```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np


#read the excel file
xls=pd.ExcelFile('North&SouthSteeles.xlsx')


df_south = pd.read_excel(xls, sheet_name="South")
df_north = pd.read_excel(xls, sheet_name="North")


# Display column names to help confirm the structure
print("South columns:", df_south.columns.tolist())
print("North columns:", df_north.columns.tolist())
```

```
, '    0 to 4 years', '    5 to 9 years', '    10 to 14 years', '    15 to 19 years', '    20 to 24 years', '    25 to 29 years', '    3
 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
```

```
# Assuming key columns include:
# - "Total private dwellings"
# - "Total - Age groups of the population - 100% data"
# - Age group columns like "5 to 9 years", "10 to 14 years", "15 to 19 years"

# Aggregate key metrics for each region

# Total population
south_total_population = df_south["Total - Age groups of the population - 100% data"].sum()
north_total_population = df_north["Total - Age groups of the population - 100% data"].sum()


# Total number of dwellings
south_total_dwellings = df_south["Total private dwellings"].sum()
north_total_dwellings = df_north["Total private dwellings"].sum()


# Estimate school-age population (using age groups: 5 to 9, 10 to 14, 15 to 19)
south_school_age = df_south[["    5 to 9 years", "    10 to 14 years", "    15 to 19 years"]].sum().sum()
north_school_age = df_north[["    5 to 9 years", "    10 to 14 years", "    15 to 19 years"]].sum().sum()


# Calculate school-age ratio (school-age children divided by total population)
south_school_ratio = south_school_age / south_total_population
north_school_ratio = north_school_age / north_total_population


print("South - Total Population:", south_total_population)
print("North - Total Population:", north_total_population)
print("South - Total Dwellings:", south_total_dwellings)
print("North - Total Dwellings:", north_total_dwellings)
print("South - School Age Population:", south_school_age)
print("North - School Age Population:", north_school_age)
print("South - School Age Ratio: {:.1%}".format(south_school_ratio))
print("North - School Age Ratio: {:.1%}".format(north_school_ratio))
```

```
South - Total Population: 44135.0
North - Total Population: 37725.0
South - Total Dwellings: 17846.0
North - Total Dwellings: 13625.0
South - School Age Population: 5485.0
North - School Age Population: 5985.0
South - School Age Ratio: 12.4%
North - School Age Ratio: 15.9%
```

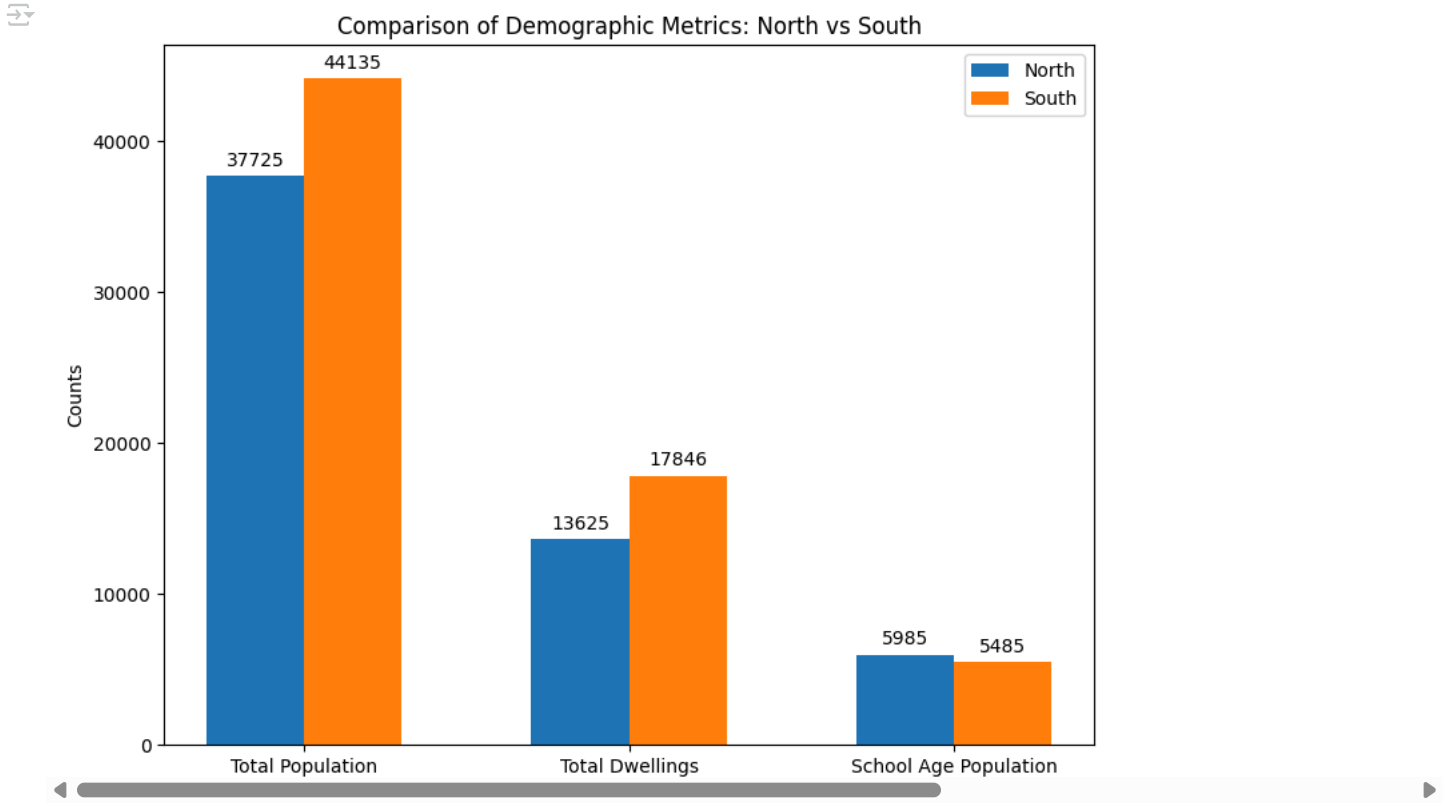
```
# Create a bar chart to compare the regions for total population, dwellings, and school-age population
categories = ["Total Population", "Total Dwellings", "School Age Population"]
north_values = [north_total_population, north_total_dwellings, north_school_age]
south_values = [south_total_population, south_total_dwellings, south_school_age]

x = np.arange(len(categories))
width = 0.3

fig, ax = plt.subplots(figsize=(8, 6))
rects1 = ax.bar(x - width/2, north_values, width, label="North")
rects2 = ax.bar(x + width/2, south_values, width, label="South")
```

```
ax.set_ylabel("Counts")
ax.set_title("Comparison of Demographic Metrics: North vs South")
ax.set_xticks(x)
ax.set_xticklabels(categories)
ax.legend()
# Attach value labels on bars for clarity
def autolabel(rects):
    for rect in rects:
        height = rect.get_height()
        ax.annotate(f'{int(height)}',
                    xy=(rect.get_x() + rect.get_width()/2, height),
                    xytext=(0, 3),  # 3 points vertical offset
                    textcoords="offset points",
                    ha='center', va='bottom')

autolabel(rects1)
autolabel(rects2)
plt.tight_layout()
plt.show()
```



South has a higher number of dwellings but fewer students, it suggests that:

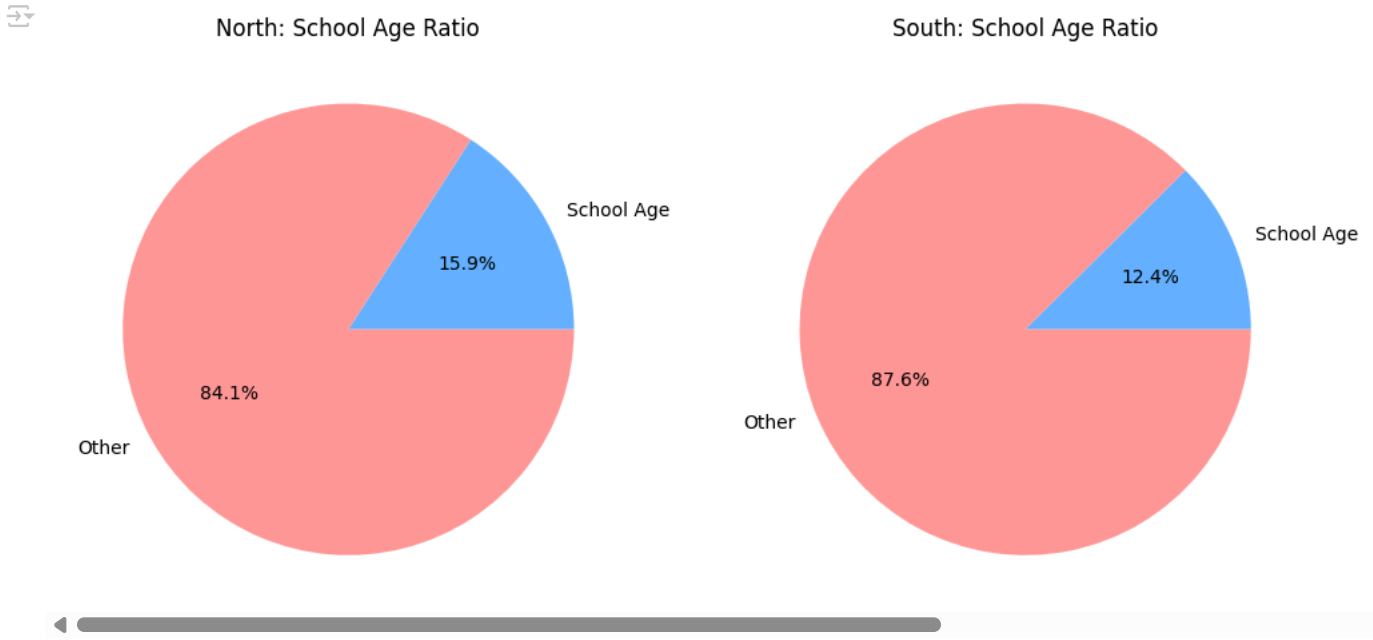
- The area may have a higher number of small households (singles, couples, or elderly) rather than families with school-age children.
- The population might be more transient, with fewer long-term residents investing in education services.
- More rental properties could indicate young professionals rather than families with school-aged children.

```
# Attach value labels on bars for clarity
def autolabel(rects):
    for rect in rects:
        height = rect.get_height()
        ax.annotate(f'{int(height)}',
                    xy=(rect.get_x() + rect.get_width()/2, height),
                    xytext=(0, 3),  # 3 points vertical offset
                    textcoords="offset points",
                    ha='center', va='bottom')

autolabel(rects1)
autolabel(rects2)
plt.tight_layout()
plt.show()
```

<Figure size 640x480 with 0 Axes>

```
# Pie charts for school-age ratio for each region
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(12, 6))
ax1.pie([north_school_ratio, 1 - north_school_ratio], labels=["School Age", "Other"], autopct="%1.1f%%", colors=["#66b3ff", "#ff9999"])
ax1.set_title("North: School Age Ratio")
ax2.pie([south_school_ratio, 1 - south_school_ratio], labels=["School Age", "Other"], autopct="%1.1f%%", colors=["#66b3ff", "#ff9999"])
ax2.set_title("South: School Age Ratio")
plt.show()
```



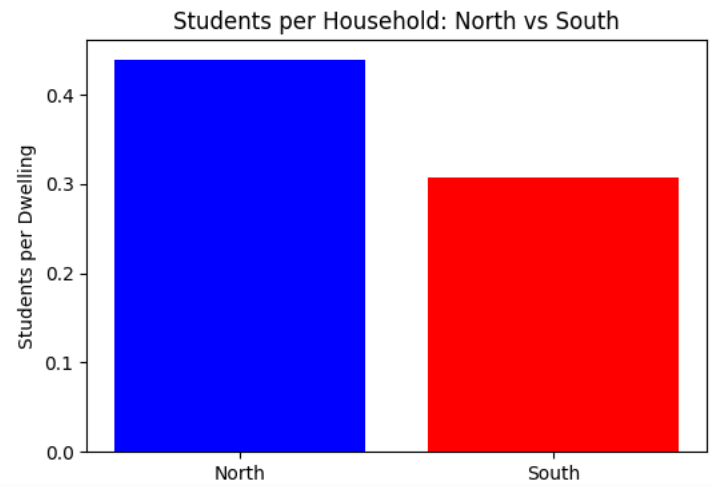
Population and Housing: By comparing the total population and number of private dwellings, you can determine if the South has a lower overall population or lower density in the areas of interest. If the South shows significantly lower numbers, it might simply have fewer households that could convert into tuition students.

School-Age Demographics: The sum of the age groups “5 to 9 years”, “10 to 14 years”, and “15 to 19 years” provides an estimate of the school-age population. A lower school-age ratio in the South relative to the North would indicate a smaller target audience for tuition services. However, if the ratio is similar but the absolute numbers are lower, the franchise might still be missing opportunities.

Visual Comparisons: The bar chart helps you see at a glance the differences in total population, dwellings, and school-age population between the two regions. Meanwhile, the pie charts highlight the relative proportions of school-age children in each region. These visualizations can guide where additional marketing efforts might yield a higher return.

```
# 3. Students per dwelling ratio
# Students per dwelling ratio
south_students_per_dwelling = south_school_age / south_total_dwellings
north_students_per_dwelling = north_school_age / north_total_dwellings

fig, ax = plt.subplots(figsize=(6, 4))
ax.bar(["North", "South"], [north_students_per_dwelling, south_students_per_dwelling], color=["blue", "red"])
ax.set_ylabel("Students per Dwelling")
ax.set_title("Students per Household: North vs South")
plt.show()
```




```
# ----- Additional Analysis: Age Group Breakdown -----
age_breakdown = ["    0 to 4 years", "    5 to 9 years", "    10 to 14 years", "    15 to 19 years"]


# Sum the age groups for each region
age_groups = ["    0 to 4 years","    5 to 9 years", "    10 to 14 years", "    15 to 19 years"]
south_school_age = df_south[age_groups].sum().sum()
north_school_age = df_north[age_groups].sum().sum()
south_age_totals = df_south[age_breakdown].sum()
north_age_totals = df_north[age_breakdown].sum()


# Create a DataFrame for easier comparison
age_df = pd.DataFrame({
    "Age Group": age_breakdown,
    "North": north_age_totals.values,
    "South": south_age_totals.values
})
print("\n----- Age Group Breakdown -----")
print(age_df)
```

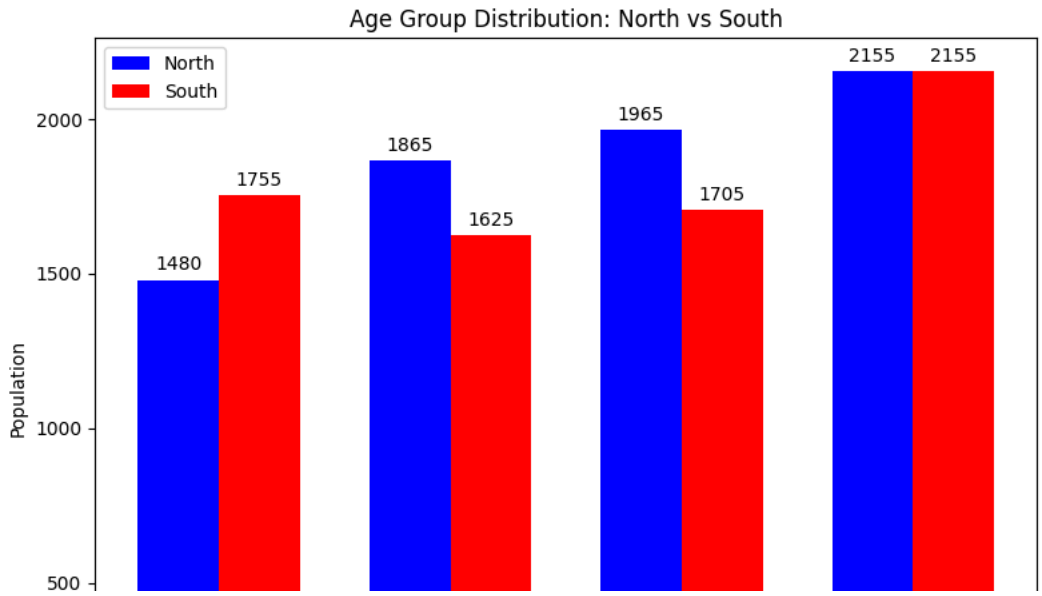
```
----- Age Group Breakdown -----
            Age Group   North   South
0        0 to 4 years  1480.0  1755.0
1        5 to 9 years  1865.0  1625.0
2      10 to 14 years  1965.0  1705.0
3      15 to 19 years  2155.0  2155.0
```

```
# Bar chart for age group breakdown
x = np.arange(len(age_breakdown))
width = 0.35

fig, ax = plt.subplots(figsize=(8, 6))
rects1 = ax.bar(x - width/2, north_age_totals, width, label="North", color="blue")
rects2 = ax.bar(x + width/2, south_age_totals, width, label="South", color="red")

ax.set_ylabel("Population")
ax.set_title("Age Group Distribution: North vs South")
ax.set_xticks(x)
ax.set_xticklabels(age_breakdown)
ax.legend()

autolabel(rects1)
autolabel(rects2)
plt.tight_layout()
plt.show()
```




```
# ----- Insight Summary -----
print("\n----- Insight Summary -----")
print("1. South has more total dwellings, suggesting more households but lower students per dwelling.")
print("2. The school-age ratio in the South is lower than in the North, which could indicate a smaller target audience for tuition.")
print("3. The age breakdown shows that the South may have fewer children in key age groups (5-19 years) relative to the North.")
print("4. This suggests that rather than just increasing advertising spend, the franchisee in the South might benefit from targeted outreach

plt.show()
```

```
----- Insight Summary -----
1. South has more total dwellings, suggesting more households but lower students per dwelling.
2. The school-age ratio in the South is lower than in the North, which could indicate a smaller target audience for tuition.
3. The age breakdown shows that the South may have fewer children in key age groups (5-19 years) relative to the North.
4. This suggests that rather than just increasing advertising spend, the franchisee in the South might benefit from targeted outreach to
```

Start coding or generate with AI.

North does not need aggressive marketing—it already has demand. South needs smarter, niche marketing—focus on fewer but high-intent parents.

Community partnerships, school sponsorships, and digital targeting are key. **bold text**