

```
import pandas as pd
import os

from google.colab import files
uploaded = files.upload()
df = pd.read_csv('data_census.csv', header='infer')
df.head(5)
```



Choose Files data\_census.csv

- **data\_census.csv**(text/csv) - 12112 bytes, last modified: 11/25/2024 - 100% done
- Saving data\_census.csv to data\_census.csv

	Province Code	City Code	Population	Households	Person's Household	Gender Ratio
0	109	1001	10,078,850	4,197,478	2.40	0.97
1	109	1188	155,695	72,882	2.14	0.98
2	109	1189	126,817	59,614	2.13	1.00
3	109	1158	235,186	108,138	2.17	0.94
4	109	1105	298,145	126,915	2.35	0.99
...	...	...	...	...	...	...
276	103	1011	63,069	27,724	2.27	0.94
277	103	1233	48,856	24,193	2.02	0.91
278	115	1001	615,250	251,478	2.45	1.01
279	115	1186	454,673	182,709	2.49	1.00
280	115	1096	160,577	68,769	2.34	1.02



Next steps:

[Generate code with df](#)



[View recommended plots](#)

[New interactive sheet](#)

## ✓ 1) Poblacion total con ProvinceCode igual a 115 o 116

```
population_115_116 = df[df['Province Code'].isin([115, 116])]['Population'].replace(',', '', regex=True)
print(f"Población total con ProvinceCode igual a 115 o 116: {population_115_116}")
```



Población total con ProvinceCode igual a 115 o 116: 5970327

## ✓ 2) Población promedio de las ciudades donde hay más hombres (GenderRatio > 1)

```
poblation_men = df[df['Gender Ratio'] > 1]['Population'].replace(',', '', regex=True).astype(int)
print(f"Población promedio de las ciudades donde hay más hombres {poblation_men}")
```

⇒ Población promedio de las ciudades donde hay más hombres (GenderRatio > 1): 419221.207692307



### 3) Lugares con más hombres (GenderRatio > 1) y menos de 2 personas por hogar

```
places_lessmen_2house = df[(df['Gender Ratio'] > 1) & (df["Person's Household"] < 2)]
print(places_lessmen_2house)
```

⇒

	Province	Code	City	Code	Population	Households	Person's Household	\
62		112		1155	20,825	11,172		1.86
149		101		1240	26,790	13,433		1.99
152		101		1017	30,066	15,485		1.94
167		117		1033	38,027	19,145		1.99
227		113		1120	43,566	22,248		1.96
253		104		1163	10,304	5,365		1.92

  

	Gender Ratio
62	1.29
149	1.27
152	1.15
167	1.03
227	1.11
253	1.15

### 4) Ordena el DataFrame en orden ascendente de 'Households' y muestra los 10 primeros

```
df_sorted_households = df.sort_values(by='Households').head(10)
print("Los 10 primeros lugares ordenados por 'Households':")
print(df_sorted_households)
```

⇒ Los 10 primeros lugares ordenados por 'Households':

	Province	Code	City	Code	Population	Households	Person's Household	\
52		112		1001	2,914,271	1,145,232		2.54
228		104		1001	2,697,791	1,160,150		2.33
254		103		1001	3,356,540	1,356,430		2.47
26		108		1001	3,517,491	1,430,441		2.46
200		114		1179	23,308	10,877		2.14
266		103		1010	252,823	100,073		2.53
256		103		1201	258,719	100,745		2.57
37		108		1042	248,177	101,090		2.46
70		107		1064	242,758	101,201		2.40
230		104		1226	246,634	101,508		2.43

  

	Gender Ratio
52	1.01
228	1.01
254	1.02
26	0.98
200	0.98

266	1.12
256	1.04
37	0.97
70	1.02
230	1.06