

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CURSO DE CIÊNCIA DA COMPUTAÇÃO

CATÁLOGO GAIA DE ESTRELAS ATÉ 23 PARSECS DO SOL
Relatório para a Disciplina de Banco Dados I

Helena Serrano Cardoso da Costa
Laura Serrano Cardoso da Costa

Dezembro de 2023

Introdução

Nosso trabalho tem como objetivo construir um banco de dados com informações, majoritariamente provenientes do Catálogo Gaia, de estrelas próximas do Sol, ou seja, distantes no máximo 23 parsecs do Sol (23 parsecs equivalem a 3.26156 anos-luz ou 7.097×10^{14} quilômetros). Além disso, também foram utilizadas informações do catálogo Hipparcos.

Hipparcos é um satélite que foi lançado pela Agência Espacial Europeia (ESA) em 18 de agosto de 1989. A missão se deu como concluída em 1993.

O gigantesco survey Gaia, utilizado neste trabalho, é um dos mais impactantes na Astronomia nos últimos anos. O telescópio Gaia foi lançado, também pela ESA, em 19 de dezembro de 2013. O primeiro catálogo proveniente da Missão Gaia com mais de um bilhão de estrelas foi publicado em 14 de setembro de 2016. Esta foi a maior pesquisa de objetos celestes até a data.



Fonte: [hipparcos](#)



Fonte: [gaia](#)

No banco de dados criado neste trabalho, foram incluídos todos os registros que puderam ser obtidos do Gaia Archive (<https://gea.esac.esa.int/archive/>), somente limitado pela capacidade de seleção da ferramenta (de 2000 registros). Trabalhamos com três arquivos de dados: um proveniente do catálogo Gaia, do tipo csv, com 1774 estrelas; um proveniente do catálogo Hipparcos, do tipo dat, com 118218 estrelas; e outro proveniente do *The Extrasolar Planet Encyclopedia* (<http://www.exoplanet.eu/>), do tipo csv, contendo 5553 exoplanetas (planetas fora do Sistema Solar).

A partir das informações contidas nestes arquivos, foi possível calcular algumas propriedades de cada estrela, como magnitude e índice de cor, e plotar essas propriedades em diagramas HR, que também foram armazenados no banco de dados.

Também relacionamos os dados do Gaia com os do catálogo de exoplanetas. Ou seja, para cada estrela, proveniente do Gaia, obtivemos as informações, como massa e semieixo maior, dos planetas que orbitam ao seu redor.

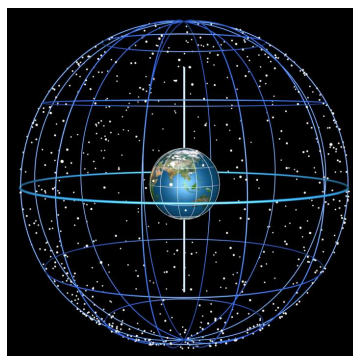
Os dados do Gaia superam, em precisão, os dados do catálogo Hipparcos, e somente algumas das 1774 estrelas contidas no Gaia também estão no Hipparcos. Dessa forma, nosso trabalho se propõe a fazer um cruzamento de dados: algumas informações mais tradicionais, que são encontradas somente no Hipparcos, foram acrescentadas às estrelas Gaia que estão, simultaneamente, nos dois catálogos. Logo, não é de interesse, neste trabalho, manter na base de dados as estrelas que estão somente no catálogo Hipparcos, mas acrescentar às estrelas provenientes do Gaia informações valiosas presentes apenas no Hipparcos.

O interesse na obtenção destas propriedades e diagramas HR de estrelas próximas do Sol é astrobiológico: dentre as estrelas selecionadas, aquelas que tiverem potencial de ter uma zona habitável (isto é, uma região em que um planeta orbitando ao seu redor possa abrigar vida) serão estudadas em mais detalhes em uma pesquisa que utilizará esta base de dados.

Conceitos de Astronomia usados no Tratamento do Dataset

Esfera celeste

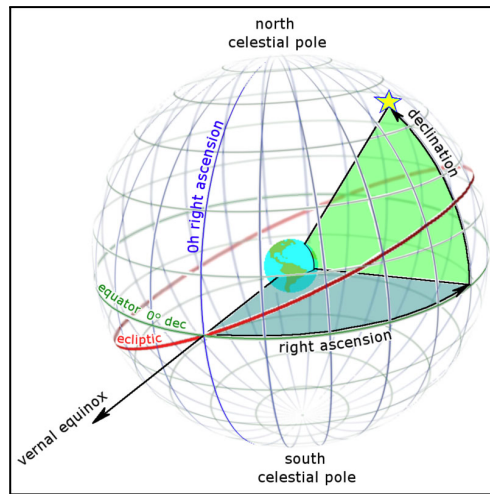
É uma esfera imaginária em volta da Terra, sem raio definido, que representa a projeção das estrelas como são vistas do nosso planeta. Os astrônomos mais antigos acreditavam que as estrelas estavam a uma mesma distância da Terra, como se estivessem “incrustadas” nesta esfera.



Fonte: [celestial_sphere](https://www.celestial-sphere.com/)

Ascensão reta

É o ângulo medido, sobre o equador celeste, do ponto vernal ao meridiano do astro. É um conceito análogo ao de latitude geográfica.



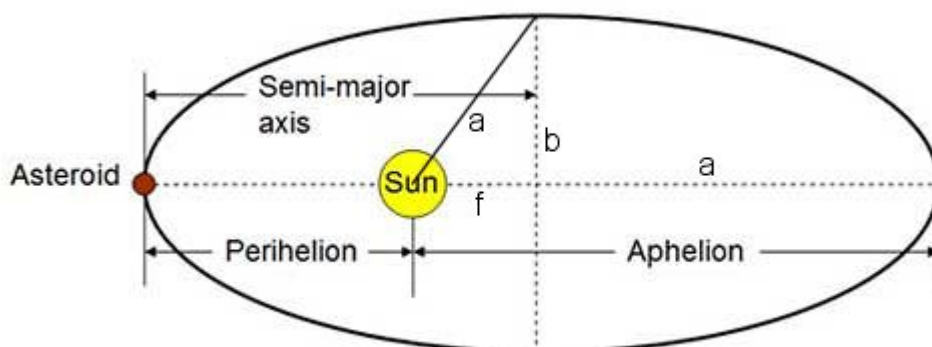
Fonte: [right ascension and declination](#)

Declinação

É a medida do ângulo, sobre o meridiano do astro, que vai do equador celeste ao astro. É um conceito análogo ao de longitude geográfica (ver figura acima).

Semi-eixo maior

Semi-eixo maior da elipse descrita pela órbita do planeta ao redor da estrela. A estrela orbitada é um dos focos da elipse.



Fonte: [semi-major-axis](#)

Datasets

Os arquivos de dados com os quais o banco de dados foi populado foram obtidos do Gaia Archive, de um arquivo pessoal ([HIP_MAIN.DAT](#)) e do *The Extrasolar Planet Encyclopedia*.

Colunas contidas no csv extraído do Gaia Archive:

record_ordinal_number (tipo *int*):

Número ordinal do registro no catálogo.

ra (tipo *numeric(65, 30)*):

Ascensão reta. É a primeira coordenada da localização da estrela.

dec (tipo *numeric(65, 30)*):

Declinação. É a segunda coordenada da localização da estrela.

designation (tipo *string*):

Designação única para cada estrela no Gaia.

ruwe (tipo *numeric(65, 30)*)

distance_gspphot (tipo *numeric(65, 30)*):

distância medida em parsecs do Sol.

distance_gspphot_lower (tipo *numeric(65, 30)*):

distância medida em parsecs do Sol, com margem de erro inferior.

distance_gspphot_upper (tipo *numeric(65, 30)*):

distância medida em parsecs do Sol, com margem de erro superior.

teff_gspphot (tipo *numeric(65, 30)*):

teff_gspphot_lower (tipo *numeric(65, 30)*):

teff_gspphot_upper (tipo *numeric(65, 30)*):

logg_gspphot (tipo *numeric(65, 30)*):

logg_gspphot_lower (tipo *numeric(65, 30)*):

logg_gspphot_upper (tipo *numeric(65, 30)*):

mh_gspphot (tipo *numeric(65, 30)*):

mh_gspphot_lower (tipo *numeric(65, 30)*):

mh_gspphot_upper (tipo *numeric(65, 30)*):

parallax (tipo *numeric(65, 30)*):

parallax_error (tipo *numeric(65, 30)*):

phot_g_mean_mag (tipo *numeric(65, 30)*):

phot_rp_mean_mag (tipo *numeric(65, 30)*):

phot_bp_mean_mag (tipo *numeric(65, 30)*):

Colunas contidas no csv HIPPARCOS_MAIN.DAT

HIP (tipo *int*):

Identificação única da estrela no Hipparcos.

Vmag (tipo *numeric(65, 30)*):

RAdeg (tipo *numeric(65, 30)*):

DEdeg (tipo *numeric(65, 30)*):

Plx (tipo *numeric(65, 30)*):

pmRA (tipo *numeric(65, 30)*):

pmDE (tipo *numeric(65, 30)*):

e_Plx (tipo *numeric(65, 30)*):

BTmag (tipo *numeric(65, 30)*):

e_BTmag (tipo *numeric(65, 30)*):

VTmag (tipo *numeric(65, 30)*):

e_VTmag (tipo *numeric(65, 30)*):

B_V (tipo *numeric(65, 30)*):

e_B_V (tipo *numeric(65, 30)*):

Colunas contidas no arquivo extraído do *The Extrasolar Planet Encyclopedia*:

name (tipo *string*):
Nome do planeta.

mass (tipo *numeric(65, 30)*):
Massa do planeta.

semi_major_axis (tipo *numeric(65, 30)*):

orbital_period (tipo *numeric(65, 30)*):

Tratamento dos Datasets

Gaia e Hipparcos

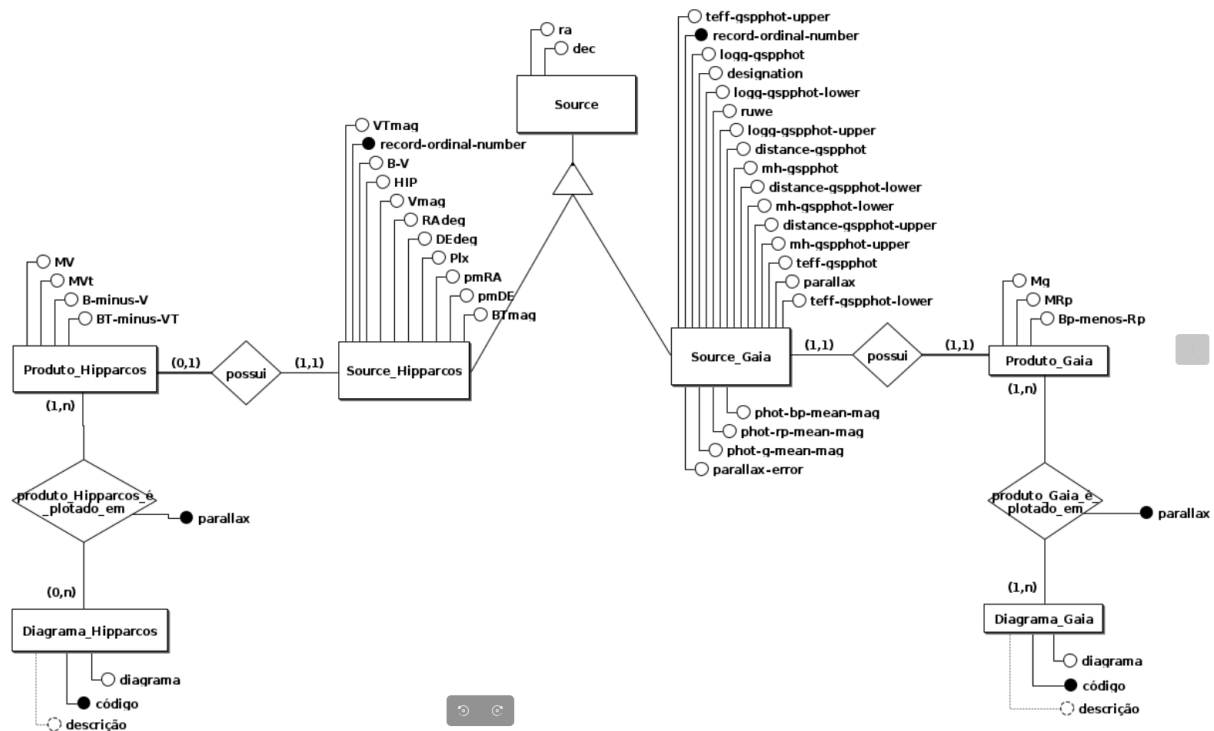
Primeiramente, fizemos o download dos arquivos nos endereços especificados anteriormente. Após isso, criamos um banco de dados auxiliar e, dentro deste banco, duas tabelas, Gaia e Hipparcos. Cada arquivo baixado foi carregado em uma destas tabelas. Foi usada uma consulta SQL *left outer join* para poder adicionar às estrelas Gaia, presentes nos dois catálogos, informações extras que estão presentes apenas no Hipparcos. O resultado desta consulta foi gravado em um arquivo no formato csv e utilizado para popular a tabela Star do banco de dados Catalogo_Gaia.

Planetas Extrasolares

Baixamos um arquivo csv do *The Extrasolar Planet Encyclopedia* e carregamos estes dados em uma tabela auxiliar. Depois, fizemos uma *left outer join* com a tabela Hipparcos, para adicionar características das estrelas aos registros de planetas. Após isto, foi feita outra *left outer join*, desta vez entre o resultado da consulta anterior e a tabela Gaia. O resultado destas duas consultas (feita na verdade através de uma única consulta com subconsulta aninhada) foi o de relacionar os registros dos planetas aos registros das estrelas no Gaia. O resultado deste processo foi gravado em um arquivo no formato csv e utilizado para popular a tabela Planet do banco de dados Catalogo_Gaia.

Projeto do Banco de Dados

Modelo Conceitual



Modelo Lógico

Star (record_ordinal_number, ra , dec, designation, ruwe, distance_gspphot, distance_gspphot_lower, distance_gspphot_upper, teff_gspphot, teff_gspphot_lower, teff_gspphot_upper, logg_gspphot, logg_gspphot_lower, logg_gspphot_upper, mh_gspphot, mh_gspphot_lower, mh_gspphot_upper, parallax, parallax_error, phot_g_mean_mag, phot_rp_mean_mag, phot_bp_mean_mag, HIP, Vmag, RAdeg, DEdeg, Plx, pmRA, pmDE, e_Plx, BTmag, e_BTmag, VTmag, e_VTmag, B_V, e_B_V)
PK(record_ordinal_number)

Property (record_ordinal_number, BT_minus_VT, B_minus_V, MV, MVt, Bp_minus_Rp, MRp, Mg, Mg_error, MRp_error, MVt_error, MV_error)
PK(record_ordinal_number)

FK(record_ordinal_number) references Star(record_ordinal_number)

Diagram(diagram, axis, parallax)

PK(parallax, axis)

Planet(name, mass, semi_major_axis, right_ascension, declination, discovered, orbital_period)
PK(name)

Generate(record_ordinal_number, parallax, axis)
PK(record_ordinal_number, parallax, axis)
FK(record_ordinal_number) references Property(record_ordinal_number)
FK(parallax, axis) references Diagram(parallax, axis)

Modelo Físico

```
create table Star (  
  record_ordinal_number int not null auto_increment primary key,  
  designation char(100),  
  ra double not null,  
  declination double not null,  
  parallax double not null,  
  parallax_error float not null,  
  pmra double not null,  
  pmdec double not null,  
  ruwe float not null,  
  phot_g_mean_mag float not null,  
  phot_bp_mean_mag float not null,  
  phot_rp_mean_mag float not null,  
  teff_gspphot float not null,  
  teff_gspphot_lower float not null,  
  teff_gspphot_upper float not null,  
  logg_gspphot float not null,  
  logg_gspphot_lower float not null,  
  logg_gspphot_upper float not null,  
  mh_gspphot float not null,  
  mh_gspphot_lower float not null,  
  mh_gspphot_upper float not null,  
  distance_gspphot float not null,  
  distance_gspphot_lower float not null,  
  distance_gspphot_upper float not null,  
  HIP long not null,  
  Vmag double,  
  RAdeg double,  
  DEdeg double,  
  Plx double,  
  BTmag double,  
  VTmag double,
```

B_V double
);

```
create table Property (  
    record_ordinal_number int not null auto_increment primary key,  
    Mg double not null,  
    MRp double not null,  
    Bp_minus_Rp double not null,  
    MVt double not null,  
    MV double not null,  
    B_minusV double not null,  
    BT_minus_VT double not null,  
    Mg_error double not null,  
    MRp_error double not null,  
    MVt_error double not null,  
    MV_error double not null,  
    foreign key (record_ordinal_number) references  
    Star(record_ordinal_number) on delete cascade  
);
```

```
create table Diagram(  
    diagram blob not null,  
    axis char(10),  
    parallax double,  
    primary key (axis, parallax)  
);
```

```
create table Generate(  
    record_ordinal_number int not null auto_increment,  
    parallax double not null,  
    axis char(10) not null,  
  
    foreign key (record_ordinal_number) references Star(record_ordinal_number)  
on delete cascade,  
    foreign key (axis, parallax) references Diagram(axis, parallax) on delete  
cascade,  
    primary key (record_ordinal_number, parallax, axis)  
);
```

```
create table Planet(  
    name char(30),  
    record_ordinal_number int null,  
    mass double,  
    semi_major_axis double,  
    RAdeg double,  
    DEdeg double,  
    foreign key (record_ordinal_number) references Star(record_ordinal_number),  
    primary key(name)  
);
```

Aplicação

Link para o repositório no GitHub com os códigos do projeto:

Distribuição do Trabalho

Os entregáveis do trabalho, listados abaixo

1. Relatório do trabalho
2. Banco de dados do MySQL
3. Aplicação Web feita com Python, Flask e HTML
4. Slides para a apresentação

foram feitos colaborativamente entre as duas integrantes do grupo, sem que houvesse separação prévia de tarefas.

Considerações Finais

Consideramos que a principal meta deste trabalho foi atingida, que era a obtenção dos seguintes diagramas:

- $M(G)$ contra $Bp-Rp$;
- $M(Rp)$ contra $Bp-Rp$;
- $M(V)$ contra $B-V$;
- $M(Vt)$ contra $BT-VT$

Assim como o relacionamento dos registros dos dois catálogos, Gaia e Hipparcos.

Referências

<https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>

https://gea.esac.esa.int/archive/documentation/GDR3/Gaia_archive/chap_datamodel/sec_dm_main_source_catalogue/ssec_dm_gaia_source.html

https://gea.esac.esa.int/archive/documentation/GDR3/Gaia_archive/chap_datamodel/sec_dm_solar_system_object_tables/ssec_dm_sso_observation.html

<https://stackoverflow.com/questions/55347322/c-program-to-convert-a-text-file-into-a-csv-file>

<https://zetcode.com/db/mysqlc/>

https://www.youtube.com/watch?v=D59J_oAGXxQ

<https://www.ufmg.br/espacodoconhecimento/a-esfera-celeste-e-a-bandeira-do-brasil/>