

Tarea-Clase-08-datos: Data Analysis sobre las Estrellas

Usando .csv y .root

Realizado desde la maquina virtual de Atlas-Open-Data-ubuntu-2020-v4 a través de Oracle VM VirtualBox.

Datos tomados de: Kaggle <https://n9.c1/v78g7>

José Miguel Ladino Méndez

Observatorio Astronómico Nacional

Universidad Nacional de Colombia - Sede Bogotá

Datos de las estrellas desde el .csv



Librerías usadas: pandas, numpy, ROOT, matplotlib.pyplot y seaborn.

Convención de los 6 tipos de estrella: Brown Dwarf=0, Red Dwarf=1, White Dwarf=2, Main Sequence=3, Supergiant=4, Hypergiant=5.

	Temperature(K)	Luminosity(L/L _o)	Radius(R/R _o)	Absolute_magnitude(Mv)	Star_type	Star_color	Spectral_Class
0	3068	0.002400	0.1700	16.12	0	Red	M
1	3042	0.000500	0.1542	16.60	0	Red	M
2	2600	0.000300	0.1020	18.70	0	Red	M
3	2800	0.000200	0.1600	16.65	0	Red	M
4	1939	0.000138	0.1030	20.06	0	Red	M
...
235	38940	374830.000000	1356.0000	-9.93	5	Blue	O
236	30839	834042.000000	1194.0000	-10.63	5	Blue	O
237	8829	537493.000000	1423.0000	-10.73	5	White	A
238	9235	404940.000000	1112.0000	-11.23	5	White	A
239	37882	294903.000000	1783.0000	-7.80	5	Blue	O

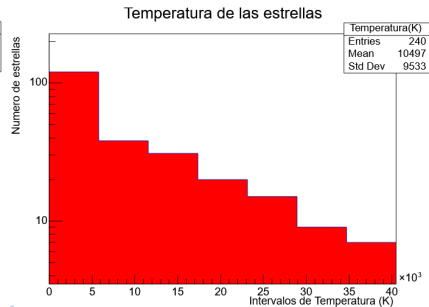
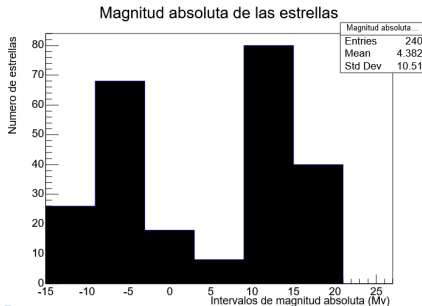
Para poder convertir el archivo .csv correctamente a .root, es necesario modificar las comas y espacios del .csv, luego se anexan índices en la primera columna para poder transformar el documento .csv a .root usando el archivo "convert_csv_to_root.ipynb".

Histograma para las magnitudes absolutas y temperaturas de las estrellas



La media de la magnitud absoluta en (Mv) de las estrellas es de 4.382 (Mv), y la desviación estándar es de 10.54 (Mv).

La media de la temperatura en (K) de las estrellas es de 20399 (K), y la desviación estándar es de 6046 (K).



El 50% de las estrellas tienen una temperatura entre 0 K y 5000 K.

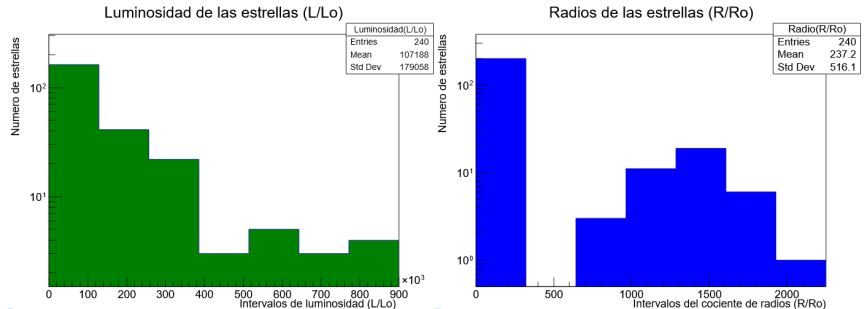
Un 33.3% de las estrellas tienen una magnitud absoluta entre 9 Mv a 15 Mv.

Histograma para las luminosidades y radios de las estrellas



La media de la Luminosidad en (L/L_o) de las estrellas (con L_o la Luminosidad del Sol) es de 107188 (L/L_o), y la desviación estándar es de 179058 (L/L_o).

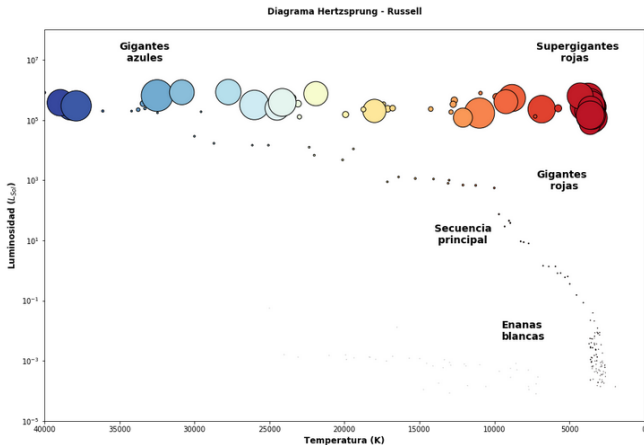
La media de los radios en (R/R_o) de las estrellas (con R_o el radio del Sol) es de 237.2 (R/R_o), y la desviación estándar es de 516.1 (R/R_o).



El 67% de las estrellas tienen una luminosidad entre 0 L/L_o y 100000 L/L_o .

El 83.3% de las estrellas tienen un radio entre 0 R/R_o y 321.3 R/R_o

Diagrama de Hertzsprung - Russell

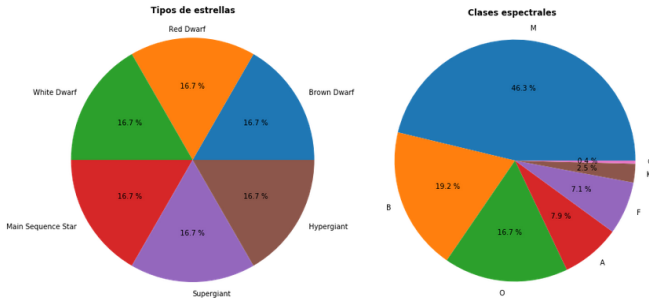


Luminosidad (L/L_{\odot}) en función de la temperatura, para las 240 estrellas, con sus respectivos radios escalados y con sus correspondientes colores, reconociéndose fácilmente los diferentes tipos de estrellas en sus respectivas regiones del diagrama.

Porcentajes para los tipos de estrellas y clases espectrales



Hay 40 estrellas de cada uno de los 6 tipos, Brown Dwarf=0, Red Dwarf=1, White Dwarf=2, Main Sequence=3, Supergiant=4, Hypergiant=5.

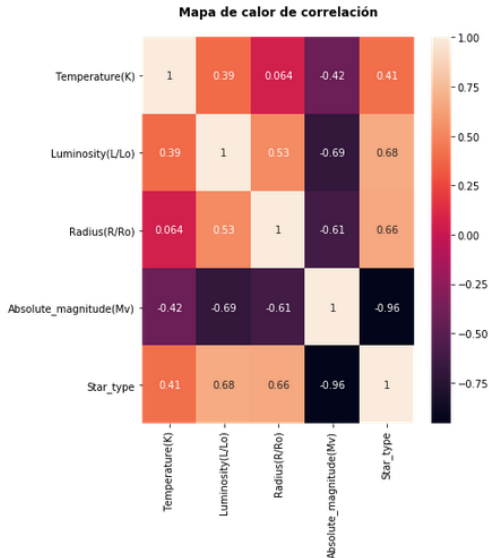


El 46.3% de las estrellas corresponden a la clase espectral M, las más abundantes en los datos, caracterizadas por ser las menos calientes y rojas. Seguida de un 19.2% de estrellas de clase espectral B, caracterizadas por ser las segundas más calientes y azules claro.

Correlaciones entre algunas variables y tipos de estrellas

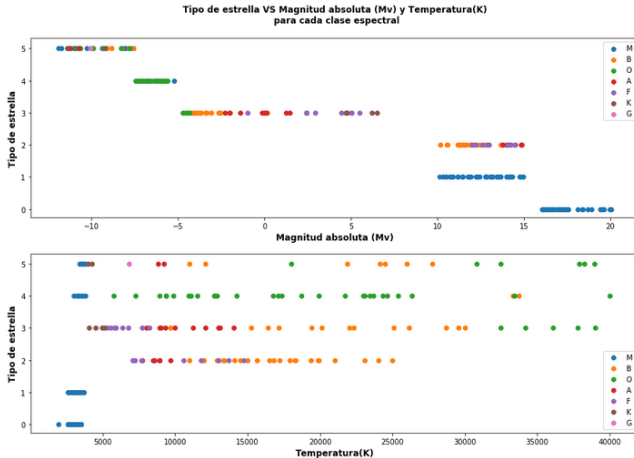


6



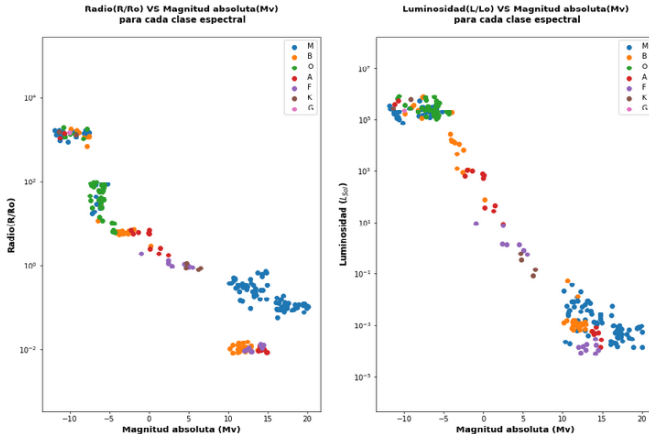
El mapa de calor de correlación entre las variables de estudio, muestra algo que se esperaba del diagrama H-R, existe una tendencia significativa entre la magnitud absoluta, la luminosidad, el tipo de estrella, su radio y su temperatura. Para ver mejor estos comportamientos se proponen los últimos dos gráficos de las últimas dos slides.

Magnitud absoluta y temperatura según: tipos de estrellas y clases espectrales



Las gráficas sugieren que hay estrellas de diferentes tipos en la misma clase espectral, se resalta que casi todas las estrellas de tipos Brown Dwarf=0 y Red Dwarf=1 y de clase espectral M tienen magnitudes absolutas y temperaturas similares.

Radios y luminosidad según: Magnitud absoluta de cada clase espectral



Las gráficas finales como era de esperar, sugieren que hay una tendencia entre los radios y luminosidades de las estrellas con respecto a su magnitud absoluta, organizando sus colores (los de la gráfica) y clases espectrales por regiones.



- ▶ Luego de instalar la maquina virtual de Atlas-Open-Data-ubuntu-2020-v4 y usarla a través de Oracle VM VirtualBox, fue posible estudiar estadísticamente a las estrellas y sus correspondientes radios, magnitudes absolutas, temperaturas y luminosidades usando pandas y ROOT.
- ▶ Se logró evidenciar que hay igual cantidad de estrellas de cada tipo en la base de datos, mientras que la mayoría de las estrellas pertenecen a la clase espectral M. También fue posible deducir que la mayoría de estas estrellas tienen una luminosidad entre 0 L/Lo y 100000 L/Lo, un radio entre 0 R/Ro y 321.3 R/Ro. Y adicionalmente se concluye que la mitad de las estrellas tienen una temperatura entre 0 K y 5000 K.
- ▶ Se logró gráficar y evidenciar las respuestas de tendencia del Diagrama Hertzsprung - Russell para diferentes tipos de estrella de distintas clases espectrales.
- ▶ La base de datos en formato .root es bastante útil para un análisis estadístico rápido y sencillo, además de que con su output interactivo se logra reconocer cada región en detalle de los plots.
- ▶ Manipular los datos en formato .csv también tiene sus beneficios, cómo el análisis de correlación de variables el cuál sugiere de manera rápida qué combinaciones de tendencias estudiar.