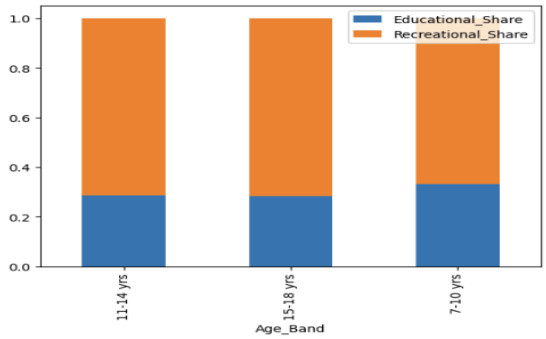
# ScreenSense:Kids' Screentime Visualization
# WEEK 3 & WEEK 4

## Univariate and Bivariate Visual Analysis:

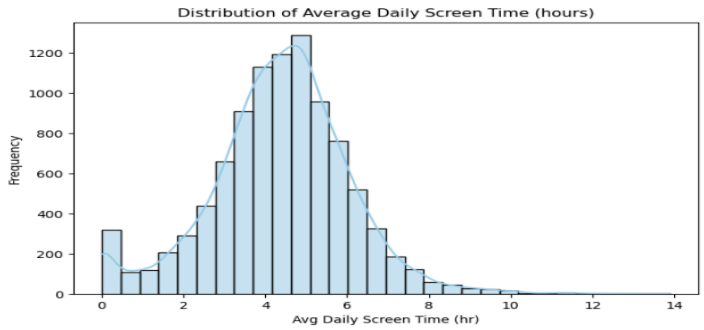
The univariate analysis shows that younger children spend a larger share of time on educational activities, while older ones focus more on recreation. Most kids have moderate daily screen time, though some exceed the recommended limit. The boxplot highlights this variation and outliers with unusually high screen usage.

```
[43]: df.groupby('Age_Band')[['Educational_Share', 'Recreational_Share']].mean().plot(kind='bar', stacked=True)
```

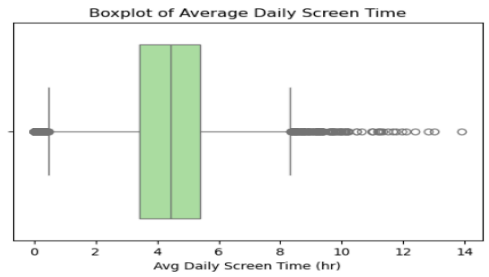
```
[43]: <Axes: xlabel='Age_Band'>
```




```
[17]: # Distribution of Avg_Daily_Screen_Time_hr
plt.figure(figsize=(8,5))
sns.histplot(df['Avg_Daily_Screen_Time_hr'], bins=30, kde=True, color='skyblue')
plt.title("Distribution of Average Daily Screen Time (hours)")
plt.xlabel("Avg Daily Screen Time (hr)")
plt.ylabel("Frequency")
plt.show()
```




```
[18]: # Boxplot
plt.figure(figsize=(6,4))
sns.boxplot(x=df['Avg_Daily_Screen_Time_hr'], color='lightgreen')
plt.title("Boxplot of Average Daily Screen Time")
plt.xlabel("Avg Daily Screen Time (hr)")
plt.show()
```

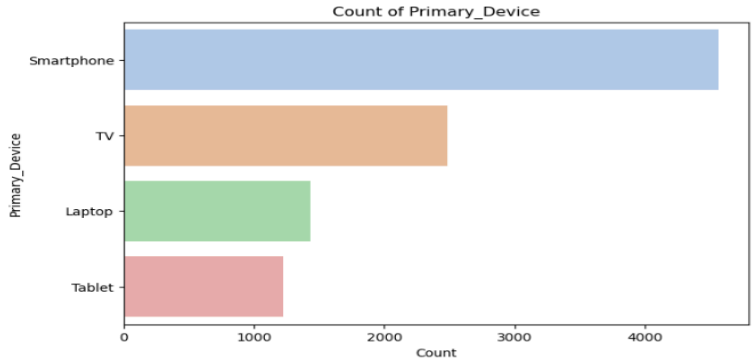


These plots show the **distribution of categorical features**:

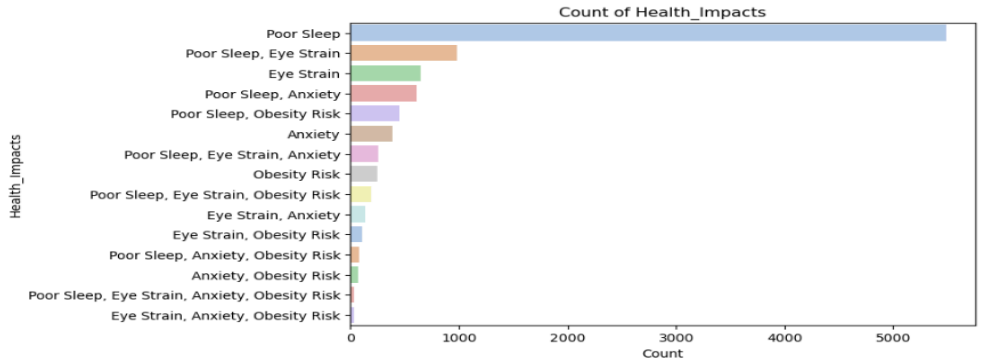
- **Primary_Device:** Identifies the most commonly used screen device among kids.
- **Health_Impacts:** Displays which health issues are reported most often due to screen use.
- **Urban_or_Rural:** Shows how many participants belong to each location type.

```
[19]: for col in ['Primary_Device', 'Health_Impacts', 'Urban_or_Rural']:
          plt.figure(figsize=(8,5))
          sns.countplot(y=df[col], order=df[col].value_counts().index, palette="pastel")
          plt.title(f"Count of {col}")
          plt.xlabel("Count")
          plt.ylabel(col)
          plt.show()
```

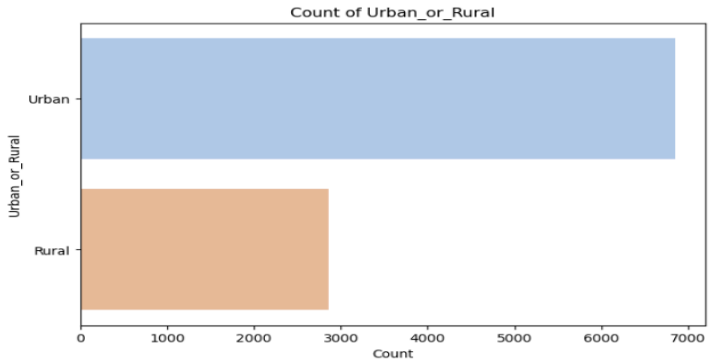
```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/337508708.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.countplot(y=df[col], order=df[col].value_counts().index, palette="pastel")
```




```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/337508708.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.countplot(y=df[col], order=df[col].value_counts().index, palette="pastel")
```



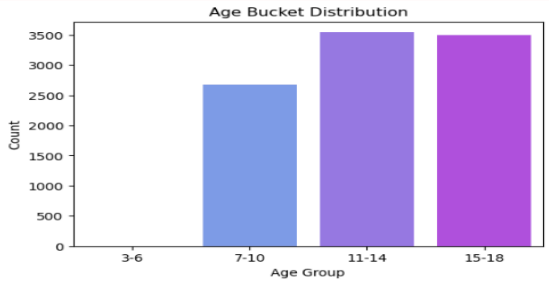

```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/337508708.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.countplot(y=df[col], order=df[col].value_counts().index, palette="pastel")
```



The age bucket analysis shows most participants fall in the 11–14 and 15–18 age groups. The count plot for exceeding screen time limits indicates that a majority of children spend more time on screens than recommended.

[20]:
```
# 🔄 Age Buckets
bins = [0, 6, 10, 14, 18]
labels = ['3-6', '7-10', '11-14', '15-18']
df['Age_Bucket'] = pd.cut(df['Age'], bins=bins, labels=labels, right=True)

plt.figure(figsize=(6,4))
sns.countplot(x=df['Age_Bucket'], palette="cool")
plt.title("Age Bucket Distribution")
plt.xlabel("Age Group")
plt.ylabel("Count")
plt.show()
```

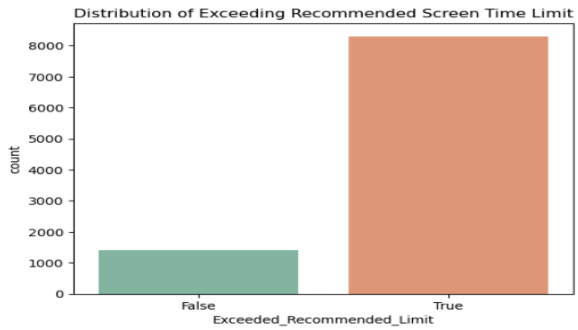
```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/4103013046.py:7: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.countplot(x=df['Age_Bucket'], palette="cool")
```




[21]:
```
sns.countplot(x=df['Exceeded_Recommended_Limit'], palette='Set2')
plt.title("Distribution of Exceeding Recommended Screen Time Limit")
```

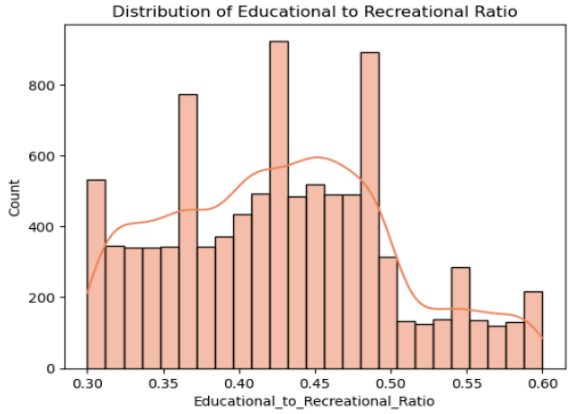
```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/2964855439.py:1: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.countplot(x=df['Exceeded_Recommended_Limit'], palette='Set2')
```

[21]:
```
Text(0.5, 1.0, 'Distribution of Exceeding Recommended Screen Time Limit')
```



This histogram shows how the Educational-to-Recreational Ratio is distributed among kids. Most values cluster below 1, indicating that children spend more time on recreational activities than on educational ones.

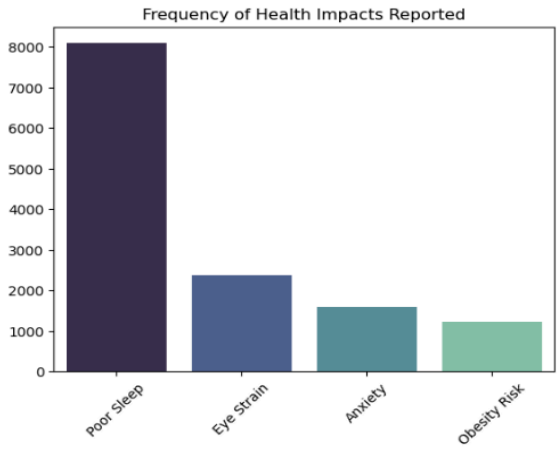
[22]:
```
Text(0.5, 1.0, 'Distribution of Educational to Recreational Ratio')
```



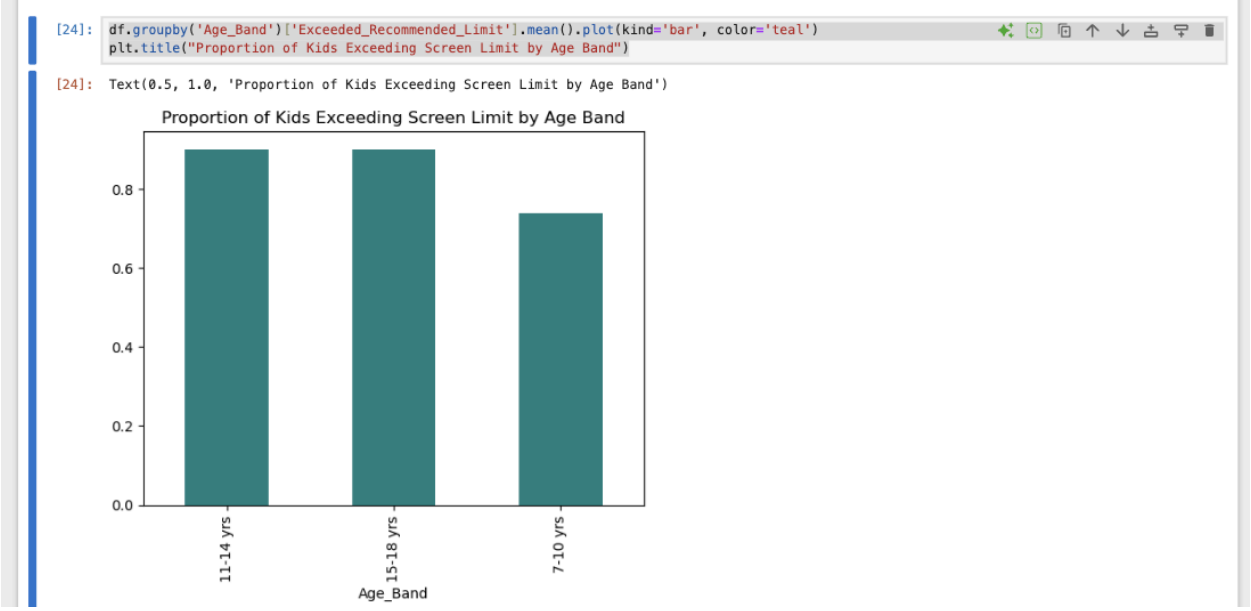

[23]:
```
from collections import Counter
from itertools import chain

health_counts = Counter(chain.from_iterable(df['Health_Impacts'].str.split(', ')))
sns.barplot(x=list(health_counts.keys()), y=list(health_counts.values()), palette='mako')
plt.title("Frequency of Health Impacts Reported")
plt.xticks(rotation=45)
```

```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/593143806.py:5: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.barplot(x=list(health_counts.keys()), y=list(health_counts.values()), palette='mako')
```

[23]:
```
([0, 1, 2, 3],
 [Text(0, 0, 'Poor Sleep'),
  Text(1, 0, 'Eye Strain'),
  Text(2, 0, 'Anxiety'),
  Text(3, 0, 'Obesity Risk')])
```



This bar chart shows the average proportion of kids exceeding the recommended screen time within each age band. It highlights that older age groups (especially teens) are more likely to go beyond healthy screen time limits compared to younger children.

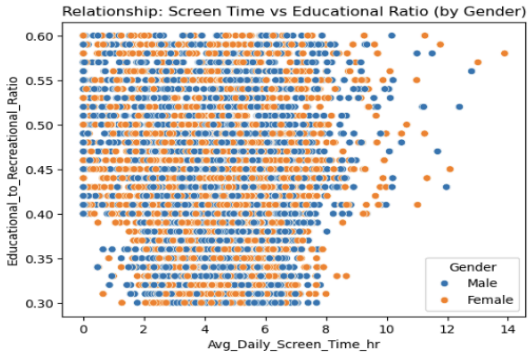



## Bivariate Analysis:-

The bivariate analysis explores relationships between key variables. The scatterplot shows that as average daily screen time increases, the educational-to-recreational ratio generally decreases, with slight gender differences — boys tend to have lower educational ratios than girls. The line plot indicates that average screen time rises steadily with age, suggesting older children spend more hours on screens. The bar chart reveals that smartphones and tablets are linked to the highest average screen time compared to other devices.

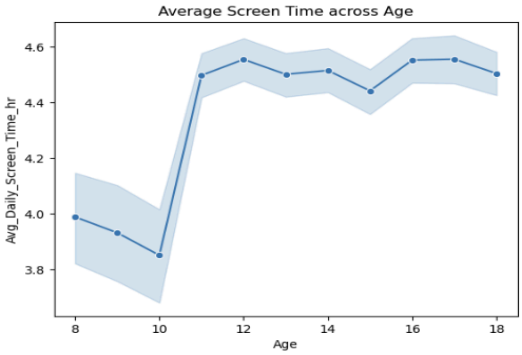
[29]:
```
#bivariate analysis
sns.scatterplot(x='Avg_Daily_Screen_Time_hr', y='Educational_to_Recreational_Ratio', data=df, hue='Gender')
plt.title("Relationship: Screen Time vs Educational Ratio (by Gender)")
```

[29]:
```
Text(0.5, 1.0, 'Relationship: Screen Time vs Educational Ratio (by Gender)')
```




[30]:
```
sns.lineplot(x='Age', y='Avg_Daily_Screen_Time_hr', data=df, marker='o')
plt.title("Average Screen Time across Age")
```

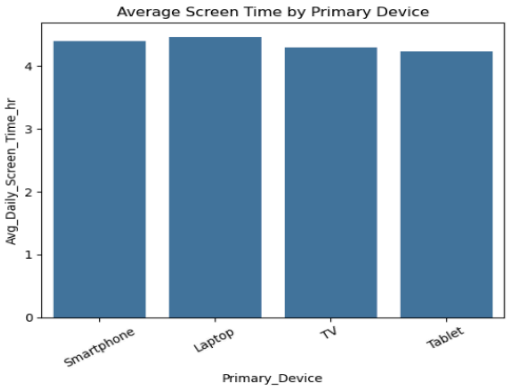
[30]:
```
Text(0.5, 1.0, 'Average Screen Time across Age')
```




[31]:
```
sns.barplot(x='Primary_Device', y='Avg_Daily_Screen_Time_hr', data=df, estimator='mean', ci=None)
plt.title("Average Screen Time by Primary Device")
plt.xticks(rotation=30)
```

```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/867479557.py:1: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.barplot(x='Primary_Device', y='Avg_Daily_Screen_Time_hr', data=df, estimator='mean', ci=None)
```

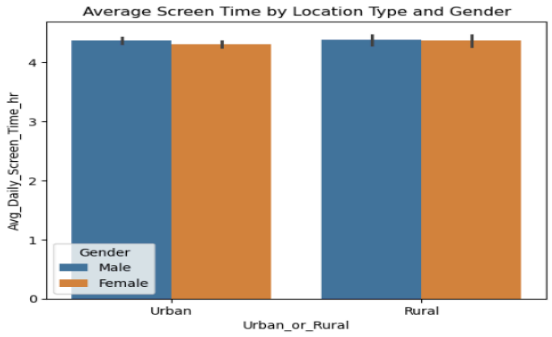
[31]:
```
([0, 1, 2, 3],
 [Text(0, 0, 'Smartphone'),
  Text(1, 0, 'Laptop'),
  Text(2, 0, 'TV'),
  Text(3, 0, 'Tablet')])
```



The bar plot shows that **urban kids generally have higher screen time** than rural ones, with **boys spending slightly more time** than girls. The boxplot reveals that **screen time increases with the number of reported health issues**, suggesting a clear link between excessive screen use and health impacts. The correlation heatmap confirms **a positive relationship between age and screen time**, while the **educational-to-recreational ratio** is slightly negatively correlated with screen time, indicating that higher screen use is mostly for recreation.

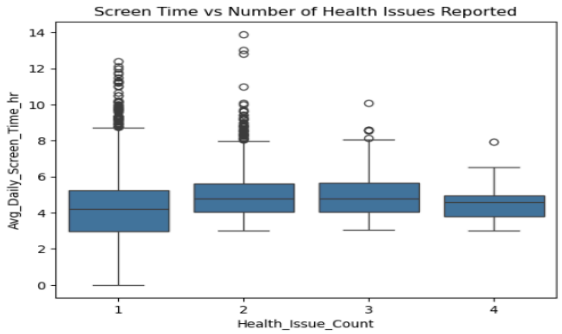
```
[32]: sns.barplot(x='Urban_or_Rural', y='Avg_Daily_Screen_Time_hr', hue='Gender', data=df)
      plt.title("Average Screen Time by Location Type and Gender")
```

```
[32]: Text(0.5, 1.0, 'Average Screen Time by Location Type and Gender')
```



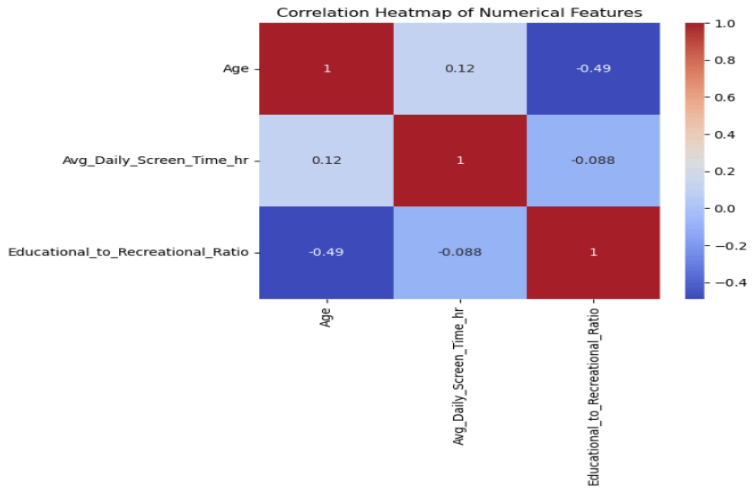

```
[33]: df['Health_Issue_Count'] = df['Health_Impacts'].apply(lambda x: len(x.split(', ')) if x != 'None' else 0)
      sns.boxplot(x='Health_Issue_Count', y='Avg_Daily_Screen_Time_hr', data=df)
      plt.title("Screen Time vs Number of Health Issues Reported")
```

```
[33]: Text(0.5, 1.0, 'Screen Time vs Number of Health Issues Reported')
```




```
[34]: sns.heatmap(df[['Age', 'Avg_Daily_Screen_Time_hr', 'Educational_to_Recreational_Ratio']].corr(),
                  annot=True, cmap='coolwarm')
      plt.title("Correlation Heatmap of Numerical Features")
```

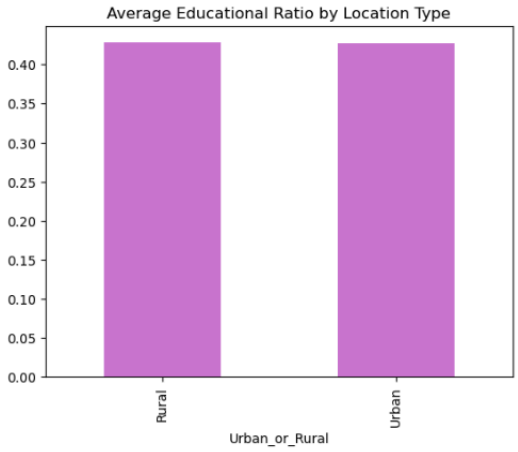
```
[34]: Text(0.5, 1.0, 'Correlation Heatmap of Numerical Features')
```



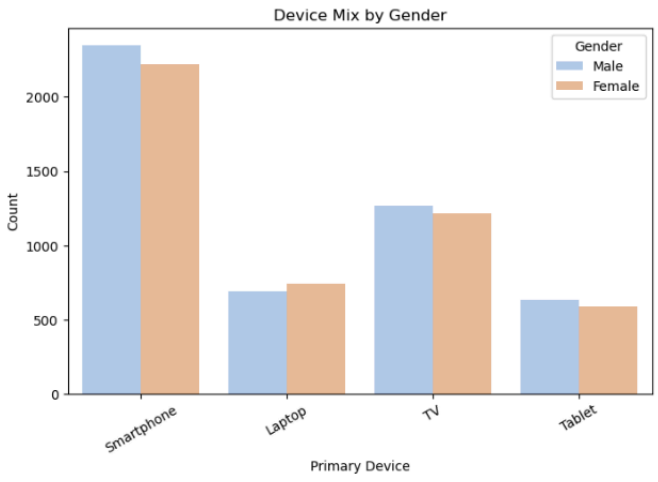
The bar chart shows that urban children have a slightly higher educational-to-recreational ratio than rural ones, suggesting more balanced screen use in urban areas. The device mix plot indicates that smartphones are the most used devices across both genders, with boys showing a slightly higher preference for gaming devices, while girls use tablets and phones more frequently.

```
[35]: df.groupby('Urban_or_Rural')['Educational_to_Recreational_Ratio'].mean().plot(kind='bar', color='orchid')
      plt.title("Average Educational Ratio by Location Type")
```

```
[35]: Text(0.5, 1.0, 'Average Educational Ratio by Location Type')
```

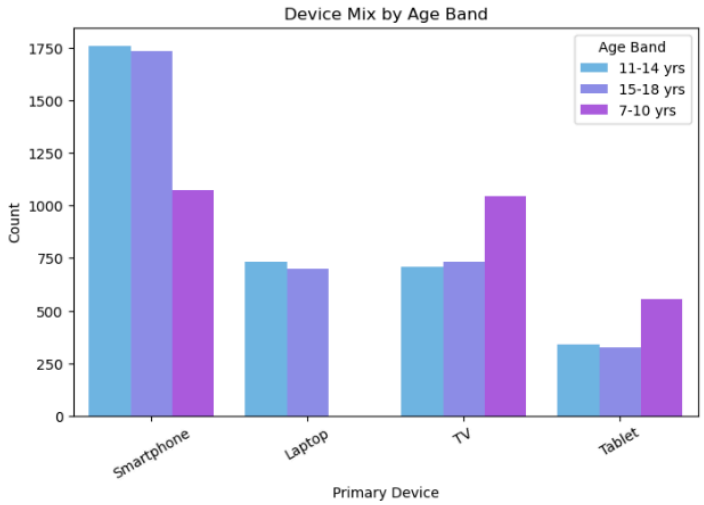



```
[36]: #Device Mix by Gender
      plt.figure(figsize=(8,5))
      sns.countplot(x='Primary_Device', hue='Gender', data=df, palette='pastel')
      plt.title("Device Mix by Gender")
      plt.xlabel("Primary Device")
      plt.ylabel("Count")
      plt.legend(title='Gender')
      plt.xticks(rotation=30)
      plt.show()
```

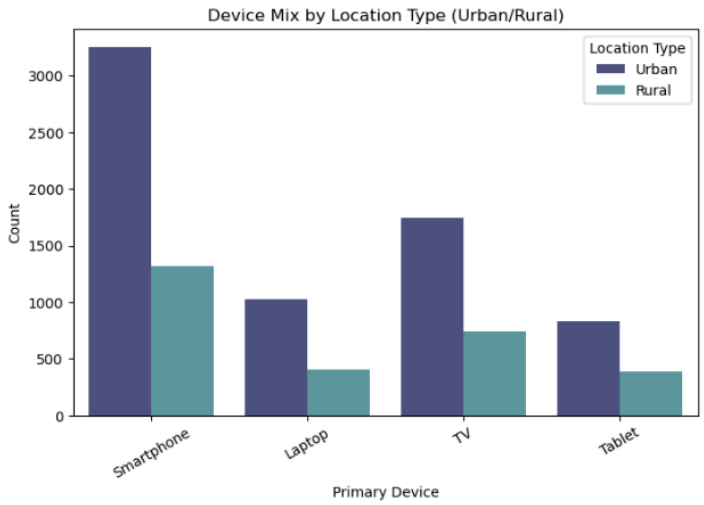


The analysis of device mix by age band shows that younger kids prefer tablets and TVs, while older children increasingly use smartphones as their main device. The urban–rural comparison reveals that urban kids use smartphones and laptops more frequently, whereas rural children rely more on shared or affordable devices like TVs. This highlights how age and location influence device preferences among children.

[37]:
```
#Device Mix by Age Band
plt.figure(figsize=(8,5))
sns.countplot(x='Primary_Device', hue='Age_Band', data=df, palette='cool')
plt.title("Device Mix by Age Band")
plt.xlabel("Primary Device")
plt.ylabel("Count")
plt.legend(title='Age Band')
plt.xticks(rotation=30)
plt.show()
```



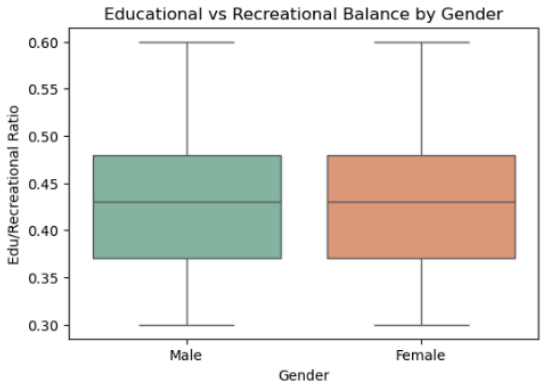

[38]:
```
#Device Mix by Urban vs Rural
plt.figure(figsize=(8,5))
sns.countplot(x='Primary_Device', hue='Urban_or_Rural', data=df, palette='mako')
plt.title("Device Mix by Location Type (Urban/Rural)")
plt.xlabel("Primary Device")
plt.ylabel("Count")
plt.legend(title='Location Type')
plt.xticks(rotation=30)
plt.show()
```



The educational-to-recreational ratio by gender shows that girls generally maintain a more balanced screen usage, spending slightly more time on educational activities than boys. Across age bands, the ratio tends to decline with age, indicating that as children grow older, their screen time shifts more toward recreational purposes rather than educational ones.

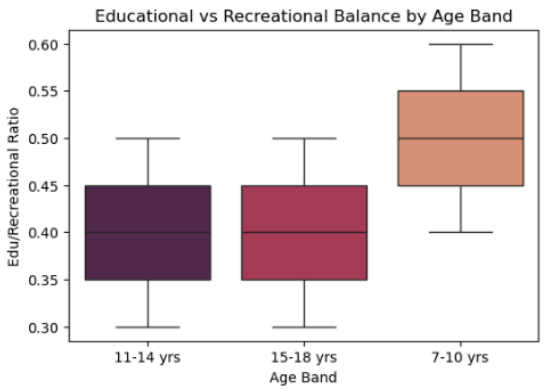
[39]:
```
#Activity Categories Across Demographics
#Educational/Recreational Ratio by Gender
plt.figure(figsize=(6,4))
sns.boxplot(x='Gender', y='Educational_to_Recreational_Ratio', data=df, palette='Set2')
plt.title("Educational vs Recreational Balance by Gender")
plt.xlabel("Gender")
plt.ylabel("Edu/Recreational Ratio")
plt.show()
```

```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/3566861620.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.boxplot(x='Gender', y='Educational_to_Recreational_Ratio', data=df, palette='Set2')
```




[40]:
```
#Educational/Recreational Ratio by Age Band
plt.figure(figsize=(6,4))
sns.boxplot(x='Age_Band', y='Educational_to_Recreational_Ratio', data=df, palette='rocket')
plt.title("Educational vs Recreational Balance by Age Band")
plt.xlabel("Age Band")
plt.ylabel("Edu/Recreational Ratio")
plt.show()
```

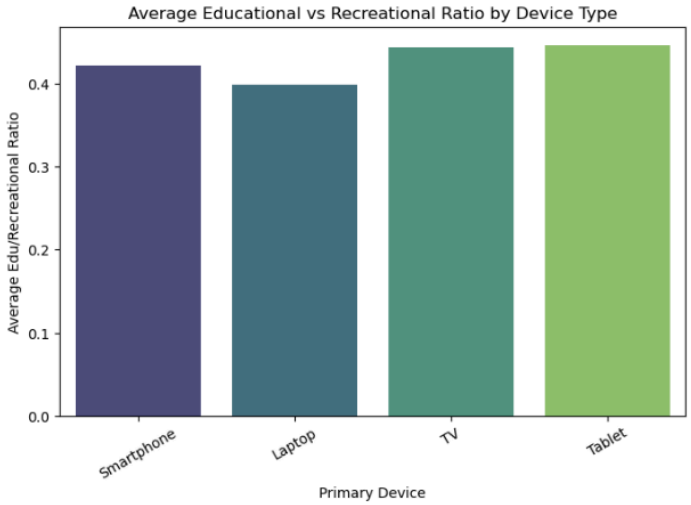
```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/3343666842.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=Fals
e` for the same effect.

  sns.boxplot(x='Age_Band', y='Educational_to_Recreational_Ratio', data=df, palette='rocket')
```



The **bar chart** shows that children using **laptops and tablets** have a higher **educational-to-recreational ratio**, suggesting these devices are used more for learning purposes, while **smartphones** are mainly used for recreation. The **summary table** highlights how **screen time and activity balance vary by age, gender, and location** — older and urban kids tend to spend **more screen hours** and have a **lower educational ratio**, especially among boys.

[41]:
```
#Combined Overview – Device vs Activity Category
plt.figure(figsize=(8,5))
sns.barplot(x='Primary_Device', y='Educational_to_Recreational_Ratio', data=df, ci=None, palette='viridis')
plt.title("Average Educational vs Recreational Ratio by Device Type")
plt.xlabel("Primary Device")
plt.ylabel("Average Edu/Recreational Ratio")
plt.xticks(rotation=30)
plt.show()
```

```
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/853713826.py:3: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.barplot(x='Primary_Device', y='Educational_to_Recreational_Ratio', data=df, ci=None, palette='viridis')
/var/folders/qs/t2124lzd51n4crc485v38zjc0000gn/T/ipykernel_1507/853713826.py:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(x='Primary_Device', y='Educational_to_Recreational_Ratio', data=df, ci=None, palette='viridis')
```




[42]:
```
summary = df.groupby(['Age_Band','Gender','Urban_or_Rural'])[['Avg_Daily_Screen_Time_hr','Educational_to_Recreational_Ratio']].mean().round(2)
display(summary)
```

| | | | Avg_Daily_Screen_Time_hr | Educational_to_Recreational_Ratio |
|---|---|---|---|---|
| **Age_Band** | **Gender** | **Urban_or_Rural** | | |
| **11-14 yrs** | **Female** | **Rural** | 4.54 | 0.4 |
| | | **Urban** | 4.48 | 0.4 |
| | **Male** | **Rural** | 4.51 | 0.4 |
| | | **Urban** | 4.54 | 0.4 |
| **15-18 yrs** | **Female** | **Rural** | 4.56 | 0.4 |
| | | **Urban** | 4.51 | 0.4 |
| | **Male** | **Rural** | 4.48 | 0.4 |
| | | **Urban** | 4.52 | 0.4 |
| **7-10 yrs** | **Female** | **Rural** | 3.89 | 0.5 |
| | | **Urban** | 3.81 | 0.5 |
| | **Male** | **Rural** | 4.09 | 0.5 |
| | | **Urban** | 3.98 | 0.5 |