Group Differences in a Rental Bike Sharing Dataset

Group Differences in a Rental Bike Sharing Dataset

About the dataset:

- **Data are from a bike sharing system; riders rent bikes for use on urban streets.**
- **Riders obtain a bike at one station, cycle, and then leave the bike at another station.**
- **Where and When?: Capital Bikeshare in Washington, D.C. for all of 2011 and 2012.**
- **Capital Bikeshare launched in late 2010, so early data inform its ongoing growth.**
- **For each day, counts of riders and measures of time and weather.**

Group Differences in a Rental Bike Sharing Dataset

Why is this of interest?

- **The use of these systems continues to expand. There are > 500 bike-sharing programs around the world which are composed of > 500,000 bicycles.**

- **These systems can give visitors and residents an option for traveling around a city beyond use of a fuel or electric powered vehicle. There can be a benefit to traffic, the environment, and health.**

- **The daily number of rider counts can be grouped by aspects of time and weather and then questions can be asked about how the groups differ. This research could be useful to the bike sharing company's managers, marketing department, and maintenance department.**

Group Differences in a Rental Bike Sharing Dataset

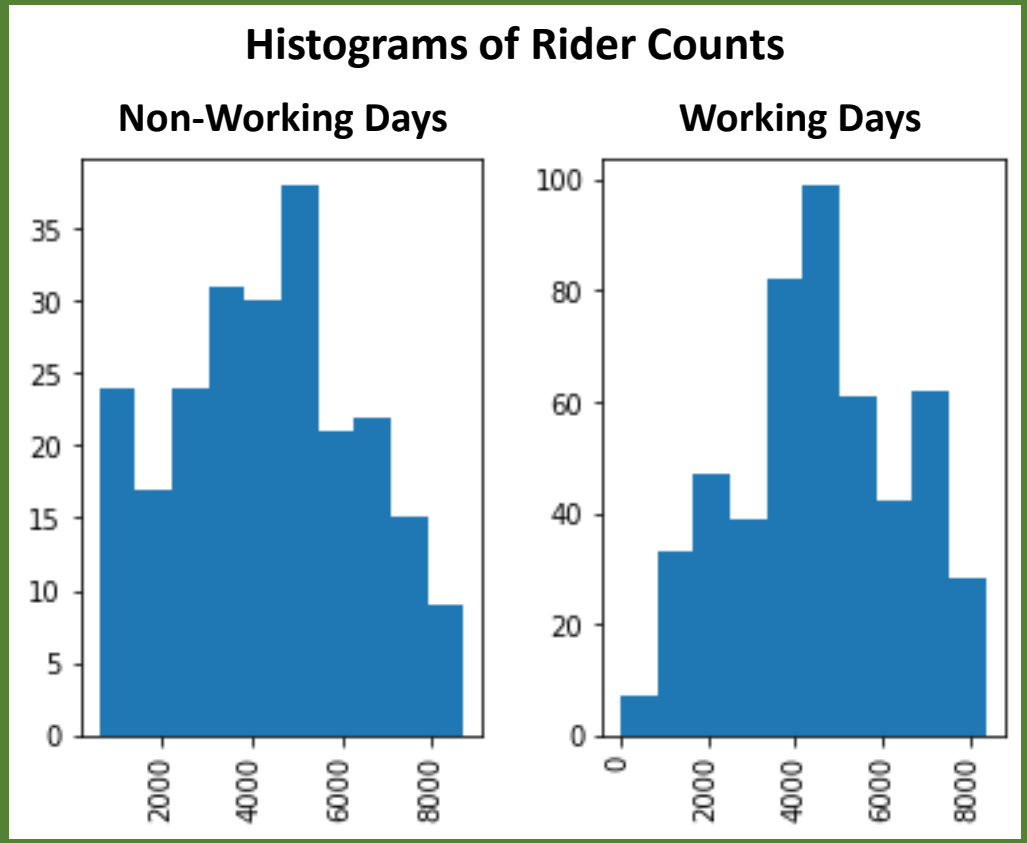
Questions asked

- Is there a difference in average rider counts on working days and non-working days?
- Is there a difference in average rider counts among the 7 days of the week?
- Is there a difference in average rider counts among the 4 seasons?
- Is there a difference in average rider counts with weather intensity?

Group Differences in a Rental Bike Sharing Dataset

Methods

- **Rider Counts on Working Days & Non-Working Days**



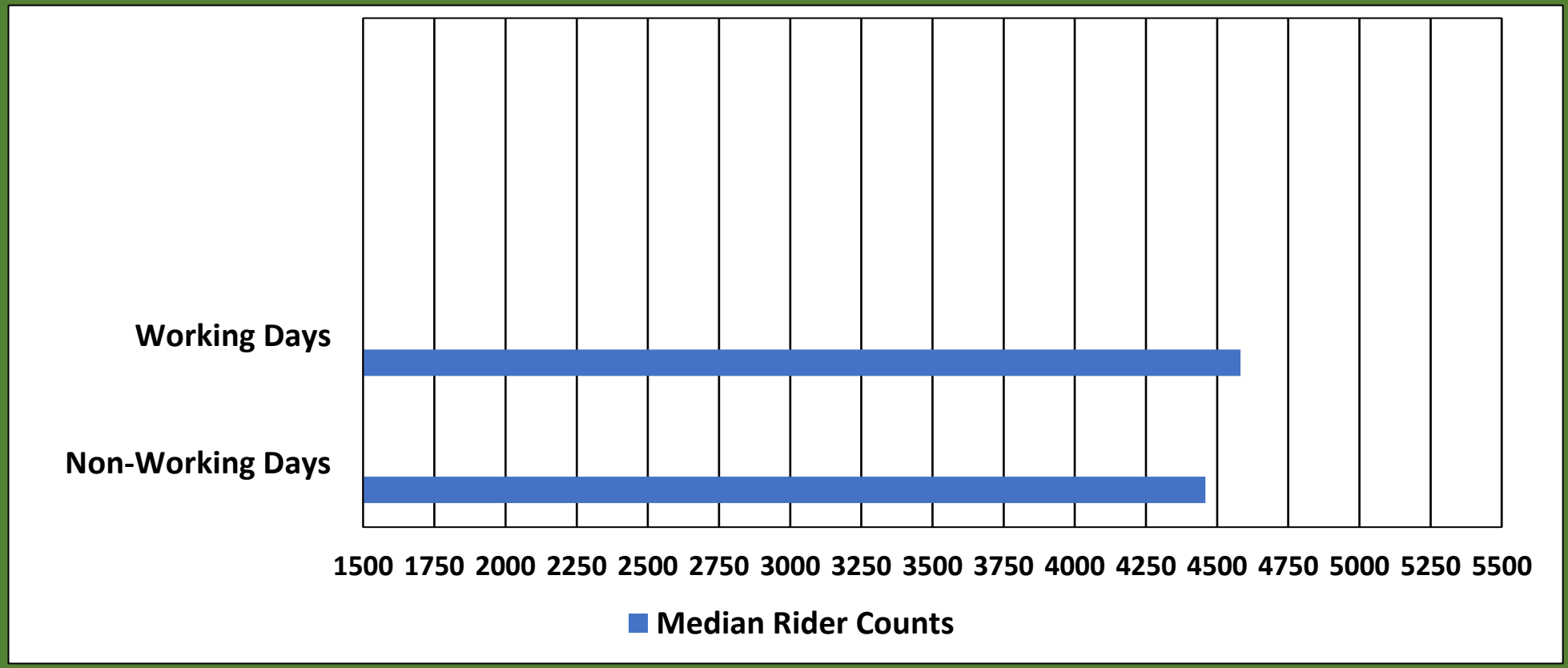
The histograms, and other tests and plots not shown, indicate the counts do not have a convincing normal distribution. A non-parametric test should be used to assess if there is a significant difference between the two groups. The Kruskal-Wallis test and Mood's median test are used.

Group Differences in a Rental Bike Sharing Dataset

Results

- **Median Rider Counts on Working Days & Non-Working Days**

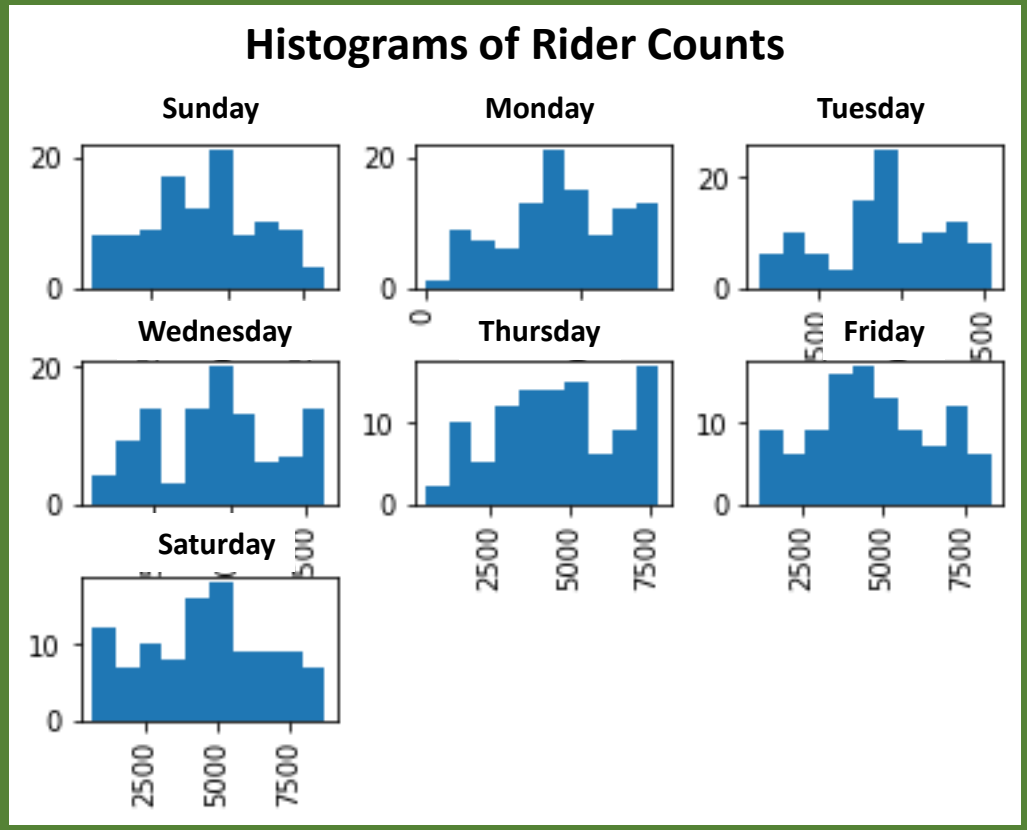
Significant Difference?	Working Days	Non-Working Days
No	4582.0	4459.0



Group Differences in a Rental Bike Sharing Dataset

Methods

- **Median Rider Counts on Different Days of Week**



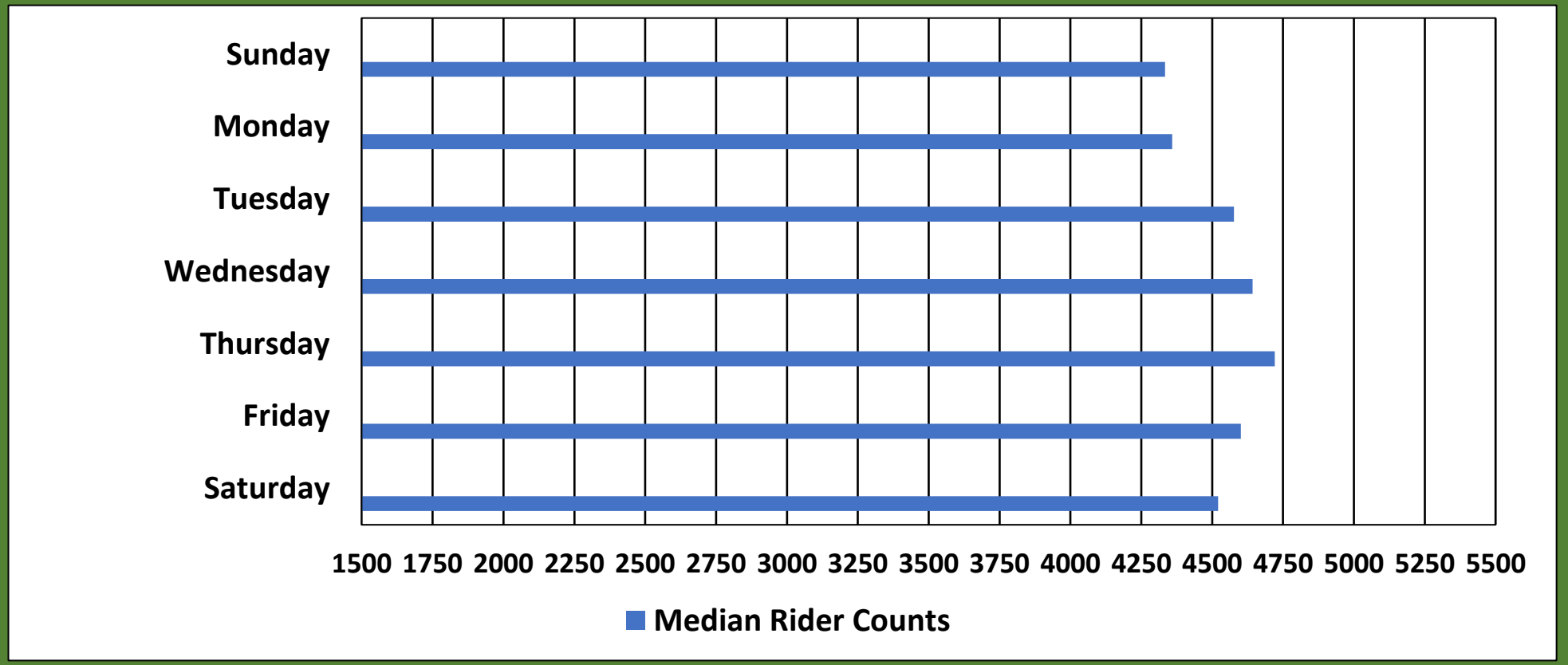
The histograms, and other tests and plots not shown, indicate the counts do not have a convincing normal distribution. A non-parametric test should be used to assess if there is a significant difference between the two groups. The Kruskal-Wallis test and Mood's median test are used.

Group Differences in a Rental Bike Sharing Dataset

Results

- **Median Rider Counts on Different Days of Week**

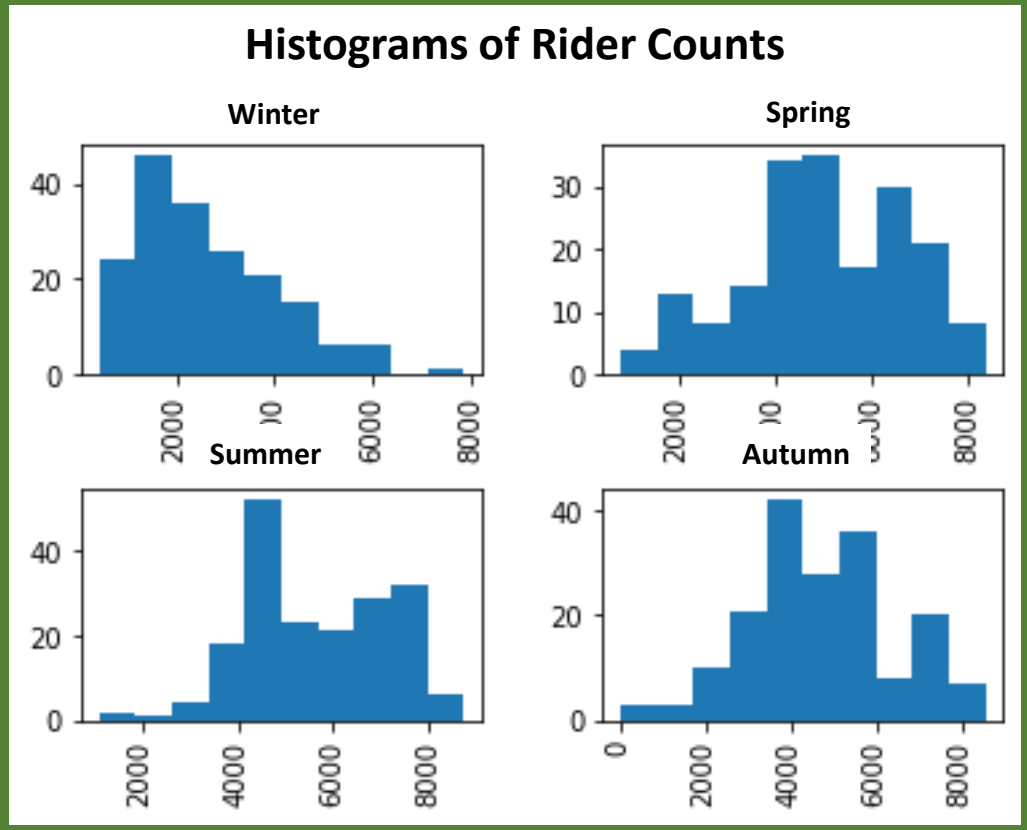
Significant Difference?	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
No	4334.0	4359.0	4576.5	4642.5	4721.0	4601.5	4521.0



Group Differences in a Rental Bike Sharing Dataset

Methods

- **Median Rider Counts in Different Seasons**



The histograms, and other tests and plots not shown, indicate the counts do not have a convincing normal distribution. A non-parametric test should be used to assess if there is a significant difference between the two groups. The Kruskal-Wallis test and Mood's median test are used.

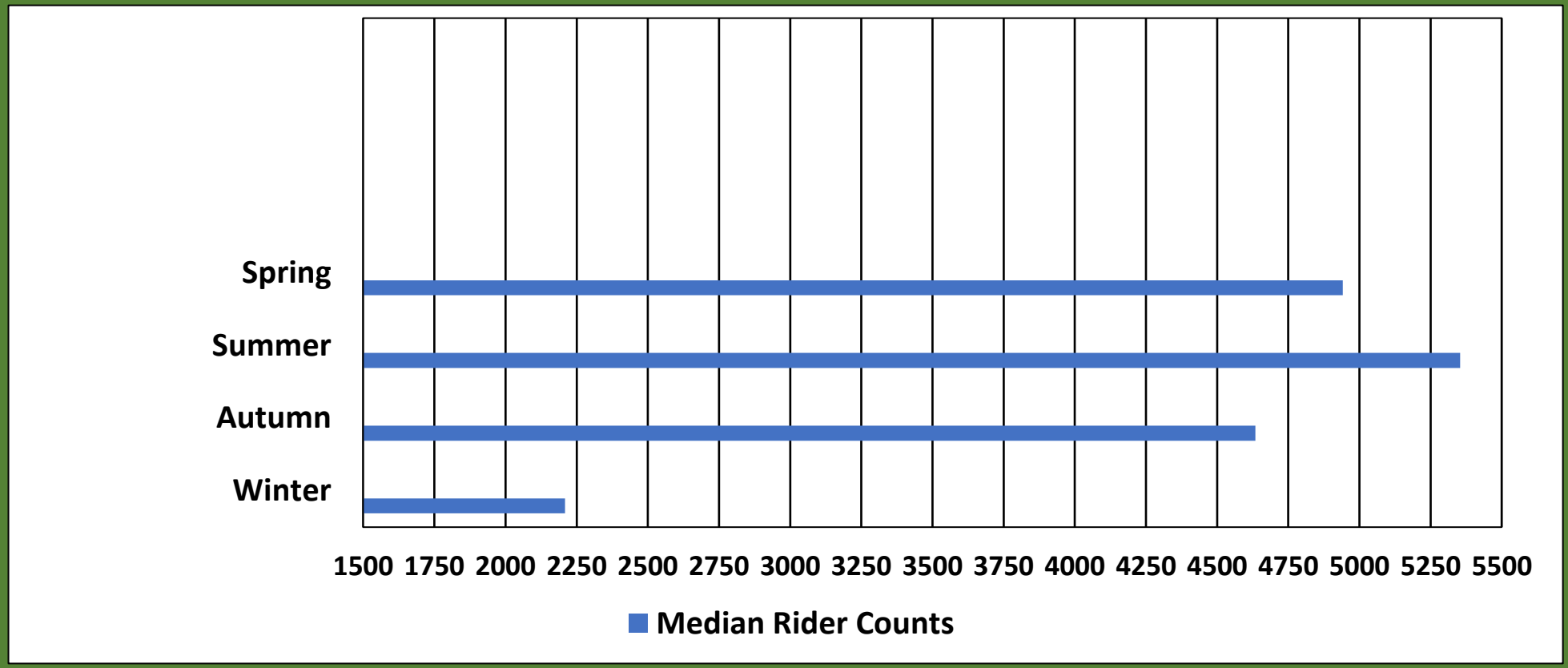
Group Differences in a Rental Bike Sharing Dataset

Results

- **Median Rider Counts in Different Seasons**

Significant Difference?	Spring	Summer	Autumn	Winter
Yes	4941.5	5353.5	4634.5	2209.0

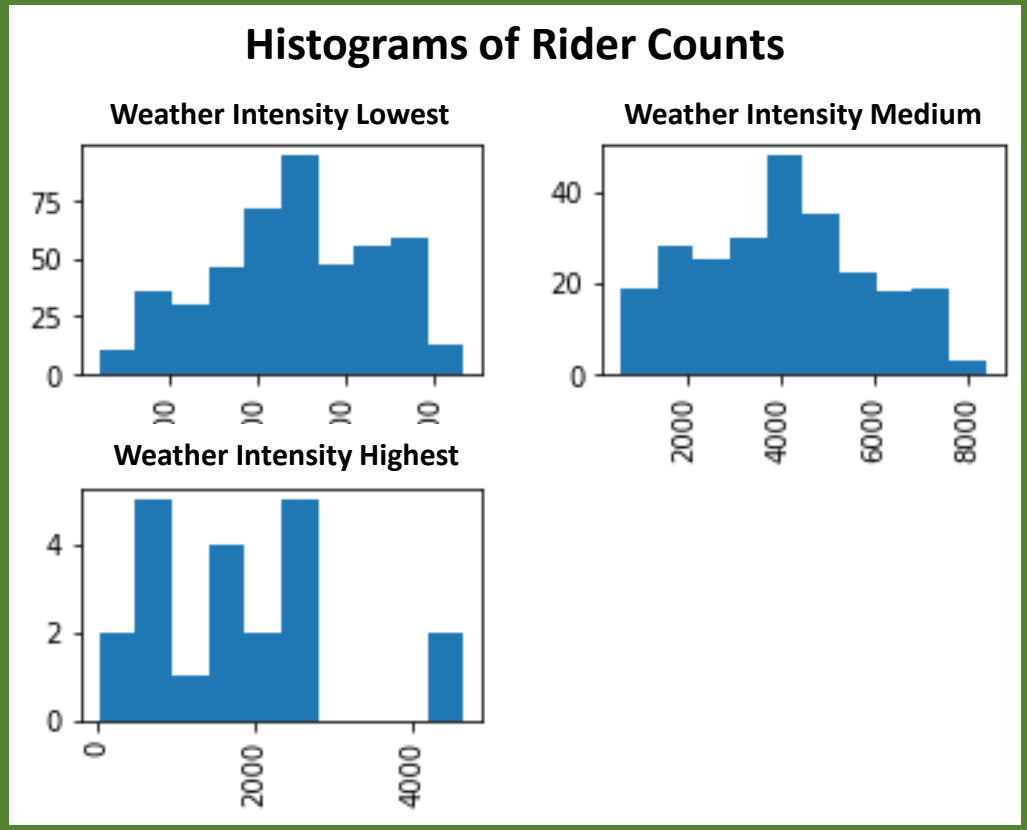
All pairs also have a significant difference except for spring and autumn pair.



Group Differences in a Rental Bike Sharing Dataset

Methods

- **Median Rider Counts in Different Weather Intensity Groups**



The histograms, and other tests and plots not shown, indicate the counts do not have a convincing normal distribution. A non-parametric test should be used to assess if there is a significant difference between the two groups. The Kruskal-Wallis test and Mood's median test are used.

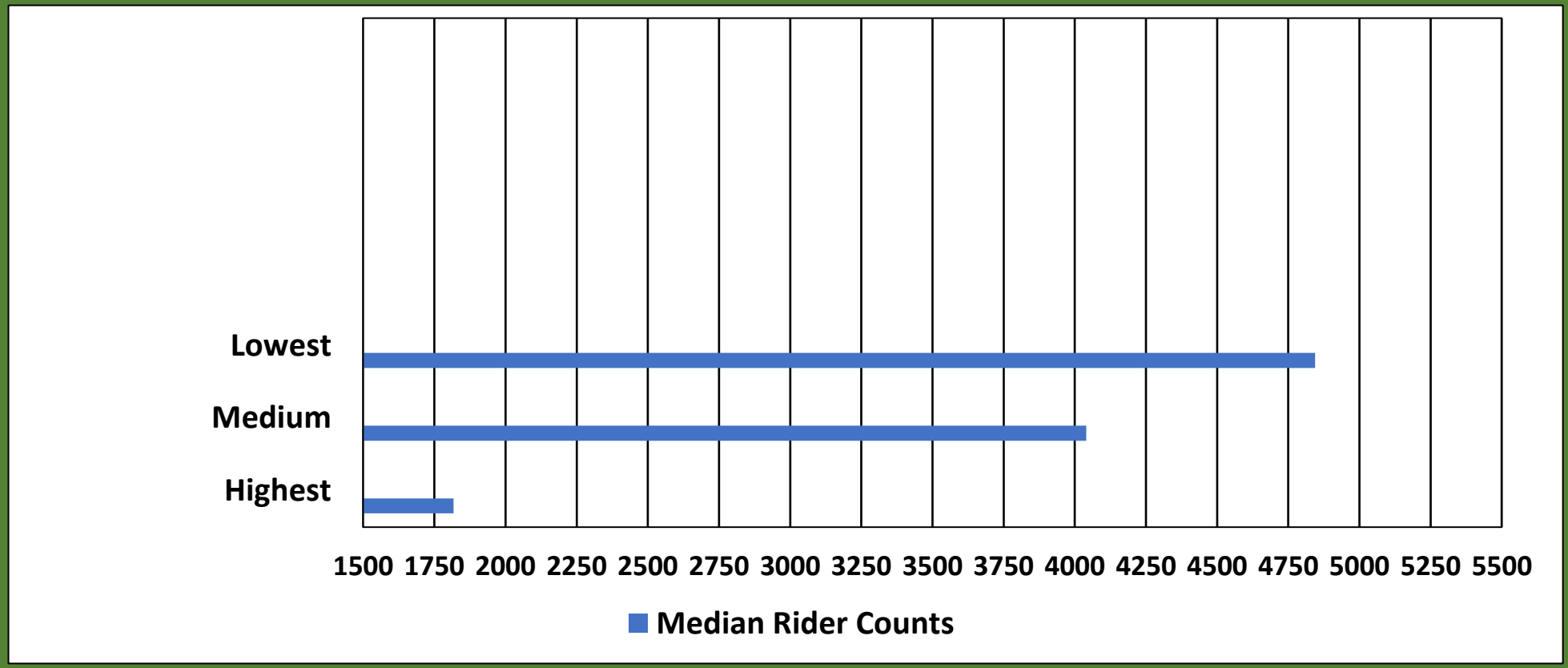
Group Differences in a Rental Bike Sharing Dataset

Results

- **Median Rider Counts in Different Weather Intensity Groups**

Significant Difference?	Weather Intensity Lowest	Weather Intensity Medium	Weather Intensity Highest
Yes	4844.0	4040.0	1817.0

All pairs also have a significant difference.



Group Differences in a Rental Bike Sharing Dataset

Recommendations

- There is no clear indication whether bikes are used more for commuting or for recreation or by DC/MD/VA residents or by visitors. Further research could be done into the hours when the bikes are used and consideration of casual and registered rider counts.

- Revenue clearly has a seasonal cycle. The company may want to promote more riding in the colder and darker times of the year, but this may be a difficult sell, and visibility safety is important.

- Winter or times of intense weather may be good times to remove bikes for maintenance.