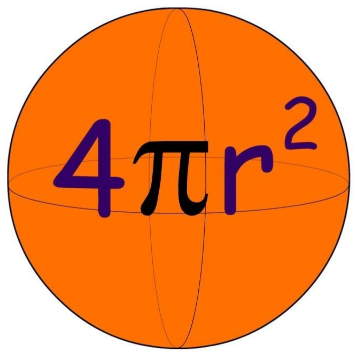
STAR DATASET TO PREDICT STAR TYPES:

1. The column **'Luminosity(L/Lo)'** in our data frame contains values of the ratio of the **Luminosity of the Star:Luminosity of the Sun.**
   And similarly column, '**Radius(R/Ro)'** contains the values of the ratio of the **Radius of the Star:Radius of the Sun.**
   Hence to find the Luminosity and Radius, let us multiply the values under the columns with **'Luminosity of the Sun'** and **'Radius of Sun'** respectively**.** We can create new columns containing the values of the Luminosities and Radii of the stars.

2. When we try to carry out calculations related to a star, we assume the star to be a perfectly spherical body, to make our calculations easily computable.
   The Surface Area of any spherical shaped object is:



3. **The formula to find the Luminosity of a star on a Main Sequence is:**

$$L = 4\pi R^2 \sigma T^4$$

L= Luminosity
R= Stellar Radius
T= Surface Temperature

σ = Stefan-Boltzmann Constant

As we are already given the **'Luminosity'** and the **'Temperature'** in the data frame, and have already computed the **'Surface Area'**, we can use this data to calculate the **'Stefan-Boltzmann Constant'**.
The value of the Stefan–Boltzmann constant is given in SI units by:

$$\sigma = 5.6703x10^{-8}\, watt / m^2 K^4$$

As per our calculations, the median value of the Stefan-Boltzmann constant is:

5.59*e-08

That is close to the value of the constant.

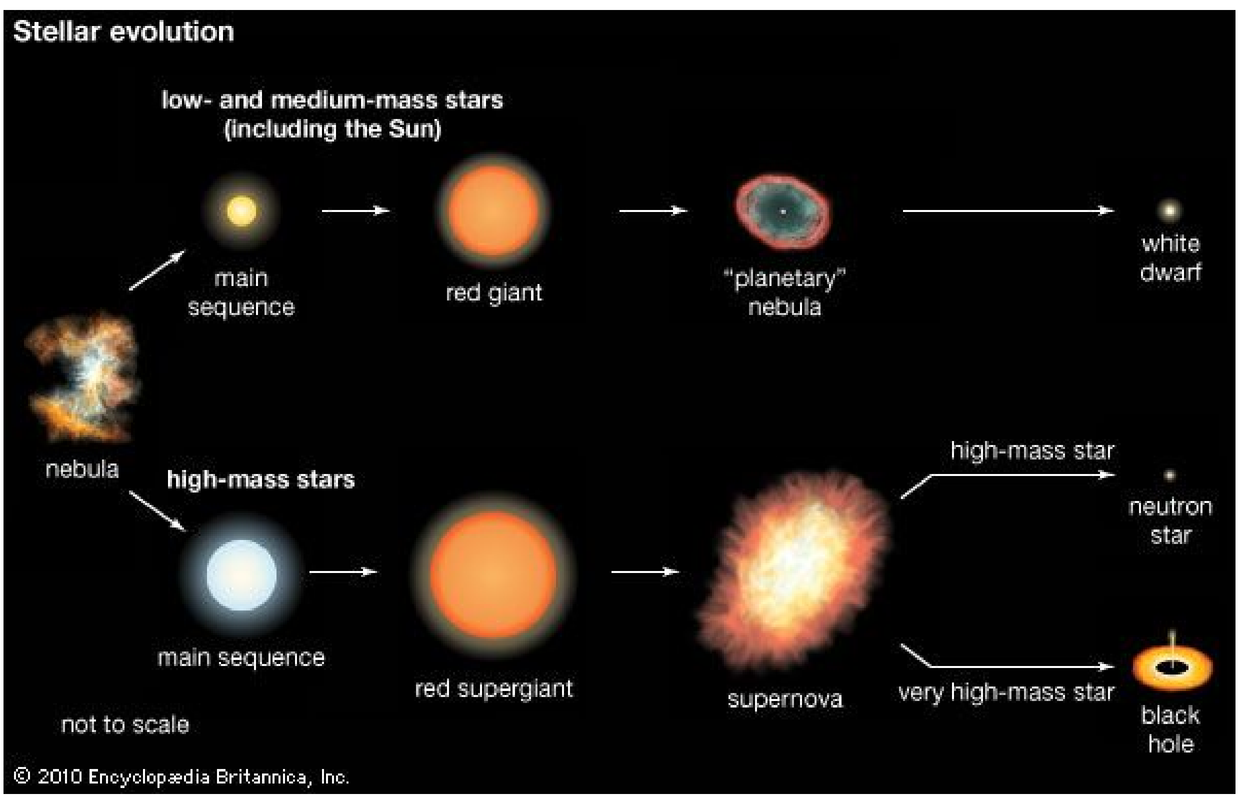
4. Stars can be classified based on their Luminosity and Temperature. We can group our data based on their type. They are usually classified into 6 categories:

- Brown Dwarf -> Star Type = 0

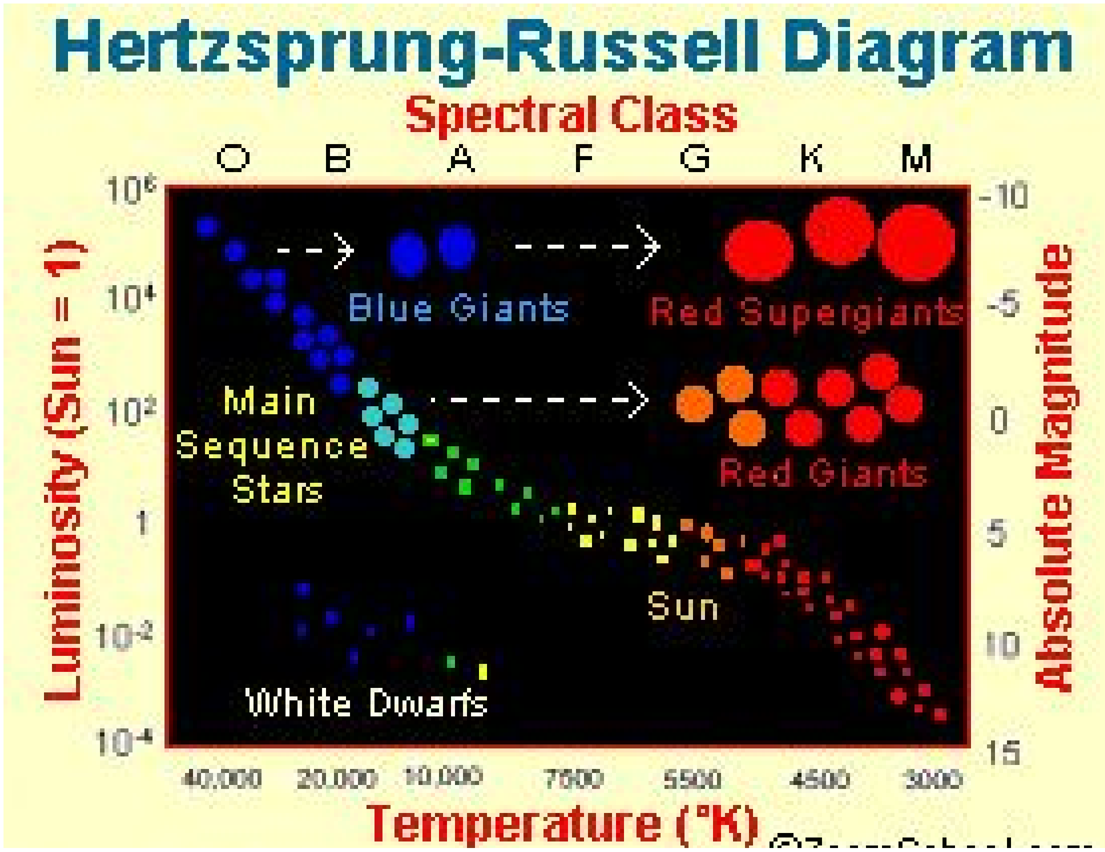
- Red Dwarf -> Star Type = 1
- White Dwarf-> Star Type = 2
- Main Sequence -> Star Type = 3
- Supergiant -> Star Type = 4
- Hypergiant -> Star Type = 5

Now, we can find the average Luminosity and Radius of the stars in each group. As we can see in the data frame 'df_g', Brown dwarfs, White and Red dwarfs have the smallest radii (as the name suggests). And Supergiants and Hypergiants have the greatest average radii.

- **Brown Dwarfs** are stars that have an extremely small radius and surface area, and can not reach the stage of nuclear fusion in their core. They do not ignite.
- **Red Dwarfs** are the smallest type of star on the main sequence. They are extremely common in the Milky Way, especially around the Sun.
- **White Dwarfs** are small in size (surface area) and also have a comparatively low luminosity. But they are very large in mass, and hence they are among the densest type of star.
- 90% of the stars in our universe are in the **Main Sequence**.Our Sun is a Main Sequence star. Main sequence stars fuse hydrogen atoms to form helium atoms in their cores.
- **Supergiants** are the most massively sized stars. There are two types of Supergiants, namely 'Red' and 'Blue'. These stars have short life spans (going as low as a few hundred thousand years).
- **Hypergiant** is a very rare kind of star. They usually shows extremely high luminosities and very high rates of mass loss.



5. The Hertzsprung - Russell (H-R) Diagram is a graph that plots **Absolute Magnitude** against the **Temperature** of the surface of the star. It is used by astronomers to classify stars according to their luminosity, spectral type, color, temperature and evolutionary stage. The diagram usually looks like this:



6. We can filter our data set to see the stars in the Main Sequence.

In [14]:

```
import pandas as pd
df= pd.read_csv('6 class csv.csv')
df.sample(n=25)
```

Out[14]:

	Temperature (K)	Luminosity(L/Lo)	Radius(R/Ro)	Absolute magnitude(Mv)	Star type	Star color	Spectral Class
28	11790	0.000150	0.01100	12.590	2	Yellowish White	F
31	30000	28840.000000	6.30000	-4.200	3	Blue-white	B
35	8052	8.700000	1.80000	2.420	3	Whitish	A
142	18290	0.001300	0.00934	12.780	2	Blue	B
43	3200	195000.000000	17.00000	-7.220	4	Red	M
37	6380	1.350000	0.98000	2.930	3	yellow-white	F
34	5800	0.810000	0.90000	5.050	3	yellow-white	F
128	2856	0.000896	0.07820	19.560	0	Red	M
111	3605	126000.000000	1124.00000	-10.810	5	Red	M
13	3628	0.005500	0.39300	10.480	1	Red	M
120	3323	0.000430	0.09120	17.160	0	Red	M
18	3192	0.003620	0.19670	13.530	1	Red	M
235	38940	374830.000000	1356.00000	-9.930	5	Blue	O
83	17200	0.000980	0.01500	12.450	2	Blue White	B
119	3780	200000.000000	1324.00000	-10.700	5	Red	M
38	5936	1.357000	1.10600	4.460	3	yellow-white	F
186	2968	0.000461	0.11900	17.450	0	Red	M
194	3523	0.005400	0.31900	12.430	1	Red	M
64	2935	0.000140	0.11600	18.890	0	Red	M
80	7100	0.000290	0.01200	14.090	2	White-Yellow	F
223	23440	537430.000000	81.00000	-5.975	4	Blue	O
126	2935	0.000870	0.09320	16.880	0	Red	M
165	7282	131000.000000	24.00000	-7.220	4	Blue	O
197	3496	0.001250	0.33600	14.940	1	Red	M
209	19360	0.001250	0.00998	11.620	2	Blue	B

In [15]:

```
#1
Sun_luminosity= 3.828 * (10**26) #Watts
df['Luminosity']=  df['Luminosity(L/Lo)']*Sun_luminosity

Sun_radius= 6.95700 * (10**8) #metres
df['Radius']= df['Radius(R/Ro)']*Sun_radius

df
```

Out[15]:

	Temperature (K)	Luminosity(L/Lo)	Radius(R/Ro)	Absolute magnitude(Mv)	Star type	Star color	Spectral Class	Lum
0	3068	0.002400	0.1700	16.12	0	Red	M	9.1872
1	3042	0.000500	0.1542	16.60	0	Red	M	1.9140
2	2600	0.000300	0.1020	18.70	0	Red	M	1.1484
3	2800	0.000200	0.1600	16.65	0	Red	M	7.6560
4	1939	0.000138	0.1030	20.06	0	Red	M	5.2826
...	...	...	...	...	...	...	...	
235	38940	374830.000000	1356.0000	-9.93	5	Blue	O	1.4348
236	30839	834042.000000	1194.0000	-10.63	5	Blue	O	3.1927
237	8829	537493.000000	1423.0000	-10.73	5	White	A	2.0575
238	9235	404940.000000	1112.0000	-11.23	5	White	A	1.5501
239	37882	294903.000000	1783.0000	-7.80	5	Blue	O	1.1288

240 rows × 9 columns

In [16]:

```
#2
import math
df['Surface Area']= 4*math.pi*( (df['Radius'])**2 )
df
```

Out[16]:

	Temperature (K)	Luminosity(L/Lo)	Radius(R/Ro)	Absolute magnitude(Mv)	Star type	Star color	Spectral Class	Lum
0	3068	0.002400	0.1700	16.12	0	Red	M	9.1872
1	3042	0.000500	0.1542	16.60	0	Red	M	1.9140
2	2600	0.000300	0.1020	18.70	0	Red	M	1.1484
3	2800	0.000200	0.1600	16.65	0	Red	M	7.6560
4	1939	0.000138	0.1030	20.06	0	Red	M	5.2826
...	...	...	...	...	...	...	...	
235	38940	374830.000000	1356.0000	-9.93	5	Blue	O	1.4348
236	30839	834042.000000	1194.0000	-10.63	5	Blue	O	3.1927
237	8829	537493.000000	1423.0000	-10.73	5	White	A	2.0575
238	9235	404940.000000	1112.0000	-11.23	5	White	A	1.5501
239	37882	294903.000000	1783.0000	-7.80	5	Blue	O	1.1288

240 rows × 10 columns

In [17]:

```
#3
# df_mainseq= df[df['Star type'] ==3]
# df['Area.Temp^4']= ((df_mainseq['Temperature (K)'])**4)*(df_mainseq['Surface Area'])
# df['Stefan-Boltzmann']= df_mainseq['Luminosity'] / ( df_mainseq['Area.Temp^4'] )
# df_mainseq
# df_mainseq['Stefan-Boltzmann'].median()
```

In [18]:

```
#4
group= df.groupby('Star type')
df_g= group.agg("mean").reset_index()
df_g
```

Out[18]:

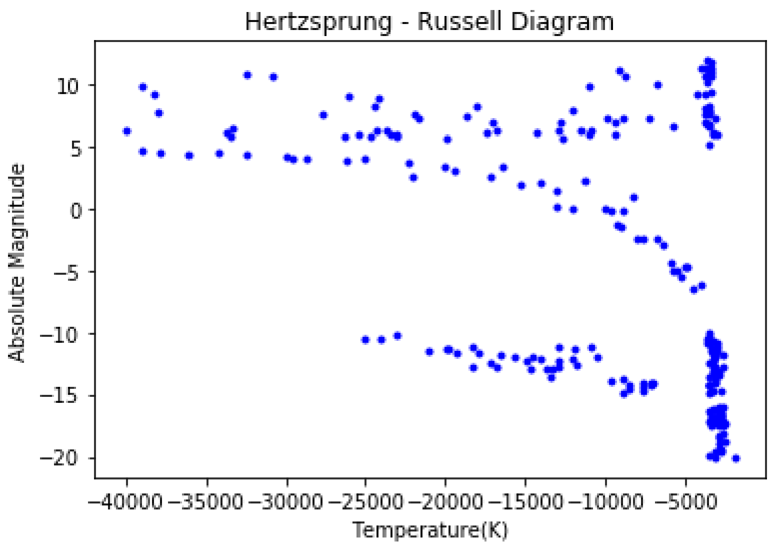
	Star type	Temperature (K)	Luminosity(L/Lo)	Radius(R/Ro)	Absolute magnitude(Mv)	Luminosity	Radi
0	0	2997.950	0.000693	0.110015	17.563500	2.653857e+23	7.653744e+
1	1	3283.825	0.005406	0.348145	12.539975	2.069321e+24	2.422045e+
2	2	13931.450	0.002434	0.010728	12.582500	9.315917e+23	7.463644e+
3	3	16018.000	32067.386275	4.430300	-0.367425	1.227540e+31	3.082160e+
4	4	15347.850	301816.250000	51.150000	-6.369925	1.155353e+32	3.558506e+
5	5	11405.700	309246.525000	1366.897500	-9.654250	1.183796e+32	9.509506e+

In [19]:

```
#5
import matplotlib.pyplot as plt
plt.plot(-(df['Temperature (K)']),-(df['Absolute magnitude(Mv)']),'b.')
plt.title('Hertzsprung - Russell Diagram')
plt.xlabel('Temperature(K)')
plt.ylabel('Absolute Magnitude')
```

Out[19]:

```
Text(0, 0.5, 'Absolute Magnitude')
```



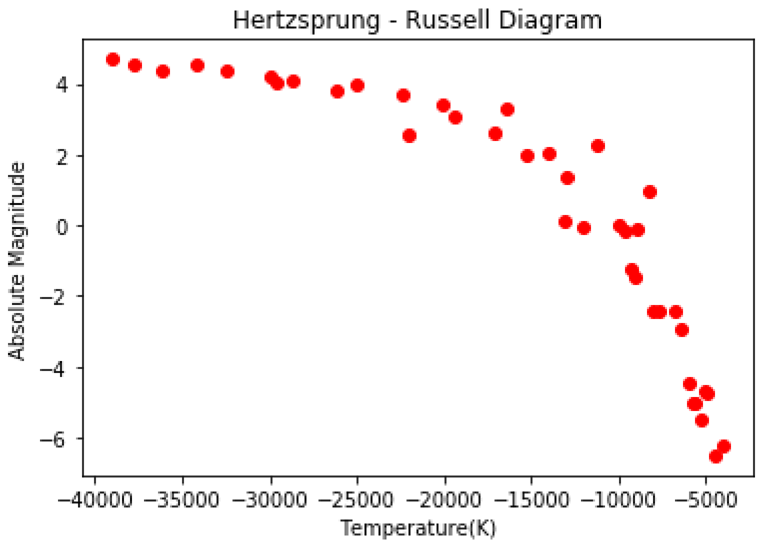
In [20]:

```
#6
#Let's see just the stars in the main sequence

plt.plot(-(df_mainseq['Temperature (K)']),-(df_mainseq['Absolute magnitude(Mv)']),'ro')
plt.title('Hertzsprung - Russell Diagram')
plt.xlabel('Temperature(K)')
plt.ylabel('Absolute Magnitude')
```

Out[20]:

```
Text(0, 0.5, 'Absolute Magnitude')
```



In [ ]: