

Hoja informativa: preparación de características

Práctica

```
# Codificación one-hot: obtén características dummy
# drop_first = True - quitar la primera columna (evitar trampas dummy)

pd.get_dummies(df['column'])
pd.get_dummies(df['column'], drop_first=True)
```

```
# Codificación ordinal

from sklearn.preprocessing import OrdinalEncoder

encoder = OrdinalEncoder()
encoder.fit(data)
data_ordinal = encoder.transform(data)

# agregar nombres de columna

data_ordinal = pd.DataFrame(encoder.transform(data), columns=data.columns)

# entrenamiento y transformación automatizados
from sklearn.preprocessing import OrdinalEncoder

encoder = OrdinalEncoder()
data_ordinal = pd.DataFrame(encoder.fit_transform(data), columns=data.columns)
```

```
# Estandarización

from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
scaler.fit(df)
df_scaled = scaler.transform(df)
```

