



Астрономическое
образование

2026, № 2

Астрономия и школа

Ключевые слова:

астрономические каталоги, анализ данных, информатика, интеграция дисциплин, Python, визуализация данных, исследовательская деятельность

Сокращения:

CSV — comma-separated values (значения, разделённые запятыми);
FTP — File Transfer Protocol (протокол передачи файлов);
IT — информационные технологии

*dyuk108@gmail.com

Игра в большие данные: работа с астрономическими каталогами на уроках информатики в старших классах

Д. Ю. Клыков*

Школа в Капотне, г. Москва

Аннотация

В работе рассказывается, как можно «оживить» уроки информатики в старших классах, используя реальные астрономические данные. Автор подготовил учебный набор данных с характеристиками ярчайших звёзд неба и разработал сценарии практических занятий. Работая на языке Python с библиотеками Pandas, Matplotlib и Plotly, учащиеся осваивают фильтрацию, обработку и визуализацию данных, в том числе воссоздают знаменитую диаграмму Герцшпрунга — Рассела. Материалы могут быть встроены в курсы информатики и физики и использованы для поддержки межпредметных связей. Набор данных, программный код и руководства опубликованы в открытом доступе и предназначены для уроков, проектной и кружковой работы, а также самостоятельных учебных исследований. Рассмотренный подход позволяет показать школьникам, что большие данные — инструмент для исследования мира, пусть даже звёздного.

1. Введение

В современной школе многие 10–11-е классы имеют профильную направленность. В ходе обучения школьники знакомятся с элементами будущей профессии и основами передовых технологий. В рамках проекта «IT-класс в московской школе» автор проводит уроки профильной информатики и предмета «Введение в специальность», а ранее также уроки физики.

IT-отрасль работает с большими массивами данных, извлекая из них полезную информацию. Подобный подход применяется и в науке, в частности в звёздной статистике. Элементы астрономии изучаются в курсе физики 11 класса. При подготовке к теме «Введение в большие данные» в 2024/2025 учебном году родилась идея в рамках углубления межпредметных связей использовать астрономические каталоги для проведения практических работ.

Результатом работы стал набор данных (датасет) со сведениями о ярчайших звёздах неба, собранными из нескольких источников, а также разработки уроков информатики с его использованием. Датасет, исходный код и учебные материалы опубликованы в [1; 2]. Их можно использовать для практических работ на уроках информатики и физики (астрономии), проектных работ, кружковых занятий и самостоятельных экспериментов любителей астрономии.

2. Исходные данные

Требовалось подготовить исходные данные так, чтобы их можно было сразу открыть и исследовать, а все необходимые характеристики звёзд уже находились в одной таблице. В качестве основных источников рассматривались «Каталог ярких звёзд» (Bright Star Catalog, BSC5 [3]) и каталог Hipparcos [4]. Некоторые их характеристики сопоставлены в табл. 1.

Поступило в редакцию	22.02.2026
После доработки	18.06.2026
В печать	08.08.2026
Опубликовано	12.08.2026

ISSN 3033-7917
journal.astroedu.ru

Таблица 1. Сравнительные характеристики каталогов BSC5 и Hipparcos для среза по видимой звёздной величине $V \leq 6.5^m$

Характеристика	BSC5	Hipparcos
Число записей в срезе	8 404	8 874
Эпоха	J2000	J1991.25
Видимая звёздная величина V	Есть	Есть
Тригонометрический параллакс	37,7 %	99,9 % (отсутствует у 4 звёзд)
Показатель цвета $B - V$	97 %	99,9 % (отсутствует у тех же 4 звёзд)
Спектральный класс	100 %	100 %

Предельную видимую звёздную величину установили равной 6.5^m : это значение приблизительно соответствует пределу чувствительности невооружённого глаза для идеального ночного неба. В полученном срезе оказалось 8 404 записи из BSC5 и 8 874 записи из Hipparcos. Разница обусловлена различием полноты (степени охвата) каталогов и попаданием в срез звёзд с блеском вблизи граничного значения (особенно это касается переменных звёзд). Также выяснилось, что некоторые двойные звёзды, занимающие по одной позиции в Hipparcos, представлены в BSC5 несколькими строками.

Причин ограничиться только видимыми глазу звёздами несколько. Во-первых, для школьных уроков требовался небольшой набор данных, который гарантированно можно обработать на слабых компьютерах. Во-вторых, связь с непосредственно наблюдаемым небом помогает сформулировать понятные учебные задачи. Более слабые звёзды можно использовать, например, при изучении Млечного Пути методом «звёздных черпков» или звёздной статистики околосолнечного пространства. Такая работа потребует обращения к более крупным каталогам, например Gaia.

Вследствие большей полноты и лучшей точности решено взять за основу каталог Hipparcos. Из него перенесены, в числе прочих:

- видимая звёздная величина V в системе Джонсона (V_{mag}),
- прямое восхождение и склонение в десятичных градусах (RA_{deg} , DE_{deg}),
- тригонометрический параллакс (Pl_x),
- показатели цвета $B - V$ и $V - I$,
- спектральный класс ($SpType$),
- каталожные номера Hipparcos (HIP) и Henry Draper Catalogue (HD), позволяющие сопоставлять записи с другими источниками.

Обозначения звёзд по Байеру (Bayer, Bayer_cst) и Флемстиду (Fl) добавлены из кросс-каталога [5]. Для каждой звезды также указано созвездие по каталогу границ созвездий [6], содержащему данные для эпохи B1875. Эти поля можно использовать, например, для сравнения созвездий по числу ярких звёзд и площади.

2.1. Получение каталогов

Наряду с домашними страницами каталогов действует Страсбургский центр астрономических данных (Centre de Données astronomiques de Strasbourg, CDS) <https://cds.unistra.fr>. Один из вариантов доступа к архиву — анонимный FTP-сервер <ftp://cdsarc.u-strasbg.fr>. Поскольку современные браузеры, как правило, не поддерживают нешифрованный FTP, предлагается использовать FTP-клиент, например FileZilla. После ввода имени пользователя anonymous и пароля, содержащего символ @, открывается дерево каталогов (рис. 1).

Сориентироваться в дереве папок помогает файл `/0/cats/cats.lis` с перечнем названий и авторов каталогов. Поиск по ключевым словам позволяет найти нужный источник и его положение на сервере. Файл с названием наподобие ReadMe обычно содержит описание каталога и структуры его полей.

3. Форматы файлов данных

Если в исходных файлах поля разделены фиксированными позициями в строке, для учебных упражнений их удобно преобразовать в формат CSV, где поля разделяет специальный символ, в большинстве случаев — запятая. С файлами в формате CSV работают библиотеки Python Pandas и Polars, многие системы управления базами данных и электронные таблицы, в частности LibreOffice Calc.

Учебные исследования можно проводить и в электронных таблицах. При открытии CSV-файла Calc показывает панель настроек (рис. 2). Разделитель обычно определяется автоматически, а кодировку следует установить вручную — Unicode (UTF-8), поскольку в названиях звёзд присутствует кириллица.

В русской локализации запятая служит десятичным разделителем, поэтому табличные программы часто ожидают точку с запятой в качестве разделителя полей. Для такого случая подготовлен отдельный вариант файла: поля разделены точкой с запятой, а дробная часть числа — запятой.

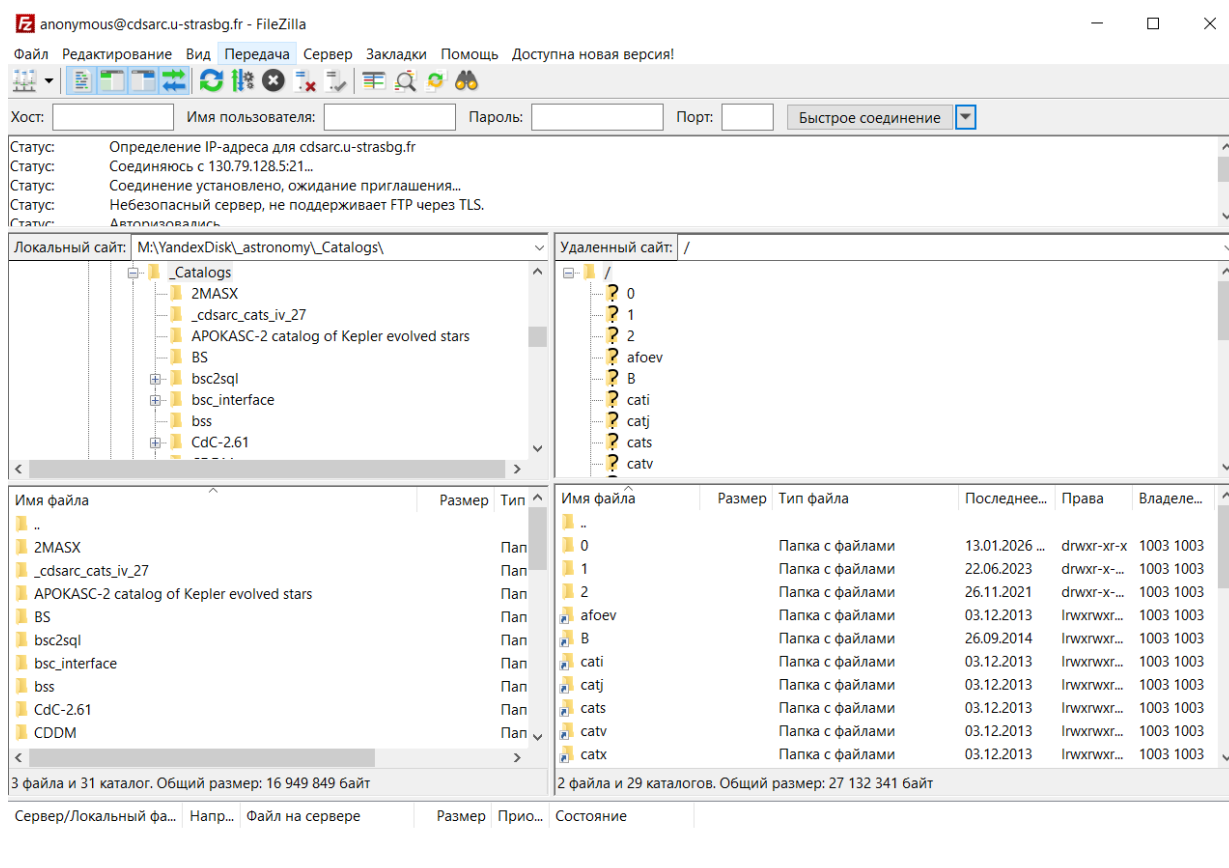


Рис. 1. Окно FileZilla с открытым FTP-архивом CDS

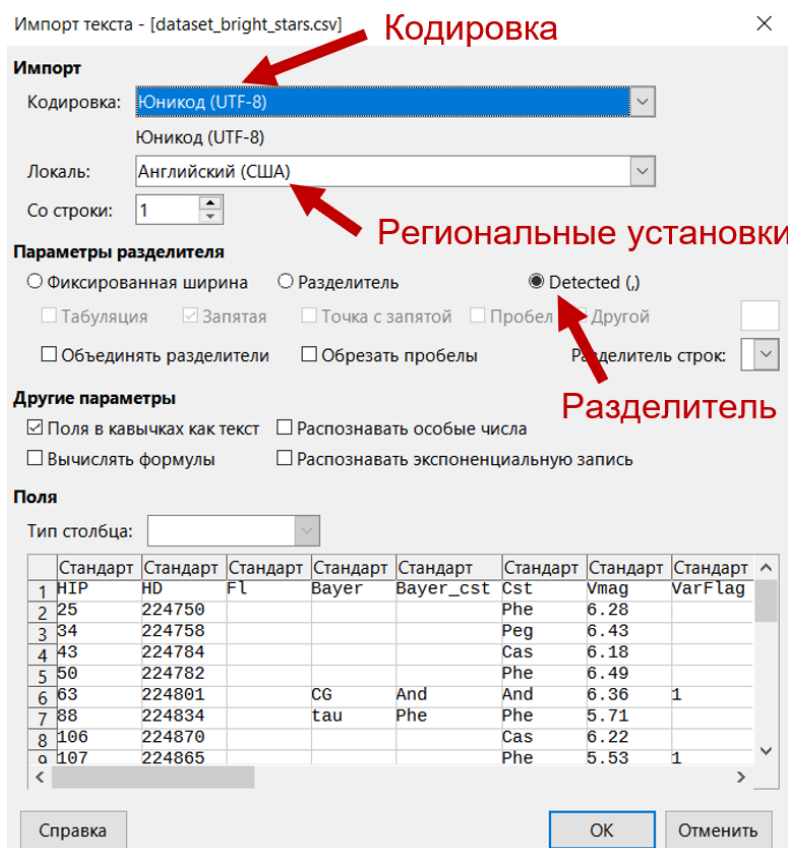


Рис. 2. Параметры импорта CSV-файла в LibreOffice Calc

В IT-классе, где автор проводил занятия, одиннадцатиклассники уже владели основами Python. Для работы с небольшим набором данных объёмом около 1 мегабайта использовалась библиотека Pandas.

Данные загружаются в память двумя строками:

```
import pandas as pd
df = pd.read_csv('dataset_bright_stars.csv',
                 encoding='utf-8')
```

Полученный объект `df` называют датафреймом: его строки соответствуют звёздам каталога, а именованные столбцы — характеристикам звёзд (рис. 3).

	HIP	HD	Fl	Bayer	Bayer_cst	Cst	Vmag	VarFlag	VarID	RAdeg
0	25	224750.0	NaN	NaN	NaN	Phe	6.28	NaN	NaN	0.079365
1	34	224758.0	NaN	NaN	NaN	Peg	6.43	NaN	NaN	0.099470
2	43	224784.0	NaN	NaN	NaN	Cas	6.18	NaN	NaN	0.129089
3	50	224782.0	NaN	NaN	NaN	Phe	6.49	NaN	NaN	0.142871
4	63	224801.0	NaN	CG	And	And	6.36	1.0	CG And	0.181750

Рис. 3. Фрагмент датафрейма в блокноте Jupyter

4. Сценарии уроков

После изучения закона излучения Планка и характеристик звёзд учащиеся переходят к программированию. Предполагается, что ранее они уже познакомились с Pandas и научились устанавливать библиотеку. Учитель раздаёт заранее подготовленные руководства с пошаговым объяснением кода, а учащиеся скачивают набор данных со страницы [1].

4.1. Фильтрация и описательная статистика

Сначала учащиеся выводят численные характеристики датафрейма. Следующие строки знакомят их с фильтрацией по условию и подсчётом числа строк:

```
print('Количество звёзд ярче 6m:',
      (df['Vmag'] <= 6).sum())
print('Из них в северном полушарии:',
      ((df['Vmag'] <= 6)
       & (df['DEdeg'] >= 0)).sum())
print('Из них в южном полушарии:',
      ((df['Vmag'] <= 6)
       & (df['DEdeg'] < 0)).sum())
```

Получаются значения 5 044, 2 400 и 2 644 соответственно: в этом наборе данных звёзд ярче 6^m в южном небесном полушарии несколько больше, чем в северном. Для звёзд ярче 2^m соотношение аналогичное: 29 звёзд в южном полушарии против 20 — в северном.

Затем отбираются звёзды с отрицательной видимой звёздной величиной, выводятся интересные столбцы и записи сортируются по возрастанию `Vmag`:

```
df_bright = df[df['Vmag'] <= 0]
df_bright[['Name_r', 'Bayer', 'Bayer_cst',
           'Vmag']].sort_values(by='Vmag')
```

В результате выводится список четырёх ярчайших звёзд ночного неба (рис. 4). На этом примере учитель объясняет шкалу видимых звёздных величин и связывает численные значения с наблюдаемой картиной неба.

	Name_r	Bayer	Bayer_cst	Vmag
2446	Сириус	alf	CMa	-1.44
2274	Канопус	alf	Car	-0.62
5234	Арктур	alf	Boo	-0.05
5367	Ригиль Кентаврус	alf	Cen	-0.01

Рис. 4. Четыре ярчайшие звезды ночного неба

4.2. Спектральные классы и диаграмма Герцшпрунга — Рассела

Следующая часть практической работы посвящена спектральным классам. С помощью библиотеки Matplotlib строится столбчатая диаграмма их распределения (рис. 5).

После этого можно перейти к обсуждению эволюции звёзд. На уроке информатики материал закрепляется построением диаграммы Герцшпрунга — Рассела: учащиеся не только видят готовую диаграмму на слайде или в учебнике, но и строят её самостоятельно по имеющимся данным.

По вертикальной оси откладывается абсолютная звёздная величина. Она связана с видимой звёздной величиной m и расстоянием r , выраженным в парсеках:

$$M = m + 5 - 5 \lg r. \quad (1)$$

В датасете параллакс p приведён в миллисекундах дуги, поэтому при вычислении расстояния используется соотношение $r = 1000/p$. Отрицательные оценки параллакса не имеют прямого физического смысла, но могут возникать при обработке измерений, когда неопределённость сопоставима с самим параллаксом. На данном примере юные программисты учатся фильтровать зашумлённые данные; записи с $p < 10$ mas отбрасывались.

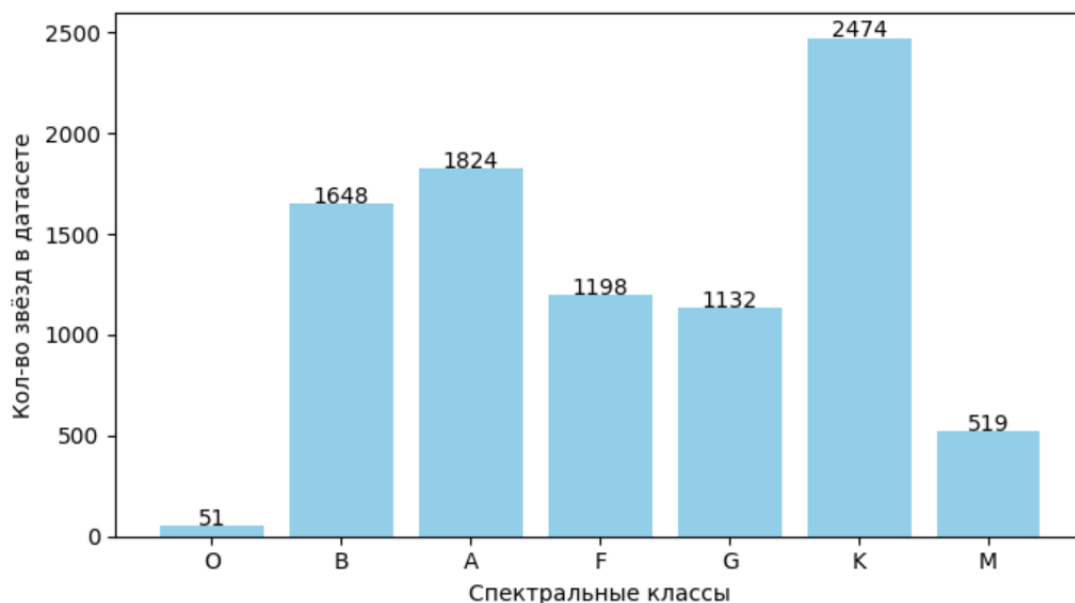


Рис. 5. Распределение звёзд с $V \leq 6.5^m$ по спектральным классам

По горизонтальной оси откладывается эффективная температура T_{eff} ; шкала логарифмическая, причём её направление инвертировано (влево). В Hipparcos такого поля нет, поэтому использовались типичные температуры спектральных классов и подклассов из модуля автора [7]. Из-за дискретности этих значений на диаграмме появились вертикальные полосы (рис. 6).

Для уточнения температур был использован каталог параметров звёздных атмосфер PASTEL [8]. Кросс-идентификация звёзд потребовала дополнительной работы, поскольку звёзды в каталоге обозначены по разным системам. В процессе возникло множество вопросов, например:

- Что делать с двойными или кратными звёздами?
- Брать ярчайший компонент.
- Приводятся несколько различных значений T_{eff} , что делать?
- Брать среднее значение.
- Как проверить корректность значений?
- По известному спектральному классу.

Разумеется, результат непригоден для использования в научном исследовании, однако вполне подходит для учебных целей.

В итоге значения T_{eff} удалось сопоставить с 62.7 % записей набора. Если температура отсутствует в PASTEL, она по-прежнему оценивается по спектральному классу. С дополненными значениями температуры диаграмма стала более информативной (рис. 7).

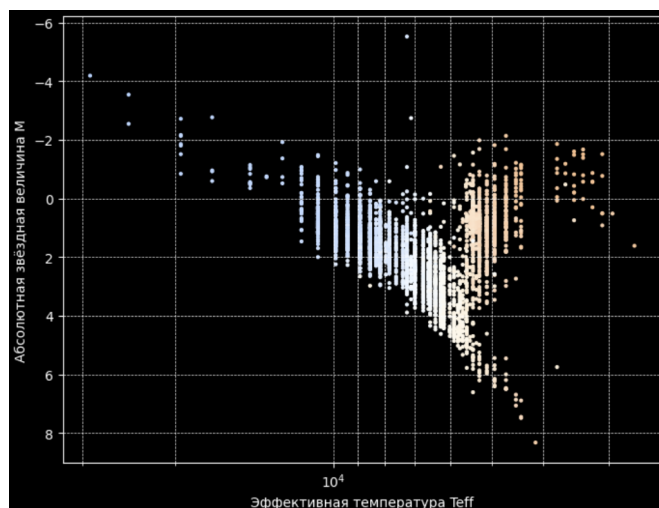


Рис. 6. Диаграмма Герцшпрунга – Рассела. T_{eff} оценена по спектральным классам

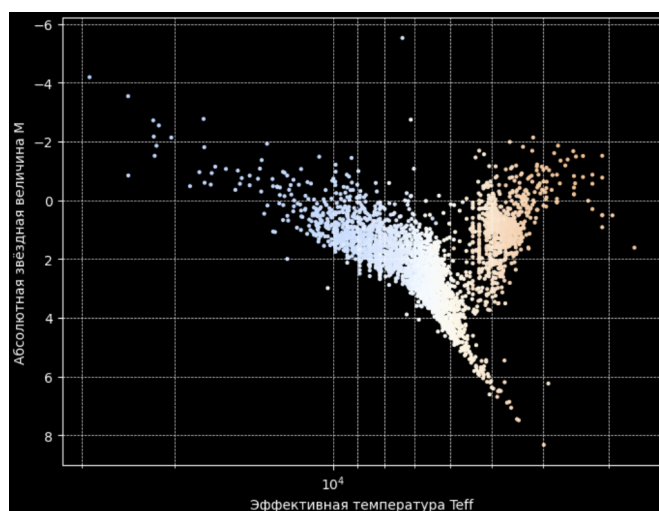


Рис. 7. Диаграмма Герцшпрунга – Рассела. T_{eff} взята из PASTEL, а при отсутствии данных оценена по спектральному классу



Рис. 8. Вид неба над южным горизонтом сентябрьским вечером на широте Москвы



Рис. 9. Учащиеся 11 класса строят диаграмму Герцшпрунга — Рассела



Рис. 10. Учащиеся 11 класса создают объёмные модели ярких звёзд

Для проведения таких уроков учитель должен обладать достаточной квалификацией как в программировании, так и в астрономии, сотрудничать с учителем физики или быть им. Ход занятия зависит также от подготовки и заинтересованности учащихся.

5. Идеи дальнейшего использования набора данных

Интерактивные диаграммы позволяют исследовать отдельные точки с помощью всплывающих подсказок. Автор создал интерактивную диаграмму Герцшпрунга — Рассела в Plotly [1]. В подсказках отображаются названия звёзд и обозначения по Байеру. Подобная визуализация может стать основой детского проекта:

небольшого интерактивного сайта, который объясняет астрономическое явление и позволяет пользователю проявить исследовательскую активность.

Другая группа работ может быть посвящена пространственному расположению светил. Объёмную модель ярких звёзд созвездия, традиционно создаваемую из подручных материалов, можно построить и на компьютере. На одном из уроков одиннадцатиклассники создавали 3D-модели в Plotly (рис. 10).

Ещё одно направление — построение карт. Автор разработал скрипт для генерации векторных звёздных карт [9]; пример приведён на рис. 8. Продолжением может стать учебный проект тайловой карты неба с использованием, например, MapLibre GL JS.

6. Заключение

Описанный опыт позволяет сформулировать следующие методические наблюдения и рекомендации:

1. Создание учителем личных проектов, связанных с программированием, не только повышает его квалификацию и, как следствие, улучшает подготовку учащихся, но и позволяет использовать их как учебный материал.

2. Способы учебного применения проекта полезно планировать до начала разработки. Необходимо публиковать результаты работы и написать проектную документацию.

3. Личный проект направлен на решение некоторой конкретной задачи или круга задач в заданной проблемной области. Межпредметный контекст помогает учащимся осваивать не только абстрактные технологии, но и предметную область, в которой эти технологии применяются, а также способствует более осознанной профессиональной ориентации.

4. При должной организации работы изложенный подход знакомит учащихся с мощными инструментами учебно-исследовательской деятельности, является активной формой приобретения знаний.

5. Подготовка таких занятий трудозатратна, однако позволяет учителю получать удовлетворение и вдохновение от творческой работы, тем самым служит профилактикой профессионального выгорания и средством самореализации.

Список литературы

- [1] Клыков Д. Ю. URL: <http://dyuk108.ru>; Личная страница автора статьи.
- [2] Клыков Д. Ю. Датасет сведений о ярчайших звёздах неба до 6.5m : Репозиторий GitHub. 2025. URL: https://github.com/dyuk108/brightstar_dataset.
- [3] Hoffleit D., Warren Jr. W. H. Bright Star Catalog, 5th Revised Edition (Preliminary Version) / HEASARC, NASA Goddard Space Flight Center. 1991. URL: <https://heasarc.gsfc.nasa.gov/W3Browse/star-catalog/bsc5p.html>.
- [4] European Space Agency. The Hipparcos and Tycho Catalogues. Noordwijk : ESA Publications Division, 1997. (ESA Special Publication ; 1200). ISBN 92-9092-399-7. URL: <https://www.cosmos.esa.int/web/hipparcos/catalogues>.
- [5] Kostjuk N. D. HD–DM–GC–HR–HIP–Bayer–Flamsteed Cross Index / Centre de Données astronomiques de Strasbourg. 2002. URL: <https://cdsarc.cds.unistra.fr/viz-bin/cat/IV/27A>.
- [6] Davenhall A. C., Leggett S. K. Catalogue of Constellation Boundary Data / Royal Observatory Edinburgh; Centre de Données astronomiques de Strasbourg. 1989. URL: <https://cdsarc.cds.unistra.fr/viz-bin/ReadMe/VI/49>.
- [7] Клыков Д. Ю. Преобразование эффективной температуры звёзд Teff в цвет RGB : Репозиторий GitHub. URL: https://github.com/dyuk108/star_colors.
- [8] The PASTEL Catalogue: 2016 Version / C. Soubiran [и др.] // Astronomy & Astrophysics. 2016. Т. 591. DOI: 10.1051/0004-6361/201628497.
- [9] Клыков Д. Ю. Скрипт для создания векторных карт звёздного неба : Репозиторий GitHub. URL: https://github.com/dyuk108/draw_skymap.

Астрономическое образование, 2026. Материал предоставлен на условиях лицензии Creative Commons Attribution-ShareAlike 4.0 (С указанием авторства — С сохранением условий). Разрешено свободно делиться (обмениваться) — копировать и распространять материал на любом носителе и в любом формате в любых целях, включая коммерческие; адаптировать (создавать производные материалы) — делать ремиксы, видоизменять и создавать новое, опираясь на этот материал в любых целях, включая коммерческие; при условии обеспечения соответствующего указания авторства, предоставления ссылки на лицензию и обозначения изменений, если таковые были сделаны. Производные материалы должны распространяться на таких же условиях.