



Almacenamiento y captura de datos

Claudio Aracena

GobLab - Universidad Adolfo Ibáñez
Chatbot Chile



Contenidos

- Captura de datos desde archivos
- Base de datos
- **Captura y almacenamiento de datos en BD**
- Captura de datos de la Web (Web scraping)
- Captura de datos de API (ej: Twitter)
- Captura y almacenamiento en arquitecturas Big data

Códigos y clase en:

<https://github.com/caracena/almacenamiento-captura-datos>

Clase de hoy



Captura y almacenamiento de datos en BD

- Captura de datos desde BD
- Almacenamiento en BD
- Procesamiento de datos en Python





Conexión y consulta a base de datos

```
import sqlite3

conn = sqlite3.connect("data/chinook.db")

cur = conn.cursor()

cur.execute("SELECT * FROM albums")

rows = cur.fetchall()
```



Consulta a base de datos con pandas

```
import pandas as pd
```

```
df = pd.read_sql("SELECT * FROM albums", conn)
```



Creación de base de datos

```
!pip install sqlalchemy
```

```
from sqlalchemy import create_engine
```

```
sql_engine = create_engine('sqlite:///data/test.db')
```

```
connection = sql_engine.raw_connection()
```

```
df.to_sql('artists_albums', connection, index=False)
```