```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns


df=pd.read_excel('/content/North&SouthSteeles.xlsx',sheet_name='South')
df.head()
```

	Dissemination area	Total private dwellings	Total - Age groups of the population - 100% data	0 to 4 years	5 to 9 years	10 to 14 years	15 to 19 years	20 to 24 years	25 to 29 years	30 to 34 years	...	30 minut
0	35204116.0	569.0	855.0	25.0	35.0	20.0	25.0	35.0	35.0	30.0	...	8
1	35204960.0	1000.0	1850.0	60.0	45.0	45.0	45.0	50.0	70.0	95.0	...	10
2	35204970.0	244.0	680.0	35.0	15.0	20.0	20.0	50.0	45.0	45.0	...	3
3	35202535.0	164.0	460.0	15.0	20.0	20.0	20.0	30.0	35.0	20.0	...	2
4	35202591.0	312.0	940.0	55.0	45.0	50.0	65.0	70.0	75.0	75.0	...	9

5 rows × 56 columns

```
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 54 entries, 0 to 53
Data columns (total 56 columns):
 #   Column
---  ------
 0   Dissemination area
 1   Total private dwellings
 2   Total - Age groups of the population - 100% data
 3   0 to 4 years
 4   5 to 9 years
 5   10 to 14 years
 6   15 to 19 years
 7   20 to 24 years
 8   25 to 29 years
 9   30 to 34 years
 10  35 to 39 years
```

```
40 to 44 years
45 to 49 years
50 to 54 years
55 to 59 years
Average age of the population
Median age of the population
Total - Census families in private households by family size - 100% data
2 persons
3 persons
4 persons
5 or more persons
Average size of census families
Average number of children in census families with children
Total - Persons not in census families in private households - 100% data
Living alone
Total - Household type - 100% data
Couple-family households
With children
Without children
Median total income of couple-with-children economic families  in 2020 ($)
Median after-tax income of couple-with-children economic families in 2020 ($)
Average family size of couple-with-children economic families
Average total income of couple-with-children economic families in 2020 ($)
Average after-tax income of couple-with-children economic families in 2020 ($)
Participation rate
Employment rate
Unemployment rate
Total - Place of work status for the employed labour force aged 15 years and over - 25% sam
Worked at home
No fixed workplace address
Usual place of work
Car, truck or van
Public transit
Less than 15 minutes
15 to 29 minutes
30 to 44 minutes
45 to 59 minutes
60 minutes and over
Total - Time leaving for work for the employed labour force aged 15 years and over with a u
Between 5 a.m. and 5:59 a.m.
Between 6 a.m. and 6:59 a.m.
```

```
print(df.describe())
```

```
count                                        52.000000
mean                                        282.788462
std                                         183.025803
min                                          65.000000
25%                                         151.250000
50%                                         195.000000
75%                                         392.500000
max                                         770.000000

       Between 5 a.m. and 5:59 a.m.  Between 6 a.m. and 6:59 a.m.  \
count                     52.000000                     52.000000
mean                      14.423077                     43.653846
std                       15.231992                     38.870264
min                        0.000000                      0.000000
25%                        0.000000                     15.000000
50%                       10.000000                     30.000000
75%                       25.000000                     61.250000
max                       70.000000                    150.000000

       Between 7 a.m. and 7:59 a.m.  Between 8 a.m. and 8:59 a.m.  \
count                     52.000000                     52.000000
mean                      55.961538                     58.365385
std                       39.170798                     40.665073
min                       10.000000                      0.000000
25%                       25.000000                     30.000000
50%                       42.500000                     45.000000
75%                       75.000000                     85.000000
max                      165.000000                    180.000000

       Between 9 a.m. and 11:59 a.m.  Between 12 p.m. and 4:59 a.m.
count                      52.000000                      52.000000
mean                       53.365385                      54.519231
std                        40.447511                      45.667808
min                         0.000000                       0.000000
25%                        25.000000                      23.750000
50%                        40.000000                      40.000000
75%                        80.000000                      76.250000
max                       190.000000                     195.000000

[8 rows x 56 columns]
```

```
df.isnull().sum()
```

	0
Dissemination area	1
Total private dwellings	1
Total - Age groups of the population - 100% data	2
0 to 4 years	2
5 to 9 years	2
10 to 14 years	2
15 to 19 years	2
20 to 24 years	2
25 to 29 years	2
30 to 34 years	2
35 to 39 years	2
40 to 44 years	2
45 to 49 years	2
50 to 54 years	2
55 to 59 years	2
Average age of the population	2
Median age of the population	2
Total - Census families in private households by family size - 100% data	2
2 persons	2
3 persons	2
4 persons	2
5 or more persons	2
Average size of census families	2
Average number of children in census families with children	2
Total - Persons not in census families in private households - 100% data	2
Living alone	2
Total - Household type - 100% data	2
Couple-family households	2
With children	2
Without children	2
Median total income of couple-with-children economic families in 2020 ($)	2
Median after-tax income of couple-with-children economic families in 2020 ($)	1
Average family size of couple-with-children economic families	2
Average total income of couple-with-children economic families in 2020 ($)	2

Average total income of couple-with-children economic families in 2020 ($)	2
Average after-tax income of couple-with-children economic families in 2020 ($)	2
Participation rate	2
Employment rate	2
Unemployment rate	2
Total - Place of work status for the employed labour force aged 15 years and over - 25% sample data	2
Worked at home	2
No fixed workplace address	2
Usual place of work	2
Car, truck or van	2
Public transit	2
Less than 15 minutes	2
15 to 29 minutes	2
30 to 44 minutes	2
45 to 59 minutes	2
60 minutes and over	2
Total - Time leaving for work for the employed labour force aged 15 years and over with a usual place of work or no fixed workplace address - 25% sample data	2
Between 5 a.m. and 5:59 a.m.	2
Between 6 a.m. and 6:59 a.m.	2
Between 7 a.m. and 7:59 a.m.	2
Between 8 a.m. and 8:59 a.m.	2
Between 9 a.m. and 11:59 a.m.	2
Between 12 p.m. and 4:59 a.m.	2

dtype: int64

```
df.columns
```

Show hidden output

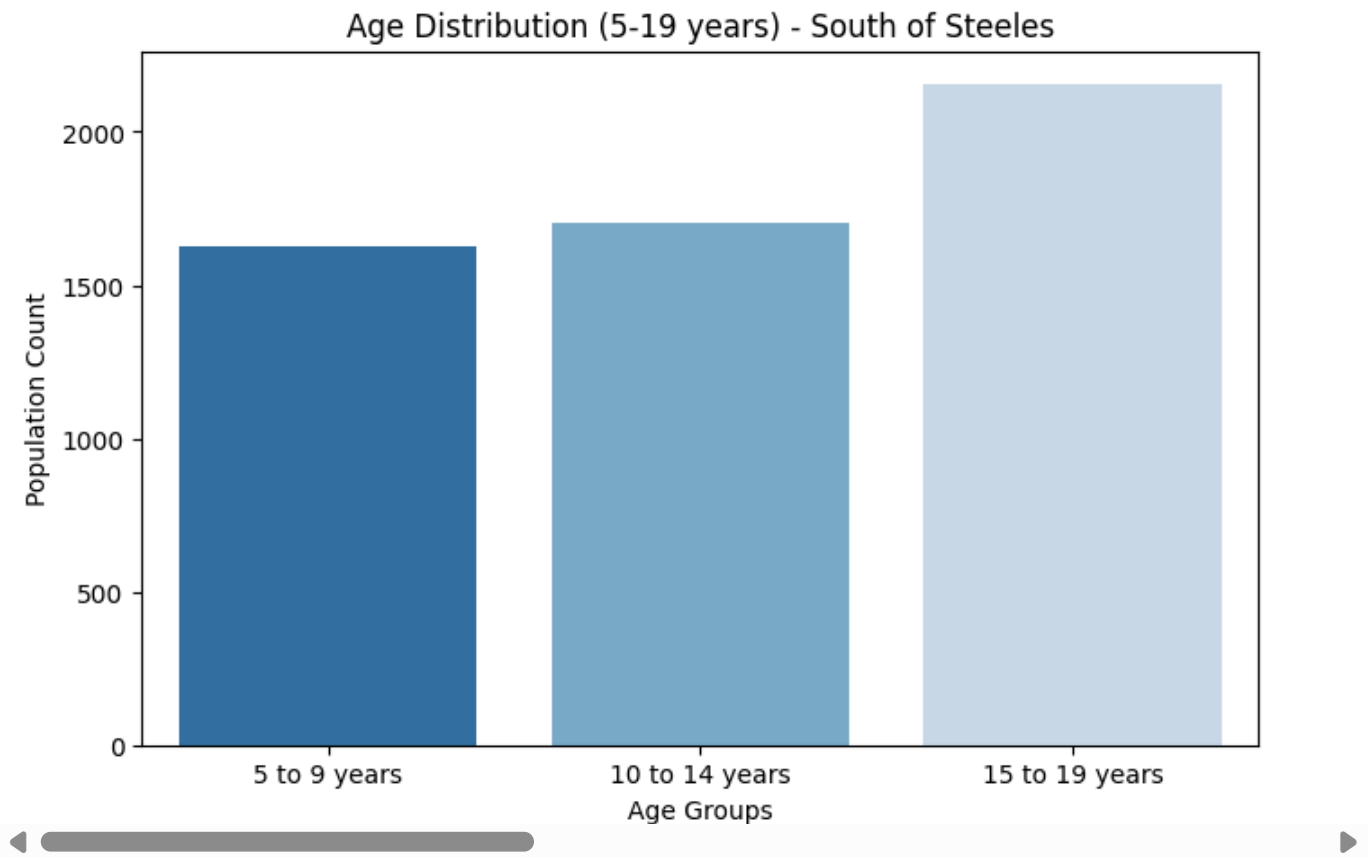
```
df.columns = df.columns.str.strip()
```

```
df.columns.tolist()
```

Show hidden output

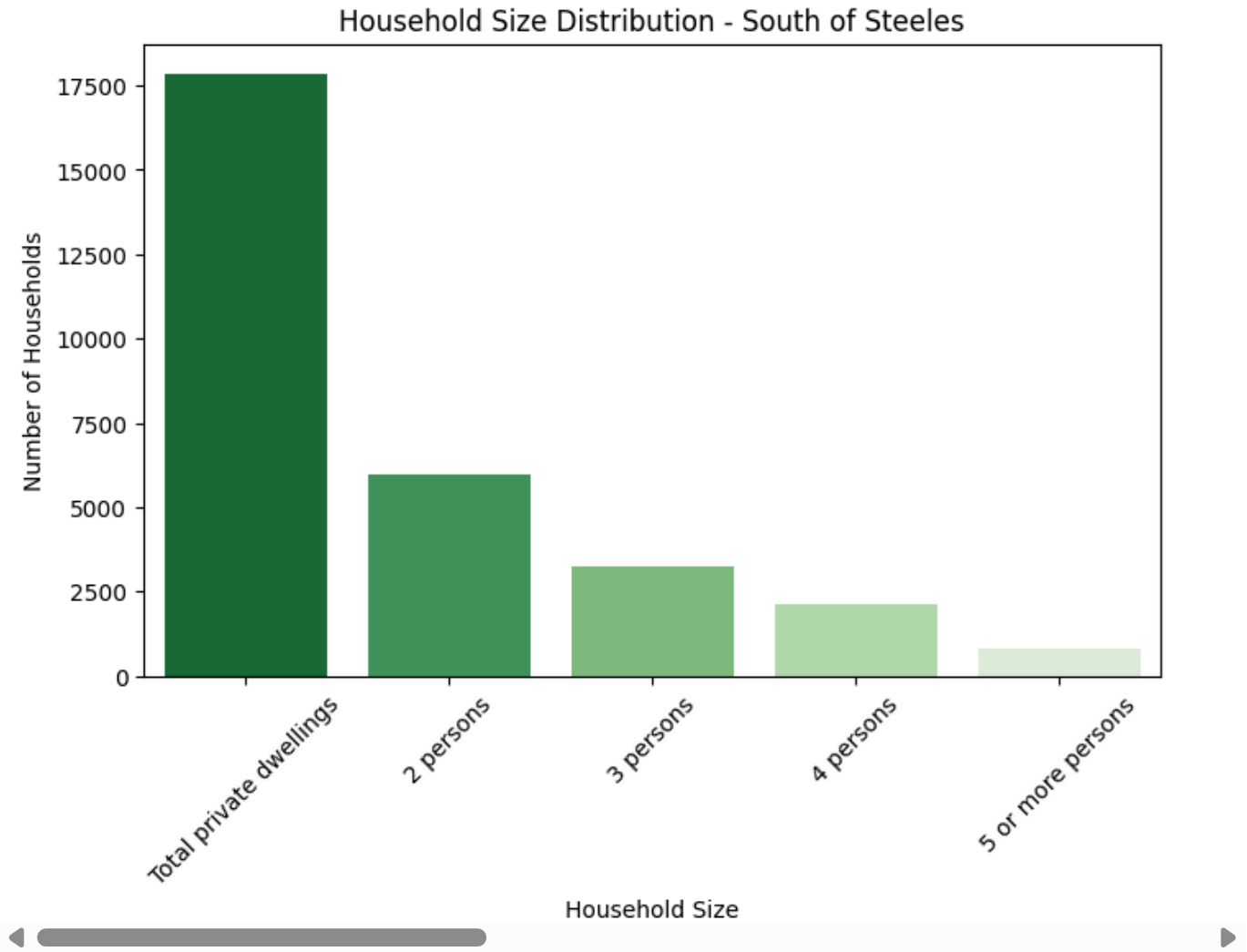
```
age_columns = [
    "5 to 9 years", "10 to 14 years", "15 to 19 years"
]

age_distribution_south = df[age_columns].sum()

plt.figure(figsize=(8, 5))
sns.barplot(x=age_distribution_south.index, y=age_distribution_south.values, palette="Blues_r")
plt.xlabel("Age Groups")
plt.ylabel("Population Count")
plt.title("Age Distribution (5-19 years) - South of Steeles")
plt.show()
```

```
<ipython-input-12-7751745341fb>:10: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the

  sns.barplot(x=age_distribution_south.index, y=age_distribution_south.values, palette="Blues_r")
```



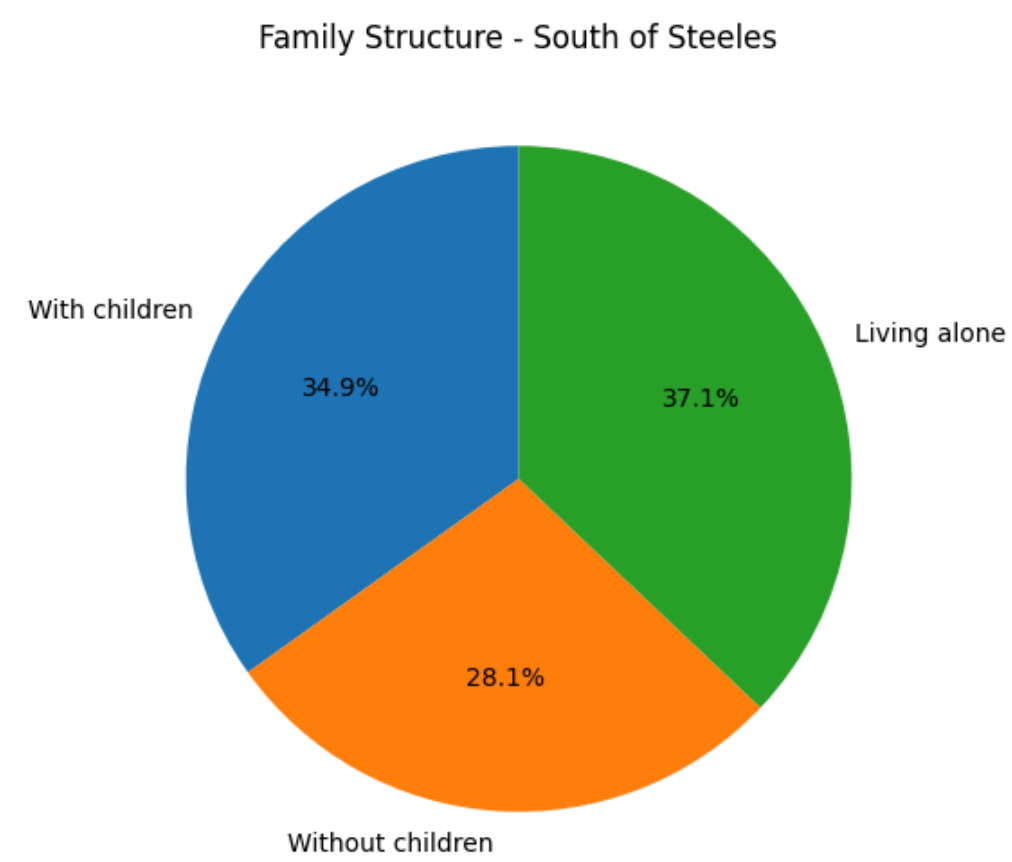
```
# Define columns related to household and family size
household_columns = ["Total private dwellings", "2 persons", "3 persons", "4 persons", "5 or more pers

# Aggregate household size data
household_distribution = df[household_columns].sum()

# Plot household size distribution
plt.figure(figsize=(8, 5))
sns.barplot(x=household_distribution.index, y=household_distribution.values, palette="Greens_r")
plt.xlabel("Household Size")
plt.ylabel("Number of Households")
plt.title("Household Size Distribution - South of Steeles")
plt.xticks(rotation=45)
plt.show()
```

```
<ipython-input-13-7ac075f55cd6>:9: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the

  sns.barplot(x=household_distribution.index, y=household_distribution.values, palette="Greens_r")
```

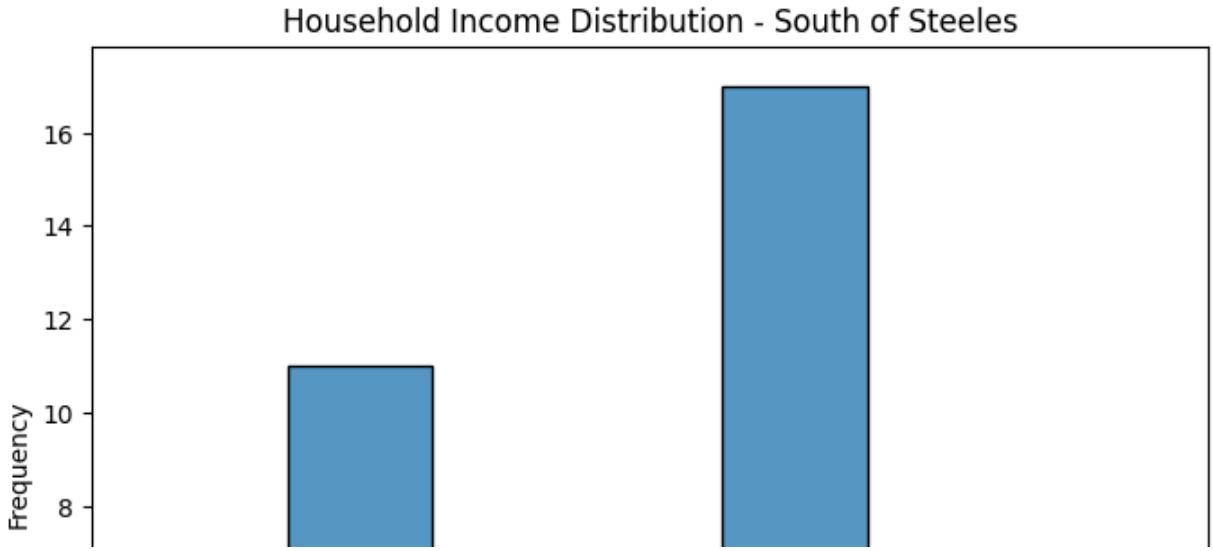



```
import matplotlib.pyplot as plt

family_types = ['With children', 'Without children', 'Living alone']
family_counts = [df['With children'].sum(), df['Without children'].sum(), df['Living alone'].sum()]
```

```
plt.figure(figsize=(6, 6))
plt.pie(family_counts, labels=family_types, autopct='%1.1f%%', startangle=90)
plt.title('Family Structure - South of Steeles')
plt.show()
```




```
import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(8, 6))
sns.histplot(data=df, x='Median total income of couple-with-children economic families  in 2020 ($)')
plt.xlabel('Household Income ($)')
plt.ylabel('Frequency')
plt.title('Household Income Distribution - South of Steeles')
plt.show()
```




```
import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(8, 6))
sns.scatterplot(data=df, x='Median total income of couple-with-children economic families  in 2020 ($)
plt.xlabel('Household Income ($)')
plt.ylabel('Family Size')
plt.title('Household Income and Family Size - South of Steeles')
plt.show()
```