

Taller 11

```
▶ #SOLUCION
import pandas as pd

# Note que el separador es "," mas no ";"
df = pd.read_csv("/content/ciudades_ecuador.csv", sep=",")

print(df.head())

valores_nulos = df.isna().sum()
print(valores_nulos)

promedio_temp = df['Temperatura Media (°C)'].mean()
print(f"\nPromedio de temperatura calculado: {promedio_temp:.2f} °C")
df['Temperatura Media (°C)'] = df['Temperatura Media (°C)'].fillna(promedio_temp)

df.to_csv('ciudades_limpio.csv', index=False)

...   Código      Ciudad  Poblacion  Altitud (msnm)  Temperatura Media (°C) \
0     UIO        Quito    1700000          2850                  15
1     GYE    Guayaquil    2800000             4                  28
2     CUE       Cuenca    400000          2560                  17
3     ATF       Ambato    350000          2575                  16
4     MEC       Manta    250000             6                  27

Costera
0    False
1    True
2   False
3   False
4    True
Código
Ciudad
Poblacion
Altitud (msnm)
Temperatura Media (°C)
Costera
dtype: int64

Promedio de temperatura calculado: 21.50 °C
```

▶ #SOLUCION

```
import pandas as pd

df = pd.read_csv('ciudades_ecuador.csv')

filtro_ciudades = df[(df['Poblacion'] > 300000) & (df['Altitud (msnm)'] < 1000)]

print(filtro_ciudades[['Ciudad', 'Poblacion', 'Altitud (msnm)']])

media_temp = df['Temperatura Media (°C)'].mean()
df['Temperatura Media (°C)'] = df['Temperatura Media (°C)'].mask(df['Temperatura Media (°C)'])

filtro_final = df[(df['Altitud (msnm)'] > 2000) | (df['Temperatura Media (°C)'] < 18)]
print(filtro_final)
```

	Ciudad	Poblacion	Altitud (msnm)	Codigo	Ciudad	Poblacion	Altitud (msnm)	Temperatura Media (°C)	Costera
1	Guayaquil	2800000	4	UIO	Quito	1700000	2850	15	False
2	CUE	400000	2560	CUE	Cuenca	400000	2560	17	False
3	ATF	350000	2575	ATF	Ambato	350000	2575	16	False
5	LOH	200000	2100	LOH	Loja	200000	2100	16	False
9	IBB	180000	2225	IBB	Ibarra	180000	2225	18	False

▶ # SOLUCION

```
import pandas as pd

df = pd.read_csv('ciudades_ecuador.csv')

df_ordenado = df.sort_values(by='Poblacion', ascending=False)

top_3_pobladas = df.nlargest(3, 'Poblacion')
top_3_menos_pobladas = df.nsmallest(3, 'Poblacion')

promedio_pob_zona = df.groupby('Costera')['Poblacion'].mean()

conteo_zonas = df['Costera'].value_counts()

df['Densidad'] = df.apply(lambda fila: fila['Poblacion'] / fila['Altitud (msnm)'], axis=1)

stats_densidad = df.agg({'Densidad': ['max', 'min']})

print(df_ordenado)
print(top_3_pobladas[['Ciudad', 'Poblacion']])
print(top_3_menos_pobladas[['Ciudad', 'Poblacion']])
print(promedio_pob_zona)
print(conteo_zonas)
print(df[['Ciudad', 'Densidad']])
print(stats_densidad)
```

```

...   Codigo      Ciudad Poblacion Altitud (msnm) Temperatura Media (°C) \
1     GYE        Guayaquil    2800000             4                  28
0     UIO        Quito       1700000            2850                15
2     CUE        Cuenca      4000000            2560                17
3     ATF        Ambato      3500000            2575                16
8     MCH        Machala     3000000             6                  27
4     MEC        Manta       2500000             6                  27
7     PVO        Portoviejo  2200000            53                  25
5     LOH        Loja        2000000            2100                16
9     IBB        Ibarra      1800000            2225                18
6     ESM        Esmeraldas 1600000            15                  26

Costera
1   True
0   False
2  False
3  False
8   True
4   True
7   True
5  False
9  False
6   True

    ... ..
         Ciudad Poblacion
*** 1  Guayaquil    2800000
0    Quito       1700000
2    Cuenca      4000000
         Ciudad Poblacion
6  Esmeraldas    160000
9    Ibarra      1800000
5    Loja        2000000
Costera
False  566000.0
True   746000.0
Name: Poblacion, dtype: float64
Costera
False  5
True   5
Name: count, dtype: int64
         Ciudad      Densidad
0    Quito      596.491228
1  Guayaquil  700000.000000
2    Cuenca     156.250000
3    Ambato     135.922330
4    Manta      41666.666667
5    Loja       95.238095
6  Esmeraldas  10666.666667
7  Portoviejo  4150.943396
8    Machala    50000.000000
9    Ibarra     80.898876
         Densidad
max  700000.000000
min   80.898876

```

Parte 4: Guardar Resultados Finales

1. Guardar el archivo en formato csv con el nombre: "ciudades_ecuador_procesadas.csv". Use como separador ";"

 #SOLUCION

```
df.to_csv('ciudades_ecuador_procesadas.csv', sep=';', index=False)
```