```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import ttest_ind,f_oneway,chi2_contingency

df = pd.read_csv('/content/bike_sharing.csv')
df
```

	datetime	season	holiday	workingday	weather	temp	atemp	humidity	windspeed	casual	registered	count
0	2011-01-01 00:00:00	1	0	0	1	9.84	14.395	81	0.0000	3	13	16
1	2011-01-01 01:00:00	1	0	0	1	9.02	13.635	80	0.0000	8	32	40
2	2011-01-01 02:00:00	1	0	0	1	9.02	13.635	80	0.0000	5	27	32
3	2011-01-01 03:00:00	1	0	0	1	9.84	14.395	75	0.0000	3	10	13
4	2011-01-01 04:00:00	1	0	0	1	9.84	14.395	75	0.0000	0	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...
10881	2012-12-19 19:00:00	4	0	1	1	15.58	19.695	50	26.0027	7	329	336
10882	2012-12-19 20:00:00	4	0	1	1	14.76	17.425	57	15.0013	10	231	241
10883	2012-12-19 21:00:00	4	0	1	1	13.94	15.910	61	15.0013	4	164	168
10884	2012-12-19 22:00:00	4	0	1	1	13.94	17.425	61	6.0032	12	117	129
10885	2012-12-19 23:00:00	4	0	1	1	13.12	16.665	66	8.9981	4	84	88

10886 rows × 12 columns

Next steps: Generate code with df | View recommended plots | New interactive sheet

INITIALS OBSERVATIONS : *

Dataset size : 10886 rows * 12 columns column types :

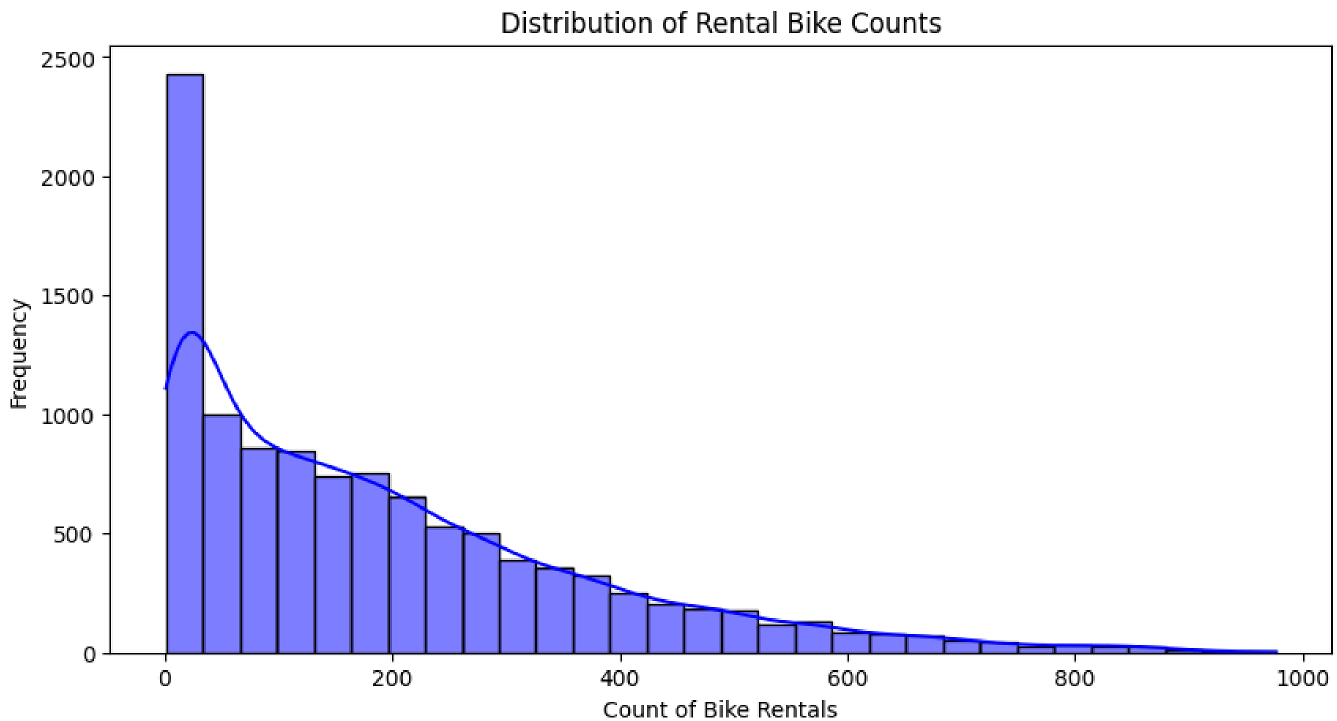
Categorical variables : season,holiday,working_day,weather

Numerical variables : temp,atemp,humidity,windspeed,casual,registered,count

```
# Histogram for count of rental bikes vs Frequency
```

```
plt.figure(figsize=(10, 5))
sns.histplot(df['count'], bins=30, kde=True, color="blue")
```

```
sns.histplot(df['count'], bins=30, kde=True, color='blue')
plt.title("Distribution of Rental Bike Counts")
plt.xlabel("Count of Bike Rentals")
plt.ylabel("Frequency")
plt.show()
```



The bike rental count (target variable) is right-skewed, meaning most rental counts are on the lower side, but there are some high-value rentals.

The peak occurs at lower rental values, indicating that demand is often moderate, with fewer high-rental instances.

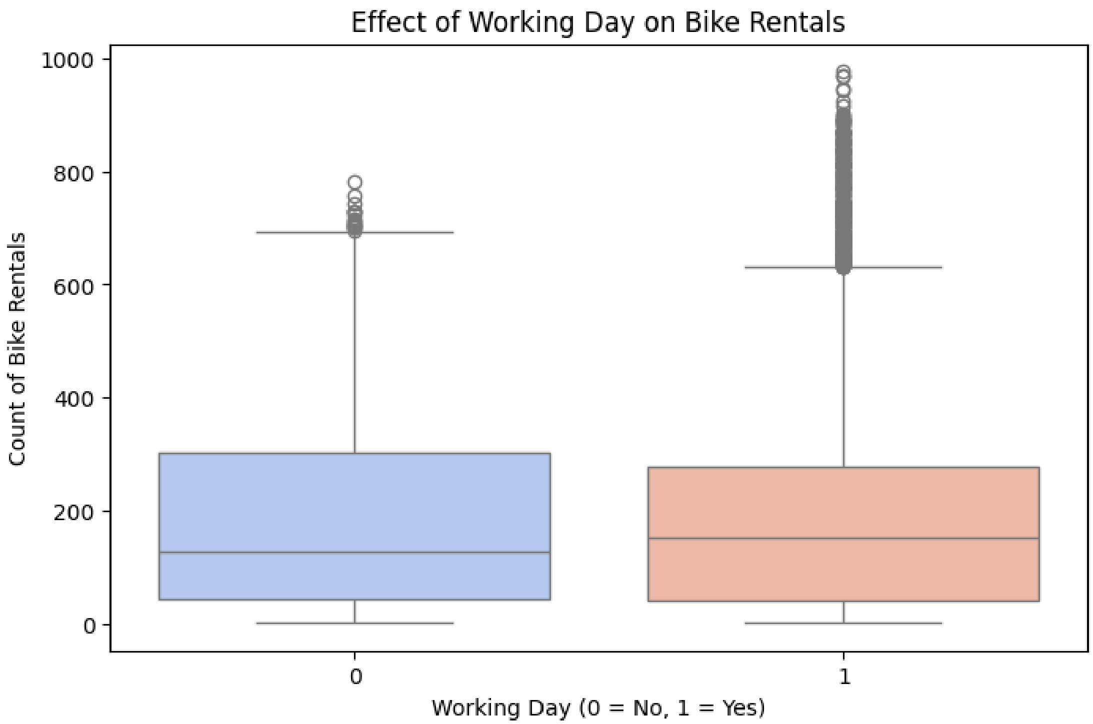
```
# Boxplot working day vs Bike Rentals

plt.figure(figsize=(8, 5))
sns.boxplot(x='workingday', y='count', data=df, palette="coolwarm")
plt.title("Effect of Working Day on Bike Rentals")
plt.xlabel("Working Day (0 = No, 1 = Yes)")
plt.ylabel("Count of Bike Rentals")
plt.show()
```

```
<ipython-input-9-979834ab3f25>: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=False` for the s

  sns.boxplot(x='workingday', y='count', data=df, palette="coolwarm")
```



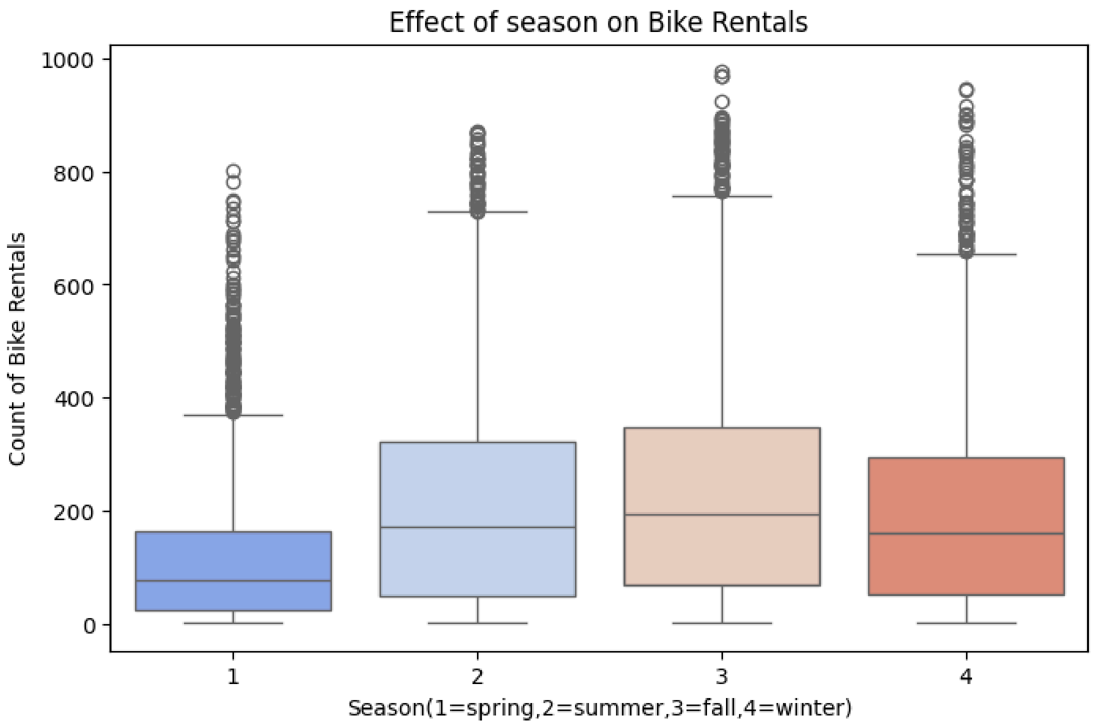

```
# Boxplot : Season vs Bike Rentals

plt.figure(figsize=(8,5))
sns.boxplot(x="season",y="count",data=df,palette="coolwarm")
plt.xlabel("Season(1=spring,2=summer,3=fall,4=winter)")
plt.ylabel("Count of Bike Rentals")
plt.title("Effect of season on Bike Rentals")
plt.show()
```

```
<ipython-input-11-47c2a27ef0a9>: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=False` for the s

  sns.boxplot(x="season",y="count",data=df,palette="coolwarm")
```



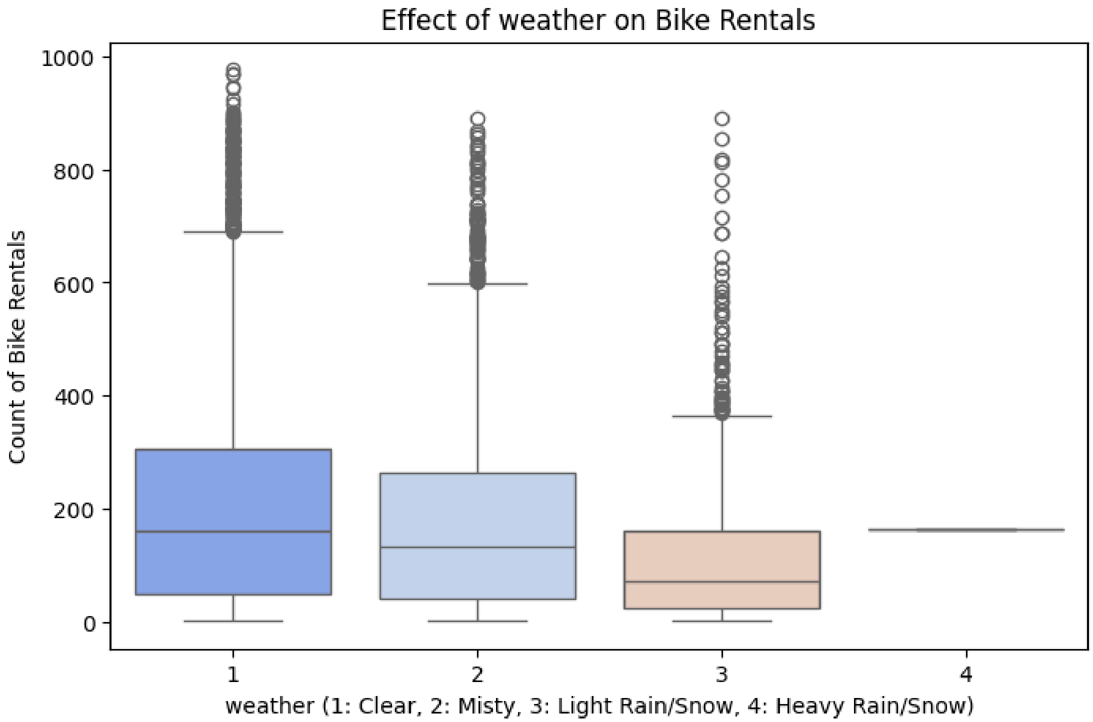

```
# Boxplot : Weather vs Bike Rentals

plt.figure(figsize=(8,5))
sns.boxplot(x="weather",y="count",data=df,palette="coolwarm")
plt.xlabel("weather (1: Clear, 2: Misty, 3: Light Rain/Snow, 4: Heavy Rain/Snow)")
plt.ylabel("Count of Bike Rentals")
plt.title("Effect of weather on Bike Rentals")
plt.show()
```

```
<ipython-input-14-6880dea65975>: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=False` for the s

  sns.boxplot(x="weather",y="count",data=df,palette="coolwarm")
```



KEY INSIGHTS ON BOXPLOTS:

1. Working days vs Rentals

   Median rentals on working days are slightly higher than non-working days.

2. Season vs Rentals

   Fall season has the highest median bike rentals while spring has the lowest median bike rentals.

3. Weather vs Rentals

   Clear weather has highest bike rentals and heavy rain/snow has lowest bike rentals

HYPOTHESIS TESTING SETUP :

1. **Working days vs Bike Rentals (2 Sample t-test)**

H0 (Null Hypothesis): Working day has no effect on the number of electric cycles rented (mean rentals are the same on working and non-working days).

Ha (Alternate Hypothesis): Working day affects the number of electric cycles rented (mean rentals are different).

```
workingday_0 = df[df['workingday'] == 0]['count']
workingday_1 = df[df['workingday'] == 1]['count']

t_stat,p_value = ttest_ind(workingday_0,workingday_1,equal_var = True)
p_value
```

```
0.22644804226361348
```

p_value for working days bike rentals is 0.22 which is higher than 0.05 we fail to reject null hypothesis.

Meaning bike rentals remain same on all days irrespective of working day or non-working days

ANOVA for season vs Bike Rentals

H0 : All seasons have same mean number of bike rentals

Ha : Atleast one season have significantly different number of rentals

```
season_groups = [df[df["season"] == i]["count"] for i in df["season"].unique()]
f_stat,p_value = f_oneway(*season_groups)
p_value
```

```
6.164843386499654e-149
```

p_value is less than 0.05 so we have to reject null hypothesis that means rentals differ significantly across seasons

ANOVA weather vs Bike Rentals

H0 : Mean number of Bike Rentals is same for all weather types

Ha : Atleast one weather conditions has significantly different number of rentals

```
weather_groups = [df[df["weather"] == i]["count"] for i in df['weather'].unique()]
f_stat,p_value = f_oneway(*weather_groups)
p_value
```

```
5.482069475935669e-42
```

p_value is less than 0.05 so we have to reject null hypothesis. Bike rentals differ significantly across weather.

Chi-Square Test for weather vs Season

H0 : weather is independent on the season.

Ha : weather is dependent on the season.

```
contingency_table = pd.crosstab(df['weather'],df['season'])
chi2_stat,p_value,dof,expected = chi2_contingency(contingency_table)
p_value
```

```
1.5499250736864862e-07
```

p_value is less than 0.05 so we reject null hypothesis.

weather is dependent on season.

RECOMMENDATIONS

1. working day affects rentals, Yulu can adjust pricing or promotions for weekends vs. weekdays.
2. season affects rentals, Yulu can increase or decrease fleet size accordingly.
3. weather affects rentals, Yulu can introduce weather-based pricing, discounts, or protective gear rental services.
4. weather depends on season, Yulu can use seasonal forecasts to predict demand trends and optimize bike availability.

Double-click (or enter) to edit

Double-click (or enter) to edit

Start coding or generate with AI.