

PLANEACIÓN DEL CURSO

Temas Selectos de Procesos Químicos-Análisis de Datos con Python

Clave: 2122075, **Grupo:** CH03
Trimestre 26-I

I. Información general

Clases: lunes , miércoles y viernes de 14:30 a 16:00 hrs. Aula: **Sala-F**.
Asesorías: jueves 13:00-15:00 hrs. (T-152). Enviar correo de confirmación.
Profesor: Dra. Abigail Marin López, Área de I.Q., T-152.
Tareas: temas.datosiq@gmail.com

II. Información sobre el programa

II.a. Objetivos del curso

Al finalizar, el alumnado será capaz de:

- Usar Python (Jupyter, NumPy, pandas) para cargar, limpiar y manipular datos experimentales.
- Explorar y visualizar datos de procesos (temperatura, presión, flujo, composición) y construir gráficos útiles para ingeniería (líneas, histogramas, diagramas de dispersión y mapas de calor).
- Analizar series temporales de procesos (remuestreo, ventanas móviles, tendencia y estacionalidad) aplicadas a datos de operación.
- Implementar modelos básicos de ML (regresión y clasificación) para problemas típicos de IQ: balance de energía, rendimiento, detección de condiciones anómalas.
- Desarrollar un proyecto aplicado de análisis de datos en un contexto de ingeniería química (secado, reactor, destilación, emisiones, etc.) priorizando reproducibilidad.

II.b. Contenido sintético

1. Introducción al análisis de datos con Python
 - 1.1. Introducción a Python como herramienta para el análisis de datos en ingeniería química.
 - 1.2. Uso de entornos interactivos (Jupyter Notebook).
 - 1.3. Sintaxis básica, tipos de datos y estructuras fundamentales.
 - 1.4. Lectura y manejo de datos experimentales en formatos comunes (CSV, Excel).
 - 1.5. NumPy: matrices y computación vectorizada. (Álgebra de vectores y matrices)
2. Manipulación y análisis de datos con pandas
 - 2.1. Indexación, filtrado, agrupación y combinación de tablas.
 - 2.2. Estadísticas descriptivas aplicada a variables de procesos.
 - 2.3. Pruebas de normalidad en datos experimentales.
 - 2.4. Análisis de correlación y covarianza entre variables de operación.

3. Acondicionamiento de datos
 - 3.1. Exploración y diagnóstico de conjuntos de datos de procesos químicos.
 - 3.2. Identificación y tratamiento de datos faltantes, duplicados y atípicos.
 - 3.3. Criterios estadísticos para detección de outliers (IQR, Z-score).
 - 3.4. Acondicionamiento de datos de sensores de proceso (temperatura, presión, flujo, conversión).
 - 3.5. Documentación y reporte del proceso de limpieza y depuración de datos.
4. Visualización de datos
 - 4.1. Uso de Matplotlib y Seaborn para la representación de datos de procesos.
 - 4.2. Gráficos de líneas, dispersión, barras y diagramas de caja.
 - 4.3. Mapas de calor para el análisis de condiciones de operación en sistemas como columnas de destilación, secado y reactores.
5. Introducción al análisis de series temporales de procesos
 - 5.1. Manejo de datos temporales: fechas, frecuencias y remuestreo.
 - 5.2. Análisis exploratorio de señales de proceso.
 - 5.3. Identificación de tendencia, estacionalidad y variaciones cíclicas.
 - 5.4. Descomposición de series temporales (STL) aplicada a datos de planta.
 - 5.5. Detección exploratoria de anomalías en señales de proceso.
6. Introducción al modelado y aprendizaje automático en procesos químicos
 - 6.1. Conceptos básicos de modelado y aprendizaje automático aplicados a ingeniería química.
 - 6.2. Uso de scikit-learn para la construcción de modelos sencillos.
 - 6.3. Modelos de regresión y clasificación básicos para datos de proceso.
 - 6.4. Evaluación de modelos mediante métricas de desempeño (MAE, RMSE, R², accuracy).
 - 6.5. Comparación entre un modelo base y un modelo mejorado para un caso de estudio de proceso químico.

II.c. Calendarización del curso

Tema	Semana
Introducción al análisis de datos con Python. Python como herramienta en ingeniería química. Entornos interactivos (Jupyter Notebook). Sintaxis básica, tipos de datos y estructuras fundamentales. Lectura y manejo de datos + NumPy.	1
Lectura de datos experimentales (CSV, Excel). Introducción a NumPy: arreglos, operaciones vectorizadas y álgebra básica aplicada a procesos. Manipulación de datos con pandas. DataFrames, indexación, filtrado, agrupación y combinación de tablas.	2
Análisis estadístico con pandas. Estadísticas descriptivas de procesos. Pruebas de normalidad en datos experimentales. Análisis de correlación y covarianza entre variables de operación.	3
Acondicionamiento de datos I: Exploración y diagnóstico de datasets de procesos químicos. Identificación y tratamiento de datos faltantes y duplicados. Acondicionamiento de datos II: Detección y tratamiento de outliers (IQR, Z-score). Acondicionamiento de datos de sensores (T, P, F, conversión). Documentación y reporte del proceso de limpieza.	4
Visualización de datos de procesos. Uso de Matplotlib y Seaborn. Gráficos de líneas, dispersión, barras, cajas y mapas de calor aplicados a destilación, secado y reactores.	5
Series temporales I: Manejo de datos temporales (fechas, frecuencias, remuestreo). Análisis exploratorio de señales de proceso.	6
	7
	8

Series temporales II: Identificación de tendencia, estacionalidad y ciclos.
Descomposición STL aplicada a datos de planta.
Introducción al modelado y aprendizaje automático.
Conceptos básicos de ML. scikit-learn.
Modelos de regresión y clasificación básicos.
Métricas de desempeño.
Comparación entre modelo base y modelo mejorado.

9
10-11

III. Evaluación del curso

- Proyecto (90%)
- Tareas (10%)

Fechas de presentación del proyecto:

- **Primer parcial:** viernes de la semana 4
- **Segundo parcial:** viernes de la semana 8
- **Presentación de Proyecto Final:** viernes de la semana 11

IV. Criterios y escalas para asignación de calificación final:

$$\begin{aligned} \mathbf{MB} &= 8.6 - 10 \\ \mathbf{B} &= 7.6 - 8.5 \\ \mathbf{S} &= 6.0 - 7.5 \end{aligned}$$

V. Bibliografía

1. Wes McKinney, “*Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*”, 2nd Edition, O’Really Media, 2017.
2. Joel Grus, “*Data Science from Scratch: First Principles with Python*”, 2nd Edition, O’Really Media, 2019.
3. Jake VanderPlas, “*Python Data Science Handbook: Essential Tools for Working with Data*”, 2nd Edition, O’Really Media, 2023.
4. Andreas C. Müller and Sarah Guido, “Introduction to machine learning with Python a guide for data scientists”, O’Really Media, 2018.
5. Mendenhall, W. *Introducción a la Probabilidad y Estadística*, 14^a ed, 2015.

VI. Notas Extra

- La evaluación final es individual. Podrán realizar las tareas y el proyecto final en equipo de máximo 3 participantes.
- Las tareas se enviarán en formato PDF al correo asignado.
- No se recibirán documentos fuera de la fecha y horario previamente establecido.