setosa
0.14	14.4	Iris-setosa
0.11	12.9	Iris-setosa
0.24	23.2	Iris-setosa

0.6	25.08	Iris-setosa
0.52	21.06	Iris-setosa
0.42	17.85	Iris-setosa
0.51	21.66	Iris-setosa
0.45	19.38	Iris-setosa
0.34	18.36	Iris-setosa
0.6	18.87	Iris-setosa
0.2	16.56	Iris-setosa
0.85	16.83	Iris-setosa
0.38	16.32	Iris-setosa
0.32	15	Iris-setosa
0.64	17	Iris-setosa
0.3	18.2	Iris-setosa
0.28	17.68	Iris-setosa
0.32	15.04	Iris-setosa
0.32	14.88	Iris-setosa
0.6	18.36	Iris-setosa

0.15	21.32	Iris-setosa
0.28	23.1	Iris-setosa
0.15	15.19	Iris-setosa
0.24	16	Iris-setosa
0.26	19.25	Iris-setosa
0.15	15.19	Iris-setosa
0.26	13.2	Iris-setosa
0.3	17.34	Iris-setosa
0.39	17.5	Iris-setosa
0.39	10.35	Iris-setosa
0.26	14.08	Iris-setosa
0.96	17.5	Iris-setosa
0.76	19.38	Iris-setosa
0.42	14.4	Iris-setosa
0.32	19.38	Iris-setosa
0.28	14.72	Iris-setosa
0.3	19.61	Iris-setosa

0.28	16.5	Iris-setosa
6.58	22.4	Iris-versicolor
6.75	20.48	Iris-versicolor
7.35	21.39	Iris-versicolor
5.2	12.65	Iris-versicolor
6.9	18.2	Iris-versicolor
5.85	15.96	Iris-versicolor
7.52	20.79	Iris-versicolor
3.3	11.76	Iris-versicolor
5.98	19.14	Iris-versicolor
5.46	14.04	Iris-versicolor
3.5	10	Iris-versicolor
6.3	17.7	Iris-versicolor
4	13.2	Iris-versicolor
6.58	17.69	Iris-versicolor
4.68	16.24	Iris-versicolor
6.16	20.77	Iris-versicolor

6.75	16.8	Iris-versicolor
4.1	15.66	Iris-versicolor
6.75	13.64	Iris-versicolor
4.29	14	Iris-versicolor
8.64	18.88	Iris-versicolor
5.2	17.08	Iris-versicolor
7.35	15.75	Iris-versicolor
5.64	17.08	Iris-versicolor
5.59	18.56	Iris-versicolor
6.16	19.8	Iris-versicolor
6.72	19.04	Iris-versicolor
8.5	20.1	Iris-versicolor
6.75	17.4	Iris-versicolor
3.5	14.82	Iris-versicolor
4.18	13.2	Iris-versicolor
3.7	13.2	Iris-versicolor
4.68	15.66	Iris-versicolor

8.16	16.2	Iris-versicolor
6.75	16.2	Iris-versicolor
7.2	20.4	Iris-versicolor
7.05	20.77	Iris-versicolor
5.72	14.49	Iris-versicolor
5.33	16.8	Iris-versicolor
5.2	13.75	Iris-versicolor
5.28	14.3	Iris-versicolor
6.44	18.3	Iris-versicolor
4.8	15.08	Iris-versicolor
3.3	11.5	Iris-versicolor
5.46	15.12	Iris-versicolor
5.04	17.1	Iris-versicolor
5.46	16.53	Iris-versicolor
5.59	17.98	Iris-versicolor
3.3	12.75	Iris-versicolor
5.33	15.96	Iris-versicolor

15	20.79	Iris-virginica
9.69	15.66	Iris-virginica
12.39	21.3	Iris-virginica
10.08	18.27	Iris-virginica
12.76	19.5	Iris-virginica
13.86	22.8	Iris-virginica
7.65	12.25	Iris-virginica
11.34	21.17	Iris-virginica
10.44	16.75	Iris-virginica
15.25	25.92	Iris-virginica
10.2	20.8	Iris-virginica
10.07	17.28	Iris-virginica
11.55	20.4	Iris-virginica
10	14.25	Iris-virginica
12.24	16.24	Iris-virginica
12.19	20.48	Iris-virginica
9.9	19.5	Iris-virginica



14.74	29.26	Iris-virginica
15.87	20.02	Iris-virginica
7.5	13.2	Iris-virginica
13.11	22.08	Iris-virginica
9.8	15.68	Iris-virginica
13.4	21.56	Iris-virginica
8.82	17.01	Iris-virginica
11.97	22.11	Iris-virginica
10.8	23.04	Iris-virginica
8.64	17.36	Iris-virginica
8.82	18.3	Iris-virginica
11.76	17.92	Iris-virginica
9.28	21.6	Iris-virginica
11.59	20.72	Iris-virginica
12.8	30.02	Iris-virginica
12.32	17.92	Iris-virginica
7.65	17.64	Iris-virginica

7.84	15.86	Iris-virginica
14.03	23.1	Iris-virginica
13.44	21.42	Iris-virginica
9.9	19.84	Iris-virginica
8.64	18	Iris-virginica
11.34	21.39	Iris-virginica
13.44	20.77	Iris-virginica
11.73	21.39	Iris-virginica
9.69	15.66	Iris-virginica
13.57	21.76	Iris-virginica
14.25	22.11	Iris-virginica
11.96	20.1	Iris-virginica
9.5	15.75	Iris-virginica
10.4	19.5	Iris-virginica
12.42	21.08	Iris-virginica
9.18	17.7	Iris-virginica

## 2. ¿Cuál de los datos podría contener errores o asignaciones de clases falsas?

Para saber qué datos podrían contener errores o asignaciones falsas se calculó la desviación estándar del área del pétalo y sépalo de cada especie y aquellos datos o instancias que estuvieran a dos desviaciones estándar de la media, significan que podrían contener errores o asignaciones falsas.

-----Sepecie Setosa-----

Area promedio del Pétalo: 0.363, Desviación Estándar del área del pétalo 0.181, Límite 0.726  
Area promedio del Sépalo: 17.209, Desviación Estándar del área del sépalo 2.918, Límite 23.045

Posibles elementos mal clasificados:

Elemento 23, área del pétalo 0.85  
Elemento 43, área del pétalo 0.96  
Elemento 44, área del pétalo 0.76

Elemento 14, área del sépalo 23.2  
Elemento 15, área del sépalo 25.08  
Elemento 33, área del sépalo 23.1

-----Sepecie Versicolor-----

Area promedio del Pétalo: 5.720, Desviación Estándar del área del pétalo 1.355, Límite 8.430  
Area promedio del Sépalo: 16.526, Desviación Estándar del área del sépalo 2.838, Límite 22.202

Posibles elementos mal clasificados:

Elemento 20, área del pétalo 8.64  
Elemento 27, área del pétalo 8.5

Elemento 0, área del sépalo 22.4

-----Sepecie Virginica-----

Area promedio del Pétalo: 11.296, Desviación Estándar del área del pétalo 2.136, Límite 15.568  
Area promedio del Sépalo: 19.685, Desviación Estándar del área del sépalo 3.424, Límite 26.533

Posibles elementos mal clasificados:

Elemento 18, área del pétalo 15.87

Elemento 17, área del sépalo 29.26  
Elemento 31, área del sépalo 30.02

3. ¿Cuál es el error causado por redondear los datos a un decimal?

El error causado de redondear a un decimal es que se podría tener a una instancia u objeto como uno mal clasificado, debido a que se pueden perder decimales y estar fuera o dentro del límite de clasificación, haciendo así que esta instancia tenga una clasificación errónea.

Además que en algoritmos de ML como el KNN esta pérdida de decimales puede tener gran importancia en los cálculos realizados, obteniendo así una clasificación errónea.

4. ¿Cuál es la correlación entre la longitud y el ancho de los pétalos en cada flor?

La correlación entre longitud y ancho de los pétalos es de: 0.9627570970509656

5. ¿Qué par de dimensiones están más correlacionadas?

El par de dimensiones que están más correlacionadas son longitud y ancho de los pétalos con valor de: 0.96

6. Ninguna de las flores en el dataset tiene un ancho de sépalo de 1.8 cm. Entonces, ¿qué longitud de sépalo se espera para una flor que tuviera 1.8 cm como ancho de sépalo?

La longitud de sépalo esperada para una flor con 1.8 cm como ancho de sépalo es de 6.105

7. ¿A qué especie pertenecería un Iris con un ancho de sépalo de 1.8 cm?

La especie a la que podría pertenecer un elemento con ancho de sépalo de 1.8cm es Iris-versicolor

8. Calcular la media aritmética, media generalizada, mediana, moda, varianza, desviación estándar, así como los valores máximos y mínimos de las áreas de los sépalos y pétalos, de cada una de las 3 especies de flores Iris. Dar una interpretación de qué representan estas medidas a nivel macro (conjunto de datos) y a nivel micro (por cada especie de flor Iris).

-----Media aritmética-----

La media del área del pétalo general es de: 5.793

La media del área del sépalo general es de: 17.807

La media del área del pétalo setosa es de: 0.363

La media del área del sépalo setosa es de: 17.209

La media del área del pétalo versicolor es de: 5.720

La media del área del sépalo versicolor es de: 16.526

La media del área del pétalo virginica es de: 11.296

La media del área del sépalo virginica es de: 19.685

-----Media generalizada-----

La media generalizada del área del pétalo general es de: 0.280

La media generalizada del área del pétalo setosa es de: 0.280

La media generalizada del área del sépalo setosa es de: 17.850

La media generalizada del área del pétalo versicolor es de: 6.580

La media generalizada del área del sépalo versicolor es de: 22.400

La media generalizada del área del pétalo virginica es de: 15.000

La media generalizada del área del sépalo virginica es de: 20.790

-----Mediana-----

La mediana del área del pétalo general es de: 5.615

La mediana del área del sépalo general es de: 17.660

Mediana pétalo setosa: 0.300

Mediana sépalo setosa: 17.000

Mediana pétalo versicolor: 5.615

Mediana sépalo versicolor: 16.385

Mediana pétalo virginica: 11.445

Mediana sépalo virginica: 20.060

-----Moda-----

La moda del área del pétalo general es de: [0.28]

La moda del área del sépalo general es de: [13.2]

La moda del petalo setosa es de: [0.28]

La moda del sepal setosa es de: [15.19, 19.38]

La moda del petalo versicolor es de: [6.75]

La moda del sepal versicolor es de: [13.2]

La moda del virginica setosa es de: [7.65, 8.64, 8.82, 9.69, 9.9, 11.34, 13.44]

La moda del virginica setosa es de: [19.5]

-----Varianza-----

La varianza del área del pétalo general es de: 22.069

La varianza del área del sépalo general es de: 11.272

Varianza pétalo setosa: 0.033

Varianza sépalo setosa: 8.515

Varianza pétalo versicolor: 1.835

Varianza sépalo versicolor: 8.055

Varianza pétalo virginica: 4.561

Varianza sépalo virginica: 11.724

-----Desviación Estándar-----

La Desviación Estándar del área del pétalo general es de: 4.698

La Desviación Estándar del área del sépalo general es de: 3.357

Desviación Estándar petalo setosa: 0.181

Desviación Estándar sepalo setosa: 2.918

Desviación Estándar petalo versicolor: 1.355

Desviación Estándar sepalo versicolor: 2.838

Desviación Estándar petalo virginica: 2.136

Desviación Estándar sepalo virginica: 3.424

-----Máximo-----

El máximo del área del pétalo general es de: 15.870

El máximo del área del sépalo general es de: 30.020

Máximo petalo setosa: 0.960

Máximo sepalo setosa: 25.080

Máximo petalo versicolor: 8.640

Máximo sepalo versicolor: 22.400

Máximo petalo virginica: 15.870

Máximo sepalo virginica: 30.020

-----Mínimo-----

El mínimo del área del pétalo general es de: 0.110

El mínimo del área del sépalo general es de: 10.000

Mínimo petalo setosa: 0.110

Mínimo sepalo setosa: 10.350

Mínimo petalo versicolor: 3.300

Mínimo sepalo versicolor: 10.000

Mínimo petalo virginica: 7.500

Mínimo sepalo virginica: 12.250