# Stars by Type

This data set was pulled from Kaggle. It also contains the following information for each sample:

(1) temperature
(2) brightness
(3) size
(4) color
(5) class

This dataset is good for classification. It categorizes data by type of star and color. Using k = 1 with KNN it has an approximate 85% accuracy with temperature and luminosity as inputs.

You can use the following code to load the data into google colaboratory:

```python
import pandas as pd
url = "https://raw.githubusercontent.com/the-codingschool/TRAIN/main/stars/stars.csv"
df = pd.read_csv(url)
```