

Evidencias de la Tarea 05

Ejercicio 1

Prácticas 9A - Extracción de Conocim...Julius AI | Your AI Data AnalystGiggs007/Py-clica06_EBCDHomeTarea05_220864_Ejercicio1y2

localhost8888/notebooks/Downloads/Practica06/Tarea05_220864_Ejercicio1y2.ipynb?

JupyterTarea05_220864_Ejercicio1y2Last Checkpoint: 3 days ago

FileEditViewRunKernelSettingsHelp

JupyterLabPython 3 (ipykernel)

Ejercicio 1: Importación, limpieza, transformación y exportación de datos (base normalizada)

[1]:# Paso 1: Conector y consultar la tabla 'patients'
import mysql.connector
import pandas as pd

conn = mysql.connector.connect(
 host="localhost",
 user="root",
 password="vargas2030", # Cambiar si tu contraseña no está vacía
 database="bd_healthcare",
 port=3308
)

query = "SELECT * FROM patients;"
df = pd.read_sql(query, conn)
df.head()

C:\Users\rsauv\AppData\Local\Temp\ipykernel_4120\1239649613.py:14: UserWarning: pandas only supports SQLAlchemy connectable (engine/connection) or database string URI or sqlite3 DBAPI2 connection. Other DBAPI2 objects are not tested. Please consider using SQLAlchemy.
df = pd.read_sql(query, conn)

[1]:

	patient_id	name	age	gender	blood_type	medical_condition
0	1	Bobby Jackson	30	Male	B-	Cancer
1	2	Leslie TBrly	62	Male	A+	Obesity
2	3	DaNin'sMth	76	Female	A-	Obesity
3	4	andrew watIS	28	Female	O+	Diabetes
4	5	adriENNE dEli	43	Female	AB+	Cancer

[2]:# Paso 2: Limpieza de datos
Verificamos duplicados y valores nulos

Prácticas 9A - Extracción de Conocim...Julius AI | Your AI Data AnalystGiggs007/Py-clica06_EBCDHomeTarea05_220864_Ejercicio1y2

localhost8888/notebooks/Downloads/Practica06/Tarea05_220864_Ejercicio1y2.ipynb?

JupyterTarea05_220864_Ejercicio1y2Last Checkpoint: 3 days ago

FileEditViewRunKernelSettingsHelp

JupyterLabPython 3 (ipykernel)

45adriENNE dEli43FemaleAB+Cancer

[2]:# Paso 2: Limpieza de datos
Verificamos duplicados y valores nulos
print("Duplicados:", df.duplicated().sum())
print("Valores nulos por columna:\n", df.isnull().sum())

Eliminamos duplicados
df = df.drop_duplicates()

Eliminamos filas con datos faltantes
df = df.dropna()

df.info()

Duplicados: 0
Valores nulos por columna:
patient_id 0
name 0
age 0
gender 0
blood_type 0
medical_condition 0
dtype: int64
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54966 entries, 0 to 54965
Data columns (total 6 columns):
Column Non-Null Count Dtype

0 patient_id 54966 non-null int64
1 name 54966 non-null object
2 age 54966 non-null int64
3 gender 54966 non-null object
4 blood_type 54966 non-null object
5 medical_condition 54966 non-null object
dtypes: int64(2), object(4)
memory usage: 2.5+ MB

[3]:# Paso 3: Transformación de datos

Prácticas 9A - Extracción de Conoci... | Julius AI [Your AI Data Analyst] | Gigg9007/Pr-ctica06_EBCD | Home | Tarea05_220864_Ejercicio1y2

localhost:8888/notebooks/Downloads/Practica06/Tarea05_220864_Ejercicio1y2.ipynb?

Jupyter Tarea05_220864_Ejercicio1y2 Last Checkpoint: 3 days ago

File Edit View Run Kernel Settings Help

memory usage: 2.5+ MB

```
[3]: # Paso 3: Transformación de datos
# Clasificar por rangos de edad en una nueva columna 'categoria_edad'

def clasificar_edad(edad):
    if edad < 18:
        return 'Menor de edad'
    elif edad <= 60:
        return 'Adulto'
    else:
        return 'Adulto mayor'

df['categoria_edad'] = df['age'].apply(clasificar_edad)
df[['age', 'categoria_edad']].head()
```

```
[3]:
```

	age	categoria_edad
0	30	Adulto
1	62	Adulto mayor
2	76	Adulto mayor
3	28	Adulto
4	43	Adulto

```
[5]: # Paso 4: Exportar a una tabla temporal 'patients_temp'
from sqlalchemy import create_engine

engine = create_engine("mysql+mysqlconnector://root:vargas2030@localhost:3308/bd_healthcare")
df.to_sql(name='patients_temp', con=engine, if_exists='replace', index=False)
print("Datos exportados correctamente a 'patients_temp'")
Datos exportados correctamente a 'patients_temp'.
```

Ejercicio 2

23°C Nublado

Búsqueda

02:32 p. m. 23/06/2025

Ejercicio 2

Prácticas 9A - Extracción de Conoci... | Julius AI [Your AI Data Analyst] | Gigg9007/Pr-ctica06_EBCD | Home | Tarea05_220864_Ejercicio1y2

localhost:8888/notebooks/Downloads/Practica06/Tarea05_220864_Ejercicio1y2.ipynb?

Jupyter Tarea05_220864_Ejercicio1y2 Last Checkpoint: 3 days ago

File Edit View Run Kernel Settings Help

memory usage: 2.5+ MB

```
[39]: # 1. Importar datos de una API - método GET
import requests
import pandas as pd

url = 'https://jsonplaceholder.typicode.com/users' # API de prueba
response = requests.get(url)
data = response.json()

# Convertir a DataFrame para facilitar el análisis
df = pd.DataFrame(data)
df.head()
```

```
[39]:
```

	id	name	username	email	address	phone	website	company
0	1	Leanne Graham	Bret	Sincere@aprilbiz	{'street': 'Yulas Light', 'suite': 'Apt. 556...', 'city': 'Newburgh', 'state': 'NY', 'zip': '12550'}	1-770-736-8031 x56442	hildegard.org	{'name': 'Romaguera-Coral', 'catchPhrase': 'Multi-tasking user interface', 'slug': 'multi-tasking'}
1	2	Ervin Howell	Antonette	Shanna@melissa.tv	{'street': 'Victor Plains', 'suite': 'Suite 87...', 'city': 'Sunnyvale', 'state': 'CA', 'zip': '95034'}	010-692-6593 x09125	anastasia.net	{'name': 'Deckow-Crist', 'catchPhrase': 'Proactive user interface', 'slug': 'proactive'}
2	3	Clementine Bauch	Samantha	Nathan@yessenia.net	{'street': 'Douglas Extension', 'suite': 'Suite 18...', 'city': 'Hettinger', 'state': 'VT', 'zip': '05643'}	1-463-123-4447	ramiro.info	{'name': 'Romaguera-Jacobson', 'catchPhrase': 'Innovative web interface', 'slug': 'innovative'}
3	4	Patricia Lebsack	Karianne	Jullianne.O'Conner@kory.org	{'street': 'Hoeger Mall', 'suite': 'Apt. 692...', 'city': 'Lynch', 'state': 'MA', 'zip': '01902'}	493-170-9623 x156	kaleb.biz	{'name': 'Robel-Corkery', 'catchPhrase': 'Multi-tier e-commerce interface', 'slug': 'multi-tier'}
4	5	Chelsey Dietrich	Kamren	Lucio_Hettinger@annie.ca	{'street': 'Skiles Walks', 'suite': 'Suite 351...', 'city': 'Kulas', 'state': 'TX', 'zip': '79465'}	(254)954-1289	demarco.info	{'name': 'Keebler LLC', 'catchPhrase': 'User-centric', 'slug': 'user-centric'}

```
[40]: # 2. Limpiar Datos
# Seleccionamos solo las columnas relevantes
df_clean = df[['id', 'name', 'email', 'phone', 'website']]
df_clean.columns = ['ID', 'Nombre', 'Email', 'Teléfono', 'SitioWeb']
df_clean.head()
```

Ejercicio 2

23°C Nublado

Búsqueda

02:34 p. m. 23/06/2025

Prácticas IA - Extracción de Conocimiento | Julius AI | Your AI Data Analyst | Gigg007/Py-chica06_EBCD | Home | Tarea05_220864_Ejercicio1y2

localhost:8888/notebooks/Downloads/Practica06/tarea05_220864_Ejercicio1y2.ipynb?

Jupyter Tarea05_220864_Ejercicio1y2 Last Checkpoint: 3 days ago

File Edit View Run Kernel Settings Help

Python 3 (ipykernel)

```
[40]: # 2. Limpieza de Datos
# Seleccionamos solo las columnas relevantes
df_clean = df[['id', 'name', 'email', 'phone', 'website']]
df_clean.columns = ['ID', 'Nombre', 'Email', 'Telefono', 'SitioWeb']
df_clean.head()
```

	ID	Nombre	Email	Telefono	SitioWeb
0	1	Leanne Graham	Sincere@april.biz	1-770-736-8031 x56442	hilegard.org
1	2	Ervin Howell	Shanna@melissa.tv	010-692-6593 x09125	anastasia.net
2	3	Clementine Bauch	Nathan@yessenia.net	1-463-123-4447	ramiro.info
3	4	Patricia Lebsack	Julianne.OConner@kory.org	493-170-9622 x156	kale.biz
4	5	Cheley Dietrich	Lucio.Hettinger@annie.ca	(254)954-1289	demarco.info

```
[41]: # 3. Transformación de Datos
# Por ejemplo, convertir los nombres en mayúsculas
df_clean['Nombre'] = df_clean['Nombre'].str.upper()
df_clean.head()
```

C:\Users\jesus\AppData\Local\Temp\ipykernel_29644\3591139229.py:13: SettingWithCopyWarning: A value is trying to be set on a copy of a slice from a DataFrame. Try using .loc[row_indexer,col_indexer] = value instead. See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df_clean['Nombre'] = df_clean['Nombre'].str.upper()
```

	ID	Nombre	Email	Telefono	SitioWeb
0	1	LEANNE GRAHAM	Sincere@april.biz	1-770-736-8031 x56442	hilegard.org
1	2	ERVIN HOWELL	Shanna@melissa.tv	010-692-6593 x09125	anastasia.net
2	3	CLEMENTINE BAUCH	Nathan@yessenia.net	1-463-123-4447	ramiro.info
3	4	PATRICIA LEBACK	Julianne.OConner@kory.org	493-170-9622 x156	kale.biz
4	5	CHELSEY DIETRICH	Lucio.Hettinger@annie.ca	(254)954-1289	demarco.info

```
[43]: # 4. Exportar Datos
# Guardar como CSV temporal
df_clean.to_csv('usuarios_API.csv', index=False)
```

```
[43]: # 4. Exportar Datos
# Guardar como CSV temporal
df_clean.to_csv('usuarios_API.csv', index=False)
```