3/15 MDS作業

> setwd("~/Downloads/MDS實做\_#01\_資料科學趨勢與工具")

> train = read.csv('MDS實作\_#01\_資料檔\_train.csv')

> library(lubridate)

> train$season <- factor(train$season, labels = c("春", "夏", "秋", "冬"))

> train$weather <- factor(train$weather, labels = c("出太陽", "陰天", "雨天", "雨天且寒冷"))

> train$hour <- factor(hour(ymd\_hms(train$datetime)))

> train$times <- as.POSIXct(strftime(ymd\_hms(train$datetime), format="%H:%M:%S"), format="%H:%M:%S")

> train$jitter\_times <- train$times+minutes(round(runif(nrow(train),min=0,max=59)))

> train$Weekday <- lubridate::wday(ymd\_hms(train$datetime), label=TRUE)

> library(ggplot2)

> library(plyr)

> weather\_prob <- ddply(train,.(season, hour),

+ summarise, Good = mean(weather == "出太陽"),

+ Normal = mean(weather == "陰天"),

+ Bad = mean(weather == "雨天"),

+ Very\_bad = mean(weather == "雨天且寒冷"))

> ggplot(train, aes(x = hour, y = Normal, colour = season)) +

+ geom\_point(data = weather\_prob, aes(group = season)) +

+ geom\_line(data = weather\_prob, aes(group = season)) +

+ scale\_x\_discrete("一天24小時分布") +

+ scale\_y\_continuous("陰天的機率") +

+ theme\_minimal() +

+ ggtitle("陰天與單車租借的關係\n在冬天時，陰天的機率最高") +

+ theme(plot.title=element\_text(size=18))