Lab 8

Lab Overview

For this lab we will look at the impact of weather on bike rentals using the (infamous) Capital Bikeshare dataset. The data set contains total number of daily bike rentals across a two year period. For each date there are also a set of weather variables. We will restrict our inference to the (astronomical) summer season with a goal of predicting total bike rentals across the 188 days.

```
bike_rentals <- read_csv('http://math.montana.edu/ahoegh/teaching/stat446/daily_bike.csv')
bike_summer <- bike_rentals %>% filter(season == 3) %>% select(cnt, dteday, yr, mnth, temp, hum, windsp
bike_summer
```

```
## # A tibble: 188 x 10
##      cnt dteday    yr  mnth  temp   hum windspeed workingday weekday
##    <dbl> <chr>  <dbl> <dbl> <dbl> <dbl>     <dbl>      <dbl>   <dbl>
##  1  4835 6/21/~     0     6 0.681 0.770    0.171           1       2
##  2  4507 6/22/~     0     6 0.733 0.708    0.172           1       3
##  3  4790 6/23/~     0     6 0.728 0.703    0.239           1       4
##  4  4991 6/24/~     0     6 0.724 0.573    0.222           1       5
##  5  5202 6/25/~     0     6 0.695 0.483    0.210           0       6
##  6  5305 6/26/~     0     6 0.68  0.513    0.0945          0       0
##  7  4708 6/27/~     0     6 0.682 0.658    0.108           1       1
##  8  4648 6/28/~     0     6 0.744 0.634    0.144           1       2
##  9  5225 6/29/~     0     6 0.728 0.498    0.262           1       3
## 10  5515 6/30/~     0     6 0.697 0.434    0.185           1       4
## # ... with 178 more rows, and 1 more variable: weathersit <dbl>
```

Data Dictionary

dteday: Date yr: Year (0: 2011, 1:2012) mnth: Month (1 to 12) holiday: weather day is holiday or not (extracted from Holiday Schedule) temp: Normalized temperature in Celsius. The values are derived via (t-t_min)/(t_max-t_min), hum: Normalized humidity. The values are divided to 100 (max) windspeed: Normalized wind speed. The values are divided to 67 (max) weekday: Day of the week workingday: If day is neither weekend nor holiday is 1, otherwise is 0. weathersit: (extracted from Freemeteo) 1: Clear, Few clouds, Partly cloudy, Partly cloudy 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist 3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds 4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog cnt: count of total rental bikes including both casual and registered

Acknowledgements:

Hadi Fanaee-T Laboratory of Artificial Intelligence and Decision Support (LIAAD), University of Porto INESC Porto, Campus da FEUP Rua Dr. Roberto Frias, 378 4200 - 465 Porto, Portugal

Original Source: http://capitalbikeshare.com/system-data

Weather Information: http://www.freemeteo.com

Holiday Schedule: http://dchr.dc.gov/page/holiday-schedule

1. (20 points)

Propose and carryout a sampling scheme with the goal of estimating the total number of bikes rented for the 188 days. The lab should consist of a short writeup that parallels the project. The only requirements are to have a point estimate and associated confidence interval for the result. Note you are welcome to use any sampling / estimation approaches, but make sure to defend those choices. You can take a total of 40 samples in whatever sampling scheme you choose.

Introduction

Brief overview of the question of interest.

Sampling Scheme

Detail the sampling and estimation approaches.

Results

Discuss the results of your sampling and estimation.

Discussion

Summary of your results, assume this is written for employees at Capital Bikeshare.