# Dataset iris

## Not compacted

Automatic testing is bad and you should feel bad.

| **label** | **variable** | **Species** | | | **test** |
| --- | --- | --- | --- | --- | --- |
| **setosa** | **versicolor** | **virginica** |
| Sepal.Length | Min / Max | 4.3 / 5.8 | 4.9 / 7.0 | 4.9 / 7.9 | p value: <0.0001  (One-way analysis of means (not assuming equal variances)) |
| Med [IQR] | 5.0 [4.8;5.2] | 5.9 [5.6;6.3] | 6.5 [6.2;6.9] |
| Mean (std) | 5.0 (0.4) | 5.9 (0.5) | 6.6 (0.6) |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |
| Sepal.Width | Min / Max | 2.3 / 4.4 | 2.0 / 3.4 | 2.2 / 3.8 | p value: <0.0001  (One-way analysis of means) |
| Med [IQR] | 3.4 [3.2;3.7] | 2.8 [2.5;3.0] | 3.0 [2.8;3.2] |
| Mean (std) | 3.4 (0.4) | 2.8 (0.3) | 3.0 (0.3) |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |
| Petal.Length | Min / Max | 1.0 / 1.9 | 3.0 / 5.1 | 4.5 / 6.9 | p value: <0.0001  (Kruskal-Wallis rank sum test) |
| Med [IQR] | 1.5 [1.4;1.6] | 4.3 [4.0;4.6] | 5.5 [5.1;5.9] |
| Mean (std) | 1.5 (0.2) | 4.3 (0.5) | 5.6 (0.6) |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |
| Petal.Width | Min / Max | 0.1 / 0.6 | 1.0 / 1.8 | 1.4 / 2.5 | p value: <0.0001  (Kruskal-Wallis rank sum test) |
| Med [IQR] | 0.2 [0.2;0.3] | 1.3 [1.2;1.5] | 2.0 [1.8;2.3] |
| Mean (std) | 0.2 (0.1) | 1.3 (0.2) | 2.0 (0.3) |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |

## Compacted

When compacting, you might want to remove the test names.

| **variable** | **Species** | | | **test** |
| --- | --- | --- | --- | --- |
| **setosa** | **versicolor** | **virginica** |
| **Sepal.Length** |  |  |  | **p value: <0.0001** |
| Min / Max | 4.3 / 5.8 | 4.9 / 7.0 | 4.9 / 7.9 |  |
| Med [IQR] | 5.0 [4.8;5.2] | 5.9 [5.6;6.3] | 6.5 [6.2;6.9] |  |
| Mean (std) | 5.0 (0.4) | 5.9 (0.5) | 6.6 (0.6) |  |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |  |
| **Sepal.Width** |  |  |  | **p value: <0.0001** |
| Min / Max | 2.3 / 4.4 | 2.0 / 3.4 | 2.2 / 3.8 |  |
| Med [IQR] | 3.4 [3.2;3.7] | 2.8 [2.5;3.0] | 3.0 [2.8;3.2] |  |
| Mean (std) | 3.4 (0.4) | 2.8 (0.3) | 3.0 (0.3) |  |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |  |
| **Petal.Length** |  |  |  | **p value: <0.0001** |
| Min / Max | 1.0 / 1.9 | 3.0 / 5.1 | 4.5 / 6.9 |  |
| Med [IQR] | 1.5 [1.4;1.6] | 4.3 [4.0;4.6] | 5.5 [5.1;5.9] |  |
| Mean (std) | 1.5 (0.2) | 4.3 (0.5) | 5.6 (0.6) |  |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |  |
| **Petal.Width** |  |  |  | **p value: <0.0001** |
| Min / Max | 0.1 / 0.6 | 1.0 / 1.8 | 1.4 / 2.5 |  |
| Med [IQR] | 0.2 [0.2;0.3] | 1.3 [1.2;1.5] | 2.0 [1.8;2.3] |  |
| Mean (std) | 0.2 (0.1) | 1.3 (0.2) | 2.0 (0.3) |  |
| N (NA) | 50 (0) | 50 (0) | 50 (0) |  |

# Dataset mtcars2

This dataset has 22 rows and 4 columns.

Look, there are labels!

| **variable** | **Transmission** | | | **Total** | **effect** |
| --- | --- | --- | --- | --- | --- |
| **auto** | **manual** | **NA** |
| **Miles/(US) gallon** |  |  |  |  | **Difference in means (t-test CI) (auto minus manual): -7.24 CI95%[-10.85 to -3.64]** |
| Min / Max | 10.4 / 24.4 | 15.0 / 33.9 | no NA | 10.4 / 33.9 |  |
| Med [IQR] | 17.3 [14.9;19.2] | 22.8 [21.0;30.4] | no NA | 19.2 [15.4;22.8] |  |
| Mean (std) | 17.1 (3.8) | 24.4 (6.2) | no NA | 20.1 (6.0) |  |
| N (NA) | 19 (0) | 13 (0) | no NA | 32 (0) |  |
| **Number of cylinders** |  |  |  |  | **No effect?** |
| 4 | 3 (27.27%) | 8 (72.73%) | 0 | 11 (34.38%) |  |
| 6 | 4 (57.14%) | 3 (42.86%) | 0 | 7 (21.88%) |  |
| 8 | 12 (85.71%) | 2 (14.29%) | 0 | 14 (43.75%) |  |
| NA | 0 | 0 | 0 | 0 |  |
| Total | 19 (59.38%) | 13 (40.62%) | 0 | 32 (100.00%) |  |
| **Displacement (cu.in.)** |  |  |  |  | **Difference in means (bootstrap CI) (auto minus manual): 146.85 CI95%[80.21 to 213.48]** |
| Min / Max | 120.1 / 472.0 | 71.1 / 351.0 | no NA | 71.1 / 472.0 |  |
| Med [IQR] | 275.8 [196.3;360.0] | 120.3 [79.0;160.0] | no NA | 196.3 [120.8;326.0] |  |
| Mean (std) | 290.4 (110.2) | 143.5 (87.2) | no NA | 230.7 (123.9) |  |
| N (NA) | 19 (0) | 13 (0) | no NA | 32 (0) |  |

# Dataset esoph

This one was compacted beforehand for some reason.

| **variable** | **value** |
| --- | --- |
| **agegp** |  |
| 25-34 | 15 (17.05%) |
| 35-44 | 15 (17.05%) |
| 45-54 | 16 (18.18%) |
| 55-64 | 16 (18.18%) |
| 65-74 | 15 (17.05%) |
| 75+ | 11 (12.50%) |
| **alcgp** |  |
| 0-39g/day | 23 (26.14%) |
| 40-79 | 23 (26.14%) |
| 80-119 | 21 (23.86%) |
| 120+ | 21 (23.86%) |
| **tobgp** |  |
| 0-9g/day | 24 (27.27%) |
| 10-19 | 24 (27.27%) |
| 20-29 | 20 (22.73%) |
| 30+ | 20 (22.73%) |
| **ncases** |  |
| Min / Max | 0 / 17.0 |
| Med [IQR] | 1.0 [0;4.0] |
| Mean (std) | 2.3 (2.8) |
| N (NA) | 88 (0) |
| **ncontrols** |  |
| Min / Max | 1.0 / 60.0 |
| Med [IQR] | 6.0 [3.0;14.0] |
| Mean (std) | 11.1 (12.7) |
| N (NA) | 88 (0) |