Observatory of Strasbourg

University of Strasbourg, France

On "J/AJ/147/129 White dwarfs within 25pc of the Sun (Sion+, 2014)"

January 26, 2024

Database Project Report

Vishrut Bezbarua

Supervised by:
Sebastien Derriere

Contents

1 Introduction

I selected *J/AJ/147/129* as the catalogue of my choice, for the purpose of this project. The associated article 'White dwarfs within 25pc of the Sun (Sion+, 2014)' [5] served as an important source to come up with interesting questions about the dataset which it analyses.

The paper [5] aims at efficiently managing and analyzing critical data related to white dwarfs within a 25 parsec radius around the Sun to explore and understand various astrophysical aspects, including the identification of the oldest white dwarfs, distribution of spectroscopic subgroups, and measurement of local space and mass density. Drawing on prior studies by Liebert et al. (1988)[3] and Oswalt et al. (1996)[4], the paper expands the dataset with new atmospheric parameters and more accurate distance measurements for a larger sample of local white dwarfs. This enhanced database facilitates a re-examination of statistical and kinematic properties, providing a reliable foundation for analysis. It aims to explore spatial distribution, space density, completeness, mass distribution, local luminosity function, and binary fraction within the 25 parsec sample. The results not only contribute to our understanding of local white dwarfs but also serve as a valuable reference for broader, less complete surveys.

The goal of this project can be considered to be more descriptive than analytical, with more focus on the variety of ways in which SQL can be used in order to extract different kinds of information. From simple queries, like printing of the first five columns to more complicated ones towards the end, the project proceeds with monotonically increasing levels of complicacy in using queries.

2 Information on the Dataset

This section elaborates on the "J/AJ/147/129" Catalogue, which presents the fractional distribution of spectroscopic subtypes, range and distribution of surface temperatures, and kinematical properties of the white dwarfs (WDs) within 25pc of the Sun. A total number of 224 WDs were selected as the sample. All columns, including the additional columns of positions generated by Vizier, were included. During the course of the project, the additional columns were deleted. For a more comprehensive understanding of the different parameters that the catalogue offers, I provide the list of parameters I use and their definitions, in the following :

- RAJ2000 : Right Ascension of the object.
- DEGJ2000 : Declination of the object.
- Name : Name of the object.
- K : Kinematic Properties
- SpT : Spectral Type.
- Teff : Effective Temperature.
- Dist : Distance from earth.
- n_Dist : Method of distance determination.
- Pm : Proper motion.
- PmPA : Position angle of proper motion vector
- Ref : Reference
- S_09 : Display the data from Sion et al.
- H_08 : Display the data from Holberg et al.
- M_C08 : Display the data from McCook et al.
- SIMBAD_name : Designation understandable by the Simbad data-base

Aladin [1] was used to plot the position WDs included in this catalogue in the sky map as shown in Figure. 1.

3 Creation of the Table

3.1 Data extraction

In this subsection I explain how data from the table of our desired VizieR catalogue has been extracted and loaded into the SQL space. The dataset is essentially made of a a single table, the extraction of which has been done using the following steps:

- Check the checkboxes for all the desired parameters.
- Set the preferences: max=unlimited, and Format=|-Separated-Values.

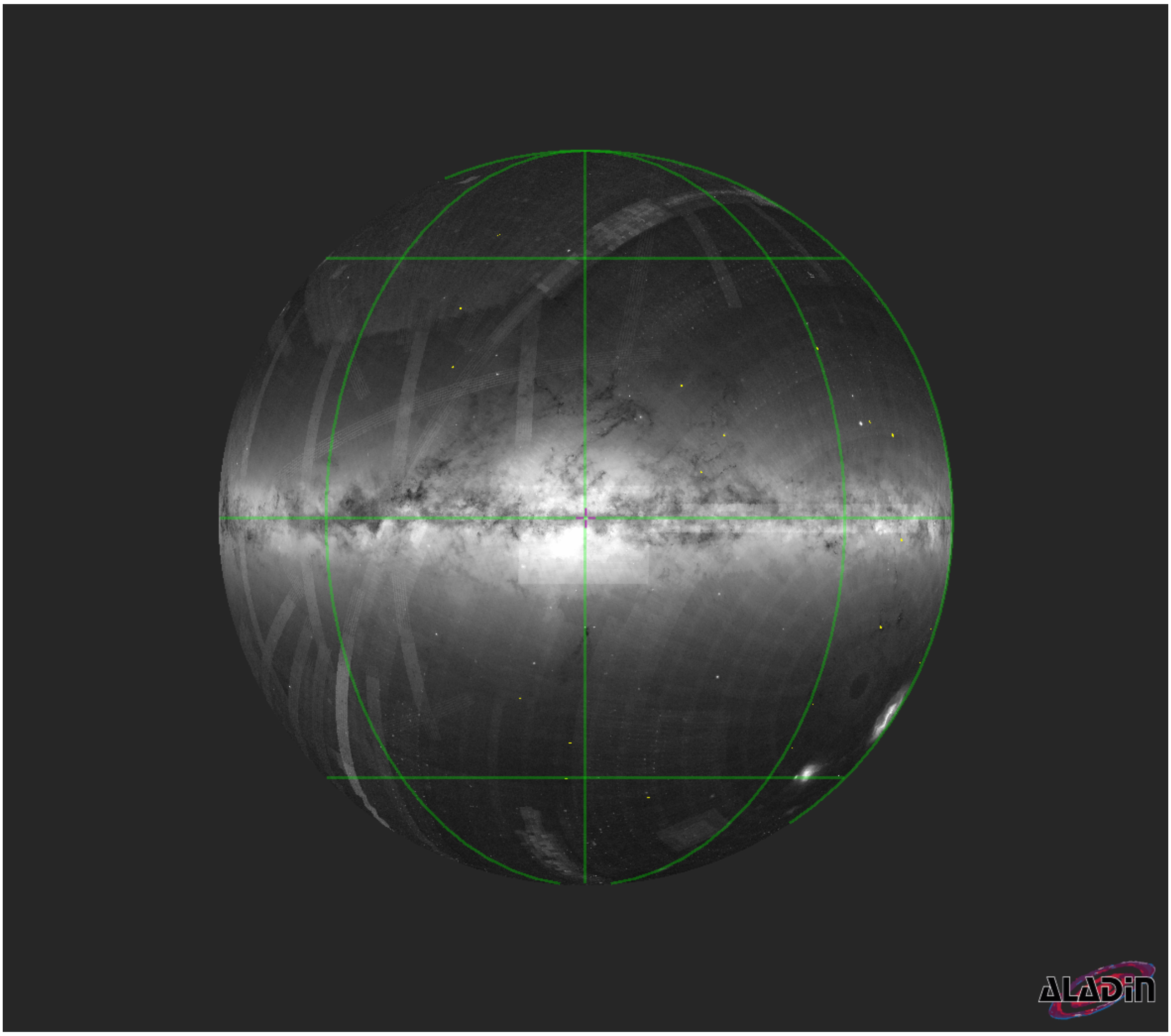

Figure 1: Positions of White Dwarfs considered in the catalogue using *Aladin*. [1]. The center of the area of interest is around.

- Save the results in a local text file.
- Remove the header and footer of the downloaded text file with a simple text editor and put them another file for reference.
- Exclude the first row, which just happens to be empty.

The following command was then used in order to create a file named "VB _ newwhitedwarfs", which would contain only the selected columns from the datafile.

```sql
CREATE TABLE VB_newwhitedwarfs (
RAJ2000 float NOT NULL,
DEGJ2000 float NOT NULL,
name char(10) NOT NULL,
K char(1) NULL,
SpT char(10) NOT NULL,
Teff int NULL,
Dist float NULL,
n_Dist char(2) NULL,
Pm float NULL,
PmPA float NULL,
Ref char(7) NULL,
S_09 int NULL,
H_08 int NULL,
MC08 int NULL,
SIMBAD_name char(23) NULL
);
```

To make sure that the table was successfully created, the following query which lists all tables in the m2 database was used:

```sql
show tables;
```

And then data corresponding to each column in the datafile was transferred to "VB _ newwhitedwarfs", using the query below.

```sql
LOAD DATA LOCAL INFILE 'white_dwarfs.dat' INTO TABLE VB_newwhitedwarfs
FIELDS TERMINATED BY '|' LINES TERMINATED BY '\n';
```

3.2 Printing 5 columns

In order to check if all the selected columns have successfully been transferred, I used the following query to print the first five rows of the table :

```sql
SELECT * FROM VB_newwhitedwarfs LIMIT 5;
```

And the output I received was :

RAJ2000	DEGJ2000	name	K	SpT	Teff	Dist	n_Dist	Pm	PmPA	Ref	S_09	H_08	MC08	SIMBAD_name
0.667	-34.2274	0000-345	K	DCP8.1	6643	13.21	p	0.7578	169.046	L20	1	1	7	WD 0000-345
2.84363	42.6781	0008+424	K	DA6.8	7380	22	sp	0.2328	191.648	L20	0	0	0	WD 0008+423
3.06167	50.4226	0009+501	K	DAH7.6	6502	11.03	p	0.715	219.92	L20	1	1	6	WD 0009+501
3.5533	-13.1836	0011-134	K	DAH8.4	5992	19.49	p	0.899	217.337	L20	1	1	7	WD 0011-134
3.45809	-71.8318	0011-721	K	DA7.8	6325	0	sp	0.326	141.3	GBD	1	0	0	WD 0011-721

5 rows in set (0.00 sec)

Figure 2: The first five tuples of the table.

3.3 Deletion of Unwanted Columns

The additional columns are : K, S_09, H_08, MC08, SIMBAD_name. In order to remove these columns, the following query was used :

```
ALTER TABLE VB_newwhitedwarfs DROP COLUMN K, DROP COLUMN S_09, DROP COLUMN H_08, DROP
COLUMN MC08, DROP COLUMN SIMBAD_name;
```

3.4 Printing the table

After removing the unwanted columns, the first five rows of the table were printed, the query used for which is the following :

```
SELECT * FROM VB_newwhitedwarfs LIMIT 5;
```

The result obtained is shown in the figure 3:

```
mysql> SELECT * FROM VB_newwhitedwarfs LIMIT 5;
+---------+---------+----------+--------+------+-------+--------+--------+---------+------+
| RAJ2000 | DEGJ2000 | name    | SpT    | Teff | Dist  | n_Dist | Pm     | PmPA    | Ref  |
+---------+---------+----------+--------+------+-------+--------+--------+---------+------+
|   0.667 | -34.2274 | 0000-345 | DCP8.1 | 6643 | 13.21 | p      | 0.7578 | 169.046 | L20  |
| 2.84363 |  42.6781 | 0008+424 | DA6.8  | 7380 |    22 | sp     | 0.2328 | 191.648 | L20  |
| 3.06167 |  50.4226 | 0009+501 | DAH7.6 | 6502 | 11.03 | p      |  0.715 |  219.92 | L20  |
|  3.5533 | -13.1836 | 0011-134 | DAH8.4 | 5992 | 19.49 | p      |  0.899 | 217.337 | L20  |
| 3.45809 | -71.8318 | 0011-721 | DA7.8  | 6325 |     0 | sp     |  0.326 |   141.3 | GBD  |
+---------+---------+----------+--------+------+-------+--------+--------+---------+------+
5 rows in set (0.00 sec)
```

Figure 3: The first five tuples of the *reduced* table.

3.5 Counting the number of objects

In order to know the total number of objects that have been used in the survey, and also whether the number is consistent in my table, the following command was used :

```
SELECT COUNT(Name) FROM VB_newwhitedwarfs;
```

This gave me the total number of rows for the column "Name" which contains the names of all the WDs.

```
+-------------+
| COUNT(Name) |
+-------------+
|         223 |
+-------------+
```

Total number of White dwarfs in the table extracted = 223, as opposed to 224, as the first line was deleted, for being empty.

4 Exploring Data

This section describes the queries I used for the purpose of exploring the different properties of the dataset. Each subsection represents the questions I asked, the contents being the queries I used to answer them and the associated results.

4.1 Display the coordinates of the White-dwarf "1820+609"

The query used for this purpose is :

```
SELECT RAJ2000, DEGJ2000 FROM VB_newwhitedwarfs WHERE Name="1820+609";
```

The output received was the following:

```
+---------+----------+
| RAJ2000 | DEGJ2000 |
+---------+----------+
| 275.333 |  61.0189 |
+---------+----------+
1 row in set (0.11 sec)
```

4.2 How many different White-Dwarfs are there in this table?

The query used to answer this question was :

```
SELECT COUNT(Distinct Name) FROM VB_newwhitedwarfs;
```

The output being :

```
+----------------------+
| COUNT(Distinct Name) |
+----------------------+
|                  222 |
+----------------------+
1 row in set (0.02 sec)
```

4.3 Which is the farthest White Dwarf in this table?

The query used to answer this question was :

```
SELECT Name, Dist FROM VB_newwhitedwarfs WHERE Dist = (SELECT MAX(Dist) FROM
VB_newwhitedwarfs);
```

And the output being:

```
+------------+------+
| Name       | Dist |
+------------+------+
| 1012+083.1 | 25.8 |
+------------+------+
1 row in set (0.11 sec)
```

Note, that the distance here is in parsecs.

5 Playing with Statistics

This section dives into the various statistical information I have extracted using the queries mentioned.

5.1 What is the percentage of "DA"type White-dwarfs in the sample, within an Teff range of "4590-25193" K, and up to a distance of 25pc?

DA white dwarfs are characterized by a surface composition that consists almost entirely of H with at most traces of other elements. They comprise about 85% of all white dwarfs. Their overwhelming presence makes them interesting types to do statistics with. Here, we find out the percentage of DA type WDs in our sample.

The query used to answer this question was :

```sql
SELECT (COUNT(DISTINCT Name)*100)/(SELECT COUNT(*) FROM VB_newwhitedwarfs)FROM
VB_newwhitedwarfs WHERE SpT REGEXP '^DA[0-9]*\\.?[0-9]*' AND Teff BETWEEN 4590
AND 25193 AND Teff IS NOT NULL  AND (Dist IS NOT NULL AND Dist != '') AND Dist < 25;
```

The output looks like this:

```
+----------------------------------------------------------------+
| (COUNT(DISTINCT Name)*100)/(SELECT COUNT(*) FROM VB_newwhitedwarfs) |
+----------------------------------------------------------------+
|                                                        59.1928 |
+----------------------------------------------------------------+
1 row in set (0.35 sec)
```

5.2 Fraction of Magnetic WDs

White dwarfs (WDs) are the last stage for the majority of stars, and more than 20% of those stars possess strong magnetic fields. The field strengths encountered range from tens of kG up to about 1000 MG, and are roughly dipolar. An interesting question addressed also in the paper, was the fraction occupied by magnetic WDs in the total sample considered, with a criterion on distance being less than 25 pc. So, I try to find the same.

I begin with finding the exact number of magnetic WDs, since the total number of the sample is not that big.

```sql
SELECT COUNT(DISTINCT Name) AS CountDistinct  FROM VB_newwhitedwarfs
WHERE SpT REGEXP '^D[A-Za-z][HP][0-9]*\\.?[0-9]*' AND Dist < 25;
```

The output of it being:

```
+---------------+
| CountDistinct |
+---------------+
|            19 |
+---------------+
1 row in set (0.01 sec)
```

which is the **exact** number mentioned in the paper [5].

To convert it into the corresponding percentage, the following query was used:

```sql
SELECT (COUNT(DISTINCT CASE WHEN SpT REGEXP '^D[A-Za-z][HP][0-9]*\\.?[0-9]*' AND
Dist < 25 THEN Name END) * 100) / COUNT(DISTINCT Name) AS Percentage_Magnetic_WDs
FROM VB_newwhitedwarfs;
```

And the output being:

```
+-------------------------+
| Percentage_Magnetic_WDs |
+-------------------------+
|                  8.5586 |
+-------------------------+
```

To see if the query really worked, I used the following query to siplay the names and Spectral Type of all the WDs that fit the criterion of the previous query :

```sql
SELECT Name, SpT FROM VB_newwhitedwarfs WHERE SpT REGEXP 'D[A-Za-z][HP][0-9]*\\.?[0-9]*'
ORDER BY SpT;
```

I order them by SpT (Spectral Type) in order to have more ease while manually counting the DA and Non-DA magnetics, as can be seen in the Table: 1, as the output.

5.3 Ratio Between DA Magnetics and Magnetic Non-DAs

An interesting fact to look into might be to see the proportion between DA-Magnetics and Magnetic Non-DAs. To find it out, I used the following query:

```sql
SELECT COUNT(DISTINCT CASE WHEN SpT REGEXP '^DA[HP][0-9]* \\.? [0-9]*' AND Dist < 25
THEN Name END) AS Magnetic_DA, COUNT(DISTINCT CASE WHEN SpT REGEXP '^D[^A][HP][0-9]*
\\.? [0-9]*' AND Dist < 25 THEN Name END) AS Magnetic_Non_DA, (COUNT(DISTINCT CASE WHEN
SpT REGEXP '^DA[HP][0-9]*\\.? [0-9]*' AND Dist < 25 THEN Name END) / COUNT(DISTINCT CA
SE WHEN SpT REGEXP '^D[^A][HP][0-9]*\\.?[0-9]*' AND Dist < 25 THEN Name END)) AS Ratio
FROM VB_newwhitedwarfs;
```

The output received was :

```
+-------------+----------------+--------+
| Magentic_DA | Magentic_Non_DA | Ratio  |
+-------------+----------------+--------+
|          12 |               7 | 1.7143 |
+-------------+----------------+--------+
1 row in set (0.00 sec)
```

6 Conclusions

This project has certainly augmented my SQL skills, showing me the variety of ways in which this potent language can be used to extract, organise and manipulate data, especially in the context of astrophysics. As mentioned in the introduction as well, the goal of this project was to put more focus on the versatility and flexibility of the SQL Language itself, than the analysis of the paper [5] used. However, there can certainly be a few things to be said about the results that I obtained.

One of my intentions also was to verify if certain statistical relations that have already been dealt with in the paper [5], do correspond well to the results I obtained. The answer is not straightforward, as some queries do manage to give me the exact results and some don't. I suspect that it could be due to the way the data in the tables are arranged in VizieR itself. I submit this reason in particular to be a potential cause of discrepancies, as I saw a difference in the total number of objects (WDs) in the paper [5] (=224), and in the VizieR tables (=223), which reduce to 222 unique objects, when used the DISTINCT command in SQL.

The primary focus of the paper remains on objects within 25 pc range of distance, notwithstanding the fact that the farthest object happens to be 25.8 pc away, as displayed upon using the query for the corresponding question. The first game with statistics starts with the question 5.1, which can exactly be found in table 2 of [5]. The result I obtained (59.19%) differs a bit from the number we get to see in the paper [5] (54%). I then move on to find out the fraction of magnetic WDs contained within the sample we have, and manage to find the exact number (=19) of distinct magnetic WDs in the sample, and the corresponding percentage (=8.5586%) as well. I encounter, however, an interesting departure from the results obtained in the paper, when it comes to finding the ratio between the DA and non-DA type magnetic WDs. I found the ratio to be 12:7, as opposed to 14:5 in the paper. One possible reason could be that my definition of a DA-type WD might be constricted and might not include any object, the SpT of which does not start with DA. As seen in the table 1 too, manually counting the DA-type WDs gives us a total of 12, as opposed to 14. It is worth mentioning also that I classify a WD as being "magnetic" if its spectral type indicates "P" or "H", where “P” stands for polarisation measurements and “H” for Zeeman splitting [2].

Name	SpT
1639+537	DAH6.7
1344+106	DAH7.1
0009+501	DAH7.6
0121-429	DAH7.9
0011-134	DAH8.4
0503-174	DAH9.5
1658+440	DAP1.7
0728+642	DAP11.1
1900+705	DAP4.2
1350-090	DAP5
0553+053	DAP8.7
1309+853	DAP9
0912+536	DCP7
0000-345	DCP8.1
1829+547	DQP8.0
0548-001	DQP8.3
1008+290	DQpec11.0
0038-226	DQpec9.3
1036-204	DQpecP10.2

Table 1: List of Magnetic WDs

I have my suspicions about the object 0038-226, as I did not mange to find conclusive evidence of it being a magnetic non-DA type WD. However, 1008+290 can safely be classified as magnetic, thanks to Berdyugin et al (2023)[6], as well as 1036-204 for being polarised.

The analysis of the results surely is an interesting exercise, but also impossible without the application of

SQL. It serves as a perfect tool to optimise data acquisition and manipulation, and was very helpful for the purpose of this project.

References

[1] F. Bonnarel et al. "The ALADIN interactive sky atlas". In: *Astronomy and Astrophysics Supplement Series* 133 (Apr. 2000), pp. 33–40.

[2] L. Ferrario, D. deMartino, and B. T. Gänsicke. "Magnetic White Dwarfs". In: *Space Sci.Rev.* 191 (2015).

[3] J. Liebert, C. Dahn, and D. Monet. "The Luminosity Function of White Dwarfs". In: *The Astronomical Journal* 332.891 (1988).

[4] T. Oswalt et al. "The White Dwarfs within 25 pc of the Sun". In: *Nature* 382.692 (1996).

[5] Edward M. Sion et al. "The White Dwarfs within 25 pc of the Sun: Kinematics and Spectroscopic Subtypes". In: *The Astronomical Journal* 147.6 (2014).

[6] Berdyugin A. V. et al. "Discovery of magnetic fields in five DC white dwarfs". In: *A & A* 670 (Feb. 2023).