



CONVOCATORIA

ExoData Challenge

Feria de Puebla 2026 · Planetario de Puebla

Objetivo

El Datathón de Exoplanetas tiene como finalidad retar a los participantes a explorar y agrupar exoplanetas mediante técnicas de aprendizaje no supervisado (clustering). Utilizando datos reales del NASA Exoplanet Archive, los equipos deberán identificar grupos significativos de exoplanetas, interpretar sus características astrofísicas y presentar sus hallazgos siguiendo la metodología CRISP-DM.

Dataset

Se utilizará la tabla *Planetary Systems Composite Parameters (PSCompPars)* del NASA Exoplanet Archive, que contiene información de **6147 exoplanetas confirmados**. Las principales variables de interés son:

Variable	Descripción	Unidad
<i>pl_rade</i>	Radio del planeta	Radios terrestres
<i>pl_bmasse</i> / <i>pl_bmassj</i>	Masa del planeta	Masas terrestres / Masas de Júpiter
<i>pl_orbper</i>	Periodo orbital	Días
<i>pl_eqt</i>	Temperatura de equilibrio	Kelvin
<i>pl_orbsmax</i>	Semi-eje mayor	Unidades astronómicas (AU)

Sin embargo, los concursantes podrán utilizar cualquier otra variable del dataset que consideren pertinente.

Nota: No todos los registros cuentan con valores para todas las variables. Parte del reto consiste en el manejo adecuado de datos faltantes.

Enlace al dataset: <https://exoplanetarchive.ipac.caltech.edu>

Estructura del Concurso

El Datathón se llevará a cabo en dos etapas:

Etapas 1: Desarrollo y Análisis

A cada equipo se le proporcionará acceso al dataset completo. Los equipos deberán:

1. Realizar un análisis exploratorio de los datos (EDA) e identificar patrones preliminares.



2. Seleccionar, justificar y aplicar al menos una técnica de clustering (por ejemplo: K-Means, DBSCAN, clustering jerárquico, Gaussian Mixture Models, entre otros).
3. Evaluar la calidad de los clusters obtenidos mediante métricas internas (Silhouette Score, Davies-Bouldin Index, etc.).
4. Interpretar los clusters en contexto astrofísico: ¿qué tipos de planetas representan? ¿Son consistentes con taxonomías conocidas (hot Jupiters, super-Tierras, mini-Neptunos, etc.)?

Etapa 2: Presentación Final

Fecha: **miércoles 29 de abril de 2026** | Sede: **Planetario de Puebla**

Cada equipo deberá preparar una presentación estructurada de acuerdo con la metodología CRISP-DM:

1. Comprensión del problema
2. Comprensión de los datos
3. Preparación de los datos
4. Modelado (clustering)
5. Evaluación de resultados
6. Conclusiones y propuesta de aplicación

La presentación deberá incluir:

- Justificación del modelo o modelos de clustering seleccionados
- Metodología de preprocesamiento y selección de variables
- Visualizaciones de los clusters obtenidos
- Interpretación astrofísica de cada grupo identificado
- Métricas de validación interna
- Reflexiones, limitaciones y propuestas de mejora

Calendario

Fecha	Actividad
27 de marzo de 2026	Publicación de la convocatoria y apertura de inscripciones
20 de abril de 2026	Cierre de inscripciones y entrega de datasets
14 - 27 de abril de 2026	Periodo de desarrollo y análisis
27 de abril de 2026	Fecha límite de entrega de resultados
29 de abril de 2026	Presentación final y premiación (Planetario de Puebla)

Evaluación y Puntaje

El jurado evaluará a los equipos en dos rubros principales. La puntuación máxima es de 100 puntos.



Rubro	Criterio	Puntos
Desarrollo y Análisis (70 pts)	Correcto uso de técnicas de clustering y validación	20
	Creatividad y solidez en la construcción del modelo	20
	Interpretación astrofísica de los clusters	30
Presentación Final (30 pts)	Claridad y estructura (metodología CRISP-DM)	10
	Calidad de la comunicación de resultados	10
	Viabilidad de propuestas y conclusiones	10
TOTAL		100

Nota importante: Dado que se trata de un problema de aprendizaje no supervisado, la evaluación pondrá especial énfasis en la capacidad del equipo para interpretar y justificar los clusters encontrados en el contexto del dominio astrofísico, más allá de las métricas numéricas.

Requisitos de Participación

- Equipos de 2 a 4 integrantes.
- Conocimientos básicos en programación y análisis de datos.
- Acceso a herramientas de análisis de datos (Python, R, o similares).
- Compromiso para participar en ambas etapas del concurso.

Inscripción

La inscripción es gratuita. Cada equipo deberá registrarse a través de la plataforma oficial del evento antes del 14 de abril de 2026. El enlace de inscripción y toda la información complementaria (dataset, materiales de apoyo, guía de contexto astrofísico) estarán disponibles en la plataforma.

Liga de registro de participantes: <https://concursosferia.yankuilotl.com.mx/>

Entrega de resultados: deberán subir una liga para acceder a su repositorio o carpeta de archivos a competir compartido desde cualquier plataforma que permita acceso público.

Premios y Reconocimientos

La premiación se llevará a cabo el 29 de abril de 2026 en el Planetario de Puebla, como parte de las actividades de la Feria de Puebla 2026.

Dudas o comentarios comunicarse al correo: equipo.stem@udlap.mx