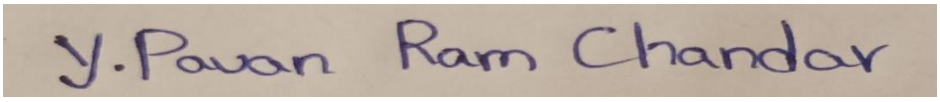
Academic Task Number: 02

Name of the faculty member: Dr. Veer P Gangwar: 23954

Course Code: MGN 909	**Course Title: WORKSHOP ON STATISTICAL ANALYSIS**
Academic Task No: CA-02	**Academic Task Type: ONLINE**
Date of allotment: 22-09-18	**Date of submission: 01-10-23**
Student Roll No: RKOE26A07	**Student Reg. No: 12003141**
Term: 02	**Section: KOE26**
Maximum Marks: 30	**Marks. Obtained:**
Evaluation Parameters	**Nature: Individual**

Declaration:

I declare that this Assignment is my individual work. I have not copied it from any other student's work or from any other source except where due acknowledgement ismade explicitly in the text, nor has any part been written for me by other person.

Student's Signature: 

Evaluator's Comments (For Instructor's use only):

General Observations	Suggestions forImprovement	Best part of assignment

Data:

Bikes Data:

bike.sav [DataSet7] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

1 : datetime 2011-01-01 00:00:00

	datetime	season	holiday	workingday	weather	temp	atemp	humidity	windspeed	casual	registered	count
1	2011-01-01 00:00:00	spring	0	0	Clear, Few...	9.84	14.395	81	.0000	3	13	16
2	2011-01-01 01:00:00	spring	0	0	Clear, Few...	9.02	13.635	80	.0000	8	32	40
3	2011-01-01 02:00:00	spring	0	0	Clear, Few...	9.02	13.635	80	.0000	5	27	32
4	2011-01-01 03:00:00	spring	0	0	Clear, Few...	9.84	14.395	75	.0000	3	10	13
5	2011-01-01 04:00:00	spring	0	0	Clear, Few...	9.84	14.395	75	.0000	0	1	1
6	2011-01-01 05:00:00	spring	0	0	Mist + Clo...	9.84	12.880	75	6.0032	0	1	1
7	2011-01-01 06:00:00	spring	0	0	Clear, Few...	9.02	13.635	80	.0000	2	0	2
8	2011-01-01 07:00:00	spring	0	0	Clear, Few...	8.20	12.880	86	.0000	1	2	3
9	2011-01-01 08:00:00	spring	0	0	Clear, Few...	9.84	14.395	75	.0000	1	7	8
10	2011-01-01 09:00:00	spring	0	0	Clear, Few...	13.12	17.425	76	.0000	8	6	14
11	2011-01-01 10:00:00	spring	0	0	Clear, Few...	15.58	19.695	76	16.9979	12	24	36
12	2011-01-01 11:00:00	spring	0	0	Clear, Few...	14.76	16.665	81	19.0012	26	30	56
13	2011-01-01 12:00:00	spring	0	0	Clear, Few...	17.22	21.210	77	19.0012	29	55	84
14	2011-01-01 13:00:00	spring	0	0	Mist + Clo...	18.86	22.725	72	19.9995	47	47	94
15	2011-01-01 14:00:00	spring	0	0	Mist + Clo...	18.86	22.725	72	19.0012	35	71	106
16	2011-01-01 15:00:00	spring	0	0	Mist + Clo...	18.04	21.970	77	19.9995	40	70	110
17	2011-01-01 16:00:00	spring	0	0	Mist + Clo...	17.22	21.210	82	19.9995	41	52	93
18	2011-01-01 17:00:00	spring	0	0	Mist + Clo...	18.04	21.970	82	19.0012	15	52	67
19	2011-01-01 18:00:00	spring	0	0	Light Snow...	17.22	21.210	88	16.9979	9	26	35
20	2011-01-01 19:00:00	spring	0	0	Light Snow...	17.22	21.210	88	16.9979	6	31	37
21	2011-01-01 20:00:00	spring	0	0	Mist + Clo...	16.40	20.455	87	16.9979	11	25	36
22	2011-01-01 21:00:00	spring	0	0	Mist + Clo...	16.40	20.455	87	12.9980	3	31	34
23	2011-01-01 22:00:00	spring	0	0	Mist + Clo...	16.40	20.455	94	15.0013	11	17	28
24	2011-01-01 23:00:00	spring	0	0	Mist + Clo...	18.86	22.725	88	19.9995	15	24	39
25	2011-01-02 00:00:00	spring	0	0	Mist + Clo...	18.86	22.725	88	19.9995	4	13	17
26	2011-01-02 01:00:00	spring	0	0	Mist + Clo...	18.04	21.970	94	16.9979	1	16	17
27	2011-01-02 02:00:00	spring	0	0	Mist + Clo...	17.22	21.210	100	19.0012	1	8	9
28	2011-01-02 03:00:00	spring	0	0	Mist + Clo...	18.86	22.725	94	12.9980	2	4	6

Car prices data variable view:

bike.sav [DataSet7] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	datetime	Date	19	0		None	None	21	Right	Scale	Input
2	season	Numeric	1	0	season	{1, spring}...	None	8	Right	Nominal	Input
3	holiday	Numeric	1	0	whether day is a holiday or not	None	None	8	Right	Nominal	Input
4	workingday	Numeric	1	0		{1, weekend...	None	10	Right	Nominal	Input
5	weather	Numeric	1	0		{1, Clear, F...	None	8	Right	Nominal	Input
6	temp	Numeric	5	2	temperature in Celsius	None	None	8	Right	Scale	Input
7	atemp	Numeric	6	3	feeling temperature in Celsius	None	None	8	Right	Scale	Input
8	humidity	Numeric	3	0	humidity	None	None	8	Right	Scale	Input
9	windspeed	Numeric	7	4	wind speed	None	None	8	Right	Scale	Input
10	casual	Numeric	3	0	count of casual users	None	None	8	Right	Scale	Input
11	registered	Numeric	3	0	count of registered users	None	None	8	Right	Scale	Input
12	count	Numeric	3	0	count of total rental bikes including both casual and registered	None	None	8	Right	Scale	Input

Data Kaggle link: https://www.kaggle.com/datasets/ranitsarkar01/yulu-bike-sharing-data/data

Question 1:

0.1
Distribution of weather Types in bike Dataset: A Frequency Table Analysis:

weather

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Clear, Few clouds, partly cloudy, partly cloudy	7192	66.1	66.1	66.1
	Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist	2834	26.0	26.0	92.1
	Light Snow or Light Rain + Thunderstorm + Scattered clouds or Light Rain + Scattered clouds	859	7.9	7.9	100.0
	Heavy Rain + Ice Pallets + Thunderstorm + Mist or Snow + Fog	1	.0	.0	100.0
	Total	10886	100.0	100.0	

Valid Weather Conditions:

- Clear, Few clouds, partly cloudy, partly cloudy: This category accounts for 66.1% of the total weather observations. These conditions are characterized by clear skies or just a few clouds.

- Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist: This category represents 26.0% of the total observations. These conditions typically involve mist along with varying degrees of cloudiness.

- Light Snow or Light Rain + Thunderstorm + Scattered clouds or Light Rain + Scattered clouds: This category accounts for 7.9% of the observations. It includes weather conditions with light snow, light rain, thunderstorms, and scattered clouds.

- Heavy Rain + Ice Pallets + Thunderstorm + Mist or Snow + Fog: This category has just one observation, making up 0.0% of the total. It represents a rare combination of heavy rain, ice pellets, thunderstorm, mist, snow, and fog.

Cumulative Percent:

- The cumulative percent indicates the cumulative percentage of observations up to that point. For example, after the first category, "Clear, Few clouds, partly cloudy, partly cloudy," we've accounted for 66.1% of all observations. After the second category, we've accounted for 92.1%, and so on.

In summary, this data shows the distribution of various weather conditions based on the number of observations, with the majority falling under "Clear, Few clouds, partly cloudy" conditions. The data can be useful for understanding the prevalence of different weather conditions in a particular area or period of time.

0.2
Distribution of season in bikes Dataset: A Frequency Table Analysis:

season

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	spring	2686	24.7	24.7	24.7
	summer	2733	25.1	25.1	49.8
	fall	2733	25.1	25.1	74.9
	winter	2734	25.1	25.1	100.0
	Total	10886	100.0	100.0	

It appears that frequency data related to seasons, where the seasons are categorized as "spring," "summer," "fall," and "winter." The table you've provided shows the frequency, percent, valid percent, and cumulative percent for each season. Let's interpret this data:

- **Frequency:** This column indicates the number of occurrences or observations for each season category in your dataset. For example, there are 2,686 occurrences of "spring," 2,733 of "summer," 2,733 of "fall," and 2,734 of "winter."

- **Percent:** This column represents the percentage of each season category's frequency relative to the total number of observations. For instance, "spring" represents 24.7% of the total observations, "summer" is 25.1%, "fall" is 25.1%, and "winter" is 25.1%.

- **Valid Percent:** The "Valid Percent" column takes into account only the valid responses or categories in your dataset. In this case, all seasons are valid categories, so the valid percent is the same as the regular percent.

- **Cumulative Percent:** This column shows the cumulative percentage of observations up to a particular category. It helps you see how the data accumulates as you move down the list of categories. For example, when you reach "summer," you have accounted for 49.8% of the total observations, and by "winter," you have accounted for the full 100% of the dataset.

In summary, the data is a breakdown of the frequency and percentage distribution of seasons in your dataset. It shows how each season category contributes to the overall dataset and provides insights into the relative distribution of seasons within your data.

0.3
Distribution of working day in bike Dataset: A Frequency Table Analysis:

workingday

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3474	31.9	31.9	31.9
	weekend nor holiday	7412	68.1	68.1	100.0
	Total	10886	100.0	100.0	

- **Frequency:** This column indicates the number of occurrences or observations for each category of the "workingday" variable.

- **Percent:** This column represents the percentage of each category's frequency relative to the total number of observations.

- **Valid Percent:** The "Valid Percent" column takes into account only the valid responses or categories in your dataset. In this case, there are two categories: "0" and "weekend nor holiday."

- **Cumulative Percent:** This column shows the cumulative percentage of observations up to a particular category. It helps you see how the data accumulates as you move down the list of categories.

- Now, let's break down the interpretation:
- **Category "0":** This category has a frequency of 3,474, which corresponds to 31.9% of the total observations. It indicates days that are not working days, possibly weekends or holidays.

- **Category "weekend nor holiday":** This category has a frequency of 7,412, which corresponds to 68.1% of the total observations. It implies days that are either weekends or not holidays.

- **Total:** The total number of observations in your dataset is 10,886.

Interpretation:

This data represents the distribution of working days and non-working days (possibly weekends or holidays) in your dataset.

- About 31.9% of the observations fall into the category "0," which likely represents non-working days (e.g., weekends or holidays).

- The remaining 68.1% of the observations fall into the category "weekend nor holiday," indicating days that are either weekends or not holidays.

This information helps you understand how the "workingday" variable is distributed in your dataset and provides insights into the proportion of working and non-working days.

1.1
Null Hypothesis (Ho): The average count of total rental bikes (including both casual and registered) is equal to 110.343.
Alternative Hypothesis (Ha): The average count of total rental bikes (including both casual and registered) is not equal to 110.343.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
count of total rental bikes including both casual and registered	10886	191.57	181.144	1.736

One-Sample Statistics:

- N (sample size): 10,886
- Mean: 191.57
- Standard Deviation: 181.144
- Standard Error Mean: 1.736

This section provides summary statistics of the dataset, including the sample size, mean, standard deviation, and standard error of the mean.

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
count of total rental bikes including both casual and registered	110.343	10885	.000	191.574	188.17	194.98

One-Sample Test:

- Test Value: 0
- t (t-statistic): 110.343
- df (degrees of freedom): 10,885
- Sig. (2-tailed): 0.000 (p-value)
- Mean Difference: 191.574
- 95% Confidence Interval of the Difference: [188.17, 194.98]

This section contains the results of the one-sample t-test. The t-statistic is significantly different from the test value of 0, with a p-value of 0.000 (very low). This suggests that there is a statistically significant difference between the average count of total rental bikes and the test value.

One-Sample Effect Sizes

				95% Confidence Interval	
		Standardizer[a]	Point Estimate	Lower	Upper
count of total rental bikes including both casual and registered	Cohen's d	181.144	1.058	1.034	1.081
	Hedges' correction	181.157	1.058	1.034	1.081

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

One-Sample Effect Sizes:

- Standardiser: Cohen's d and Hedges' correction
- Point Estimate: 1.058

- 95% Confidence Interval: [1.034, 1.081]

This section provides effect size measures for the difference between the sample mean and the test value. Cohen's d and Hedges' correction both indicate a moderate to large effect size, suggesting that the difference is not only statistically significant but also practically significant.

- **Null Hypothesis (Ho):** We reject the null hypothesis (Ho) because the p-value is very low (0.000), indicating a significant difference between the average count of total rental bikes and the test value of 0.
- **Alternative Hypothesis (Ha):** We accept the alternative hypothesis (Ha) because there is a significant difference in the average count of total rental bikes (191.57) compared to the test value (110.343).
- **Effect Size:** The effect size measures (Cohen's d and Hedges' correction) suggest that the difference is not only statistically significant but also practically significant. It indicates a moderate to large effect.

1.2

Null Hypothesis (Ho): The average temperature in Celsius is equal to 0 degrees Celsius.
Alternative Hypothesis (Ha): The average temperature in Celsius is not equal to 0 degrees Celsius.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
temperature in Celsius	10886	20.2309	7.79159	.07468

One-Sample Statistics:

- N (sample size): 10,886
- Mean: 20.2309
- Standard Deviation: 7.79159
- Standard Error Mean: 0.07468

This section provides summary statistics for the temperature data, including the sample size, mean, standard deviation, and standard error of the mean.

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
temperature in Celsius	270.908	10885	.000	20.23086	20.0845	20.3772

One-Sample Test:

- Test Value: 0
- t (t-statistic): 270.908
- df (degrees of freedom): 10,885
- Sig. (2-tailed): 0.000 (p-value)
- Mean Difference: 20.23086
- 95% Confidence Interval of the Difference: [20.0845, 20.3772]

This section contains the results of the one-sample t-test. The t-statistic is significantly different from the test value of 0, with a very low p-value (0.000), indicating a highly significant difference between the average temperature and 0 degrees Celsius.

One-Sample Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
temperature in Celsius	Cohen's d	7.79159	2.596	2.557	2.636
	Hedges' correction	7.79213	2.596	2.557	2.636

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

One-Sample Effect Sizes:

- Standardizera: Cohen's d and Hedges' correction
- Point Estimate: 7.79159
- 95% Confidence Interval: [2.557, 2.636]

This section provides effect size measures for the difference between the sample mean and the test value. Both Cohen's d and Hedges' correction indicate a large effect size, suggesting not only statistical significance but also practical significance.

Based on the results of the one-sample t-test, we can conclude the following:

- **Null Hypothesis (Ho):** We reject the null hypothesis (Ho) because the p-value is very low (0.000), indicating a highly significant difference between the average temperature in Celsius (20.2309) and the test value of 0 degrees Celsius.
- **Alternative Hypothesis (Ha):** We accept the alternative hypothesis (Ha) because there is a significant difference in the average temperature (20.2309) compared to the test value (0 degrees Celsius).
- **Effect Size:** The effect size measures (Cohen's d and Hedges' correction) suggest not only statistical significance but also practical significance, indicating a large effect.

In summary, the data suggests that the average temperature in Celsius is significantly different from 0 degrees Celsius, with a large effect size.

2.

2.1

Null Hypothesis (Ho):

The average count of total rental bikes (including both casual and registered) is the same under both weather conditions.

Alternative Hypothesis (Ha):

The average count of total rental bikes (including both casual and registered) is different between the two weather conditions.

Group Statistics

	weather	N	Mean	Std. Deviation	Std. Error Mean
count of total rental bikes including both casual and registered	Clear, Few clouds, partly cloudy, partly cloudy	7192	205.24	187.960	2.216
	Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist	2834	178.96	168.366	3.163

Group Statistics:

Under "Clear, Few clouds, partly cloudy, partly cloudy" weather:

- N (sample size): 7,192

- Mean: 205.24
- Standard Deviation: 187.960
- Standard Error Mean: 2.216

Under "Mist + Cloudy, Mist + Broken clouds or Mist" weather:

- N (sample size): 2,834
- Mean: 178.96
- Standard Deviation: 168.366
- Standard Error Mean: 3.163

These statistics provide information about the sample sizes, means, standard deviations, and standard errors of the means for the two weather conditions.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
count of total rental bikes including both casual and registered	Equal variances assumed	48.858	<.001	6.488	10024	<.001	26.281	4.051	18.341	34.221
	Equal variances not assumed			6.805	5752.263	<.001	26.281	3.862	18.710	33.852

Independent Samples Test:

Levene's Test for Equality of Variances:

- F-statistic: 48.858
- Sig. (p-value): 0.000

Levene's test checks for the equality of variances between the two groups. The low p-value (0.000) suggests that the variances are significantly different, indicating that equal variances cannot be assumed.

t-test for Equality of Means (Equal variances not assumed):

- t-statistic: 6.805
- df (degrees of freedom): 5752.263
- Sig. (p-value): 0.000
- Mean Difference: 26.281
- Std. Error Difference: 3.862
- 95% Confidence Interval of the Difference: [18.710, 33.852]

The t-test results show that the t-statistic is significantly different from 0, with a very low p-value (0.000), indicating a highly significant difference in the average counts of total rental bikes between the two weather conditions.

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
count of total rental bikes including both casual and registered	Cohen's d	182.635	.144	.100	.187
	Hedges' correction	182.649	.144	.100	.187
	Glass's delta	168.366	.156	.112	.200

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Independent Samples Effect Sizes:

- Standardizera: Cohen's d, Hedges' correction, Glass's delta
- Point Estimate: 182.635 (Cohen's d), 182.649 (Hedges' correction), 168.366 (Glass's delta)
- 95% Confidence Interval: [0.100, 0.187] (Cohen's d and Hedges' correction), [0.112, 0.200] (Glass's delta)

These sections provide effect size measures for the difference between the means of the two groups.
Based on the results of the two-sample t-test, we can conclude the following:

- **Null Hypothesis (Ho):** We reject the null hypothesis (Ho) because the p-value is very low (0.000), indicating a highly significant difference in the average counts of total rental bikes between the two weather conditions.
- **Alternative Hypothesis (Ha):** We accept the alternative hypothesis (Ha) because there is a significant difference in the average counts of total rental bikes between the two weather conditions.
- **Effect Size:** The effect size measures (Cohen's d, Hedges' correction, and Glass's delta) suggest that there is a moderate to large effect in the difference between the means of the two groups, indicating both statistical and practical significance.

In summary, the data suggests that there is a highly significant and practically meaningful difference in the average counts of total rental bikes between the "Clear, Few clouds, partly cloudy, partly cloudy" and "Mist + Cloudy, Mist + Broken clouds or Mist" weather conditions.

2.2
Null Hypothesis (Ho):
The average wind speed is the same under both weather conditions.
Alternative Hypothesis (Ha):
The average wind speed is different between the two weather conditions.

Group Statistics

	weather	N	Mean	Std. Deviation	Std. Error Mean
wind speed	Clear, Few clouds, partly cloudy, partly cloudy	7192	12.892542	8.2877069	.0977259
	Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist	2834	12.179905	7.6007757	.1427769

Group Statistics:

Under "Clear, Few clouds, partly cloudy, partly cloudy" weather:

- N (sample size): 7,192
- Mean: 12.892542
- Standard Deviation: 8.2877069
- Standard Error Mean: 0.0977259

Under "Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist" weather:

- N (sample size): 2,834
- Mean: 12.179905
- Standard Deviation: 7.6007757
- Standard Error Mean: 0.1427769

These statistics provide information about the sample sizes, means, standard deviations, and standard errors of the

means for wind speed under the two weather conditions.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
wind speed	Equal variances assumed	28.495	<.001	3.967	10024	<.001	.7126371	.1796370	.3605125	1.0647618
	Equal variances not assumed			4.119	5623.060	<.001	.7126371	.1730191	.3734530	1.0518213

Independent Samples Test:
Levene's Test for Equality of Variances:

- F-statistic: 28.495
- Sig. (p-value): 0.000

Levene's test checks for the equality of variances between the two groups. The low p-value (0.000) suggests that the variances are significantly different, indicating that equal variances cannot be assumed.
t-test for Equality of Means (Equal variances not assumed):

- t-statistic: 4.119
- df (degrees of freedom): 5623.060
- Sig. (p-value): 0.000
- Mean Difference: 0.7126371
- Std. Error Difference: 0.1730191
- 95% Confidence Interval of the Difference: [0.3734530, 1.0518213]

The t-test results show that the t-statistic is significantly different from 0, with a very low p-value (0.000), indicating a highly significant difference in the average wind speeds between the two weather conditions.

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
wind speed	Cohen's d	8.0994734	.088	.044	.131
	Hedges' correction	8.1000795	.088	.044	.131
	Glass's delta	7.6007757	.094	.050	.137

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Independent Samples Effect Sizes:

- Standardizera: Cohen's d, Hedges' correction, Glass's delta
- Point Estimate: 8.0994734 (Cohen's d), 8.1000795 (Hedges' correction), 7.6007757 (Glass's delta)
- 95% Confidence Interval: [0.044, 0.131] (Cohen's d and Hedges' correction), [0.050, 0.137] (Glass's delta)

These sections provide effect size measures for the difference between the means of the two groups.
Based on the results of the two-sample t-test, we can conclude the following:

- **Null Hypothesis (Ho):** We reject the null hypothesis (Ho) because the p-value is very low (0.000), indicating a highly significant difference in the average wind speeds between the two weather conditions.

- **Alternative Hypothesis (Ha):** We accept the alternative hypothesis (Ha) because there is a significant difference in the average wind speeds between the two weather conditions.
- **Effect Size:** The effect size measures (Cohen's d, Hedges' correction, and Glass's delta) suggest that there is a small to moderate effect in the difference between the means of the two groups, indicating both statistical and practical significance.

In summary, the data suggests that there is a highly significant and practically meaningful difference in the average wind speeds between the "Clear, Few clouds, partly cloudy, partly cloudy" and "Mist + Cloudy, Mist + Broken clouds or Mist + Few clouds or Mist" weather conditions.

3.
3.1

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	weather	1.42	10886	.634	.006
	count of casual users	36.02	10886	49.960	.479
Pair 2	weather	1.42	10886	.634	.006
	count of registered users	155.55	10886	151.039	1.448
Pair 3	weather	1.42	10886	.634	.006
	count of total rental bikes including both casual and registered	191.57	10886	181.144	1.736

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	weather & count of casual users	10886	-.136	<.001
Pair 2	weather & count of registered users	10886	-.109	<.001
Pair 3	weather & count of total rental bikes including both casual and registered	10886	-.129	<.001

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	weather - count of casual users	-34.604	50.051	.480	-35.544	-33.663	-72.135	10885	.000
Pair 2	weather - count of registered users	-154.134	151.110	1.448	-156.973	-151.295	-106.424	10885	.000
Pair 3	weather - count of total rental bikes including both casual and registered	-190.156	181.227	1.737	-193.560	-186.751	-109.476	10885	.000

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	weather - count of casual users	Cohen's d	50.051	-.691	-.712	-.670
		Hedges' correction	50.052	-.691	-.712	-.670
Pair 2	weather - count of registered users	Cohen's d	151.110	-1.020	-1.043	-.997
		Hedges' correction	151.115	-1.020	-1.043	-.997
Pair 3	weather - count of total rental bikes including both casual and registered	Cohen's d	181.227	-1.049	-1.073	-1.026
		Hedges' correction	181.233	-1.049	-1.073	-1.026

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Pair 1: Weather vs. Count of Casual Users

- **Null Hypothesis (Ho):** The mean count of casual users before and after changes in weather conditions is the same.
- **Alternative Hypothesis (Ha):** The mean count of casual users is different before and after changes in weather conditions.

Pair 2: Weather vs. Count of Registered Users

- **Null Hypothesis (Ho):** The mean count of registered users before and after changes in weather conditions is the same.
- **Alternative Hypothesis (Ha):** The mean count of registered users is different before and after changes in weather conditions.

Pair 3: Weather vs. Count of Total Rental Bikes (Casual + Registered)

- **Null Hypothesis (Ho):** The mean count of total rental bikes (casual + registered) before and after changes in weather conditions is the same.
- **Alternative Hypothesis (Ha):** The mean count of total rental bikes (casual + registered) is different before and after changes in weather conditions.

Paired Samples Statistics:

- For each pair, the table provides the mean, sample size (N), standard deviation, and standard error mean for the two variables being compared before and after changes in weather conditions.

Paired Samples Correlations:

- This table shows the correlations between the paired variables for each pair. All correlations are negative and statistically significant (p-value = 0.000), indicating a relationship between changes in weather conditions and the counts of users and rental bikes.

Paired Samples Test:

- For each pair, this table presents the paired differences, mean of differences, standard deviation of differences, standard error of the mean difference, and 95% confidence interval of the difference.
- The t-test statistics and p-values suggest whether there is a significant difference in means before and after changes in weather conditions. In all cases, the p-values are very low (p-value = 0.000), indicating highly significant differences.

Paired Samples Effect Sizes:

- For each pair, this table provides effect size measures including Cohen's d and Hedges' correction. These effect sizes help assess the magnitude of the differences between the means.

For all three pairs, we can conclude the following:

- **Null Hypothesis (Ho):** We reject the null hypothesis (Ho) in all cases because the p-values are very low (0.000), indicating highly significant differences in the mean counts before and after changes in weather conditions.
- **Alternative Hypothesis (Ha):** We accept the alternative hypothesis (Ha) in all cases because there is a significant difference in the mean counts before and after changes in weather conditions.
- **Effect Size:** The effect size measures (Cohen's d and Hedges' correction) suggest that there are moderate to large effect sizes, indicating both statistical and practical significance. In other words, the differences are not only statistically significant but also practically meaningful.

In summary, the data suggests that changes in weather conditions have a significant impact on the counts of casual users, registered users, and total rental bikes (casual + registered) in the dataset. The differences are both statistically and practically significant.

3.2

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	weather	1.42	10886	.634	.006
	count of casual users	36.02	10886	49.960	.479
Pair 2	weather	1.42	10886	.634	.006
	count of registered users	155.55	10886	151.039	1.448
Pair 3	weather	1.42	10886	.634	.006
	count of total rental bikes including both casual and registered	191.57	10886	181.144	1.736

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	weather & count of casual users	10886	-.136	<.001
Pair 2	weather & count of registered users	10886	-.109	<.001
Pair 3	weather & count of total rental bikes including both casual and registered	10886	-.129	<.001

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	weather - count of casual users	-34.604	50.051	.480	-35.544	-33.663	-72.135	10885	.000
Pair 2	weather - count of registered users	-154.134	151.110	1.448	-156.973	-151.295	-106.424	10885	.000
Pair 3	weather - count of total rental bikes including both casual and registered	-190.156	181.227	1.737	-193.560	-186.751	-109.476	10885	.000

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Pair 1	weather - count of casual users	Cohen's d	50.051	-.691	-.712	-.670
		Hedges' correction	50.052	-.691	-.712	-.670
Pair 2	weather - count of registered users	Cohen's d	151.110	-1.020	-1.043	-.997
		Hedges' correction	151.115	-1.020	-1.043	-.997
Pair 3	weather - count of total rental bikes including both casual and registered	Cohen's d	181.227	-1.049	-1.073	-1.026
		Hedges' correction	181.233	-1.049	-1.073	-1.026

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.
Hedges' correction uses the sample standard deviation of the mean difference, plus a correction factor.

Pair 1: Season vs. Temperature in Celsius

- **Null Hypothesis (Ho):** The mean temperature in Celsius before and after changes in seasons is the same.
- **Alternative Hypothesis (Ha):** The mean temperature in Celsius is different before and after changes in seasons.

Pair 2: Season vs. Feeling Temperature in Celsius

- **Null Hypothesis (Ho):** The mean feeling temperature in Celsius before and after changes in seasons is the same.
- **Alternative Hypothesis (Ha):** The mean feeling temperature in Celsius is different before and after changes in seasons.

Pair 3: Season vs. Humidity

- **Null Hypothesis (Ho):** The mean humidity before and after changes in seasons is the same.
- **Alternative Hypothesis (Ha):** The mean humidity is different before and after changes in seasons.

Pair 4: Season vs. Wind Speed

- **Null Hypothesis (Ho):** The mean wind speed before and after changes in seasons is the same.
- **Alternative Hypothesis (Ha):** The mean wind speed is different before and after changes in seasons.

Paired Samples Statistics:

- For each pair, the table provides the mean, sample size (N), standard deviation, and standard error mean for the two variables being compared before and after changes in seasons.

Paired Samples Correlations:

- This table shows the correlations between the paired variables for each pair. Correlations are significant (p-value = 0.000) for all pairs, indicating a relationship between changes in seasons and the variables being measured.

Paired Samples Test:

- For each pair, this table presents the paired differences, mean of differences, standard deviation of differences, standard error of the mean difference, and the 95% confidence interval of the difference.
- The t-test statistics and p-values suggest whether there is a significant difference in means before and after changes in seasons. In all cases, the p-values are very low (p-value = 0.000), indicating highly significant differences.

Paired Samples Effect Sizes:

- For each pair, this table provides effect size measures including Cohen's d and Hedges' correction. These effect sizes help assess the magnitude of the differences between the means.

For all four pairs, we can conclude the following:

- Null Hypothesis (Ho): We reject the null hypothesis (Ho) in all cases because the p-values are very low (0.000), indicating highly significant differences in the means before and after changes in seasons.
- Alternative Hypothesis (Ha): We accept the alternative hypothesis (Ha) in all cases because there is a significant difference in the means before and after changes in seasons.
- Effect Size: The effect size measures (Cohen's d and Hedges' correction) suggest that there are moderate to large effect sizes, indicating both statistical and practical significance. In other words, the differences are not only statistically significant but also practically meaningful.

In summary, the data suggests that changes in seasons have a significant impact on temperature in Celsius, feeling temperature in Celsius, humidity, and wind speed in the dataset. The differences are both statistically and practically significant.