Ejecutar la elección de las columnas

* 10 binarias (polyatomic categorical and seven numerical variables)
* Media cero variancia uno (centered) == PCA
* Split (random)

PLanet Name

Number of stars

Numbers of planets

Discovery Metyhod

Discover Facility

Parttition

Teoria (modelos centrados, supervisados o no supervisados, variables cuantitativas o cualitativas)

Diferentes modelos

Clustering (por ejemplo un modelo circular o eluipsoides y cuando es mejor un y el otro)

Supervisado modelo predictivo

Modelo no supervisado el que clasifica (tiempo)

PCA====

Classification tree el ejercicio 2

Slide 4 Machine learning

Correo Daniel

<https://exoplanetarchive.ipac.caltech.edu/docs/data.html>

ID

Name

1 bynari categorical variable

Controversial Flag (integer class)

2 polyatomic categorical variable (or more)

Discoverymethod

Discovery Year

Discovery Facility

Spectral Type

Seven numerical variables

# COLUMN sy\_snum: Number of Stars

# COLUMN sy\_pnum: Number of Planets

Orbital Period [days]

Planet Radius [Earth Radius]

Planet Mass or Mass\*sin(i) [Earth Mass]

Eccentricity

Equilibrium Temperature [K]

Distance (parsec)

Orbit Semi-Major Axis