

Proyecto de análisis de galaxias

Centro Universitario de Guadalajara

José Manuel Alejandro González Campos

Proyecto Final



UNIVERSIDAD DE
GUADALAJARA
Red Universitaria e Institución Benemérita de Jalisco



Programación
Juan Manuel Nava Cervantes
26 de abril del 2025

1. Introducción

Este proyecto tiene como propósito analizar datos estructurales de galaxias a partir de una base de datos en Excel que contiene información sobre parámetros observacionales. A través de Python se implementaron transformaciones matemáticas, cálculos estadísticos, visualizaciones gráficas y una regresión lineal para comprender la relación entre el radio efectivo de las galaxias y su brillo superficial.

2. Carga y preparación de los datos

Se importaron bibliotecas esenciales como pandas, numpy, matplotlib, seaborn, scipy.stats y sklearn. A partir del archivo `galaxias_data.xlsx`, se cargaron las columnas: `raefcorkpg` (radio efectivo), `error_raefcorkpg`, `muecorg` (brillo), y `error_muecorg`.

3. Transformación de datos

Se aplicó una transformación logarítmica base 10 al radio efectivo para facilitar la linealidad del análisis:
 $\log_{10}(\text{raefcorkpg})$.

4. Medidas de tendencia central y dispersión

Se calcularon moda, media, mediana, varianza, desviación estándar y covarianza entre $\log(\text{raefcorkpg})$ y `muecorg`.

5. Visualización de los datos

A continuación se muestran dos visualizaciones generadas: un diagrama de dispersión y una regresión lineal.

Figura 1: Dispersión

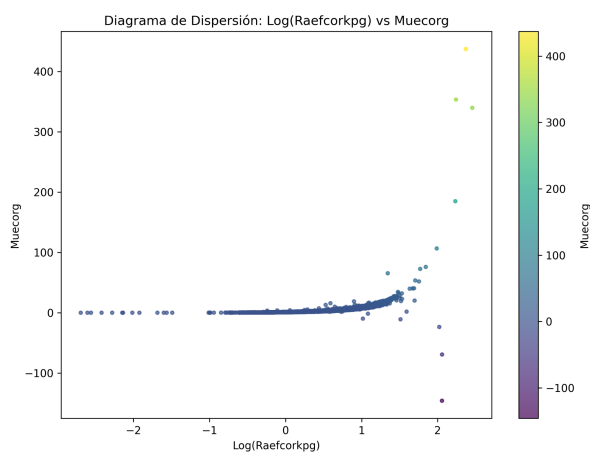
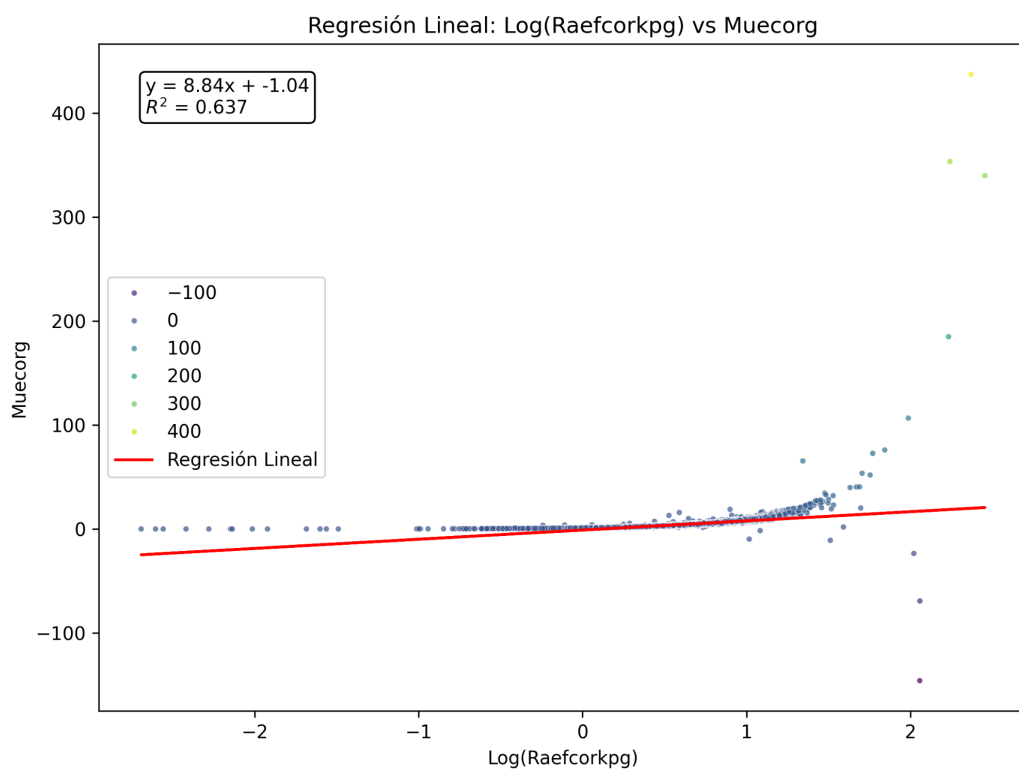


Figura 2: Regresión Lineal



6. Regresión lineal

Se ajustó un modelo lineal **muecorg = b0 + b1 * log(raefcorkpg)**, y se obtuvo R^2 como métrica de ajuste.

Se utilizó `LinearRegression()` de `sklearn` para calcular la pendiente (b1), el intercepto (b0) y el R^2 . Estos se representaron visualmente sobre la gráfica.

7. Resultados obtenidos

Moda: 0.3374550171438442

Media: 0.6656852075119191

Mediana: 0.6962279718240181

Varianza: 0.05993133443555184

Desviación estándar: 0.2448087711573093

Covarianza: 0.529546267705918

Ecuación de la regresión: $y = 8.8359x + -1.0415$

R^2 : 0.6374

8. Conclusiones

Se confirmó una relación entre el log del radio efectivo y el brillo superficial. El análisis demostró que Python permite caracterizar estas propiedades de manera eficiente.

9. Anexos

- estadisticas.txt
- datos_modelo.csv
- dispersogram.png / pdf
- regresion_lineal.png / pdf