Bike sharing data

Group 5

(Kristina Yamkovoy, Zhixing Zhao, Yueyang Zhang, Xueying He, Hernisha Radia)

**Introduction**

The success and availability of bike sharing services depends highly on the usage of such services. Bikes must be available when and where people need them; otherwise, the rental service will not be used to its full capacity, and the provider will not be able to offset operation and maintenance costs. By understanding customers’ usage of the bike sharing service, the provider can optimize their profit by ensuring that bikes are available when needed, and thus provide a better service for consumers.

**Data**

The data is provided on both daily and hourly time scales. The outcome of interest is the count of bikes rented - by registered users, casual users, and both. Additionally, information on the time of rental is provided, including the date, season, day of the week, whether or not the day is a holiday, and whether or not the day is a weekday. Information on weather includes temperature, humidity, windspeed, and general quality of weather.

**Questions**

Interesting questions raised by this data include:

* How does weather affect rental counts? Do only some weather variables have an effect?
* How does time of day affect rental counts? Are there more rentals during rush hour? If there is a relationship between rental count and time of day, does it hold on both weekdays and weekends?
* If weather is related to rental counts? What about vice versa?
* Are people more intend to ride bikes during holiday or not? Which holiday has most rentals? How does weather affect this?
* Is there a difference in number of people using shared bikes between weekends and weekdays? Compare registered vs. casual.
* Is there any relationship between the number of registered users and not registered users? Are registered users more likely to use the bike for multiple times?
* Look into correlation between variables (collinearity)
  + Run regression with all variables separately - see which have higher p-values
  + Variable selection (automatic or manual)