

# Aplicaciones en astronomía: las Hyades

Informe

*Nicolás Abuchar*  
*Franco Betteo*  
*Milena Dotta*  
*Francisco Valentini*  
*Octubre de 2018*

## Índice

<b>1. Comprensión del dominio</b>	<b>1</b>
<b>2. Comprensión de los datos</b>	<b>2</b>
<b>3. Preparación de los datos</b>	<b>7</b>
<b>4. Modelado</b>	<b>8</b>
<b>5. Evaluación</b>	<b>14</b>
<b>6. Implementación</b>	<b>17</b>
<b>7. Anexo</b>	<b>17</b>

## 1. Comprensión del dominio

### 1.1. Información general del dominio

En el presente informe se presenta un análisis realizado en base a los datos recolectados por la misión espacial Hipparcos, los cuales se volcaron en dos catálogos con distinta calidad de información, Hipparcos y Tycho, y un listado de estrellas que los astrónomos asignaron a las Hyades (cluster abierto de estrellas).

Este último cluster es el objeto de interés central del presente estudio. Se trata de un cluster abierto cuyas estrellas se originaron a partir de la misma nube molecular. Estas estrellas tienden a tener características químicas y edades similares, lo cual hace posible agruparlas utilizando ciertas variables. Tiene unas 300 a 400 miembros, con un núcleo de 17,6 años luz de diámetro con estrellas cercanas entre sí, rodeada por otro grupo de estrellas más separadas entre sí.

### 1.2. Recursos

El análisis se llevó a cabo a partir de los datos disponibles en dos catálogos de estrellas, Hipparcos y Tycho, un listado de estrellas pertenecientes a las Hyades y un archivo que permite cruzar las estrellas identificadas como Hyades con estrellas registradas en el catálogo de Hipparcos (ver Sección 2).

Se usó hardware de uso personal con un procesador Intel Core i7 y 8 GB de memoria RAM. El análisis se llevó a cabo en lenguaje R en su versión 3.5.0, con la asistencia de la IDE RStudio y del conjunto de librerías *tidyverse*. Asimismo, para las tareas de clustering y su respectivo diagnóstico se usaron las librerías *cluster*, *factoextra*, *clustertend* y *dbscan*.

### 1.3. Objetivo de data mining

El objetivo principal del trabajo es encontrar dentro de los catálogos Hipparcos y Tycho estrellas que podrían pertenecer a las Hyades. Esta identificación de estrellas candidatas se realizó mediante técnicas de clustering: las estrellas pertenecientes al mismo agrupamiento que una cantidad significativa de las ya identificadas como Hyades son consideradas candidatas. Se realizaron asimismo las siguientes actividades adicionales relacionadas al objetivo general: un estudio de las características diferenciales de las estrellas candidatas de ambos catálogos, una identificación de estrellas candidatas en Hipparcos a partir de métodos de clustering difuso y por densidad, una exportación de un modelo de K-Medias con la especificación PMML desde R hacia Python y un análisis de datos faltantes de paralaje en Tycho.

### 1.4. Plan del proyecto

No corresponde.

## 2. Comprensión de los datos

### 2.1. Recolección inicial de los datos

Los datos analizados provienen de mediciones realizadas por la misión espacial Hipparcos de la Agencia Espacial Europea entre 1989 y 1993, los cuales se volcaron en dos catálogos con distinta calidad de información: Hipparcos y Tycho. Los datos efectivamente usados corresponden a un subconjunto de estas mediciones provistas en archivo Excel junto a la consigna del informe. La información de este archivo fue importada al entorno R por medio de la librería *readxl* de *tidyverse*.

### 2.2. Descripción de los datos

El catálogo Hipparcos cuenta con 2655 estrellas, sobre las cuales se midieron los atributos descritos en el Cuadro 1. Por su parte, Tycho tiene 16258 registros definidos por los atributos presentados en el Cuadro 2. 2402 de las estrellas de Hipparcos se encuentran identificadas en Tycho. Se sabe de la existencia de 50 Hyades, de las cuales 49 están presentes en Tycho y la totalidad en Hipparcos.

Cuadro 1: Descripción del catálogo Hipparcos

Nombre	Descripción	Tipo
HIP	Identificador	character
RA_J2000	Ascensión recta	double
DE_J2000	Declinación	double
Plx	Paralaje	double
pmRA	Movimiento propio en ascensión recta	double
pmDE	Movimiento propio en declinación	double
Vmag	Magnitud en banda V de Johnson	double
B-V	Color BV de Johnson	double

Cuadro 2: Descripción del catálogo Tycho

Nombre	Descripción	Tipo
recno	Identificador	character
TYCID1	ID de región según catálogo Guide Star	character
TYCID2	ID corriente dentro de la región	character
TYCID3	ID específico de Tycho	character
RA_J2000_24	Ascensión recta	double
DE_J2000	Declinación	double
pmRA	Movimiento propio en ascensión recta	double
pmDE	Movimiento propio en declinación	double
BT	$(B-V)/0.85 + VT$	double
VT	$V + 0.09*(BT-VT)$	double
V	Magnitud en banda V de Johnson	double
B-V	Color BV de Johnson	double
HD	Identificador de Hyades	character
HIP	Identificador de Hipparcos	character
Plx	Paralaje	double

## 2.3. Exploración de los datos

Se realizó un análisis exploratorio de los catálogos Hipparcos y Tycho por separado, primero a nivel agregado y luego diferenciando por pertenencia confirmada al cluster Hyades.

Para esta etapa fue necesario llevar adelante las tareas de preprocesamiento descritas en las Secciones 3.1 y 3.2.

### 2.3.1. Hipparcos

En la Figura 1 resulta interesante ver que las variables no presentan correlaciones muy fuertes, destacándose únicamente la relación negativa entre las dos métricas de movimiento propio. A su vez el paralaje correlaciona positivamente con  $pmRA$  y negativamente con  $pmDE$ .

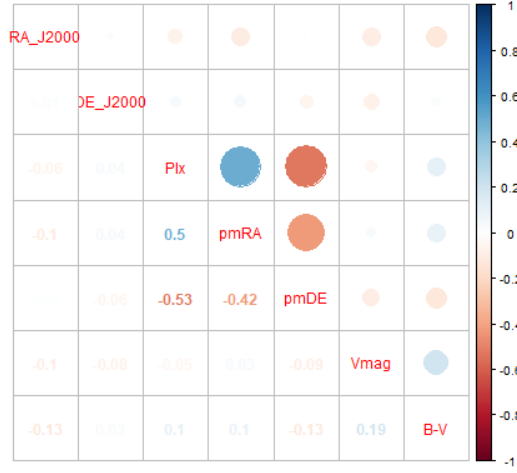


Figura 1: Correlograma (Hipparcos)

En los gráficos de densidad de la Figura 2 se puede observar que la mayoría de las medianas de las variables son distintas aunque se encuentran próximas entre un grupo y otro. La dispersión es menor por lo general en las Hyades. Esto último resulta razonable ya que se supone que son similares entre ellas.

Algunos casos destacables son la variable de temperatura  $B-V$ , que presenta una mediana similar entre grupos; sin embargo las Hyades tienen sus valores concentrados mucho menos dispersas respecto al resto de las estrellas. Por otra parte, la variable de magnitud  $Vmag$  presenta una dispersión dentro de todo comparable entre ambos grupos, pero con una mediana claramente distinta entre ellos, siendo menor la correspondiente a las Hyades.

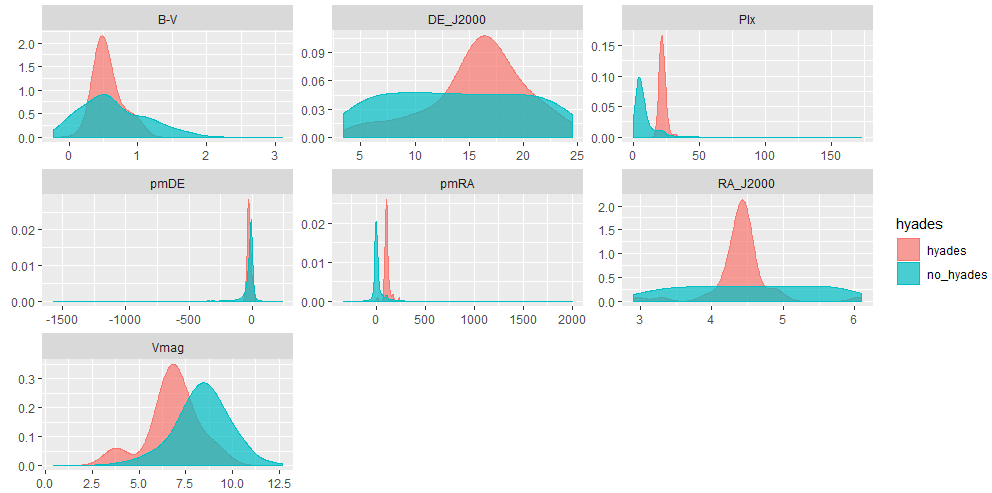


Figura 2: Densidades de las variables (Hipparcos)

### 2.3.2. Tycho

Al observar las correlaciones entre variables de Tycho en la Figura 3 se puede ver que las variables  $pmRA$  y  $pmDE$ , que en Hipparcos mostraban una correlación relevante, aquí no figuran con valores altos.

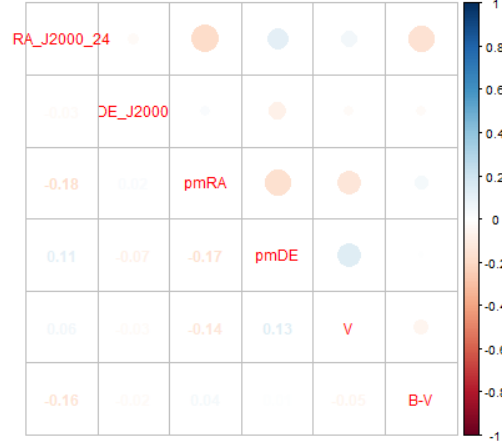


Figura 3: Correlograma (Tycho)

Por otra parte, como se observa en la Figura 4 se registran evidencias más fuertes de la diferencia entre las estrellas pertenecientes a las Hyades y el resto. Las medianas y las dispersiones presentan diferencias mayores, a la vez que se observa una gran cantidad de outliers univariados para las estrellas no pertenecientes a las Hyades.

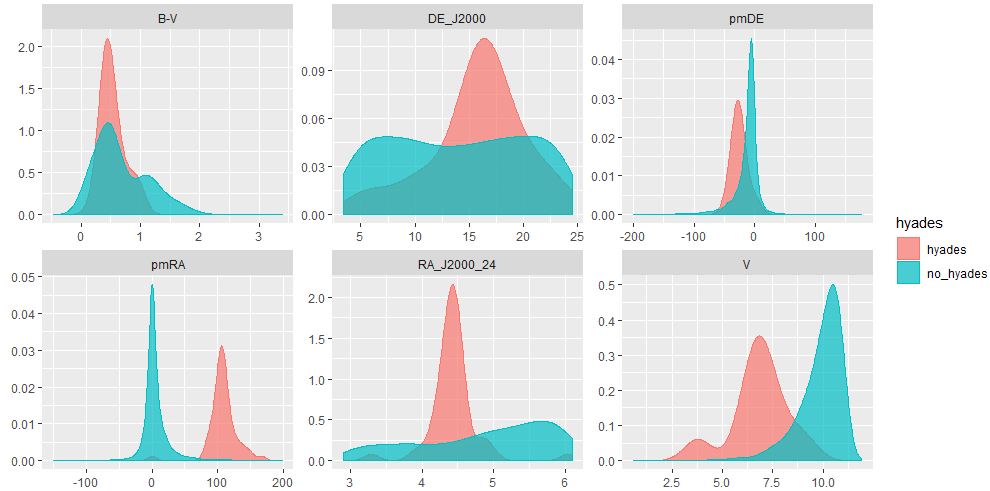


Figura 4: Densidades de las variables (Tycho)

## 2.4. Verificación de la calidad de los datos

El análisis de la calidad de los datos consistió en la detección y análisis de datos faltantes. En Hipparcos únicamente la variable  $B-V$  cuenta con registros faltantes, en particular en 15 de los 2655 casos. En Tycho el atributo  $Plx$  presenta 14005 instancias faltantes, lo cual representa el 86.14 % de los registros.

Se estudió si existe algún patrón en los datos que explique la presencia o ausencia de datos faltantes en el atributo  $Plx$  (paralaje) de Tycho.

En primer lugar se realizó un análisis gráfico comparando el comportamiento de los registros con faltantes en el paralaje y los que no. El análisis de las densidades comparadas (Figura 5) reveló que las estrellas con faltantes en el paralaje presentan valores más elevados en las variables  $BT$ ,  $V$  y  $VT$ , a la vez que se observa una menor dispersión en los valores de dichas variables para las estrellas que tuvieron faltantes.

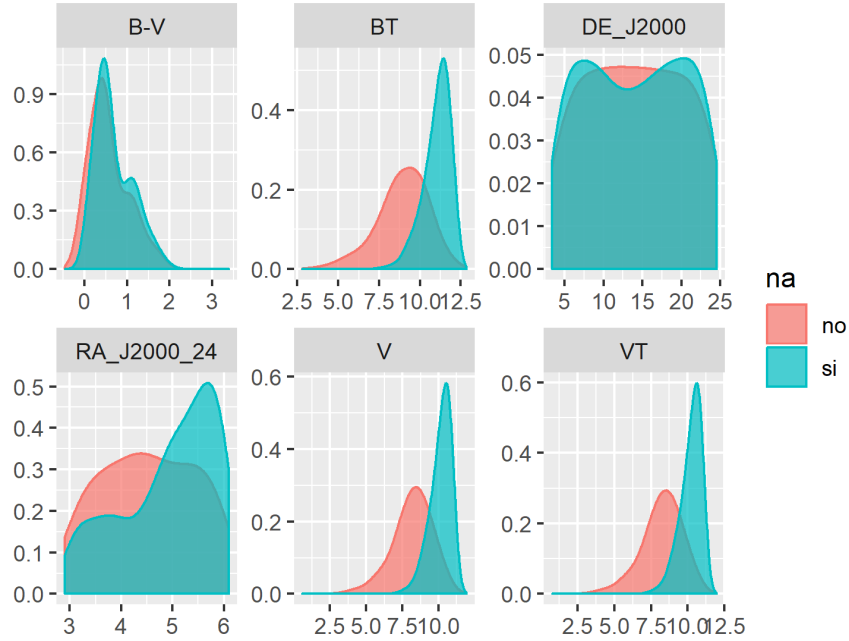


Figura 5: Densidades de las variables según presencia de datos de paralaje (Tycho)

Luego se realizó un test de diferencia de medias para ver si la diferencia de medias entre las estrellas con faltantes en el paralaje y las que no lo tienen resulta estadísticamente significativa. De esta análisis resultó que las medias de todas las variables son estadísticamente significativas exceptuando la variable  $DE\_J2000$ .

A continuación se realizó un test para evaluar si los faltantes son MCAR (Missing Completely At Random) utilizando el algoritmo *LittleMCAR* del paquete *BaylorEdPsych*. En estos casos, el mecanismo de faltantes no depende ni de la variable de interés ni de otra variable que es observada en el dataset. El resultado del test rechazó la hipótesis nula de que los datos faltantes son MCAR. Por lo tanto, deberían ser MAR (Missing at Random) o Nonignorable missing data.

Se realizó una prueba empírica para ver si los datos existentes logran predecir la presencia o no de faltantes. Se generó un modelo predictivo sencillo utilizando la implementación de Random Forest del paquete *randomForest*. Se entrenó al modelo con el 80 % del dataset y se lo testeó con el 20 % restante. Se obtuvo un Accuracy de 0.92 y un AUC del 0.78. Si bien se parte de una no-information rate de 0.86, los datos parecerían tener significatividad a la hora de determinar la presencia o no de faltantes en los datos de paralaje.

En cuanto a la imputación de datos, por tratarse de una variable con un muy elevado porcentaje de faltantes, se consideró que realizar una no sería adecuado. Sin embargo, se optó por probar realizar una imputación

mediante una regresión lineal. La misma se construyó usando las variables  $RA\_J2000$ ,  $DE\_J200$  y  $BT$  (realizando previamente un análisis de significatividad y correlación entre las variables). En la Figura 6 se puede observar la diferencia entre los valores predichos y los valores originales. La misma resulta bastante notoria por lo que se descarta la imputación usando regresión como una opción. Este resultado era esperable debido a la gran cantidad de faltantes que presenta la variable.

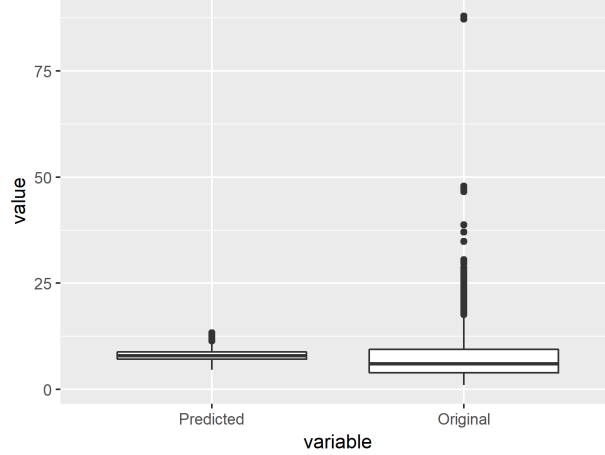


Figura 6: Imputación de faltantes de Plx (Tycho) con regresión lineal: valores ajustados vs. observados

Se concluye entonces que los datos podrían explicar la ausencia de datos faltantes en el paralaje, pero debido a la gran cantidad de NAs resulta inviable realizar algún tipo de imputación al valor del paralaje.

### 3. Preparación de los datos

#### 3.1. Selección de los datos

Para el análisis exploratorio y de clustering, en el caso de Hipparcos no se tuvo en cuenta la variable  $HIP$  por tratarse del atributo identificador. Por su parte, los atributos  $TYCID1$ ,  $TYCID2$ ,  $TYCID3$ ,  $HD$  y  $HIP$  de Tycho fueron descartados porque refieren a la variable identificadora o por ser fracciones del nombre de la estrella y no atributos útiles para clustering. Asimismo, no se consideraron las variables  $BT$  y  $VT$ , ya que constituyen combinaciones lineales de otros atributos, como se describe en el Cuadro 2. Cabe destacar que el campo identificador  $HIP$  fue usado posteriormente para realizar una identificación cruzada que permitió descartar de Tycho los candidatos ya identificados como resultado del clustering sobre Hipparcos.

#### 3.2. Limpieza de datos

La principal tarea de limpieza fue el tratamiento de los datos faltantes. En el caso de la variable  $B-V$  del catálogo Hipparcos se optó por reemplazar los datos faltantes por la mediana, debido a que solo representaban el 0.56 % de los casos.

En cambio, la variable Plx de Tycho fue omitida antes de realizar el análisis de clustering por la alta proporción de casos faltantes que presentó (86.14 %).

#### 3.3. Construcción de los datos

En ambos datasets se normalizaron las variables a partir de una transformación minmax. No se generaron nuevos registros ni atributos.

### 3.4. Integración de los datos

No corresponde ya que se trabajó con cada catálogo por separado. Sin embargo, cabe destacar que para evitar obtener candidatos de Tycho que ya hayan sido identificados en los agrupamientos de Hipparcos, se omitieron las estrellas correspondientes en Tycho usando el campo de identificación cruzada *HIP* disponible en el dataset.

### 3.5. Formateo de datos

No corresponde.

## 4. Modelado

Antes de llevar adelante los agrupamientos, se computó el índice de Hopkins de tendencia a la clusterización en ambos catálogos, fijando una cantidad de puntos en el espacio al azar equivalente al 10 % de cada dataset. El indicador arrojó un valor de 0.1048 en Hipparcos y 0.1083 en Tycho, lo cual indica tendencias significativas al agrupamiento, y entonces que intentar clusterizar las estrellas es razonable.

### 4.1. Selección de las técnicas de modelado

Las siguientes técnicas de clustering fueron implementadas para identificar candidatos: K-Medias (sobre ambos catálogos), clustering difuso (sobre Hipparcos) y DBSCAN (solo sobre Hipparcos—ver Sección 4.3.3). En el caso de Tycho se removieron del dataset las estrellas ya identificadas como candidatas en Hipparcos antes de implementar los algoritmos. En todas las ejecuciones se usó como medida de disimilitud la distancia euclidia.

### 4.2. Construcción de los modelos

#### 4.2.1. K-Medias

El algoritmo K-Medias se corrió usando su implementación `stats::kmeans()` disponible en R base. En cada ejecución del algoritmo se escogieron 20 asignaciones distintas de los centroides, seleccionándose aquella que minimizara la suma de cuadrados dentro de los grupos.

Para definir la cantidad de agrupamientos óptima en Hipparcos se usaron dos criterios: la maximización del Silhouette promedio ( $S$ ) y la búsqueda de un punto de quiebre en el scree-plot de la Suma de Cuadrados Dentro de los grupos ( $SCD$ ). Para tal motivo, se generaron agrupamientos de K-Medias para 10 valores posibles de  $K$ , entre 1 y 10. Los resultados se visualizan en la figura 7. Como se observa, la cantidad óptima de agrupamientos bajo estos criterios es 4.

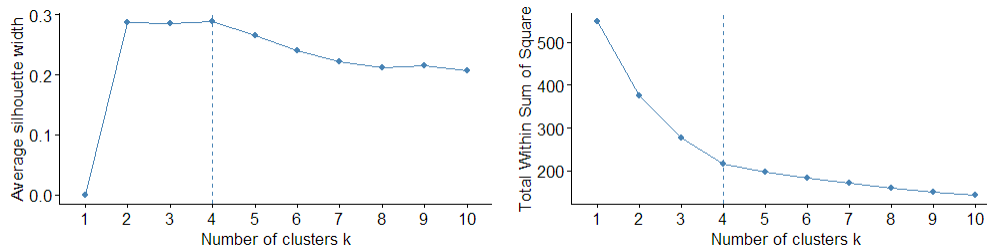


Figura 7: Selección de  $K$  para K-Medias (Hipparcos)



Una vez fijado el K en su valor óptimo, se obtuvo la representación del modelo en lenguaje PMML (ver Anexo). La misma fue importada y ejecutada en Python 2.7 usando el módulo *augustus*, obteniendo los mismos resultados que en R.

En el catálogo Tycho se implementó K-Medias y se usó en primera instancia el mismo criterio para determinar el número óptimo de clusters, que resultó ser igual a 3, como se presenta en la Figura 8.

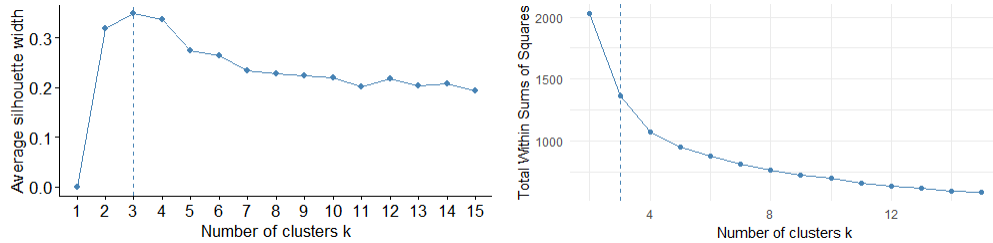


Figura 8: Selección de K para K-Medias (Tycho)

Sin embargo, un agrupamiento con  $K=3$  da lugar a la identificación de un número elevado de candidatas, debido a la gran cantidad de observaciones en Tycho. Por tal motivo, se probaron otros valores de K, y se eligió  $K=15$ , que da lugar a una cantidad razonable de estrellas candidatas (ver Sección 4.3.1).

#### 4.2.2. Clustering difuso

El algoritmo de clustering difuso se corrió usando la implementación *fanny()* de la librería *cluster* para R.

Para facilitar la comparación con los modelos de K-Medias, se seleccionó un K igual al identificado como óptimo para K-Medias en ambos catálogos. Por otra parte, se fijó el parámetro de *fuzziness* en 12.

#### 4.2.3. DBSCAN

Para ejecutar clustering por densidad se usó el algoritmo DBSCAN, en particular la implementación *dbscan()* de la librería *dbscan* para R.

El parámetro *minPts* (umbral de objetos cercanos para identificar objetos semilla) fue fijado en un valor de 5. Para determinar el parámetro *eps* (radio para identificar objetos cercanos) se ordenaron los puntos según su distancia al 5to vecino y se graficó dicha distancia (ver Figura 9).

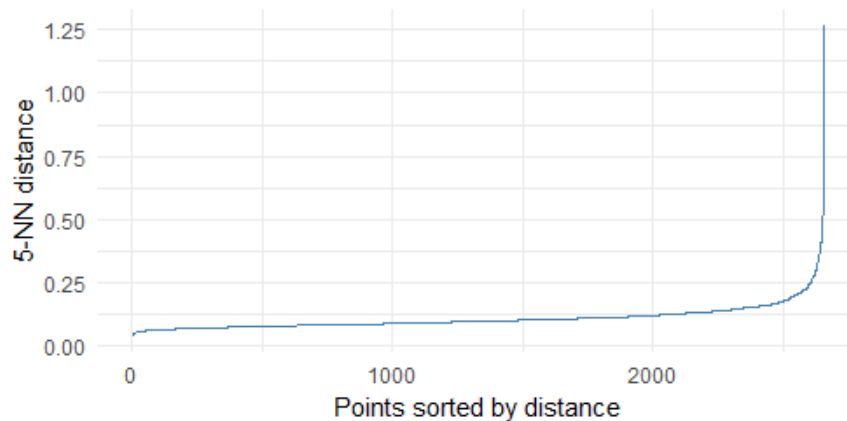


Figura 9: Selección de eps para DBSCAN (Hipparcos)

El objetivo es hallar un punto de quiebre a partir del cual las distancias incrementan significativamente la tasa de crecimiento. Este punto se encontró en una distancia aproximadamente igual a 0.2, de modo que se fijó a *eps* en este valor.

### 4.3. Evaluación de los modelos

#### 4.3.1. K-Medias

En el Cuadro 3 se observan la distribución de los clusters generados en función de la pertenencia al grupo de estrellas Hyades, para el catálogo Hipparcos. Bajo esta técnica fue posible identificar 622 estrellas candidatas en Hipparcos con un nivel de confianza moderado ya que 37 de las 50 Hyades se hallan agrupadas junto a ellas. Este es el principal criterio que define la bondad del modelo.

Cuadro 3: Distribución de estrellas Hyades en clusters de K-Medias (Hipparcos)

Cluster	Hyades	
	FALSE	TRUE
1	658	6
2	664	37
3	661	5
4	622	2

El modelo también fue evaluado mediante el cálculo de Silhouette y su gráfico asociado, que se observa en la Figura 10. Por último, en la Figura 11 se presenta un diagrama de dispersión de las estrellas sobre las dos primeras componentes principales.

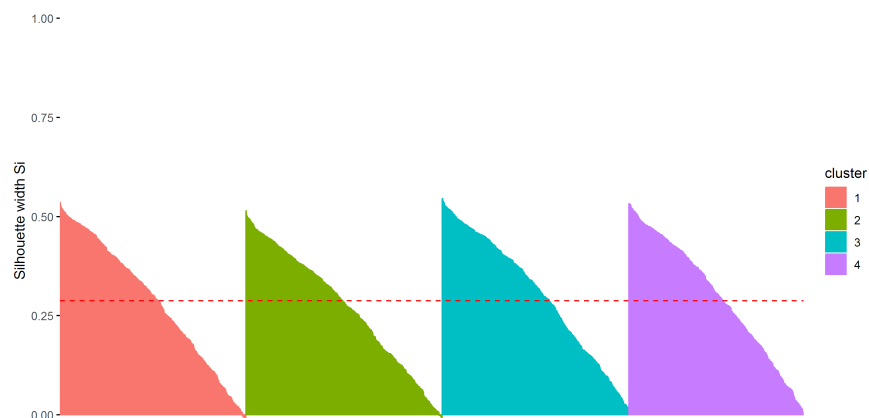


Figura 10: Gráfico de Silhouette K-Medias K=4 (Hipparcos)

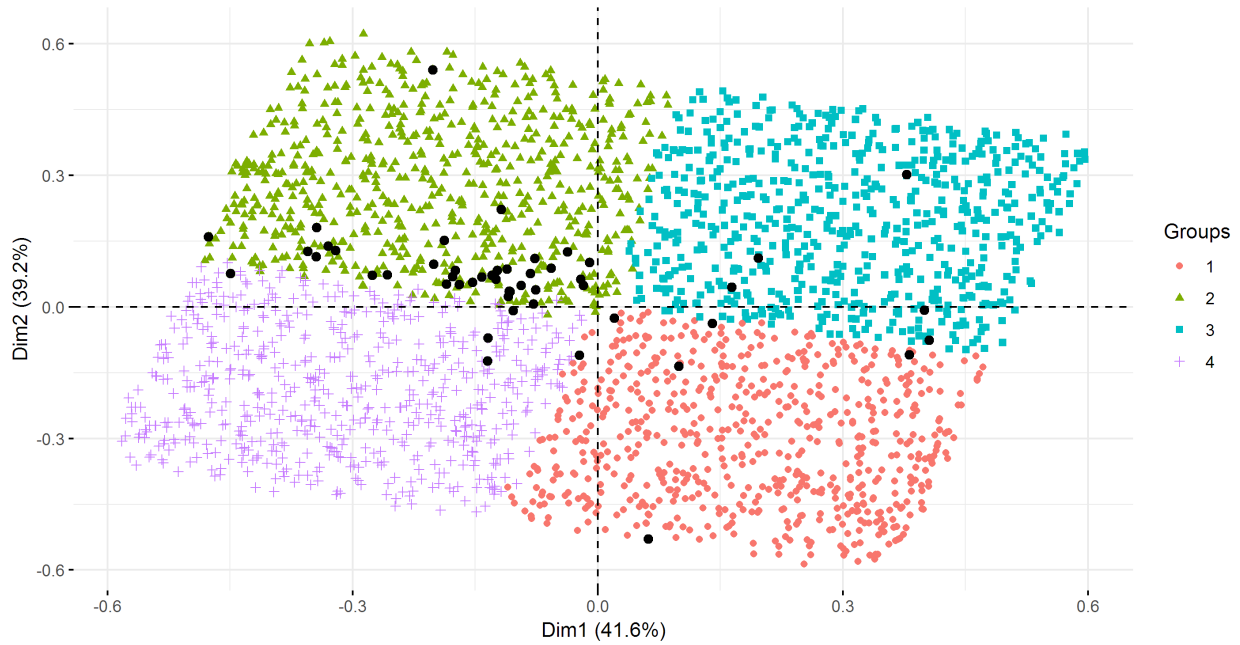


Figura 11: Dos dimensiones de PCA: K-Means K=4 (Hipparcos - datos transformados con minmax)

Vemos que los datos quedaron separados en cuatro cuadrantes bien definidos al graficar las dos primeras componentes principales. Estas explican de manera conjunta el 80 % de la variabilidad. Queda reflejado que el cluster 2 (superior izquierdo) contiene a la mayoría de las observaciones que forman parte de Hyades. La mayoría de estas se encuentran cerca y a lo largo de la frontera inferior a excepción de una que se encuentra en el extremo superior dando cuenta que la primera dimensión forma un primer corte importante entre Hyades y no Hyades, y que dentro de las primeras igualmente hay variabilidad. Por otro lado la proyección en la segunda componente está mucho más concentrada para las Hyades aunque con gran mayoría en valores positivos. Las estrellas Hyades que quedaron en algún otro cluster se encuentran en buena medida cerca de los límites del cluster 2 a excepción de 5 observaciones bastante alejadas en base a estas componentes graficadas.

En lo que respecta a Tycho, el Cuadro 4 indica la distribución de los agrupamientos en función de la pertenencia al cluster Hyades. Fue posible identificar 257 estrellas candidatas en Tycho, que se encontraron agrupadas junto a 35 de las Hyades ya conocidas. Sumando las candidatas de ambos catálogos, se cuenta con 879 estrellas candidatas. Esta lista se presenta en el Anexo.

Cuadro 4: Distribución de estrellas Hyades en clusters de K-Medias (Tycho)

Cluster	Hyades	
	FALSE	TRUE
1	735	0
2	1194	9
3	1453	0
4	1057	0
5	1312	0
6	1116	0
7	1382	0
8	804	3
9	850	0
10	825	0
11	1294	0
12	1177	1
13	257	35
14	792	0
15	1354	1

#### 4.3.2. Clustering difuso

Luego de realizar los agrupamientos difusos en Hipparcos, se registró la máxima probabilidad de pertenencia a cada grupo—el histograma correspondiente se presenta en la Figura 12. Se observa que un alta concentración de valores entre por encima de 0.8, lo que indica que las asignaciones son razonablemente unívocas.

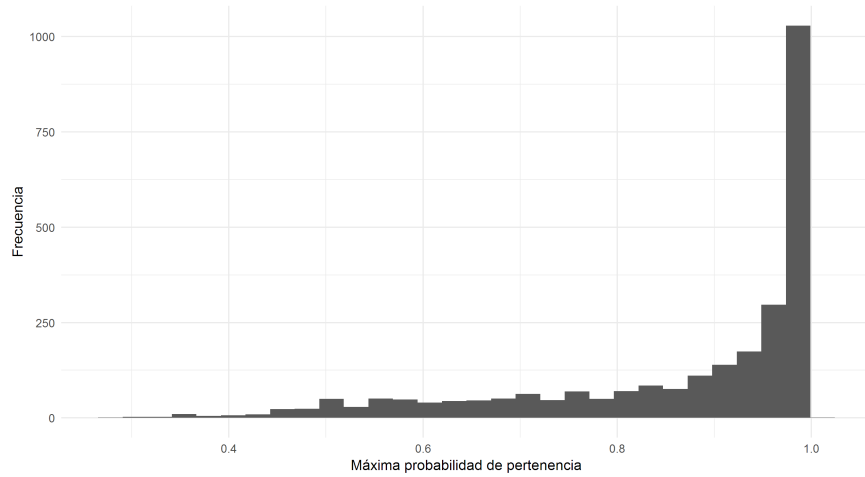


Figura 12: Histograma: máxima probabilidad de pertenencia- Fuzziness=1.2 (Hipparcos)

Luego de la clusterización, se definió que aquellas estrellas con una probabilidad máxima menor a 0.6 no fueran agrupadas en ningún cluster, de modo que son incluidas como cluster 0. El Cuadro 5 indica la distribución de los clusters en función de la pertenencia a Hyades. Se observa que una parte no despreciable las 50 estrellas Hyades no puede ser asignada unívocamente a ningún cluster bajo este criterio.

Cuadro 5: Distribución de estrellas Hyades en clusters difusos (Hipparcos)

Cluster	Hyades	
	FALSE	TRUE
0	258	13
1	600	4
2	581	28
3	597	3
4	569	2

Los clusters obtenidos para Hipparcos según las dos primeras componentes principales del dataset se visualizan en la Figura 13.

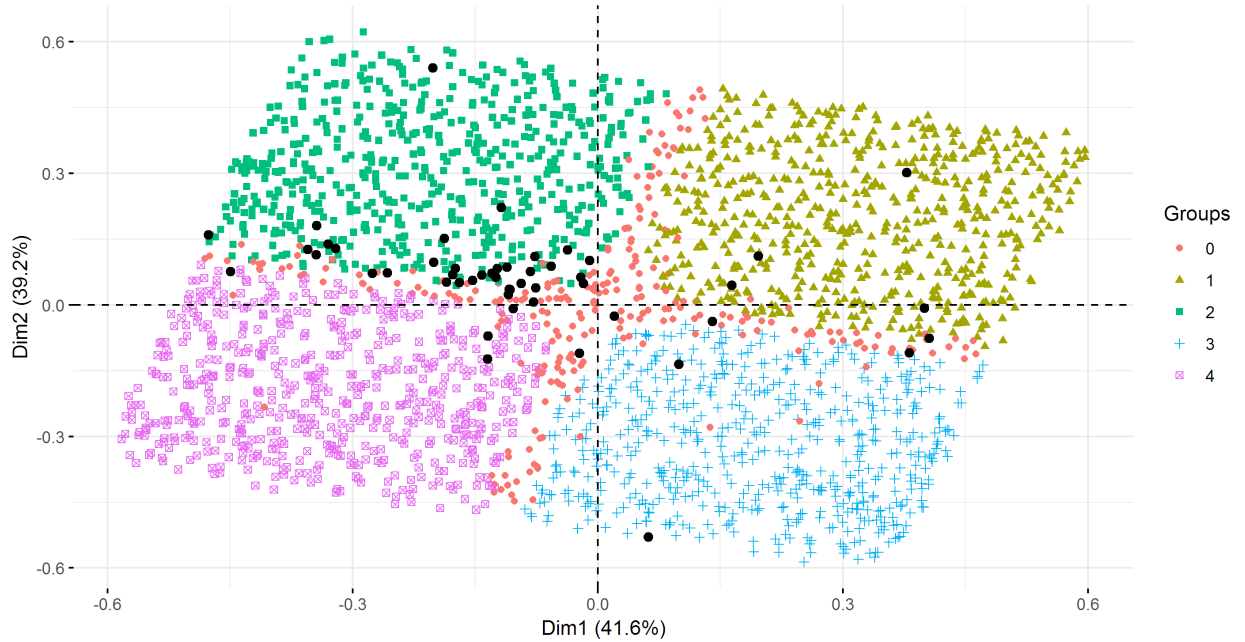


Figura 13: Dos dimensiones de PCA: Clustering difuso K=4 (Hipparcos)

Como se puede observar, los clusters definidos quedaron muy similares a los vistos en K-Means. Sin embargo la región central corresponde al cluster 0 donde el grado de certeza de la pertenencia a un cluster es reducido. Las conclusiones son básicamente las mismas que para la Figura 11 aunque notando que algunas de las observaciones Hyades que antes caían en el cluster 2 ahora no pueden ser clasificadas concluyentemente (cluster 0). De la misma manera podemos ver que observaciones que antes pertenecían a otros clusters también caen en el cluster 0, notando que tampoco es concluyente que pertenezcan todas a otro cluster realmente.

Se intentó ejecutar un clustering difuso sobre Tycho descontando las candidatas del cluster con mayor cantidad de Hyades según el Cuadro 5, pero no fue posible para  $K > 2$  debido a las limitaciones del hardware.

### 4.3.3. DBSCAN

El Cuadro 6 indica la distribución de los clusters de DBSCAN en función de la pertenencia a Hyades. El cluster 0 corresponde a las estrellas clasificadas como ruido.

Cuadro 6: Distribución de estrellas Hyades en clusters de DBSCAN (Hipparcos)

Cluster	Hyades	
	FALSE	TRUE
0	59	0
1	2546	50

Se observa que se obtuvo un solo gran cluster que incluye a las Hyades, mientras que el 2.22 % de las estrellas son clasificadas como ruido. Las características del dataset, puestas de relieve en los gráficos de PCA, implican que no es posible hallar agrupamientos relevantes por densidad, de modo que se descarta este método para hallar estrellas candidatas.

## 5. Evaluación

### 5.1. Evaluación de resultados

Se usan las candidatas resultantes de K-Means. En la Figura 14 se visualiza la posición de las candidatas en términos de  $RA$  y  $DE$ .

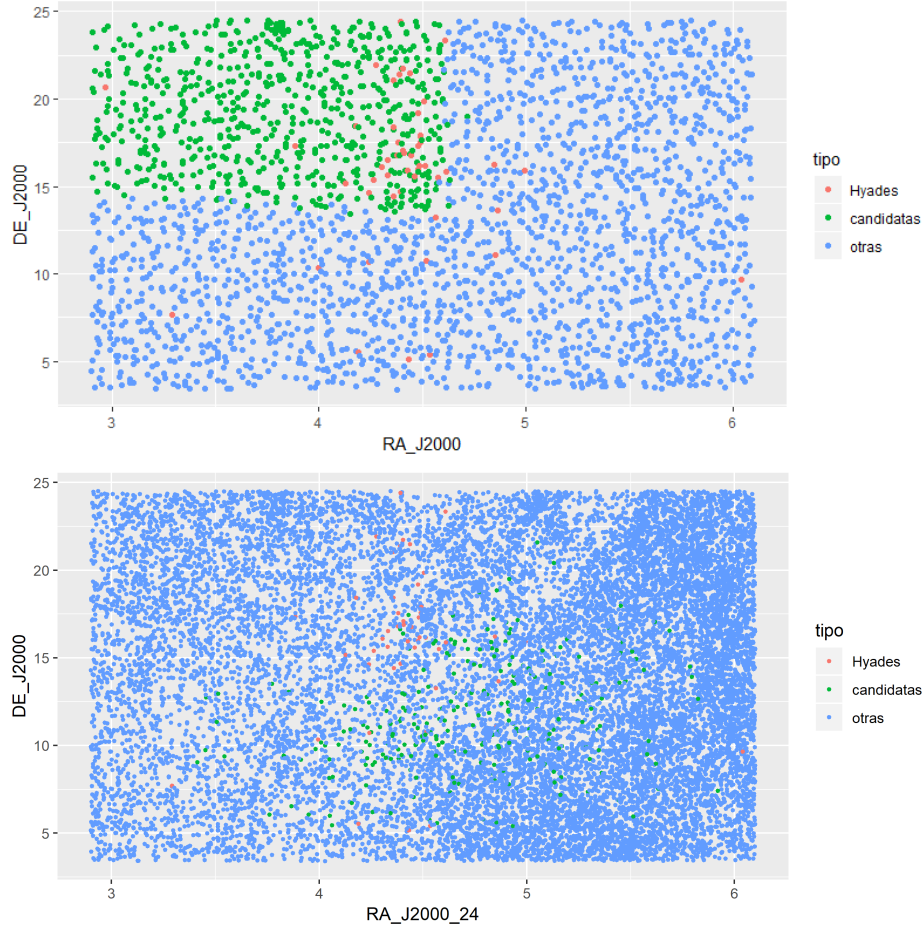


Figura 14: Candidatas de K-Means (Hipparcos -arriba- y Tycho-abajo-)

Se observa que en Hipparcos las candidatas quedan muy bien delimitadas desde el punto de vista de las variables de posición. La selección incluye el núcleo duro de Hyades gráficamente y se extiende abarcando a dos de las alejadas. Al menos para estas dos variables, la cantidad de candidatas identificada parece excesiva – es decir que la región de candidatas incluye zonas donde no hay Hyades confirmadas cerca. Por otra parte, la región de corte queda próxima a ciertas Hyades que finalmente fueron agrupadas en otro cluster.

En Tycho, por su parte, no se distingue una región totalmente aislada en términos de las variables de posición; más bien, las candidatas se encuentran entremezcladas con otras estrellas no candidatas. El algoritmo pareciera captar ciertas candidatas cercanas a la región de corte de Hipparcos, pero que formaban parte de otros clusters. A su vez, algunas candidatas se encuentran en la zona inferior derecha mostrando una clara diferencia en las características de posición del cluster obtenido para el catálogo Tycho respecto al de Hipparcos.

En la Figura 15 se comparan las candidatas que surgen de cada catálogo de forma univariada, para los atributos disponibles en ambos datasets.

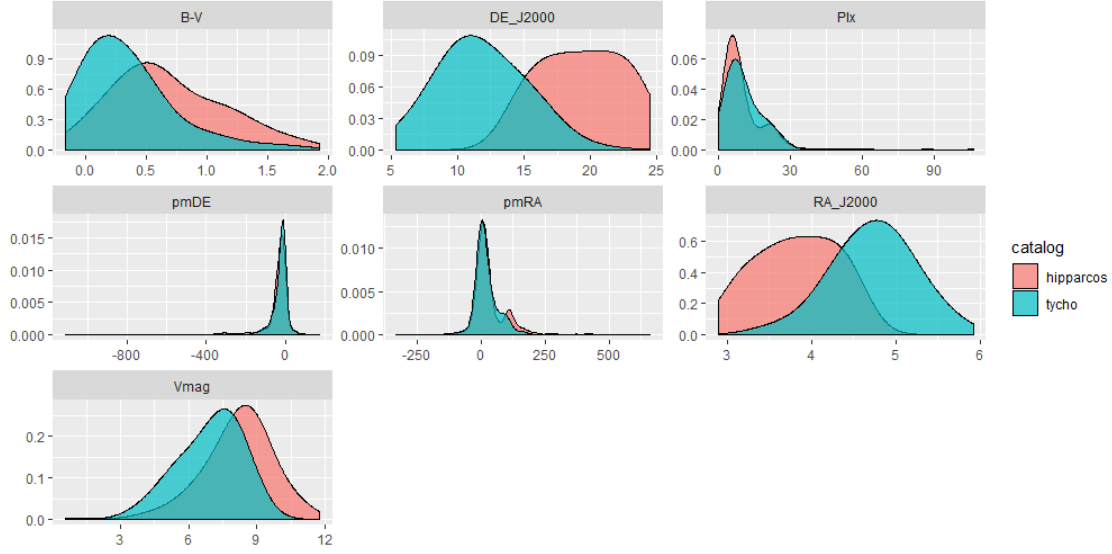


Figura 15: Comparación de candidatas

Se registran algunas diferencias claras: en primer lugar, las candidatas de los catálogos difieren significativamente en términos de posición de las estrellas, estando concentradas en distintos valores tanto  $DE$  como  $RA$  (esto mismo se observa en la Figura 14). En segundo lugar, se destaca la diferencia en la densidad para  $B-V$ , teniendo las candidatas de Tycho una concentración mayor en valores más bajos que Hipparcos, pero presentando una cola más pesada a derecha. Por último, la magnitud  $Vmag$  presenta una distribución de forma similar pero aparentemente centrada en valores levemente mayores en Hipparcos. El resto de las variables compartidas no presenta mayores diferencias.

La posición de las candidatas en términos de  $RA$  y  $DE$  se observa en la Figura 16.

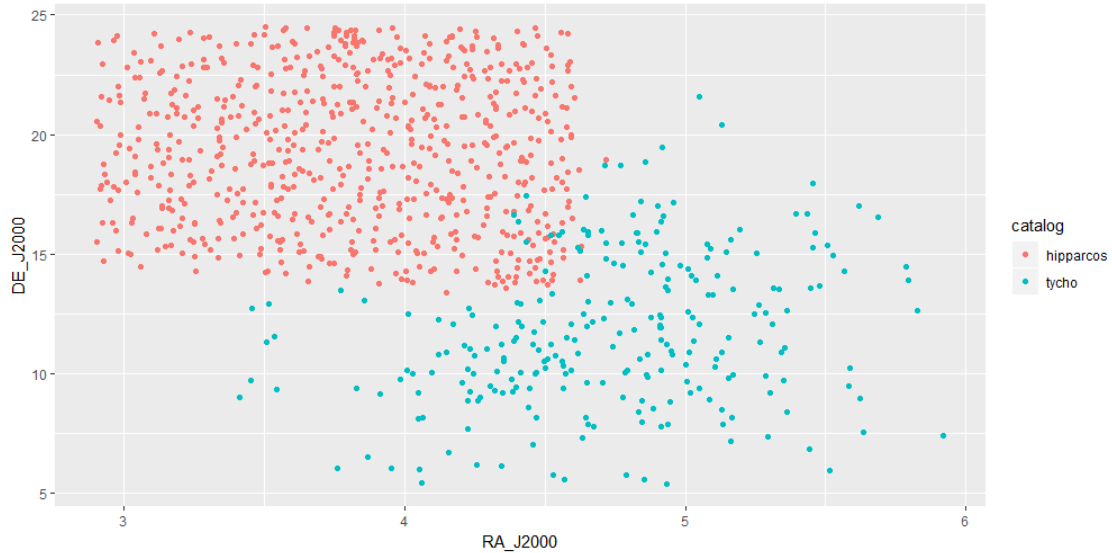


Figura 16: Posición de candidatas

Este gráfico refleja más claramente lo observado en las Figuras 14 y 15 al quitar los puntos de las estrellas que no son candidatas en ningún caso. Los catálogos presentan distintas agregaciones en términos de posición,



Hipparcos captando todo el cuadrante superior izquierdo mientras que en Tycho (descartadas las candidatas de Hipparcos) se incluyen como candidatas muchas de la región central y otros cuadrantes.

## 5.2. Proceso de revisión

No corresponde.

## 5.3. Futuras etapas

En Hipparcos la posición de las estrellas determinó un área muy delimitada mientras que en Tycho esto no sucedió, intercalando candidatas en áreas muy saturadas con no candidatas. Un elemento que quizás haya tenido impacto es que a la hora de realizar clusters para Tycho se descartaron todas las candidatas que surgieron de Hipparcos, por ende sacándole el peso que tenía esa región. Podrían a futuro realizarse los clusters en orden inverso para comparar resultados en términos de candidatas y con énfasis en este fenómeno. Sería necesario además observar la configuración de los clusters generados en términos de otras variables, además de las de posición. Por último, se podrían implementar técnicas para determinar la relevancia de cada variable en los agrupamientos encontrados.

## 6. Implementación

No corresponde.

## 7. Anexo

Especificación PMML del modelo K-Medias óptimo para Hipparcos:

```
## [1] "<?xml version='1.0' encoding='UTF-8'?"
## [2] "<!--(Comment generated by ADAPA) PMML processed by ADAPA (Version : 4.1)-->"
## [3] "<PMML version='4.1' xmlns='http://www.dmg.org/PMML-4.1'>"
## [4] "<Header copyright='Copyright (c) 2018 Franco' description='KMeans cluster model'>"
## [5] "  <Extension name='user' value='Franco' extender='Rattle/PMML'>"
## [6] "  <Application name='Rattle/PMML' version='1.4'>"
## [7] "  <Timestamp>2018-10-20 15:12:16</Timestamp>"
## [8] "</Header>"
## [9] "<DataDictionary numberOfFields='7'>"
## [10] "  <DataField name='RA_J2000' optype='continuous' dataType='double'>"
## [11] "  <DataField name='DE_J2000' optype='continuous' dataType='double'>"
## [12] "  <DataField name='Plx' optype='continuous' dataType='double'>"
## [13] "  <DataField name='pmRA' optype='continuous' dataType='double'>"
## [14] "  <DataField name='pmDE' optype='continuous' dataType='double'>"
## [15] "  <DataField name='Vmag' optype='continuous' dataType='double'>"
## [16] "  <DataField name='B-V' optype='continuous' dataType='double'>"
## [17] "</DataDictionary>"
## [18] "<ClusteringModel modelName='KMeans_Model' functionName='clustering' algorithmName='KMeans: Hartigan and Wong' modelClass='centerBased' numberOfClusters='4'>"
## [19] "  <MiningSchema>"
## [20] "    <MiningField name='RA_J2000' invalidValueTreatment='returnInvalid'>"
## [21] "    <MiningField name='DE_J2000' invalidValueTreatment='returnInvalid'>"
## [22] "    <MiningField name='Plx' invalidValueTreatment='returnInvalid'>"
## [23] "    <MiningField name='pmRA' invalidValueTreatment='returnInvalid'>"
## [24] "    <MiningField name='pmDE' invalidValueTreatment='returnInvalid'>"
## [25] "    <MiningField name='Vmag' invalidValueTreatment='returnInvalid'>"
## [26] "    <MiningField name='B-V' invalidValueTreatment='returnInvalid'>"
## [27] "  </MiningSchema>"
## [28] "  <Output>"
## [29] "    <OutputField name='predictedValue' feature='predictedValue' dataType='string' optype='categorical'>"
## [30] "  </Output>"
## [31] "  <ComparisonMeasure kind='distance'>"
## [32] "    <squaredEuclidean/>"
## [33] "  </ComparisonMeasure>"
## [34] "  <ClusteringField field='RA_J2000' compareFunction='absDiff'>"
## [35] "  <ClusteringField field='DE_J2000' compareFunction='absDiff'>"
## [36] "  <ClusteringField field='Plx' compareFunction='absDiff'>"
## [37] "  <ClusteringField field='pmRA' compareFunction='absDiff'>"
## [38] "  <ClusteringField field='pmDE' compareFunction='absDiff'>"
## [39] "  <ClusteringField field='Vmag' compareFunction='absDiff'>"
## [40] "  <ClusteringField field='B-V' compareFunction='absDiff'>"
## [41] "  <Cluster name='1' size='664' id='1'>"
## [42] "    <Array n='7' type='real'>0.755139189555844 0.241935538818172 0.0465383676529563 0.147097817335246 0.853029818521207 0.635088264606337 0.232131307183622</Array>"
## [43] "  </Cluster>"
## [44] "  <Cluster name='2' size='701' id='2'>"
## [45] "    <Array n='7' type='real'>0.296491467762999 0.750656736693898 0.0619431822639149 0.157868264857416 0.848902346065933 0.628525296142985 0.271688350691279</Array>"
## [46] "  </Cluster>"
## [47] "  <Cluster name='3' size='666' id='3'>"
```

```

## [48] "    <Array n=7 type='real'>0.252429273848283 0.247773003287282 0.0491943173274577 0.152934106441195 0.854962479643253 0.679990801612424 0.278616586511323</Array>"
## [49] "  </Cluster>"
## [50] "  <Cluster name='4' size='624' id='4'>"
## [51] "    <Array n=7 type='real'>0.771209924685494 0.751556400348525 0.0452115106311922 0.14787504778024 0.851043347132821 0.641072891072892 0.248897484094852</Array>"
## [52] "  </Cluster>"
## [53] " </ClusteringModel>"
## [54] "</PMML>"

```

Lista de estrellas candidatas:

catalog	id_hipparcos	id_tycho
hipparcos	13532	102081
hipparcos	13533	101620
hipparcos	13547	151361
hipparcos	13571	102069
hipparcos	13579	101856
hipparcos	13589	102070
hipparcos	13600	101875
hipparcos	13601	NA
hipparcos	13627	151362
hipparcos	13631	NA
hipparcos	13643	52352
hipparcos	13654	101881
hipparcos	13702	NA
hipparcos	13723	102049
hipparcos	13752	NA
hipparcos	13786	101615
hipparcos	13801	151402
hipparcos	13824	101866
hipparcos	13846	101641
hipparcos	13855	101844
hipparcos	13867	151442
hipparcos	13870	101861
hipparcos	13871	101648
hipparcos	13890	102050
hipparcos	13892	102075
hipparcos	13893	101840
hipparcos	13913	101806
hipparcos	13914	NA
hipparcos	14021	101869
hipparcos	14024	151413
hipparcos	14048	101649
hipparcos	14062	101810
hipparcos	14069	52333
hipparcos	14089	101686
hipparcos	14098	101683
hipparcos	14129	151426
hipparcos	14132	101926
hipparcos	14140	101952
hipparcos	14141	101935
hipparcos	14152	102117
hipparcos	14183	101947
hipparcos	14201	101922
hipparcos	14203	102150
hipparcos	14211	102124
hipparcos	14230	151401
hipparcos	14236	52363
hipparcos	14292	102137
hipparcos	14310	102135
hipparcos	14318	NA
hipparcos	14378	102158
hipparcos	14391	101918
hipparcos	14394	151476
hipparcos	14403	101957
hipparcos	14451	151446
hipparcos	14481	102148
hipparcos	14484	101968
hipparcos	14493	101919
hipparcos	14497	101716
hipparcos	14500	101678
hipparcos	14514	101910
hipparcos	14531	151496
hipparcos	14550	101668
hipparcos	14554	151453
hipparcos	14567	101660
hipparcos	14569	101689
hipparcos	14572	151490
hipparcos	14586	101909
hipparcos	14614	NA
hipparcos	14649	102169
hipparcos	14672	101889
hipparcos	14675	101884
hipparcos	14679	101719
hipparcos	14682	101664
hipparcos	14683	101904
hipparcos	14689	102107
hipparcos	14705	NA
hipparcos	14715	102138
hipparcos	14721	101969
hipparcos	14727	102106
hipparcos	14735	101671

hipparcos	14751	101718
hipparcos	14767	102134
hipparcos	14807	102139
hipparcos	14809	102122
hipparcos	14810	102170
hipparcos	14818	102097
hipparcos	14838	101971
hipparcos	14847	151443
hipparcos	14863	101666
hipparcos	14901	102336
hipparcos	14908	151469
hipparcos	14909	102197
hipparcos	14910	102176
hipparcos	14929	102555
hipparcos	14971	102349
hipparcos	14974	102701
hipparcos	14993	102724
hipparcos	15038	102733
hipparcos	15048	102554
hipparcos	15052	102364
hipparcos	15057	151515
hipparcos	15062	NA
hipparcos	15064	102175
hipparcos	15072	151512
hipparcos	15098	102323
hipparcos	15100	102518
hipparcos	15109	102186
hipparcos	15110	102734
hipparcos	15129	102187
hipparcos	15134	102177
hipparcos	15137	52772
hipparcos	15153	52754
hipparcos	15189	102705
hipparcos	15198	102703
hipparcos	15227	102551
hipparcos	15249	NA
hipparcos	15267	151535
hipparcos	15268	151543
hipparcos	15278	151499
hipparcos	15283	102192
hipparcos	15341	151553
hipparcos	15380	151517
hipparcos	15381	102334
hipparcos	15394	NA
hipparcos	15412	102328
hipparcos	15506	102552
hipparcos	15509	102533
hipparcos	15513	102706
hipparcos	15519	102370
hipparcos	15525	102216
hipparcos	15536	52749
hipparcos	15546	102221
hipparcos	15550	NA
hipparcos	15552	151524
hipparcos	15555	102589
hipparcos	15564	102218
hipparcos	15566	102572
hipparcos	15586	151528
hipparcos	15588	102381
hipparcos	15597	102567
hipparcos	15608	102741
hipparcos	15612	102566
hipparcos	15627	102787
hipparcos	15640	152101
hipparcos	15651	102749
hipparcos	15683	NA
hipparcos	15724	102386
hipparcos	15737	102788
hipparcos	15766	102769
hipparcos	15777	102211
hipparcos	15805	102226
hipparcos	15808	102384
hipparcos	15823	102215
hipparcos	15838	102610
hipparcos	15844	NA
hipparcos	15854	102420
hipparcos	15870	102759
hipparcos	15882	102771
hipparcos	15894	52809
hipparcos	15908	102786
hipparcos	15939	102747
hipparcos	15951	102223
hipparcos	15954	102404
hipparcos	15961	NA
hipparcos	15996	102593
hipparcos	16072	NA
hipparcos	16077	102594
hipparcos	16109	102591
hipparcos	16117	102224
hipparcos	16119	152122
hipparcos	16132	152124
hipparcos	16133	102757

hipparcos	16136	102586
hipparcos	16143	102614
hipparcos	16143	102655
hipparcos	16144	NA
hipparcos	16150	NA
hipparcos	16152	102424
hipparcos	16174	102805
hipparcos	16176	102279
hipparcos	16181	152153
hipparcos	16205	102612
hipparcos	16221	52815
hipparcos	16241	102277
hipparcos	16246	102276
hipparcos	16268	102247
hipparcos	16271	102436
hipparcos	16295	152171
hipparcos	16304	52816
hipparcos	16305	152475
hipparcos	16329	102641
hipparcos	16359	52800
hipparcos	16374	102625
hipparcos	16393	102251
hipparcos	16401	NA
hipparcos	16405	NA
hipparcos	16414	102442
hipparcos	16416	102429
hipparcos	16423	102817
hipparcos	16428	152155
hipparcos	16440	152206
hipparcos	16441	102624
hipparcos	16469	102269
hipparcos	16505	102629
hipparcos	16512	102453
hipparcos	16529	152212
hipparcos	16532	102263
hipparcos	16538	102271
hipparcos	16551	102642
hipparcos	16568	102260
hipparcos	16571	102258
hipparcos	16593	52794
hipparcos	16598	102443
hipparcos	16613	102448
hipparcos	16618	102821
hipparcos	16627	102261
hipparcos	16639	152173
hipparcos	16641	NA
hipparcos	16643	102812
hipparcos	16659	152192
hipparcos	16662	102634
hipparcos	16664	152488
hipparcos	16668	152156
hipparcos	16671	53242
hipparcos	16674	102650
hipparcos	16701	102637
hipparcos	16703	102819
hipparcos	16706	102449
hipparcos	16707	102646
hipparcos	16737	102647
hipparcos	16753	152169
hipparcos	16784	152178
hipparcos	16788	NA
hipparcos	16802	102503
hipparcos	16808	102692
hipparcos	16811	102493
hipparcos	16813	102661
hipparcos	16824	152176
hipparcos	16828	102838
hipparcos	16831	102299
hipparcos	16859	152185
hipparcos	16865	102303
hipparcos	16891	102282
hipparcos	16896	102475
hipparcos	16908	102854
hipparcos	16924	102319
hipparcos	16944	102468
hipparcos	16950	102664
hipparcos	16961	102694
hipparcos	16979	102823
hipparcos	16985	102283
hipparcos	17000	152237
hipparcos	17009	102659
hipparcos	17026	102865
hipparcos	17043	102860
hipparcos	17044	152522
hipparcos	17049	53284
hipparcos	17058	102313
hipparcos	17068	53259
hipparcos	17076	NA
hipparcos	17091	152223
hipparcos	17121	102295
hipparcos	17148	102856
hipparcos	17174	102462

hipparcos	17225	152263
hipparcos	17232	53247
hipparcos	17234	102690
hipparcos	17281	53258
hipparcos	17282	102472
hipparcos	17289	152264
hipparcos	17291	102478
hipparcos	17300	53251
hipparcos	17305	152252
hipparcos	17309	102695
hipparcos	17316	102839
hipparcos	17317	102828
hipparcos	17318	102317
hipparcos	17325	102678
hipparcos	17336	102502
hipparcos	17340	102837
hipparcos	17401	152251
hipparcos	17405	NA
hipparcos	17408	102668
hipparcos	17413	102302
hipparcos	17414	NA
hipparcos	17421	102674
hipparcos	17426	53268
hipparcos	17453	103344
hipparcos	17466	103047
hipparcos	17481	103340
hipparcos	17489	152275
hipparcos	17497	152257
hipparcos	17499	152276
hipparcos	17503	103165
hipparcos	17505	53297
hipparcos	17511	103313
hipparcos	17531	152565
hipparcos	17552	103328
hipparcos	17572	152248
hipparcos	17573	152274
hipparcos	17586	103028
hipparcos	17589	103333
hipparcos	17607	NA
hipparcos	17608	152362
hipparcos	17610	102883
hipparcos	17614	103034
hipparcos	17657	103025
hipparcos	17684	152336
hipparcos	17692	152331
hipparcos	17694	152342
hipparcos	17702	152360
hipparcos	17704	152327
hipparcos	17708	102887
hipparcos	17714	53302
hipparcos	17726	53320
hipparcos	17740	103040
hipparcos	17746	103029
hipparcos	17763	103327
hipparcos	17776	152365
hipparcos	17791	152323
hipparcos	17794	102869
hipparcos	17803	NA
hipparcos	17804	102874
hipparcos	17828	152349
hipparcos	17830	102881
hipparcos	17832	152363
hipparcos	17836	53288
hipparcos	17837	103324
hipparcos	17847	152361
hipparcos	17851	152358
hipparcos	17862	152629
hipparcos	17864	102876
hipparcos	17892	152277
hipparcos	17900	152364
hipparcos	17909	103308
hipparcos	17920	103173
hipparcos	17921	103348
hipparcos	17923	152309
hipparcos	17928	152278
hipparcos	17944	102890
hipparcos	17960	NA
hipparcos	17961	103162
hipparcos	17962	103023
hipparcos	17983	103043
hipparcos	17999	152316
hipparcos	18018	152315
hipparcos	18025	103154
hipparcos	18059	103302
hipparcos	18091	103228
hipparcos	18096	103072
hipparcos	18097	NA
hipparcos	18110	103216
hipparcos	18116	103081
hipparcos	18158	NA
hipparcos	18165	103223

hipparcos	18176	102909
hipparcos	18181	152296
hipparcos	18222	103068
hipparcos	18227	102939
hipparcos	18228	102908
hipparcos	18244	NA
hipparcos	18263	NA
hipparcos	18266	NA
hipparcos	18267	NA
hipparcos	18274	53309
hipparcos	18283	103069
hipparcos	18296	153160
hipparcos	18304	103080
hipparcos	18326	103064
hipparcos	18327	103076
hipparcos	18346	153157
hipparcos	18347	NA
hipparcos	18354	103062
hipparcos	18405	103390
hipparcos	18409	103190
hipparcos	18431	153149
hipparcos	18433	NA
hipparcos	18459	103364
hipparcos	18471	103399
hipparcos	18473	103199
hipparcos	18474	102932
hipparcos	18485	153154
hipparcos	18490	103357
hipparcos	18502	103375
hipparcos	18506	103050
hipparcos	18508	153359
hipparcos	18531	153179
hipparcos	18542	102925
hipparcos	18544	103397
hipparcos	18551	53379
hipparcos	18559	153147
hipparcos	18576	102903
hipparcos	18595	103061
hipparcos	18609	103058
hipparcos	18620	153153
hipparcos	18625	153186
hipparcos	18626	102924
hipparcos	18634	53374
hipparcos	18643	NA
hipparcos	18684	103424
hipparcos	18692	53381
hipparcos	18696	153170
hipparcos	18717	103102
hipparcos	18718	53349
hipparcos	18719	103245
hipparcos	18721	103244
hipparcos	18730	102953
hipparcos	18735	103115
hipparcos	18748	153156
hipparcos	18748	153190
hipparcos	18754	102964
hipparcos	18762	103252
hipparcos	18784	103267
hipparcos	18787	103251
hipparcos	18792	103234
hipparcos	18803	103416
hipparcos	18806	53389
hipparcos	18820	103273
hipparcos	18823	103092
hipparcos	18833	103401
hipparcos	18843	102950
hipparcos	18849	103101
hipparcos	18864	102955
hipparcos	18902	53360
hipparcos	18909	103104
hipparcos	18946	103257
hipparcos	18955	153192
hipparcos	18976	103433
hipparcos	18995	153223
hipparcos	18998	102954
hipparcos	19002	103434
hipparcos	19009	153238
hipparcos	19032	102968
hipparcos	19036	103106
hipparcos	19038	103435
hipparcos	19054	103430
hipparcos	19072	153205
hipparcos	19075	NA
hipparcos	19076	NA
hipparcos	19078	153201
hipparcos	19082	NA
hipparcos	19085	53823
hipparcos	19093	103271
hipparcos	19098	103094
hipparcos	19120	102949
hipparcos	19146	103429
hipparcos	19148	102979

hipparcos	19157	53835
hipparcos	19163	103410
hipparcos	19175	103259
hipparcos	19176	103242
hipparcos	19187	103108
hipparcos	19207	102942
hipparcos	19216	103427
hipparcos	19262	103421
hipparcos	19263	102966
hipparcos	19281	102959
hipparcos	19284	103116
hipparcos	19299	103448
hipparcos	19308	103292
hipparcos	19358	103137
hipparcos	19361	153228
hipparcos	19363	NA
hipparcos	19367	103277
hipparcos	19376	53825
hipparcos	19388	103295
hipparcos	19401	103140
hipparcos	19403	103282
hipparcos	19417	153207
hipparcos	19420	103121
hipparcos	19436	153225
hipparcos	19451	103013
hipparcos	19459	153226
hipparcos	19464	153229
hipparcos	19470	103117
hipparcos	19472	NA
hipparcos	19484	153193
hipparcos	19499	102989
hipparcos	19519	103008
hipparcos	19526	53830
hipparcos	19529	103000
hipparcos	19542	103453
hipparcos	19548	103002
hipparcos	19549	103127
hipparcos	19558	103452
hipparcos	19586	153450
hipparcos	19591	153250
hipparcos	19615	103439
hipparcos	19616	103471
hipparcos	19617	53833
hipparcos	19634	153264
hipparcos	19641	103151
hipparcos	19651	103146
hipparcos	19662	153266
hipparcos	19665	103470
hipparcos	19669	103274
hipparcos	19672	103473
hipparcos	19696	102992
hipparcos	19702	153285
hipparcos	19745	153287
hipparcos	19763	103288
hipparcos	19772	103290
hipparcos	19785	153256
hipparcos	19789	103461
hipparcos	19793	153260
hipparcos	19797	NA
hipparcos	19822	102993
hipparcos	19829	153248
hipparcos	19834	NA
hipparcos	19847	103130
hipparcos	19853	153245
hipparcos	19862	NA
hipparcos	19870	103465
hipparcos	19960	NA
hipparcos	19975	103815
hipparcos	19981	153303
hipparcos	19990	103859
hipparcos	19997	NA
hipparcos	20011	103478
hipparcos	20012	NA
hipparcos	20018	103821
hipparcos	20019	103632
hipparcos	20025	53862
hipparcos	20038	104052
hipparcos	20041	103638
hipparcos	20056	103639
hipparcos	20071	153291
hipparcos	20084	103813
hipparcos	20086	103652
hipparcos	20087	104080
hipparcos	20130	103823
hipparcos	20136	104042
hipparcos	20139	153307
hipparcos	20146	103641
hipparcos	20150	53844
hipparcos	20171	104081
hipparcos	20175	103819
hipparcos	20181	104034
hipparcos	20182	53855

hipparcos	20186	104079
hipparcos	20197	NA
hipparcos	20209	153321
hipparcos	20219	53867
hipparcos	20226	153331
hipparcos	20237	103827
hipparcos	20255	103661
hipparcos	20261	103500
hipparcos	20270	153335
hipparcos	20278	104069
hipparcos	20284	53869
hipparcos	20290	103482
hipparcos	20336	103650
hipparcos	20352	103480
hipparcos	20369	103829
hipparcos	20371	53894
hipparcos	20389	103491
hipparcos	20390	103830
hipparcos	20391	104049
hipparcos	20400	53913
hipparcos	20412	153315
hipparcos	20417	NA
hipparcos	20440	103488
hipparcos	20482	103853
hipparcos	20484	103498
hipparcos	20485	103494
hipparcos	20493	104082
hipparcos	20508	53901
hipparcos	20513	153342
hipparcos	20526	NA
hipparcos	20527	53905
hipparcos	20530	53886
hipparcos	20531	153351
hipparcos	20533	153352
hipparcos	20542	103709
hipparcos	20553	53874
hipparcos	20563	103708
hipparcos	20580	153348
hipparcos	20600	103867
hipparcos	20605	NA
hipparcos	20614	103896
hipparcos	20631	153318
hipparcos	20635	104124
hipparcos	20641	104125
hipparcos	20644	153328
hipparcos	20646	103874
hipparcos	20648	103711
hipparcos	20679	NA
hipparcos	20680	104101
hipparcos	20686	103868
hipparcos	20711	153354
hipparcos	20713	103551
hipparcos	20719	103506
hipparcos	20732	53911
hipparcos	20740	53902
hipparcos	20756	103677
hipparcos	20769	103862
hipparcos	20769	103897
hipparcos	20774	103884
hipparcos	20780	103710
hipparcos	20782	NA
hipparcos	20785	104113
hipparcos	20789	153353
hipparcos	20810	NA
hipparcos	20827	53872
hipparcos	20834	NA
hipparcos	20842	104126
hipparcos	20850	53888
hipparcos	20859	104093
hipparcos	20866	153878
hipparcos	20873	53941
hipparcos	20877	103543
hipparcos	20890	103888
hipparcos	20894	103547
hipparcos	20897	104116
hipparcos	20898	53936
hipparcos	20917	104090
hipparcos	20952	53935
hipparcos	20971	103700
hipparcos	20978	103536
hipparcos	20994	104117
hipparcos	20995	103550
hipparcos	21013	153880
hipparcos	21019	103865
hipparcos	21029	103544
hipparcos	21036	53942
hipparcos	21039	103549
hipparcos	21082	103552
hipparcos	21089	NA
hipparcos	21090	53914
hipparcos	21094	153881



hipparcos	21099	103878
hipparcos	21112	53933
hipparcos	21115	104083
hipparcos	21120	153894
hipparcos	21123	103699
hipparcos	21132	53934
hipparcos	21137	103548
hipparcos	21138	NA
hipparcos	21158	104159
hipparcos	21173	103737
hipparcos	21240	103909
hipparcos	21246	153877
hipparcos	21251	103715
hipparcos	21256	104170
hipparcos	21257	103584
hipparcos	21261	103933
hipparcos	21273	53940
hipparcos	21280	103571
hipparcos	21314	103929
hipparcos	21318	104176
hipparcos	21332	103738
hipparcos	21341	153899
hipparcos	21347	153875
hipparcos	21348	153882
hipparcos	21350	103729
hipparcos	21359	103720
hipparcos	21365	103752
hipparcos	21380	153890
hipparcos	21399	104177
hipparcos	21403	103915
hipparcos	21408	103903
hipparcos	21421	NA
hipparcos	21436	104149
hipparcos	21517	103725
hipparcos	21533	54776
hipparcos	21554	103574
hipparcos	21932	NA
tycho	17563	4945
tycho	18481	4956
tycho	18089	4997
tycho	19411	5443
tycho	18957	5466
tycho	18907	5487
tycho	19859	5526
tycho	20794	5578
tycho	20268	5632
tycho	21295	6285
tycho	22549	6486
tycho	22923	6545
tycho	21135	6598
tycho	21575	6660
tycho	22245	6734
tycho	24021	8410
tycho	24681	8426
tycho	25453	8552
tycho	25813	9658
tycho	27989	9858
tycho	16083	52545
tycho	16511	52546
tycho	15900	52547
tycho	16478	52666
tycho	16322	52667
tycho	16095	52834
tycho	16369	52844
tycho	17907	52920
tycho	18297	52923
tycho	18618	52985
tycho	18893	52986
tycho	18805	52991
tycho	18887	53038
tycho	18975	53039
tycho	18715	53175
tycho	18724	53229
tycho	17617	53310
tycho	18033	53321
tycho	20165	53424
tycho	19718	53436
tycho	19719	53441
tycho	NA	53451
tycho	19860	53453
tycho	19740	53454
tycho	20458	53458
tycho	20715	53460
tycho	20281	53462
tycho	20852	53535
tycho	19103	53551
tycho	19217	53556
tycho	NA	53558
tycho	NA	53569
tycho	20188	53576
tycho	19821	53580
tycho	NA	53584

tycho	19720	53588
tycho	20089	53592
tycho	19663	53593
tycho	19750	53597
tycho	19799	53599
tycho	20597	53603
tycho	NA	53612
tycho	20830	53621
tycho	20447	53626
tycho	20318	53627
tycho	20845	53631
tycho	20734	53634
tycho	NA	53635
tycho	20804	53637
tycho	20522	53638
tycho	21091	53641
tycho	NA	53646
tycho	21308	53648
tycho	NA	53649
tycho	20963	53657
tycho	21266	53659
tycho	NA	53660
tycho	20985	53670
tycho	NA	53675
tycho	21269	53687
tycho	NA	53688
tycho	21402	53689
tycho	19481	53690
tycho	19223	53702
tycho	19736	53715
tycho	20174	53716
tycho	19786	53725
tycho	20826	53744
tycho	20472	53749
tycho	20616	53758
tycho	20511	53759
tycho	20540	53766
tycho	20595	53773
tycho	20558	53777
tycho	21335	53795
tycho	20960	53800
tycho	21395	53804
tycho	20901	53805
tycho	NA	53925
tycho	NA	53938
tycho	NA	53980
tycho	21626	54007
tycho	21670	54009
tycho	22461	54037
tycho	22221	54050
tycho	NA	54083
tycho	22509	54091
tycho	22984	54108
tycho	22938	54119
tycho	22834	54155
tycho	22722	54156
tycho	NA	54191
tycho	23666	54234
tycho	21506	54299
tycho	NA	54325
tycho	21882	54334
tycho	22258	54367
tycho	22475	54371
tycho	22239	54378
tycho	22044	54406
tycho	22927	54426
tycho	NA	54437
tycho	23035	54438
tycho	NA	54439
tycho	22616	54456
tycho	22845	54482
tycho	22597	54483
tycho	23239	54506
tycho	23303	54513
tycho	NA	54540
tycho	NA	54548
tycho	23494	54558
tycho	21672	54564
tycho	21650	54569
tycho	21904	54573
tycho	21446	54585
tycho	21589	54591
tycho	21735	54592
tycho	22386	54595
tycho	22327	54597
tycho	22021	54602
tycho	22267	54612
tycho	22157	54637
tycho	22826	54652
tycho	NA	54670
tycho	22823	54675
tycho	22682	54676

tycho	NA	54682
tycho	22831	54683
tycho	22833	54684
tycho	23382	54716
tycho	23301	54718
tycho	NA	54730
tycho	23492	54737
tycho	23404	54751
tycho	21654	54768
tycho	21942	54791
tycho	22208	54813
tycho	22063	54814
tycho	22942	54843
tycho	22853	54857
tycho	22914	54864
tycho	22667	54884
tycho	22957	54886
tycho	23621	54896
tycho	23629	54910
tycho	23443	54912
tycho	NA	54932
tycho	23286	54935
tycho	23710	54942
tycho	23159	54957
tycho	23161	54958
tycho	23896	55036
tycho	24031	55050
tycho	23879	55053
tycho	24730	55064
tycho	25041	55155
tycho	23770	55327
tycho	23866	55336
tycho	NA	55351
tycho	23983	55377
tycho	24632	55386
tycho	24998	55394
tycho	24921	55399
tycho	24971	55420
tycho	26215	55641
tycho	26176	55642
tycho	23977	55673
tycho	23980	55698
tycho	24633	55715
tycho	24527	55719
tycho	24452	55721
tycho	NA	55737
tycho	25037	55744
tycho	24555	55759
tycho	23789	55945
tycho	24059	55975
tycho	24883	55985
tycho	24789	56032
tycho	25638	56083
tycho	25470	56087
tycho	25886	56158
tycho	26093	56175
tycho	26487	56208
tycho	26414	56266
tycho	27511	57138
tycho	27316	57540
tycho	27364	57542
tycho	NA	103490
tycho	NA	103511
tycho	NA	103522
tycho	NA	103525
tycho	21543	103559
tycho	NA	103578
tycho	NA	103596
tycho	NA	103597
tycho	21588	103598
tycho	21683	103599
tycho	21673	103600
tycho	21873	103611
tycho	NA	103617
tycho	22203	103627
tycho	NA	103702
tycho	21630	103746
tycho	21923	103759
tycho	22176	103807
tycho	22422	104257
tycho	22505	104265
tycho	NA	104268
tycho	22347	104274
tycho	22773	104275
tycho	NA	104277
tycho	22846	104284
tycho	22566	104289
tycho	22441	104291
tycho	22913	104292
tycho	23939	104370
tycho	23669	104425

tycho	24197	104451
tycho	24010	104452
tycho	23607	104453
tycho	24475	104528
tycho	22496	104538
tycho	22782	104568
tycho	23043	104615
tycho	22850	104808
tycho	22565	104813
tycho	23871	104895
tycho	23497	105059
tycho	25216	105180
tycho	25410	105184
tycho	25502	105218
tycho	25555	105247
tycho	25790	105326
tycho	26777	105353
tycho	25499	105585
tycho	26382	105707

---